WORKSHOP ON THE PRODUCTION OF SWEPT-AREA ESTIMATES FOR ALL HAULS IN DATRAS FOR BIODIVERSITY ASSESSMENTS (WKSAE-DATRAS)

VOLUME 3 | ISSUE 74

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 74

WORKSHOP ON THE PRODUCTION OF SWEPT-AREA ESTIMATES FOR ALL HAULS IN DATRAS FOR BIODIVERSITY ASSESSMENTS (WKSAE-DATRAS)

Recommended format for purpose of citation:

ICES. 2021. Workshop on the production of swept-area estimates for all hauls in DATRAS for biodiversity assessments (WKSAE-DATRAS).
ICES Scientific Reports. 3:74. 77 pp. https://doi.org/10.17895/ices.pub.8232

Editors

Kai Ulrich Wieland

Authors

Juan Carlos Arronte • Francisco Baldó • Patrik Börjesson • Finlay Burns • Corina Chaves • Ruadhán Gillespie-Mules • Ailbhe Kavanagh • Ruth Kelly • Jed Kempf • Cecilia Kvaavik • Kim Ludwig • Valentina Melli • Meadhbh Moriarty • Anna Rindorf • Lara Salvany • Sonia Seixas • Anne Sell • Vaishav Soni Morgane Travers-Trolet • Francisco Velasco Guevara • Kai Ulrich Wieland • Hongru Zhai

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

Contents

i Executive summary

The workshop on the production of swept-area estimates for all hauls in DATRAS for biodiversity assessments (WKSAE-DATRAS) considered three groups of surveys for which data are submitted to the Database of Trawl Surveys (DATRAS): various Beam Trawl Surveys, the Northeast Atlantic International Bottom Trawl Survey (Northeast Atlantic IBTS), and the North Sea International Bottom Trawl Survey (North Sea IBTS).

All countries contributing to the above-mentioned surveys were represented by at least one participant during the workshop, apart from the Netherlands and Norway.

The main objectives of the workshop were to establish tow-by-tow swept-area estimates for time-series as far back in time as possible, compare different approaches for the estimates of missing observations, and harmonize the resulting dataseries for biodiversity assessments.

For all of the surveys considered, problems with data quality were detected. This included the Beam Trawl Surveys but was most pronounced for the North Sea IBTS. Outliers and potential erroneous data were listed for reporting back to the respective national institutes. In particular, missing observations or algorithms affected wing spread-based swept-area, which is needed in several applications.

This workshop compared the Marine Scotland Science-MSS/OSPAR approach, which includes a data quality check for the information needed for the calculation of swept-area, and the DATRAS approach, which depends solely on correctly reported data from the national institutes. Larger data gaps were identified, in particular for several years of the North Sea IBTS. For those surveys, it is proposed that the best possible way forward at this moment is to use estimates based on the MSS/OSPAR approach.

However, if dubious records (i.e. extreme outliers) were identified by the MSS/OSPAR and no other information was available, values (e.g. speed over ground or the depth at which a change from short to long sweeps should have happened) were taken from the manual. However, experience has shown that the survey manuals are not followed in all instances, and so persistent country-specific and survey-specific deviations may occur.

The national institutes are encouraged to check, correct, and fill in missing survey data through re-submissions to DATRAS. It is recommended that DATRAS data quality control on data submission is extended for the information needed for the calculation of swept-area (e.g. distance, depth, door spread, and wing spread) and that this is done in close cooperation between the ICES Data Centre and the respective ICES survey working groups, WGBEAM (Working Group on Beam Trawl Surveys) and IBTSWG (International Bottom Trawl Survey Working Group).

ii Expert group information

Expert group name	Workshop on the production of swept-area estimates for all hauls in DATRAS for biodiversity assessments (WKSAE-DATRAS)
Expert group cycle	Annual
Year cycle started	2021
Reporting year in cycle	1/1
Chair	Kai Wieland, Denmark
Meeting venue and dates	31 May–4 June 2021, online meeting, (19 participants)

1 Introduction

Workshop on the production of swept-area estimates for all hauls in DATRAS for biodiversity assessments (WKSAE-DATRAS)

1.1 Assessment of biological diversity and foodwebs in marine environments–policy context

Fish constitute a major fraction of the marine biota in marine ecosystems and have an essential role in foodweb structure and function. Fish communities are key components of marine biodiversity and marine foodwebs. Reporting and assessment of fish communities are essential to evaluate the environmental status across the Northeast Atlantic waters covered by the Marine Strategy Framework Directive (MSFD).

The OSPAR (Oslo/Paris convention for the Protection of the Marine Environment of the North-East Atlantic) Commission ensures a common approach to the management of the human activities that affect the OSPAR maritime area and to the assessment of the marine environment. Furthermore, OSPAR facilitates the coordinated implementation of the MSFD, to achieve good environmental status (GES) according to 11 descriptors, including biological diversity (D1) and foodweb structure and function (D4).

A suite of indicators is being developed by OSPAR for the Quality Status Report 2023[1] to assess fish population and foodwebs. Fish community common indicators (potential indicators not included) are:

- Recovery in the population abundance of sensitive fish species (FC1)
- Proportion of large fish (Large Fish Index) (FC2)

Foodweb common indicators (potential indicators not included) are:

- Size composition in fish communities (FW3)
- Change in average trophic level of marine predators in the Bay of Biscay (FW4)

These indicators use data from scientific groundfish surveys conducted through ICES.

1.2 Groundfish surveys

Groundfish surveys intended to sample commercial fish species populations to support formal stock assessments under the European Union's Common Fisheries Policy (CFP) can also be used to monitor and assess the status of the broader fish community.

However, the use of groundfish survey data for environmental assessments was not the original purpose and the detail and resolution level of the data varies across historic time-series. Data collection and detail of reporting for non-commercial fish are in general poorer.

Data from groundfish surveys are free and available for download from the DATRAS database portal on the ICES website. The national institutes initially check data prior to submission to ICES in a specific format (see section 3). A further screening process is applied at ICES before the data are accepted and incorporated into the DATRAS database. The full screening process is

[1] https://www.ospar.org/work-areas/cross-cutting-issues/qsr2023

valid only for data from 2004 onwards meaning that some of the more historical data have not been subject to the same level of quality assurance.

The list of surveys, areas, and time-series covered by groundfish surveys in WKSAE-DATRAS is summarized in Table 1.1.

Table 1.1. List of individual surveys considered in WKSAE-DATRAS.

DATRAS identifier	Survey Acronym In OSPAR product	Country	Start Year	End Year	Vessels	Quarter	Gear Type	Sub-region	Data Source
IBTSQ1	GNSIntOT1	International	1983	2021	Multiple Ships	1	Otter (GOV)	Greater North Sea	DATRAS
IBTSQ3	GNSIntOT3	International	1998	2021	Multiple Ships	3	Otter (GOV)	Greater North Sea	DATRAS
FR CGFS	GNSFraOT4	France	1988	2020	Thalassa II, Gwen Drez	4	Otter (GOV)	Greater North Sea	DATRAS
SCOWCGFSQ1(1)	CSScoOT1	Scotland	1985	2021	Scotia II/III	1	Otter (GOV)	Celtic Sea	DATRAS
SCOWCGFSQ4(1)	CSScoOT4	Scotland	1990	2021	Scotia II/III	4	Otter (GOV)	Celtic Sea	DATRAS
IE-IGFS	CSIreOT4	Ireland	2003	2020	Celtic Explorer	4	Otter (GOV)	Celtic Sea	DATRAS
NIGFSQ1	CSNIrOT1	Northern Ireland	1992	2020	Lough Foyle/ Corystes	1	Otter (ROT)	Celtic Sea	DATRAS
NIGFSQ4	CSNIrOT4	Northern Ireland	1992	2020	Lough Foyle/ Corystes	4	Otter (ROT)	Celtic Sea	NDB (92-07) DATRAS (08-15)
EVHOE	CS/BBFraOT4	France	1997	2020	Thalassa II	4	Otter (GOV)	Celtic Sea/Bay of Biscay	NDB (92-07) DATRAS (08-15)
SP-North	BBIC(n)SpaOT4	Spain	1990	2020	Cornide de Saavedra Miguel Oliver	4	Otter (BACA)	Bay of Biscay and Iberian Coast	NDB
SP-ARSA	BBIC(s)SpaOT1	Spain	1992	2020	Cornide de Saavedra Miguel Oliver	1	Otter (BACA)	Bay of Biscay and Iberian Coast	NDB
SP-ARSA	BBIC(s)SpaOT4	Spain	1997	2020	Cornide de Saavedra Miguel Oliver	4	Otter (BACA)	Bay of Biscay and Iberian Coast	NDB
PT-IBTS	BBICPorOT4	Portugal	2002	2018	Capricornio, Noruega	4	Otter (NCT)	Bay of Biscay and Iberian Coast	DATRAS
SCOROC(2)	WAScoOT3	Scotland	1999	2020	Scotia II/III	3	Otter (GOV)	Wider Atlantic	DATRAS
SP-PORC	WASpaOT3	Spain	2001	2020	Vizconde de Eza	3	Otter (PBACA)	Wider Atlantic	NDB
BTS	GNSNetBT3	Netherlands	1987	2020	Isis, Tridens II	3	Beam (8m)	Greater North Sea	DATRAS
BTS	GNSEngBT3	England	1990	2020	Carhelmar, Corystes, Endeavour	3	Beam (4m)	Greater North Sea	DATRAS
BTS	GNSGerBT3	Germany	2002	2020	Solea I/II	3	Beam (7m)	Greater North Sea	DATRAS
BTS	CSEngBT3	England	1993	2020	Corystes, Endeavour	3	Beam (4m)	Celtic Sea	DATRAS
BTS	Not included	Belgium	2004	2020	RV Belgica, The Ramblers	3	Beam (4m)	Greater North Sea	DATRAS

(1) Scottish surveys on the west coast started a new series in 2011 (SCOWCGFSQ4), previous surveys (1985-2010) are in DATRAS as SWC-IBTS

(2) Rockall surveys started a new series in 2011 (SCOROC), previous surveys (1999-2009) are in DATRAS as ROCKALL

The International Bottom Trawl Survey Working Group (IBTSWG) and Working Group on Beam Trawl Surveys (WGBEAM) are the umbrella working groups at ICES coordinating the surveys listed in Table 1.1. The national surveys included under IBTSWG or WGBEAM follow the same specified sampling protocol and tow the same gear at a specific speed for a determined amount of time.

However, there is some degree of decision-making power at the national level. Individual countries are responsible for cleaning the data before submission to ICES and selecting the procedure used to estimate missing values.

Previous assessments of fish community and foodwebs by OSPAR (Intermediate Assessment 2017) used an approach to calculate swept-area indices described in section 2. The haul level (Sampling Information) from this product was rerun by this workshop, to update the outputs to include the most recent surveys and data changes on DATRAS (see section 6).

1.3 Swept-area calculations

Swept-area (tonnes per km^2) based indices can be used to calculate ecological indicators of biodiversity of fish communities and foodwebs. However, there is high variability of the area swept by trawls primarily linked to variation in tow speed, depth, and door and wing spread separation. On occasion, the information required to estimate the area swept by trawl is missing, especially in the early time-series. Statistical modelling is then necessary to estimate the missing values. This workshop primarily considered the following data fields, in relation to the calculation of swept-area products: Sample Location, Depth, Sweep Length, Haul Duration, Groundspeed, Towed Distance, and Otter Trawl Geometry (e.g. door and wing spread). The fields included for each survey and biological information are presented in Table 1.2 and Table 1.3.

Table 1.2. Sampling information required in each survey (OSPAR version).

Field		Unit	Description
HaulID	A27		Unique haul identifier (SurveyAcronym/Ship/Year/HaulNo)[1] (*H*)
Survey-Acronym	A13		Unique survey identifier (SubregionCountryGearTypeQuarter: e.g. GNSNedBT3)
Ship	A4		Unique vessel identifier (e.g. SCO3: *Scotia* III)
GearType	A4		Unique gear type code (BT = Beam Trawl, OT = Otter Trawl)
Gear	A6		Unique gear code (e.g. GOV = Grande Oerverture Verticale)
YearShot	S		Year that gear was shot[2]
MonthShot	S		Month that gear was shot[2]
DayShot	S		Day that gear was shot[2]
TimeShot	S	GMT	Time that gear was shot (in format HHMM)[3]
HaulDur(min)	S	min	Duration of fishing operation[4]
ShootLat(decdeg)	N	Deg.	Latitude in decimal degrees of the haul shoot position[5]
ShootLong(decdeg)	N	Deg.	Longitude in decimal degrees of the haul shoot position[5]
ICESStSq	A12		ICES statistical rectangle where gear was shot
SurvStratum	A12		Stratum tag for stratified surveys[6]
Depth(m)	N	m	Depth tag assigned to the haul[7]
Distance(km)	N	km	Tow distance[8] ($d_{H,TOW}$)
WingSpread(m)	N	m	Mean distance between the wings during fishing operation[9,12] ($d_{H,WING}$)
DoorSpread(m)	N	m	Mean distance between the doors during fishing operation[10,13] ($d_{H,DOOR}$)
NetOpen(m)	N	m	Mean head-line height above seabed during fishing operation[11,14] ($d_{H,HEIGHT}$)
WingSwptArea(sqkm)	N	km^2	Area of seabed swept by the net[15] ($A_{H,WING} = d_{H,TOW} \times d_{H,WING}$)
WingSwptVol_CorF	N		Multiplier ($1 / d_{H,HEIGHT}$): converts to 'density by wing-swept volume'[16]
DoorSwptArea_CorF	N		Multiplier ($d_{H,WING} / d_{H,DOOR}$): converts to 'density by door-swept area'[17]
DoorSwptVol_CorF	N		Multiplier ($d_{H,WING} / (d_{H,DOOR} \times d_{H,HEIGHT})$): converts to 'density by door-swept volume'[18]

Table 1.3. Biological information (OSPAR version).

Field	Unit	Description
HaulID		Unique haul identifier (SurveyAcronym/Ship/Year/HaulNo)[1] (*H*)
SpeciesSciName		Unique species name for each species sampled across the NE Atlantic[2] (*S*)
FishLength(cm)	cm	Integer numbers indicating fish length to the 'cm below'[3] (*L*)
IndivFishWght(g)	g	Estimated weight of individual fish of specified species and length[4] ($W_{S,L}$)
Number		Total number of fish of specified species and length in the catch[5] ($N_{S,L,H}$)
DensAbund(N_sqkm)	km^{-2}	Abundance density estimate[6,8] ($D_{nos,S,L,H} = N_{S,L,H} / A_{H,WING}$)
DensBiom(kg_Sqkm)	kg km^{-2}	Biomass density estimate[7,8] ($D_{biom,S,L,H} = (N_{S,L,H} \times W_{S,L}) / A_{H,WING}$)

The data collected comprises the number of each species of fish sampled in each trawl, measured to defined length categories. By dividing these species catch numbers-at length by the area swept on each sampling occasion, the catch data are converted to standardized estimates of fish density-at-length, by species, at each sampling location i.e. the data product. Length-weight relationships to convert numbers at length to biomass at length data, for all well-sampled species (list to be supplied), are also required in order to split the size-frequency data and calculate the biomass of large fish vs. small fish (i.e. above and below length threshold) for the Large Fish Indicator. The biomass of fish within trophic guilds also requires biomass at size data since small and large individuals of fish species may have very different diets.

1.4 Aims and expectations of the workshop

WKSAE-DATRAS was tasked to discuss and agree on several remaining issues:

1. Quality checks: For NS-IBTS, BTS and NEA-IBTS, there are issues in current products that need to be checked with reported data.

2. Time-series: The DATRAS scripts currently contain conditions based on time- series. WKSAE-DATRAS needs to discuss and agree on which time-series to be included.
3. Target species: An analysis and calculation for all species would be too time-consuming for DATRAS using current methods requested by BTS. WKSAE-DATRAS needs to decide what species to work with for BTS surveys.
4. Year ranges: WKSAE-DATRAS needs to decide on whether quarters can be merged within years, or data should be split between quarters when estimated missing gear parameters in the Northeast Atlantic IBTS.
5. Criteria for missing observations: WKSAE-DATRAS needs to discuss and agree to what to do when data are missing, options are to use MSS approach (e.g. fill missing values from information from other data fields or values given in survey manuals) or not use those survey points in the calculation of indices (e.g. Norway data for NS-IBTS and Portugal data for NS-NEA)?
6. R script: Needs to be reviewed by WKSAE-DATRAS.
7. What is next after SWKM2 product by WS or DS? Do we aim for species based WS and DS calculation for NS- and NEA-IBTS?
8. DATRAS documentation needs to be reviewed and updated. Revision and merge all surveys into one document.
9. Future products: It is important to know where we are heading to. Our aim is to start developing indicators.

2 MSS/OSPAR approach for calculating swept-area

To provide background information for the WKSAE-DATRAS workshop, Dr Moriarty outlined the protocols used in the development of the 'Groundfish Survey Monitoring and Assessment Data Products for the Northeast Atlantic' by Marine Scotland Science (MSS) with OSPAR in 2017.

The decision-making process for assessing haul parameters in the formal Quality Assurance–Quality Audit (QAQA) undertaken by Marine Scotland Science is summarized below:

Nineteen groundfish surveys were subjected to a comprehensive QAQA protocol documented in Moriarty *et al.* (2017); Greenstreet and Moriarty (2017a); and Greenstreet and Moriarty (2017b), with additional information available in Moriarty *et al.* (2019). The QAQA process was originally applied to European groundfish survey data to ensure their adequacy to support Marine Strategy Framework Directive (MSFD) needs. Source data were downloaded from the ICES DATRAS portal (ICES, 2017) where available, or when not available on DATRAS, data were provided directly by the national institutes involved. The overall aim was to produce fully QAQA compliant, groundfish survey 'data products' that could provide the basis for assessments of the groundfish component of marine ecosystems across the entire Northeast Atlantic region. These surveys provide temporal coverage of between 10y and 35y and a spatial coverage spanning much of the Northeast Atlantic (Figure 2.1).

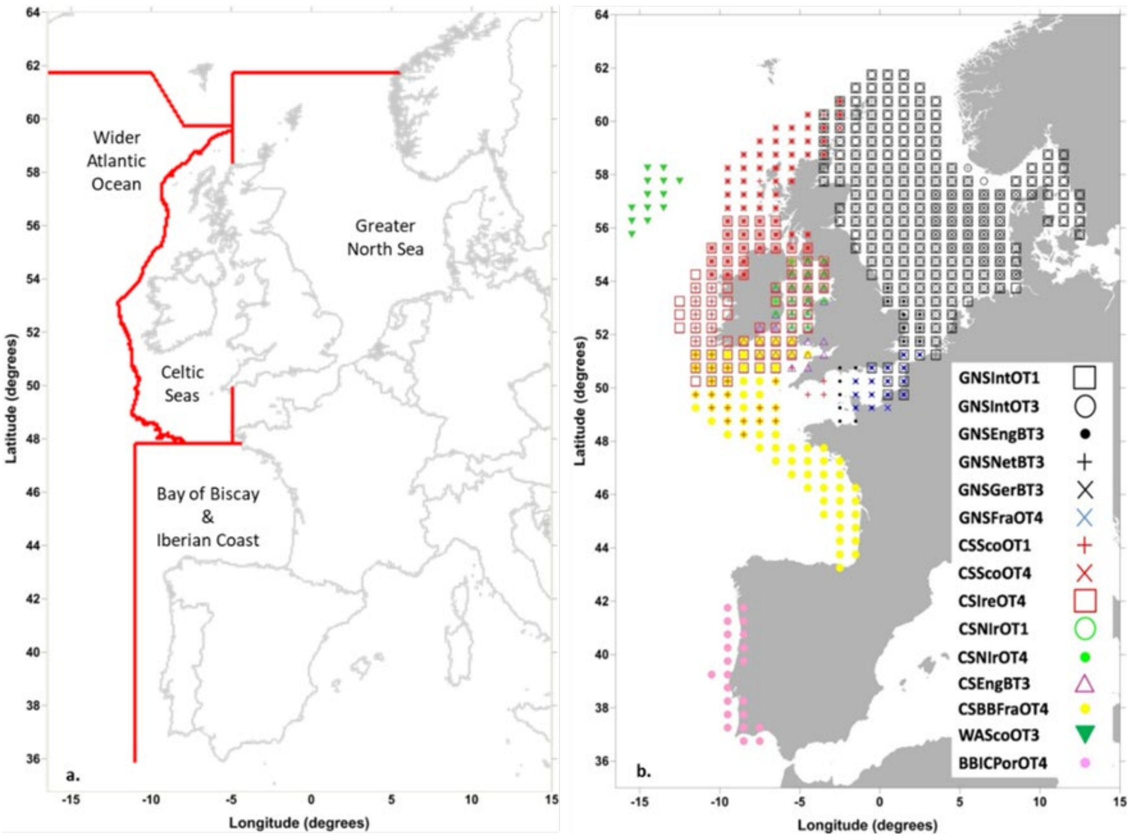


Figure 2.1. The subregions of the Northeast Atlantic Ocean Region; The Greater North Sea (GNS; including the Kattegat and the English Channel), the Celtic Seas (CS), the Bay of Biscay (BB), and Iberian Coast (IC), and the Wider Atlantic (WA) Ocean (Shapefile source: OSPAR website). Right: Survey coverage of the 15 published datasets across the Northeast Atlantic. Survey acronyms follow a consistent formula of region, country, gear (OT = otter trawl, BT = beam trawl) and quarter (Figure source: Moriarty *et al.*, 2019).

Data were processed following the protocol summarized in Figure 2.2 to derive the data products (green box). The three blue oval steps constitute the main quality assurance part of the protocol; the individual decision-making processes contained in these steps are summarized in Table 1.1. The orange and lilac 'review' box steps, along with the detailed documentation describing the whole QAQA protocol (Moriarty, *et al.*, 2017; Greenstreet and Moriarty, 2017a; 2017b), create the quality audit.

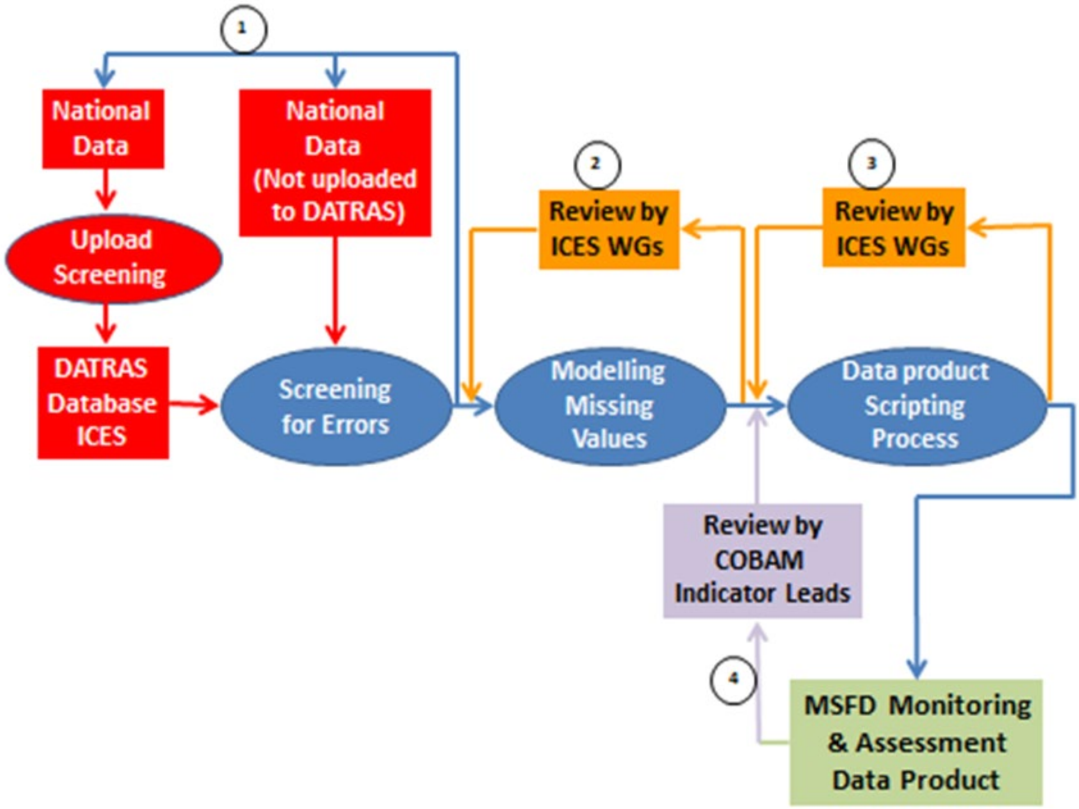


Figure 2.2. Overview of the groundfish survey monitoring and assessment process relevant to the ICES community. Numbers highlight the different feedback loops following consultation with national data providers (1), ICES Working Groups (2 and 3), and indicator leads (4) (Figure source: Moriarty *et al.*, 2019).

The screening process involved examining all relevant parameter values for outlier and missing values. Relevant parameters considered included: Sample Location, Depth, Sweep Length, Haul Duration, Groundspeed, Towed Distance and Otter Trawl Geometry. Where values were absent, and perhaps never recorded, models were developed for each parameter so that missing values could be filled by modelled estimates (Moriarty *et al.*, 2017). Potential data errors were referred back to relevant NDPs for checking (feedback loop 1, Figure 2.2). Three outcomes were possible:

1. Datum was confirmed to be correct and simply an outlier,
2. Datum was deemed to be 'erroneous', or,
3. Datum was deemed to be 'incorrect'.

'Erroneous data' were a consequence of imperfect data archiving: a typo. These were corrected simply by editing the archived values and re-uploading the revised national data to DATRAS. 'Incorrect data' were more difficult to rectify; here archived values matched original values recorded at source. If mistakes occurred, they happened at source and it was no longer possible to establish whether the value in question was a data error or a correct but outlier value. In these instances, it was necessary to decide whether the value in question had sufficient credibility as to be possible, or whether the recorded value was so unlikely that it must be considered wrong.

Clear criteria were defined and described in various decision-making flow charts such as the example given in Figure 2.3.

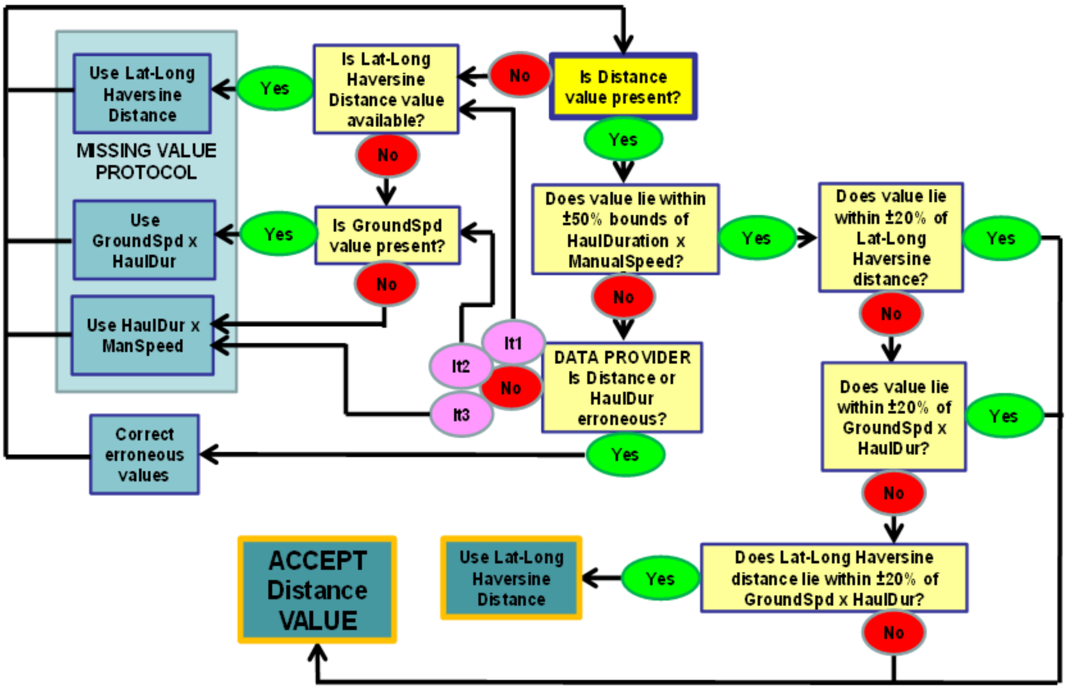


Figure 2.3. An example flow chart illustrating the steps involved in assessing the validity of recorded towed distance values, estimating missing values and replacing incorrect data from Moriarty *et al.* (2017) (Figure source: Moriarty *et al.*, 2019).

Flow charts were developed for each haul parameter needed to calculated swept-area estimates based on the expert judgement from the ICES survey working groups, the OSPAR indicator leads and the data product authors. Where the datum was deemed to be 'incorrect', so extreme an outlier as to not be possible, these data were deleted and a 'missing value' procedure was employed to replace them with modelled estimates (Moriarty *et al.*, 2017).

Replacing 'incorrect' and 'missing' values in this way was considered to be preferable to the alternative of simply deleting the records concerned for three main reasons.

1. Individual parameter values often affected other data.
2. Deletion of missing or incorrect data would impart bias
3. Missing or incorrect data were more common in the early years of most surveys; deletion of these data could have compromised time-series longevity.

Table 2.1. Summary of issues identified in the groundfish survey data stored on the DATRAS portal or on national databases and approaches adopted to address these. (Table source: Moriarty *et al.*, 2019).

Issue	Solution
Haul positions missing or same as shoot position.	Haul position deleted if same as shoot position. Georeferencing dependent on shoot position.
Shoot/Haul positions outside reported ICES statistical rectangle.	If position correct, ICES rectangle adjusted. If ICES rectangle correct, position altered to rectangle midpoint.
Reported depths checked against bathymetry map. Deviation of ± 50% checked.	Erroneous values corrected, otherwise all recorded depths considered correct.
Missing depth data (1% of samples).	Depth from bathymetry map at trawl location assigned.

Issue	Solution
Missing sweep-length data (40% of samples).	Available data suggested close adherence to survey manuals. Missing values filled with manual recommendations.
Extreme haul duration values.	Invariably correct or erroneous. If erroneous, corrected accordingly. No missing values.
Missing groundspeed data (38% of samples). Incorrect groundspeed value recorded.	Groundspeed estimated from one of two possible models using Quarter, Vessel, and Gear as factors.
Missing/incorrect towed distance data.	Estimated as: 1) Haversine distance between shoot/haul positions (15.1% samples); 2) function of tow duration x groundspeed (7.3% samples); function of tow duration and manual recommended groundspeed (0.2% samples).
Missing/incorrect wing-spread values (44% of samples).	Estimated using one of four models using door-spread, depth, and gear as factors or using value stipulated in relevant survey manual.
Missing/incorrect door-spread values (29% of samples).	Estimated using one of four models using wing-spread, depth, and gear as factors or using value stipulated in relevant survey manual.
Missing/incorrect net opening values (20% of samples).	Estimated using one of three models using depth, gear as factors or using value stipulated in relevant survey manual.
Mix of accepted and historic species names and/or synonyms.	All species assigned their unique 'accepted ' WoRMSaphia code
Species recorded outside known geographic range.	Referred to data provider for checking. Erroneous identifications corrected. Otherwise, if supported by evidence ID retained, if no supporting evidence, species ID replaced with genus/family ID code and subsequently changed to most likely Species ID code using kNN procedure (see below).
Multiple length measurement types (total length, fork length, pre-anal length, etc.) and length measurement units (cm, mm) used.	All lengths converted to 'total length' measured to 1 cm below.
Recorded length outside known minimum and maximum length range for the species recorded.	Referred to data for check and erroneous species ID or length measurements altered. Otherwise extreme lengths retained if supported by taxonomic evidence or length > $0.6L_{min}$ or < $1.4L_{max}$. If no supporting evidence and length < $0.6L_{min}$ or > $1.4L_{max}$, species ID assumed correct and length altered to 1.1 L_{max} or 0.9 L_{min} as appropriate.
Multiple abundance measures used.	All abundances altered to actual numbers in the catch, then numbers per square kilometre of area swept in the trawl determined.
Recorded species ID code is not a species-level code, is either a genus-level or family-level code.	On a survey by survey basis, kNN procedure applied to assign most likely species-level code, or to replace all species-level codes in the genus or family to the coarser taxonomic resolution genus-level or family-level ID code.
No numbers at length data recorded, just a species count.	On a survey by survey basis, kNN procedure applied to assign most likely length frequency distribution.

2.1 Data and code accessibility

These Groundfish Survey Monitoring and Assessment Data Products and documentation are available from Marine Scotland's data publishing portal[2] the and the code is accessible on GitHub[3]. It is the degree of consultation with NDPs, the extent of review by experts involved in survey operations, data management, and assessment analysis (4 feedback loops, Figure 2.2), and the documentation describing the process (Moriarty *et al.*, 2017; Greenstreet and Moriarty 2017a; 2017b), that separates the data products described here from those produced previously by individual scientists pursuing personal research programmes.

2.2 Ongoing work

As part of the WKSAE-DATRAS workshop, the Sampling Information from the Groundfish Survey Monitoring and Assessment Data Products was rerun to include the most up-to-date data from DATRAS. Details of this process are given in section 6 below, and a code repository for further development of this product by the ICES community was created on GitHub at: https://github.com/ices-tools-dev/MSFD-QA-GFSM-A-DP.

2.3 References

Greenstreet, S.P.R., and Moriarty, M. 2017a. OSPAR Interim Assessment 2017 Fish Indicator Manual (Relating to Version 2 of the Groundfish Survey Monitoring and Assessment Data Product). Scottish Marine and Freshwater Science, Vol. 8 No. 17: 83 pp. https://doi.org/10.7489/1985-1.

Greenstreet, S.P.R. and Moriarty, M. 2017b. Manual for Version 3 of the Groundfish Survey Monitoring and Assessment Data Product. Scottish Marine and Freshwater Science, Vol. 8 No. 18: 77 pp. https://doi.org/10.7489/1986-1.

ICES Database of Trawl Surveys (DATRAS), (2017). ICES, Copenhagen.

Moriarty, M., Greenstreet, S.P.R. and Rasmussen, J. (2017). Derivation of Groundfish Survey Monitoring and Assessment Data Products for the Northeast Atlantic. Scottish Marine and Freshwater Science, Vol. 8 No. 16: 240 pp. https://doi.org/10.7489/1984-1.

Moriarty, M., Greenstreet, S.P.R., Rasmussen, J. and de Boois, I. 2019. Assessing the state of demersal fish to address formal ecosystem-based management needs: making fisheries independent trawl survey data 'fit for purpose'. Frontiers in Marine Science, Vol. 6: 162 pp. https://doi.org/10.3389/fmars.2019.00162.

[2] https://data.marine.gov.scot

[3] https://github.com/MarineScotlandScience/MSFD-QA-GFSM-A-DP/releases

3 ICES DATRAS approach for estimating missing data of swept-area

DATRAS calculates the products from DATRAS exchange data based on provided algorithms by national experts.

3.1 Background

In 2015, ICES Data Centre and the NS-IBTS working group joined forces to create DATRAS swept-area products for the North Sea, later WGBEAM joined and also the IBTSWG working group for NEA-IBTS. Currently, the process involves a FlexFile for every survey and includes all surveys and participating countries. NS-IBTS and NEA-IBTS FlexFile is based on Haul based information, for the Beam Trawl Surveys, the FlexFile is based on species and length instead of by haul, and is called cpue by swept-area. An R code to calculate the swept-area based products is also under development.

3.2 Parameters considered for each survey

The national experts from each country are responsible for the data and have full control to modify data and/or change the algorithm. The ICES Data centre apply the changes in SQL and then ask countries to check the product directly from the download page. A list of the parameters used in the calculation is shown in Table 3.1. The Beam Surveys parameters are different because it is based on HH and cpue is an additional column.

Table 3.1. List of parameters included in every survey.

IBTS	IBTS –NE Atl	BEAM
[ShootLat]	[ShootLat]	[Gear]
[ShootLong]	[ShootLong]	[DataType]
[HaulLat]	[HaulLat]	[NoPerHaul]
[HaulLong]	[HaulLong]	[HaulDur]
[SweepLngt]	[SweepLngt]	[Distance]
[Depth]	[Depth]	[LngtCode]
[HaulVal]	[HaulVal]	[LngtClass]
[HaulDur]	[HaulDur]	[SubFactor]
[Warplngt]	[Warplngt]	[HLNoAtLngt]
[DoorSpread]	[DoorSpread]	
[WingSpread]	[WingSpread]	
[Distance]	[Distance]	
[GroundSpeed]	[GroundSpeed]	

3.3 FlexFile calculated fields for NS-IBTS and NEA-IBTS

An example of the FlexFile for NS-IBTS and NEA-IBTS is provided below (Table 3.2). Data are calculated when raw data are missing and algorithms are applied to fill in missing values. The final data products from the FlexFile are SweptAreaDSKM2, SweptAreaWSKM2 and also Cal_Distance, Cal_Doorspread and Cal_ Wingspread.

Table 3.2. Example of how a FlexFile for calculated fields for NS-IBTS and NEA-IBTS. Exchange data are in black and calculated data are in green. The flags are observedvs.calculated value. DSKM2 (door spread/km²) WSKM2 (wing spread/km²)

Survey	*RecordType*	*Quarter*	*Country*	*Gear*	*Ship*	*HaulNo*	*Year*	*Month*	*Day*	*TimeShot*	*DepthStratum*	*HaulDur*	*DayNight*	*ShootLat*	*ShootLong*	*StatRec*
NS-IBTS	HH	1	DK	GOV	26D4	1	2021	2	2	1114	-9	30	D	57.8413	10.2951	44G0
NS-IBTS	HH	1	NL	GOV	64T2	1	2021	1	26	738	-9	30	D	54.3641	5.6065	37F5
NS-IBTS	HH	1	FR	GOV	35HT	1	2021	1	20	822	-9	31	D	50.9162	1.6161	30F1
NS-IBTS	HH	1	SE	GOV	77SE	1	2021	1	20	931	-9	30	D	57.8872	11.1885	44G1
NS-IBTS	HH	1	DE	GOV	06NI	1	2021	1	27	726	-9	30	D	54.619	6.8908	38F6

ICESArea	SweepLngt	Depth	HaulVal	DataType	Warplngt	DoorSpread	WingSpread	Distance	GroundSpeed	**Cal_DoorSpread**	**DSflag**	**Cal_WingSpread**	**WSflag**	**Cal_Distance**	**DistanceFlag**
27.3.a	60	83	V	R	360	92.6		3663	4	**92.6**	**O**	**24.9**	**C**	**3663**	**O**
27.4.b	60	42	V	R	235	68		3362	4	**68**	**O**	**17**	**C**	**3362**	**O**
27.7.d	50	25	V	C	99	45	15.3	3585	3.7	**45**	**O**	**15.3**	**O**	**3585**	**O**
27.3.a	60	64	V	C	200	78	20	3389	3.6	**78**	**O**	**20**	**O**	**3389**	**O**
27.4.b	50	39	V	C	175	62		3702	4	**62**	**O**	**17.4**	**C**	**3702**	**O**

SweptAreaDSKM2	*SweptAreaWSKM2*
0.3392	0.0914
0.2286	0.0571
0.1613	0.0549
0.2643	0.0678
0.2295	0.0644

3.4 FlexFile calculated fields for Beam Trawl Surveys (BEAM)

Calculated products for BEAM are different because it has some additional calculation fields such as cpue (Table 3.3).

Table 3.3. Example Swept-area and cpue data products for BEAM surveys.

Survey	Quarter	Country	Ship	Gear	SweepLngt	GearEx	DoorType	StNo	HaulNo	Year	Month	Day	TimeShot	DepthStratum	HaulDur	DayNight	ShootLat	ShootLong	HaulLat
BTS	3	GB	7.40E+10	BT4A	-9	-9	-9	1	75	2020	7	21	735	-9	20	D	49.427	-0.0602	49.4488

HaulLong	StatRec	Depth	HaulVal	StdSpecRecCode	BySpecRecCode	DataType	Netopening	Rigging	Tickler	Distance	Warplngt	TowDir	WindDir	WindSpeed	SwellDir	SwellHeight	ICESArea
-0.0587	2.70E+10	18	V	1	1	R	-9	FM	0	2340	54	352	100	12	-9	-9	27.7.d

AphiaID	Species	SpecVal	Sex	SubFactor	LngtClass	HLNoAtLngt	NoPerHaul	**BeamWidth**	**DistanceDerived**	**CPUE_number_per_hour**	**SweptArea_km2**	**CPUE_number_per_km2**
105814	Sciliorhinus caniculus	1	F	1	230	1	1	**4**	**2340**	**3**	**0.01**	**106.84**
105814	Scyliorhinus canicula	1	F	1	270	1	1	**4**	**2340**	**3**	**0.01**	**106.84**

4 Available DATRAS time-series

4.1 Beam Trawl Surveys

The year ranges from BEAM depend on which survey is used and are listed in Table 4.1. For further details see Annex 5.

Table 4.1. Year ranges for the BEAM surveys (BTS, SNS, DYFS and BTS-VIII).

Years	BTS-BE	BTS-DE	BTS-GB	BTS-NL	BTS-VIII	DYFS	SNS
1985				x			
1987				x			
1988				x			
1989				x			
1990			x	x			
1991			x	x			
1992			x	x			
1993		x	x	x			
1994		x	x	x			
1995		x	x	x			
1996			x	x			
1997		x	x	x			
1998		x	x	x			
1999		x	x	x			
2000		x	x	x			
2001		x	x	x			
2002		x	x	x		x	x
2003		x	x	x		x	
2004	x	x	x	x		x	x
2005		x	x	x		x	x
2006			x	x		x	x
2007	x	x	x	x		x	x
2008		x	x	x		x	x
2009		x	x	x		x	x
2010	x	x	x	x		x	x
2011	x	x	x	x	x	x	x
2012	x	x	x	x	x	x	x
2013	x	x	x	x	x	x	x
2014	x	x	x	x	x	x	x
2015	x	x	x	x	x	x	x
2016	x	x	x	x	x	x	x
2017	x	x	x	x	x	x	x
2018	x	x	x	x	x	x	x
2019	x	x	x	x	x	x	x
2020			x		x	x	x

4.2 Northeast Atlantic IBTS surveys

National experts from the NEA-IBTS have provided year ranges for each survey to the revised and the current time-series for the different surveys is listed in Table 4.2.

Table 4.2. Summary of surveys on the Northeast Atlantic IBTS surveys area with time-series with the number of hauls present in the FlexFiles for each of them.

Years	EVHOE	FR-CGFS	IE-IAMS	IE-IGFS	NIGFS (1,2)	ROCKALL	SCOROC	SCOWCGFS(1)	SP-ARSA(1,4)	SP-NORTH	SP-PORC	SWC-IBTS(1)
2004	138			159					48		79	145
2005	143			140	1	38			23		80	149
2006	127			168	2	32			23		88	134
2007	145			171	2	42			78		98	151
2008	147			166	53	37			82		83	124
2009	135			164	121	41			83		80	131
2010	139			176	120				80		80	59
2011	151			159	119		45	112	82		85	
2012	130			172	119		36	130	70		79	
2013	140			176	112		31	92	83		85	
2014	155			170	113		47	121	85		80	
2015	148	73		147	121		42	120	86	136	85 (165)(3)	
2016	157	73	107	172	124		48	123	89	134	85 (165)(3)	
2017	25	66	109	149	120		41	117	89	136	80	
2018	155	73	116	153	122		41	116	86	130	83	
2019	149	65	129	161	122		44	124	89	130	79	
2020	156	59	70	127	116		40	113	89	123	84	

1) **In the case of NIGFS, SCOWCGFS, SWC-IBTS and SP-ARSA columns data from Q1 and Q4 are mixed.**
2) **NIGFS 08 only Q1, 09–20 Q1–4.**
3) **SP-PORC 2015 and 2016 are doubled; just additional hauls are unique, while the standard hauls have two double records copies.**
4) **SP-ARSA 07–20: Are the two quarters, 1–4 in just one cell. 04 from two quarters but incomplete, 05 only from Q1, 06 from Q4.**

Within the Terms of Reference, WKSAE-DATRAS is to propose strategies to reduce missing data in the crucial variables to estimate swept-area as a prerequisite for biodiversity assessments. One of the problems with these estimations is that the needed parameters: door spread, wing spread, and vertical opening information (for swept volume) are not available for all hauls.

These parameters are usually related with depth and are also related between them, so it is possible to use regression models to estimate the door spread for a given vessel and gear with the data of the depth at what the haul occurs, a similar relation with depth exists with the wing spread and the net opening.

Finally, in some surveys there are numerous occasions in which only one of the parameters, Door spread or wing spread, are available (either because only one set of net sensors was available or because the sensors do not worked properly, see Table 4.3 and Table 4.4).

Table 4.3. Number of records with missing data on wing spread of hauls present in the FlexFiles per survey and year (- are years with no data at all because there was no survey or are not considered due to different gears vessels).

Years	EVHOE	FR-CGFS	IE-IAMS	IE-IGFS	NIGFS	ROCKALL	SCOROC	SCOWCGFS	SP-ARSA	SP-NORTH	SP-PORC	SWC-IBTS
2004	107	-	-	4	-	-	-	-	0	-	75	1
2005	8	-	-	77	1	0	-	-	23	-	83	2
2006	15	-	-	26	2	0	-	-	6	-	89	15
2007	58	-	-	12	2	0	-	-	15	-	89	1
2008	3	-	-	32	54	0	-	-	0	-	68	6
2009	107	-	-	34	122	1	-	-	0	-	84	2
2010	17	-	-	27	121	-	-	-	0	-	82	2
2011	151	-	-	28	119	-	2	0	0	-	88	-
2012	130	-	-	2	126	-	0	2	0	-	79	-
2013	7	-	-	6	112	-	0	0	1	-	89	-
2014	155	-	-	3	115	-	0	3	0	139	83	-
2015	14	76	-	4	103	-	0	0	0	5	85	-
2016	2	0	0	6	66	-	0	0	0	14	49	-
2017	0	0	0	5	120	-	0	0	3	138	6	-
2018	0	0	0	1	121	-	0	1	1	132	88	-
2019	0	0	2	4	103	-	0	1	0	5	2	-
2020	0	0	0	1	116	-	0	2	89	8	7	-

Table 4.4. Number of records with missing data on door spread of hauls present in the FlexFiles per survey and year (- are years with no data at all because there was no survey or are not considered due to different gears vessels).

Years	EVHOE	FR-CGFS	IE-IAMS	IE-IGFS	NIGFS	ROCKALL	SCOROC	SCOWCGFS	SP-ARSA	SP-NORTH	SP-PORC	SWC-IBTS
2004	10	-	-	0	-	-	-	-	48	-	15	1
2005	18	-	-	0	1	1	-	-	23	-	10	3
2006	10	-	-	0	1	0	-	-	23	-	11	14
2007	55	-	-	0	0	0	-	-	78	-	7	7
2008	7	-	-	0	1	0	-	-	82	-	24	6
2009	96	-	-	0	0	1	-	-	83	-	2	6
2010	22	-	-	0	0	-	-	-	80	-	0	2
2011	151	-	-	0	0	-	2	4	82	-	3	-
2012	130	-	-	1	0	-	0	3	70	-	1	-
2013	2	-	-	1	0	-	1	1	80	-	4	-
2014	0	-	-	3	0	-	0	3	16	14	3	-
2015	0	0	-	0	0	-	1	0	86	5	0	-
2016	9	0	0	0	0	-	0	0	1	13	1	-
2017	26	0	0	1	1	-	1	0	3	3	2	-
2018	0	0	0	12	0	-	0	0	1	3	2	-
2019	9	0	0	2	0	-	0	0	0	3	2	-
2020	10	0	0	0	0	-	0	0	0	8	5	-

Considering this aim, plots are presented for:

- Wing spread vs. the door spread;
- Wing spread and depth;
- Door spread and depth;
- Net vertical opening and depth.

In the case of wing spread vs. door spread a linear regression model is estimated in R using lm() as lm.WingVsDoor<-lm(WingSpread~DoorSpread).

While in the rest of the relationships of parameters vs. depth a non-linear logarithmic model: nls(WingSpread~a1+b1*log(Depth) is estimated in R.

Figures show relationships between gear parameters for each survey. A solid blue line shows fitted models. Dashed lines show 95% confidence intervals. Years of data for each model vary based on data availability and model fitting considerations and are shown in the bottom left of each plot. Model equations are fitted excluding the 2020 data year (except for door spread equations in the Northern Ireland Groundfish Surveys, see below), and 2020 data points are shown as filled circles for reference.

There are three surveys (Scottish Western Coast Groundfish survey, Northern Ireland Groundfish Survey and the Spanish Survey on the Gulf of Cádiz), that are usually performed biannually, with surveys in the 1st and the 4th quarters, in these case models presented are fitted combining the results from both quarters combined.

4.2.1 West Coast Scottish Groundfish Survey (Q1 and Q4) and quarters combined

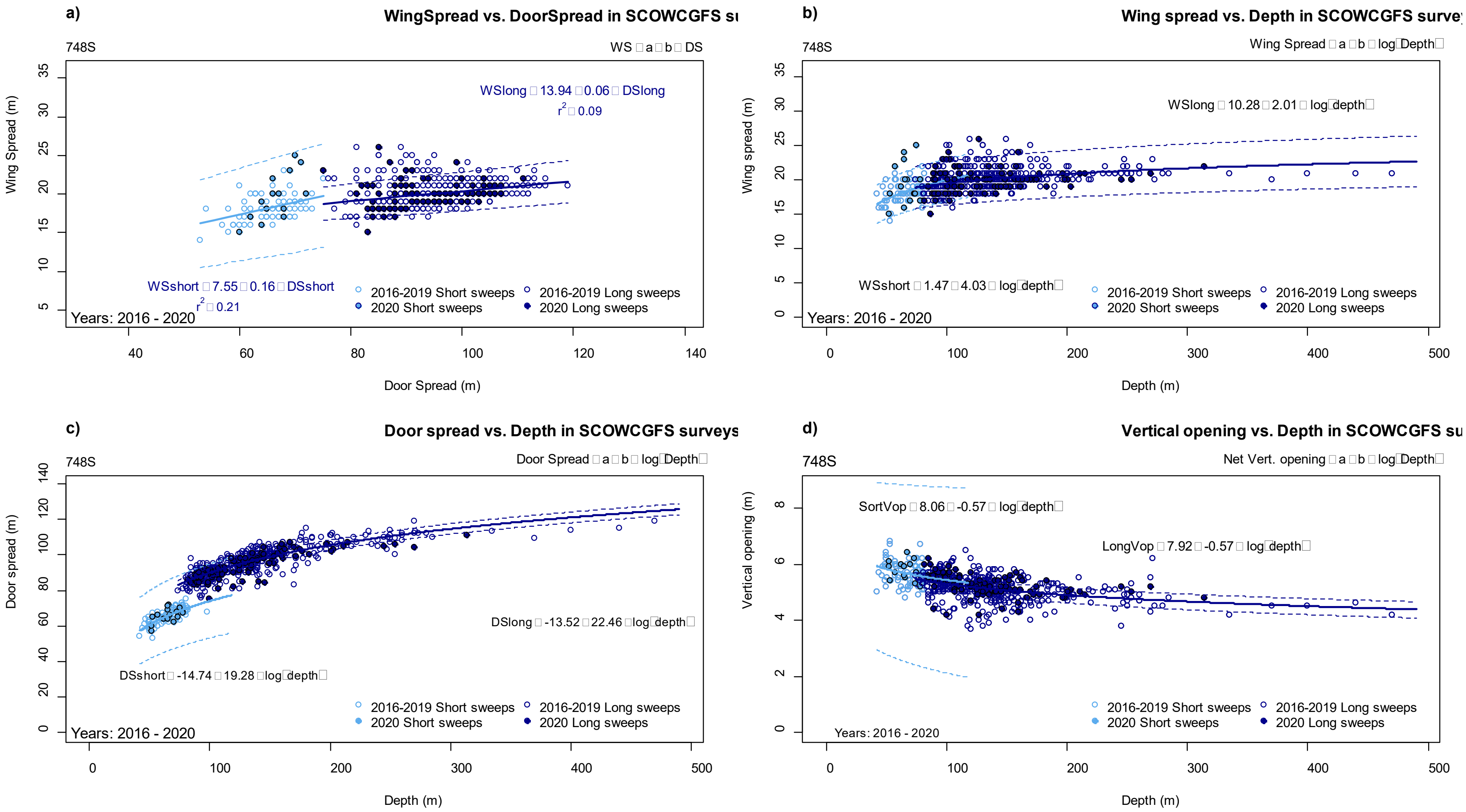


Figure 4.1. Scottish groundfish survey on the west coast (1st and 4th quarters combined) gear parameters showing relationships between a) wing spread vs. door spread, b) wing spread vs. depth, c) door spread vs. depth and d) net vertical opening and depth. The solid light blue line shows fitted models when using short sweeps while dark blue lines so models with long sweeps. Dashed lines show 95% confidence intervals. Years of data for each parameter are 2016–2020 and model fitting are done excluding the last year shown.

4.2.2 Scottish Groundfish survey on Rockall Q3

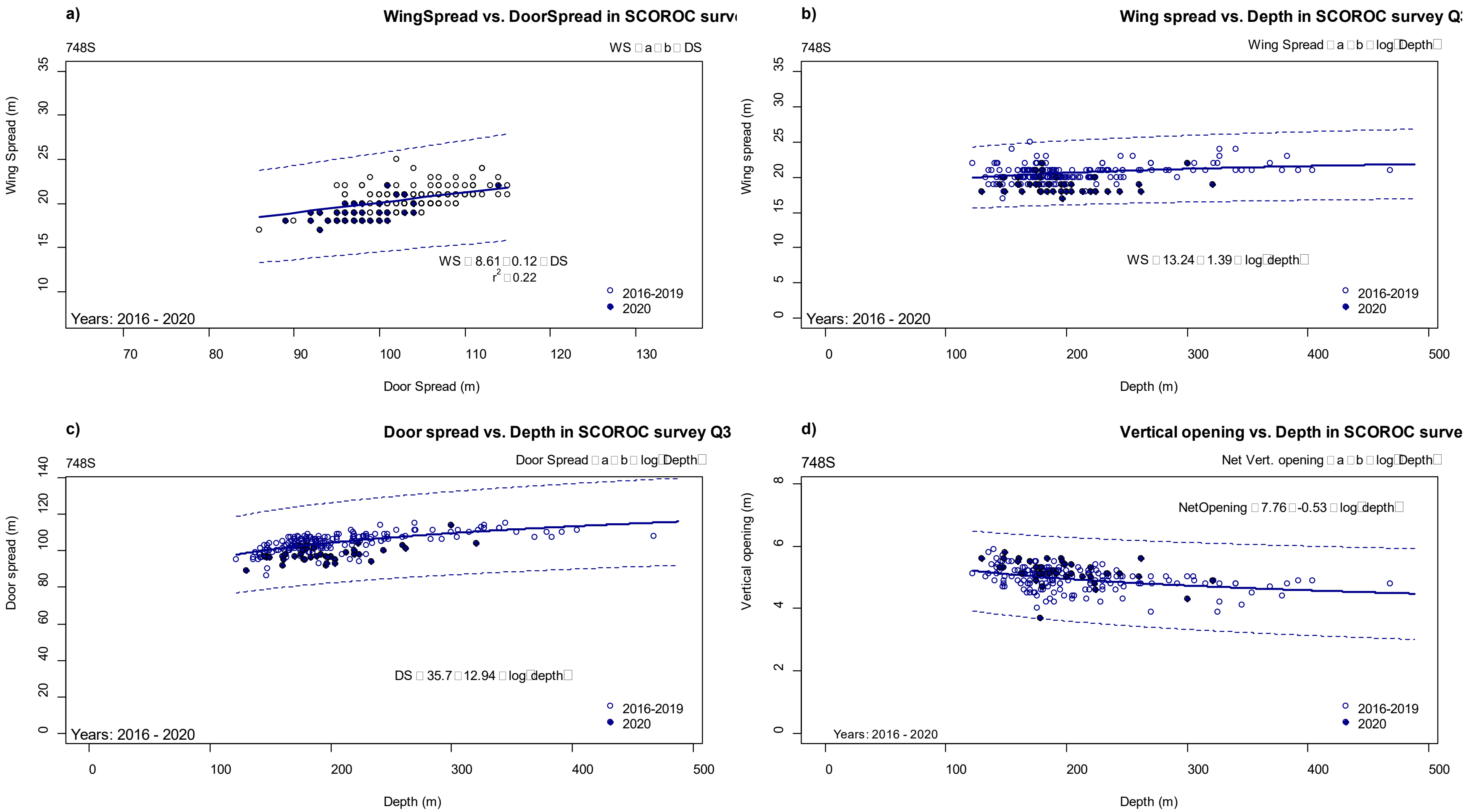


Figure 4.2. Scottish groundfish survey on Rockall 3rd quarter gear parameters, showing relationships between a) wing spread vs. door spread, b) wing spread vs. depth, c) door spread vs. depth and d) net vertical opening and depth. Solid light blue line shows fit.

4.2.3 Northern Ireland Groundfish Survey (Q1 and Q4) and quarters combined

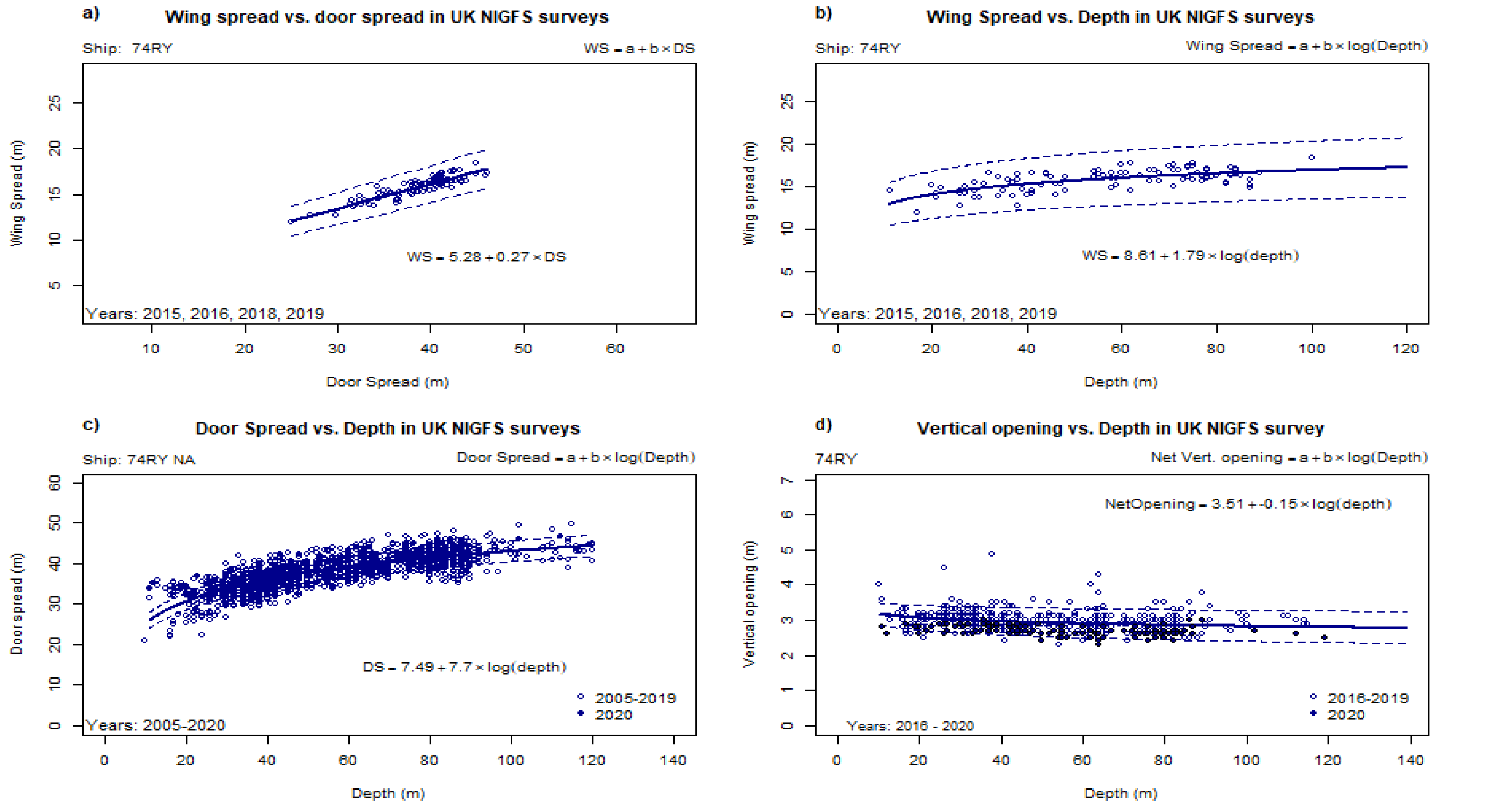


Figure 4.3. Northern Ireland groundfish survey combining data from 1st and 4th quarters gear parameters showing relationships between a) wing spread vs. door spread, b) wing spread vs. depth, c) door spread vs. depth and d) net vertical opening and depth. A solid dark blue line shows fitted models. Dashed lines show 95% confidence intervals. Years shown used to fit models for each parameter depend on data available in DATRAS and excluding the last year shown.

For NIGFS a warp length plot is not shown because DATRAS does not contain warp length data for this survey. Wing spread models are based on the years 2015, 2016, 2018 and 2019 because these were the only years for which wing spread data were available. Similarly, vertical net opening information is available for both quarters only from 2016–2020.

Door spread-depth relationships for NIGFS are based on the full DATRAS time-series (2005–2020). This represents a slight change from the time frame used in the Manual of the IBTS North Eastern Atlantic Surveys (2017), although the fitted equation is similar. The reason for this change is due to a slight trend towards a higher door spread in shallow depths over time (see Figs below). This shift is considered minor and adding 'Year' to the fitted equation only increased R2 by only 0.05. For model parsimony, because only very shallow hauls are affected, and consistency across regions it was therefore decided to continue to use the equation Door spread = $\alpha + \beta \times \log(\text{Depth})$, but to use the entire timespan to better reflect the earlier and later years in the time-series.

4.2.4 Irish Groundfish Survey Q4

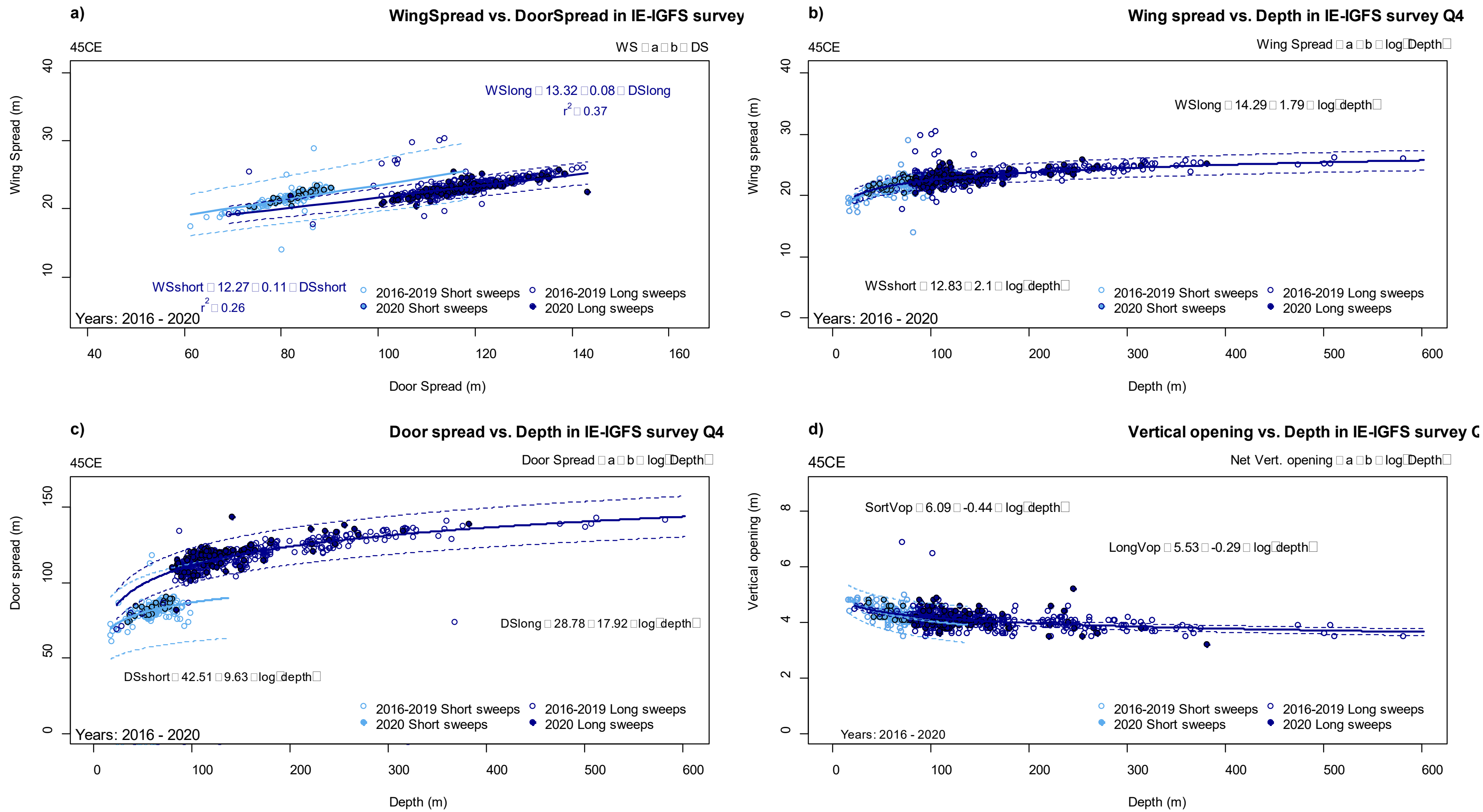


Figure 4.4. Irish groundfish survey 4th quarter gear parameters showing relationships between a) wing spread vs. door spread, b) wing spread vs. depth, c) door spread vs. depth and d) net vertical opening and depth. A solid light blue line shows fitted models when using short sweeps while dark blue lines so models with long sweeps. Dashed lines show 95% confidence intervals. Years of data for each parameter are 2016–2020 and model fitting are done excluding the last year shown.

4.2.5 Irish anglerfish and megrim Groundfish Survey Q1

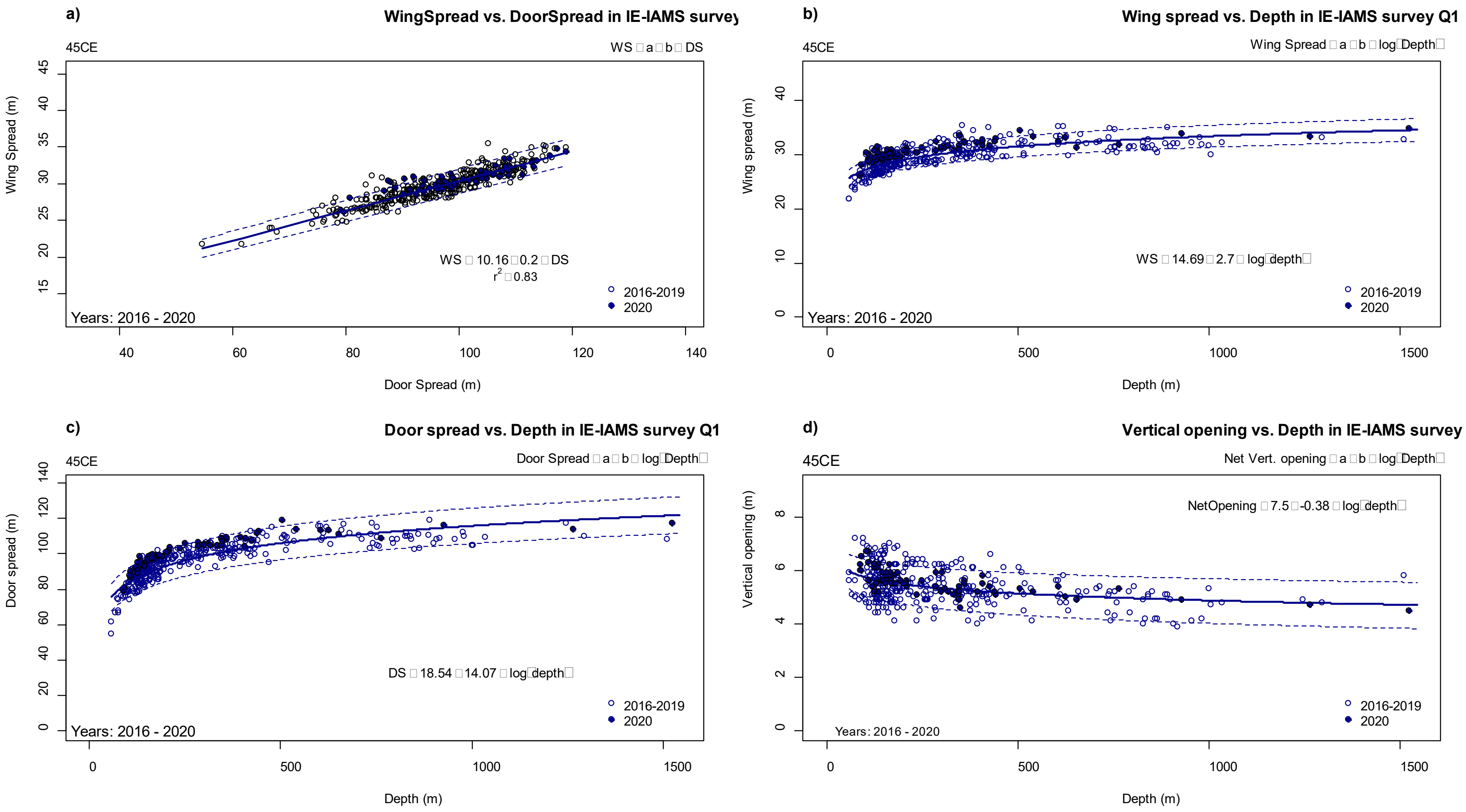


Figure 4.5. Irish angler and megrim groundfish survey 1st quarter gear parameters showing relationships between a) wing spread vs. door spread, b) wing spread vs. depth, c) door spread vs. depth, and d) net vertical opening and depth. Solid dark blue line shows fitted models. Dashed lines show 95% confidence intervals. Years of data for each parameter are 2016–2020 and model fitting are done excluding the last year shown.

4.2.6 Spanish Groundfish Survey on Porcupine Bank Q3

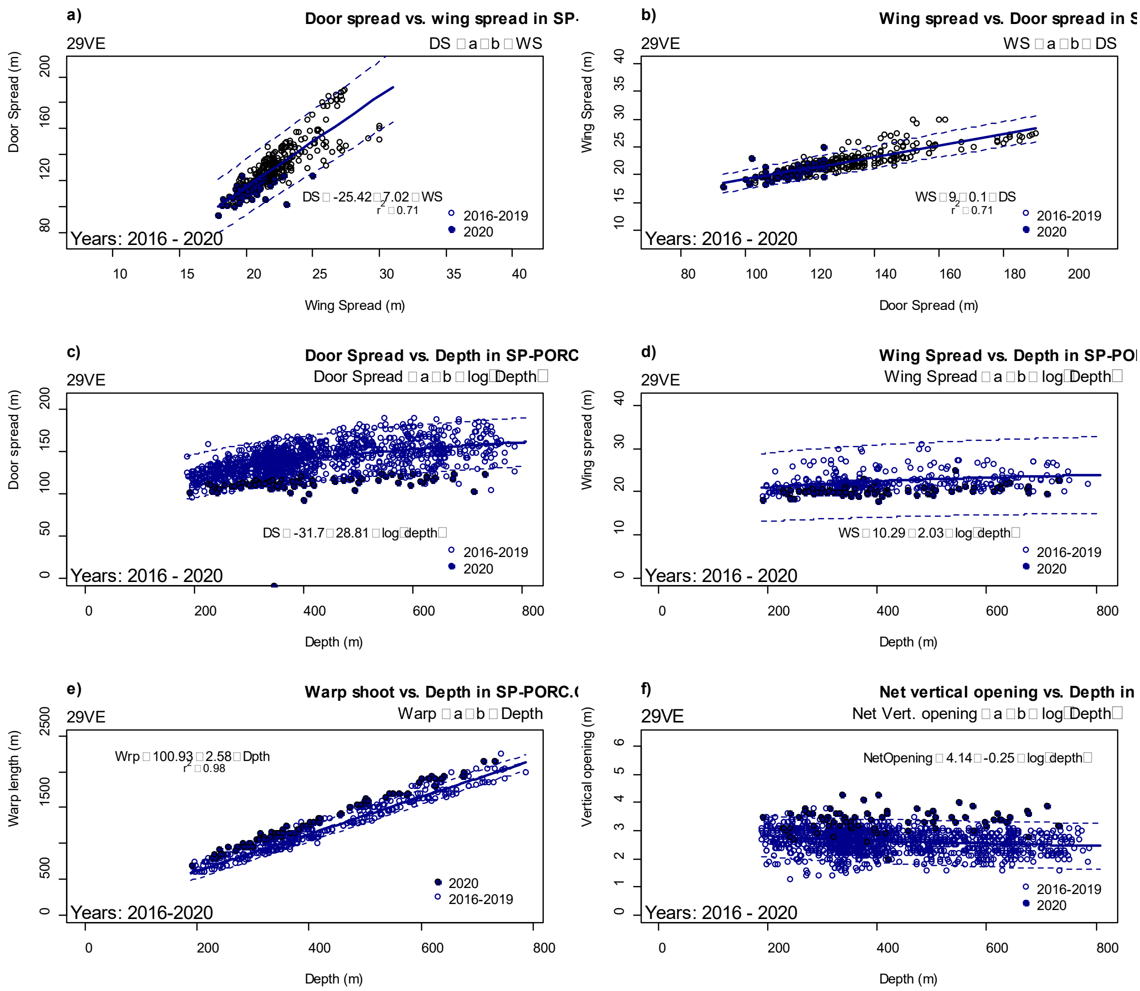


Figure 4.6. Spanish groundfish survey on the Porcupine bank 3rd quarter gear parameters showing relationships between a) door spread vs. wing spread, b) wing spread vs. door spread, c) door spread vs. depth, d) wing spread vs. depth, d) warp shot vs. depth and f) net vertical opening and depth. A solid blue line shows fitted models. Dashed lines show 95% confidence intervals. Years of data for each parameter are 2014–2020 and model fitting are done excluding the last year shown.

4.2.7 French Groundfish Survey on the Channel Q4

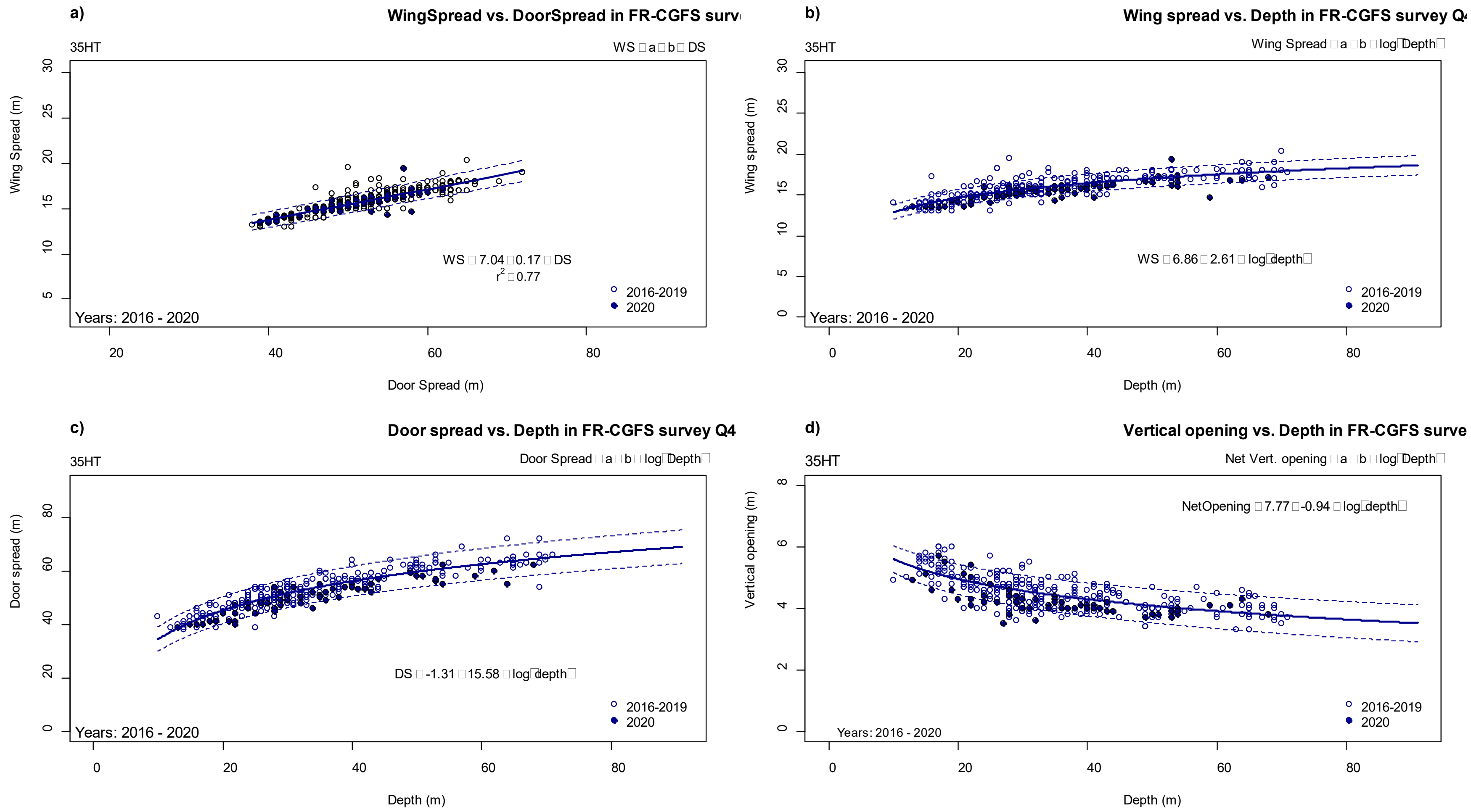


Figure 4.7. French groundfish survey on the channel 4th quarter gear parameters showing relationships between a) wing spread vs. door spread, b) wing spread vs. depth, c) door spread vs. depth and d) net vertical opening and depth. A solid dark blue line shows fitted models. Dashed lines show 95% confidence intervals. Years of data for each parameter are 2016–2020 and model fitting are done excluding the last year shown.

4.2.8 French EVHOE Survey Q4

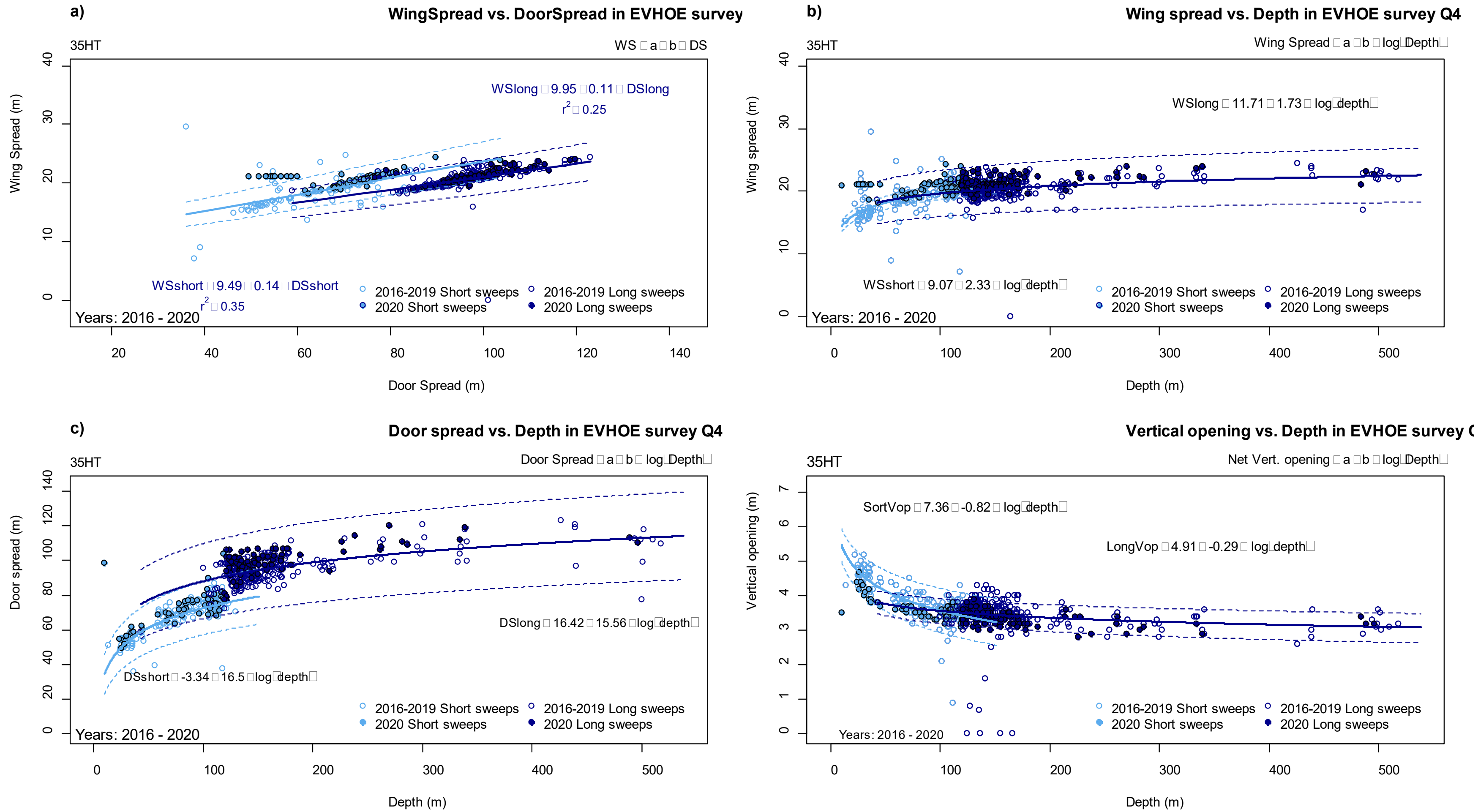


Figure 4.8. French EVHOE groundfish survey on the 4th quarter gear parameters showing relationships between a) wing spread vs. door spread, b) wing spread vs. depth, c) door spread vs. depth and d) net vertical opening and depth. A solid light blue line shows fitted models when using short sweeps while dark blue lines so models with long sweeps. Dashed lines show 95% confidence intervals. Years of data for each parameter are 2016–2020 and model fitting are done excluding the last year shown.

4.2.9 Spain Northern Shelf Groundfish Survey Q4

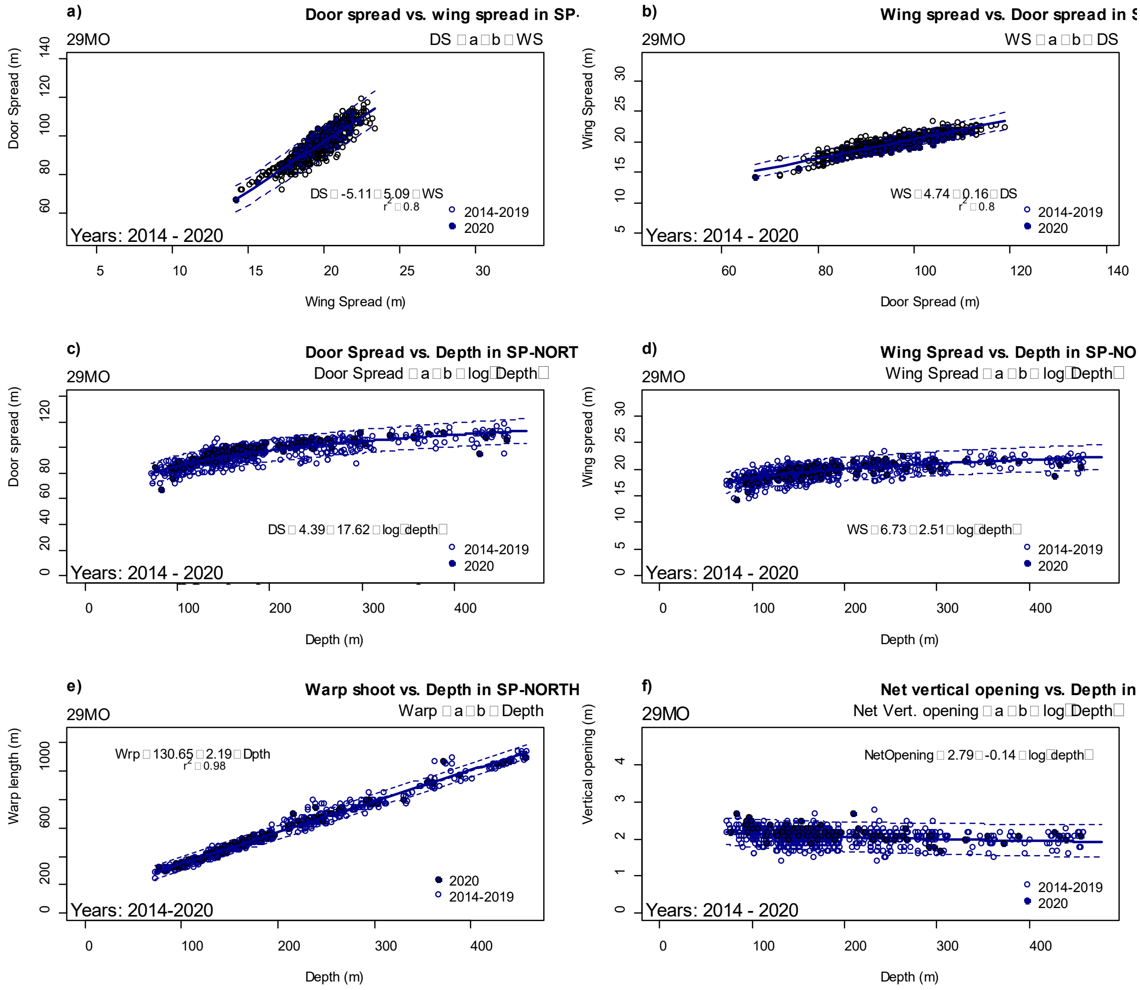


Figure 4.9. Groundfish survey on the northern Spanish shelf (SP-NGFS) 4th quarter gear parameters showing relationships between a) door spread vs. wing spread, b) wing spread vs. door spread, c) door spread vs. depth, d) wing spread vs. depth, d) warp shot vs. depth and f) net vertical opening and depth. A solid blue line shows fitted models. Dashed lines show 95% confidence intervals. Years of data for each parameter are 2014–2020 and model fitting are done excluding the last year shown.

4.2.10 Portuguese Groundfish Survey Q4

For the last years PT-IBTS-Q4 has made the effort to get geometry data but was not able to have used a DoorSpread sensor, therefore only WingSpread and Net Opening for years 2005, 2014, 2015, 2017 and 2018, and also only for NCT gear. The second gear, CAR is only used with an old ship, RV “Capricórnio” or when NCT is severely damaged and unusable.

Manual data for PT-IBTS nets are, for NCT, WingSpread = 15.10 m, DoorSpread = 45.70 m and Net Opening = 4.6 m, while for CAR parameters to be used are set as : WingSpread = 24 m, DoorSpread = 60m and NetOpening = 2.2 m.

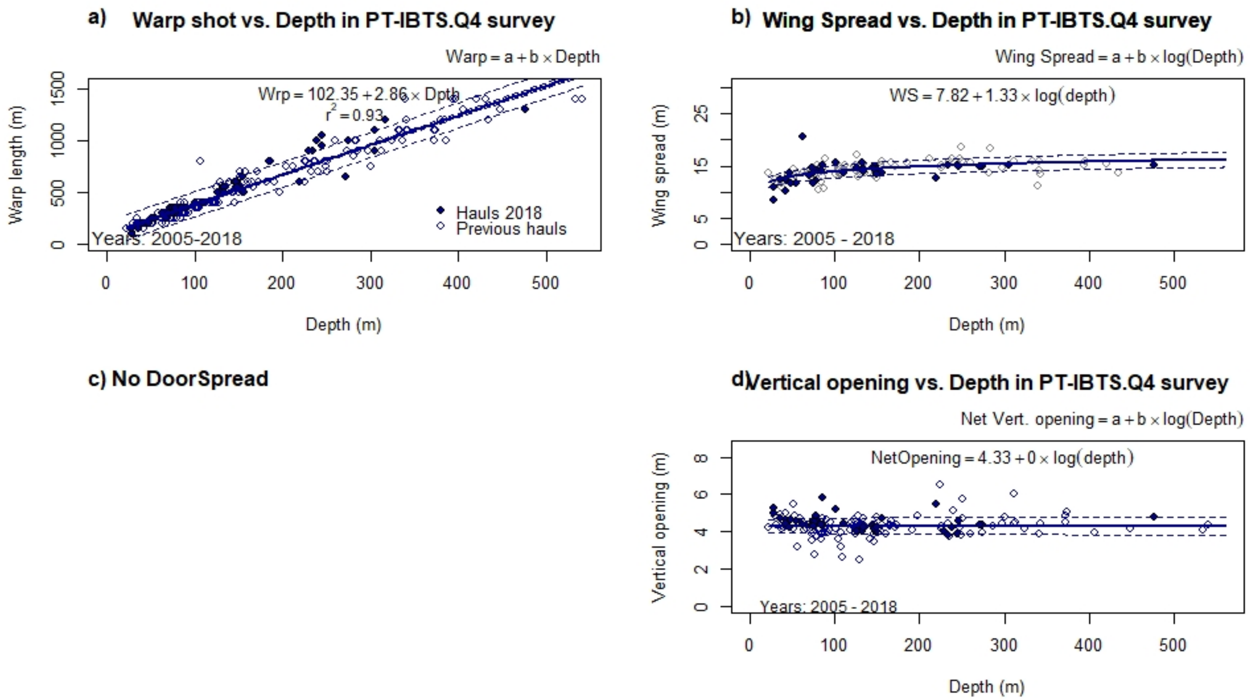


Figure 4.10. Portuguese PT-IBTS Quarter 4 Groundfish Survey gear parameters showing relationships between a) warp shot and depth, b) wing spread and depth, d) vertical opening and depth. A solid blue line shows fitted models. Dashed lines show 95% confidence intervals. Years of data for each model vary based on PT-IBTS data availability and model fitting considerations and are shown in the bottom left of each plot.

4.2.11 Spain–Gulf of Cádiz (ARSA) Groundfish Survey (Q1 and Q4) and quarters combined

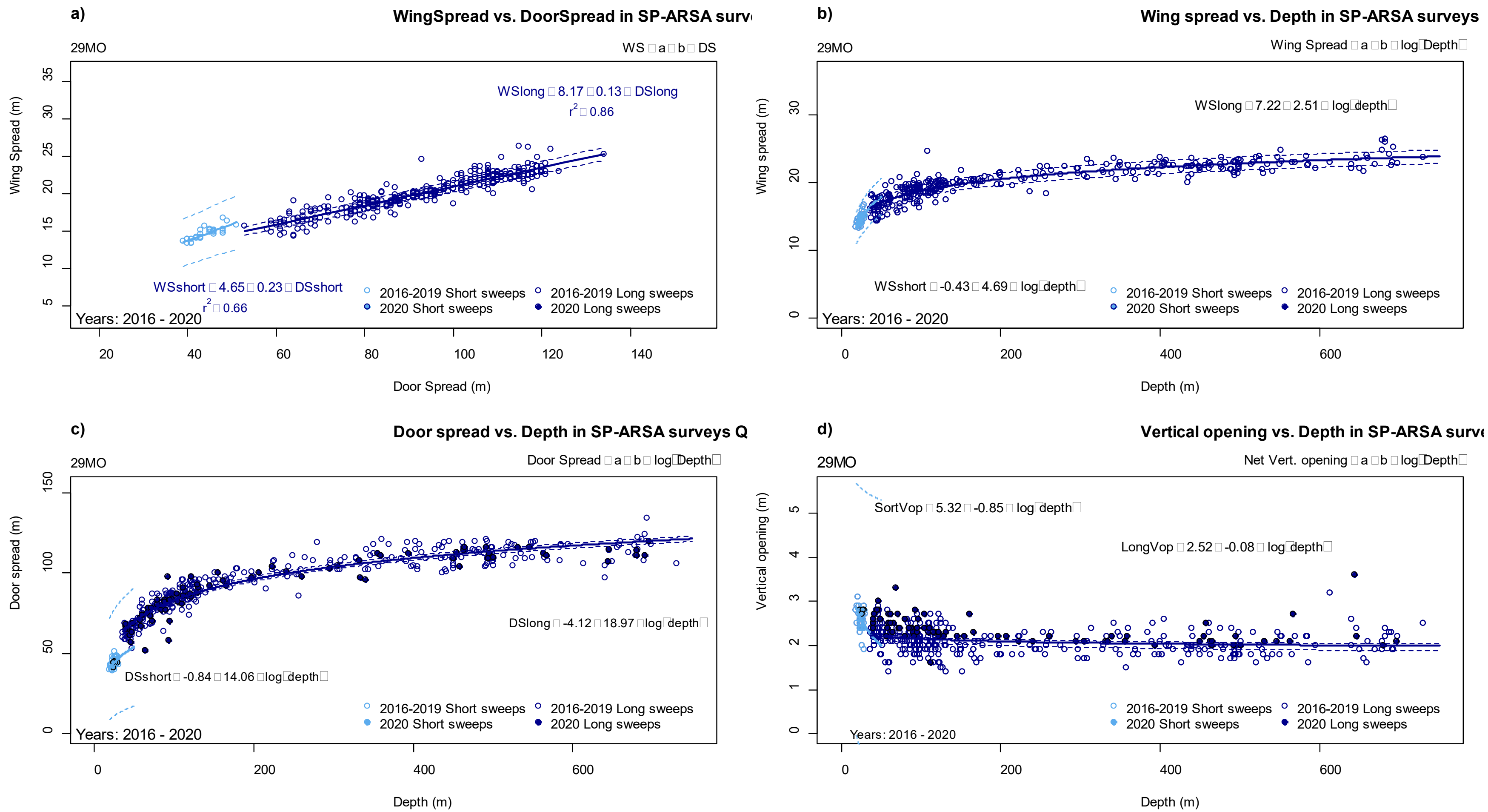


Figure 4.11. SP-ARSA Groundfish Survey on the Gulf of Cádiz, both quarters (1 and 4) gear parameters showing relationships between a) wing spread vs. door spread, b) wing spread and depth, c) door spread vs. depth, d) vertical opening and depth. A solid blue line shows fitted models. Dashed lines show 95% confidence intervals. Years of data for each parameter are 2016–2020 and model fitting are done excluding the last year shown.

4.3 North Sea IBTS

During the IBTSWG meeting in 2013, the national representatives for the North Sea IBTS agreed to clean up their data required for the calculation of swept-area back in time to 2004 and provide country and vessel specific algorithms for estimating missing values (ICES 2013). Most of this work was done until the IBTSWG meeting in 2015, and based on the results the ICES Data Centre established a so-called FlexFile (see section 3) containing the information for swept-area based on door spread and for swept-area based on wing spread covering the period from 2004 to present (ICES 2019). The DATRAS dataset, although not thoroughly cleaned in every case, has also been available for a Marine Science Scotland project in which the time-series has recently been updated for the period 1983 to 2019 (see section 2).

However, the most recent version of FlexFile still contains some erroneous data and, more importantly, parts or even entire surveys are missing because either basic information hasn't been submitted to DATRAS (Table 4.5) or country and vessel specific algorithms are lacking (Table 4.6). Hence, a series of actions have to be performed (Table 4.7) before the FlexFile can be completed.

Table 4.5. Missing basic information in DATRAS exchange data (HH records) by country and vessel for the period 2004 to present (DATRAS download from 7 May 2021).

			Hauls with							
HH records (Total number of hauls: 12505)			missing distance		missing ground speed		missing door spread		missing wing spread	
Country	Vessel		n	%	n	%	n	%	n	%
FR	35HT	Thalassa	0	0.00	348	26.40	70	5.31	78	5.92
	64T2	Tridens	0	0.00	5	100.00	0	0.00	5	100.00
DE	06NI	Walther Herwig	6	0.39	0	0.00	99	6.51	279	18.34
	26D4	Dana	0	0.00	0	0.00	0	0.00	52	100.00
DK	26D4	Dana	465	28.79	0	0.00	39	2.41	1372	84.95
GB	74E9	Endeavour	0	0.00	0	0.00	61	4.74	158	12.29
GB-SCT	748S	Scotia	0	0.00	0	0.00	9	0.36	30	1.22
NL	64T2	Tridens	18	1.85	0	0.00	65	6.68	973	100.00
	74E9	Endeavour	1	1.05	0	0.00	3	3.16	33	34.74
NO	58G2	GO Sars	82	16.27	82	16.27	86	17.06	389	77.18
	58J3	Johan Hjort	132	27.85	132	27.85	134	28.27	474	100.00
	58UO	Kristine Bonnevie	0	0.00	55	25.70	18	8.41	143	66.82
	58AA	Håkon Mosby	271	100.00	271	100.00	271	100.00	271	100.00
	74E9	Endeavour	0	0.00	0	0.00	13	32.50	40	100.00
SE	26D4	Dana	0	0.00	0	0.00	2	0.26	357	46.36
	77AR	Argos	0	0.00	0	0.00	6	0.91	662	100.00
	77MA	Mimer	0	0.00	0	0.00	4	9.30	43	100.00
	77SE	Svea	0	0.00	0	0.00	0	0.00	0	0.00
	Overall:		975	7.80	893	7.14	880	7.04	5359	42.85

Table 4.6. Missing information in the NS-IBTS DATRAS FlexFile by country and vessel for the period 2004 to present (Flexfile download from 1 June 2021; *: blanks and zero values excluded, ?: missing depth, -: missing distance and/or algorithm).

FlexFile (Total number of hauls: 11961)			SweptAreaDSKM2: Number of			SweptAreaWSKM2: Number of			Number of valid
Country	Vessel		valid records *	missing hauls		valid estimates *	missing hauls		hauls in HH records
FR	35HT	Thalassa	1293	25	?	1318	0		1318
	64T2	Tridens	5	0		5	0		5
DE	06NI	Walther Herwig	1521	0		1521	0		1521
	26D4	Dana	110	0		110	0		110
DK	26D4	Dana	1615	0		1615	0		1615
GB	74E9	Endeavour	1286	0		1286	0		1286
GB-SCT	748S	Scotia	2468	0		2468	0		2468
NL	64T2	Tridens	973	0		973	0		973
	74E9	Endeavour	95	0		95	0		95
NO	58G2	GO Sars	423	81	-	0	504	-	504
	58J3	Johan Hjort	330	144	-	0	474	-	474
	58UO	Kristine Bonnevie	0	214	-	0	214	-	214
	58AA	Håkon Mosby	0	271	-	0	271	-	271
	74E9	Endeavour	0	40	-	0	40	-	40
SE	26D4	Dana	770	0		770	0		770
	77AR	Argos	662	0		662	0		662
	77MA	Mimer	43	0		43	0		43
	77SE	Svea	136	0		136	0		136
	Overall:		11729	776		893	11612		12505

Table 4.7. Status of available swept-area information in the NS-IBTS DATRAS for WKSAE and required actions. Missing information NS-IBTS, update for WKSAE based on DATRAS download of HH records 7/5-2021 and FlexFile version 1/6-2021 (All ok: no obvious outliers detected but missing values can occur).

a) Q1 surveys 2004 to 2021

Year(s) Q1	Country/Vessel missing in FlexFile	Distance	Door spread	Other (WS, depth, SOG etc.)	Action needed
2004-2008	NOR/Håkon Mosby: completely missing in FlexFile	NOR: Not measured, Neither haul end position nor SOG available, i.e. distance cannot be calculated	NOR: no data in actual HH records, previous submissions with 3 missing observations (for tows with long (110m) sweep length only	NOR: Sweeplngt missing, no WS data in actual HH	NOR: to check what is possible to provide the missing information and provide updated algorithms for estimating DS and WS dependent on short/long sweeps and strapping/no strapping
2009	NOR/GO Sars: no swept-area estimates in FlexFile	NOR: no data in actual HH records, Neither haul end position nor SOG available, i.e. distance cannot be calculated	NOR: no data in actual HH records, only calculated values in FlexFile	NOR: no WS data in actual HH records, Sweeplngt missing	NOR: to check whether information on distance, haul end position and/or SOG can be made available
2010-2011	All ok in HH records, missing values calculated in FlexFile except: NOR/GO Sars: WS missing FRA/Thalassa: DS missing for 25 hauls	All ok	All ok	NOR: no WS data in actual HH records FRA 2010: missing depth for 25 hauls	NOR: to provide algorithm for calculation of missing WS ICES Data Centre: estimate depth and SweptAreaDS for the 25 missing for 1Q2010
2012	NOR/GO Sars: Swept-area missing in FlexFile for all tows	NOR: Not measured, Neither haul end position nor SOG available, i.e. distance cannot be calculated	All ok	NOR: no WS data in actual HH records	NOR: to check whether information on distance, haul end position and/or SOG can be made available

Year(s) Q1	Country/Vessel missing in FlexFile	Distance	Door spread	Other (WS, depth, SOG etc.)	Action needed
2013-2016	All ok in HH records, missing values calculated in FlexFile except: NOR/GO Sars: WS missing	All ok	All ok except: GER 2015: 1 outlier*	NOR: no WS data in actual HH records SWE 2014: 1 outlier for WS*	NOR: to provide algorithm for calculation for missing WS
2017	NOR/ENDW (Endeavour) completely missing in FlexFile (although some information available in HH records for distance and door spread)	All ok	All ok except: NOR: 2 outliers	NOR: no WS data in actual HH records	NOR: to check 2 outliers DS=1 and DS=102, re-submission of HH records), NOR: to provide algorithm for calculation for missing DS and WS Data Centre to add NOR/Endeavour in FlexFile
2018–2021	All ok in HH records, missing values calculated in FlexFile except: NOR/GO Sars: WS missing	All ok in HH records	All ok in HH records All ok in HH records except: GER 2021: 1 outlier*	NOR: no WS data in actual HH records SCO 2018: 1 outlier for depth (280m)*	NOR: to provide algorithm for calculation for missing WS

(*: data have been corrected during WKSAE or will be corrected soon including new submission of exchange data)

b) Q3* surveys 2004 to 2020[4]

Year(s) Q3	Country/Vessel missing in FlexFile	Distance	Door spread	Other (WS, depth, SOG etc.)	Action needed
2004-2005	NOR/Håkon Mosby: surveys completely missing in FlexFile, Q4 tows (2004) missing in FlexFile	NOR: Not measured, Neither haul end position nor SOG available, i.e. distance cannot be calculated	NOR: no data in actual HH records	NOR: no WS data in HH records, Sweeplngt missing	NOR: to check what is possible to provide the missing information and provide algorithms for estimating DS and WS dependent on short/long sweeps and strapping/no strapping

4

Year(s) Q3	Country/Vessel missing in FlexFile	Distance	Door spread	Other (WS, depth, SOG etc.)	Action needed
2006- 2008	NOR/Johan Hjort: Swept-area missing in FlexFile for all tows, Q2 tows (2006 and 2007) missing in FlexFile	NOR: Not measured, Neither haul end position nor SOG available, i.e. distance cannot be calculated	NOR: no door spread data in HH records	NOR: no WS data in HH records, Sweeplngt missing	NOR: to check what is possible to provide the missing information and check short/long sweeps and strapping/no strapping
2009	All ok	All ok	All ok		
2010- 2011	NOR/Johan Hjort: WS missing in FlexFile, Q2 tows (2011) missing in FlexFile	All ok	All ok	NOR: no WS data in HH records	NOR: to check algorithms for estimating DS and WS dependent on short/long sweeps and strapping/no strapping
2012	NOR/Johan Hjort: WS missing in FlexFile	All ok	All ok	NOR: 1 outlier for depth (415m), no WS data in HH records	NOR: to check outlier for depth, provide algorithm for estimating WS
2013	NOR/Johan Hjort: DS missing for 1 tow WS missing for all tows	All ok	All ok except: NOR: 1 missing value (**not** calculated in FlexFile) at which most likely strapping was used		NOR: to provide algorithms for estimating DS and WS for short sweeps dependent on strapping/no strapping
2014	NOR/Johan Hjort: WS missing in FlexFile	All ok except: NOR: 1 potential outlier (haul 362: 1236 m for 30 min tow)	All ok	NOR: no WS data in HH records	NOR:to check distance for haul 362, NOR:to provide algorithms for estimating DS and WS for short sweeps dependent on strapping/no strapping
2015	NOR/Johan Hjort: WS missing in FlexFile Note: 7 valid tows were conducted in Q2 but area labelled Q3 in HH records	All ok	All ok	NOR: no WS data in HH records	NOR:to provide algorithms for estimating WS for short sweeps dependent on strapping/no strapping

Year(s) Q3	Country/Vessel missing in FlexFile	Distance	Door spread	Other (WS, depth, SOG etc.)	Action needed
2016	NOR/Johan Hjort: 1 erroneous value for DS (calculated)in FlexFile WS missing in FlexFile	All ok	All ok	NOR: no WS data in HH records,	Data Centre: to correct NOR/Johan Hjort erroneous DS value NOR:to provide algorithms for estimating WS for short sweeps dependent on strapping/no strapping
2017	NOR/Kristine Bonnevie:completely missing in FlexFile(although data in HH records for distance and door spread)	All ok	All ok	NOR: no WS data in HH records	NOR: to provide algorithm for estimating DS and WS values
2018	NOR/Kristine Bonnevie:completely missing in FlexFile(although data in HH records for distance, door and wing spread)	All ok	All ok	NOR: missing DS and WS data	NOR: to provide algorithm for estimating DS and WS values
2019	NOR/Kristine Bonnevie:completely missing in FlexFile(although data in HH records for distance, door and wing spread)	All ok	All ok	NOR: several outliers for WS	NOR to check WS data and to provide algorithm for missing WS values
2020	NOR/ Kristine Bonnevie: completely missing in FlexFile although data in HH records for distance and door spread	All ok	All ok	NOR: no WS data in HH records	NOR to provide algorithm for missing WS values

All countries made considerable effort to improve the quality of the data needed for the calculation of swept-area in recent years, and for both, door spread and wing spread the number of missing observations declined drastically since 2004 (Figure 4.12). This was most successfully for door spread whereas a high amount of wing spread data are still missing. The reason for the latter is that some countries are still lacking additional distance sensors for measuring wing spread which works sufficiently well with their trawl monitoring systems or because the received data are quite noisy (Figure 4.13) and false recordings cannot always easily be filtered out. Furthermore, wing spread is directly linked to door spread and can be estimated from door spread using country/vessel linear relationships once the data are cleaned and separated for sweep length category (Figure 4.13).

The North Sea IBTS provides indices as fishery-independent information for stock assessment for target species. Absolute estimates are not required for this aspect. Currently, stock assessments use indices in number/hour and for a few species (cod, plaice) the country/vessel differences in trawl geometry are included as covariates in GAM models (Berg *et al.,* 2014). Using swept-area based indices would also allow accounting for the differences in trawl geometry between countries and vessels affecting catchability. Here, swept-area based on door appears preferable to swept-area based on wing spread (ICES 2019) considering the high amount of values, which have to be imputed based on more or less well-defined algorithms for estimating wing spread (see section 5 for some examples).

Preliminary results indicate that there is no obvious effect of depth on the ratio of door and wing spread based swept-area and that the ratio has been stable over time (Figure 4.14). Hence, door spread based swept-area can also be used for non-herding species for which otherwise inconsistencies in length and, subsequently, age distribution as used in stock assessment may occur. This consideration, however, warrants more in-depth analyses (see section 6 for further discussion).

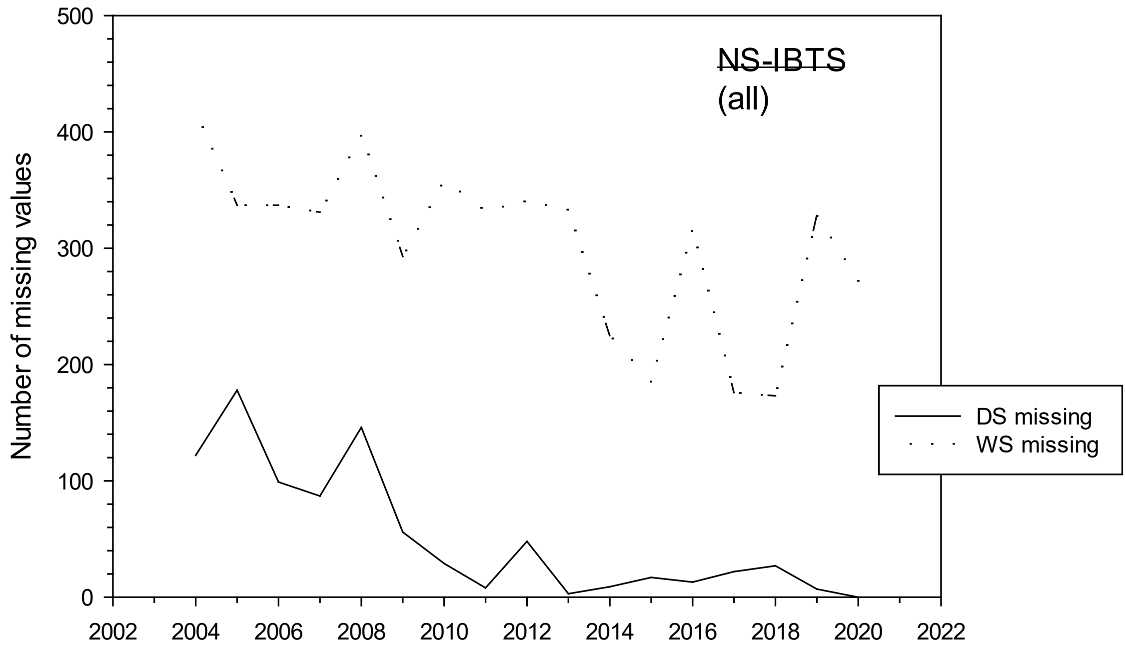


Figure 4.12. Number of missing observations of door and wing spread in the NS-IBTS 2004–2020 (HH records download 7/5-2021, data for both quarters pooled).

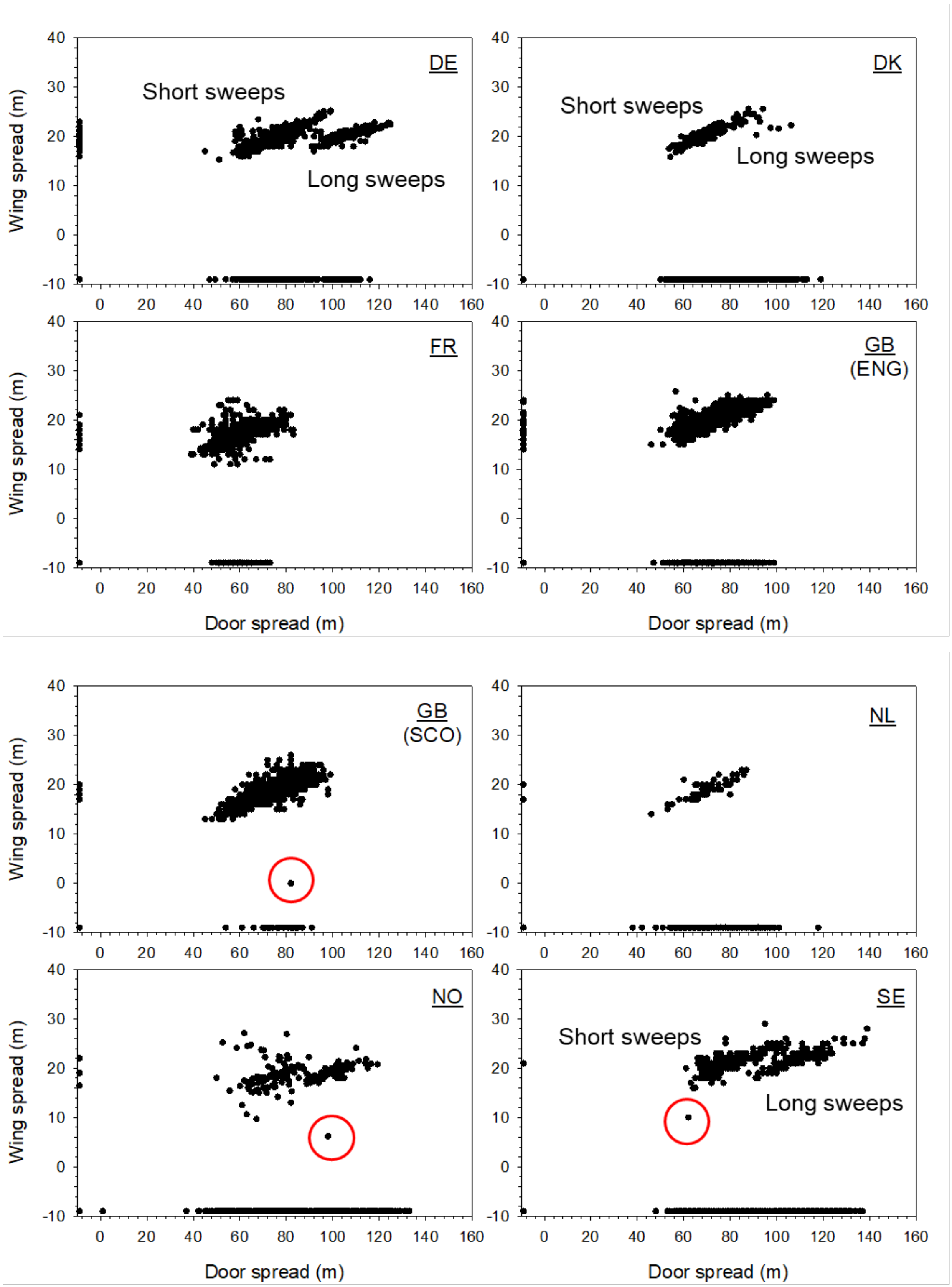


Figure 4.13. Relationship between door and wing spread by country (NS-IBTS HH records 2004–2021, download 7/5-2021).

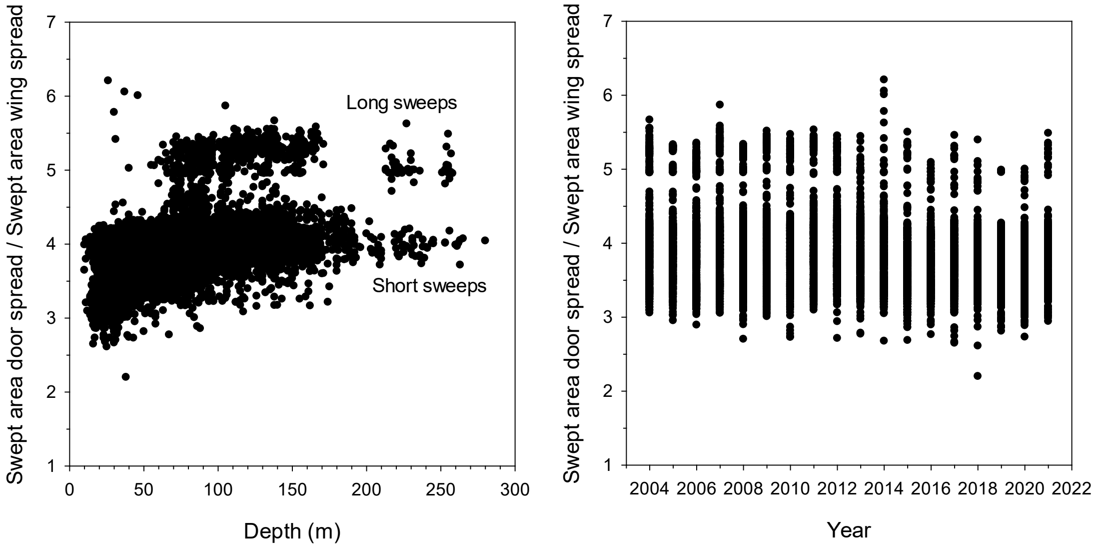


Figure 4.14. Ratio of door and wing spread swept-area in relation to depth and year (NS-IBTS FlexFile, download 1/6-2021).

4.4 References

Berg, C. W., Nielsen, A., Kristensen, K. (2014). Evaluation of alternative age-based methods for estimating relative abundance from survey data in relation to assessment models. Fisheries research, 151, 91–99.

ICES 2013. ICES. 2013. Report of the International Bottom Trawl Survey Working Group (IBTSWG), 8–12 April 2013, Lisbon, Portugal. ICES CM 2013/SSGESST:10. 272 pp.

ICES 2019. Workshop on Methods to develop a swept-area based effort index (WKSABI). ICES Scientific Reports. 1:3. 24 pp. http://doi.org/10.17895/ices.pub.4902.

5 Data quality checks and examples for estimating missing values comparing MSS/OSPRAR and ICES approach: how to handle remaining missing and erroneous values?

Several outliers and potential erroneous values were detected in the data available in DATRAS during WKSAE. This included both historical and data from the most recent years, i.e. after 2014 (see section 4). The national representatives in the respective survey working groups are encouraged to check their data required for the calculation of swept-area prior to submission to DATRAS, and it is recommended that the ICES Data Centre establish tighter quality control during the submission process (Annex 7 Recommendations).

5.1 Beam Trawl Surveys

The available beam trawl data in DATRAS contains 26 996 valid hauls from 1985 to 2020, but at present only offshore data from the North Sea are included in the updated OSPAR product (Table 5.1 and Figure 5.1). The original OSPAR product also included data from the Celtic Sea, which has not yet been included in the update. Neither the French beam trawl survey in the Bay of Biscay (BTS-VIII) nor the inshore beam trawl surveys (DYFS and SNS) were included in the original OSPAR product but will hopefully be added to the updated product at a later stage.

Table 5.1. Summary of DATRAS beam trawl survey data.

	In DATRAS back to	valid hauls	additional hauls	Invalid hauls	OSPAR (updated)
BTS	1985[1]	16 492	44	446	9023
BTS-VIII	2011[2]	595	25	2	
DYFS	2002[3]	7946		229	
SNS	1985[4]	1933		22	

[1] Starting year; running since [2]2007, [3]1970 and [4]1969, respectively (ICES, 2021).

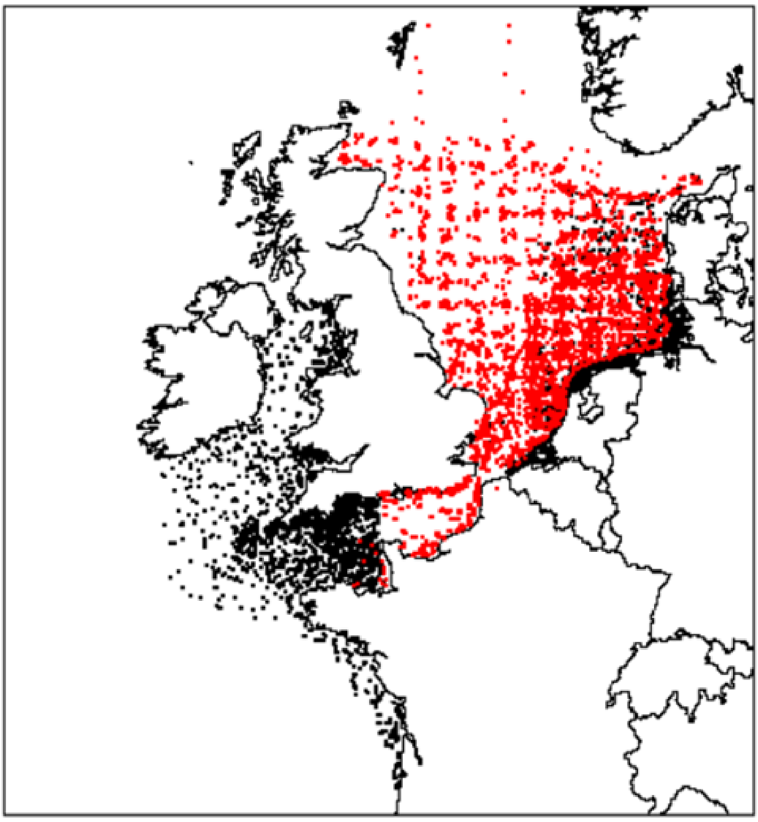


Figure 5.1. Spatial distribution of beam trawl hauls in DATRAS 1985–2020 (in black) overlaid with hauls included in the updated OSPAR product (in red).

The working group on Beam trawls surveys list three ways to calculate swept-area (ICES, 2021) which are used to produce the DATRAS swept-area estimate:

1. Swept-area in km² = beam width * distance/106
2. Swept-area in km² = beam width * (haul duration/60) * fishing speed * 1852/106
3. Calculate distance based on shooting and hauling position

If the distance is available, the first approach is used; when distance is missing, the second approach is the recommended method. When fishing speed is missing, the standard values from the WGBEAM manual (ICES 2019) can be used, or, distance can be calculated using the third approach based on shooting and hauling positions. The OSPAR product uses a similar for calculating swept-area but also includes additional checks of outliers and missing/duplicated values. As a result, the OSPAR product neither includes reported or derived zeros, or extreme outliers, in the swept-area calculations. The protocol applied also results in that approach 3 can be selected before approach 2.

6821 (75.6%) of the beam trawl hauls in the updated OSPAR product had available distances. 53 outliers (0.8%) were detected and replaced with calculated values resulting in different swept-area estimates compared to the DATRAS product (Figure 5.2). There was also a small consistent difference in calculated swept-area for the German data between the two products because different beam lengths had been applied in the calculations (7.2 m in the OSPAR product vs. 7 m in DATRAS).

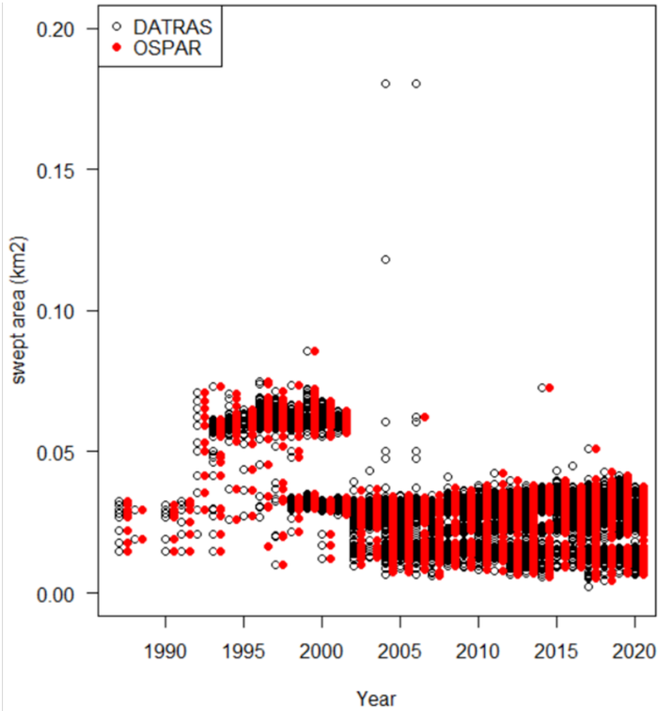


Figure 5.2. Comparison of swept-area estimates in the DATRAS and OSPAR product for cases with available distance in the database. The higher values in the early period are due to longer haul durations in the Dutch survey 1992 - 2001.

For the remaining 2202 hauls, swept-area was calculated based on hauls duration and fishing speed or from positional data. Again, most of the calculated swept-areas were similar but 76 outliers (3.5%) had been replaced with calculated values in the updated OSPAR product (Figure 5.3).

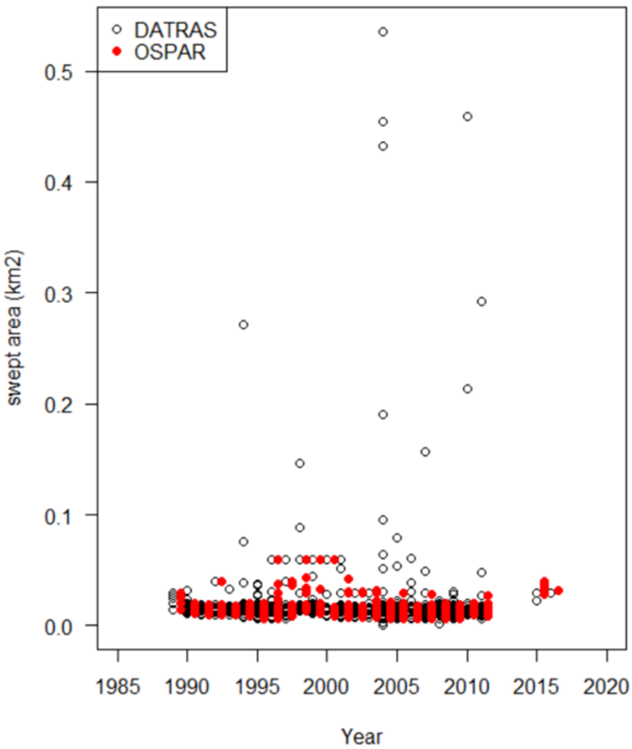


Figure 5.3. Comparison of swept area estimates in the DATRAS and OSPAR product for cases missing distance in the database.

Although the total number of outliers in the DATRAS beam trawl data appears to be low it would be useful to implement a protocol for detection, flagging and replacement of obvious outliers in the DATRAS swept-area product similar to that used for the OSPAR product.

5.1.1 References

ICES. 2019. Manual for the Offshore Beam Trawl Surveys, Version 3.4, April 2019, Working Group on Beam Trawl Surveys. 54 pp. http://doi.org/10.17895/ices.pub.5353.

ICES. 2021. Working Group on Beam Trawl Surveys (WGBEAM). ICES Scientific Reports. 3:46. 89 pp. https://doi.org/10.17895/ices.pub.8114.

5.2 Northeast Atlantic and North Sea IBTS

The ICES DATRAS algorithm for estimation missing door and wing spread for the single national surveys under the Northeast Atlantic umbrella were established first just prior to WKSAE and followed by several revisions during the workshop. Thus, no comparison between the ICES DATRAS FlexFiles and the algorithm used in the MSS/OSPRAR approach has been made.

The ICES DATRAS FlexFiles for the North Sea IBTS were established using algorithms provided by the national representatives based on data from the years 2004 to 2014/2015. However, some survey participants introduced several changes after this period:

- Modification of vessel-specific warp-depth ratio
- Drop of using long sweeps in Q1 surveys (all countries except Norway and Sweden)
- Replacement of nylon with polyethene trawls (Sweden, Norway, Denmark)
- Replacement of kite with Vonin flyers (Denmark, Germany when using RV Dana in 1Q2019 and 1Q2020)
- Introduction of new vessels (Norway: RV Kristine Bonnevie in Q3 since 2017; Sweden: RV Svea in both quarters since Q3 in 2020)
- New country vessel combinations (Sweden and Germany on RV Dana; usually using their own trawls).

None of these changes should have severely influenced gear geometry for a given country or vessel but it had so far not been checked whether the existing algorithms are appropriate cover the actual conditions. Hence, a few examples, including a comparison between the DATRAS and the MSS/OSPRA were presented at WKSAE. Here, the ICES FlexFile algorithms specified in Annex 5: and the MSS model parameters for Moriarty *et al.* (2017, see section 2) were used.

5.2.1 DE / RV Dana

The ICES FlexFile estimates are too low and the MSS algorithm for Walter Herwig usually used by Germany does not fit either. The MSS algorithm for DK Dana (short sweeps) gives values too low for depths shallower than 100 m and thus a new regression should be applied (Figure 5.4).

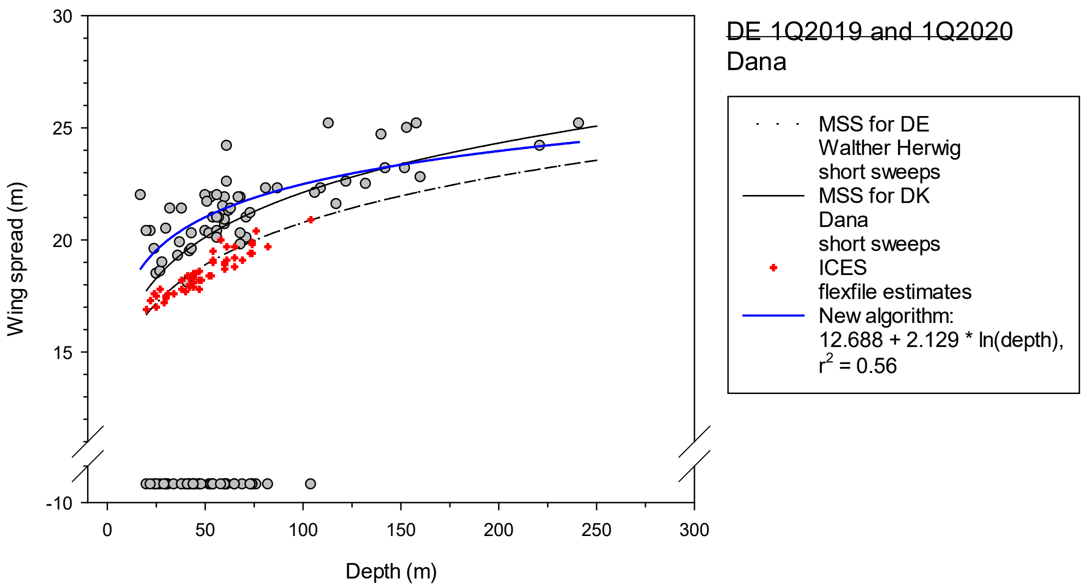


Figure 5.4. DE / RV Dana: Wing spread in relation to depth (Q1 2019 and Q1 2020, only short sweeps used).

5.2.2 DK / RV Dana

Both, the ICES and the MSS algorithms overestimate wing spread for larger depths (Figure 5.5) and new versions for short sweeps should be established preferably when more observations for depths > 90 m have become available.

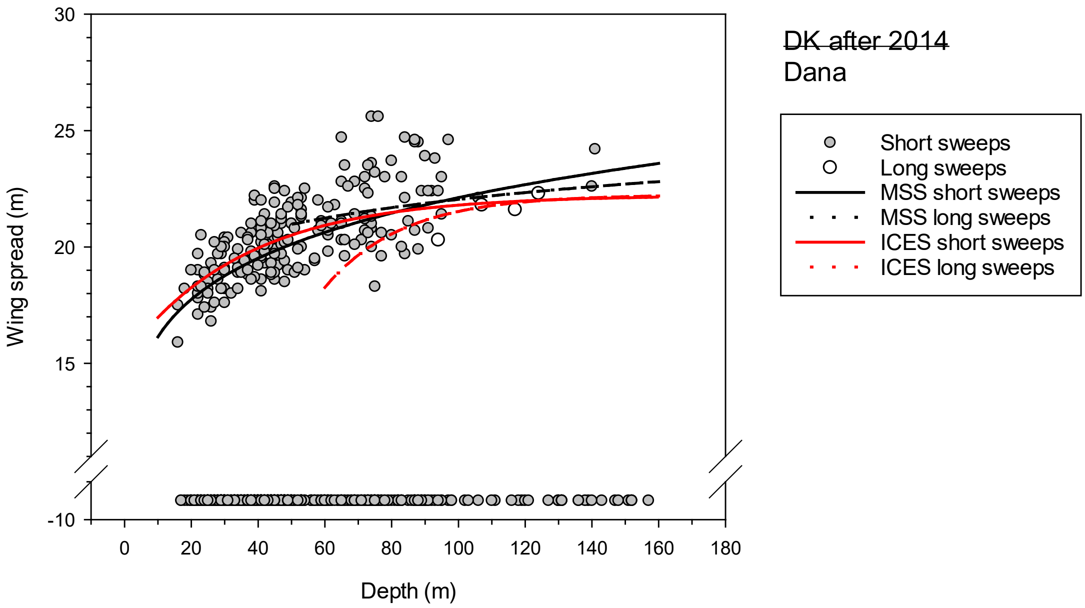
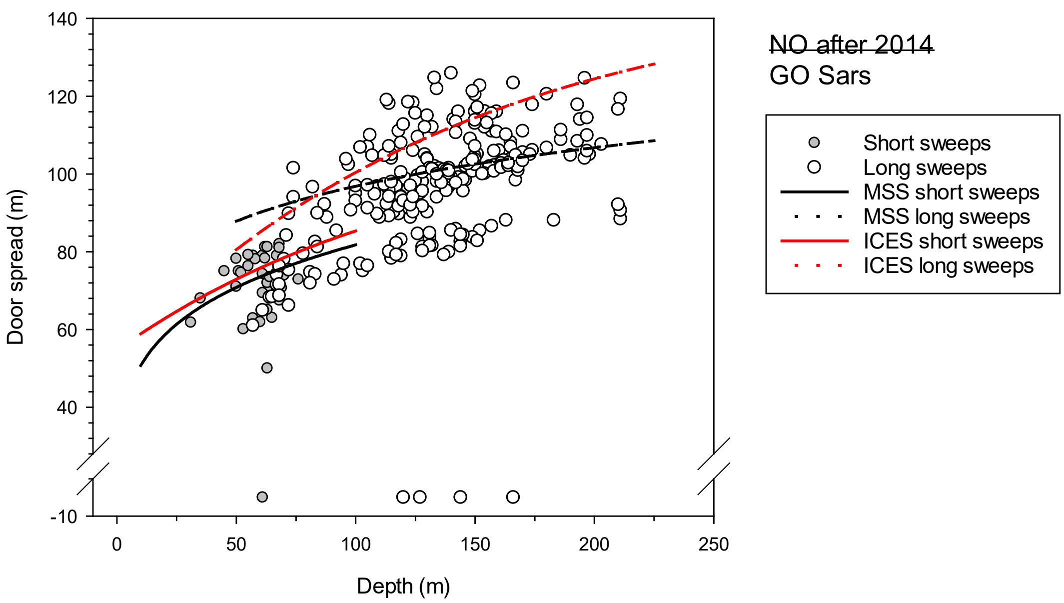


Figure 5.5. DK / RV Dana: Wing spread in relation to depth, Q1 and Q3 (long sweeps were used in 1Q 2015 for the last time).

5.2.3 NO / GO Sars

The ICES and the MSS algorithms are similar for short sweeps but differ considerably for long sweeps. However, the few missing values 1Q2015 and 1Q2018 can reasonably well be estimated with the existing ICES algorithms (Figure 5.6).



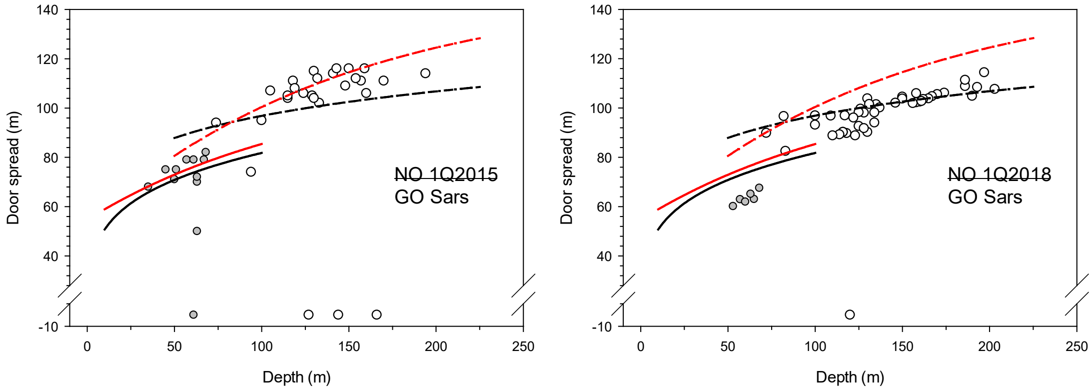


Figure 5.6. NO / GO Sars: Door spread in relation to depth, 1Q 2015 to 1Q2021.

5.2.4 NO / Kristine Bonnevie

For this vessel, neither ICES nor MSS has a specific algorithm. The MSS version for "all other vessels" is not suitable for estimating the missing door spread and a new regression is suggested (Figure 5.7). Similar, a new algorithm should be established for the wing spread calculation once the data have been cleaned for outliers and one obvious erroneous observation.

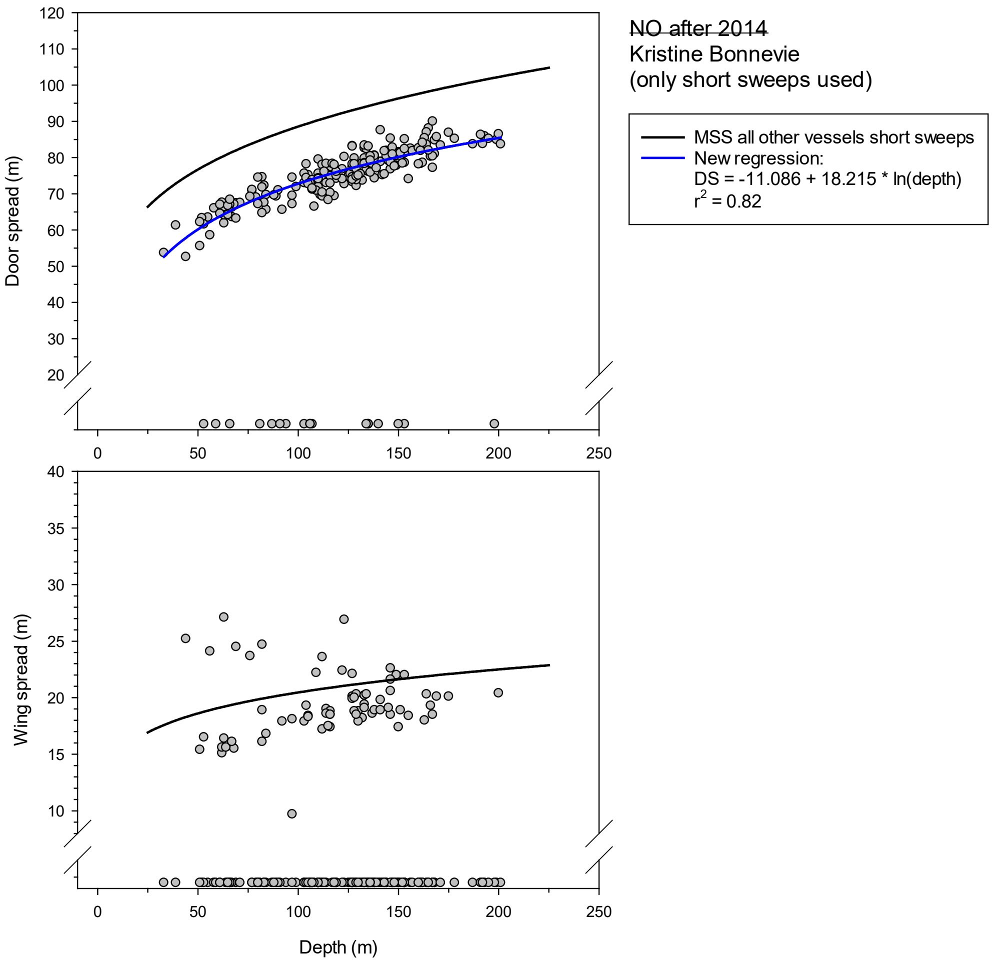


Figure 5.7. NO /Kristine Bonnevie: Door spread in relation to depth, 3Q 2017 to 3Q 2020 and wing spread in relation to depth, 3Q 2018 and 3Q 2019 (no wing spread data for 3Q 2017 and 3Q 2020 in HH record download 7/5-2021).

Until the gaps in the ICES DATRAS FlexFile are filled and for extending the time-series back in time, the use of the MSS approach is currently the only solution but this should then be based on an updated set of parameters. However, this requires an assumption on other variables for which observations are missing in the actual exchange data, e.g. speed over ground, and it appears preferable that the missing information is submitted to DATRAS and the ICES Data Centre by the national representatives as soon as possible. This is urgently needed before swept-area based indices from the North Sea IBTS can be provided for stock assessments.

6 Update to the Groundfish Survey Monitoring and Assessment Data Products 'Sampling Information' data product

In order to provide an up-to-date comparison to the DATRAS swept-area, Flexfiles a rerun and the MSS/OSPAR Groundfish Survey Monitoring and Assessment Data Product Sampling Information was conducted using the current DATRAS datasets.

This provides an alternative methodology with which to compare the DATRAS FlexFile outputs, and has the added benefit of providing a longer time-series than the current DATRAS FlexFile for many regions. Conversely, the DATRAS FlexFile data product is based on survey specific equations provided by survey coordinators, and may therefore provide more accurate estimates of missing data for the period covered by these data products.

This data update to the Groundfish Survey Monitoring and Assessment Data Products includes only the 'Sampling Information' product (see section 2 for details), based on the haul information (HH) files in DATRAS, and not the 'Biological Information' product which was beyond the remit and scope of the WKSAE workshop. DATRAS datasets were downloaded on either the 8 and 9th of June 2021 and reference copies of the downloaded data, code and data product are provided for reproducibility in the SharePoint 'Data' folder of WKSAE-DATRAS[5]

Note that this data product includes all valid hauls which met with the quality assurance criteria applied by Moriarty *et al.* (2017). However, the Standard Survey Area protocols which exclude hauls from locations that do not have consistent temporal sampling in the original MSS/OSPAR product have not been applied (i.e. this data product includes the full spatial extent of the survey data). It is expected that data users will make their own decisions about which hauls and the spatial extent to include in future workshops (e.g. WKABSENS) and other analyses.

These updated code files from the reanalysis are provided on the ices-dev-tools GitHub page: https://github.com/ices-tools-dev/MSFD-QA-GFSM-A-DP. It is intended that this repository should be a working location for future development of this product by ICES Working Groups and collaborators going forward.

[5] https://community.ices.dk/ExpertGroups/WKSAE_DATRAS/_layouts/15/start.aspx#/SitePages/WKSAE-DATRAS.aspx?RootFolder=%2FExpertGroups%2FWKSAE%5FDATRAS%2F2021%5FMeeting%5FDocuments%2F06%2E%20Data&FolderCTID=0x012000395388E6B91BD24D9A8C9C4D87BF4C1B&View=%7B2128A940%2DFE68%2D485F%2DBBFE%2D03FECE8127D7%7D

Table 6.1. DATRAS datasets included in the rerun product. Values show the number of hauls per year included in the final Sampling Information product. DATRAS survey names refer to the current dataset names as downloaded from the DATRAS data centre, GFSM survey names are given below for continuity with previous data product versions. * The Belgian BTS data were excluded from previous data product versions due to insufficient data, but are included in this version. The survey acronym "BEBTS_not_in_previous" in given to these in the data product.

DATRAS name	BTS	EVHOE	FR-CGFS	IE-IGFS	NIGFS	NS-IBTS	PT-IBTS	ROCKALL	SCOROC	SCOWCGFS	SWC-IBTS
GFSM name	GNSNetBT3/ GNSGerBT3/ GNSEngBT3*	CSBBFraOT4	GNSFraOT4	CSIreOT4	CSNIrOT1/ CSNIrOT4	GNSIntOT1/ GNSIntOT3	BBICPorOT4	WAScoOT3	WAScoOT3	CSScoOT1/ CSScoOT4	CSScoOT1/ CSScoOT4
1983	-	-	-	-	-	380	-	-	-	-	-
1984	-	-	-	-	-	459	-	-	-	-	-
1985	-	-	-	-	-	515	-	-	-	-	59
1986	-	-	-	-	-	526	-	-	-	-	38
1987	64	-	-	-	-	540	-	-	-	-	50
1988	70	-	66	-	-	404	-	-	-	-	50
1989	82	-	61	-	-	426	-	-	-	-	46
1990	181	-	75	-	-	379	-	-	-	-	90
1991	198	-	81	-	-	424	-	-	-	-	105
1992	179	-	60	-	79	344	-	-	-	-	75
1993	178	-	65	-	89	374	-	-	-	-	85
1994	173	-	88	-	82	363	-	-	-	-	77
1995	194	-	89	-	77	338	-	-	-	-	101
1996	234	-	61	-	82	329	-	-	-	-	106
1997	218	129	90	-	85	363	-	-	-	-	122
1998	218	125	83	-	89	677	-	-	-	-	111
1999	245	119	102	-	87	721	-	41	-	-	121
2000	250	120	101	-	85	701	-	-	-	-	137
2001	252	151	108	-	100	771	-	44	-	-	134
2002	311	152	98	-	113	759	66	29	-	-	141
2003	305	148	96	150	119	746	-	60	-	-	153

DATRAS name	BTS	EVHOE	FR-CGFS	IE-IGFS	NIGFS	NS-IBTS	PT-IBTS	ROCKALL	SCOROC	SCOWCGFS	SWC-IBTS
GFSM name	**GNSNetBT3/ GNSGerBT3/ GNSEngBT3***	**CSBBFraOT4**	**GNSFraOT4**	**CSIreOT4**	**CSNIrOT1/ CSNIrOT4**	**GNSIntOT1/ GNSIntOT3**	**BBICPorOT4**	**WAScoOT3**	**WAScoOT3**	**CSScoOT1/ CSScoOT4**	**CSScoOT1/ CSScoOT4**
2004	389	138	92	159	102	721	-	-	-	-	145
2005	336	143	108	140	94	726	89	38	-	-	149
2006	299	127	106	168	90	709	88	32	-	-	132
2007	360	145	96	171	102	697	96	42	-	-	147
2008	333	147	101	166	53	719	87	37	-	-	123
2009	361	135	100	164	120	682	93	41	-	-	127
2010	320	139	87	176	120	694	87	-	-	-	58
2011	323	151	103	159	117	723	86	-	45	112	-
2012	346	129	94	172	118	708	-	-	36	130	-
2013	351	140	93	176	112	711	93	-	31	92	-
2014	289	155	94	170	112	665	81	-	47	121	-
2015	297	148	73	147	120	726	90	-	42	120	-
2016	300	157	73	172	124	741	85	-	48	123	-
2017	330	25	66	149	118	714	88	-	41	117	-
2018	354	155	73	153	120	713	49	-	41	116	-
2019	341	148	65	161	121	696	-	-	44	124	-
2020	342	155	59	127	115	695	-	-	40	113	-
2021	-	-	-	-	-	374	-	-	-	63	-

Data product data field description

Product data fields are as per the previous product, see section 2 Table 2.1 above (from Moriarty *et al.*, 2017).

There are two exceptions to this:

1. the HaulID is now given as (SurveyAcronym/Ship/Year/HaulNo/CountryCode/StationNo) as the previous HaulID format was not unique in the NS-IBTS surveys or NIGFS in the new product, and
2. a column 'SurveyDATRAS' containing the DATRAS survey names has been added as the final column in the data product.

See Annex 6 for details of re-estimated equations used for missing values in groundspeed, door spread, wing spread and net opening.

6.1 References

Moriarty, M., Greenstreet, S.P.R and Rasmussen (2017) Derivation of Groundfish Survey Monitoring and Assessment Data Products for the Northeast Atlantic Area. Scottish Marine and Freshwater Science Report Vol 8 No 16

7 Swept-area based on door spread or wing spread. What fits best for which purpose?

Calculations of swept-area may rely on either door spread or wing spread for the dimension of the width of the trawl path. The argument for choosing wing spread would be that this reflects the area covered during the trawl, whereas door spread is advocated when a herding effect through the sweeps is expected, which would funnel the fish from the entire area between the doors into the net opening.

Reasons for choosing one of the options may either be technical or based on the biology of the species targeted with the respective survey. The differences between the dimensions may be considerable. The ratio between both area calculations often lies beyond a factor of 3, equivalent to a threefold (or 1/3, respectively) abundance estimate for the respective species.

While it appears unlikely that 100% of the individuals of a herding species encountered between wings and doors will be captured, and zero individuals of a non-herding species encountered in this region will end up in the net, one of the two dimensions will still be more relevant in each case. For example, Reid *et al.* (2007) observed for anglerfish (*Lophius spp*.) that all fish in the path of the net (i.e. between the wings) were captured, whereas more than half of the fish between the wings and the doors were not. They, therefore, concluded that this genus did not appear to herd and many of the encounters with the wires were passive. Hence, in such a case, wing-based calculations of swept-area would be appropriate.

Existing assessment procedures relying on survey data for deriving abundance indices or estimates for commercial species assessments differ in the dimension chosen. In this, the Celtic Sea is an ecoregion where both types are in use: for example, a wing spread-based survey index is used in the stock assessment for plaice (in stratum 27.7h–k) while door spread-based abundance indices are being used in the WGCSE stock assessments for cod (in 27.7e–k), haddock (27.7b–k) and whiting (27.7b–ce–k). In all of these cases, the decision between the two parameters has been taken during the benchmark for the respective stock assessment.

Several aspects may influence the decision to apply wing spread or door spread-based calculations of swept-area, respectively:

7.1 Technical reasons

A technical reason to decide for one or the other measurement of the width of the trawl path may be the pragmatic selection of the one, for which more complete or reliable data exist in the relevant survey time-series. For example. The North Sea IBTS has a much more complete series of door spread values, compared to wing spread measurements (see section 4.3).

Another technical reason to select one of the dimensions would be to keep the time-series consistent within stock assessments, in order to be able to focus on temporal changes in the abundance indices of the assessed species.

Independent of, and potentially contrary to these technical considerations, analyses of abundance estimates for biodiversity assessments should focus on the dimension that best reflects the biology of the species in the assemblage considered. The same argument goes for data entering ecosystem models involving various species or groups of fish. This could mean that for individual species within the same analyses, the choice for either wing spread or door spread-based calculations of swept-area should be based on being biologically meaningful, which may result in a mix of both parameters.

7.2 Biological reasons

7.2.1 Herding effect

Whether or not a species' abundance would best be estimated by door spread- or wing spread-based calculations of the swept-area depends on the species-specific behavioural responses to the approaching trawl. The anterior spreading components of the trawl (i.e. doors and sweeps) trig-ger avoidance behaviour in fish species that can lead them into the trawl path (i.e. wing spread area), effectively increasing the catching efficiency of the trawl ("herding effect"; re-viewed in Winger *et al.,* 2010). In general, herding is more likely to occur for species that actively swim away from a threat (e.g. an approaching trawl) rather than using camouflage as an anti-predator strategy. Observations in the field showed that demersal and pelagic fish tend to react at great distances from the approaching trawl components, while benthic fish such as flatfish or skates tend to keep their position until direct or near physical contact with the gear components (Main and Sangster, 1981; Ryer *et al.,* 2010). As a result of the difference in timing of the response, demersal and pelagic species can be herded with relatively short sweeps (20 m), whereas benthic species requires longer sweeps (100–400 m) to be herded efficiently (Ryer *et al.,* 2010; Winger *et al.,* 2010). Therefore, the extent to which herding occurs depends on the design of the survey trawl, specifically through the length of the sweeps (Winger *et al.*, 2010).

Herding can also vary within species according to the size of the individuals and seasonally ac-cording to the water temperature. This occurs because for a fish to be herded by the sweeps, this has to have sufficient time and swimming capacity to move into the trawl path. Therefore, all parameters that affect swimming capacity (e.g. body length, temperature) and herding time (e.g. towing speed, sweeps length) will affect the catch efficiency of the trawl (Beamish, 1966; Winger *et al.*, 1999; Winger *et al.*, 2010).

In several surveys, sweeps of different lengths were used (and partly are still in use), depending on depth or season (compare Tables 4.1.3.2 and 4.1.3.3 in Moriarty *et al.,* 2017). This does not only affect the net geometry (see section 4.3) but also the catchability of species through alteration of the herding effect. A bias of length distributions and species composition can therefore only be avoided using a trawl that has no or short (< 20 m) sweeps. WKSAE-DATRAS, therefore, en-courages survey groups to consider this aspect when deciding on the replacement of the GOV with a new bottom trawl for the IBTS.

Additionally, the herding process has been shown to be mainly vision-dependent and to cease at low light levels (Wardle, 1993; Kim and Wardle, 1998). Therefore, fishing conducted by night or at depths below the level of light penetration (approximately 200 m) is unlikely to effectively herd fish into the trawl path. However, this aspect may be less important for those surveys under consideration, which are conducted during daytime only, or where the survey area covers habi-tats of less than 200 m depth.

Yet, even for the IBTS early on in the time-series, some countries did initially conduct night-time tows. These tows are classified as valid tows and can be identified by the day/night flag in DATRAS. For the NS-IBTS, a depth limit of 200 m was applied to the majority of the survey area (except Skagerrak, Division 3.a). This has been extended to 250 m with the latest update of the survey manual (ICES, 2020) for the unification of the subareas and to qualify a few tows from the northeast-ern part of the North Sea. However, the majority of stations are shallower than 200 m deep (see section 4 Figure 4.12).

Stations for the surveys in the eastern North Atlantic partly target deep habitats, e.g. on Porcupine Bank or in the Gulf of Cádiz, but in most regions, stations with < 200 m depth dominate (for details see the respective survey manuals in ICES, 2017).

Moreover, a disruption of the herding effect has also been observed when trawling was conducted in shallow waters after an intense storm event (Wieland *et al.*, 2011), such temporary conditions are less suitable to be considered in the selection between wind spread and door spread.

The current survey design related implications of herding for the use of door or wing spread based swept-area abundance indices are summarized in Table 7.1

Table 7.1. Consequences on density estimates at length when adopting the wing spread or door-spread to calculate swept-area.

Measure used to estimate Swept-area	Herding	No Herding
Wing spread	overestimation of density; length-based bias(large individuals)	correct
Door spread	underestimation of density;length-based bias(small individuals)	underestimation of density

Unfortunately, at present, the response of individual species to the doors and sweeps, hence the herding effect, has been investigated mostly for the main commercial species, whereas much less information is available for low value and non-commercial fish species.

Therefore, for the choice between door spread and wing spread for a multispecies dataset, e.g. during analyses of biodiversity, a concrete species-specific recommendation would be needed. Since we are not aware of such a list, we propose here the first approximation to solving this problem (Table 7.2). This allocation may serve—and should be viewed—as a starting point for future refinements as additional information on the herding behaviour becomes available through either literature research or new catchability trials at sea.

While for particular investigations, researchers may want to conduct analyses with not only species-specific but even haul-based decisions on whether to use wing spread- or door spread-based swept-area calculations, this would be beyond the scope of WKSAE-DATRAS. At the same time, we see the need for further studies investigating the effect of differential catchability of species for analyses based upon swept-area metrics.

Table 7.2. Proposed allocation of species/groups for the calculation of door spread- vs. wind spread-based swept-area calculations. Upper section: Species for which concrete recommendations exist; lower section: considerations for species without specific recommendations (The proposed reference measures for swept-area estimates refer to sweep length of maximum 100 m, 4 kn towing speed and daytime hauls at depths shallower than 200 m).

Species / group	Herding Use door spread	Non-herding Use wing spread	Reference / Explanation
Species-specific decisions			
Haddock (*Melanogrammusaeglefinus*)	X		WGCSE; ref's in Fraser *et al.* (2007)
Whiting (*Merlangiusmerlangus*)	X		WGCSE; ref's in Fraser *et al.* (2007)
Cod (*Gadusmorhua*)	X		WGCSE

Species / group	Herding Use door spread	Non-herding Use wing spread	Reference / Explanation
Saithe (*Pollachiusvirens*)	X		Main and Sangster, 1981
Plaice (*Pleuronectesplatessa*)		X	WGCSE; ref's in Fraser *et al.* (2007)
Angler - *Lophius* spp.		X	Reid *et al.* (2007)
Dab *(Limandalimanda)*		X	Main and Sangster, 1981
Lemon sole (*Microstomuskitt*)		X	Main and Sangster, 1981
Northern rock sole (*Lepidopsettapolyxystra*)		X	Rose, 1996
Halibut – *Hippoglossus* spp.	X		Rose, 1996
Group-wise decisions, where species-specific information is lacking			
Slow-swimming species (relative to trawl speed)		X	
Camouflages species, which tend to hide, rather than escape		X	
Pelagic species, tending to escape upwards		X	
Benthic species, tending to burrow in sediment		X	
Schooling species	X		
Commercially targeted species where fleet uses OTB type gear to achieve herding	X		Proxy, if no experimental data are available (note if length of sweeps is comparable to survey)

7.3 References

Beamish, F.W.H., 1966. Swimming endurance of some Northwest Atlantic fishes. Journal of Fisheries Research Board of Canada, 23: 341–347.

Fraser, H. M., Greenstreet, S. P., and Piet, G. J. 2007. Taking account of catchability in groundfish survey trawls: implications for estimating demersal fish biomass. ICES Journal of Marine Science, 64(9), 1800–1819.

ICES.2017. Manual of the IBTS North Eastern Atlantic Surveys. Series of ICES Survey Protocols SISP 15. 92 pp. http://doi.org/10.17895/ices.pub.3519.

ICES. 2020. Manual for the North Sea International Bottom Trawl Surveys. Series of ICES Survey Protocols SISP 10-IBTS 10, Revision 11. 102 pp. http://doi.org/10.17895/ices.pub.7562.

Kim, Y.H., and Wardle, C.S., 1998a. Modelling the visual stimulus of towed fishing gear. Fisheries research, 34: 165–177.

Main, J., and Sangster, G., 1981. A study of the fish capture process in a bottom trawl by direct observation from a towed underwater vehicle. Scottish fisheries research report No. 23, 24 pp.

Moriarty, M., Greenstreet, S. P. R., and Rasmussen, J. 2017. Derivation of groundfish survey moni-toring and assessment data products for the Northeast Atlantic area. Scottish Marine and Freshwater Science, 8(16), 240.

Reid, D. G., Allen, V. J., Bova, D. J., Jones, E. G., Kynoch, R. J., Peach, K. J., Fernandes, P. G., and Turrell, W. R. 2007. Anglerfish catchability for swept-area abundance estimates in a new survey trawl. – ICES Journal of Marine Science, 64: 1503–1511.

Rose, C.S., 1996. Behavior of North Pacific groundfish encountering trawls: applications to re-duce bycatch. In: Alaska Sea Grant Program (Ed.), Solving Bycatch: Considerations for Today and Tomorrow. Alaska Sea Grant Report AK-56-96-03, pp. 235–241.

Ryer, C.H., Rose, C.S., and Iseri, P.J., 2010. Flatfish herding behavior in response to trawl sweeps: a comparison of diel responses to conventional sweeps and elevated sweeps. Fishery Bulle-tin, 108: 145–154.

Wardle, C.S., 1993. Fish behaviour and fishing gear. In: Pitcher, T.J. (Ed.), Behavior of Teleost Fishes, pp. 609–643. Chapman & Hall, London.

Wieland, K., H.J. Olesen, E.M. Fenger Pedersen and J.E. Beyer (2011): Potential bias in estimates of abundance and distribution of North Sea cod (Gadusmorhua) due to strong winds prevail-ing prior or during a survey. Fish. Res. 110: 325–330.

Winger, P. D., He, P., and Walsh, S.J., 1999. Swimming endurance of American plaice (Hippoglossoidesplatessoides) and its role in fish capture. ICES Journal of Marine Science, 56: 252–265.

Winger, P. D., Eayrs, S., and Glass, C. W. 2010. Fish behaviour near bottom trawls. In Behavior of Marine Fishes: Capture Processes and Conservation Challenges, pp. 67–102. Ed. by P. He. Wiley-Blackwell, Ames, IA. 392 p.

Annex 1: List of participants

Name	Institute	Country
Ailbhe Kavanagh	Marine Institute	Ireland
Anna Rindorf	Technical University of Denmark, National Institute of Aquatic Resources (DTU Aqua), Lyngby	Denmark
Anne Sell	Thünen Institute of Sea Fisheries	Germany
Corina Chaves	Portuguese Institute for Sea and Atmosphere (IPMA)	Portugal
Finlay Burns	Marine Scotland Science (MSS)	United Kingdom
Francisco Baldó	Spanish Institute of Oceanography (IEO), Cádiz	Spain
Francisco Velasco Guevara	Spanish Institute of Oceanography (IEO), Santander	Spain
Hongru Zhai	Swedish University of Agricultural Sciences (SLU)	Sweden
Jed Kempf	Marine Institute	Ireland
Juan Carlos Arronte	Spanish Institute of Oceanography (IEO), Santander	Spain
Kai Ulrich Wieland (chair)	Technical University of Denmark, National Institute of Aquatic Resources (DTU Aqua), Hirtshals	Denmark
Kim Ludwig	Thünen Institute of Sea Fisheries	Germany
Meadhbh Moriarty	Marine Scotland Science (MSS)	United Kingdom
Morgane Travers-Trolet	French Research Institute for Exploitation of the Sea (Ifremer)	France
Patrik Börjesson	Swedish University of Agricultural Sciences (SLU)	Sweden
Ruadhán Gillespie-Mules	Marine Scotland Science (MSS)	United Kingdom
Ruth Kelly	School of Natural Sciences, Trinity College	United Kingdom
Sonia Seixas	Science and Technology Department, Aberta University	Portugal
Valentina Melli	Technical University of Denmark, National Institute of Aquatic Resources (DTU Aqua), Hirtshals	Denmark

Annex 2: Resolutions

2020/WK/EOSG05 The **Workshop on the production of swept area estimates for all hauls in DATRAS for biodiversity assessments** (WKSAE_DATRAS), chaired by Kai Wieland*, Denmark, will be established and meet online 31 May–4 June 2021 to:

a) Harmonize the selection on surveys and time series available in DATRAS for biodiversity assessments by:
   i. Checking and validating the calculations of missing data of the variables related to the swept-area effort estimates by some countries/countries in need;
   ii. Proposing common strategies to reduce missing data in the crucial variables for biodiversity assessments;
   iii. Defining common calculations, when possible, across surveys and countries, and perform a quality check against the observations from the most recent year(s);
   iv. Building on previous work (WKSABI) to define species groups for which the swept-area estimates should be based on door spread or wing spread and those for which swept-area may not be used;

b) Develop a script to calculate swept-area indices for biodiversity assessments;

c) Calculate swept-area indices and create a data product as input to OSPAR common indicators for fish and food webs;

d) Update the DATRAS calculation document to include reference to the new data product and fields used for biodiversity assessments.

In the first part of the workshop, national experts and members of survey groups together with the ICES Data Centre will work to clean up the data (obtain and validate missing algorithms) for the agreed NE Atlantic bottom and beam trawl surveys and time-series (see Table 1 in supporting information).

WKSAE-DATRAS will report 30th July 2021 for the attention of the Advisory Committee and the Ecosys-tem Observation Steering Group.

Annex 3: Agenda

31 May13:00-16:00 **Presentation of participants**

Introduction, ToR's, Code of conduct and conflict of interest,

- Lara

Agenda

- Kai

MSS/OSPRA approach for calculating swept area

(Background Docs on SharePoint: SMFS_0818_0.pdf, SMFS_0816.pdf)

- Meadhbh

IBTS approach for estimating missing input data for the calculation of swept area

(Background docs on SharePoint: WKSABI report, IBTSWG2020 Annex 7 NS.pdf, IBTSWG2020 Annex 8 NeAtl.pdf)

- Vaishav

1 June 10:00-12:00 **Available data in DATRAS – Beam Trawl surveys**

- Vaishav

DATRAS flexfile for NeAtl IBTS

- Vaishav

13:00-15:00 **DATRAS flexfile for NS-IBTS**

Format, SQL and R-codes

- Vaishav, Cecilia

Selected examples for estimating missing values comparing MSS and IBTS approach

- Kai

How to handle remaining missing and erroneous values?

2 June 10:00-12:00	**Data availability, defining time series and output data formats**
	Distribution of writing tasks for report
	Swept area based on door spread or wing spread: What fits best for which purpose? a) for Stock assessments - Kai
Afternoon	**Update documentation for DATRAS swept area calculation product,**
(no plenary)	**Data analysis and report writing**
3 June 10:00-12:00	**Data check on Beam Trawl distance and Swept Area** - Patrik
	Swept area based on door spread or wing spread: What fits best for which purpose? b) for other purposes, e.g. OSPAR indicators - Valentina, Anne
	Recommendations
	Distribution of remaining writing tasks for report
Afternoon	**Update algorithms and documentation for swept area calculations,**
(no plenary)	**Report writing**
4 June 10:00-12:00	**Presenting parts for report**
	Adjourn

Times are in CEST (Central European Summer Time)

Annex 4: List of presentations

Eugene Nixon and Lara Salvany: ICES workshop on the production of swept-area estimates for all hauls in DATRAS for biodiversity assessments (WKSAE-DATRAS). 10 pp.

Meadhbh Moriarty: MSS/OSAR Approach for calculating swept-area. 18 pp.

Vaishav Sony: DATRAS Swept-area product calculation. 12 pp.

Kai Wieland: Estimating missing values for door and wing spread. Comparing algorithms—a few examples from the North Sea IBTS in recent years. 8 pp.

Kai Wieland: Swept-area based on door or wing spread. What is most appropriate and needed? 5 pp.

Patrik Börjesson: Data check on Beam Trawl distance and Swept-area comparing OSPAR and DATRAS datasets. 5 pp.

Annex 5: Documentation of DATRAS swept-area calculations

Beam Trawl Surveys

Beam Trawl Survey includes 4 surveys (inshore and offshore): BTS, SNS, DYFS, BTS-VII. They cover mostly Q3 and Q4.

Chapter 4 of the 2021 WGBEAM document (https://doi.org/10.17895/ices.pub.8114) is used to calculate swept-area-based product in DATRAS.

2 different formula to calculate swept-area:

1. Gear expression indicating width of beam. Uses latitude and longitude
2. Based on fishing speed and haul duration

Steps to calculate swept-area-based product:

- It follows the same logic. The first step is a standardization from subsample to catch per unit per effort.
- Beam width is defined by the gear code.
- For the distance a standard formula is used with haul duration. In absence of haul duration data, it can still be calculated using latitude and longitude
- Where there's no catch hauls, a cpue values of 0 is assigned.

In summary,

- Swept-area in km2= 2*beam width*distance/106
- Swept-area in km2= 2*beam width*(haul duration/60)*fishing speed*1852/106

WKSAE-DATRAS discussed the following regarding BEAM surveys:

- Whether fishing speed (ground speed) was always present or not. Speed is the only variable to the otherwise constant parameters. Concerns arise for cases where value is not 0 but it is the same value all through the data and still is not informative.
- It requires that Beam calculation follows the same way as NS-IBTS, which based on HH data and HL data not need to consider in the calculation it means calculation is not a species-specific transformation for beam trawl. It would be easier to have the HH file finished and then 2 options in the HH for example swept-area wing and swept-area-door.
- WGBEAM requested 2 extra columns based on HL data with catch-per unit-per effort (cpue) per swept-area in the same file. WKSAE-DATRAS do not think this is useful as a standard HH file can be used by other groups so HL based data not necessary.
- Only 2 decimals are provided in cpue_number as compared with 4 decimals on other gears, so it need to implement.

SQL steps to calculate swept-area calculation

Fetch data from DATRAS. First HH records with list of all ships. Select survey, quarter, year and ship

1. Length class normalization
2. Combine HH and HL data

Results are multiplied to calculate distance if not present.

3. If subsample, it is converted to cpue (standard procedure that is not relevant to WKSAE-DATRAS and WKABSENS. Can be removed for OSPAR data products)
4. Apply Swept-area Km2 formula
5. Fill null holes
6. Result stores in single file

WGBEAM report 2021 is published and information on calculation procedure is available. It is done for every species at the moment so it is a very large file. For BEAM surveys, only distance matters so need to make sure it is correctly calculated.

WKSAE discussed process and needs for OSPAR data product. Agreement was to create an additional product without HL combined, with 4 decimals and compare it to the OSPAR product as a length-based estimate.

Northeast Atlantic and North Sea IBTS

NS-IBTS is based on IBTSWG 2015 report[6], and NEA-IBTS is based on NEA manual 2017 (SISP)[7].

A new document can be found on the DATRAS webpage under the document page[8] and called "Swept-area (km2) algorithms". This new document will replace the old document called "NS-IBTSsweptareakm2algorithms" that contained all the formulas used to calculate swept-area for each country only from the NS-IBTS. Since every country have different conditions for year ranges, the documentation is not clear for users at the moment, especially because of some of the information are in duplicates, therefore on the recommendation of WKSAE-DATRAS , there document has been changed to make it more user-friendly and up-to-date, including both NS- and NEA-IBTS.

The same parameters and logic as are used for both NS - and NEA-IBTS, but they are only haul based (HH) with no length information. The algorithms have different values applied to each survey, using the same model. The NEA-IBTS manual contains the master document with the specific conditions for each survey (see link further up). A SQL script is run for each individual survey and uses only valid hauls, also flagging calculated/observed values observed value present for wing and door spread. For NS-IBTS there is a separate routine for every country in the SQL script, but it follows the same logic as NEA-IBTS. For distance, there are 2 formulas: straight distance lines or based on ground speed.

WKSAE-DATRAS noted that the main difference with the MSS approach, is that MSS applies a data cleaning to remove obvious errors prior to estimating the parameters. Both approaches fill missing data, but only MSS detects and eliminates outliers. ICES Data Centre communicates the issues to the countries but there is a time-lag before that will be corrected.

One possibility would be to include the columns of estimated values in the FlexFile file, this equation of entirely estimated values for all hauls could then be used to see differences between expected and calculated values. What the WKSAE-DATRAS group's proposal is, if data back to 1983 is requested, is still not finalized. The reliability will be poorer, but maybe it could be an approximation rather than assuming an average swept-area for all hauls.

[6] https://www.ices.dk/sites/pub/Publication Reports/Expert Group Report/SSGIEOM/2015/2015 IBTSWG Annex 7.pdf

[7] https://www.ices.dk/sites/pub/Publication%20Reports/ICES%20Survey%20Protocols%20(SISP)/SISP%2015%20NeAtl%20IBTS%20Survey.pdf

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

However, changes in gear and/or methodology on specific surveys may hinder the assumption that the relationship holds and feedback with data submitters may be required. Some survey documentation changes in the manual.

R code for NS-IBTS and NEA-IBTS swept-area calculation

Before and during the WKSAE-DATRAS meeting two R scripts were produced by the ICES Data Centre together with WKSAE-DATRAS participants, to mimic the SQL algorithms for the swept-area data products for both NS - and NEA-IBTS.

So far, the R codes have been produced with the algorithms, values and underlying conditions from each survey and/or country, using the information provided by the national experts.

WKSAE-DATRAS recommends that these codes need to be tested further and modified, both by the national experts themselves and the ICES Data Centre so that they are fully compatible with the swept-area-based products from the DATRAS SQL procedure. For now, these scripts are readily available at the WKSAE-DATRAS GitHub page[9]. When testing has been completed, the R script will be readily available on DATRAS own GitHub page for all to use[10].

Swept-area algorithms and values

The tables below list the values used for the algorithms used for each country/survey in NS-IBTS (Table 1) and NEA-IBTS (Table 2). These are the same values, algorithms and conditions that are used in SQL and the R code to calculate swept-area products for NS and NEA IBTS.

[9] https://github.com/ices-eg/WKSAE-DATRAS

[10] https://github.com/ices-tools-prod/DATRAS

Table 1. Algorithms for NS-IBTS.

Country	Ship	Quarter	Initial Year	Final Year	Dependent Variable(x)	Priority	Independent Var 1 (y)	Independent Var 2 (z)	Independent Var 3 (q)	Sweep Length Min	Sweep Length Max	Formula	a	b	c	d
DK			2004	current year	DoorSpread	1	Depth			0	60	x1= a+b*EXP(c*y)	79.386	-33.695	-0.028	
DK			2004	current year	DoorSpread	1	Depth			else	else	x2= a+b*EXP(c*y)	104.5	-316.682	-0.043	
DK			2004	current year	WingSpread	1	DoorSpread			0	60	x= a+b*y	5.867	0.206		
DK			2004	current year	WingSpread	1	DoorSpread			else	else	x= a+b*y	4.9	0.166		
NL			2003	2004	DoorSpread	1	Depth	Warplngt				x=a*log10(y)+b*log10(z)+c	29.544	14.116	-3.456	
NL			2005	2014	DoorSpread	1	Depth	Warplngt				x=a*log10(y)+b*log10(z)+c	31.165	0.2974	29.321	
NL			2015	2016	DoorSpread	1	Depth	Warplngt				x=a*log10(y)+b*log10(z)+c	28.947	23.372	-32.476	
NL			2017	current year	DoorSpread	1	Depth	Warplngt				x=a*log10(y)+b*log10(z)+c	15.842	30.868	-24.793	
GB-SCT			2005	current year	DoorSpread	1	Warplngt					x=a*log(y)+b	24.481	-60.895		
GB-SCT			2005	current year	DoorSpread	2	WingSpread					x=(a*y)+b	4.3277	-3.784		
GB-SCT			2005	current year	WingSpread	1	DoorSpread	Warplngt				x=a*log(z)+b	4.6235	-7.3296		
GB-SCT			2005	current year	WingSpread	2	DoorSpread					x=(a*y)+b	0.1909	4.011		
FR			2004	current year	DoorSpread	1	Depth					x=a+b*y	47.548	0.296		
FR			2004	current year	WingSpread	1	DoorSpread					x=a+b*y	9.4306	0.131		
FR			2005	current year	WingSpread	2	Depth					x=a+b*y	15.72	0.038		
SE			2004	current year	DoorSpread	1	Depth			0	60	x1=a*log(y)+b	13.706	26.853		
SE			2004	current year	DoorSpread	1	Warplngt			else	else	x2=a*log(y)+b	29.489	-67.157		
SE			2004	current year	WingSpread	2	Depth			0	60	x=a*log(x1)+b	15.78	-48.248		
SE			2004	current year	WingSpread	2	Depth			else	else	x=a*log(x1)+b	21.231	-77.605		
SE			2004	current year	WingSpread	2	Warplngt			0	60	x=a*log(x2)+b	15.78	-48.284		
SE			2004	current year	WingSpread	2	Warplngt			else	else	x=a*log(x2)+b	21.231	-77.605		
GB			2004	current year (- 2006)	DoorSpread	3	Average Depth					x= a*log(y)+b	15.031	12.6399		
GB			2004	current year (- 2006)	DoorSpread	2	Warplngt					x= a*log(y)+b	21.78	-47.2		
GB			2004	current year (- 2006)	DoorSpread	1	WingSpread					x=a*y+b	4.616	-15.966		
GB			2004	current year (- 2006)	WingSpread	3	Average Depth					x= a*log(y)+b	2.9249	7.43486		
GB			2004	current year (- 2006)	WingSpread	2	Warplngt					x= a*log(y)+b	4.074	-3.137		
GB			2004	current year (- 2006)	WingSpread	1	DoorSpread					x=a*y+b	0.1869	5.7416		
GB			2006	2006	DoorSpread	3	Average Depth					x= a*log(y)+b	12.468	17.5865		
GB			2006	2006	DoorSpread	2	Warplngt					x= a*log(y)+b	16.442	-25.4727		
GB			2006	2006	DoorSpread	1	WingSpread					x=a*y+b	3.8182	-11.9066		
GB			2006	2006	WingSpread	3	Average Depth					x= a*log(y)+b	3.1495	8.2192		
GB			2006	2006	WingSpread	2	Warplngt					x= a*log(y)+b	4.1885	-2.8637		
GB			2006	2006	WingSpread	1	DoorSpread					x=a*y+b	0.2242	5.7889		
NO	58G2	1	2009	2015	DoorSpread	1	Depth			0	60	x=a+b*y+(c*y^2)	54.84	0.41	-0.001	
NO	58G2	1	2009	2015	DoorSpread	1	Depth			else	else	x=a+b*y+(c*y^2)	55.7	0.56	-0.001	
NO	58G2	1	2015	2015	WingSpread	1	DoorSpread			0	60	x=a+b*y	40.074	1.9259		
NO	58G2	1	2015	2015	WingSpread	1	DoorSpread			else	else	x=a+b*y	-23.41	6.931		
NO	JHJ	3	2006	2013 (minus 2009, 2012)	DoorSpread	1	Depth			0	60	x=a+b*y+(c*y^2)	64.94	0.152	-1.99	
NO	JHJ; 58UO; HAV	3	2010	2013 (minus 2012)	DoorSpread	1	Depth			0	60	x=a+b*y+(c*y^2)	64.94	0.152	-1.99	
DE			2004	current year	DoorSpread	1	WingSpread	Depth		0	50	x=a+b*y+c*log(z)	-7.456	3.616	3.124	
DE			2004	current year	DoorSpread	1	WingSpread	Depth		else	else	x=a+b*y+c*log(z)	-7.935	5.123	2.366	
DE			2004	current year	DoorSpread	2	Warplngt	Depth		0	50	x=a+b*log(y)+c*log(z)	-0.441	10.009	4.768	
DE			2004	current year	WingSpread	1	DoorSpread	Warplngt	Depth	0	50	x=a+b*y+c*log(z)+d*log(q)	3.359	0.095	1.391	0.261
DE			2004	current year	WingSpread	1	DoorSpread	Warplngt	Depth	esle	else	x=a+b*y+c*log(z)+d*log(q)	3.087	0.118	0.445	0.368
DE			2004	current year	WingSpread	2	Warplngt	Depth		0	50	x=a+b*log(y)+c*log(z)	3.317	2.341	0.713	

Table 2. Algorithms for NEA-IBTS.

Country	Survey	Ship	Initial Year	Final Year	Dependent Variable(x)	Priority	Independent Variable1(y)	Sweep Length Min	Sweep Length Max	Formula	a	b
GB-SCT	ROCKALL		2005	2009	Doorspread		Depth			x=a+b*log(y)	-23.35	21.27
GB-SCT	ROCKALL		2005	2009	Wingspread		Depth			x=a+b*log(y)	10.16	2.01
GB-SCT	SCOROC		2016	current	Doorspread		Depth			x=a+b*log(y)	28.57	14.07
GB-SCT	SCOROC		2016	current	Wingspread		Depth			x=a+b*log(y)	12.43	1.48
GB-SCT	SWC-IBTS		2004	2010	Doorspread		Depth			x=a+b*log(y)	-23.35	21.27
GB-SCT	SWC-IBTS		2004	2010	Wingspread		Depth			x=a+b*log(y)	10.16	2.01
GB-SCT	SCOWCGFS		2016	current	Doorspread		Depth	0	60	x=a+b*log(y)	-14.74	19.28
GB-SCT	SCOWCGFS		2016	current	Doorspread		Depth	else	else	x=a+b*log(y)	-13.52	22.46
GB-SCT	SCOWCGFS		2016	current	Wingspread		Depth	0	60	x=a+b*log(y)	1.47	4.03
GB-SCT	SCOWCGFS		2016	current	Wingspread		Depth	else	else	x=a+b*log(y)	10.28	2.01
GB-NIR	NIGFS		2005	current	Doorspread	1	Depth			x=a+b*log(y)	7.49	7.70
GB-NIR	NIGFS		2005	current	Wingspread	1	Doorspread			x=a+b*y	5.28	0.27
GB-NIR	NIGFS		2005	current	Wingspread	2	Depth			x=a+b*log(y)	8.61	1.79
FR	FR-CGFS	35TH	2015	current	Doorspread	1	Depth			x=a+b*log(y)	-0.96	15.38
FR	FR-CGFS	35TH	2015	current	Wingspread	1	Doorspread			x=a+b*log(y)	6.62	0.17
FR	FR-CGFS	35TH	2015	current	Wingspread	2	Depth			x=a+b*log(y)	7.13	2.51
FR	EVHOE		2016	current	Doorspread	1	Depth	0	60	x=a+b*log(y)	6.29	14.46
FR	EVHOE		2016	current	Doorspread	1	Depth	else	else	x=a+b*log(y)	-15.08	21.78
FR	EVHOE		2016	current	Wingspread	1	Depth	0	60	x=a+b*log(y)	14.39	1.23
FR	EVHOE		2016	current	Wingspread	1	Depth	else	else	x=a+b*log(y)	11.03	1.92
SP	SP-PORC		2016	current	Doorspread	1	Depth			x=a+b*log(y)	-56.94	32.32
SP	SP-PORC		2016	current	Wingspread	1	Depth			x=a+b*log(y)	4.31	2.95
SP	SP-NORTH		2016	current	Doorspread	1	Depth			x=a+b*log(y)	13.42	16.25
SP	SP-NORTH		2016	current	Wingspread	1	Depth			x=a+b*log(y)	7.1	2.46
SP	SP-ARSA		2016	current	Doorspread		Depth	0	60	x=a+b*log(y)	-0.84	14.06
SP	SP-ARSA		2016	current	Doorspread		Depth	else	else	x=a+b*log(y)	-4.12	18.97
SP	SP-ARSA		2016	current	Wingspread		Depth	0	60	x=a+b*log(y)	-0.43	4.69
SP	SP-ARSA		2016	current	Wingspread		Depth	else	else	x=a+b*log(y)	7.22	2.51
IE	IE-IGFS		2016	current	Doorspread	1	Depth	0	60	x=a+b*log(y)	38	10.83
IE	IE-IGFS		2016	current	Doorspread	1	Depth	else	else	x=a+b*log(y)	29.76	17.8
IE	IE-IGFS		2016	current	Wingspread	1	Depth	0	60	x=a+b*log(y)	12.69	2.13
IE	IE-IGFS		2016	current	Wingspread	1	Depth	else	else	x=a+b*log(y)	14.2	1.8
IE	IE-AIMS		2015	current	Doorspread	1	Depth			x=a+b*log(y)	16.06	14.66
IE	IE-AIMS		2015	current	Wingspread	1	Depth			x=a+b*log(y)	14.49	2.75

Annex 6: Summary of refitted models in the re-estimation of Groundfish Survey Monitoring and Assessment Data Product for sampling information

This annex details the updates to the fitted equations in the Groundfish Survey Monitoring and Assessment Data Product, which resulted from the inclusion of new data from DATRAS in the reanalysis.

Details of data processing and protocols are given in Moriarty *et al.* (2017), equation numbers here refer the reader to the same equations in Moriarty *et al.* (2017)[11]. See section 6 for further details of rerun, and locations of relevant code and data archival.

Equ. 4.1.5.1. Groundspeed Model 1 – for vessels where groundspeed was recorded

```
lm(formula = GroundSpeed ~ Quarter:Ship:Gear, data = hauls)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.2596 -0.0901  0.0000  0.1164  4.3164 

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)    
(Intercept)                3.714874   0.003830 969.956  < 2e-16 ***
Quarter:Ship29CS:GearBAK  -0.212668   0.003324 -63.987  < 2e-16 ***
Quarter:Ship29MO:GearBAK  -0.202790   0.003335 -60.801  < 2e-16 ***
Quarter:Ship11BE:GearBT4A  0.114925   0.007677  14.971  < 2e-16 ***
Quarter:Ship74E9:GearBT4A  0.072679   0.003666  19.823  < 2e-16 ***
Quarter:Ship06S1:GearBT7   0.088246   0.008065  10.941  < 2e-16 ***
Quarter:Ship64SS:GearBT8   0.095042   0.002085  45.578  < 2e-16 ***
Quarter:Ship64T2:GearBT8   0.095042   0.002212  42.960  < 2e-16 ***
Quarter:Ship06NI:GearGOV   0.133075   0.003368  39.515  < 2e-16 ***
Quarter:Ship26D4:GearGOV   0.061966   0.002515  24.638  < 2e-16 ***
Quarter:Ship35HT:GearGOV   0.023599   0.001563  15.101  < 2e-16 ***
Quarter:Ship35TH:GearGOV   0.258701   0.017869  14.478  < 2e-16 ***
Quarter:Ship45CE:GearGOV   0.042184   0.001481  28.488  < 2e-16 ***
Quarter:Ship58EJ:GearGOV   0.114756   0.016019   7.164 8.02e-13 ***
Quarter:Ship58G2:GearGOV   0.214747   0.012409  17.306  < 2e-16 ***
Quarter:Ship58J3:GearGOV  -0.017546   0.004622  -3.796 0.000147 ***
Quarter:Ship58UO:GearGOV   0.091666   0.006555  13.983  < 2e-16 ***
Quarter:Ship64T2:GearGOV   0.285126   0.007794  36.584  < 2e-16 ***
Quarter:Ship748S:GearGOV  -0.006205   0.001667  -3.722 0.000198 ***
Quarter:Ship749S:GearGOV   0.132068   0.005358  24.650  < 2e-16 ***
Quarter:Ship74E9:GearGOV   0.071279   0.002594  27.482  < 2e-16 ***
Quarter:Ship77AR:GearGOV   0.010928   0.003575   3.057 0.002238 ** 
Quarter:Ship77MA:GearGOV  -0.017200   0.037174  -0.463 0.643601    
Quarter:Ship77SE:GearGOV  -0.010810   0.011026  -0.980 0.326922    
Quarter:Ship68NA:GearNCT  -0.067301   0.002009 -33.500  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2425 on 33242 degrees of freedom
  (17885 observations deleted due to missingness)
Multiple R-squared:  0.4019,	Adjusted R-squared:  0.4015 
F-statistic: 893.5 on 25 and 33242 DF,  p-value: < 2.2e-16
```

[11] Moriarty, M., Greenstreet, S.P.R. and Rasmussen, J. (2017). Derivation of Groundfish Survey Monitoring and Assessment Data Products for the Northeast Atlantic. Scottish Marine and Freshwater Science, Vol. 8 No. 16: 240 pp. https://doi.org/10.7489/1984-1.

Equ. 4.1.5.2. Groundspeed model 2

```
Call:
lm(formula = GroundSpeed ~ Quarter:Gear, data = hauls)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.1546 -0.1185  0.0000  0.1805  4.3719 

Coefficients:
                  Estimate Std. Error  t value Pr(>|t|)    
(Intercept)       3.789703   0.003450 1098.320  < 2e-16 ***
Quarter:GearBAK  -0.229707   0.002591  -88.659  < 2e-16 ***
Quarter:GearBT4A  0.054956   0.003559   15.441  < 2e-16 ***
Quarter:GearBT7  -0.023398   0.003009   -7.777 7.64e-15 ***
Quarter:GearBT8   0.070099   0.001744   40.198  < 2e-16 ***
Quarter:GearGOV   0.009605   0.001258    7.632 2.37e-14 ***
Quarter:GearNCT  -0.086008   0.002087  -41.204  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.261 on 33261 degrees of freedom
  (17885 observations deleted due to missingness)
Multiple R-squared:  0.3068,	Adjusted R-squared:  0.3067 
F-statistic:  2454 on 6 and 33261 DF,  p-value: < 2.2e-16
```

4.1.7.1.1.1. The GOV Otter Trawl – Wingspread

```
Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's me
thod [
lmerModLmerTest]
Formula: log(WingSpread) ~ LogDepthCenter:EstSweepCat + (1 | Ship:EstSweepCat
)
   Data: train

     AIC      BIC   logLik deviance df.resid 
-37816.6 -37778.1  18913.3 -37826.6    16283 

Scaled residuals: 
     Min       1Q   Median       3Q      Max 
-14.4880  -0.5825  -0.0202   0.5991   7.3347 

Random effects:
 Groups           Name        Variance Std.Dev.
Ship:EstSweepCat (Intercept) 0.004676 0.06838 
 Residual                    0.005704 0.07552 
Number of obs: 16288, groups:  Ship:EstSweepCat, 19

Fixed effects:
                                Estimate Std. Error        df t value Pr(>|t
|)    
(Intercept)                    2.988e+00  1.580e-02 1.881e+01  189.06   <2e-
16 ***
LogDepthCenter:EstSweepCatlong  8.553e-02  2.909e-03 1.628e+04   29.40   <2e-
16 ***
LogDepthCenter:EstSweepCatshort 1.401e-01  1.445e-03 1.629e+04   96.95   <2e-
16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
                    (Intr) LgDpthCntr:EstSwpCtl
LgDpthCntr:EstSwpCtl -0.030                     
LgDpthCntr:EstSwpCts  0.018 -0.001              
```

4.1.8.1.1.1. The GOV Otter Trawl – Doorspread

```
Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's me
thod [
lmerModLmerTest]
Formula: log(DoorSpread) ~ LogDepthCenter:EstSweepCat + (1 | Ship:EstSweepCat
)
   Data: train

     AIC      BIC   logLik deviance df.resid
-38132.0 -38093.5  19071.0 -38142.0    16283

Scaled residuals:
    Min      1Q  Median      3Q     Max
-9.9195 -0.5809  0.0287  0.5872 10.6707

Random effects:
 Groups           Name        Variance Std.Dev.
Ship:EstSweepCat (Intercept) 0.014782 0.12158
 Residual                    0.005587 0.07474
Number of obs: 16288, groups:  Ship:EstSweepCat, 19

Fixed effects:
                                 Estimate Std. Error        df t value Pr(>|t
|)
(Intercept)                     4.446e+00  2.796e-02 1.900e+01  159.04   <2e-
16 ***
LogDepthCenter:EstSweepCatlong  1.534e-01  2.881e-03 1.628e+04   53.25   <2e-
16 ***
LogDepthCenter:EstSweepCatshort 2.169e-01  1.431e-03 1.628e+04  151.64   <2e-
16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
                     (Intr) LgDpthCntr:EstSwpCtl
LgDpthCntr:EstSwpCtl -0.017
LgDpthCntr:EstSwpCts  0.010  0.000
```

Equ 4.1.7.1.2.1 – Doorspread of 'ROT' gear

```
Call:
lm(formula = log(DoorSpread[hauls$Gear == "ROT"]) ~ log(Depth[hauls$Ge
ar == "ROT"]), data = hauls)

Residuals:
     Min       1Q   Median       3Q      Max
-0.35722 -0.04561  0.00522  0.05423  0.26998

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                   2.743854   0.016574  165.56   <2e-16 *
**
log(Depth[hauls$Gear == "ROT"]) 0.220215   0.004148   53.09   <2e-16 *
**
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.07917 on 1988 degrees of freedom
  (955 observations deleted due to missingness)
Multiple R-squared:  0.5864,   Adjusted R-squared:  0.5862
F-statistic:  2819 on 1 and 1988 DF, p-value: < 2.2e-16
```

Equ 4.1.7.1.2.1 Wingspread of 'ROT' gear

Here the equation from the Moriarty *et al.* 2017 is fitted without change.

i.e. Wingspread = exp (0.3859096 + 0.6483112(ln *D*oorspread))

4.1.7.3.1.1. Net opening GOV trawl

```
Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's me
thod [
lmerModLmerTest]
Formula: log(Netopening) ~ LogDepthCenter:EstSweepCat + (1 | Ship:EstSweepCat
)
   Data: train

     AIC      BIC   logLik deviance df.resid
-22846.2 -22807.7  11428.1 -22856.2    16283

Scaled residuals:
    Min      1Q  Median      3Q     Max
-6.5685 -0.5570  0.0305  0.6094  5.0993

Random effects:
 Groups           Name        Variance Std.Dev.
Ship:EstSweepCat (Intercept) 0.01827  0.1352
 Residual                    0.01429  0.1195
Number of obs: 16288, groups:  Ship:EstSweepCat, 19

Fixed effects:
                                  Estimate Std. Error        df t value Pr(>
|t|)
(Intercept)                      1.490e+00  3.116e-02  1.891e+01   47.81   <2
e-16 ***
LogDepthCenter:EstSweepCatlong  -7.053e-02  4.606e-03  1.629e+04  -15.31   <2
e-16 ***
LogDepthCenter:EstSweepCatshort -1.462e-01  2.288e-03  1.628e+04  -63.89   <2
e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
                     (Intr) LgDpthCntr:EstSwpCtl
LgDpthCntr:EstSwpCtl -0.024
LgDpthCntr:EstSwpCts  0.015  0.000
```

Annex 7: Reviewer reports

Review of suitability of ICES workshop outputs to support advice in response to OSPAR request

Jens Rasmussen

OSPAR has requested advice from ICES to provide swept area estimates and abundance index outputs for all otter and beam trawl surveys in the Northeast Atlantic and regional seas based on the ICES Database of trawl surveys (DATRAS). The aim is to provide input to the assessment of OSPAR common indicators. Two specific components are detailed in order of priority

1. Swept area estimates for all hauls in the DATRAS database

2. Annual estimates of abundance of all species identified as sensitive, listed in the ICES WGBIODIV 2020 outputs – later refined to a collection of regional assessment lists per ecoregion defined during the ICES workshop WKCOFIBYC.

ICES has conducted two workshops to address approaches required to develop the advice: WKSAE-DATRAS – Workshop on the production of swept-area estimates for all hauls in DATRAS for biodiversity assessment and WKABSENS – Workshop on the production of annual estimates of abundance of sensitive species. This review examines the reports from the two workshops and assesses their suitability as basis for ICES advice to the OSPAR request. Given one report was formatted, but not published, and the other in draft, the main focus of the review is on methods and outputs rather than formatting and editorial issues.

The two workshops had terms of references to directly develop scripts and data products, but also extended to look at securing future data quality, methods, and documentation.

At the time of reporting, there was no (publicly or accessible) available data product from WKSAE-DATRAS or WKABSENS, so the review is based on the content of the reports and descriptions of methodology rather than detailed inspection of the data products themselves.

WKSAE-DATRAS focusses on the first part of the advice request – to generate swept area estimates for all relevant trawls in the DATRAS database. The report does however also highlight several considerations that needs to be made when using the estimates to determine abundance of species caught in the trawls.

The WKSAE-DATRAS report in particular addresses the challenge that there is a significant amount of haul records in the DATRAS database that does not include all the data required to calculate swept areas directly. Swept areas are estimated both for door and wind spreads of trawls, but many hauls in the database do not contain the data to provide both. Especially wing spreads are more challenging as many member countries participating in the surveys do not have sensor packages that allow recording of both door and wing positions on trawls.

Across the north sea IBTS survey, there are considerable variability in reporting between countries. Efforts to include more data has greatly increased across member countries in recent years (as summarised in Fig 4.12 in the report), but these were examined in detail during the workshop to identify best possible remedial action to increase inclusion of data.

Beam trawl surveys are overall treated differently with more fixed dimensions, although there are some minor variability in rigging between countries. WKSAE-DATRAS updated and included new data, covering the years 1985 to 2020. The updated beam trawl part of the output focuses on the requested offshore spatial coverage.

For each of the surveys, WKSAE-DATRAS examined relationships between door and wing spreads and relevant parameters such as towed distance, duration etc. to make it possible to derive and fill in missing estimates that the current DATRAS flexfile format would otherwise leave blank. Extracting an up to date dataset, methods from standard flex files, updated flex file approach and the previous approach for the OSPAR product were performed, and the updated data product include the same full spatial extent, and is aligned to the same format with two exception according to the report.

Overall, WKSAE-DATRAS has merged considerations from several previous approaches to provide the best possible estimates, examining as precise relationships as are possible, building upon work in other expert groups (WKBEAM, IBTSWG). Based on the information in the report, and the adoption of merged principles from previous exercises in expert groups and publication, the WKSAE-DATRAS output product for estimated swept areas is deemed suitable for use in advice.

The second part of the advice request, abundance indices of sensitive species, is addressed in the draft WKABSENS report.

The first step is to identify the sensitive species to include, and establish the criteria for inclusion or exclusion of records for which reliable swept area estimates can be obtained (from the first product)

While the original request from OSPAR refers to the list of sensitive species from WGBIODIV 2020 but it is understood that relevant staff from OSPAR were in agreement to utilise and updated list – Regional Assessment List in Annex 3 of the ICES Workshop on Fish Conversation and Bycatch Relevance (WKCOFIBYC). WKABSENS provide clear information on when species was classified as sensitive, based on the method presented in Rindorf *et al* (2020)[12] but including some additional species from some previous methods.

A full list of species that were considered is included in the report together with detailed information on additional selection criteria and count of the number of occurrences within the DATRAS dataset.

Rather surprisingly, WKABSENS has performed its own swept area calculation, using a separate method from the WKSAE-DATRAS product. There is no mention of the WKSAE-DATRAS report or data product in the report. It is not clear if the WKSAE product was not available at the time of WKABSENS. Nor is it clear which of the two products are intended to be presented as the first part of the OSPAR request – but providing one approach to deliver the swept area product and then using another approach to derive the abundance indices seems confusing since there is no comparison between the two products at the time of writing this report.

The remaining evaluation of the WKABSENS draft report is based on the assumption that this issue is addressed in the ICES advice process, and instead focuses on the approach to determine the abundance indices.

The report does detail a set of criteria for inclusion of data (Section 3.3), but they are highly reliant on users being experienced not only in the structure of DATRAS data, but also have additional business knowledge on the rationale for introduce some of the criteria (For example, removing hauls from NIGFS survey prior to 2007, since it is later referenced as 2006-2020 in table 3.5.1). This makes it somewhat difficult to interpret how much data were discarded/deselected as part of this process, and no summary is provided.

[12] Rindorf, A., Gislason, H., Burns, F., Ellis, J.R., Reid, D. 2020.Are fish sensitive to trawling recovering in the Northeast Atlantic? Journal of Applied Ecology 57: 1936-1947. doi: 10.1111/1365-2664.13693

Subsequently, there are clear and unambiguous classification of species, and description of approaches for providing abundance indices where appropriate. The criteria for inclusion or exclusion of surveys for a given species are clear and well considered.

Gear groupings used for the subsequent modelling approach are detailed, and clearly defined.

The statistical analysis and modelling efforts in WKABSENS has to deal with a challenging set of data that often include a wide range of zero observations (e.g. zero inflated species), sparse recordings, or variable reporting levels over time. Thus, there will inevitably be strong limitations to the number of species that can be fully assessed.

WKABSENS reports on a total of 55 species, inclusive of species for which stock assessments are available. These 55 species were based on the lists from different working groups, and exclude species that were observed 10 times or less. It isn't entirely clear if this is 10 times within individual surveys (in which case their rarity is variable since the number of hauls in different surveys vary considerably) or across all available data.

An initial binomial model approach allow comparison of the full time series of presence-absence information with two different reference period, providing an indication of whether rarer species are declining, stable, or increasing. This approach provide a rapid overview and extends the species list slightly compared to the use of existing indices and the GAM analyses also carried out by WKABSENS.

The two GAM model approaches increases the demand of data availability to 100 non-zero abundance hauls of a given species or population (several species divided into subpopulations), resulting in a reduction to 50 eligible species. Additional consideration about the reduction of habitat scale and time series fragmentation are relevant for the application of GAMs. However, for the majority of species, the full spatio-temporal model with gear interactions could be utilised, with a few using a simpler model without gear interactions.

Results from a VAST mode was applied to a single species, but it is not really clear from the report was used for anything or is considered a candidate method for further work.

Overall this work leads to an extensive table of the 50 sensitive species with abundance indices originating either from the GAM models or existing ICES assessments.

The existing ICES assessments retrieved from SAG is covering 16 of the 50 species, with 34 other sensitive species. This is a large amount of work, and the report provides a good methodology session to explain the application of models.

It would be beneficial to have a clearer summary of each step of inclusion criteria of data on each level from survey, haul, and species. The information is in there for the most, but is somewhat scattered, and the rationale for the criteria is not always explained.

Overall, the outputs of the two workshops are comprehensive, and represents a large amount of work by collective experts. My main concern is the deviation in using two different approaches in determining the swept area indices across the two products. It may be a very minor issue, but there appears to have been no comparison, so the production of advice on this basis should be done with some caution.

ILVO
Flanders Research Institute for Agriculture, Fisheries and Food

Animal Sciences
Ankerstraat 1
8400 Oostende, Belgium
T +32 59 56 98 75
www.ilvo.vlaanderen.be

Review of the WKSAE-DATRAS and WKABSENS reports

By Dr. Noémi Van Bogaert (ILVO), Ir. Heleen Raat (ILVO), MSc. Ir. Loes Vandecasteele (UGent)

Request from *(name organisation)*	ICES
Contact within organisation *(name, email, tel.)*	Lara Salvany
Request received	6/8/2021
Outcome of request required by client	1-2 page report with review comments
Request title	Review of the WKSAE-DATRAS and WKABSENS reports
Deadline date (requested)	Friday 20/8/2021

Flanders
is agriculture and fisheries

1 WKSAE-DATRAS report

The ToRs of the WKSAE-DATRAS were as follows:

a) harmonize the selection of surveys and time series available in DATRAS for biodiversity assessments;
b) develop a script to calculate swept-area indices for biodiversity assessments;
c) calculate swept-area indices and create a data product as input to OSPAR common indicators for fish and foodwebs;
d) update the DATRAS calculation document to include reference to the new data product and field for biodiversity assessments.

During the working group the Marine Scotland Science (MSS)/OSPAR approach was adopted (described in Chapter 2) which includes data quality checks for the information needed for the calculation of swept-area, and the DATRAS approach (described in Chapter 3), which depends solely on correctly reported data from the national institutes. Several large data gaps were identified, in particular for several years of the North Sea IBTS, for which the authors propose the MSS/OSPAR approach as the way forward.

Overall, the research conducted during the workshop to produce this report addresses the initial ToRS well and is scientifically sound. However, t**he structure of the report could be improved to increase the readability of the text** by shortly introducing the different Chapters/topics somewhere in the Introduction. Additionally, text could be added on how the original ToRs/objectives were addressed and what the link is with the outcomes presented in the report. Other minor comments/questions have been added into the report and in Table 1 below. Furthermore, editorial corrections were made directly in the report using track changes.

Table 1: Detailed comments/questions on the WKSAE-DATRAS report (also marked in the document as comments)

Page number	Comments WKSAE-DATRAS report
2	Please add and discuss relevant parameters for Beam Trawl Surveys (*i.e.* beam width and distance)
3	Please explain footnotes of table in description under table or add reference to OSPAR version
3-4	Please describe the structure of the report; what will be presented in the different chapters? How do the outcomes link up to the original ToRs (maybe this could be added somewhere at the end near the Conclusions?
4	Elaboration on bullet point number 3: 'target species' seems to be missing from the text?
4	Include full terminology when mentioned first time in text with abbreviation in brackets
6	Include full terminology when mentioned first time in text with abbreviation in brackets
7	Figure 2.3: explain meaning of purple circles in the flow chart
8	Table 2.1: - not clear what is meant with 'sweep-length' - mention standard groundspeed can be obtained from survey manuals - explain why reasoning behind methodology for swept area calculation is not in line with the advice from WGBEAM 2021
9	Reference style slightly differs between individual references
11	Table 3.3: please provide sufficient detailed information, so CPUE calculations can be reproduced if desired by readers.
12	Table 4.1: - Selection details are not provided in Annex 5? Please explain reasoning behind year selection procedure. For example BE-BTS: why are the years 2005, 2006, 2008, 2009 and 2020 left out? - Not clear which DYFS surveys are included.
14	Table 4.3: is it possible to include similar table for the Beam Trawl Surveys?
16	Check if Figure referencing is correct
16	Check version and R reference
17, 44	How shorts vs. long? Add in brackets?
18	Add data years in caption similar to Figure 4.1
19	Add data years in caption similar to Figure 4.1
20	Please explain 'warp length'
30	Table 4.5: please explain duplicates and missing (Belgian) data?
32	Add description of abbreviation when mentioned first time
37	Could you add information in the caption on what the red circle indicates?
38	Make reference style consistent (brackets or no brackets etc.)

Flanders
is agriculture and fisheries

39	Please be more consistent in the structuring of the text, especially with the order of the surveys in each chapter.
39	Table 5.1: please specify which BTS and DYFS surveys are included?
45	Figure windows with one caption should be together on one single page (Figure 5.6) Add R2 to regression lines on Figure 5.6
48	Reference? Extra info?
49	Table 6.1: Please explain * in table footnote. Why is the "BEBTS_not_in_previous" not included in the list as stated in the caption?

2 WKABSENS report

The ToRs of the WKABSENS report were as follows:

a) Consider the applicability of the ICES list of sensitive fish species for OSPAR FC1. Split the list into:
   - Species with existing ICES assessments (including reference points);
   - Species for which no ICES assessments currently exist;

b) Calculate swept-area indices and create a data product as input to OSPAR common indicators for fish and food webs;
   - Define criteria to select surveys and time-series for analysis;
   - Discuss and agree in algorithm(s) to infill missing data at genus or family level;
   - Agree on the approach to estimate single species population abundance density per year;

c) Discuss and agree on criteria of data capable to support formal assessment; The selection of assessment units will be informed by available information, data and knowledge from other ICES expert groups.

d) Calculate individual survey-based assessments for all sensitive species and create a data product to OSPAR informed by Table 2 Biological information, see Annex 2.

In Chapter 2 criteria for a species to be listed as sensitive were formulated and depending on the occurrence of the species in DATRAS, they were categorized into four groups (Chapter 3). Criteria for excluding specific surveys were discussed as well (Subtitle 3.5). Different approaches (e.g., binomial models, GAMs, VAST) to abundance assessment of sensitive species were explored in Chapter 4 and reference was made to existing stock assessment methods developed in other ICES working groups. For the selection of assessment units distributional maps and selected reports from assessment working groups were reviewed (Chapter 5). Lastly, individual survey-based indices for all sensitive species are presented in Annex 2 and discussed in Chapter 6.

Main comments:

1) While the main ToRs have been addressed by the working group, the readability of the report could be improved by sketching the outline of the different Chapters somewhere in the beginning of the text (or at the beginning of each Chapter) and **linking the different Chapters to each other and with the original ToRs/objectives**.
2) **More detail could be added to the Executive summary** (e.g. which methods were used to calculate abundance indices? What are their limitations)? to make it more informative. The executive summary should reflect the main outcomes from each Chapter/ToR.
3) **Chapter "4.4. VAST" lacks information/interpretation on the Figures that are presented**. In contrast to the binomial method & the GAMs which are well-explained we're missing info on why/when this method could be useful and why/when not (limitations).
4) **Review the Reference list:** see track changes in Table 2 below.

We have added other minor comments/questions to the report (and in Table 2). Furthermore, editorial corrections were made directly in the report using track changes.

Flanders
is agriculture and fisheries

Table 2: Detailed comments/questions on the WKABSENS report

Page number	Comments WKABSENS report
Ex Sum	Remove instructions when doing final review
Ex Sum	Should the heading of this page be "acronym 2021". If not? Leave it blank.
Ex Sum	Spell out numbers lower than ten in full in text
Ex Sum	Add something about the models as these make up a large part of the report e.g. "Different statistical modelling approaches (binomial, GAMs, VAST) were evaluated" or "Abundance indices were estimated using General Additive Models (GAMs) according to the approach by Berg et al. (2014)"
1	Footnote still needs to be added to the *
1	Explain where FC1 indicator stands for.
2	Explain acronym when mentioned the first time
2	Something was wrong with the sentence, check if I altered this correctly
2	Greenstreet et al. (2012)a or b? check with reference list
2	See comment above
3	when selecting all surveys and years/quarters etc. in DATRAS? Could you add a bit more info on how these numbers were obtained?
4	2012a or b? See also the other Greenstreet references in the rest of the document
9	Add meaning of acronyms when mentioned the first time
9	DATRAS is capitalized
9	Explain why beam trawls are not included in the FlexFiles
12	Make reference to parameters used in equation in text below
14	Explain acronym when mentioned the first time
15	Why? I see that NIGFS (<2007) hauls have been included in Tables 3.1.2, 3.5.1
16	Add time range?
18	First time GAMs are mentioned; refer to section 4.3.1 where background on this modelling approach is provided and add original reference: Hastie T, Tibshirani R (1990) Generalized additive models. Chapman and Hall, London
18	Explain acronyms when mentioned the first time
23	Define low occurrence? Is it the same criterion as below 4.2 ("Species that were observed ten times or less over the whole time period were excluded." ?) or is it different?
23-24	Table 4.1.1 When 'existing stock assessments use selected DATRAS data only' is stated, please explain more into detail why the GAM model is more appropriate than ICES assessment?
25	Check lay-out of the first row in this Table
27	No reference for this Figure?
29	Does the colour coding in the first column indicate anything? If not, remove. If they mean something please add more info into the Table's caption
30	Spell out numbers when placed at the beginning of a sentence
31	Is this the official jargon?
34	This should be Table 4.2.2?
35	Please write these abbreviations consistently with or without capitals throughout the text, tables and figures
35	There are currently no yellow highlights in the table, please change
37	Please add reference where the term GES comes from
38	The same is true for the GAM approach?
39	Wrong Figure number, I adjusted this
40	Please add in caption what the red line means (same comment for the following figures below)
43	Add this reference please: Hastie T, Tibshirani R (1990) Generalized additive models. Chapman and Hall, London
43	See previous comment on spatial autocorrelation and dependency. Mention these? Maybe add something in the Discussion section of this part 4.3.4. An alternative to GAM could be INLA
43	Can you add a bit more detail on the data exploration and (violation of) assumptions (if there) for this model
43	Could you add the name of the package is it "mgcv"?)? As well as the author
44	Add reference
44	Put Latin names in italic
51	Latin species names should be in italic
53	Mention considered time periods again in caption
57	Spell out fully first time
57	Add time period (as the first year cannot be read clearly from the graph)
57	Add version R
58	Same comment as above. This picture is not easy to read (small axes), maybe consider to break up in time periods instead of showing all years

Flanders
is agriculture and fisheries

58	Lacking reference to figures in text. Difficult to understand what we see
60	Figures 4.4.1-4.4.6 are not discussed anywhere in the main text. A short conclusion would help the reader in understanding the main outcomes of the VAST approach. Please also explain why the VAST model is described here? It wasn't mentioned as an approach in paragraph 4.1.
60	Axes unreadable, same comment as above. Please also add the time period in brackets
61	Table 5.1: please provide references upon which evidence is based.
62	Reference?
62	Refere to Innoray (WUR) & Raywatch (ILVO) projects?
62	'DATRAS'?
64 (References)	I rearranged the ICES references in a chronological order. Check list again
64	Double entry removed

Flanders
is agriculture and fisheries