Scottish Government
Riaghaltas na h-Alba
gov.scot

Manual for Version 3 of the Groundfish Survey Monitoring and Assessment Data Product

Scottish Marine and Freshwater Science Vol 8 No 18

S P R Greenstreet and M Moriarty

marine scotland

Manual for Version 3 of the Groundfish Survey Monitoring and Assessment Data Product

Scottish Marine and Freshwater Science Report Vol 8 No 18

S P R Greenstreet and M Moriarty

Published by Marine Scotland
ISSN: 2043-7722
DOI: 10.7489/1986-1

Marine Scotland is the directorate of the Scottish Government responsible for the integrated management of Scotland's seas. Marine Scotland Science (formerly Fisheries Research Services) provides expert scientific and technical advice on marine and fisheries issues. Scottish Marine and Freshwater Science is a series of reports that publishes results of research and monitoring carried out by Marine Scotland Science. It also publishes the results of marine and freshwater scientific work that has been carried out for Marine Scotland under external commission. These reports are not subject to formal external peer-review.

This report presents the results of marine and freshwater scientific work carried out by Marine Scotland Scinece.

Table of Contents

Manual for Version 3 of the Groundfish Survey Monitoring and Assessment Data Product

Simon P R Greenstreet and Meadhbh Moriarty
Marine Scotland, Marine Laboratory
375 Victoria Road, Aberdeen AB11 9DB

Executive Summary

Coastal European Union (EU) Member States (MS) operate groundfish surveys that cover almost the entire continental shelf of the Northeast Atlantic off Western Europe. Some survey work is also carried out in deeper shelf-edge waters and on sea mounts and plateaus further west in the wider Atlantic Ocean. These surveys use different vessels and different fishing gears, are undertaken at different times of year following various methodological approaches. The data are recorded and archived in a variety of different formats. The data obtained from many of these surveys are routinely uploaded and stored on an open access data portal on the ICES website: the DATRAS portal. The DATRAS portal allows for the uploading of data collected, recorded and archived in a variety of different forms and formats, but the resulting database structure required to accommodate this is far more complicated, and potentially confusing for users, than is really necessary, with many fields in the database simply explaining how other fields should be interpreted. Furthermore, the data stored on the DATRAS database have long been known to be affected by various data quality issues. These have not been addressed in a single co-ordinated process, rather it has been left to individual data users to make any corrections that they deem necessary, effectively resulting in multiple different versions being in circulation of what is in fact a single DATRAS data set.

The MSFD requires formal assessment of the state of marine ecosystems in EU waters. Fish communities constitute a key component in the structure and functioning of marine ecosystems, so formal assessment of the state of the fish component is, therefore, mandatory. To meet their MSFD obligations, EU MS have invariably nominated their groundfish surveys as part of their marine monitoring programmes. However, to meet the obligation for formal assessment at the required regional seas spatial scale, these issues of data standardisation and quality assurance noted in the groundfish survey data stored on DATRAS all need to be resolved. The Groundfish Survey Monitoring and Assessment (GSMA) data product sets out to address these issues; to derive a single set of fully standardised and quality assured data products for all the surveys operating in the Northeast Atlantic.

Two previous reports have been published by Marine Scotland, each linked to an earlier version of the data product. Moriarty et al. (2017) describe in detail the approach and the protocols used to derive each survey's data product. This report was essentially linked to the Version 1 GSMA data product, which, together with the Moriarty et al. (2017) report, was subjected to review by several ICES working groups, The OSPAR Fish and Cephalopod Technical Expert Group (Greenstreet and Moriarty, 2017), and by the national Data Providers. Issues raised were subsequently addressed leading to release of the Version 2 GSMA data product. This process is described by Greenstreet and Moriarty (2017). It was the Version 2 GSMA data product that was analysed for the OSPAR Interim Assessment 2017 (IA2017). However, further issues with the Version 2 data product were identified by some national Data Providers and by the indicator leads responsible for carrying out the assessments. It was not possible to deal with these issues in time for the IA2017, but they have since been addressed to derive a third version of data product. This report describes the process used to produce the Version 3 GSMA data product.

1. Introduction

The Marine Strategy Framework Directive (MSFD) places an obligation on European Union (EU) Member States (MS) to assess the state of fish species and communities, and the role of fish in marine food webs, as part of their aspirations to achieve good environmental status in the marine environment. Coastal EU MS have nominated their groundfish surveys as part of their monitoring programme commitments under the MSFD. For the most part, the data obtained from these surveys is uploaded to, and maintained on, an open access database, the DATRAS data portal housed on the ICES website. However, numerous issues with these groundfish survey data have been identified over the years (Daan, 2001; ter Hofstede and Daan, 2006; 2008). If these data are to support formal assessment mechanisms required by the MSFD, then a more rigorous approach to data standardisation and quality assurance is required. Moriarty et al. (2017) describe 19 surveys operating in waters throughout the OSPAR/MSFD Northeast Atlantic region. They go on to describe a set of rigorous quality assurance and data standardisation protocols that they applied to these data to derive what has subsequently been termed, the Version 1 Groundfish Survey Monitoring and Assessment (GSMA) data product.

The MSFD requires the first formal assessment of status (following the initial assessments in 2012) to be delivered in 2018. It requires status to be assessed at both regional (Northeast Atlantic) and subregional (e.g. Greater North Sea, Celtic Seas, Bay of Biscay and Iberian Coast, Wider Atlantic) spatial scales and charges the Regional Seas organisations with the task of coordinating the efforts of EU MS with coastlines bordering each subregion and region. In fulfilling this role, OSPAR have undertaken an Interim Assessment 2017 (IA2017) at the end of 2016 into the beginning of 2017. To meet the needs of scientists responsible for assessing the state of fish species and communities, and their role in food web structure and functioning, the Version 1 GSMA data product was revised to address known problems identified in this version of the data product by ICES survey working groups and following review by the national Data Providers, and to include further uploads of data (additional years) to the DATRAS portal. The issues addressed and major difference between the Version 1 and Version 2 data products are described by Greenstreet and Moriarty (2017).

The Version 2 GSMA data product was subsequently used in the OSPAR IA2017. However, in the weeks following the release of the Version 2 data product, several national Data Providers informed us of a number of issues with the data product that they had just noticed. There was no time to deal with these issues before the data was needed for analysis in the IA2017 process. It was felt that, for the most part, actual assessment outcomes based on three biodiversity indicators (FC1 - abundance of a suite of sensitive species; FC2 - the Large Fish Indicator (LFI); FC3 - the mean maximum length of fish in the community) and two food web indicators (FW3 - Typical Length (TyL); FW5 - Mean Trophic Level) would be largely unaffected by the data issues identified. However, it was also felt that leaving these issues in the Version 2 GSMA data product unresolved was also undesirable, given the intention that this data product be used as the main data source for future development of fish community indicators and the assessment process.

A Version 3 GSMA data product was, therefore, developed with the intention of addressing the specific issues raised. These were:

1. A subsampling raising factor issue, which would have resulted in underestimation of the density-at-length, particularly of smaller sized fish of generally abundant species (which was when sub-sampling was most often practiced), in the last two or three years of Scottish data in the CSScoOT1, CSScoOT4, GNSIntOT1 and GNSIntOT3 surveys. This error was linked to a software bug when using the electronic measuring boards which were introduced in recent surveys. The impact of this error was judged to be most influential in the two Celtic Seas surveys, and to have affected the LFI and TyL indicators most.
2. A subsampling raising factor rounding error was detected in the R code and this has been addressed. Again the data most likely to have been affected are the Scottish Celtic Seas and Greater North Sea surveys.
3. Corruption of the unique haul identifier code in the four Spanish surveys: BBIC(n)OT4; BBIC(s)OT1; BBIC(s)OT4; and WASpaOT3. This led to the exclusion of many trawl samples from the data products derived for these surveys because the samples affected were deemed to be invalid hauls because they could not be linked to any biological data. This error had no impact on the OSPAR IA2017 process because these indicators were not accepted as common indicators in Spanish waters, so no assessments were made.
4. An issue with the wing-spread estimation routine when the research vessel "*Gwendolyn*" was used in the GNSFraOT4 survey. The estimation model used in the Version 1 and Version 2 data products gave wing spread estimates of 14 m to 19 m, but the national Data Provider deemed these to be far too high. Instead a fixed width of 10 m is now assumed for all such samples. The effect of this change will be to have increased estimates of fish density at length by a factor of between 1.4 and 2.0.

5. The addition of three years of early data at the start of the GNSGerBT3 survey.
6. General updating of many surveys by the addition of data for the most recent surveys to the DATRAS data portal. Surveys updated with more recent data included GNSGerBT3 (2016), GNSNetBT3 (2016), GNSEngBT3 (2016), GNSIntOT1 (2017), GNSIntOT3 (2016), GNSFraOT4 (2016), CSScoOT4 (2016), CSIreOT4 (2016), CSNIrOT1 (2016), CSNIrOT4 (2016), CSBBFraOT4 (2016); and WAScoOT3 (2016).

This report, therefore, describes the Version 3 GSMA data product. It provides diagnostic plots exploring the variability in, and relationships between, key parameter values. Charts are provided showing the location of all trawl samples collected in each survey's standard monitoring programme (excluding samples collected before survey methodologies were fully established and excluding extreme short and extreme long trawl tow durations). Two criteria are applied to define a standard survey area for each survey. The criteria are designed to ensure that the area included within each survey's standard survey area is sampled reasonably frequently throughout the whole of each survey time series. The consequences of applying these criteria are fully explained. For each survey a brief history of the changes made between each version of the data product is presented.

2. Data Product Structure

The GSMA data product consists of two primary data tables for each survey, Sampling Information and Biological Information (see Moriarty et al. (2017a) for full details regarding the sources and derivation of the GSMA data product).

2.1. Sampling Information

Two sampling information data products are available. Both contain data only deemed to constitute "*standard monitoring programme*" data. Thus, only samples collected once the surveys had become fully established and routine, and collected using a standardised regular procedure (e.g. a single specified fish gear, fished for a specified time, at a specified speed) are included. For some surveys, earlier years have been excluded if survey procedures were considered not to have been fully standardised at the start of the full survey time series. Like-wise, trawl samples with duration less than 13 minutes or greater than 66 minutes have been omitted, as have samples collected using a non-standard fish gear (see Moriarty et al. 2017 for further details). The difference between the two files is that the second file consists of a sub-set of the first, being just the trawl samples collected from within those rectangles deemed to constitute the "*standard survey area*". Rectangles making up each survey's *standard survey area* have to meet two criteria:

1. They must have been sampled in at least 50% of years of the survey *standard monitoring programme* time series. Thus if the time series is 20 years long, only rectangles sampled in at least ten years are deemed part of the *standard survey area*; samples collected in rectangles not meeting this criterion have been omitted from the standard survey area file.
2. They must be sampled at least once in both the start and end periods of the time series, where these periods are defined as 20% of the times series duration. Thus in a 20 year time series, only rectangles sampled in at least ten years, and at least once in the first four years and once in the last four years of the times series will be deemed part of the *standard survey area*; samples collected in rectangles not meeting this second criterion have also been omitted from the standard survey area file.

Moriarty et al., (2017) introduced and applied the first of these criteria. The second was added by Greenstreet and Moriarty (2017) when deriving the Version 2 data products used in the OSPAR IA2017.

In defining these periods, the result of the division involved has always been rounded up. Thus, for example, in a 17 year time series, 50% is 8.5, so this has been rounded up to nine. The 20% start and end periods would arithmetically both be 3.4 years long, so again this has been rounded up to four years. Thus in such a time series, running perhaps from 1999 to 2015, to be included in the *standard survey area*, rectangles would have to have been sampled in at least nine years, and at least once in both the four year start and end periods, 1999 to 2002 and 2012 to 2015.

Both types of file have been included because the method for defining the *standard survey area* is based on whether ICES rectangles to be included meet the specified sampling criteria or not. This is appropriate where ICES rectangles constitute the basis for the survey design, which is generally the case in the northern parts of the Northeast Atlantic area, but may be deemed less appropriate in southern parts of the Northeast Atlantic area where the coastal continental shelf is much narrower and surveys tend to be stratified along depth bands, not ICES rectangles. Thus the *standard survey area* data products might be more appropriate to support the assessment of environmental/fish community status in the Greater North Sea subregion and northern parts of the Celtics Seas subregion, but the full *standard monitoring programme* data product might be more useful in the southern Celtic

Seas and in the Bay of Biscay and Iberian Coast subregion. The provision of both types of product was also requested as part of the consultation process with ICES and the various data providers.

The structure of the "Sampling Information" data files is illustrated in the table below. The column 'Type' indicates the recommended field format for a fixed format file structure; thus 'A27' indicates an alphanumeric field 27 characters in length, 'S' indicates a small value integer numeric field, 'L' indicates a large value integer numeric field, and 'N' indicates a real number numeric field. Where appropriate, units are indicated for numeric fields. A brief description of each field is given with further detail provided in the notes following the table.

Field	Type	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)
HaulDur(min)	S	min	Duration of fishing operation[3]
ShootLat(decdeg)	N	Deg.	Latitude in decimal degrees of the haul shoot position[4]
ShootLong(decdeg)	N	Deg.	Longitude in decimal degrees of the haul shoot position[4]
ICESStSq	A12		ICES statistical rectangle where gear was shot
SurvStratum	A12		Stratum tag for stratified surveys[5]
Depth(m)	N	m	Depth tag assigned to the haul[6]
Distance(km)	N	km	Tow distance[7] ($d_{H,TOW}$)
WingSpread(m)	N	m	Mean distance between the wings during fishing operation[8,11] ($d_{H,WING}$)
DoorSpread(m)	N	m	Mean distance between the doors during fishing operation[9,12] ($d_{H,DOOR}$)
NetOpen(m)	N	m	Mean head-line height above seabed during fishing operation[10,13] ($d_{H,HEIGHT}$)
WingSwptArea(sqkm)	N	km^2	Area of seabed swept by the net[14] ($A_{H,WING} = d_{H,TOW} \times (d_{H,WING}/1000)$)
WingSwptVol_CorF	N		Multiplier ($1 / (d_{H,HEIGHT}/1000)$): converts to 'density by wing-swept volume'[15]
DoorSwptArea_CorF	N		Multiplier ($d_{H,WING} / d_{H,DOOR}$): converts to 'density by door-swept area'[16]
DoorSwptVol_CorF	N		Multiplier ($d_{H,WING} / (d_{H,DOOR} \times (d_{H,HEIGHT}/1000))$): converts to 'density by door-swept volume'[17]

Notes for Sampling Information

1. This is a unique tag assigned to each haul. Using the survey acronym avoids conflict where the same haul number is used by more than one survey participant. Using 'Ship' avoids conflict where more than one vessel might be used in the survey by the same country. 'Haul No' is the same haul number used in the original national data set so hauls can still be related to original data.
2. All date components kept separate so that queries can be run on any individual component.
3. Time of hauling can be established by adding haul duration to time ("HourShot" & "MinShot") of shooting.
4. This is the latitudinal and longitudinal position in decimal degrees (e.g. 56.4333°N -01.7895°W) where the haul was shot. Ideally a mid-trawl position would be given, but haul positions were frequently missing. Only the shoot position was supplied for all hauls, although in some instances, this is an arbitrary position as it coincides with the central point of the nominal ICES statistical rectangle where actual shoot position data were missing or incorrect.
5. This will be the same as the ICES statistical rectangle (identical to "ICESStSq") where ICES statistical rectangles constitute the survey strata (e.g. the North Sea IBTS).
6. Each haul will have a depth assignation. In most cases this is real data, either an average depth during the fishing operation, or a depth at the shoot position. But where depth data were absent in the original data, this will have been estimated. See Moriarty et al. (2017) for full details.
7. This is the distance along the seabed that the trawl was towed. The values in this field will have been derived through several different procedures. See Moriarty et al. (2017) for full details.
8. This is the mean distance between the wings of the net while the gear was towed between the shoot and haul positions. The values in this field will have been derived through several different procedures. See Moriarty et al. (2017) for full details.
9. This is the mean distance between the trawl doors while the gear was towed between the shoot and haul positions. The values in this field will have been derived through several different procedures. See Moriarty et al. (2017) for full details.
10. This is the mean height of the net headline above the seabed while the gear was towed between the shoot and haul positions. The values in this field will

have been derived through several different procedures. See Moriarty et al. (2017) for full details.

11. For a beam trawl survey, the value in this field will be the width of the beam trawl.
12. For a beam trawl survey, this field is not strictly applicable. The value in this field will again be the width of the beam trawl, and so identical to the value in the “WingSpread(m)” field.
13. For a beam trawl survey, the value in this field will be the height of the beam trawl.
14. The ‘standard’ density values provided in the Biological Information are based on the area of seabed swept by the net, as this is deemed most appropriate for the majority of species sampled (Fraser et al., 2007). If for any reason these standard density data are considered inappropriate, then these ‘standard’ density estimates can be adjusted by multiplying them by an appropriate correction factor. Likely correction factors required are given in next three fields. Dividing by 1000 converts the wing spread distance in m to the equivalent distance in km.
15. For pelagic fish species, or even perhaps some bentho-pelagic species, densities based on the volume of water filtered by the net could be deemed to be more appropriate for some indicators. Multiplying the ‘standard’ density estimates in the Biological Information database by this correction factor will provide the required adjustment. Dividing by 1000 converts the headline height distance in m to the equivalent distance in km.
16. For the majority of demersal fish species, the area swept by the net is the appropriate swept area to use to estimate density. Only for haddock and whiting is there evidence of substantial herding by the trawl doors, such that wing swept densities infer an apparent catchability in the trawl of >1. Density estimates for species deemed likely to be herded by the trawl doors could be considered more appropriate; if so then multiplying the ‘standard’ density estimates in the Biological Information database by this correction factor will provide the required adjustment. There is no need to divide both measurements by 1000 to convert to km as this would simply cancel out.
17. Pelagic species might also be considered likely to be herded by the trawls doors, and as stated above, volume-filtered density estimates could be deemed more appropriate. Where both considerations are deemed pertinent, multiplying the ‘standard’ density estimates in the Biological Information database by this correction factor will provide the required adjustment. The first ratio, wing distance : door distance, needs no conversion to km as these would simply cancel. Only the headline height distance needs to be divided by 1000 therefore.

2.2. Biological Information

Again two types of Biological information file are provided, and since both type of file are provided to accompany each type of Sampling information file, this gives a total of four Biological information files for each survey. The principal file type has the kNN label. Data in these files include the results of the application of the k-Nearest-Neighbour (kNN) procedure to resolve species density information to species density-at-length information (fish not measured, density data is for the whole catch of the species in question) and to resolve coarse taxonomic resolution level (genus or family-level, e.g Callionymus) density-at-length data to constituent species level (e.g. *Callionymus lyra*, *C. maculatus*, and *C. reticulatus*) density-at-length data (see Moriarty et al. (2017) for more details). The second file type, given the tag baseline, holds the original unresolved data (species density data and coarse taxonomic resolution level density-at-length data. Data in the baseline file types have, therefore, undergone the full quality assurance process described by Moriarty et al. (2017), with the exception of the application of the k-NN procedure. Again the provision of the baseline file type was in response to feedback from the data providers and from potential users of the data product.

The data file structure described in the table below relates primarily to the kNN file type. Some fields are absent in the baseline file type. The notes indicate which fields are missing and explains why.

Field		Unit	Description
HaulID	A27		Unique haul identifier (SurveyAcronym/Ship/Year/HaulNo)[1] (*H*)
SpeciesSciName	A45		Unique species name for each species sampled across the NE Atlantic[2] (*S*)
FishLength(cm)	S	cm	Integer numbers indicating fish length to the 'cm below'[3] (*L*)
IndivFishWght(g)	N	g	Estimated weight of individual fish of specified species and length[4] ($W_{S,L}$)
Number	N		Total number of fish of specified species and length in the catch[5] ($N_{S,L,H}$)
DensAbund(N_sqkm)	N	km^{-2}	Abundance density estimate[6,8] ($D_{nos,S,L,H} = N_{S,L,H} / A_{H,WING}$)
DensBiom(kg_Sqkm)	N	kg km^{-2}	Biomass density estimate[7,8] ($D_{biom,S,L,H} = (N_{S,L,H} \times W_{S,L}) / A_{H,WING}$)

Notes for Biological Information

1. This is a unique tag assigned to each haul. This field is identical to the field with the same name in the Sampling Information data table. This is the relational field linking these two tables.
2. Species names are the accepted scientific name as defined in the World Register of Marine Species (WoRMS). In the baseline file type, this field is

simply called SciName, because not all identification tags in the database are to species taxonomic resolution level.

3. All lengths in the data base are “to the cm below”: all fish of 11.0 to 11.9 cm, therefore, assigned a length of 11 cm. Effectively, therefore, this is an integer field.
4. This is the mean weight of an individual fish of specified species and length derived from a weight at length relationship of the form $W_{S,L} = \alpha_S L^{\beta_S}$. Since all recorded lengths are to “the cm below”, the individual mean weights for each length class of each species are calculated for the half-centimetre; e.g. specified weight for a fish of recorded length 11 cm is the weight calculated for a fish of 11.5 cm from the weight at length relationship, this being the probable mean length of all fish between 11.0 and 11.9 cm. This field is missing in the baseline file type because species-specific weight at length relationships could not be applied where the fish in question have either not been identified to species, or measured to a length category, or both.
5. This is the number of fish of specified species and length obtained in the trawl sample. This is either the actual count or an estimate derived from the raising of a known sub-sample.
6. This is the local point abundance density estimate, the number of fish of species (*S*) and length (*L*) per square kilometre estimated at the spatial location of trawl sample (*H*). This is obtained by dividing the species total catch number at length ($N_{S,L,H}$) by the area swept by the net ($A_{H,WING}$).
7. This is the local point biomass density estimate, the biomass of fish of species (*S*) and length (*L*) per square kilometre estimated at the spatial location of trawl sample (*H*). This is obtained by dividing the species total catch weight at length ($N_{S,L,H}$ x $W_{S,L,H}$) by the area swept by the net ($A_{H,WING}$). This field is missing in the baseline file type because species-specific weight at length relationships could not be applied where the fish in question have either not been identified to species, or measured to a length category, or both. Thus estimates of individual fish weight could not always be determined.
8. As detailed above, if other density estimates are required (e.g. density as number/biomass per cubic metre of water filtered by the net, density as number/biomass per square metre of seabed swept by the gear, density as number/biomass per cubic metre of water filtered by the gear), then these density estimates need to be multiplied by one of the three correction factors given in the Sampling Information table for the haul in question.

3. Overview of Data Products

This section presents a summary overview of the OSPAR Groundfish Survey Monitoring and Assessment data products. Full details regarding the derivation of the data products are given in Moriarty et al. (2017). Greenstreet and Moriarty (2017) update this document describing the differences between the initial Version 1 data products and the Version 2 products used in the OSPAR IA2017.

3.1. Sampling Information

For several surveys the number of rectangles sampled in 50% of years in the time series, and/or in the 20% of years that constituted the early and late periods of the survey time series, was markedly lower than the number of rectangles actually sampled at any time by the survey. Moriarty et al. (2017) concluded that there was, therefore, a need to define a *standard survey area* for each survey. To be included within a survey *standard survey area*, rectangles had to be sampled in at least 50% of years in which the survey was carried out and sampled at least once in at least one year in both the start (early) and end (late) periods of the time series, each period representing at least 20% of the full time series. For example, consider a 30 year time series, 1986 to 2015. To be included in the *standard survey area* each rectangle would have to be sampled in at least 15 years, and at least once in the two six-year periods 1986 to 1991 and 2010 to 2015, to meet both criteria. In some instances, by excluding the earliest years of the survey, generally the years prior to the survey in question becoming fully established, the spatial extent of the *standard survey area* could be increased; more rectangles would end up meeting the two inclusion criteria.

Table 1 summarises the data content of each groundfish survey contributing to the data product and illustrates the differences between the full *standard monitoring program* and *standard survey area* data products for each survey. Table 2 indicates the range (minimum and maximum values) of some of the key parameters held in the Sampling Information tables for each survey. The information in this table relates to the full *standard monitoring program* survey data products. A final concluding table (Table 3) summarises the outcome of the processes described to determine the final *standard survey area* data products.

Then in a series of individual survey subsections key sampling information is displayed in two figures. The first figure in each survey subsection illustrates key relationships between selected trawl sample information parameters and pertinent information regarding the survey in question, including:

i. The relationship between tow duration and tow distance, with the bounds representing the minimum and maximum acceptable tow speeds shown;
ii. The relationship between tow duration and the area swept between the wings of the trawl;
iii. The relationship between tow distance and the area swept between the wings of the trawl;
iv. The frequency distribution of tow duration;
v. The number of trawl samples collected in each year of the survey;
vi. The number of ICES statistical rectangles sampled by the survey in each year;
vii. The number of years that each individual ICES statistical rectangle was sampled over the course of the survey;
viii. Box and whisker plots of median tow speed (upper and lower quartiles indicated by the box, 95% of data range by the whisker, and outliers as dots);
ix. Box and whisker plots of the median multiplier value to convert wing-swept area density to wing-swept volume density (upper and lower quartiles indicated by the box, 95% of data range by the whisker, and outliers as dots);
x. Box and whisker plots of the median multiplier value to convert wing-swept area density to door-swept area density (upper and lower quartiles indicated by the box, 95% of data range by the whisker, and outliers as dots);
xi. Box and whisker plots of the median multiplier value to convert wing-swept area density to door-swept volume density (upper and lower quartiles indicated by the box, 95% of data range by the whisker, and outliers as dots);
xii. The relationship between wing-spread and door-spread;
xiii. The relationship between wing-spread and net-opening;
xiv. The relationship between door-spread and net-opening.

Note that where single values for WingSpread, DoorSpread, and NetOpen have been assumed, as for the beam trawl surveys, plots ix to xiv will simply show a single datum. The second figure shows the locations of all the trawl samples plotted on a chart for each survey.

Next the steps taken to derive *standard survey area* data product are outlined. For each survey, the data excluded in reducing the *standard monitoring program* data products to the *standard survey area* data products are described, along with the reasons for exclusion. Finally, a plot showing the survey's *standard survey area* is provided.

Deriving a standard survey area for each survey, and on occasion determining the optimal survey period that provided the best compromise between survey temporal range and spatial coverage, resulted in the exclusion of 1989 samples across the 19 surveys addressed to date. This represented a 4.6% loss of data, from the full standard monitoring programme data set (which already excludes any trawl samples of non-standard duration, and trawl samples collected before individual survey protocols were fully standardized: see Moriarty et al. 2017 for details) which was deemed necessary in order to derive a standard monitoring data product. The original data set of 43,383 samples was reduced to the monitoring data product containing data obtained from 41,394 otter and beam trawl samples. The monitoring data product holds data collected from 370 ICES statistical rectangles across the OSPAR Northeast Atlantic Region. These figures exclude the survey data for the four Spanish surveys, which still have unresolved problems (see Sections 3.1.14, 3.1.15, 3.1.16 and 3.1.19). Rectangles around the Iberan coast and on the Porcupine Bank in the Wider Atlantic, included in the GSMA Version 2 data product used in the OSPAR IA2017 are currently not included in this Version 3 data product.

Survey	Start Year	End Year	Breaks	No. Years	Number of Trawl Samples		Number of ICES Statistical Rectangles	
					Full SMP	SSA only	Full SMP	SSA only
GNSGerBT3	1998	2016	2000, 2006	17	908	820	30	19
GNSNetBT3	1999	2016	None	18	2665	2542	123	91
GNSEngBT3	1990	2016	None	27	2467	2337	31	15
GNSIntOT1	1983	2017	None	35	13892	13555	195	171
GNSIntOT3	1998	2016	None	19	6254	6170	182	168
GNSFraOT4	1988	2016	None	29	2544	2513	16	15
CSEngBT3	1993	2015	None	23	2445	2378	32	23
CSScoOT1	1985	2016	None	32	1795	1492	69	39
CSScoOT4	1995	2016	2010	21	1370	1071	105	42
CSIreOT4	2003	2016	None	14	2290	2119	78	52
CSNIrOT1	1992	2016	None	25	1234	1125	19	12
CSNIrOT4	1992	2016	2008	24	1242	1106	19	12
CSBBFraOT4	1997	2016	None	20	2798	2742	74	66
BBIC(n)SpaOT4			None					
BBIC(s)SpaOT1			1996, 2003					
BBIC(s)SpaOT4			2013					
BBICPorOT4	2002	2014	2003, 2004, 2012	10	866	839	22	18
WAScoOT3	1999	2016	2000, 2004, 2010	15	613	585	13	8
WASpaOT3			None					

Table 1: Basic survey information. Number of records (equals the number of unique HaulIDs and unique Year/Month/Day/Time/Ship combinations). Time series start and end years, breaks in the time series if any, number of years in the time series. The total number of trawl samples collected and the total number of ICES statistical rectangles sampled in the whole standard monitoring programme (defined start and end years and standardised trawl duration of between 13 to 66 minutes) and from within the standard survey area (ICES statistical rectangles sampled in at least 50% of years that the survey was undertaken and at least once in both the start and end phases of the standard monitoring programme, in turn defined as periods of time equivalent to 20% of the number of years that the survey was undertaken).

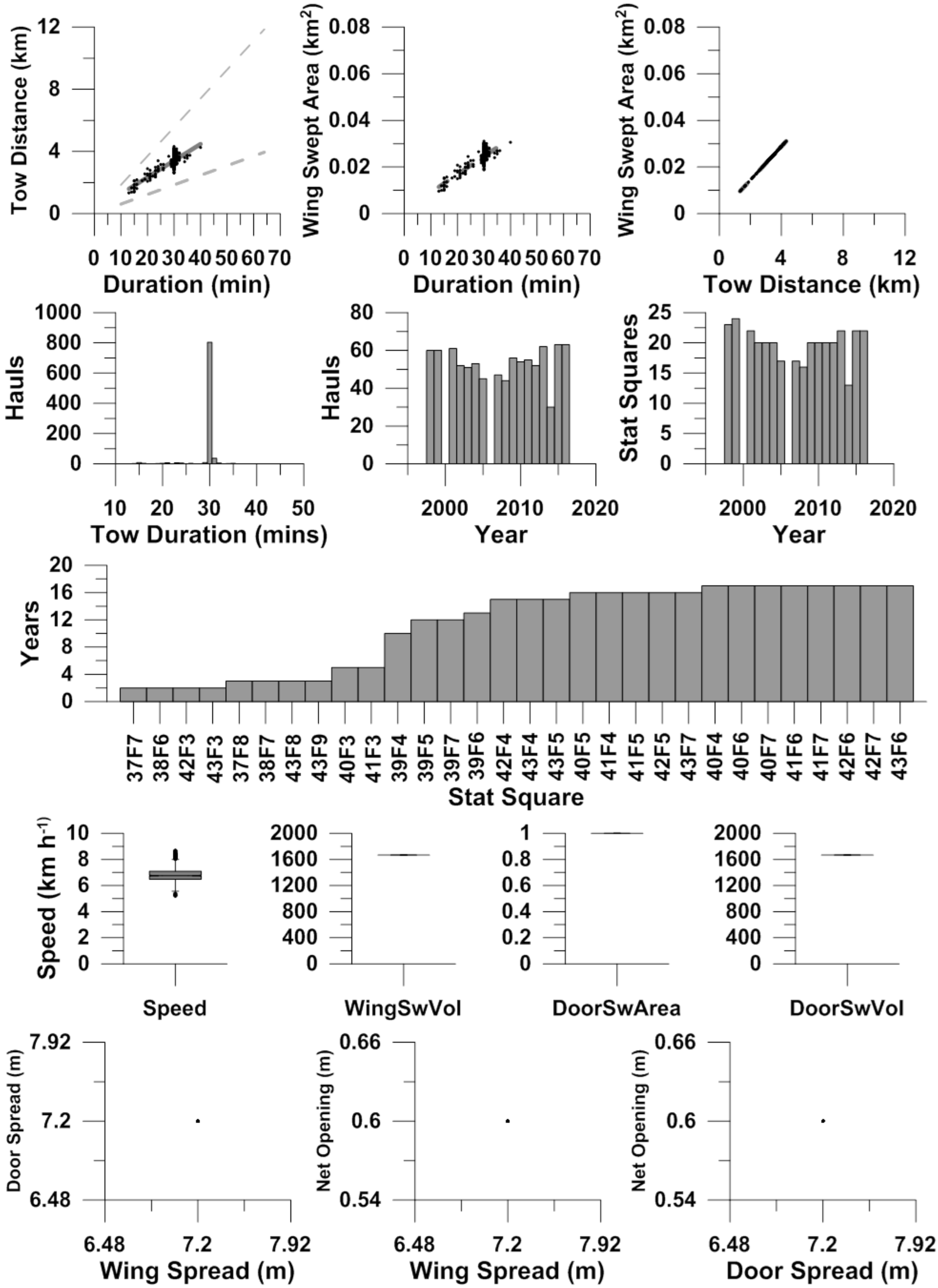
Survey	Lat. south	Lat. north	Long. west	Long. east	Wing min.	Wing max.	Door min.	Door max.	NetO min.	NetO max.	SwptA min.	SwptA max.
GNSGerBT3	54.1840	57.4897	3.5958	9.3153	7.20	7.20	7.20	7.20	0.60	0.60	0.00962	0.03107
GNSNetBT3	51.4713	60.7518	-3.3223	8.2068	8.00	8.00	8.00	8.00	0.80	0.80	0.01203	0.08584
GNSEngBT3	48.7183	53.5917	-2.8433	2.7958	4.00	4.00	4.00	4.00	0.53	0.53	0.00593	0.03024
GNSIntOT1	49.6013	61.5955	-3.9700	12.8764	10.00	29.00	38.00	143.00	1.90	9.70	0.02220	0.21765
GNSIntOT3	51.5298	61.8833	-3.9640	12.8754	13.00	27.08	46.00	110.00	1.80	7.90	0.02172	0.11343
GNSFraOT4	49.3167	51.2567	-1.8030	2.4100	10.00	19.00	40.00	85.82	2.10	6.00	0.01018	0.06575
CSEngBT3	50.5350	54.8033	-8.0067	-3.1390	4.00	4.00	4.00	4.00	0.53	0.53	0.00445	0.02650
CSScoOT1	53.4833	60.6000	-10.3192	-2.0167	10.00	26.00	40.00	129.00	3.30	7.10	0.02715	0.21638
CSScoOT4	48.0667	60.6333	-11.5500	-2.6833	13.00	25.00	50.00	118.00	2.80	6.70	0.02470	0.21119
CSIreOT4	50.0100	56.4900	-12.8410	-3.4590	12.00	31.00	30.00	153.00	3.00	7.00	0.03043	0.11575
CSNIrOT1	52.2923	54.7310	-6.1108	-3.4867	10.69	18.01	21.45	47.90	2.60	3.40	0.02073	0.10868
CSNIrOT4	52.2924	54.8112	-6.1068	-3.4750	10.84	18.40	21.90	46.60	2.40	4.00	0.01836	0.12418
CSBBFraOT4	43.3938	51.8321	-11.3465	-1.2567	13.00	30.00	47.00	120.10	2.50	6.20	0.02853	0.10581
BBIC(n)SpaOT4												
BBIC(s)SpaOT1												
BBIC(s)SpaOT4												
BBICPorOT4	36.7883	41.8133	-10.1183	-7.4216	15.10	15.10	45.70	45.70	4.60	4.60	0.01640	0.07829
WAScoOT3	55.8203	58.3190	-15.9157	-12.9860	17.00	26.00	82.00	119.00	3.20	6.20	0.02673	0.11286
WASpaOT3												

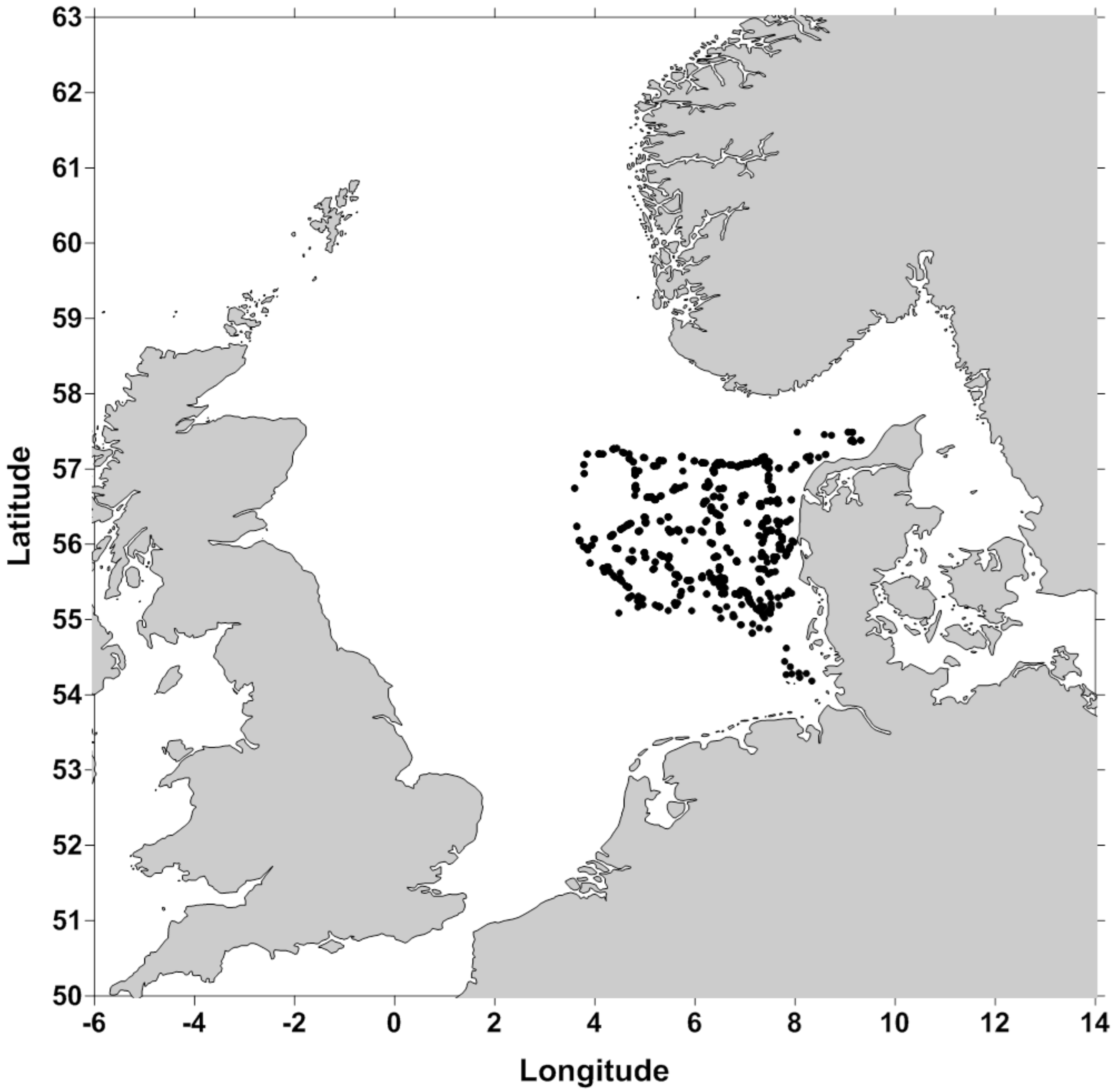
Table 2: Minimum (Min) and maximum (Max) Wing-spread (Wing), Door-spread (Door), Net-opening (NetO), Swept area (SwptA), Latitude (Lat: min = south, max = north) and Longitude (Long: min = west, max = east) values recorded in the database in each survey.

Survey	Samples	Start	End	Years	Rectangles
GNSGerBT3	820	1998	2016	17	19
GNSNetBT3	2542	1999	2016	18	91
GNSEngBT3	2337	1990	2016	27	15
GNSIntOT1	13555	1983	2017	35	171
GNSIntOT3	6170	1998	2016	19	168
GNSFraOT4	2513	1988	2016	29	15
CSScoOT1	2378	1993	2015	23	23
CSScoOT4	1492	1985	2016	32	39
CSIreOT4	1071	1995	2016	21	42
CSNIrOT1	2119	2003	2016	14	52
CSNIrOT4	1125	1992	2016	25	12
CSEngBT3	1106	1992	2016	24	12
CSBBFraOT4	2742	1997	2016	20	66
BBIC(n)SpaOT4					
BBIC(s)SpaOT1					
BBIC(s)SpaOT4					
BBICPorOT4	839	2002	2014	10	18
WAScoOT3	585	1999	2016	15	8
WASpaOT3					

Table 3: Summary of data held in the version 3 OSPAR Groundfish Survey Monitoring and Assessment *standard survey area* data products for each survey.

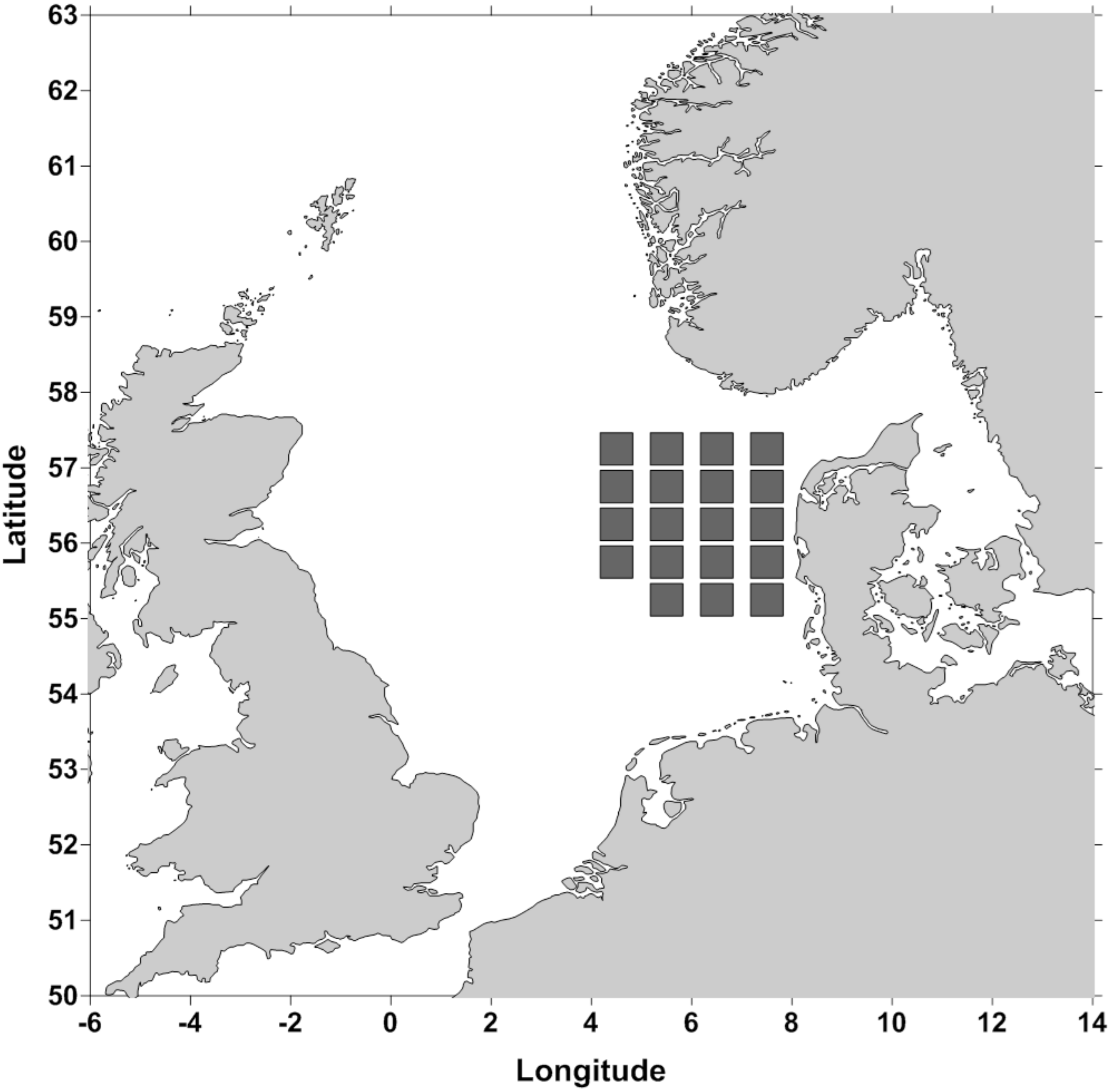
3.1.1. GNSGerBT3



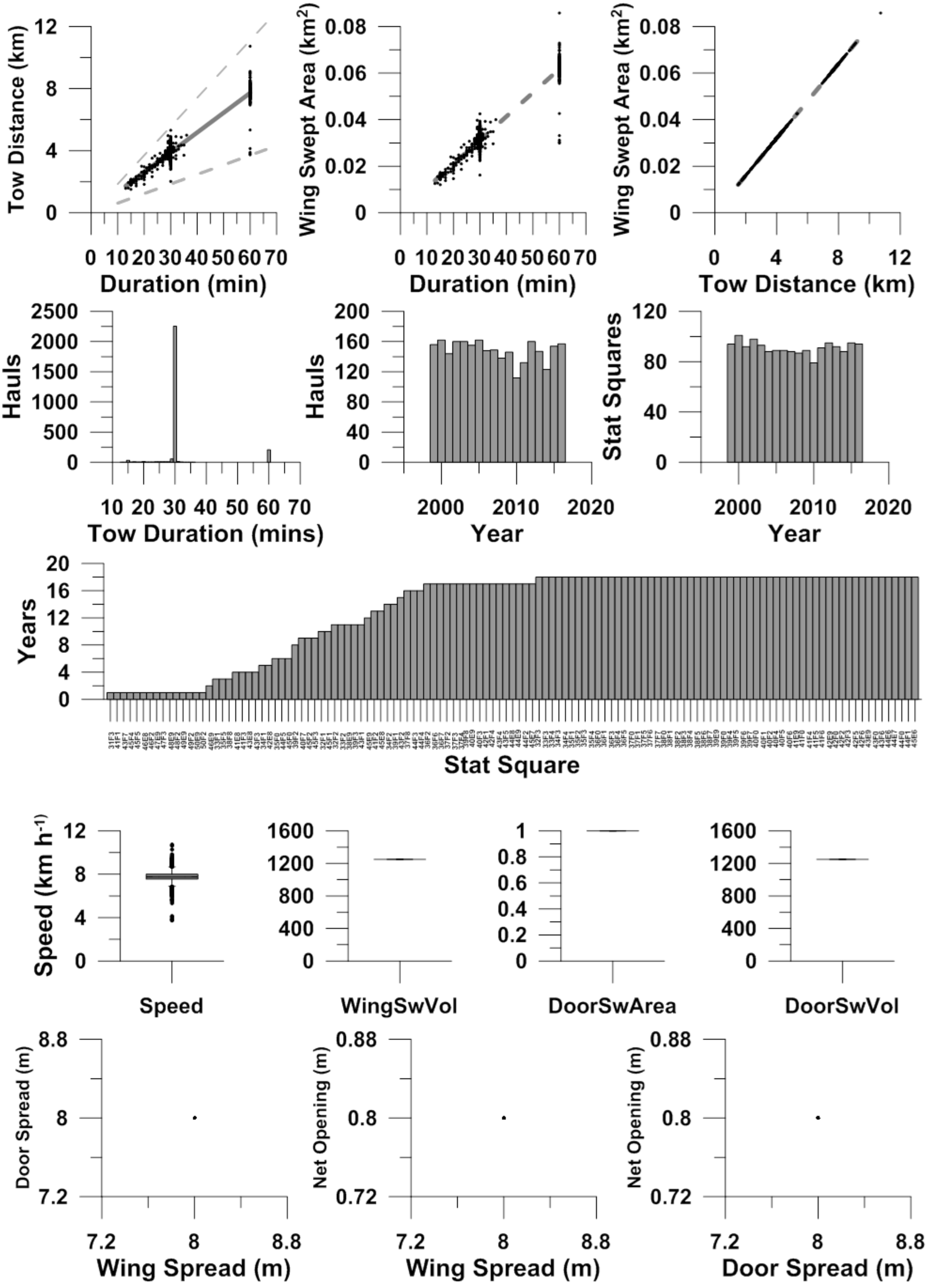


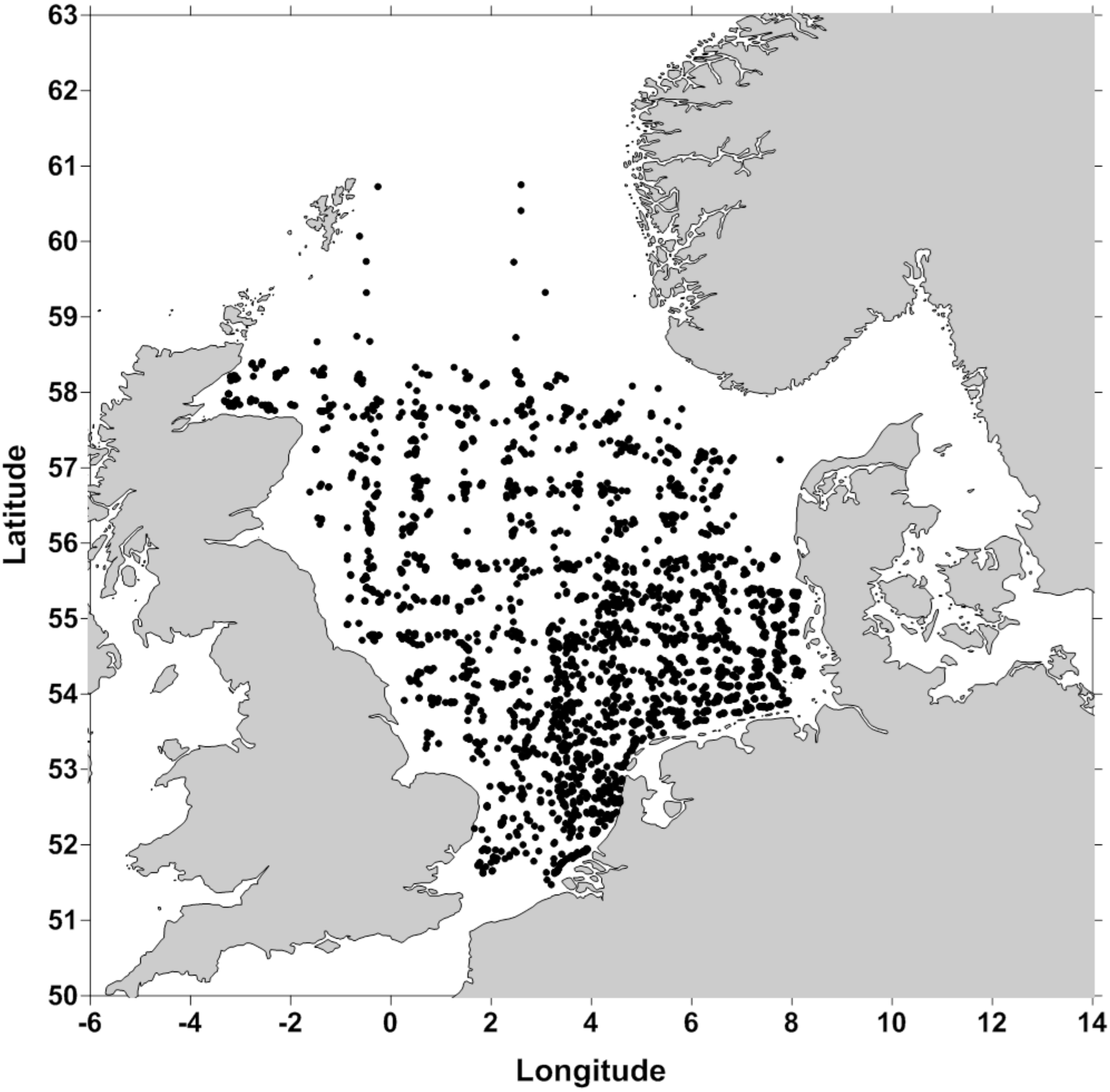
Overview of the cleaned GNSGerBT3 data revealed no inconsistency with the conclusions presented in Moriarty et al. (2017). Moriarty et al. (2017) only analysed data from 2003 to 2015. Subsequently, data for 2002 was uploaded to the DATRAS database, giving a time series of 2002 to 2015, and these data were included in the Version 2 data product analysed for the OSPAR IA2017 (Greenstreet and Moriarty 2017). More recently, the data provider has added an additional four years of data onto DATRAS: 1998, 1999, 2001 and 2016, no data in 2000. These recent additions of data to the ICES DATRAS database means that the survey now spans the period 1998 to 2016 with breaks in 2000 and 2006, giving 17 years of data.

The 50% of years rule requires ICES statistical rectangles to be sampled in nine years or more to be included in the *standard survey area*. Of the 30 rectangles sampled over the course of the entire survey (an increase of four rectangles over the Version 2 data product used for the OSPAR IA2017), only 20 met this criterion. Data for the ten rectangles (37F7, 37F8, 38F6, 38F7, 40F3, 41F3, 42F3, 43F3, 43F8 and 43F9) failing the 50% of years rule were excluded amounting to 68 hauls. A further rectangle, 39F4, failed the criterion of having to be sampled at least once in the four year period 1998-2001 that constituted the early 20%ile of the time series. Data for this rectangle, amounting to 20 hauls, were also excluded. All remaining 19 rectangles met the criterion of having been sampled at least once in the four year period 2013 to 2016 that constituted the late 20%ile of the time series. Rectangle 39F4 was first sampled in 2004 and because Version 2 of the database used in the IA2017 only contained GNSGerBT3 data from 2002 onwards, this was sufficiently early to meet the 20%ile criterion of the Version 2 time series (Greenstreet and Moriarty, 2017). This rectangle was, therefore, actually included in the IA2017 assessment based on the GNSGerBT3 *standard survey area* Version 2 data product. In this Version 3 data product, the original data set of 908 GNSGerBT3 samples was reduced to 840 records collected in the 17 years, 1998 to 1999, 2001 to 2005 and 2007 to 2016 from 19 ICES statistical rectangles. The potential increase of three additional years of data across 19 rectangles, associated with the inclusion of the new early data added to the database, was deemed to outweigh the loss of 13 years of data from the one rectangle, 39F4 now excluded from the *standard survey area*, particularly since this rectangle is also covered by the GNSNetBT3 survey *standard survey area*. The rectangle lost is the one missing in the southwest corner of five by four square of rectangles shown in the figure below.


63
62
61
60
59
58
57
56
55
54
53
52
51
50
Latitude
-6
-4
-2
0
2
4
6
8
10
12
14
Longitude

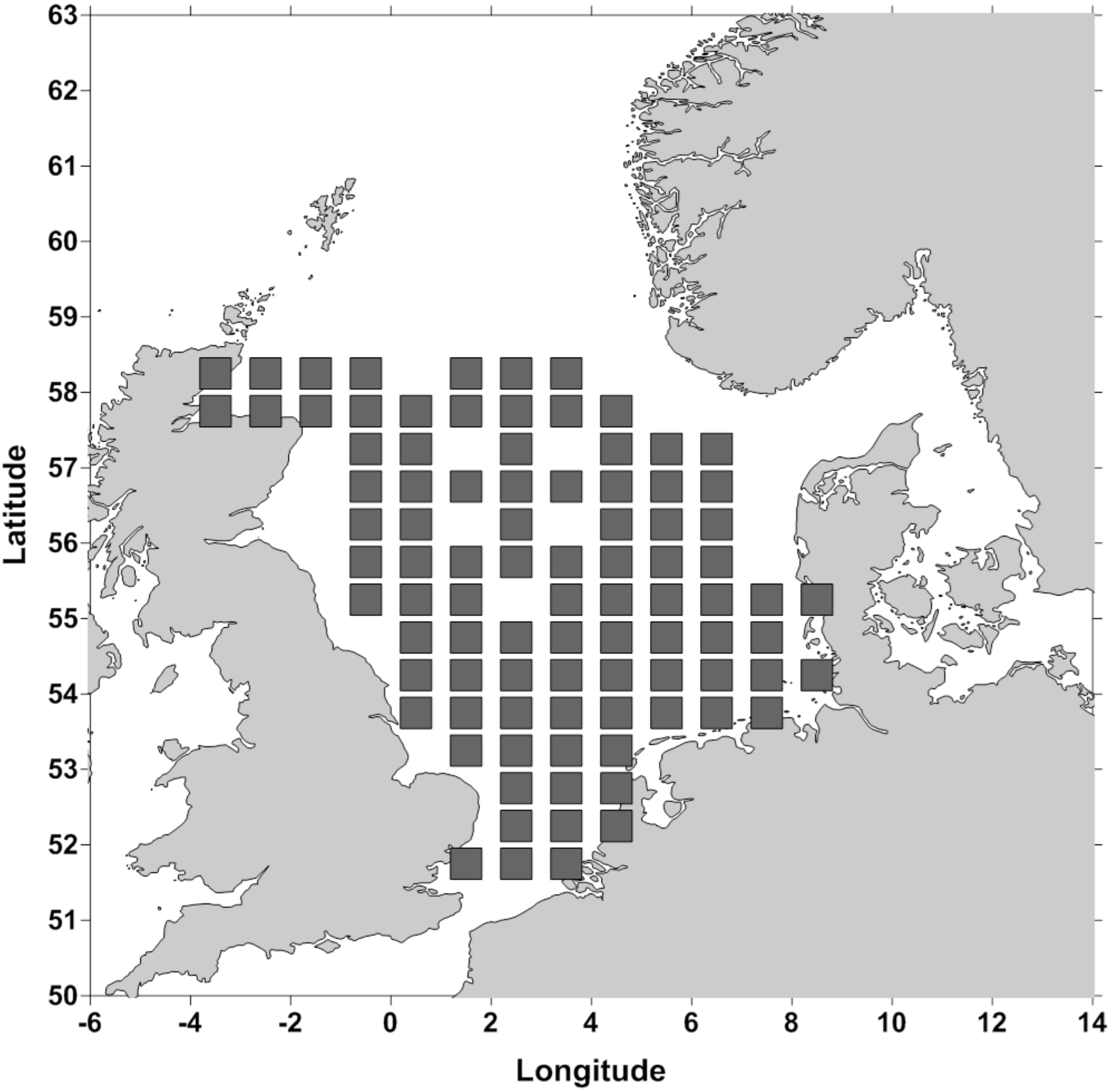
3.1.2. GNSNetBT3



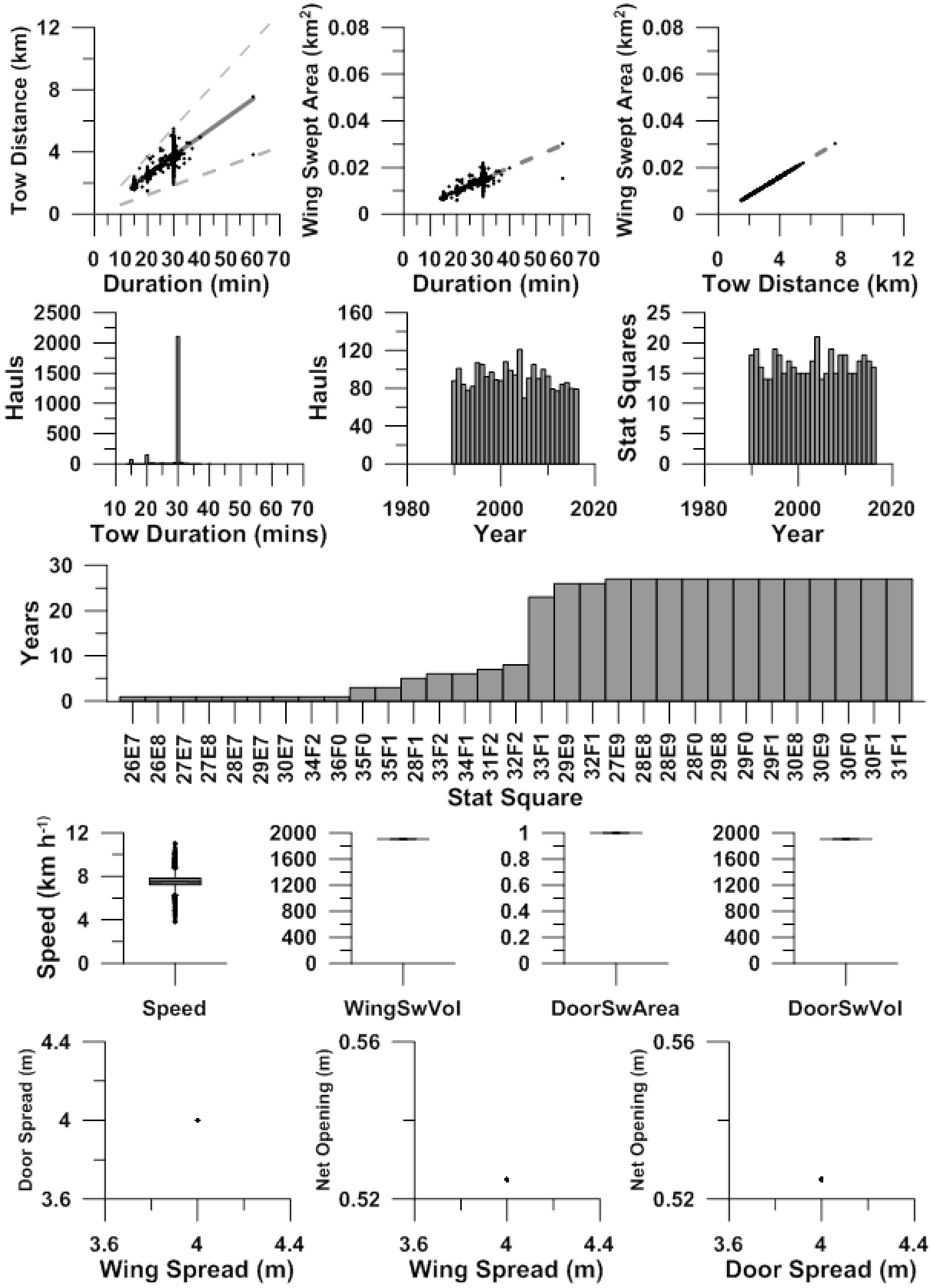


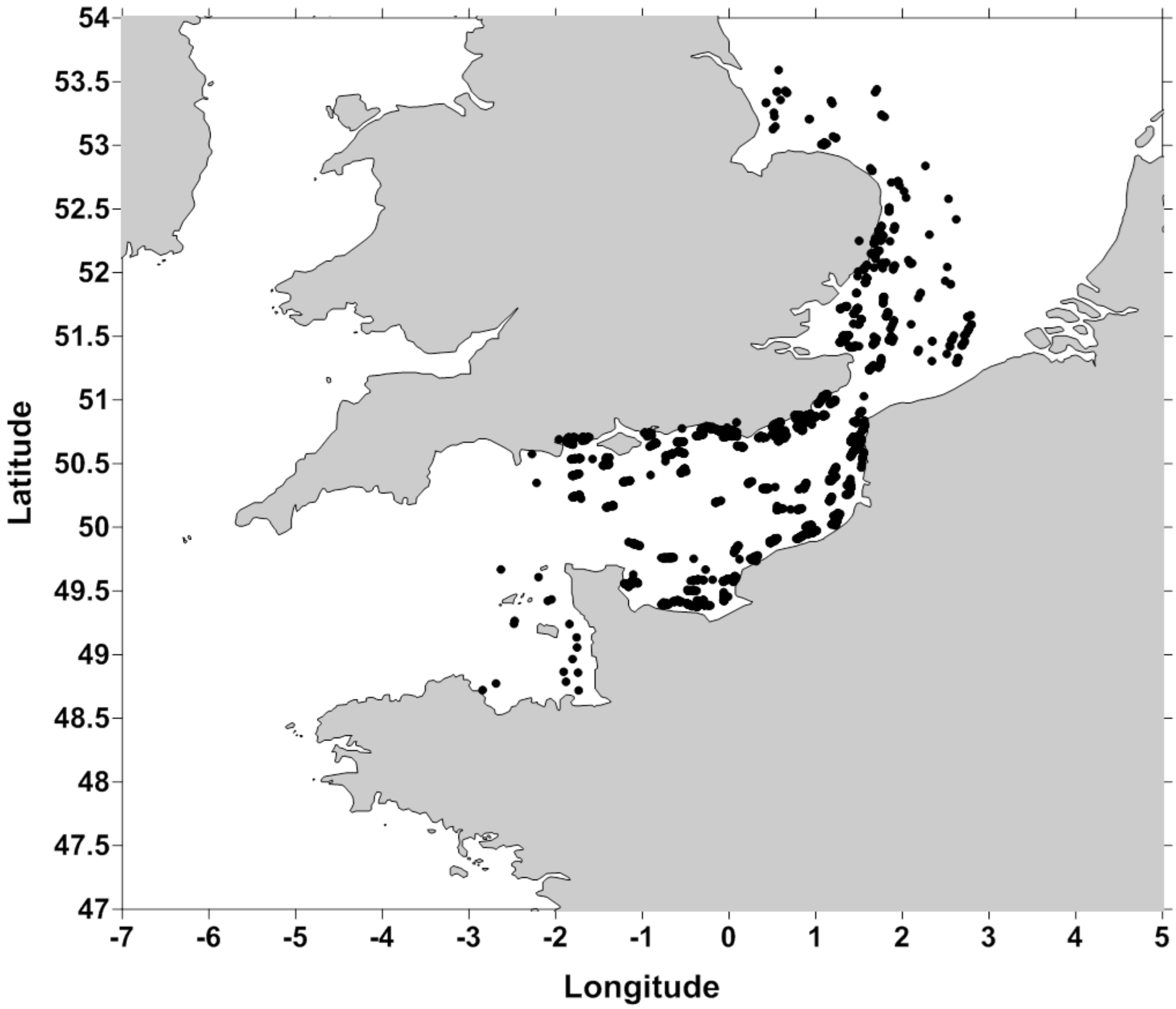
Moriarty et al. (2017) suggest that the GNSNetBT3 survey really became established in 2000, and with this start date, the 50% of years rule implied a *standard survey area* of 94 ICES statistical rectangles. However, overview of the Version 2 data set suggested that both the number of samples collected and the number of rectangles sampled in each year had actually stabilised by 1999. But with a start date of 1999, the 50% of years rule only gave a *standard survey area* of 91 rectangles. A start date of 1999, therefore, provided a 6.3% increase in temporal range at the cost of a 3.3% decrease in spatial range over a start date of 2000. Consequently, a start date of 1999 was adopted, which the data provider also agreed was the more appropriate. This gave a 17 year times series spanning 1999 to 2015 without any breaks. Application of the second criterion ensuring that rectangles were also sampled in the early and late 20%ile periods of the time series further reduced the Version 2 standard survey area to 89 statistical rectangles (Greenstreet and Moriarty 2017).

In this Version 3 data product, the time series has been increased to 18 years by the recent addition of 2016 data to the ICES DATRAS database. The 50% of years rule, therefore, infers that ICES statistical rectangles must be sampled in nine or more years to be included in the SSA; 29 rectangles (31F3, 33F1, 34F1, 35F0, 35F5, 38F8, 39F2, 41E8, 41F1, 41F3, 42E8, 43E8, 43F3, 43F7, 44F5, 45F0, 45F4, 45F5, 46E8, 46E9, 46F2, 47E9, 47F3, 48E9, 48F2, 49E9, 49F2, 50E9 and 50F2) failed this criterion and exclusion of the hauls sampled in these rectangles resulted in the deletion of 80 hauls. A further three rectangles (38E9, 40F7 and 43F1) failed the early 20%ile criterion by not being sampled within the four year period 1999-2002 at the start of the time series. Exclusion of the data associated with these rectangles resulted in the loss of a further 43 hauls. All remaining rectangles met the late 20%ile period criterion. From an initial total of 2665 hauls for the standard monitoring program, the GNSNetBT3 *standard survey area* data product, therefore, held data for 2542 hauls collected over 18 years from 91 ICES statistical rectangles, an increase of two rectangles, 45F2 and 45F3 located at the eastern end of the northern most row of rectangles sampled, over the Version 2 data product used in the OSPAR IA2017 (Greenstreet and Moriarty 2017).

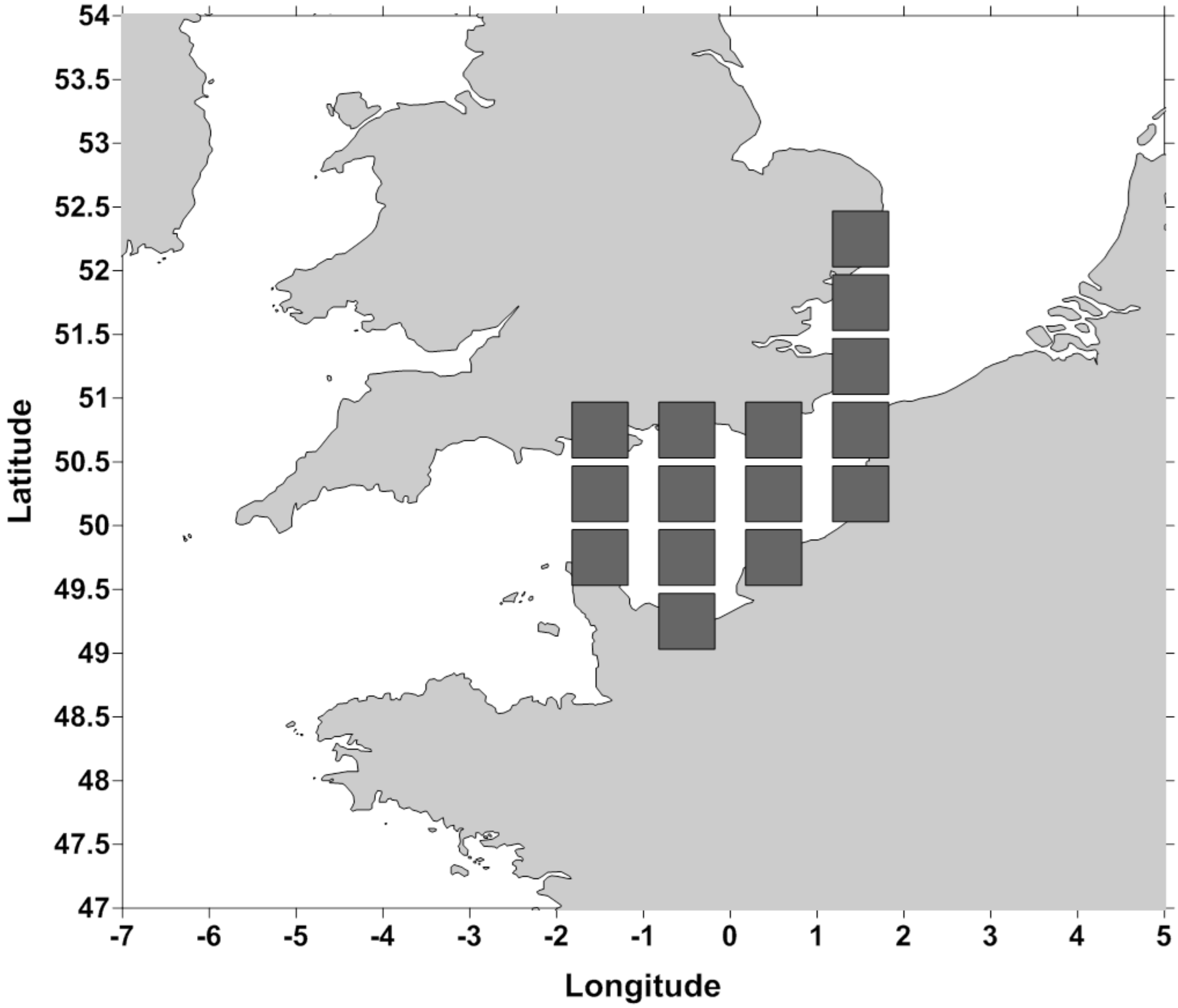

63
62
61
60
59
58
57
56
55
54
53
52
51
50
Latitude
-6
-4
-2
0
2
4
6
8
10
12
14
Longitude

3.1.3. GNSEngBT3

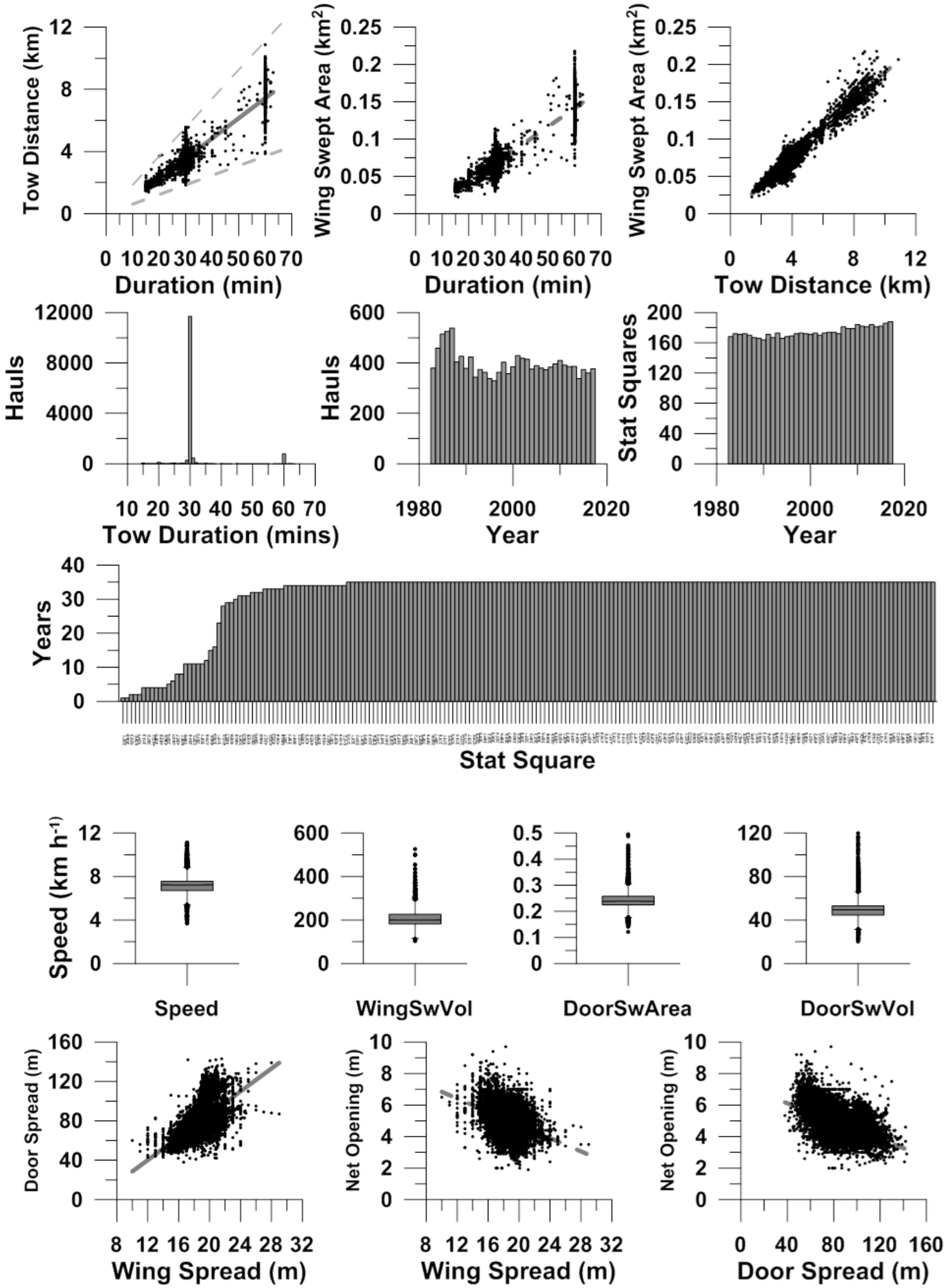


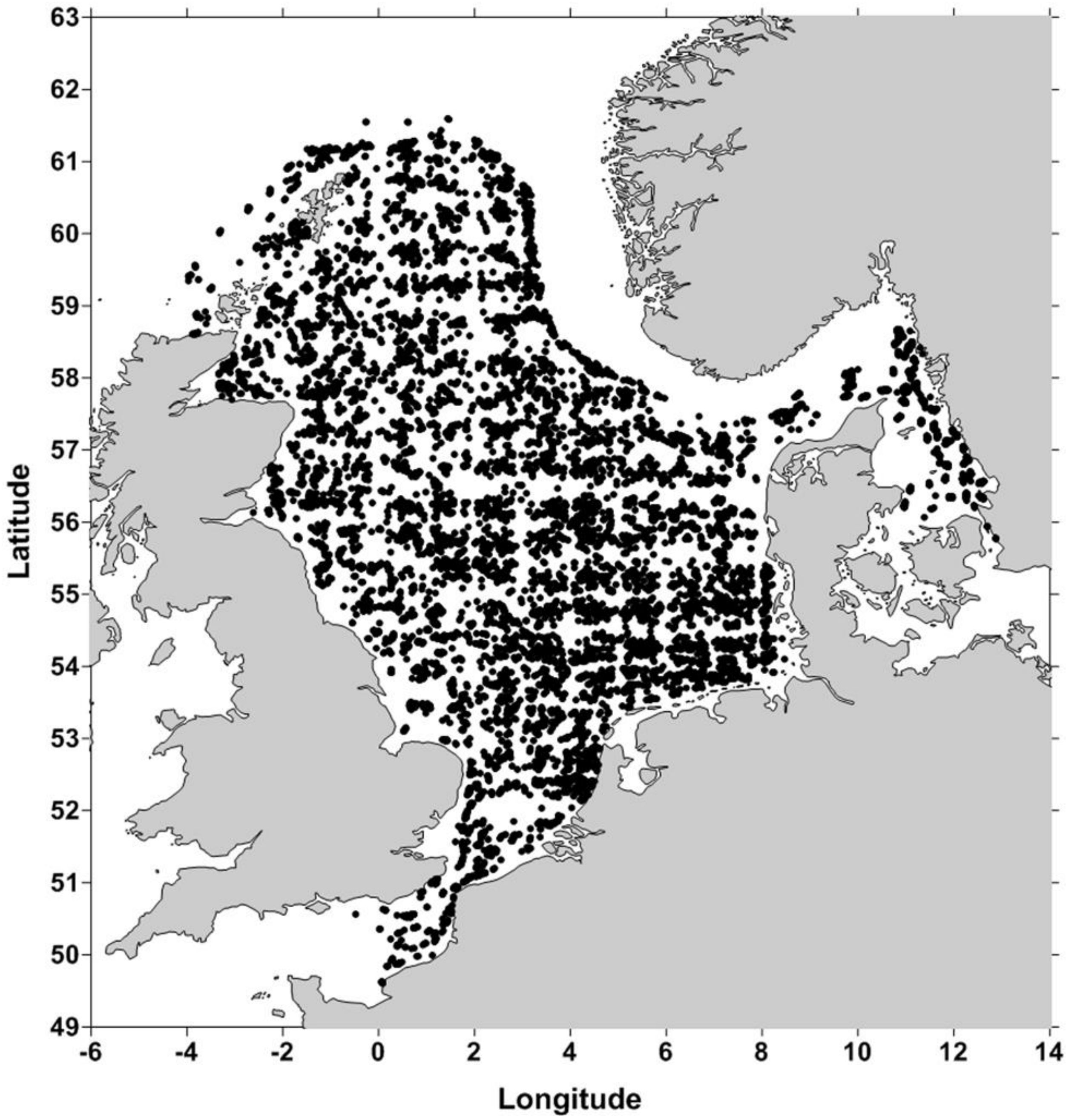


Overview of the GNSEngBT3 data revealed no inconsistency with the conclusions presented in Moriarty et al. (2017). With the recent addition of 2016 data to the DATRAS database, the survey now spans the period 1990 to 2016 with no breaks, giving 27 years of data; an increase of one year (2016) on the Version 2 data product used for the OSPAR IA2017 (Greenstreet and Moriarty, 2017), which in turn was an increase of one year (2015) over the Version 1 data derived by Moriarty et al. (2017). The 50% of years rule therefore required ICES statistical rectangles to be sampled in 14 years or more to be included in the *standard survey area*. Of the 31 rectangles sampled over the course of the entire survey (an increase of two rectangles over the Version 2 data product), 16 (26E7, 26E8, 27E7, 27E8, 28E7, 28F1, 29E7, 30E7, 31F2, 32F2, 33F2, 34F1, 34F2, 35F0, 35F1 and 36F0) failed this criterion causing data for 130 hauls to be excluded from the GNSEngBT3 *standard survey area* data product. The remaining 15 rectangles all met the second criterion of being sampled at least once in the six year periods (1990-1995 and 2011-2016) that constituted the early and late 20%iles of the time series. There was actually no difference between the Version 2 and Version 3 GNSEngBT3 *standard survey areas*, both covered the same 15 ICES statistical rectangles; the two additional rectangles sampled in 2016 both failed the 50% of years rule. The original *standard monitoring program* data set of 2467 samples was reduced to 2337 records collected in the 27 year period 1990 to 2016 from a *standard survey area* of 15 ICES statistical rectangles.

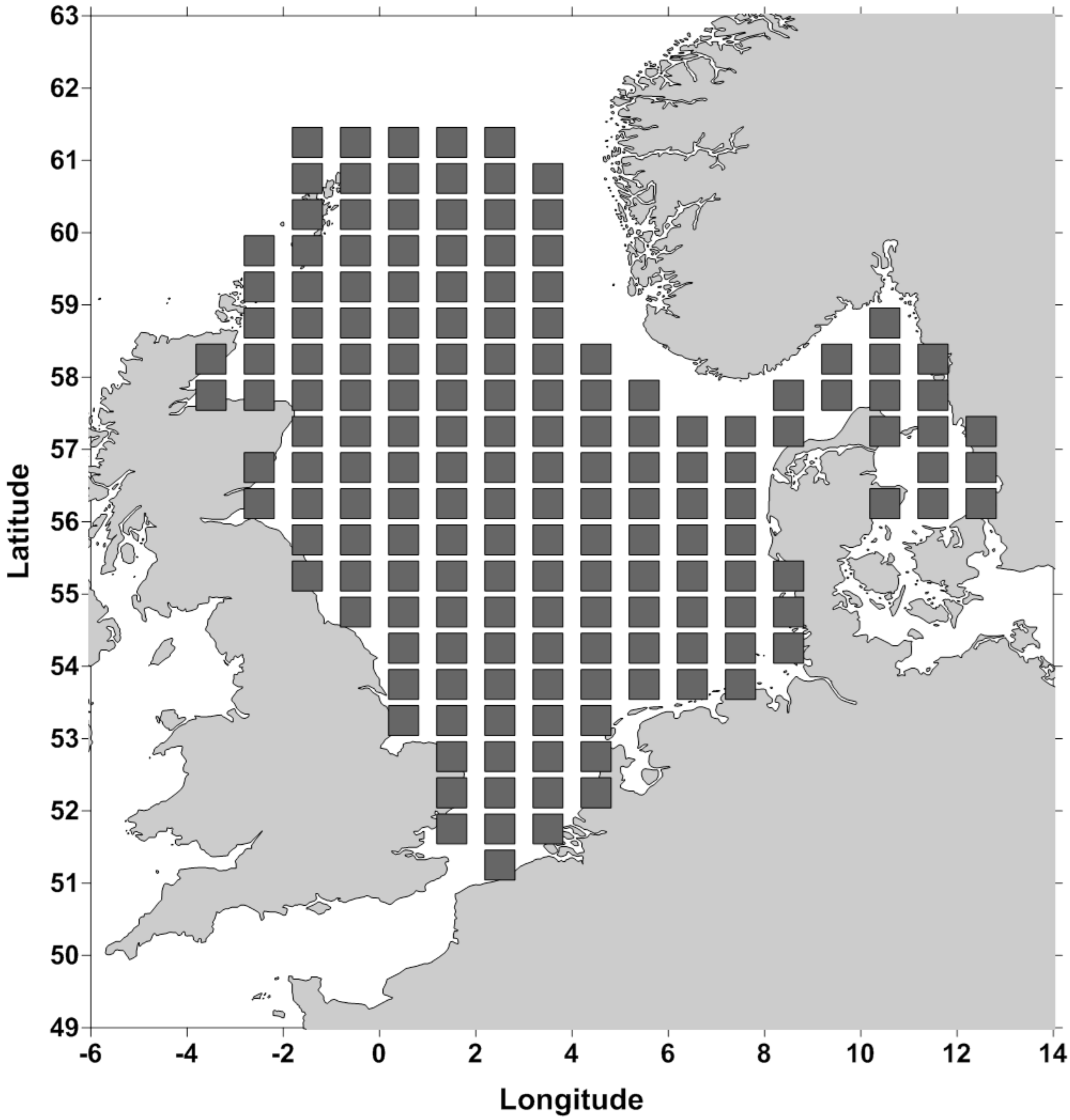


3.1.4. GNSIntOT1

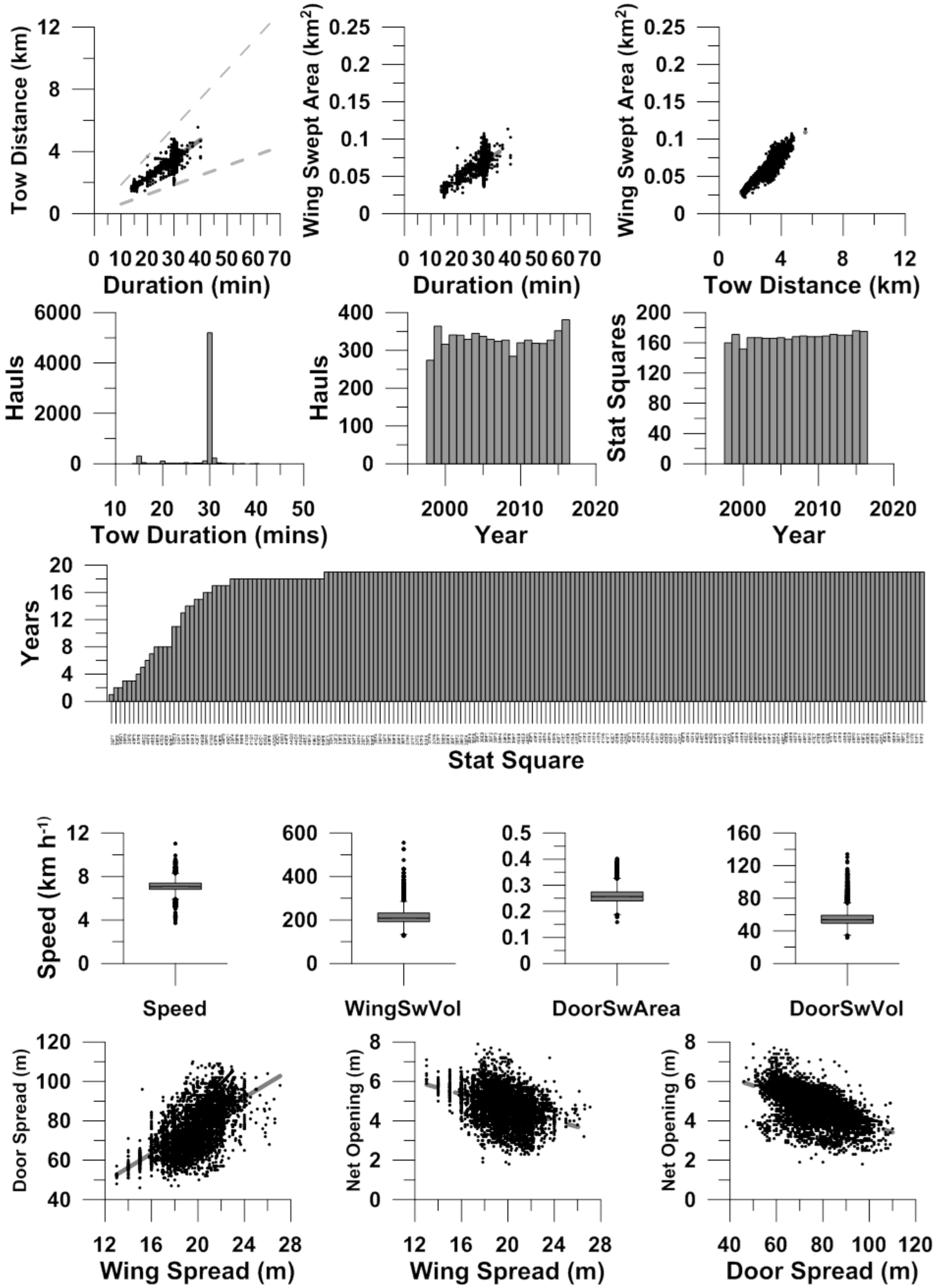


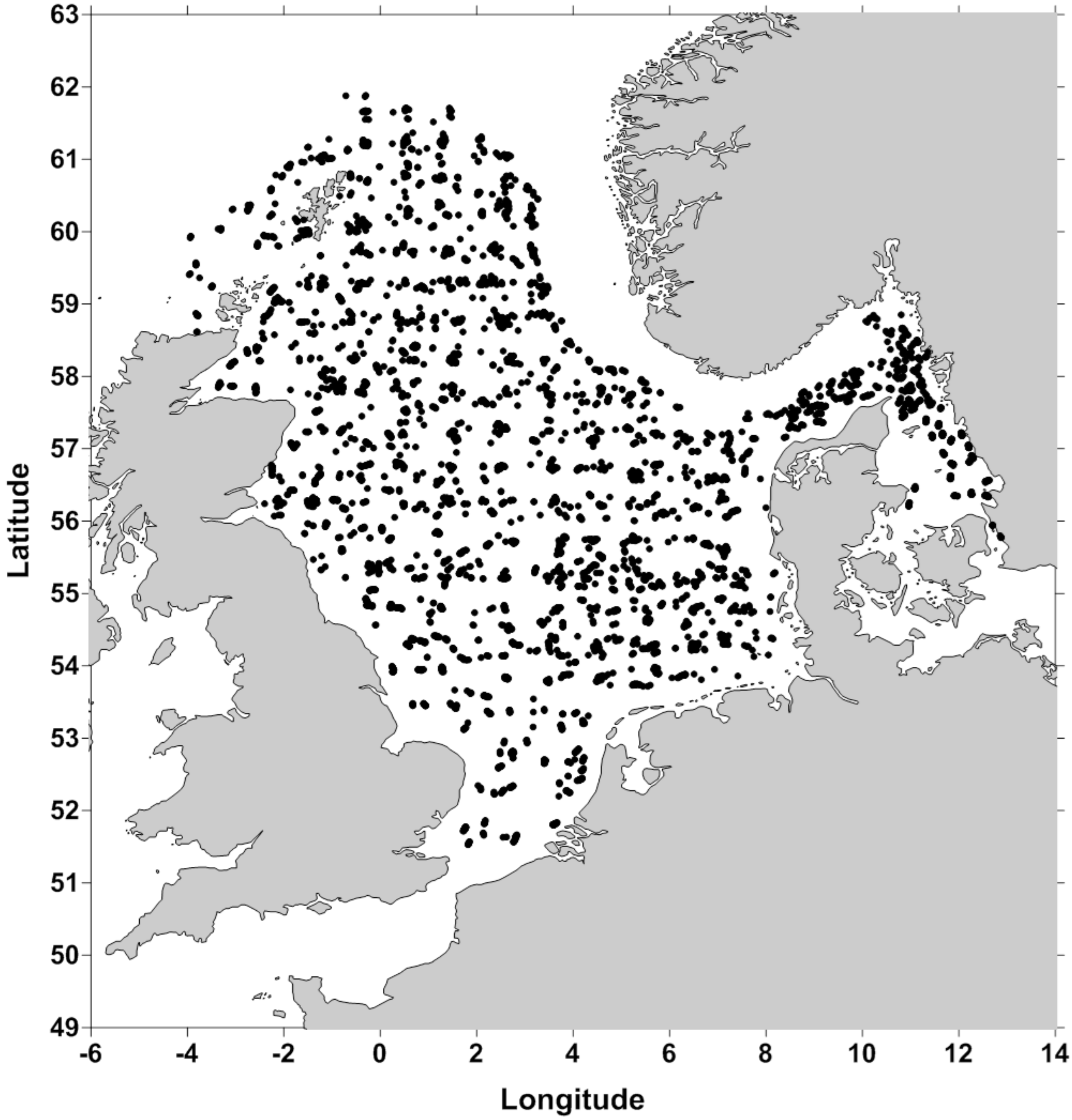


The data set available to Moriarty et al. (2017) for derivation of the data product covered the period 1983 to 2015. The Version 2 data product used in the OSPAR IA2017 was extended by one year, amounting to a 34 year time series spanning 1983 to 2016 with no breaks. For Version 3, the time series has again been extended by the addition of 2017 data, therefore, giving a 35 year times series. The 50% of years rule, therefore, required ICES statistical rectangles to be sampled in 18 years or more to be included in the *standard survey area*. Of the 195 rectangles sampled over the course of the entire survey, 172 met this criterion. Data for the 23 rectangles (28F0, 28F1, 29F0, 29F1, 30E9, 30F0, 30F1, 31F3, 35F5, 36F8, 37E9, 40G2, 43F9, 46E6, 46G1, 47E6, 48E6, 49E6, 49E7, 50E7, 52E9, 52F0 and 52F1) failing the 50% of years rule were excluded, amounting to 280 hauls. These 23 rectangles were sampled infrequently throughout the time series. One further rectangle (31F1) also failed the second criterion of having to be sampled in the seven year period (1983-1989) that constituted the early 20%ile of the time series. Data for 31F1, amounting to a further 57 samples, were also excluded from the data product reducing the total number of GNSIntOT1 trawl samples in the Version 3 *standard survey area* data product to 13555 from the 13516 samples in the full *standard monitoring program* data set. The Version 3 *standard survey area* data product, therefore, spanned a 35 year time series, from 1983 to 2017, and were obtained from an *standard survey area* of 171 ICES statistical rectangles.



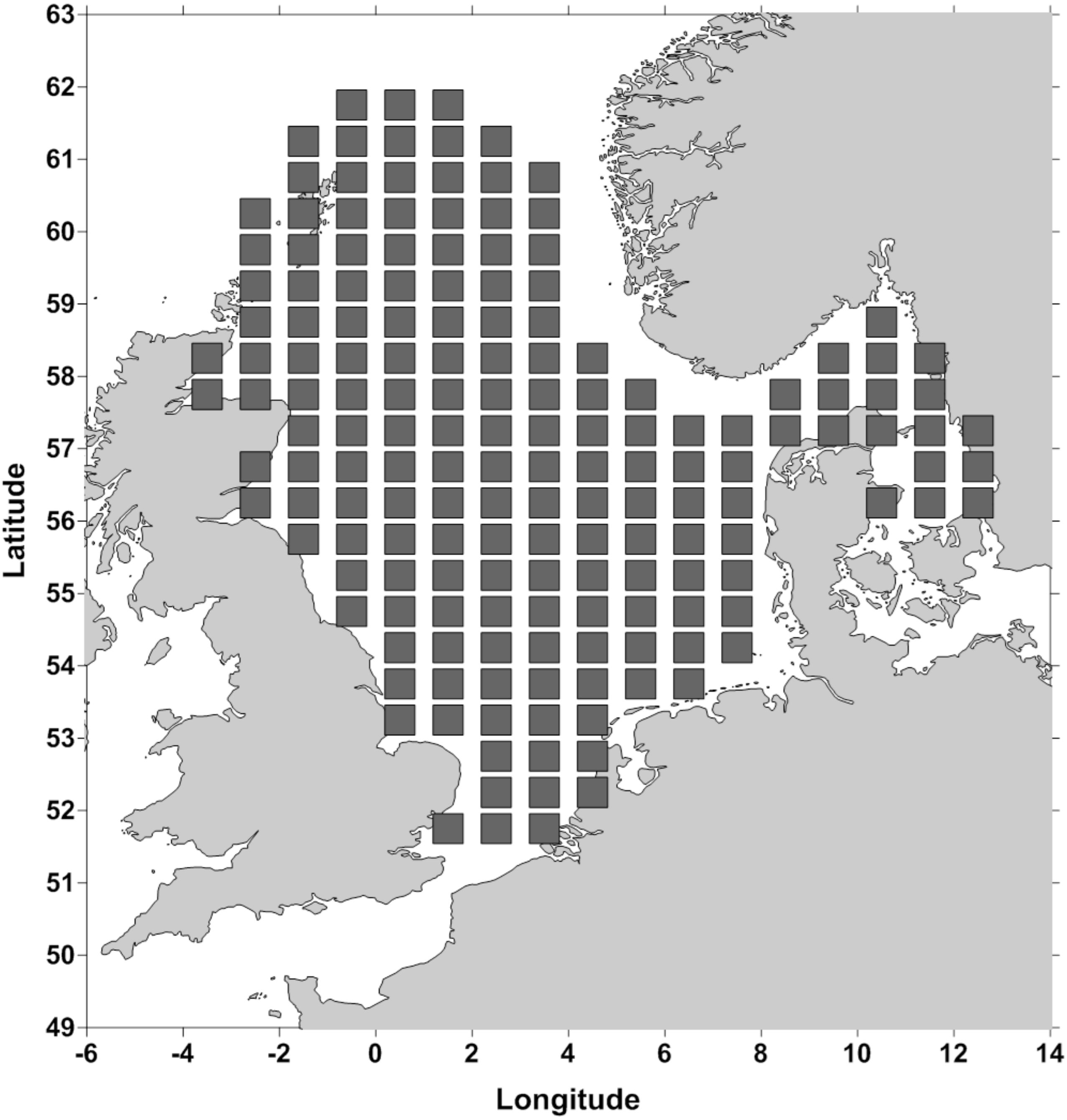
3.1.5. GNSIntOT3



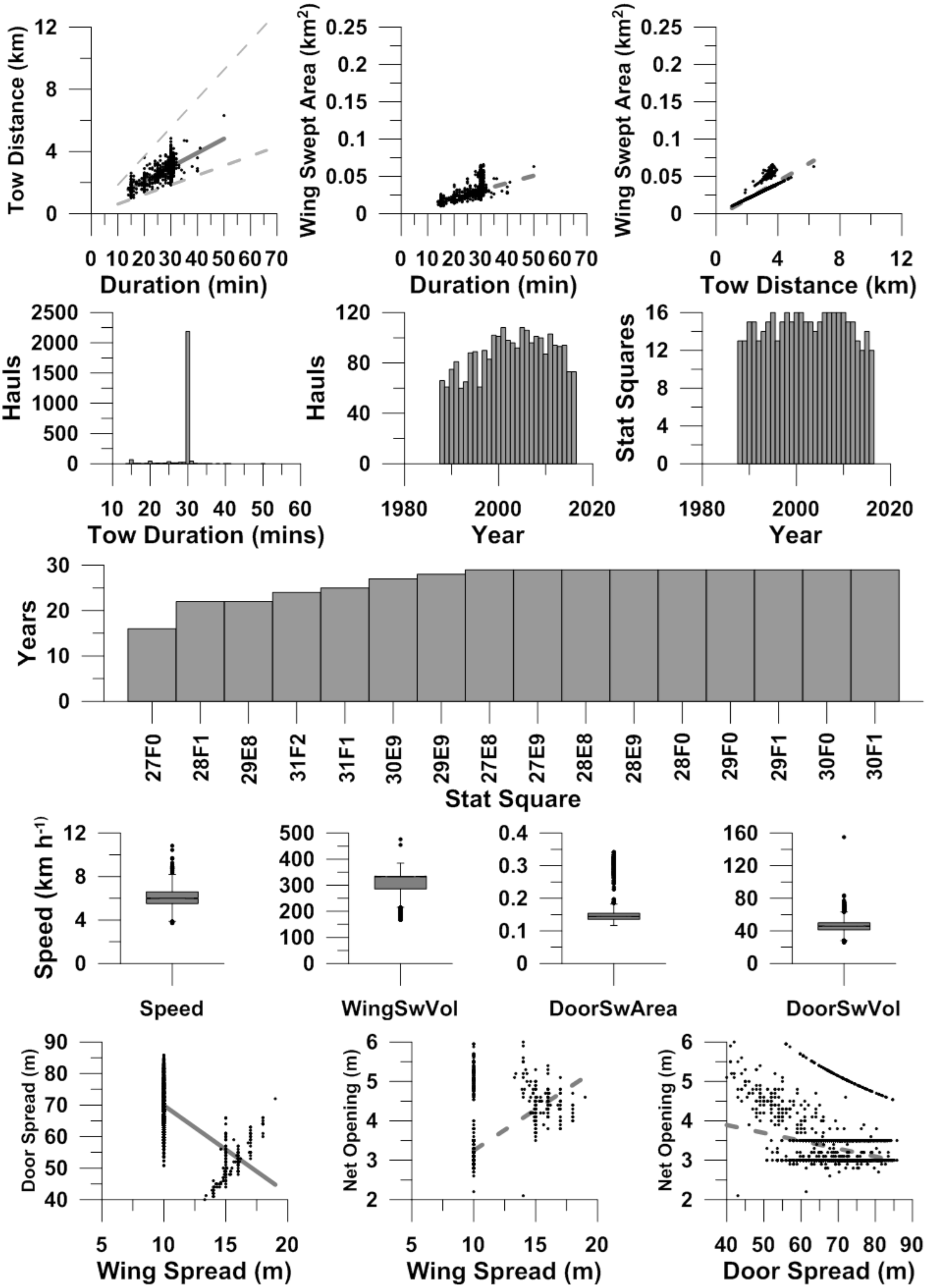


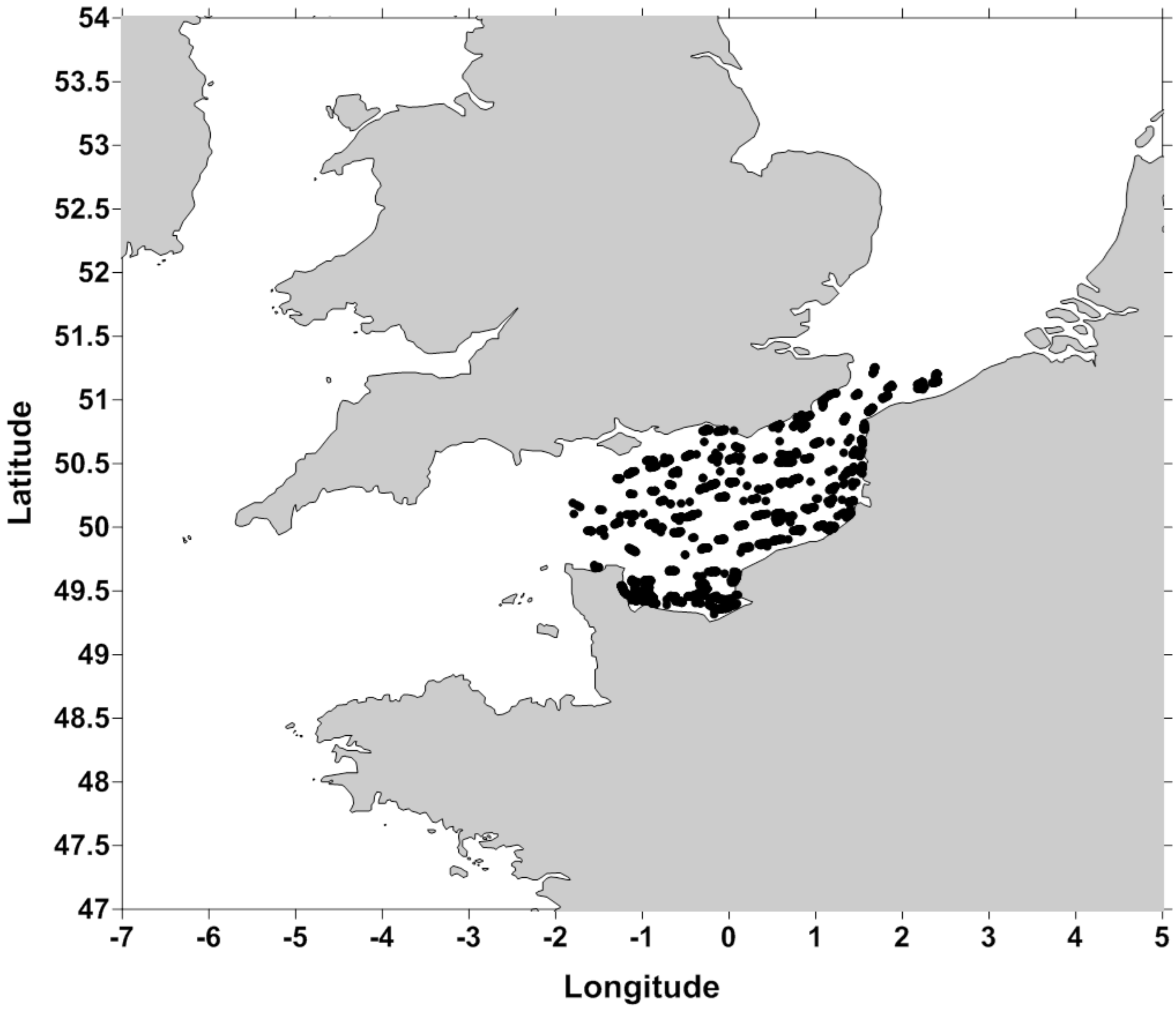
The data set available to Moriarty et al. (2017) covered the period 1998 to 2014. The Version 2 data product used for the OSPAR IA2017 was extended by one year, amounting to a 18 year time series spanning 1998 to 2015 with no breaks (Greenstreet and Moriarty, 2017). Data on the DATRAS portal have subsequently been updated by the addition of data collected in 2016, now giving a 19 year time series, 1998 to 2016, with no breaks. The 50% of years rule, therefore, required ICES statistical rectangles to be sampled in ten years or more to be included in the SSA. Of the 182 rectangles sampled over the course of the entire survey, 168 met this criterion. Data for the 14 rectangles (36F7, 37F8, 38F8, 39E8, 39F8, 40G2, 44F6, 45F5, 46E6, 46G1, 47E6, 48E6, 49E6 and 50E7) failing the 50% of years rule were excluded, amounting to 84 hauls. These 14 rectangles were sampled infrequently throughout the time series. All 168 rectangles that met the 50% of years rule also met the criteria of having been sampled in the four year periods 1998-2001 and 2013-2016 that constituted the early and late 20%iles of the time series. Thus, of the original 6254 samples in the *standard monitoring program* data set, 6170 were retained in the *standard survey area* data set collected over the 19 year period 1998 to 2016 from a *standard survey area* of 168 ICES statistical rectangles.

Comparison of the data diagnostic plot shown here with the one published previously by Greenstreet et al. (2017) describing the Version 2 data product reveal less marked outliers in the door spead v wing spread and door spread v net opening plots. This reflects corrections made to the door spread data held on the DATRAS portal by the national Data Provider.


63
62
61
60
59
58
57
56
55
54
53
52
51
50
49
Latitude
-6
-4
-2
0
2
4
6
8
10
12
14
Longitude

3.1.6. GNSFraOT4

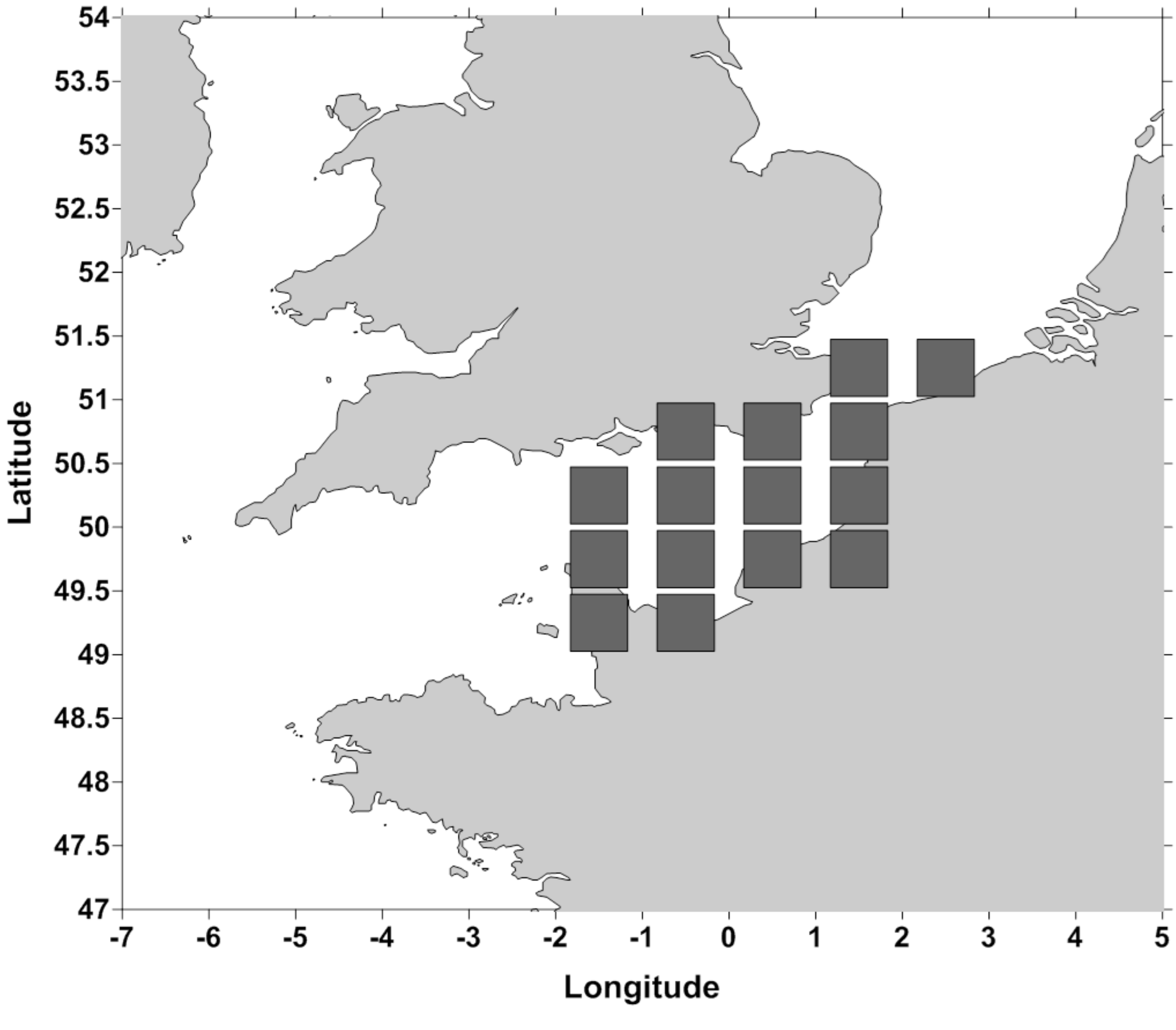




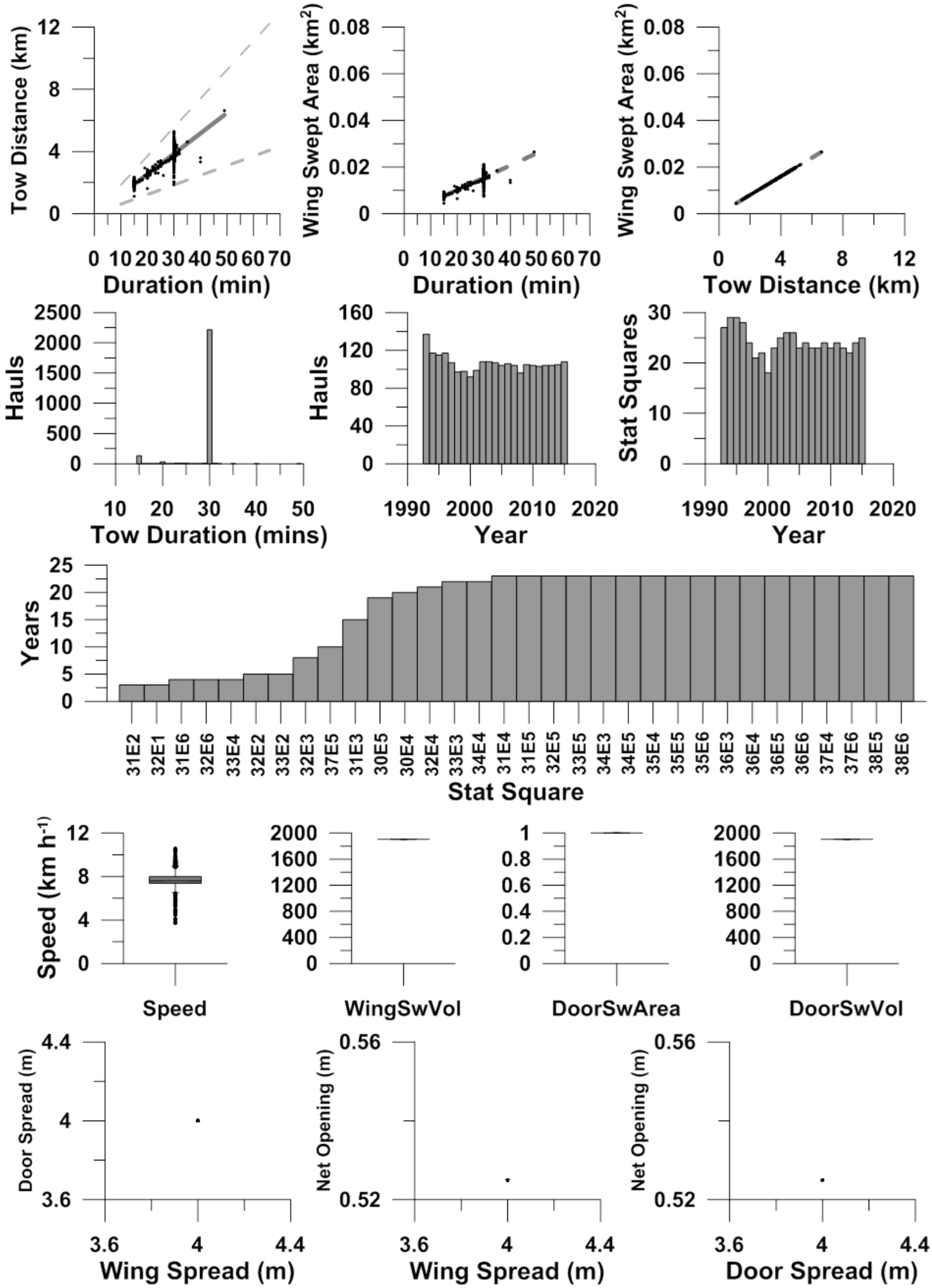
The GNSFraOT4 data set available to Moriarty et al. (2017) covered the period 1988 to 2014. The Version 2 data product used in the OSPAR IA2017 included data added for 2015 (Greenstreet and Moriarty, 2017). Data for a further year, 2016, has subsequently been added to the ICES DATRAS database and is, therefore ,now included in the Version 3 data product. Consequently, the time series now covers a 29 year period, 1988 to 2016, with no breaks. The 50% of years rule, therefore, required ICES statistical rectangles to be sampled in 15 years or more to be included in the *standard survey area*. In fact all 16 rectangles ever sampled by the survey met this criterion. However one rectangle, 27F0, was first sampled in 1995, and so failed the criterion for being sampled in the first six years, 1988 to 1993, that represented the early 20%ile of the time series. Data for this rectangle, amounting to 31 samples, were, therefore, excluded from the data product leaving 2513 GNSFraOT4 records in the *standard survey area* time series out of the original 2544 records in the *standard monitoring program* time series. The *standard survey area* time series, therefore, spans a 29 year period between 1988 and 2016 consisting of

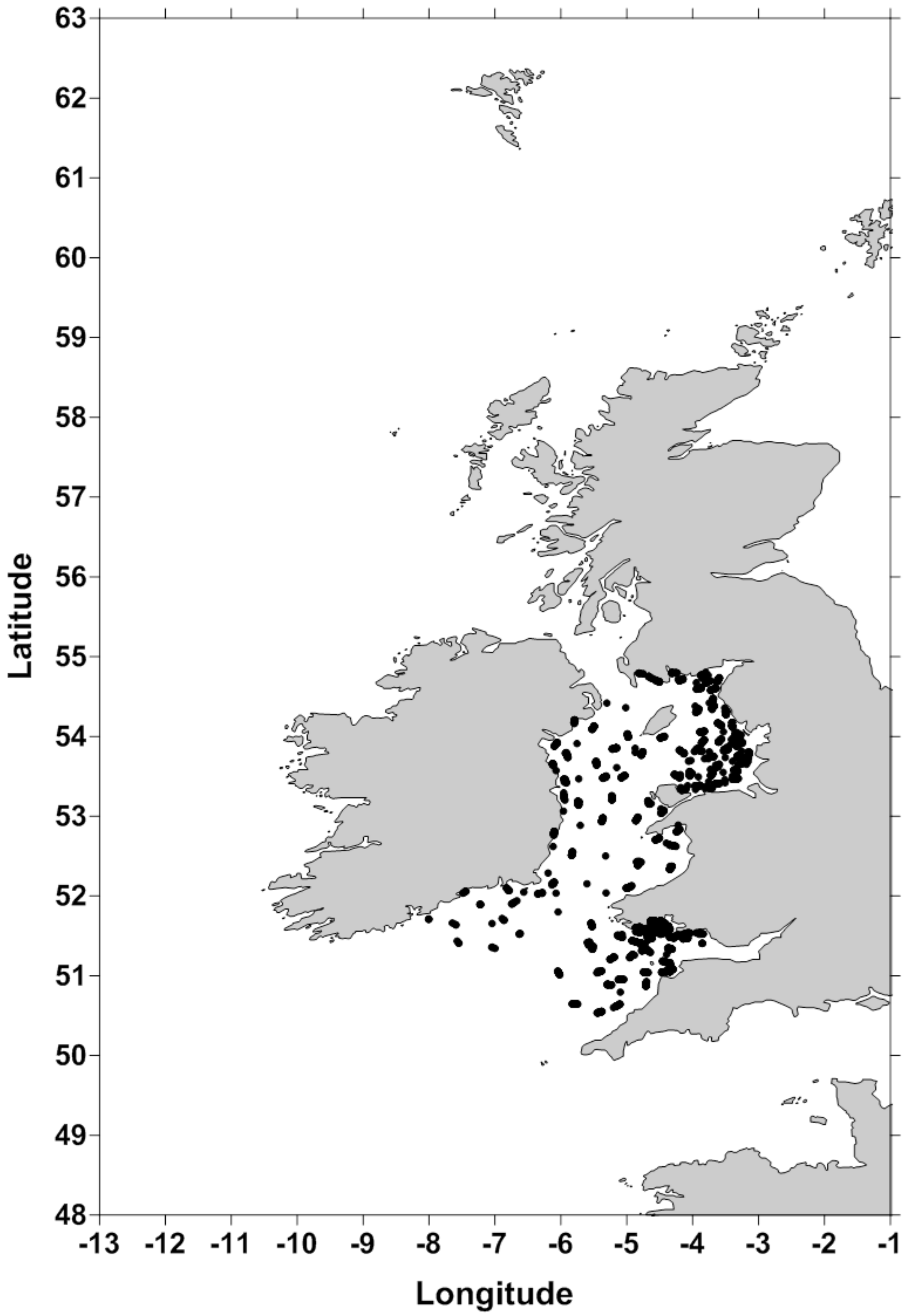
2513 trawl sample records data collected from a *standard survey area* of 15 ICES statistical rectangles.

Comparison of the data diagnostic plot shown here with the one published previously by Greenstreet et al. (2017) describing the Version 2 data product reveal marked differences in some of the plots. This is due to revisions made in the wing spread and door spread estimation models, and the use of a fixed wing spread value of 10 m for samples collected by the RV "*Gwendolyn*".

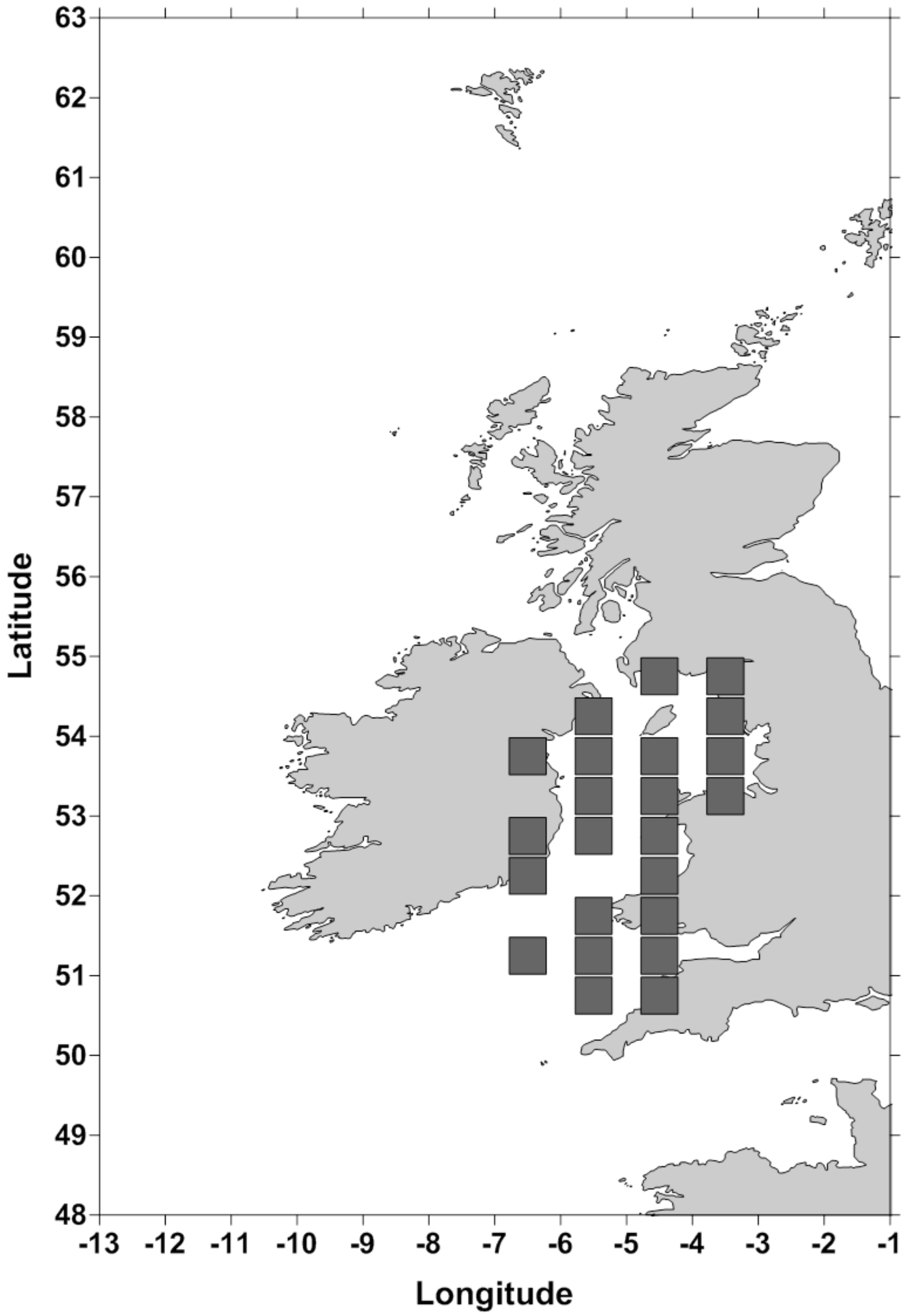


3.1.7. CSEngBT3

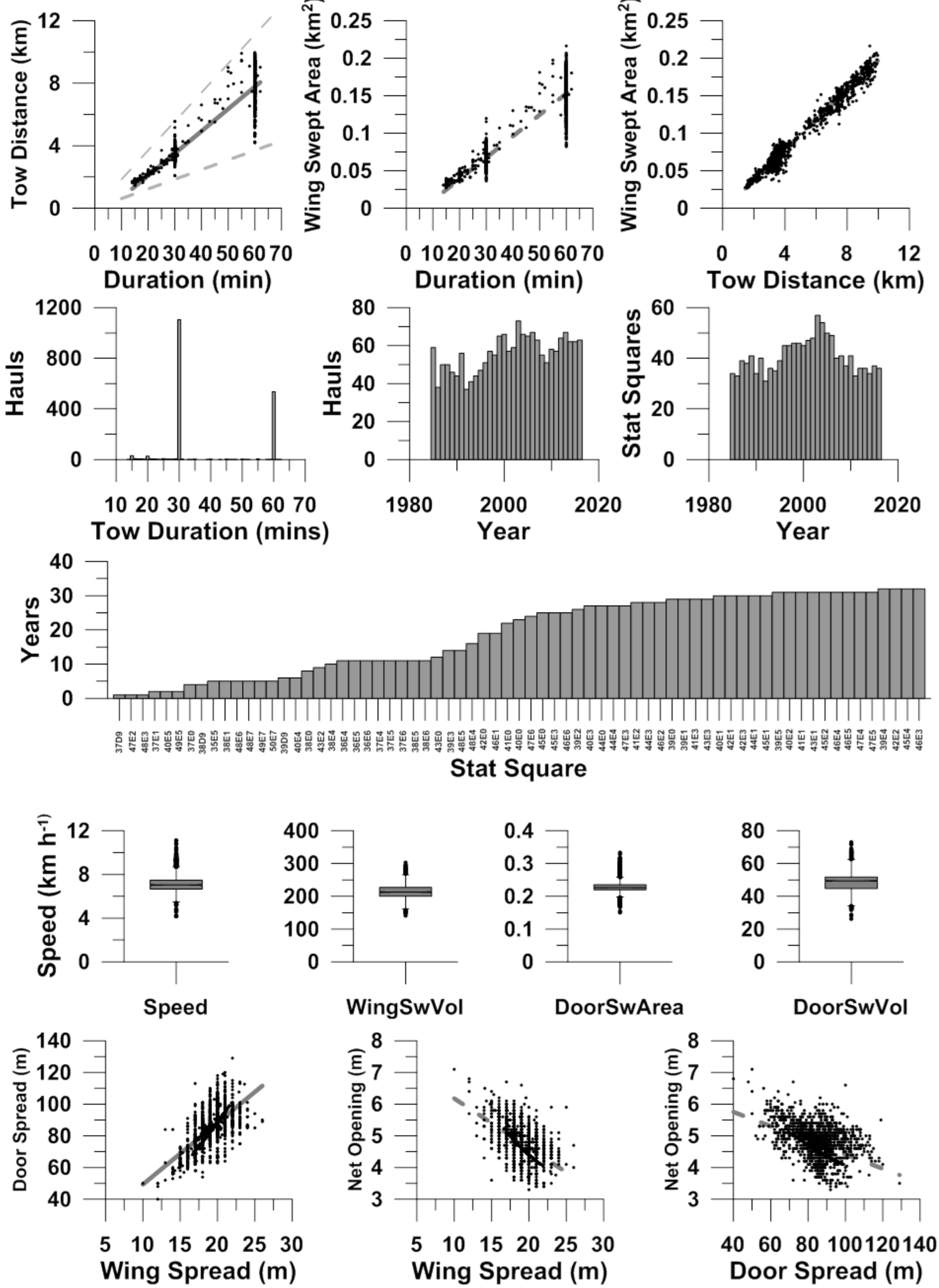


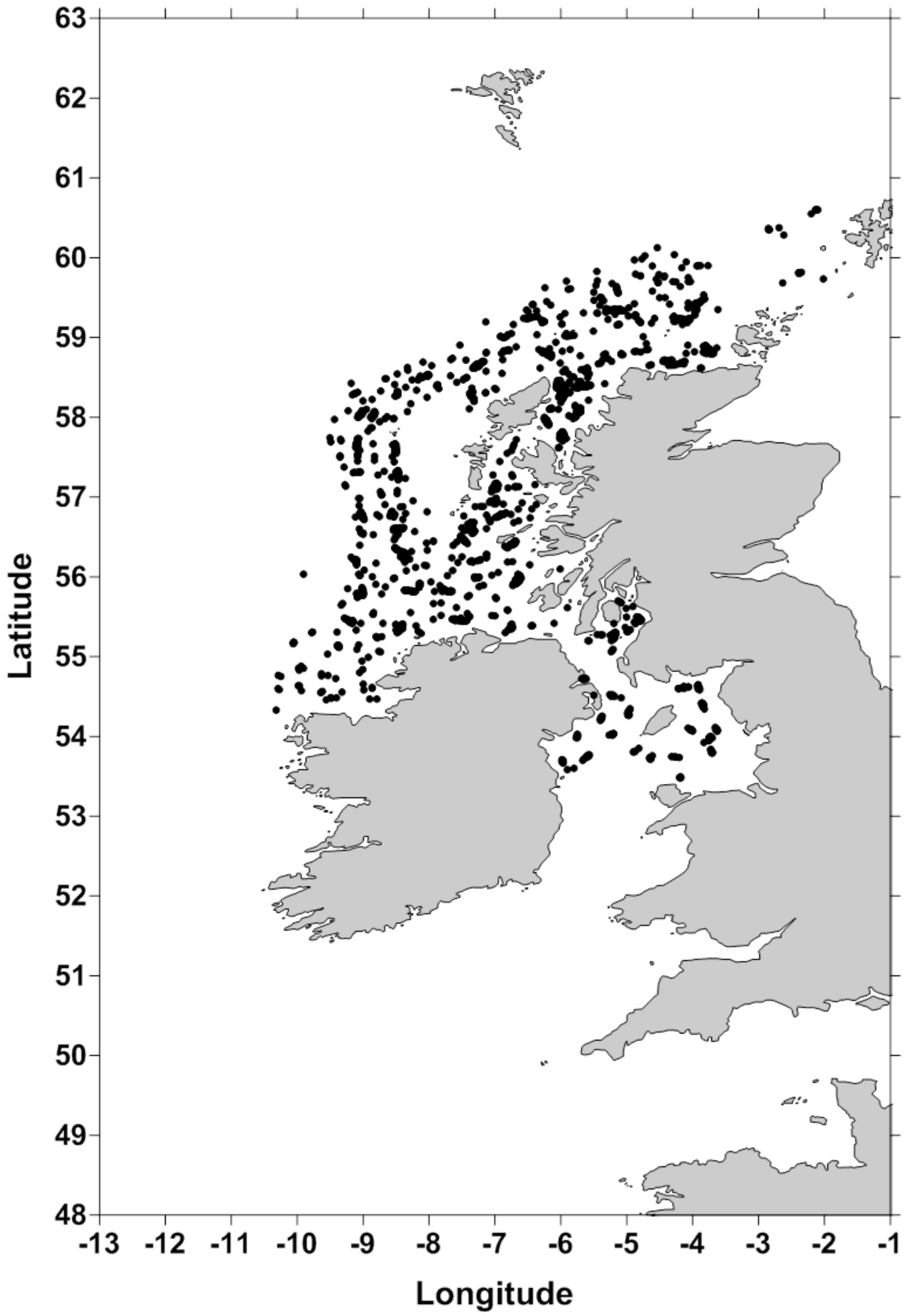


When Moriarty et al. (2017) first derived the Version 1 data product for the CSEngBT3 survey, the data available for download from the DATRAS portal for this survey spanned a 23 year period between 1993 and 2015 with data collected from a total of 32 ICES statistical rectangles. Data for this survey on the DATRAS database have not been updated since. The Version 2 data set was, therefore, unchanged over Version 1 (Greenstreet and Moriarty, 2017) and this Version 3 data set is unchanged over Version 2. Of the 32 rectangles sampled by the survey, only 23 met the 50% of years rule, being sampled in 12 years or more. All 23 rectangles met the second criterion having been sampled in the five year periods, 1993-1997 and 2011-2015, which constituted the early and late 20%iles of the time series. Data for the nine ICES statistical rectangles (31E2, 31E6, 32E1, 32E2, 32E3, 32E6, 33E2, 33E4 and 37E5) that failed the 50% of years rule were excluded, amounting to 67 samples collected by the full *standard monitoring program*. Excluding these samples reduced the original *standard monitoring program* of 2445 samples to 2378 trawl samples collected from a *standard survey area* of 23 ICES statistical rectangles over a 23 year period 1993 to 2015.

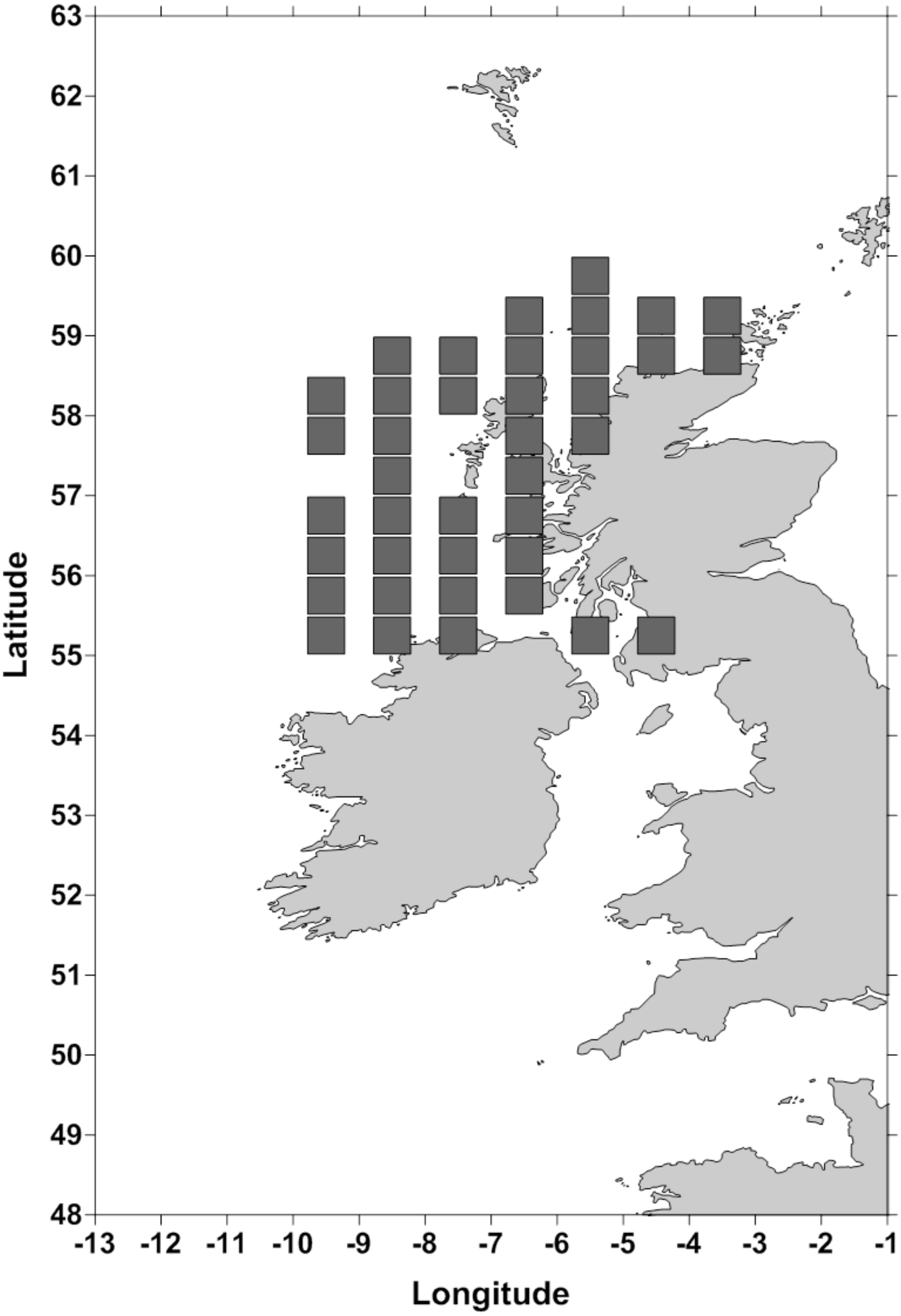


3.1.8. CSScoOT1



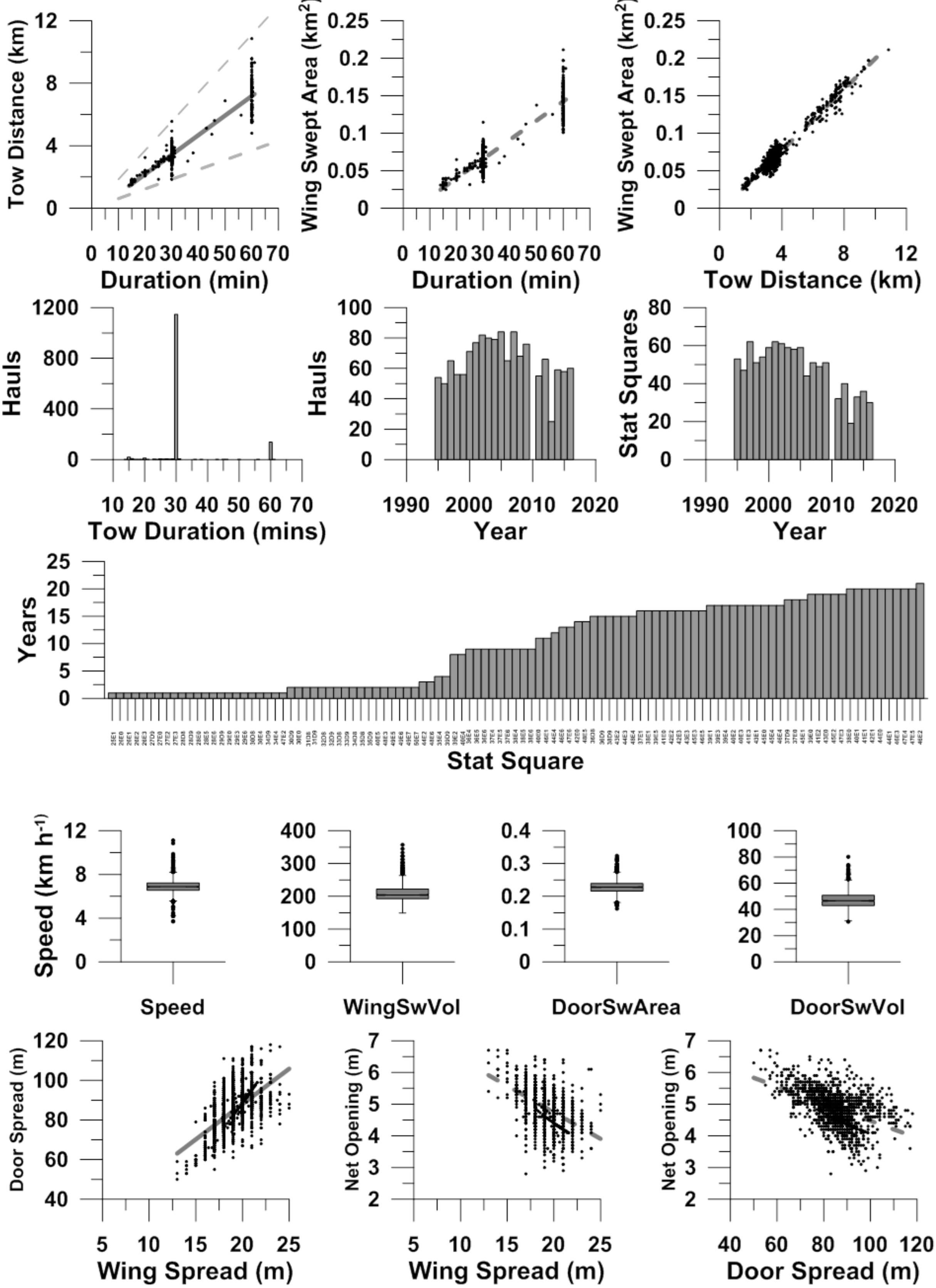


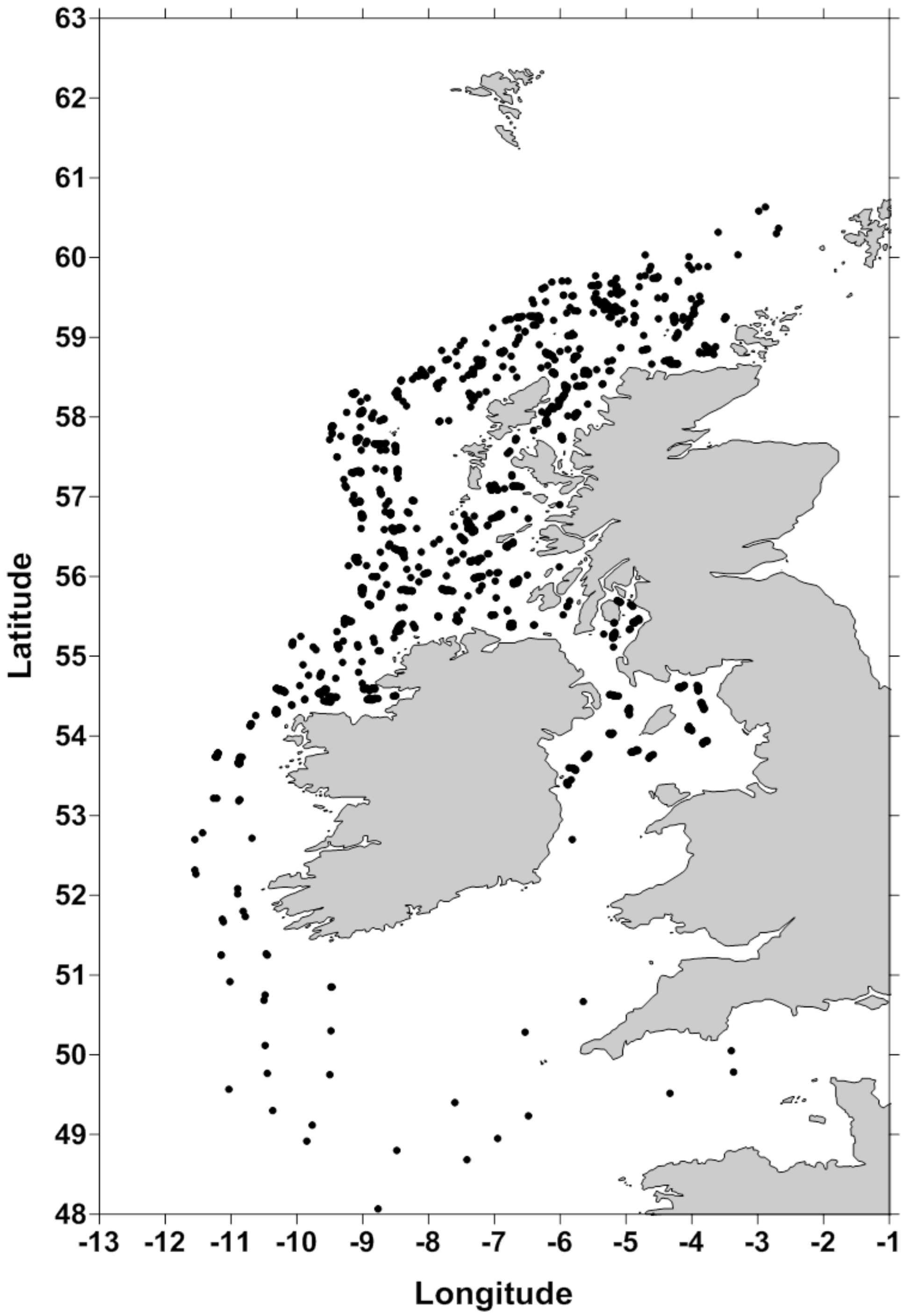
The data available on DATRAS for derivation of the data product spanned the period 1985 to 2015 (Moriarty et al., 2017). These data showed this survey to have been highly variable during its 31 year history, with a pronounced change from 60 minutes to 30 minutes duration hauls occurring in the late 1990s. Major changes in spatial coverage have also occurred but this is largely controlled through application of the 50% of years rule. Overview of the data product suggested that one option might be to restrict the time series to just the period 1999 onwards, but examination of the data showed that this had minimal impact in enhancing consistency of spatial coverage, yet the cost in temporal range was considerable. In determining the *standard survey area*, the full *standard monitoring program* time series was assessed. Version 2 of the data product, used to support the OSPAR IA20127, included data for 2016 added to the DATRAS database subsequently to Moriarty's et al. (2017) initial analysis (Greenstreet and Moriarty, 2017). To date no further years of data have been added so this Version 3 data product still covers the same 32 year period 1985 to 2016 as the Version 2 data product. Data for the 30 ICES statistical rectangles (35E5, 36E4, 36E5, 36E6, 37D9, 37E0, 37E1, 37E4, 37E5, 37E6, 38D9, 38E0, 38E1, 38E4, 38E5, 38E6, 39D9, 39E3, 40E4, 40E5, 43E0, 43E2, 47E2, 48E3, 48E5, 48E6, 48E7, 49E5, 49E7 and 50E7) that failed the 50% of years rule (needing to have been sampled at least once in at least 16 years) were excluded, amounting to 303 samples collected throughout the time series. The 39 rectangles that met the 50% of years rule also met the criterion for having been sampled in the start and finish 20%ile of time series seven year periods (1985-1991 and 2010 to 2016). However, two rectangles 46E6 and 47E6 were last sampled in 2010 and given the change in survey design that occurred in 2011, it is quite possible that these rectangles might not be sampled again and at the next assessment they would, therefore, drop out of the *standard survey area*. Exclusion of the 303 samples collected in the 30 rectangles that were insufficiently sampled reduced the original full CSScoOT1 *standard monitoring program* data set from 1795 to 1492 samples spanning 32 years, from 1985 to 2016, collected across a *standard survey area* of 39 ICES statistical rectangles.



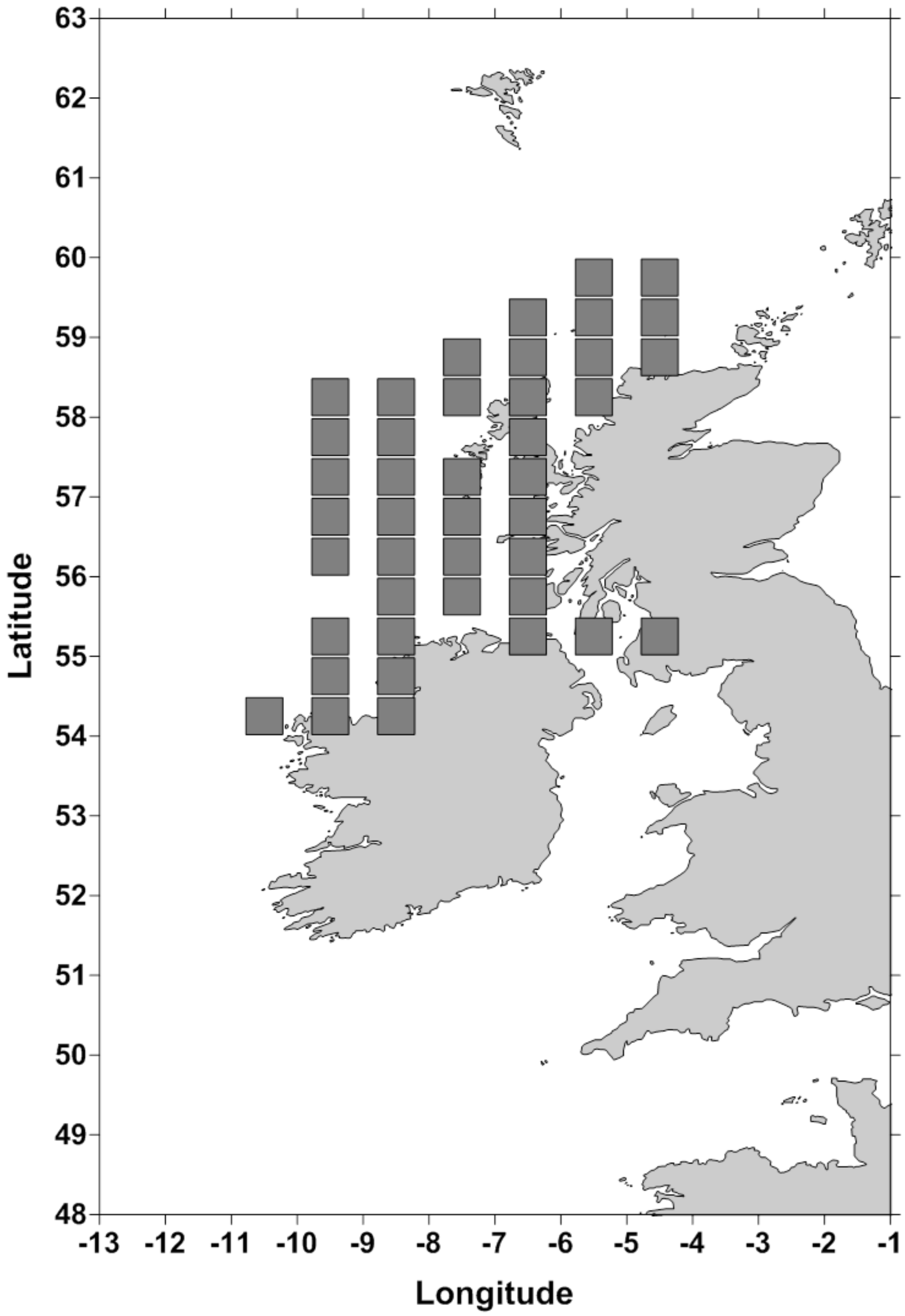
63
62
61
60
59
58
57
56
55
54
53
52
51
50
49
48
Latitude
-13
-12
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
Longitude

3.1.9. CSScoOT4

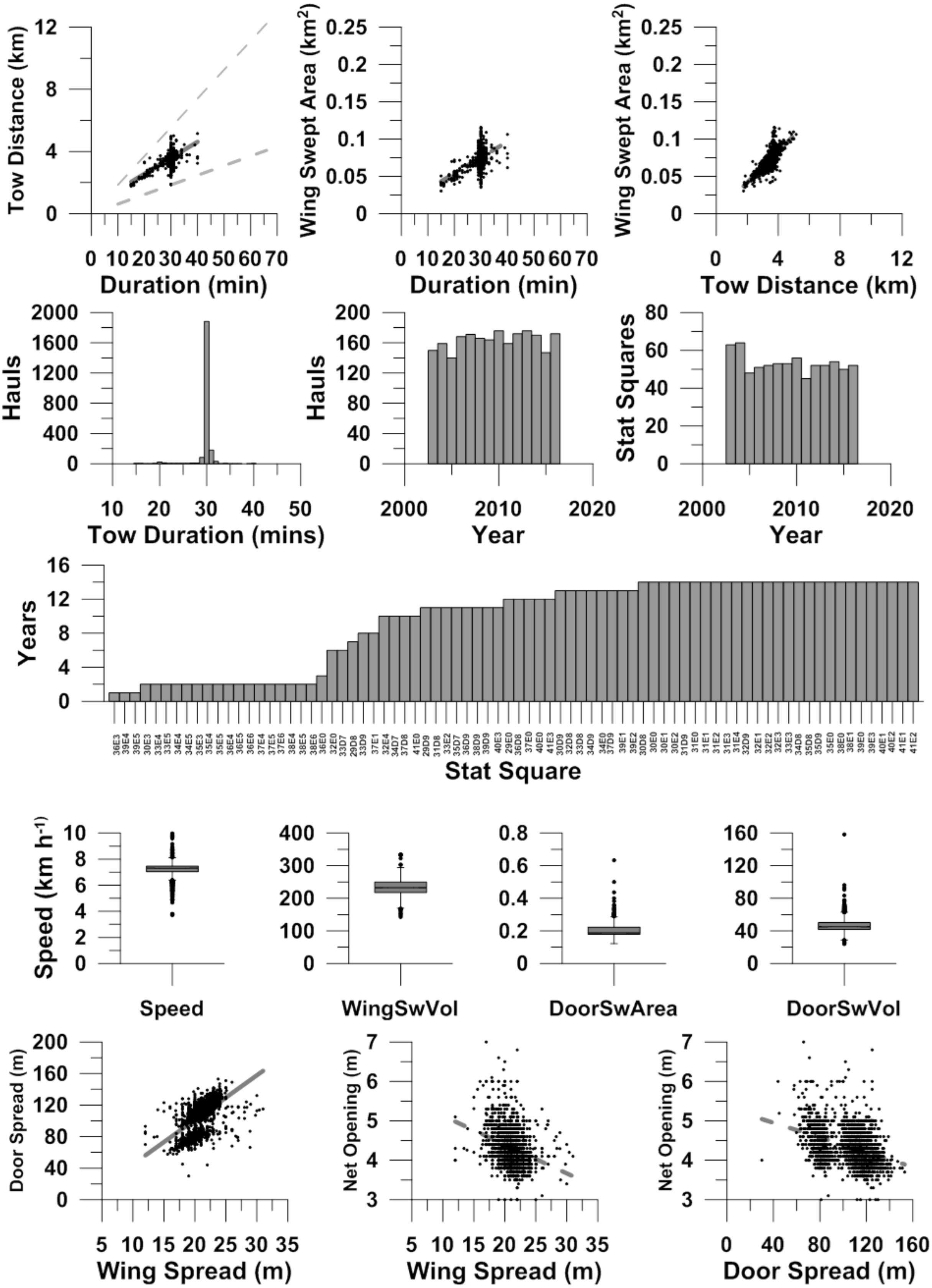


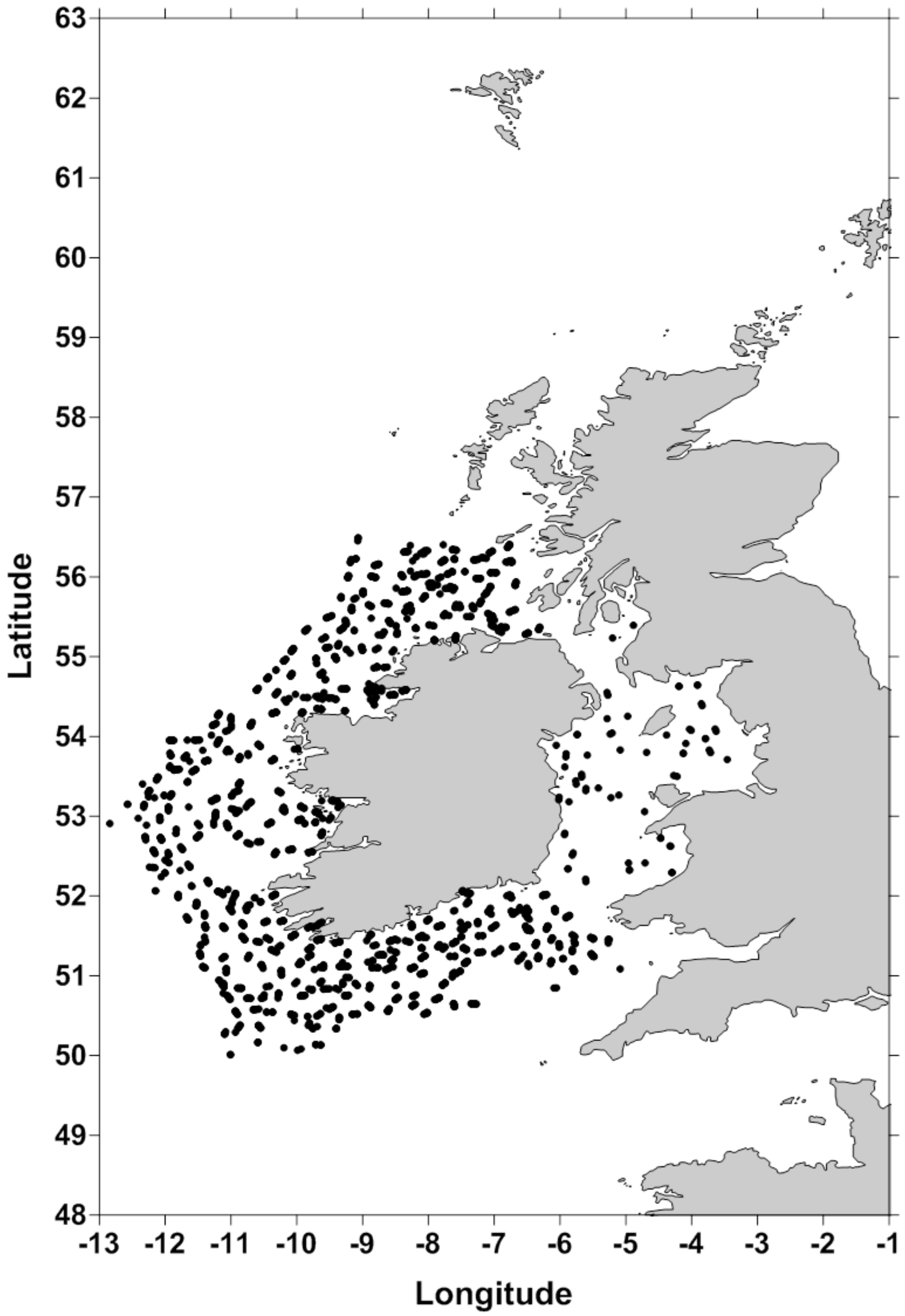


This survey has also undergone considerable evolution, particularly in the early years, which resulted in considerable variation in spatial coverage. Consequently, although data from 1990 onwards were available for download from the DATRAS portal, Moriarty et al. (2017) concluded that the most appropriate start date for the survey data product was 1995; this date providing the optimal compromise between survey time duration and spatial coverage. Given the history of the development of this survey, the national Data Provider also concurred that this was a more appropriate start date. In reviewing this survey to produce the Version 2 data product used in the OSPAR IA2017, Greenstreet and Moriarty (2017) saw no reason to break from these previous decisions. At the time when Moriarty et al. (2017) derived the Version 1 data product, the survey standard monitoring programme, therefore, spanned the period 1995 to 2014, but by the time the Version 2 data product was derived, data for 2015 had been uploaded to the DATRAS database (Greenstreet and Moriarty, 2017). More recently data for 2016 have also been uploaded for to the DATRAS database so that this Version 3 CSScoOT4 survey data product now includes 21 years of data spanning the period 1995 to 2016 but, as with both previous version, with a break in 2010 when no survey was undertaken. Of 105 ICES statistical rectangles ever sampled in the full *standard monitoring program*, 55 failed the 50% of years rule (25E1, 26E0, 26E1, 26E2, 26E3, 27D9, 27E0, 27E2, 27E3, 28D8, 28D9, 28E0, 28E5, 28E6, 29D9, 29E0, 29E3, 29E6, 30D8, 30D9, 30E0, 30E4, 31D8, 31D9, 32D8, 32D9, 33D8, 33D9, 34D8, 34D9, 34E4, 35D8, 35D9, 35E4, 36E4, 36E5, 36E6, 37E4, 37E5, 37E6, 38E4, 38E5, 38E6, 39D9, 39E2, 40E4, 40E5, 44E2, 47E2, 48E3, 48E6, 49E5, 49E6, 49E7 and 50E7). Excluding samples collected in these rectangles resulted in the loss of 190 samples. One rectangle (46E1) was not sampled in the early five year time series 20%ile period between 1995 and 1999 and seven rectangles (36D8, 36D9, 38D9, 40E0, 44E4, 46E6 and 47E6) were not sampled in the late five year time series 20%ile period of 2012 to 2016, resulting in the further exclusion of 11 and 98 samples respectively. Thus from the *standard monitoring program* data set of 1370 samples, 1071 were retained collected from a *standard survey area* of 42 ICES statistical rectangles over the period 1995 to 2016 with a break in 2010, so 21 years of data.


63
62
61
60
59
58
57
56
55
54
53
52
51
50
49
48
Latitude
-13
-12
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
Longitude

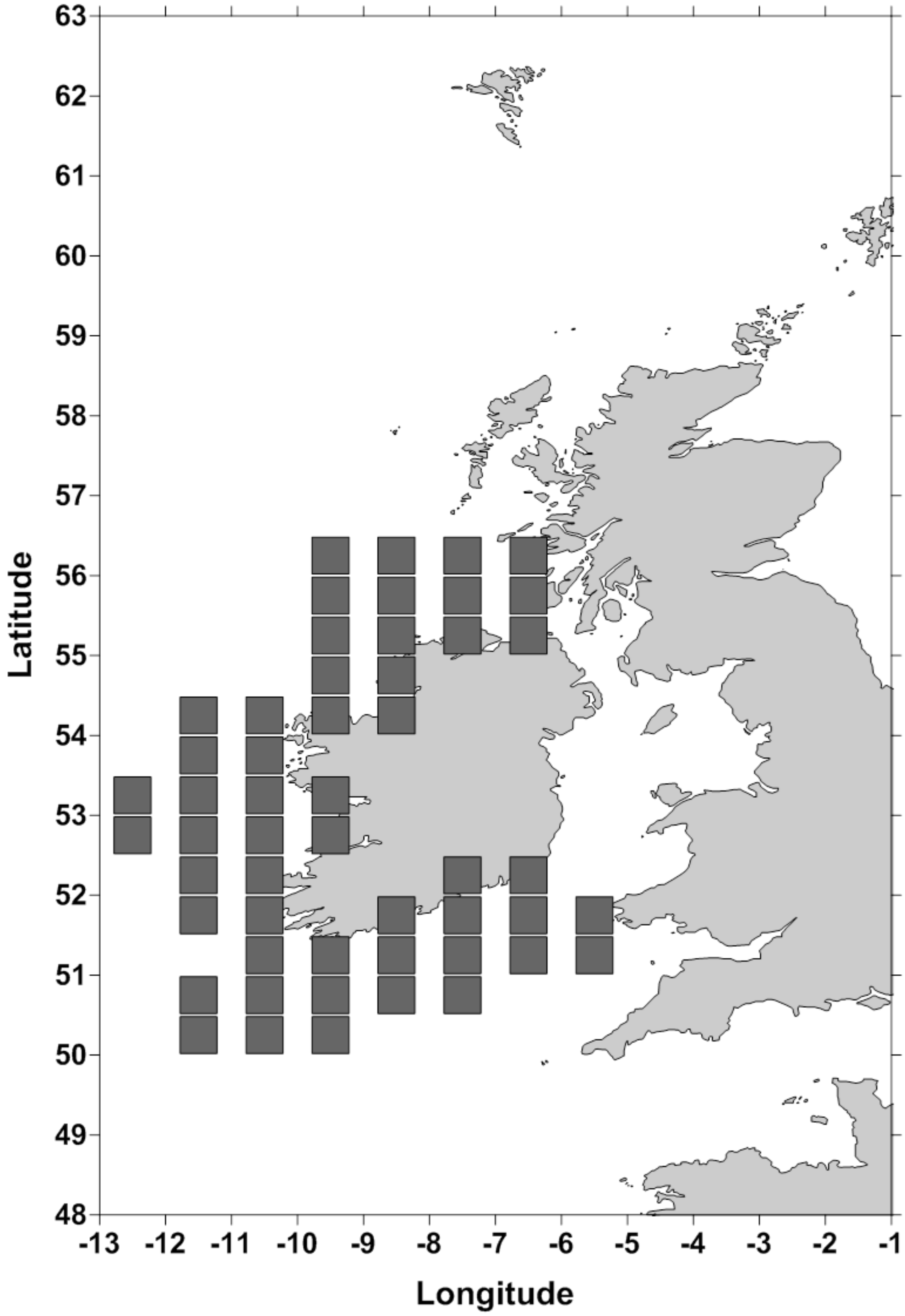
3.1.10. CSIreOT4





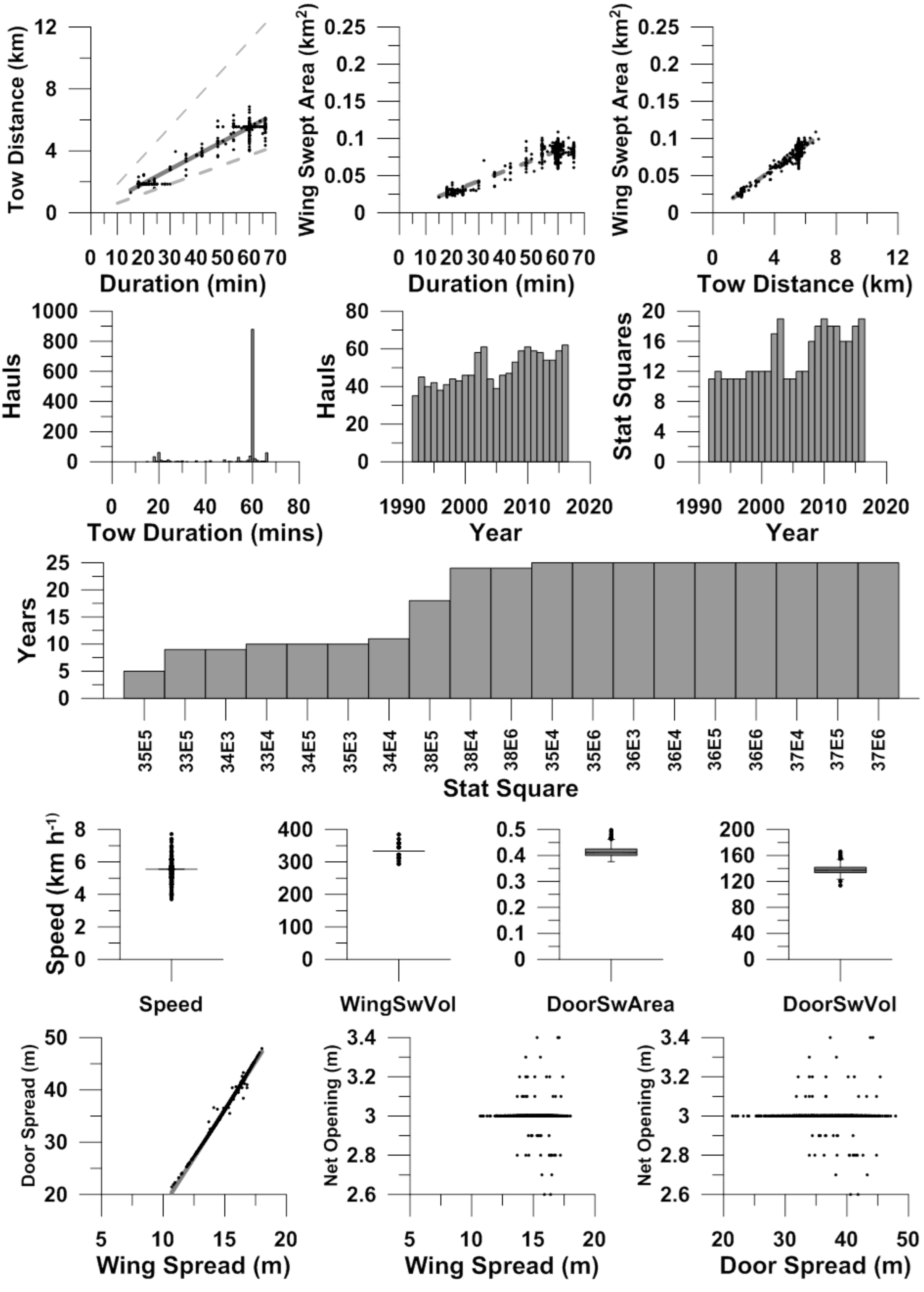
Moriarty et al. (2017) concluded that standardised monitoring in the CSIreOT4 survey only commenced in 2003. Data available for derivation of the Version 1 data product, therefore, spanned the 12 year period 2003 to 2014. By the time the Version 2 data product used in the OSPAR IA2017 was developed, data for 2015 had also been uploaded onto the DATRAS database, giving a 13 year time series, 2003 to 2015 (Greenstreet and Moriarty, 2017). Data for 2016 have subsequently added to the DATRAS database and are now, therefore, included in this Version 3 data product.

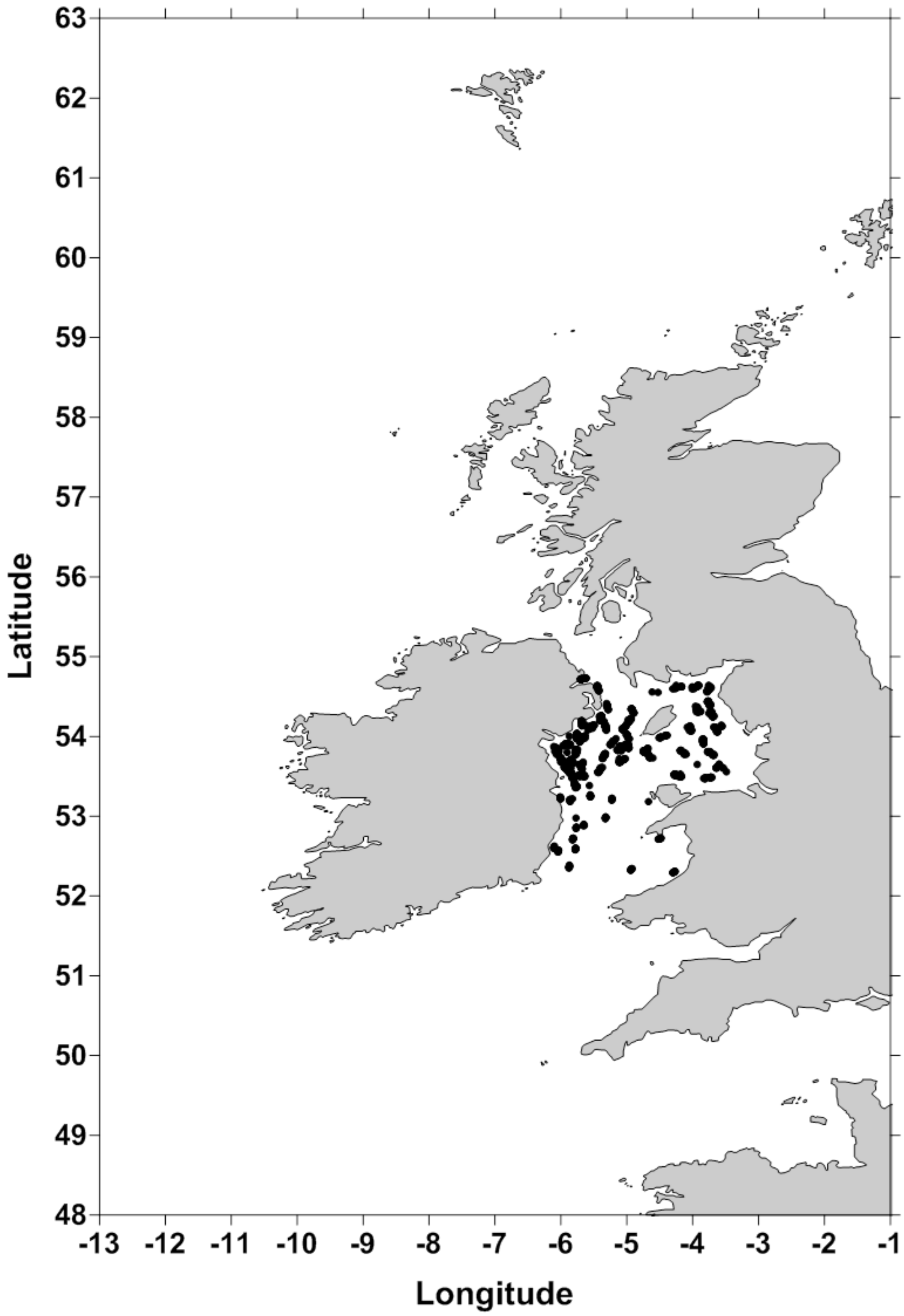
The *standard monitoring program* collected 2290 samples from a total of 78 ICES statistical rectangles over a 14 year period 2003 to 2016. However, 23 of these rectangles (30E3, 32E0, 33D7, 33E4, 33E5, 34E4, 34E5, 35E3, 35E4, 35E5, 36E0, 36E3, 36E4, 36E5, 36E6, 37E4, 37E5, 37E6, 38E4, 38E5, 38E6, 39E4 and 39E5) failed the 50% of years rule, being sampled in fewer than seven years. Exclusion of samples collected in these rectangles resulted in the loss of 94 records. A further three rectangles (31D8, 38D9 and 39D9) were not sampled in the early 20%ile of the time series period 2003 to 2005 resulting in the exclusion of an additional 77 samples. All remaining rectangles met the criterion of having been sampled in the late 20%ile of the time series period 2014 to 2016. The Version 3 *standard survey area* CSIreOT4 data product, therefore, includes data for 2119 samples collected over the 14 year period 2003 to 2016 from 52 ICES statistical rectangles. This is an increase of one ICES statistical rectangle over the version data product used in the OSPAR IA2017 (Greenstreet and Moriarty, 2017), the rectangle now meeting both criteria for inclusion, 29D8, being located at the extreme southwest corner of the Version 3 standard survey area.



63
62
61
60
59
58
57
56
55
54
53
52
51
50
49
48
Latitude
-13
-12
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
Longitude

3.1.11. CSNIrOT1

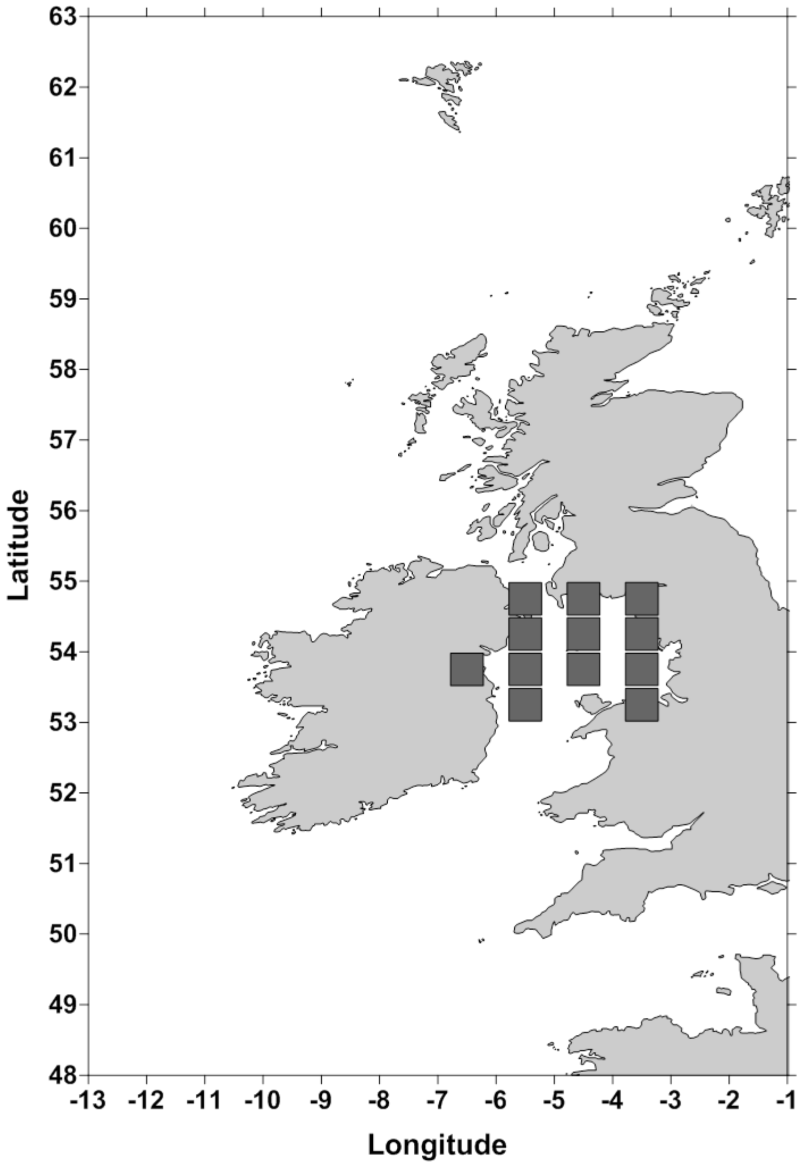




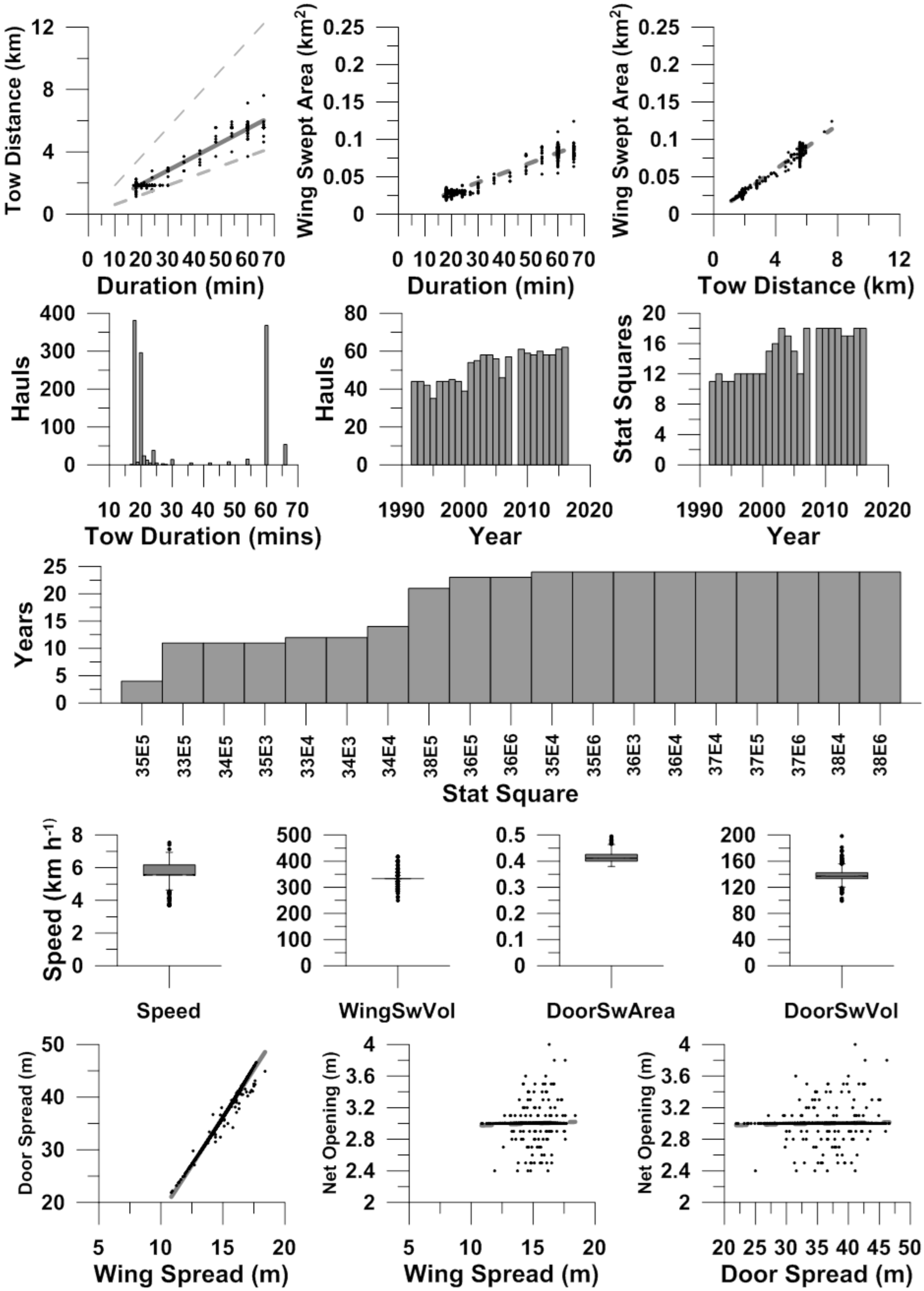
The CSNIrOT1 data available on the DATRAS portal available to Moriarty et al. (2017) for derivation of the Version 1 data product only started from 2008 onwards up to 2014. Data for 2015 were subsequently uploaded to the DATRAS database and earlier data covering the period 1992 to 2007 were provided directly from the data provider. These additional data were all included in Version 2 of the groundfish survey monitoring and assessment data product used in the OSPAR IA2017 (Greenstreet and Moriarty, 2017). Even more recently, data for 2016 have been added to the ICES DATRAS database so that the CSNIrO1 survey time series for the Version 3 data product, therefore, now spans a 25 year period between 1992 and 2015 with *standard monitoring program* consisting of data for 1234 trawl samples collected from a total of 19 ICES statistical rectangles. However, only 12 of these rectangles met the 50% of years rule, being sampled in 13 years or more. Data for the 7 ICES statistical rectangles (33E4, 33E5, 34E3, 34E4, 34E5, 35E3 and 35E5) that failed the 50% of years rule were excluded, amounting to 109 samples. All remaining 12 rectangles met the time series early and late 20%ile rules having been sampled in the 5 year periods 1992-1996 and 2012-2016.

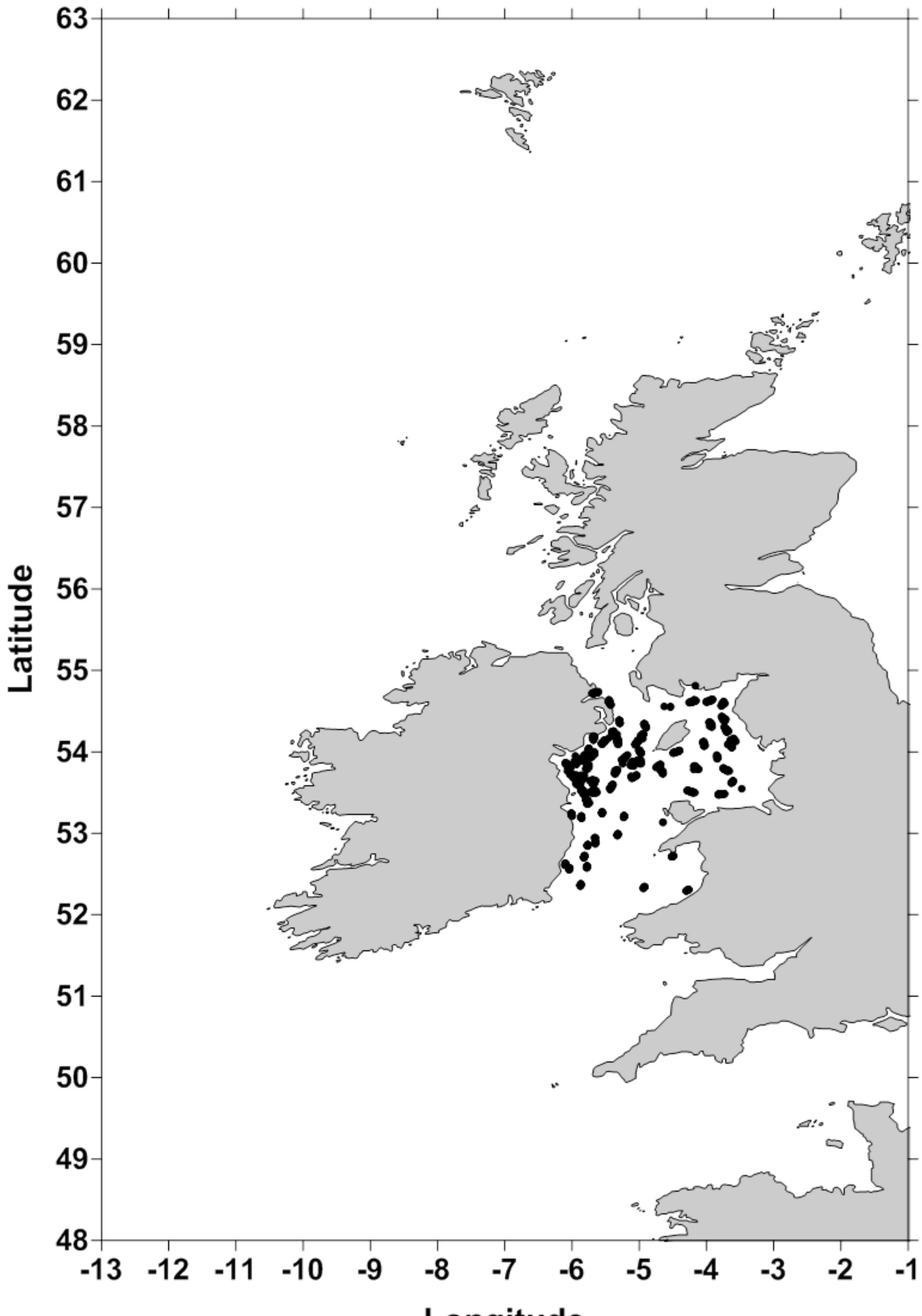
Exclusion of these samples reduced the *standard monitoring program* CSNIrOT1 data set from 1234 records to a *standard survey area* data set consisting of data from 1125 trawl samples collected over the 25 year period 1992 to 2016 from a *standard survey area* of 12 ICES statistical rectangles. However, data for the period 1992 to 2007 were still not available for download from the DATRAS data portal. These data were still only represented in this Version 3 data product because the national Data Provider has kindly supplied them to us directly. Unlike the data downloaded from DATRAS, these data have, therefore, not been subjected to exactly the same quality assurance process as that described by Moriarty et al. (2017) because they have not gone through the standard DATRAS upload scrutiny procedure. This caveat should be borne in mind by users of the CSNIrOT1 Version 3 data product, given the issues encountered with all four Spanish survey data sets (see Sections 3.1.14, 3.1.15, 3.1.16 and 3.1.19), the data for which are also not routinely uploaded to DATRAS, and which have, therefore, been subjected to a similar quality assurance process as these early Northern Irish data.

Comparison of the data diagnostic plot shown here with the one published previously by Greenstreet et al. (2017) describing the Version 2 data product reveal marked differences in some of the plots. This reflects the fact that in the most recent year of survey, net geometry monitoring equipment was fitted to the gear, where in previous years this was absent. So the 2016 data involved less modelling of wing spread, door spread and net opening, with the consequence that these actual observed data are more variable that the modelled estimates.


63
62
61
60
59
58
57
56
55
54
53
52
51
50
49
48
Latitude
-13
-12
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
Longitude

3.1.12. CSNIrOT4

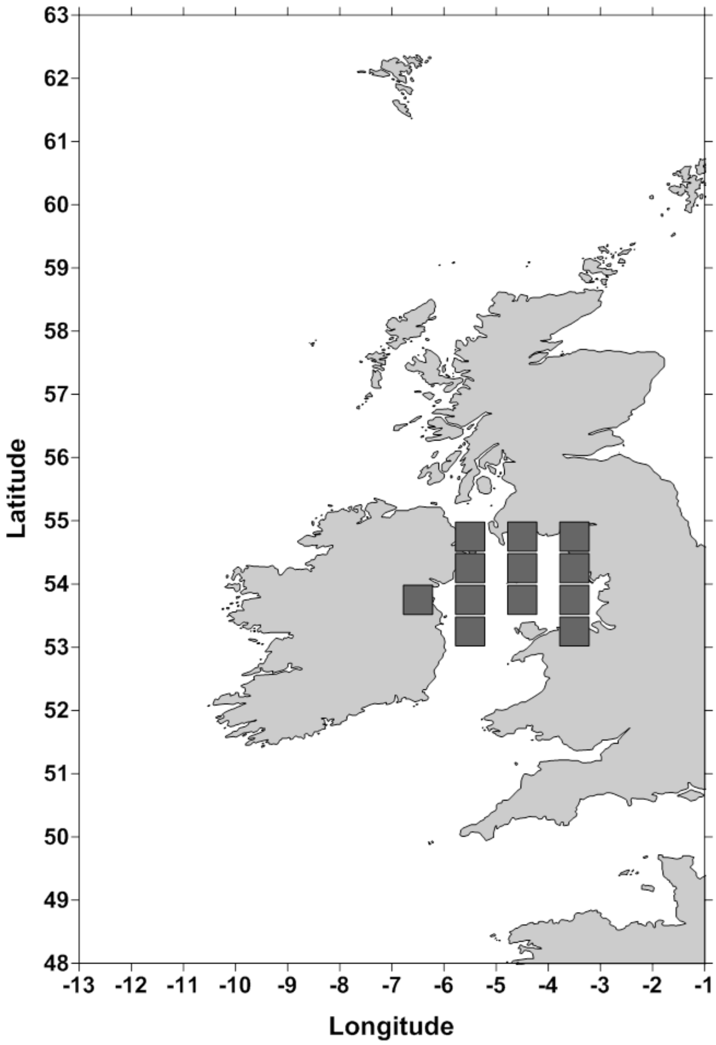




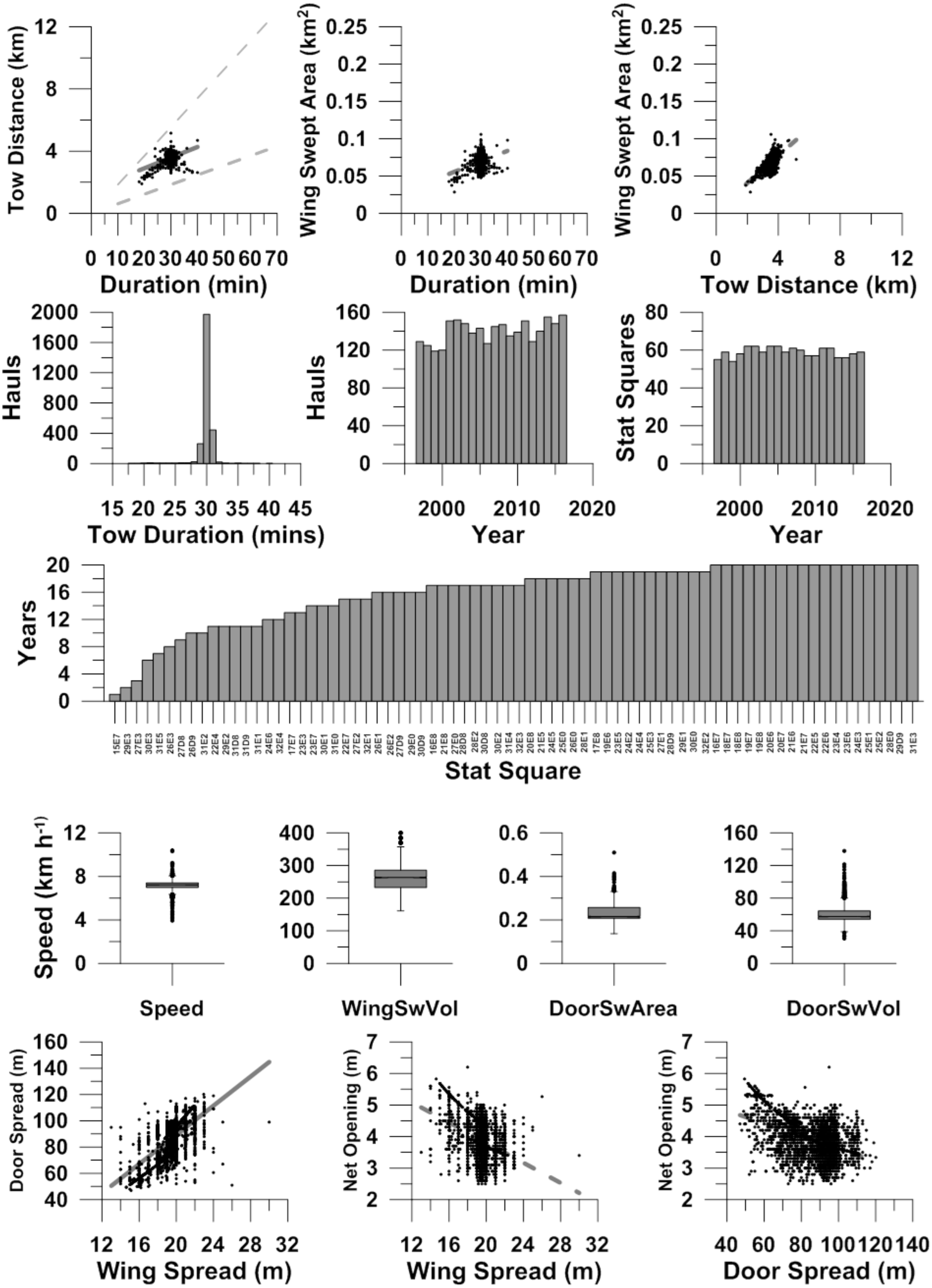
The CSNIrOT4 data available on the DATRAS portal available to Moriarty et al. (2017) for derivation of the Version 1 data product only started from 2009 onwards up to 2014. Data for 2015 were subsequently uploaded to the DATRAS database and earlier data covering the period 1992 to 2007 were provided directly from the data provider. These additional data were all included in Version 2 of the groundfish survey monitoring and assessment data product used in the OSPAR IA2017 (Greenstreet and Moriarty, 2017). Even more recently, data for 2016 have been added to the ICES DATRAS database. However, the CSNIrO1 survey time series for the Version 3 data product only contains data for 24 years from the 25 year period 1992 to 2016 because no survey took place in 2008. The *standard monitoring program* included data for 1242 trawl samples collected from a total of 19 ICES statistical rectangles, but 4 of these (33E5, 34E5, 35E3 and 35E5) failed the 50% of years rule, being sampled in fewer than 12 years. Exclusion of samples collected in these four rectangles resulted in the loss of data for 47 trawl samples. A further three rectangles (33E4, 34E3 and 34E4) failed the early 20%ile of the time series rule by not being sampled in the five year period 1992-1996. Exclusion of samples collected in these three rectangles resulted in the loss of data for a further 89 trawl samples. All remaining 12 rectangles met the time series late 20%ile rule having been sampled in the five year period 2012-2016.

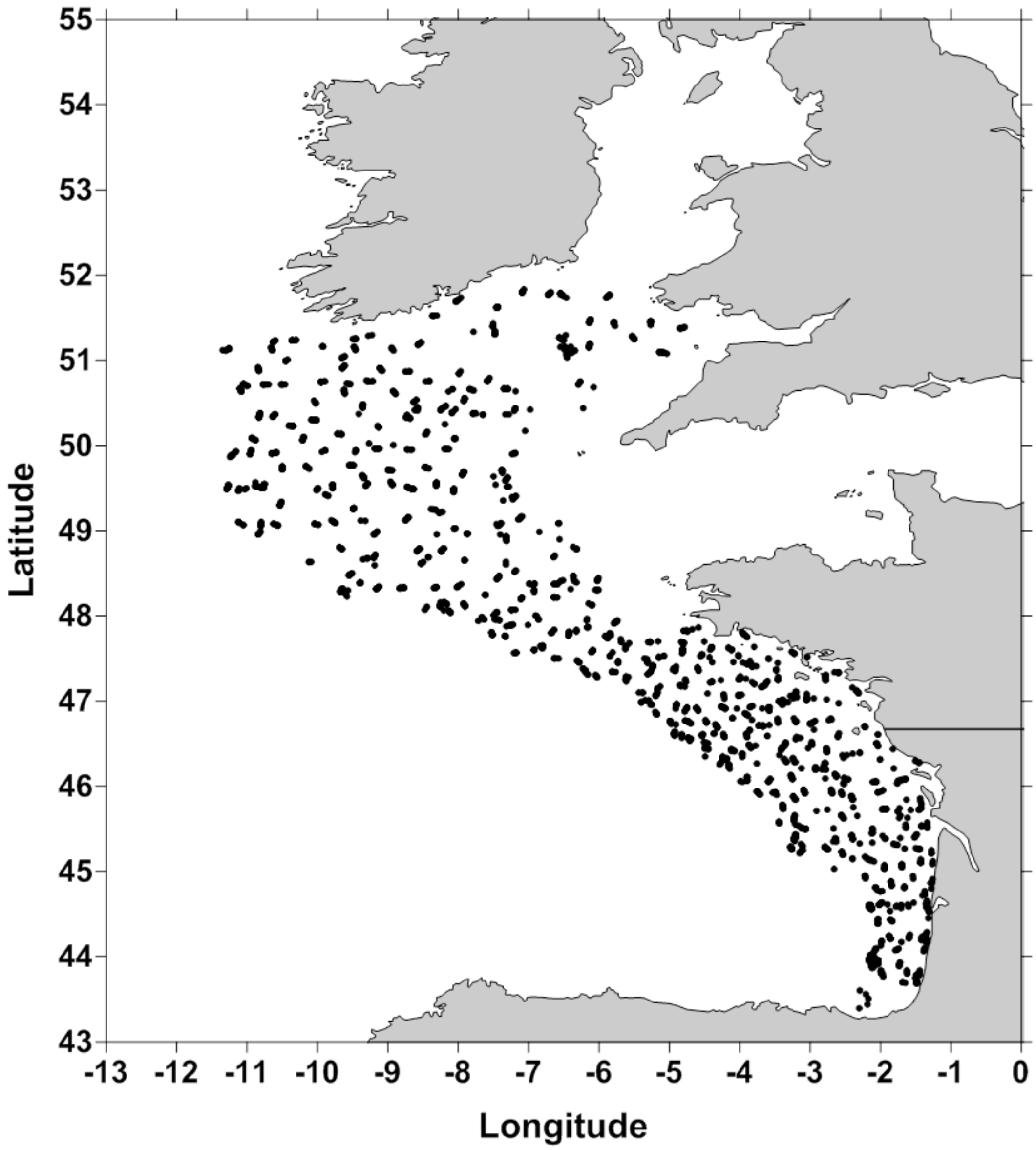
Exclusion of these samples reduced the *standard monitoring program* CSNIrOT4 data set from 1242 records to a *standard survey area* data set consisting of data from 1106 trawl samples collected over 24 years between 1992 and 2016 (with no survey in 2008) from a *standard survey area* of 12 ICES statistical rectangles. However, data for the period 1992 to 2007 were still not available for download from the DATRAS data portal. These data were still only represented in this Version 3 data product because the national Data Provider has kindly supplied them to us directly. Unlike the data downloaded from DATRAS, these data have, therefore, not been subjected to exactly the same quality assurance process as that described by Moriarty et al. (2017) because they have not gone through the standard DATRAS upload scrutiny procedure. This caveat should be borne in mind by users of the CSNIrOT4 Version 3 data product, given the issues encountered with all four Spanish survey data sets (see Sections 3.1.14, 3.1.15, 3.1.16 and 3.1.19), the data for which are also not routinely uploaded to DATRAS, and which have, therefore, been subjected to a similar quality assurance process as these early Northern Irish data.

Comparison of the data diagnostic plot shown here with the one published previously by Greenstreet et al. (2017) describing the Version 2 data product reveal marked differences in some of the plots. This reflects the fact that in the most recent year of survey, net geometry monitoring equipment was fitted to the gear, where in previous years this was absent. So the 2016 data involved less modelling of wing spread, door spread and net opening, with the consequence that these actual observed data are more variable that the modelled estimates.

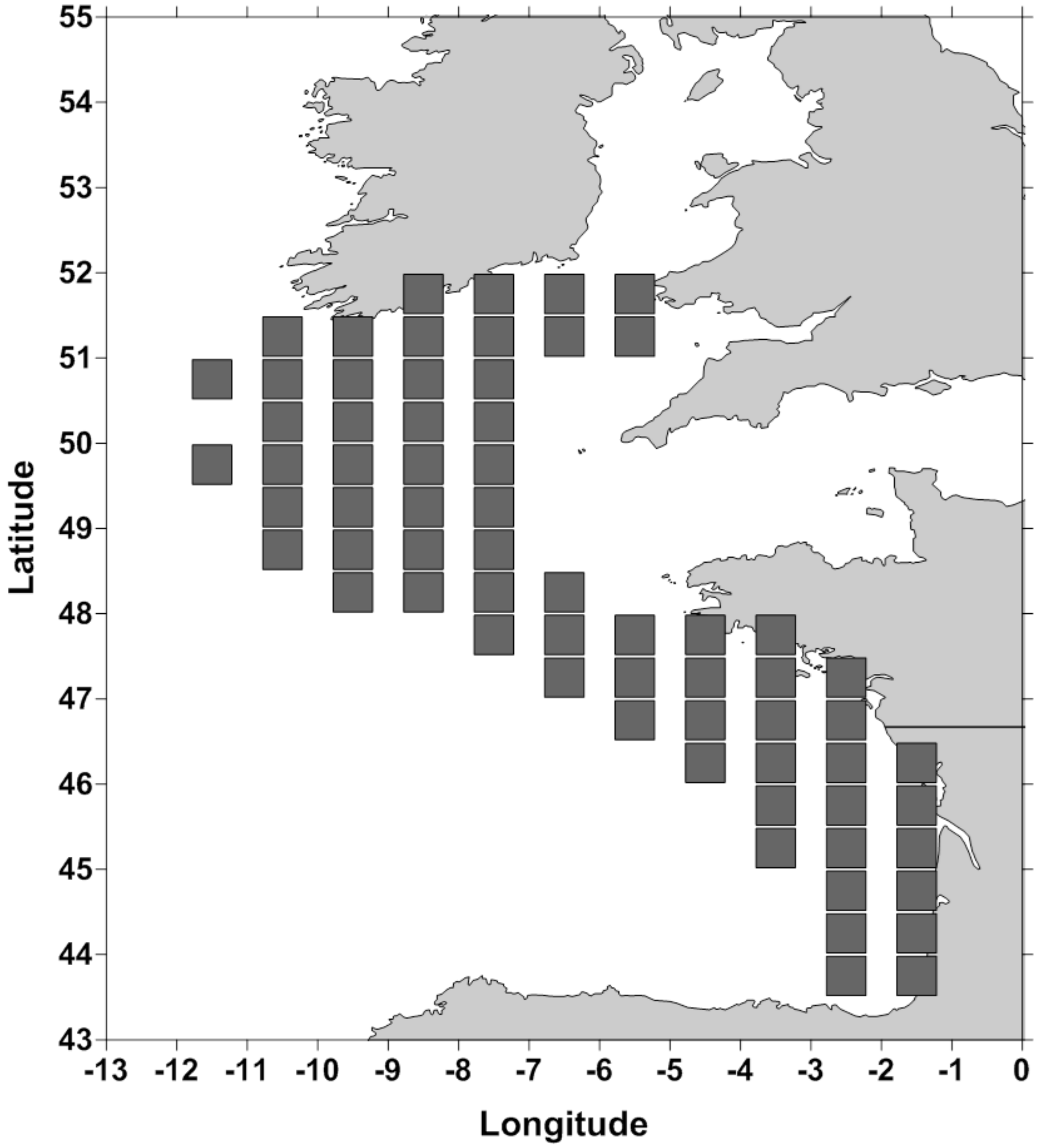

63
62
61
60
59
58
57
56
55
54
53
52
51
50
49
48
Latitude
-13
-12
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
Longitude

3.1.13. CSBBFraOT4





In deriving the initial data product, Moriarty et al. (2017) had access to CSBBFraOT4 data for the period 1997 to 2014. The Version 2 data product used for the OSPAR IA2017 included data for 2015 that had subsequently been added to the DATRAS database (Greenstreet and Moriarty, 2017). More recently data for 2016 has also been added to the DATRAS database so that the Version 3 data product now spans a 20 year period between 1997 and 2016, with no breaks. The *standard monitoring program* data product includes data for 2798 trawl samples collected from a total of 74 ICES statistical rectangles. Seven rectangles (15E7, 26E3, 27D8, 27E3, 29E3, 30E3 and 31E5) were not sampled at least once in at least ten years and so failed the 50% of years rule, resulting in data for 45 samples being excluded. One rectangle to the south of Ireland that was excluded from the Version 2 data product by this rule now met the 50% of years criterion and so was included in the Version 3 standard survey area data product. One further rectangle, 31D8 at the extreme northwest of the standard survey area, was last sampled in 2012 and so failed the time series late 20%ile four year (2013-2016) period rule. This rectangle had been included in the Version 2 standard survey area data product because it met the then 2012-2015 20% end period criterion. Excluding data for this rectangle removed a further 11 samples. One rectangle, 24E6, that failed this rule when constructing the Version 2 data product, and so was excluded from the *standard survey area* data set (Greenstreet and Moriarty, 2017), was actually sampled again in 2016, so once more became included in this Version 3 *standard survey area* data product. This rectangle lies off Lorient on the coast of France in southern Brittany. These exclusions reduced the CSBBFraOT4 *standard survey area* data product to 2742 trawl samples collected over the 20 year period 1997 to 2016 from a *standard survey area* of 66 ICES statistical rectangles.



3.1.14. BBIC(n)SpaOT4

A major issue with the Version 2 data product used in the OSPAR IA2017 was reported by the national Data Provider. The number of trawl samples in the BBIC(n)SpaOT4 standard monitoring programme, and, therefore, available for selection in the standard survey area data product, was far fewer than had been received in the original data submission obtained directly from the Data Provider. This error was identified as having been caused by corruption in many records of one of the fields used to generate the unique haul identifier codes. Because these haul identifier codes provide the relational link between the Sampling Information and Biological Information files, this meant that for the trawl samples affected, no related abundance at length data could be found in the Biological Information.

In the DATRAS database, it is conceivable that trawl samples might be recorded as valid in the Sampling Information data, but have no associated species abundance at length at length data in the Biological Information files. This is because the only information recorded for the haul was age at length data. Our data products do not include age at length data, so there was no need for such trawl samples to be retained in the data product, and every reason to exclude them to facilitate more straight forward determination samples size (number of trawl samples) when analysing the data for the OSPAR IA2017, or for any other analysis. The code used to generate the data products therefore removed any trawl samples recorded in the Sampling Information files for which there was no associated species abundance at length data in the Biological Information files. When applied to the BBIC(n)SpaOT4 data, this routine removed all the trawl samples where the unique haul identifier code was corrupted.

Had the BBIC(n)SpaOT4 survey data been uploaded routinely to the DATRAS database, the DATRAS upload data checking procedures would have caught this anomaly. Our R code used to derive the data product was primarily constructed to process data downloaded from DATRAS; the code, therefore, assumed that these checks had already been made. This highlights the sort of problems that can occur when building automated data quality assurance routines when the data involved are obtained from multiple sources. The issue with the BBIC(n)SpaOT4 survey data underlines the fact that, in an ideal situation, groundfish survey data to be used in OSPAR and other MSFD assessments of fish species and community status should, in the first instance, all be routinely uploaded to the ICES DATRAS database.

In attempting to construct the Version 3 data product for the BBIC(n)SpaOT4 survey, having addressed the unique haul identifier miss-match problem, a further issue was detected. Rather than delay release of the Version 3 data products for the remainder of the surveys, the decision was taken to exclude the BBIC(n)SpaOT4 survey for the time being. The BBIC(n)SpaOT4 survey data product could be released at a later date, should there be sufficient call for it.

3.1.15. BBIC(s)SpaOT1

A major issue with the Version 2 data product used in the OSPAR IA2017 was reported by the national Data Provider. The number of trawl samples in the BBIC(s)SpaOT1 standard monitoring programme, and, therefore, available for selection in the standard survey area data product, was far fewer than had been received in the original data submission obtained directly from the Data Provider. This error was identified as having been caused by corruption in many records of one of the fields used to generate the unique haul identifier codes. Because these haul identifier codes provide the relational link between the Sampling Information and Biological Information files, this meant that for the trawl samples affected, no related abundance at length data could be found in the Biological Information.

In the DATRAS database, it is conceivable that trawl samples might be recorded as valid in the Sampling Information data, but have no associated species abundance at length at length data in the Biological Information files. This is because the only information recorded for the haul was age at length data. Our data products do not include age at length data, so there was no need for such trawl samples to be retained in the data product, and every reason to exclude them to facilitate more straight forward determination samples size (number of trawl samples) when analysing the data for the OSPAR IA2017, or for any other analysis. The code used to generate the data products, therefore, removed any trawl samples recorded in the Sampling Information files for which there was no associated species abundance at length data in the Biological Information files. When applied to the BBIC(s)SpaOT1 data, this routine removed all the trawl samples where the unique haul identifier code was corrupted.

Had the BBIC(s)SpaOT1 survey data been uploaded routinely to the DATRAS database, the DATRAS upload data checking procedures would have caught this anomaly. Our R code used to derive the data product was primarily constructed to process data downloaded from DATRAS; the code, therefore, assumed that these checks had already been made. This highlights the sort of problems that can occur when building automated data quality assurance routines when the data involved are obtained from multiple sources. The issue with the BBIC(s)SpaOT1 survey data underlines the fact that, in an ideal situation, groundfish survey data to be used in

OSPAR and other MSFD assessments of fish species and community status should, in the first instance, all be routinely uploaded to the ICES DATRAS database.

In attempting to construct the Version 3 data product for the BBIC(s)SpaOT1 survey, having addressed the unique haul identifier miss-match problem, a further issue was detected. Rather than delay release of the Version 3 data products for the remainder of the surveys, the decision was taken to exclude the BBIC(s)SpaOT1 survey for the time being. The BBIC(s)SpaOT1 survey data product could be released at a later date, should there be sufficient call for it.

3.1.16. BBIC(s)SpaOT4

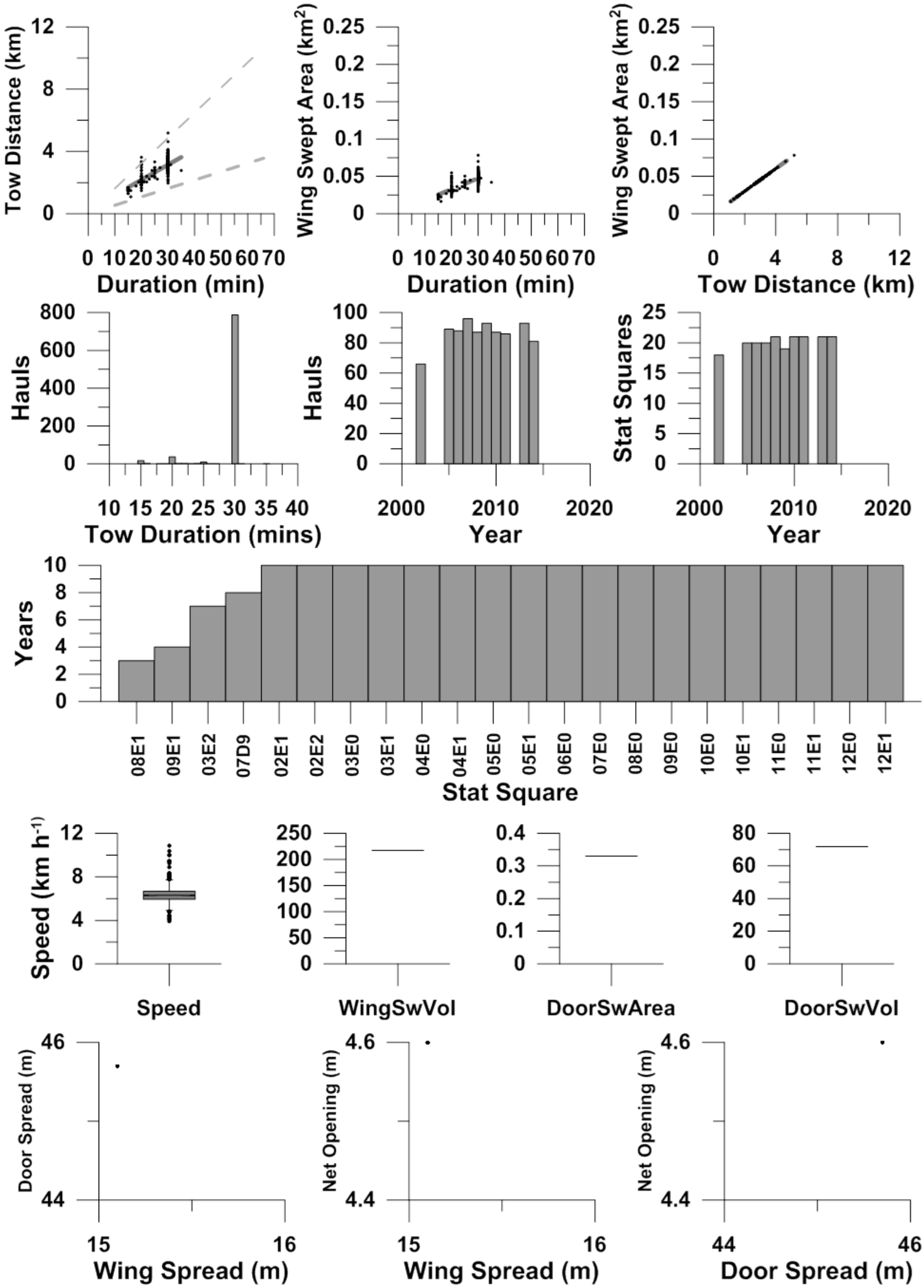
A major issue with the Version 2 data product used in the OSPAR IA2017 was reported by the national Data Provider. The number of trawl samples in the BBIC(s)SpaOT4 standard monitoring programme, and, therefore, available for selection in the standard survey area data product, was far fewer than had been received in the original data submission obtained directly from the Data Provider. This error was identified as having been caused by corruption in many records of one of the fields used to generate the unique haul identifier codes. Because these haul identifier codes provide the relational link between the Sampling Information and Biological Information files, this meant that for the trawl samples affected, no related abundance at length data could be found in the Biological Information.

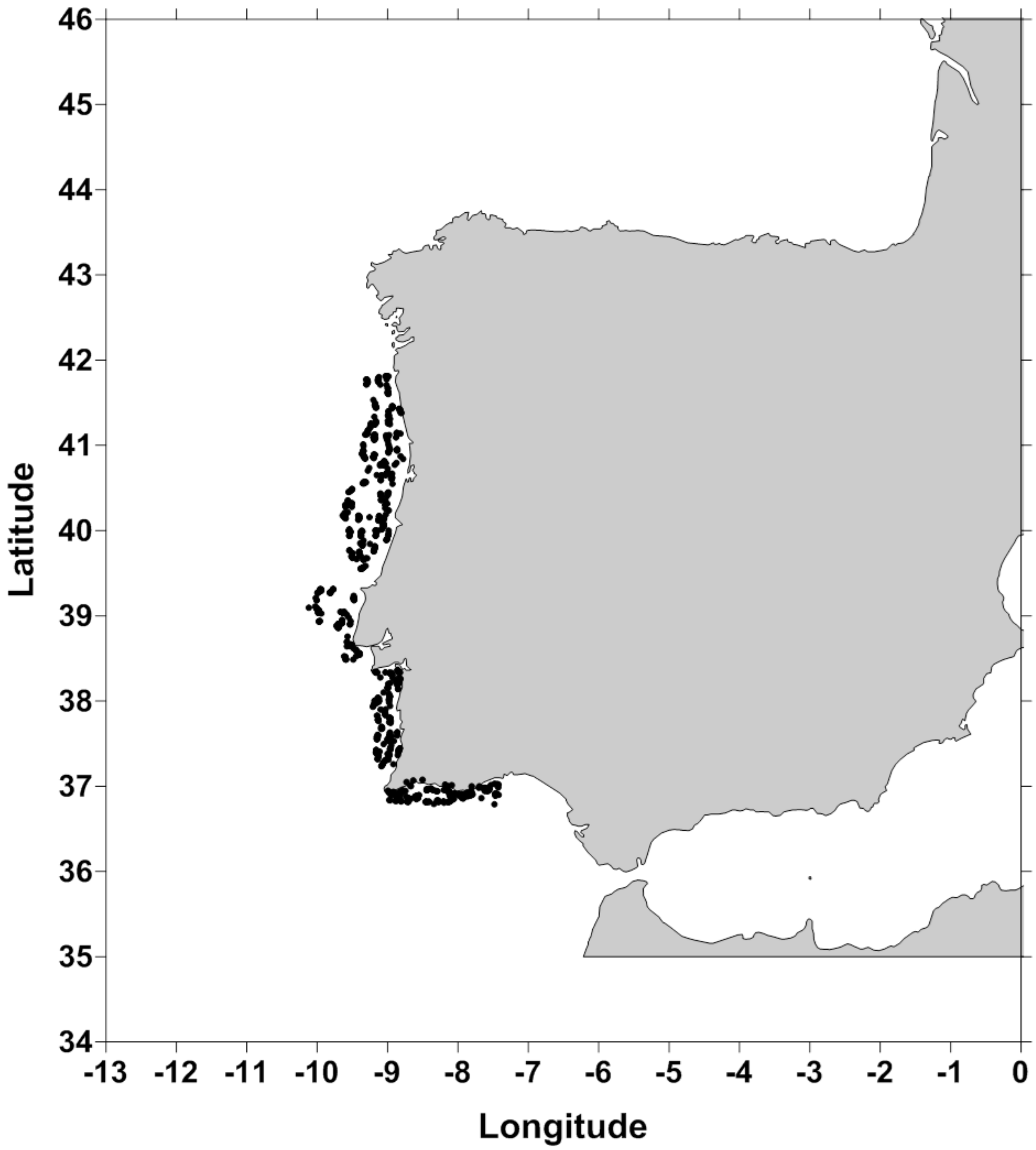
In the DATRAS database, it is conceivable that trawl samples might be recorded as valid in the Sampling Information data, but have no associated species abundance at length at length data in the Biological Information files. This is because the only information recorded for the haul was age at length data. Our data products do not include age at length data, so there was no need for such trawl samples to be retained in the data product, and every reason to exclude them to facilitate more straight forward determination samples size (number of trawl samples) when analysing the data for the OSPAR IA2017, or for any other analysis. The code used to generate the data products therefore removed any trawl samples recorded in the Sampling Information files for which there was no associated species abundance at length data in the Biological Information files. When applied to the BBIC(s)SpaOT4 data, this routine removed all the trawl samples where the unique haul identifier code was corrupted.

Had the BBIC(s)SpaOT4 survey data been uploaded routinely to the DATRAS database, the DATRAS upload data checking procedures would have caught this anomaly. Our R code used to derive the data product was primarily constructed to process data downloaded from DATRAS; the code therefore assumed that these checks had already been made. This highlights the sort of problems that can occur when building automated data quality assurance routines when the data involved are obtained from multiple sources. The issue with the BBIC(s)SpaOT4 survey data underlines the fact that, in an ideal situation, groundfish survey data to be used in OSPAR and other MSFD assessments of fish species and community status should, in the first instance, all be routinely uploaded to the ICES DATRAS database.

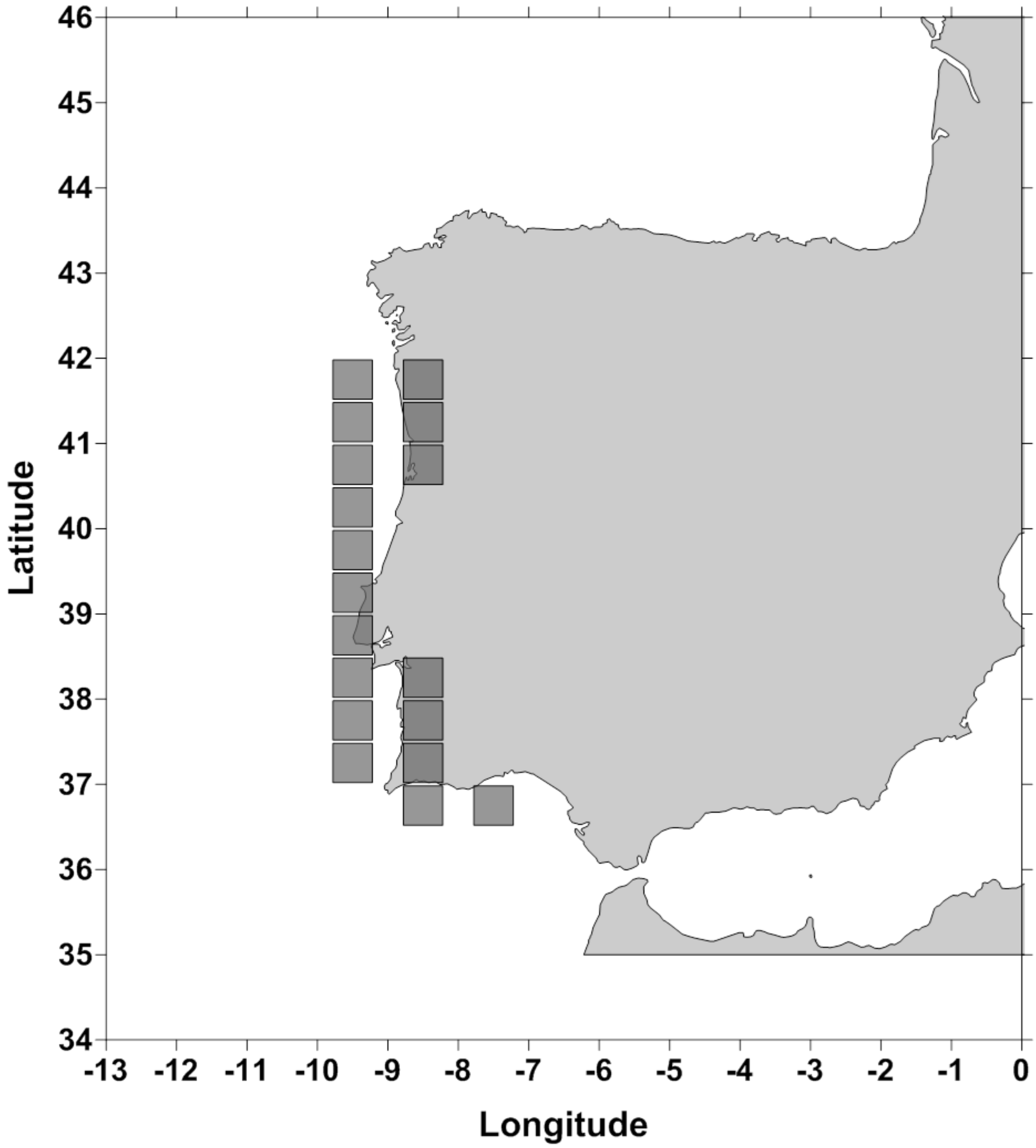
In attempting to construct the Version 3 data product for the BBIC(s)SpaOT4 survey, having addressed the unique haul identifier miss-match problem, a further issue was detected. Rather than delay release of the Version 3 data products for the remainder of the surveys, the decision was taken to exclude the BBIC(s)SpaOT4 survey for the time being. The BBIC(s)SpaOT4 survey data product could be released at a later date, should there be sufficient call for it.

3.1.17. BBICPorOT4

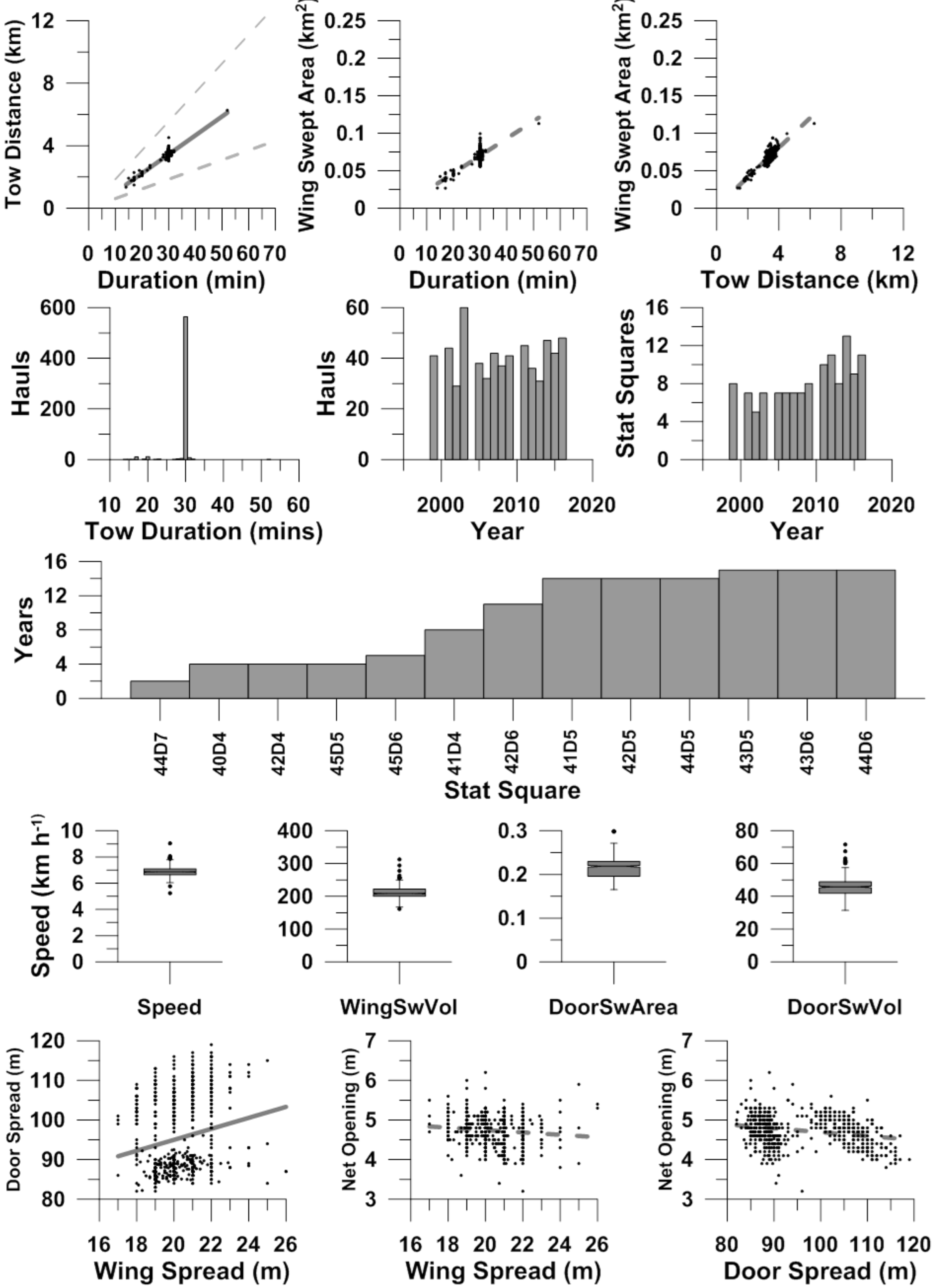


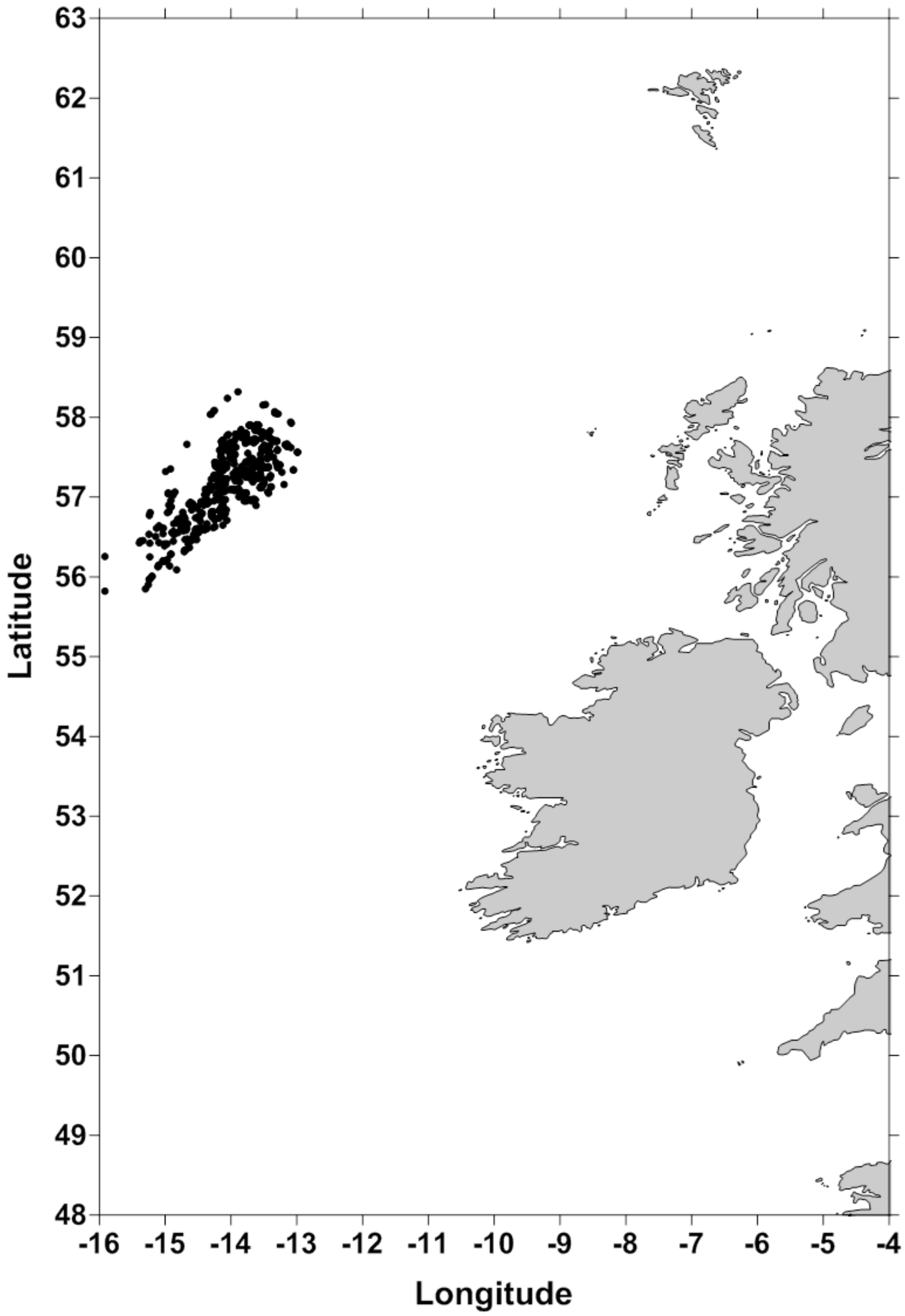


The BBICPorOT4 data set consisted of ten years of data collected over the period between 2002 and 2014 with no survey undertaken in 2003, 2004 and 2012. This survey data set does not appear to have been updated (other than the corrections made as part of the process to derive the data products) since the time of the original download. No new data have been added and the time series covered in the Version 1 data product (Moriarty et al., 2017), the Version 2 data product used in the OSPAR IA21017 (Greenstreet and Moriarty, 2017), and in this Version 3 data product are all the same. Data were collected from 22 ICES statistical rectangles, but only 20 of these met the 50% of years rule, requiring the exclusion of nine samples collected in rectangles 08E1 and 09E1. In addition, two rectangles (03E2 and 07D9) were not sampled in the time series start 20%ile four year (2002-2003) period and these data were also excluded (18 samples). These exclusions reduced the original BBICPorOT4 *standard monitoring program* data set of 866 samples to 839 samples collected in ten years spanning the period 2002 to 2014 from a *standard survey area* of 18 ICES statistical rectangles.

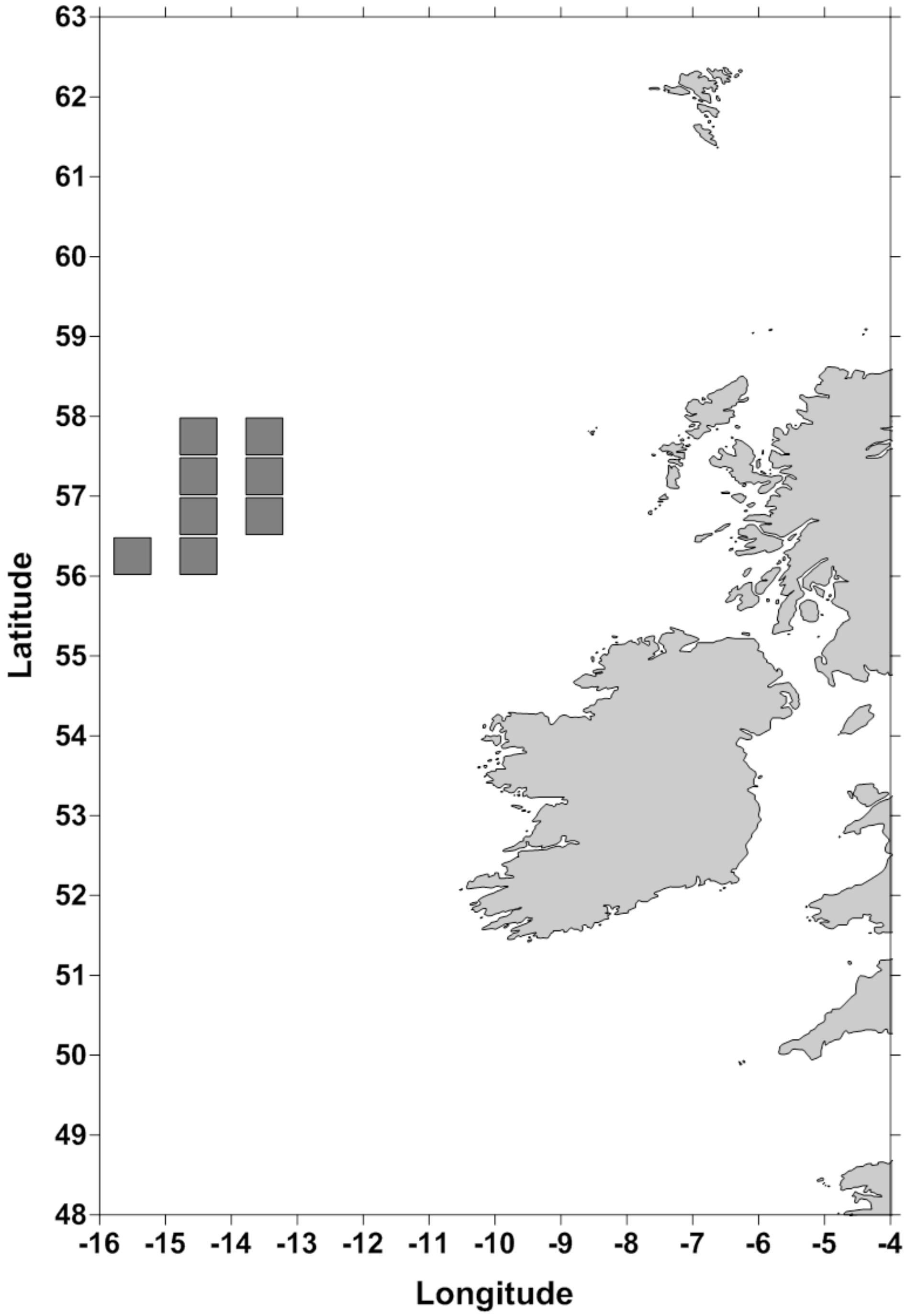


3.1.18. WAScoOT3





The Version 2 WAScoOT3 data product used for the OSPAR IA2017 consisted of 14 years of data collected over the period between 1999 and 2015, but with no survey undertaken in 2000, 2004 and 2010 (Greenstreet and Moriarty, 2017). This was the same time series of data available to Moriarty et al. (2017) for the original derivation of the Version 1 data product. However, data for 2016 have subsequently been added to the ICES DATRAS database and these are included in this Version 3 data product, giving a time series spanning the period 1999 to 2016, but with just 15 years of data; the same years are missing. The *standard monitoring program* data set consisted of 613 samples collected from 13 ICES statistical rectangles, but only eight of these met the 50% of years rule of having been sampled at least once in at least eight years. Excluding the 28 samples collected in five rectangles (40D4, 42D4, 44D7, 45D5 and 45D6) failing this criterion left a total of 585 samples in the WAScoOT3 *standard survey area* data product collected in 15 years between 1999 and 2016 from an *standard survey area* of 8 ICES statistical rectangles. These eight rectangles all met the criteria of having been sampled at least once in the three year periods 1999-2001 and 2014-2016 that represented the early and late 20%iles of the time series.



3.1.19. WASpaOT3

A major issue with the Version 2 data product used in the OSPAR IA2017 was reported by the national Data Provider. The number of trawl samples in the WASpaOT3 standard monitoring programme, and, therefore, available for selection in the standard survey area data product, was far fewer than had been received in the original data submission obtained directly from the Data Provider. This error was identified as having been caused by corruption in many records of one of the fields used to generate the unique haul identifier codes. Because these haul identifier codes provide the relational link between the Sampling Information and Biological Information files, this meant that for the trawl samples affected, no related abundance at length data could be found in the Biological Information.

In the DATRAS database, it is conceivable that trawl samples might be recorded as valid in the Sampling Information data, but have no associated species abundance at length at length data in the Biological Information files. This is because the only information recorded for the haul was age at length data. Our data products do not include age at length data, so there was no need for such trawl samples to be retained in the data product, and every reason to exclude them to facilitate more straight forward determination samples size (number of trawl samples) when analysing the data for the OSPAR IA2017, or for any other analysis. The code used to generate the data products, therefore, removed any trawl samples recorded in the Sampling Information files for which there was no associated species abundance at length data in the Biological Information files. When applied to the WASpaOT3 data, this routine removed all the trawl samples where the unique haul identifier code was corrupted.

Had the WASpaOT3 survey data been uploaded routinely to the DATRAS database, the DATRAS upload data checking procedures would have caught this anomaly. Our R code used to derive the data product was primarily constructed to process data downloaded from DATRAS; the code therefore assumed that these checks had already been made. This highlights the sort of problems that can occur when building automated data quality assurance routines when the data involved are obtained from multiple sources. The issue with the WASpaOT3 survey data underlines the fact that, in an ideal situation, groundfish survey data to be used in OSPAR and other MSFD assessments of fish species and community status should, in the first instance, all be routinely uploaded to the ICES DATRAS database.

In attempting to construct the Version 3 data product for the WASpaOT3 survey, having addressed the unique haul identifier miss-match problem, a further issue was detected. Rather than delay release of the Version 3 data products for the remainder of the surveys, the decision was taken to exclude the WASpaOT3 survey for the time being. WASpaOT3 survey data product could be released at a later date, should there be sufficient call for it.

3.2. Biological Information

Generally the kNN procedure was able to partition genus-level and family-level identification codes in the Biological Information to the most probable species-level codes. However, in some instances the data were insufficient to permit this and the kNN analysis suggested that the most appropriate course of action was to aggregate all species-level identifications down to the next lower taxonomic level that the data could support. Thus for all surveys, all sandeel and all goby species level identifications had to be aggregated down to the family-level IDcodes Ammodytidae and Gobiidae respectively. Family-level codes had also to be applied to all fish identified as belonging to the Scorpaenidae and Zoarcidae families in the CSIreOT4 survey. In the two Northern Irish surveys, CSNIrOT1 and CSNIrOT4, all fish identified as belonging to the genus Alosa, Argentina, Salmo and Syngnathus had also to be aggregated down to the coarser genus-level IDcodes. The genus level code Diaphus was used in the CSBBFraOT4 survey and could not be resolved to species-level because no other fish belonging to this genus were sampled.

4. Acknowledgments

We could not have undertaken this project without the huge help and support of the numerous data providers who submit groundfish survey data to the ICES DATRAS data portal. We are particularly indebted to Finlay Burns (Marine Scotland Science, Scotland), Gary Burt and Jim Ellis (CEFAS, England), Mathieu Lundy (Agri-Food and Biosciences Institute, Northern Ireland), David Stokes (Marine Institute, Ireland), Barbara Bland (Swedish University of Agricultural Sciences), Matthias Kloppmann and Anne Sell (Thünen Institute of Sea Fisheries, Germany), Jenifer Devine and Trine Haugen (Institute of Marine Research, Norway), Yves Vérin, Michele Salaun, Franck Coppin and Morgane Travers (Ifremer, France), Kai Ulrich Wieland (DTU Aqua, Denmark), Ingeborg de Boois and Ralf van Hal (Wageningen Marine Research, The Netherlands), Francisco Velasco (Instituto Español de Oceanografía, Spain), Corina Chaves (Instituto Portugues do Mar e da Atmosfera, Portugal) for their unlimited patience and unswerving support and assistance with all out numerous queries. We would also like to thank members of the ICES working groups WGIBTS, WGBEAM, and WGMSFDemo for their feedback and other contributions to this project. We are also grateful to members of the ICES

Secretariat, particularly Mark Dickey-Collas, Neil Holdsworth, Vaishav Soni, for their willingness to acknowledge the need for this work and their assistance with seeing it through. We thank Rob Fryer for all his help and statistical advice, and we are also very grateful to colleagues at Marine Scotland Science, particularly Bill Turrell, Matt Gubbins and Colin Moffatt, for the support provided when the going got tough! Finally, we are grateful to all our colleagues involved in the OSPAR IA2017 process for their patience in waiting for the groundfish survey monitoring and assessment data products to be delivered.

5. References

Daan, N. 2001. The IBTS database: a plea for quality control. ICES Document CM 2001/T: 03. 19 pp.

Fraser, H.M., Greenstreet, S.P.R., 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: 1800–1819.

Moriarty, M., Greenstreet, S.P.R. and Rasmussen, J. 2017. Derivation of Groundfish Monitoring and Assessment Data Products for the Northeast Atlantic Area. Scottish Marine and Freshwater Science Report. (in prep).

Greenstreet, S.P.R. and Moriarty, M. 2017. OSPAR Interim Assessment Fish Indicator Data Manual (Relating to Version 2 of the Groundfish Survey Monitoring and Assessment Data Product). Scottish Marine and Freshwater Science Report. (in prep).

ter Hofstede, R. and Daan, N. 2006. Quality check surveys: DATRAS North Sea IBTS. In Report of the International Bottom Trawl survey Working Group (IBTSWG) 27-31 march 2006, Lysekil, Sweden. ICES CM 2006/RCM:03, Ref ACFM, 298pp.

ter Hofstede, R. and Daan, N. 2008. A proposal for a consistent use of the North Sea IBTS data. ICES CM 2008/R:25 6pp.

Marine Scotland Science
Marine Laboratory
375 Victoria Road
Aberdeen
AB11 9DB

Copies of this report are available from the Marine Scotland website at:
www.gov.scot/marinescotland