THE ATLANTIC-WIDE RESEARCH PROGRAMME FOR BLUEFIN TUNA (GBYP Phase 11)

SHORT-TERM CONTRACT (ICCAT GBYP 03/2022)

Design- and model-based inference to estimate density, abundance and biomass of bluefin tuna. Reanalysis of 2017-2021 Aerial Surveys of Region A

Final Report

June 2022

CREEM, University of St Andrews

M Chudzinska, L Burt, S Buckland

This project is co-funded by the European Union

Please cite this report as: Chudzinska *et al.* (2022) Design-based inference to estimate density, abundance and biomass of bluefin tuna. Reanalysis of 2017-2021 Aerial Surveys of Region A. Report number CREEM-2022-03. Provided to ICCAT, July 2022 (Unpublished).

Document control

Please consider this document as uncontrolled copy when printed.

Version	Date	Reason for issue	Prepared by	Checked by
DRAFT v1	27.06.2022	Outline of report	MC/LB	SB
FINAL	30.06.2022	Review by co-authors	MC/LB/SB	MC/LB/SB
REVISED	07.09.2022	Review by ICCAT		MC/LB

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Summary

Information was required on the distribution, biomass and abundance of spawning stock (adult individuals) of bluefin tuna (BFT) in the Balearic Sea (Region A). Aerial surveys took place between late May and early August in years 2017-2019 and, recently, in 2021. Line transect distance sampling methods (Buckland *et al.*, 2001) have been applied to the data collected 2017-2019 surveys to estimate tuna abundance (e.g. Chudzinska *et al.*, 2021). One consideration in the survey design has been whether the Region used completely covered the spawning regions. In 2021, the survey concentrated on the core area of Region A (A-core) and an area surrounding A-core (A-outer) and so estimates could be compared between these two areas. This outer area (excluding the core area) was twice the size of the A-core. The transect design differed between these two regions resulting in distance flown by the plane 2.5 times greater in A-core than in A-outer. The estimation of distance to the detected schools differed between 2021 and previous years as well as no small schools were detected in 2021. These difference might have affected the probability of detection and encounter rate.

Two approaches to estimating abundance have been used: design-based methods (Buckland *et al.*, 2001) and model-based methods (e.g. Hedley and Buckland, 2004). Design-based methods estimate a constant density within a survey block, where as model-based methods allow density and abundance to be estimated as a function of location and environment, allowing density to vary spatially throughout a region. The objective was to assess the feasibility of using model-based methods to estimate tuna abundance (Burt *et al.*, 2021) and in the two areas of interest and compare estimates with previous years.

There were too few sightings to use data from 2021 only (8 and 12 in A-core and A-outer, respectively) and so data from 2017-2019 were included for both analyses. In these previous years only the A-core area was surveyed.

A range of models including various covariates and detection functions were fitted to the 2017-2021 data, both excluding A-outer and also including A-outer. The final models included company and school size as explanatory variables, as it was for 2017-2019 models (Chudzinska *et al.*, 2021). The estimated abundance in A-core in 2021 was 26,300 BFT (CI: 9,620 - 71,920) when only sightings in A-core were included in the model, and 26,110 BFT (CI: 9 590 - 71 130) when A-core and A-outer were included. These are lower estimates than in 2019. The estimated abundance in A-outer was 80,990 BFT (CI: 26,860 - 244,170).

Two models were fitted in the model-based approach, one to describe the number of groups and the other to describe group size. To illustrate this approach a limited set of potential explanatory variables were used, such as sea surface temperature on the day of the survey, the difference in sea surface temperature between day of the survey and 10 days before and depth, year and location. The selected models explained only small fractions of variation in density of groups and group sizes and there are large uncertainties around the estimated values. These data present considerable challenges for modellers and further analyses may benefit from careful consideration of environmental covariates.

Introduction

In order to estimate density and abundance of bluefin tuna (BFT; *Thunnus thynnus*) in the Balearic Sea (Region A), a series of aerial surveys were undertaken. Surveys took place between late May and early August in years 2010-2011, 2013, 2015, 2017-2019 and 2021 in Region A. If the spawning region is not covered completely, the abundance of tuna will be underestimated. Previous analyses (Burt *et al.* 2021, see Figure 6) suggested that this may be the case, in Region A for which the analysis was undertaken. In 2021 two series of surveys were conducted: one for within the core area of Region A (hereafter A-core) and one in a region surrounding the core region (A-outer). Note that here, A-outer refers to an area excluding A-core. In this report, we re-analyse the data incorporating the newest survey from 2021 in Region A to estimate density, abundance and biomass for BFT and compare the estimates for these two areas (A-core and A-outer). For all surveys, line transect distance sampling (DS) methods (Buckland *et al.* 2001) were used; the planes flew along pre-determined transects, or tracklines, and trained observers searched for animals, recording relevant information when an animal, or group of animals (here schools), was detected. DS methods estimate average density of animals within the region of interest and so there can be a step-change in density between regions. In reality, it is likely that density changes gradually between contiguous regions and in response to environmental conditions and location. Model-based methods (Hedley and Buckland, 2004) estimate density which can vary in response to location/habitat and so predictions can be made based on the model and values of variables included in the model. This has the potential to provide insight into environmental drivers of tuna density. The predicted surface may indicate regions of higher and lower density. If density is estimated to be high on the edges of the region, this may indicate that a substantial fraction of the population of interest is outside the region. The aims of this report are, therefore:

- To update density, biomass, and abundance estimates of BFT in the Balearic Sea in A-core following the design-based approach including data from 2021 (Task 1)
- To compare estimates of density, and abundance of BFT between A-core and A-outer using both design- and model-based approaches (Task 2).

Methods

Task 1 (using data from A-core only) was conducted in two steps: a. data from 2021 survey in A-core only was assessed prior to using distance sampling (DS) analysis methods (Buckland *et al.* 2001). b. we re-analysed the 2017-2019 model to include the data from A-core in 2021, following the same methodology as the previous step.

Task 2 (using data from A-core and A-outer) was conducted in three steps: a. data from 2021 survey was assessed prior to using DS method applied to A-core and A-outer areas b. then we re-analysed the 2017-2019 model to include the new sightings from 2021 for two areas (A-core and A-outer), following the same methodology. In both of these steps, A-core and A-outer are treated as separate areas so the density, abundance and biomass estimates are calculated separately for these two areas. c. As a final step, we applied model-based inference to all 2021 sightings to estimate density and abundance for A-core and A-outer, using the detection function obtained in the previous two steps.

2021 Survey design

In June - July 2021, surveys were conducted within A-core (as defined in Canadas and Vazquez (2020)) and A-outer (Figure 1). The A-outer is twice the size of the A-core (Table 1). Transects were aligned approximately north to south within the survey region although in 2021 a few transects were aligned in other directions. The coverage of transects in A-core was higher than in A-outer (Figure 1).

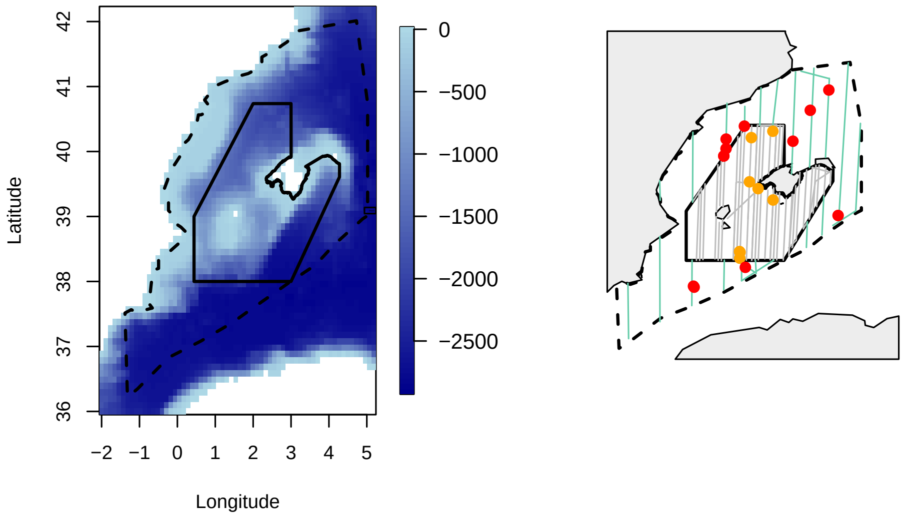


Figure 1. Right panel shows depth [m] within the study site. Right panel: Overview of the transects and sightings in Region A. Transects and sightings in A-core (solid line) are marked in grey and yellow, respectively. Transects and sightings in A-outer (dashed line) are marked in green and red, respectively.

Table 1. Areas [km^2] of A-core and A-outer in Region A.

Region	Area [km2]
A-core	61837
A-outer	123743

2021 Search protocol

Observers travelled along the transects on-board a plane; there were two types of observers on board: scientific and professional. Detailed description of the survey conducted in 2021 is given in Quevreus and Quiquempois (2021). On detecting BFT, the observers recorded the angle to the detection when plane was perpendicular to sighting, school size, age composition, biomass as well as other information. Environmental conditions were also recorded along each transect (e.g. Beaufort sea state, visibility). For the purpose of this report, only sightings of school consisting of adult individuals spotted by professional observers were taken into account (see Canadas and Vazquez (2020) for description of the distinction between adults and juveniles).

Statistical methods - Distance sampling (DS)

Line transect distance sampling (DS) analysis methods (Buckland *et al.* 2001), applied as described in Chudzinska *et al.* (2021), was used to estimate individual density and abundance.

Survey effort

Survey effort was calculated from the start and end locations of the effort when observers were searching (on-effort).

Perpendicular distance calculation

The perpendicular distances of detections to the trackline, x, were required to estimate the probability of detection. These were calculated following the approach of Canadas and Vazquez (2020) by analysing the GPS tracks of the plane to manually calculate the distance between schools and transects.
However, this proved difficult to implement (see Results) and so an alternative method was used: using the angle of declination when abeam, θ, and plane altitude at the time, A:

$$x = tan((90 - \theta) * \pi/180) * A$$

Probability of detection

Two critical assumptions of DS methods are that all schools on the transect centre line (i.e., at zero perpendicular distance) are detected with certainty and that distance measurements are exact. Given these assumptions, the distribution of perpendicular distances is used to model how the probability of detection decreases with increasing distance from the trackline.

The probability of detection, p, was estimated from a detection function model fitted to the observed distribution of perpendicular distances using the exact distances for fish/school. Perpendicular distances were right truncated, where required, to avoid a long tail in the detection function, as well as left truncated, where required, to account for lower detection on the transect centre line. Left truncation is a common practice for aerial surveys, due to difficulties in searching directly underneath the plane, especially when the plane does not have a bubble window, which was not the case in the studied survey years. Perpendicular distances were truncated at 1,500 m to avoid a long tail in the detection function. The choice of this truncation distance was based on visual inspection of fitted detection function, comparison with truncation distance used for previous years (2017-2019 models), results of Cramer-von Mises test and distribution of probabilities of detection for each model.

Two forms of the detection function were considered: a hazard rate and a half normal.

The effect of a range of covariates was incorporated into the detection function: year (as a factor), region (A-core or A-outer for Task 2 were treated as two separate regions), type of airplane, company conducting surveys (a factor with two levels: the company, Airmed, surveyed Region A in the years 2017-2019 and ActionAir in 2021), sea state, and various combinations of these listed covariates. As detectability is frequently a function not only of distance, but also school size (large schools are easier to see than small schools), then schools in the sample are likely to be larger than schools in the entire population. We therefore included school size (on a logarithmic scale) as a covariate for all models. The sizes of detected schools varied between 1 and 15,000 individuals of BFT between 2017-2021 (see Appendix B in Chudzinska *et al.* (2021)).

The effect of the above covariates was incorporated into the detection function model by setting the scale parameter in the model to be an exponential function of the covariates (Marques and Buckland 2004). Thus, the covariates could affect the rate at which detection probability decreases as a function of distance, but not the shape of the detection function. Adjustment terms were not included in this case.

The form that resulted in the smallest Akaike Information Criterion (AIC) was selected. Visual inspection of fitted functions, quantile-quantile plots, results of Cramer-von Mises test, estimated probability of detection and coefficient of variation were also taken into account (see Buckland *et al.* 2001 for details of detection function models and model selection methods).

Density, abundance and biomass

Detections and search effort were pooled within each survey to obtain encounter rates ($\frac{n}{L}$), and hence obtain estimates of density and abundance, by year (for 2017-2021 combined models). Estimates averaged overall surveys (weighted by survey effort) were also obtained. To estimate biomass, estimated biomass of the schools, instead of school size, was substituted in the final models used to estimate abundance and density.

Analyses were performed in R (R Core Team, 2019) using the Distance library (Miller *et al.* 2019b).

Statistical methods - model-based inference

The count method of Hedley and Buckland (2004) was implemented to model the trend in spatial distribution in BFT. A common approach is to model the number of individuals in a small section of effort as a function of location and environmental descriptors. However, due to the nature of these data, where the range of group size can be from few to thousands (1 - 3000 individuals in the data used in this model), a multi-step process was implemented:

- a model was fitted to the number of BFT groups ("group encounter rate model")
- a model was fitted to group size ("group size model")
- predictions from these two models were multiplied to produce a surface of BFT abundance.

For the purpose of this report, the same sightings, as used for DS, were considered: only sightings of schools consisting of adult individuals spotted by professional observers (for 2021 data).

Modelling the number of groups

The response variable was the estimated number of groups in a small section (segment) i of track line (of length l_i), $\hat{N}_i$, calculated using a Horvitz-Thompson-type estimator (Horvitz and Thompson, 1952) as follows:

$$\hat{N}_i = \sum_{r=1}^{R_i} \frac{1}{\hat{p}_r}$$

where R_i is the number of detected groups in segment i. The parameter $\hat{p}_r$ is the estimated probability of detection for group r in segment i; this was estimated using distance sampling (DS) methods (Buckland *et al.*, 2001) (see below for details).

The lengths of track lines were calculated from the recorded positions (i.e. latitude and longitude), when observers were on search effort. In addition, only groups sighted when observers were on search effort were included in the analysis. The target length of segments was 10 km but segments varied from this because of breaks in search effort.

Probability of detection

The probability of detection was estimated from the detection function modelled as described in *Statistical methods - Distance sampling (DS)* section.

Model specification

The estimated numbers of BFT groups per segment along the transect lines were used to estimate group abundance in the region of interest. This approach models spatial trend in the density and allows it to vary throughout the region of interest.

The number of groups in each segment (with known area) was estimated, $\hat{N}_i$ where i indicates an individual segment, and this formed the response variable in the statistical model.

Counts are often modelled using a Poisson distribution, however, these data were over dispersed (i.e. more variable than expected for Poisson distributed data), and, therefore, we assumed a Tweedie distribution, which allows more flexibility, for the counts.

The mean (μ_i) was modelled with location, habitat and temporal variables as candidate explanatory variables represented as follows.

$$\mu_i = \exp(\log_e(a_i) + \beta_0 + \sum_{j=1}^{J} \beta_j F_{ij} + \sum_{k=1}^{K} s_k(D_{ik}) + s_l(X_i, Y_i))$$

where

- $\log_e(a_i)$ is an offset term (a term with known regression coefficient) that corresponds to the area of each segment ($a_i = 2wl_i$ where w is the strip width and l_i is the length of each segment i),
- β_0 is an intercept,
- $\beta_j F_{ij}$ represent factor terms (e.g. year) with β_j representing the regression coefficients for the jth factor variable,
- $s_k(D_{ik})$ represent one dimensional smooth terms (e.g. depth)
- $s_l(x_i, y_i)$ represents a two-dimensional smooth term of location (determined for each segment i by X_i and Y_i).

The models were fitted using generalised additive models in the R package `dsm` (Miller *et al.*, 2013; Miller *et al.*, 2019a).

Candidate explanatory variables

The available candidate explanatory variables were:

- year (*year*, as a factor with 4 levels),
- location of the segment (x, y; as kilometers from a reference point to the midpoint of the segment), fitted as a 2-dimensional term (as shown in the equation above) and also with an interaction with *year*,
- depth in metres (*depth*), fitted as a 1-dimensional term,
- daily sea surface temperature (SST) for the day of the survey, fitted as a 1-dimensional term,
- difference between SST on the survey date and the SST 10 days prior to the survey date ($SSTd10$), fitted as a 1-dimensional term.

Longitude and latitude were transformed into a distance (in km) east (x) and north (y), respectively, from a reference point in the survey region (1.5^{o}W, 36^{o}N). This was to ensure that a unit change in the north-south direction was the same as a unit change in the east-west direction. Location is unlikely to determine BFT distribution but it acts as a proxy for other unmeasured or unknown variables that will determine BFT distribution.

Depth was obtained from ETOPO1, a 1-arc-minute global relief model (https://www.ngdc.noaa.gov/mgg/global/global.html).

SST for the survey day and 10 days prior was obtained from https://coastwatch.pfeg.noaa.gov/erddap/index.html (from the dataset with identifier "ncdcOisst21Agg_LonPM180") (Figures 2A and B).

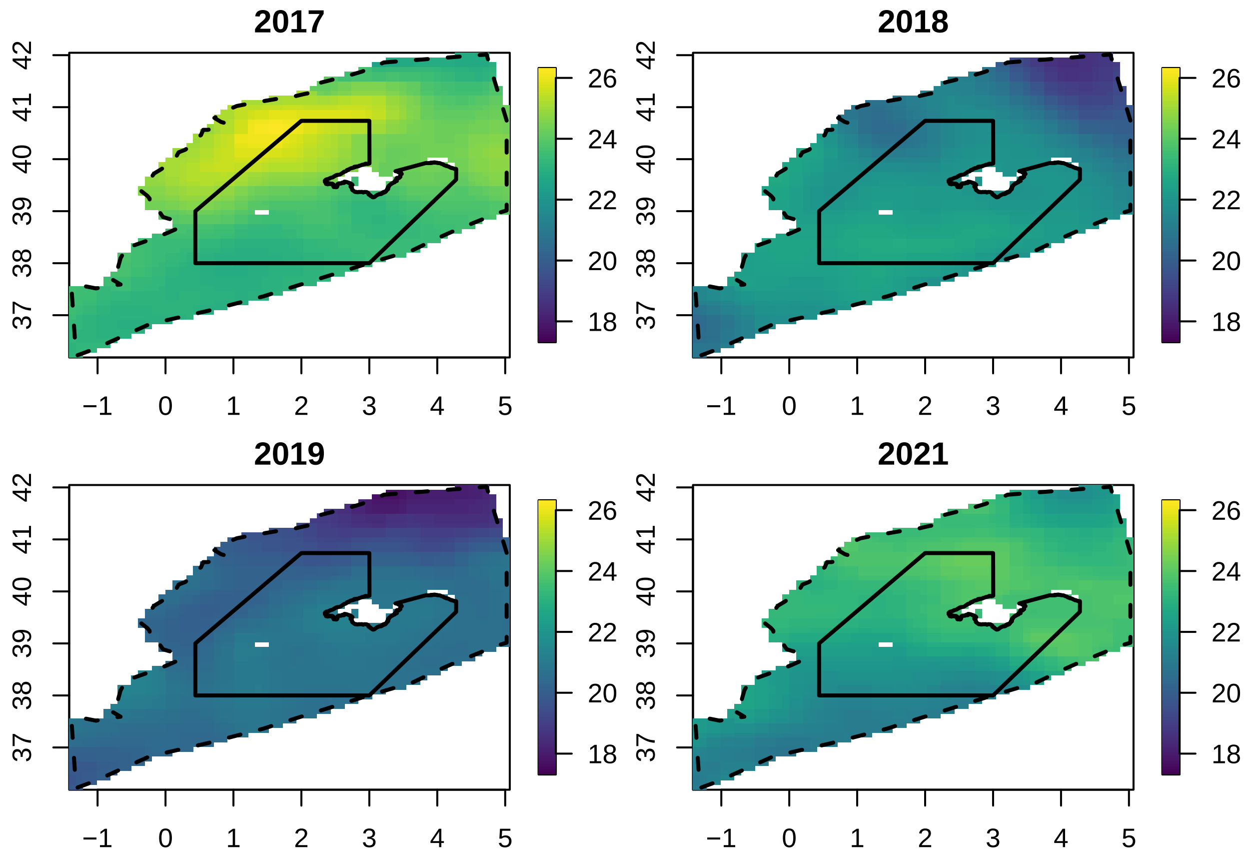


Figure 2A. Sea surface temperatures (^oC) on 15 June in 2017, 2018 and 2019, representing the approximate mid point of the surveys.

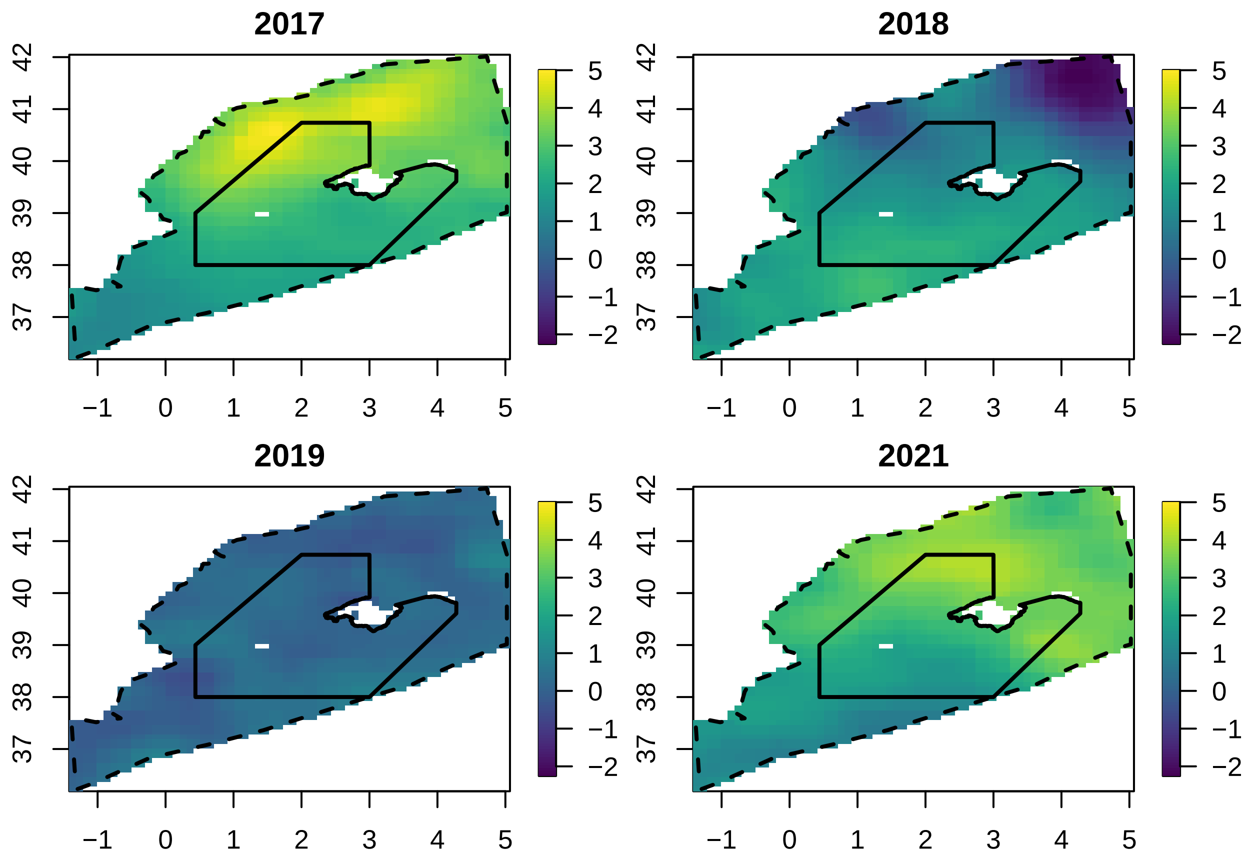


Figure 2B. Difference in sea surface temperatures (^{o}C) between 15 June and 10 days earlier (5 June)

The group encounter rate model included an interaction between the factor *year* and location, or *year* and $SSTd10$, such that smooths were calculated as a difference from a reference level (similar to regression coefficients for factors) (implemented using the option `bs="fs"`). If years were similar, this is a more parsimonious approach (i.e. fewer parameters estimated) than if separate smooths were calculated for each year.

One-dimensional terms were fitted with a modification to the smoothing penalty such that the term could effectively be excluded (i.e. the argument `bs="ts"` was included in the model specification) (Pedersen *et al.*, 2019). Having determined which of these candidate variables explained the most variation in the response, the model was examined and terms removed if necessary.

Term selection started by including all candidate variables in the model ("full model") and model reduction was based on effective degrees of freedom (edf), variables with edf < 1 were removed from the model ("shrinkage method"). Two full models were tested: one including location and year as interaction and one without any interactions. Model checking was based on quantile-quantile (q-q) plots, and looking at residuals. Comparison between final candidate models was based on AIC. If AIC values were similar, the simpler model was chosen. The restricted maximum likelihood smoothing parameter estimation method was used for all models.

The models were fitted in the R package `mgcv` (Wood, 2017) via the `dsm` package (Miller *et al.*; 2019a).

Modelling group sizes

In this step, only sightings data were included and the response variable in the generalized additive model (GAM) was recorded group size; this was modelled as a function of the candidate variables described above. Since a group of size zero is impossible, a potentially useful distribution to describe these data is the zero-truncated Poisson, possibly including over dispersion. However in practice, specifying these distributions created errors in the model fitting and so after some trial and error, the square root of the group sizes were modelled as a normal distribution (even though this is not theoretically correct).

Estimating density and abundance

Using the selected model, predicted number of groups was calculated for a grid of points (the prediction grid) from the encounter rate model, with associated area and known values for the explanatory variables. We chose 15th of June as prediction day (to obtain SST and $SSTd10$) because it was an approximate midpoint of the surveys over the four years (Table 1; Figures 2A and B shows SST on this date). Similarly, group size was predicted over the same grid using the group size model. The two surfaces were then multiplied together resulting in the estimated number of individuals per cell.

Total abundance for the Region was estimated by summing predicted number of individuals over all grid points in the region of interest (i.e. A-core and A-outer shown in Figure 1). An average estimated group size in the region of interest was obtained from the average of the predicted group sizes in the regions. Estimates were obtained A-core for all years and A-outer for 2021.

Uncertainty in the abundance estimate was obtained by combining uncertainty from GAM parameters, the detection function parameters and the group size model. We obtained this for the two GAM-based models via posterior simulation (Wood, 2017) by repeatedly sampling possible predictive surfaces from the encounter rate and group size models and taking their product for each cell, thus summary statistics over the samples (variances, etc.) encapsulate the model-based uncertainty from each model. In our final estimates of uncertainty, we also included uncertainty from the detection function via the delta method (summing squared coefficients of variation (CV)). This assumed independence between the spatial processes and the detection processes, but as there was little spatial pattern in group size, this seems to be justifiable. Confidence intervals (CI) for abundance were obtained using log-normal 95% CI (Buckland *et al.*, 2001).

Results

Survey effort, perpendicular distances and number of detections in 2021

In 2021, 9313 km were covered on search effort during 22 days of surveys. Due to different design of transect routes, the search effort was longer in A-core than in A-outer, even though A-outer was twice the size of A-core (Tables 1 and 2, Figure 1).

In 2021, 25 sightings of BFT were recorded by the professional observers: 10 within A-core, and 15 in A-outer. All 15 sightings in A-outer were on-effort and 8 (out of 10) in the A-core. All sightings within A-core and 12 sightings in A-outer were of schools consisting of adult individuals. Off-effort sightings and sightings of juvenile schools were removed from the analysis resulting in 8 and 12 sightings left for further analysis in A-core and A-outer, respectively (Figure 1, Table 2).

Perpendicular distances of only 4 (out of 25) sightings could be manually recalculated based on GPS track of the plane. The remaining sightings either had no GPS track provided, or the shape of the track was not clear enough to estimate the precise position of the schools. Therefore, perpendicular distances calculated on the angle of declination and plane altitude were used for the final analysis.

Most of the detections in 2021 in the A-outer were north from the core area of Region A (Figure 1), which is overlapping with estimated areas of high density of BFT as shown in Burt *et al.*, (2021).

Table 2. Search effort and number of sightings for each of the studied areas of Region A. The presented values apply to all surveys used in the final models but before right or left truncation.

Region	Effort [km]	Sightings
A-core	6716	8
A-outer	2697	12

Design-based approach

Analysis of 2021 data based on A-core only

Eight sightings were not enough for the design-based approach using only 2021 sightings from A-core and so this step was not pursued further.

Analysis of 2017-2021 data using data from A-core only

The same truncation distance (1,500 m) as for the 2017-2019 model (Chudzinska *et al.* (2021)) was used for 2017-2021 model resulting in inclusion of 7 (out of 8) sightings from A-core in 2021. No left truncation was needed for this data set.

A half-normal detection function including school size and type of airplane had the lowest AIC out of all models tested for 2017-2021 (Table 3). However, the half-normal detection function including school size and company was comparable in terms of AIC and CVs and had higher probability of detection. As the latter covariates were also used in the final model for the 2017-2019 data (Chudzinska *et al.* (2021)), this model was chosen as the final model for the 2017-2021 data (Table 3).

Table 3. List of the covariates used for a given model, key function, p values for Cramer-von Mises test, estimated average probability of detection (p), coefficient of variation (p.CV), and AICc for models fitted to 2017-2021 period using sightings in A-core only. Only models which converged are presented.

Model	Key function	p value of C-vM	p.CV	p	Delta AIC
log(School size) + Airplane	Half-normal	0.4290265	0.3179478	0.1211359	0.000000
log(School size) + Company	Half-normal	0.5044709	0.3143543	0.2268308	1.066812
log(School size) + Company	Hazard-rate	0.0134348	0.4097296	0.1141373	4.555392
log(School size) + Sea state	Half-normal	0.2049768	0.3423757	0.1103064	6.416082
log(School size)	Half-normal	0.2521504	0.3556266	0.1320423	7.205540
log(School size) + Year	Half-normal	0.2384650	0.3493375	0.1036864	10.021980
log(School size) + Sea state	Hazard-rate	0.0592313	0.3977068	0.1331438	14.223466
log(School size)	Hazard-rate	0.0455648	0.4209946	0.1169873	16.633228
log(School size) + Year	Hazard-rate	0.0296486	0.4239833	0.1067453	17.575857
No covariates	Hazard-rate	0.9055207	0.2518737	0.1223590	43.187830
No covariates	Half-normal	0.8281440	0.3234038	0.1263391	47.512301

The detection function of the final model is shown in Figure 3. The estimates abundance, density, and biomass for each year as well as uncertainties associated to these estimates are given in Table 4.

Table 4. Summary of results for period 2017-2021 for Region A (using sightings from A-core in all years for fitting the detection function): detection probability (p), Search effort (km), number of schools within truncation distance (n), encounter rate (ER, schools/km) and coefficient of variation (ER.CV), individual density (N-D, fish/km^2) and coefficient of variation (N-D.CV), individual abundance (N, in thousands), coefficient of variation (N.CV) and lower (N-LCL) and upper (N-UCL) limits of the 95% confidence interval for N, expected school size (N-ES),CV (N-ES.CV), biomass (B, tonnes), CV of B (B.CV) and lower (B-LCL) and upper (B-UCL) limits of the 95% confidence interval for B, biomass density (B-D, kg of fish/km^2), coefficient of variation (B-D.CV), and expected school biomass (B-ES, kg) and CV of B-ES (B-ES.CV).

Survey	p	Effort	n	ER	ER.CV	N-D	N-D.CV	N	N.CV	N-LCL
A-2017	0.17	4949.538	18	0.0036	0.30	0.82	0.44	50.79	0.44	22.22
A-2018	0.24	6092.870	24	0.0039	0.21	1.34	0.31	83.08	0.31	46.19
A-2019	0.23	5574.084	20	0.0036	0.24	1.23	0.38	76.30	0.38	37.37
A-2021	0.86	6715.943	7	0.0010	0.53	0.43	0.54	26.30	0.54	9.62

Survey	N-UCL	N-ES	N_ES.CV	B	B.CV	B-LCL	B-UCL	B-D	B-D.CV	B-ES	B-ES.CV
A-2017	116.10	118.56	0.56	8072.21	0.45	3467.16	18793.67	130.54	0.45	19.22	0.57
A-2018	149.43	245.48	0.39	13470.74	0.31	7427.84	24429.85	217.84	0.31	39.60	0.41
A-2019	155.77	234.42	0.42	11648.99	0.38	5670.44	23930.97	188.38	0.38	36.01	0.43
A-2021	71.92	1050.99	0.17	4716.59	0.53	1751.02	12704.74	76.27	0.53	210.92	0.16

Analysis of 2017-2021 data, including 2021 data from A-core and A-outer

The same truncation distance (1500 m) as for the 2017-2019 model (Chudzinska *et al.* (2021)) was used for 2017-2021 model (including sightings in A-core for all years and for A-outer in 2021). This resulted in inclusion of 7 (out of 12) sightings from A-outer in 2021 and 7 (out of 8) sightings from A-core (Table 6). No left truncation was needed for this data set.

Similarly as for model including sightings from A-core only, a half-normal model including school size and type of airplane had lowest AIC out of all models tested for 2017-2021 (Table 5). However, half-normal model including school size and company was comparable in terms of AIC and CVs and higher probability of detection (Table 5). As the latter covariates were also used in the final model for 2017-2019 data (Chudzinska *et al.* (2021)) and final model based on sightings from A-core only (Table 3), this model was chosen as the final model for 2017-2021 (Table 5). The detection function of the final model is shown in Figure 3.The estimates abundance, density, and biomass for each year as well as uncertainties associated to these estimates are given in Table 6.

Table 5. List of the covariates used for a given model, key function, p values for Cramer-von Mises test, estimated average probability of detection (p), coefficient of variation (p.CV), and AICc for models fitted to 2017-2021 period (sightings from A-core in all years and in 2021 from A-outer). Only models which converged are presented.

Model	Key function	p value of C-vM	p.CV	p	Delta AIC
log(School size) + Airplane	Half-normal	0.4433097	0.3292302	0.1199375	0.000000
log(School size) + Company	Half-normal	0.5373894	0.3255147	0.2177224	1.113029
log(School size) + Airplane	Hazard-rate	0.0147119	0.4290804	0.1144207	3.038070
log(School size) + Company	Hazard-rate	0.0182285	0.4238127	0.1056962	4.489117
log(School size) + Region	Hazard-rate	0.4216316	0.3776141	0.1302215	6.694598
log(School size) + Sea state	Half-normal	0.2416946	0.3537154	0.1020325	7.008126
log(School size)	Half-normal	0.2519820	0.3677845	0.1312578	8.043231
log(School size) + Region	Half-normal	0.8986057	0.3063189	0.1008225	9.068470
log(School size) + Sea state	Hazard-rate	0.0080574	0.4288004	0.1160319	11.486259
log(School size) + Year	Half-normal	0.2504814	0.3666013	0.1081919	13.780603
log(School size)	Hazard-rate	0.0170951	0.4469406	0.1226102	16.810349
log(School size) + Year	Hazard-rate	0.0245816	0.4420553	0.1177722	21.861491
No covariates	Hazard-rate	0.8569841	0.2709736	0.1256065	47.130685
No covariates	Half-normal	0.8025839	0.3369596	0.1507420	52.259059

Table 6. Summary of results for period 2017-2021 for region A (including sightings in A-core from all years and A-outer in 2021 for fitting the detection function): detection probability (p), Search effort (km), number of schools within truncation distance (n), encounter rate (ER, schools/km) and coefficient of variation (ER.CV), individual density (N-D, fish/km^2) and coefficient of variation (N-D.CV), individual abundance (N, in thousands), coefficient of variation (N.CV) and lower (N-LCL) and upper (N-UCL) limits of the 95% confidence interval for N, expected school size (N-ES),CV (N-ES.CV), biomass (B, tonnes), CV of B (B.CV) and lower (B-LCL) and upper (B-UCL) limits of the 95% confidence interval for B, biomass density (B-D, kg of fish/km^2), coefficient of variation (B-D.CV), and expected school biomass (B-ES, kg) and CV of B-ES (B-ES.CV).

Survey	p	Effort	n	ER	ER.CV	N-D	N-D.CV	N	N.CV	N-LCL
A-2017	0.17	4949.538	18	0.0036	0.30	0.82	0.44	50.44	0.44	22.06
A-2018	0.24	6092.870	24	0.0039	0.21	1.33	0.31	82.49	0.31	45.83
A-2019	0.23	5574.084	20	0.0036	0.24	1.23	0.38	75.79	0.38	37.11
A-2021-core	0.87	6715.943	7	0.0010	0.53	0.42	0.54	26.11	0.54	9.59
A-2021-outer	0.85	2697.090	7	0.0026	0.51	0.65	0.58	80.99	0.58	26.86

Survey	N-UCL	N-ES	N_ES.CV	B	B.CV	B-LCL	B-UCL	B-D	B-D.CV	B-ES	B-ES.CV
A-2017	115.33	117.24	0.56	8058.97	0.45	3461.35	18763.49	130.33	0.45	19.17	0.57
A-2018	148.48	243.01	0.40	13447.28	0.31	7414.02	24390.18	217.46	0.31	39.50	0.41
A-2019	154.78	232.23	0.42	11630.23	0.38	5661.00	23893.69	188.08	0.38	35.94	0.43
A-2021-core	71.13	1052.41	0.17	4714.40	0.53	1750.48	12696.82	76.24	0.53	210.95	0.16
A-2021-outer	244.17	641.40	0.13	14039.41	0.51	5166.12	38153.40	113.46	0.51	125.03	0.19

Comparison of density, abundance and biomass between models based on A-core only and models including A-outer

The detection functions for the two data sets, 2017-2021 A-core only and 2017-2021 A-core and A-outer, are shown in Figure 3. The estimates of abundance, biomass, density and a range of uncertainty measures are shown in Tables 5 and 6.

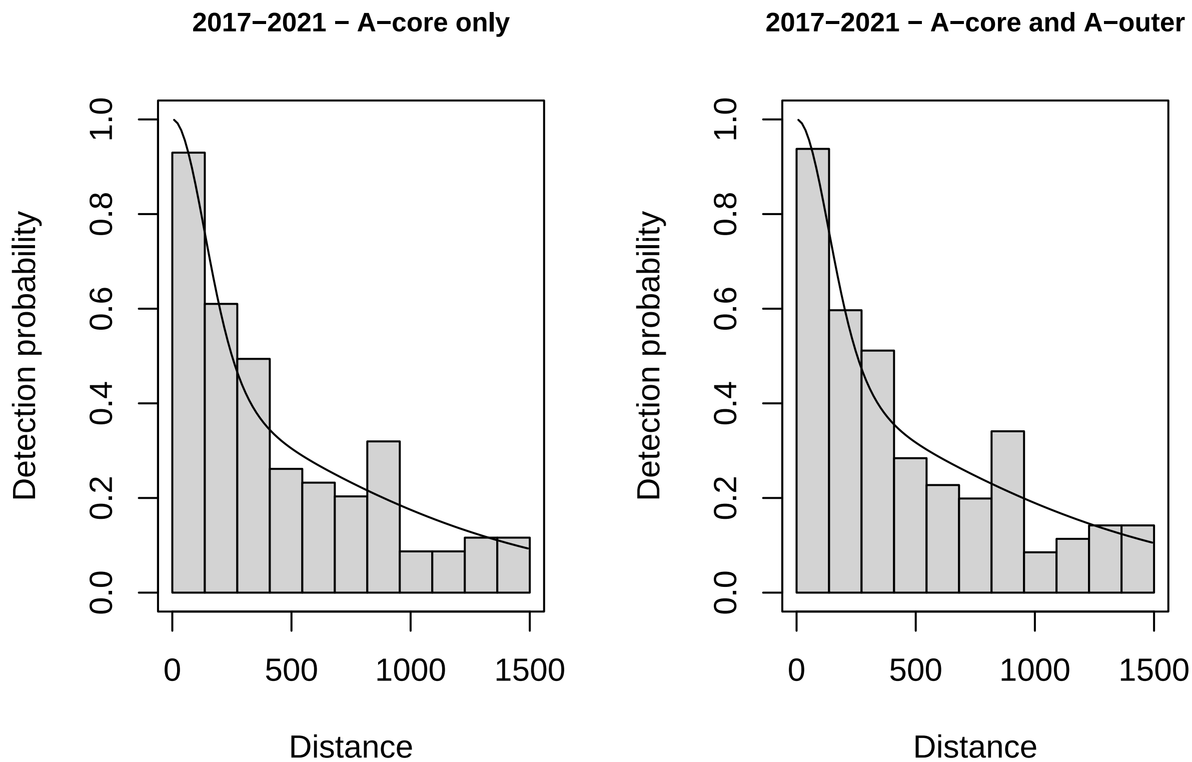


Figure 3. Average estimated detection function (black line) for 2017-2021 which included sightings from 2017-2021 in A-core only (left panel) and 2017-2021 in A-core plus A-outer in 2021 (right panel).

Both new models (2017-2021 from A-core only and 2017-2021 including 2021 sightings from A-core and A-outer) show the same estimates of abundance, biomass, density and biomass density for 2017-2019 period as the model for this period shown in Chudzinska *et al.* (2021) (Figures 4-7). As A-outer is twice the size of A-core, estimated abundance and biomass is higher in the A-outer and cannot be directly compared (Figures 4 and 5). Fish density is, however, comparable between these two areas (Figure 6). Biomass density is slightly higher in A-outer than in the A-core.

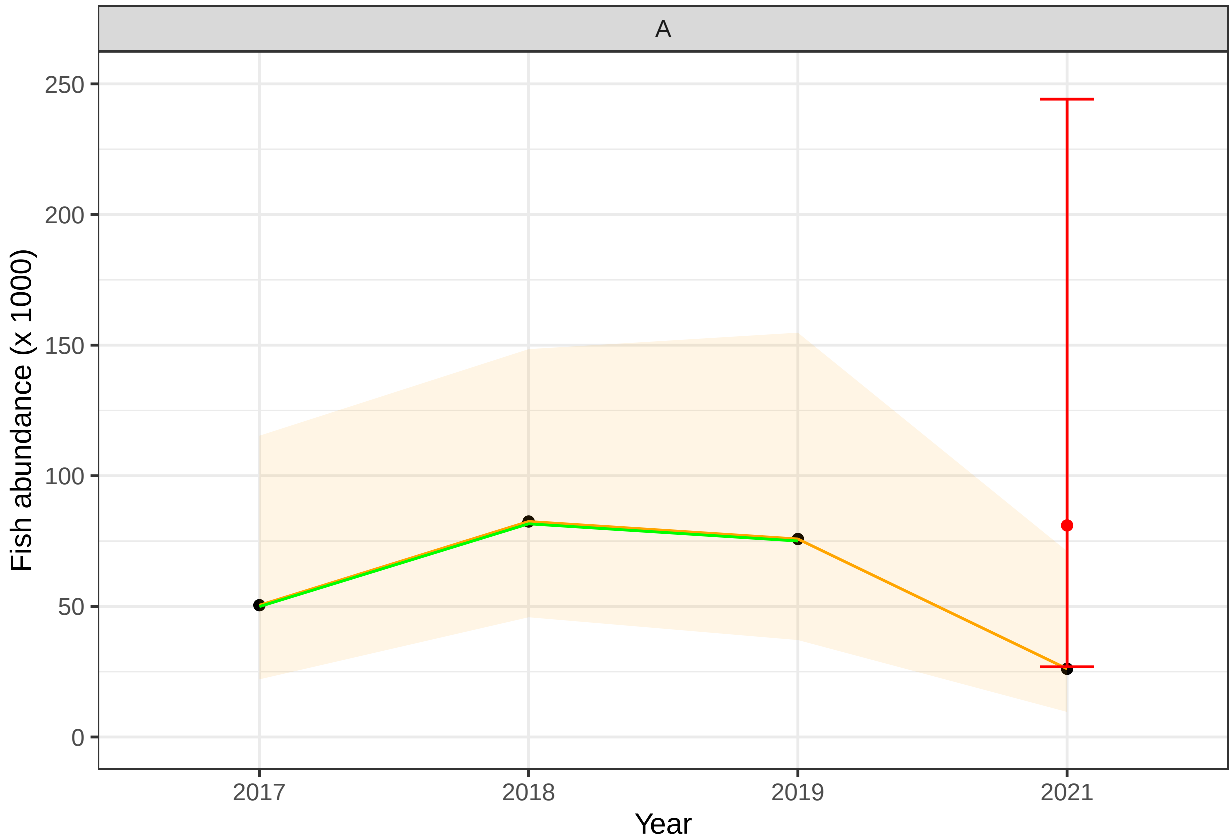


Figure 4. Estimated abundance of BFT for surveyed years in A-core in years 2017-2021. Orange ribbon shows upper and lower confidence limits of the 95% confidence interval for the 2017-2021 models. For comparison, estimated abundance based on 2017-2019 data only (Chudzinska *et al.*, (2021)) are shown in green. For year 2021 abundance for A-outer is plotted in red with point being mean estimate and error bars showing confidence intervals. Note that the area of A-core and A-outer differs and, therefore, the abundance estimates are not directly comparable.

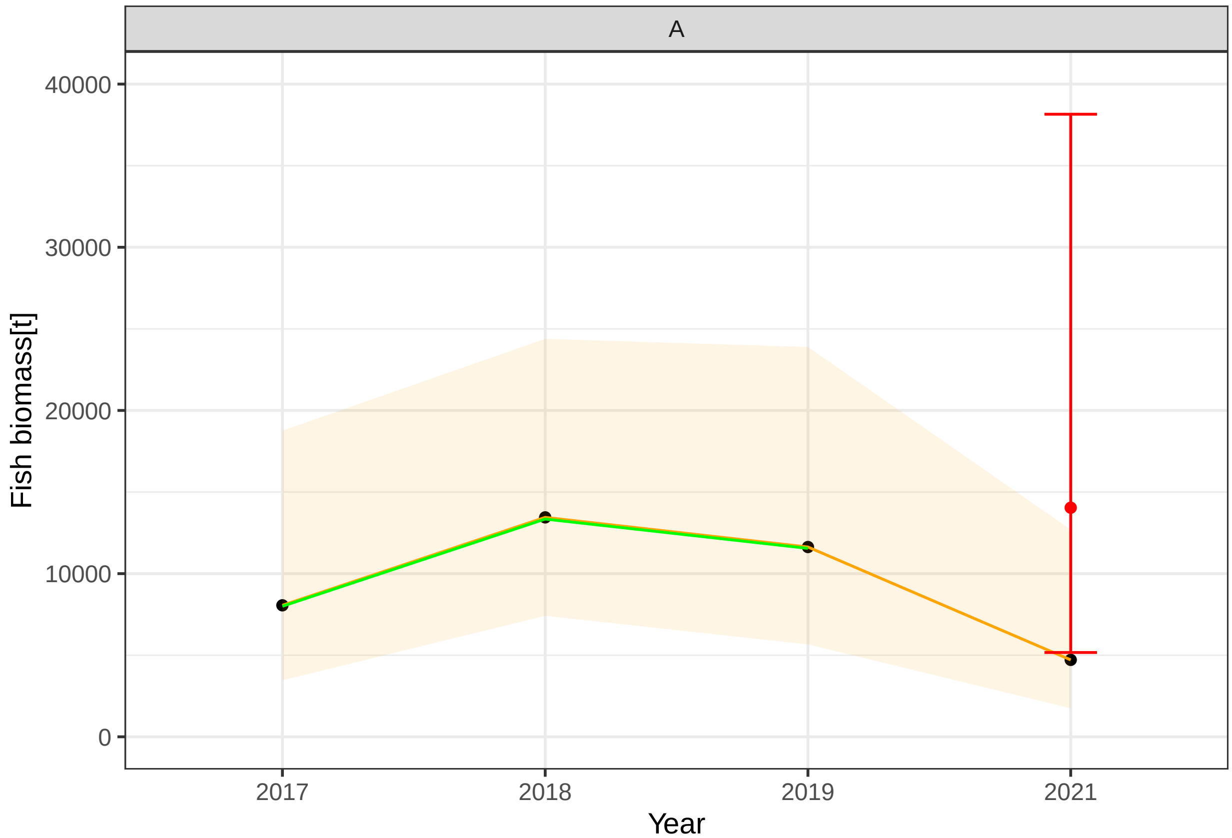


Figure 5. Estimated biomass (t) of BFT for surveyed years in Region A-core in years 2017-2021. Orange ribbon show upper and lower confidence limits of the 95% confidence interval for the 2017-2021 models. For comparison, estimated biomass based on 2017-2019 data only (Chudzinska *et al.* (2021)) are also shown in green. For year 2021 biomass for A-outer is plotted in red with point being mean estimate and error bars showing confidence intervals. Note that the area of A-core and A-outer differs and the biomass estimates are, therefore, not directly comparable.

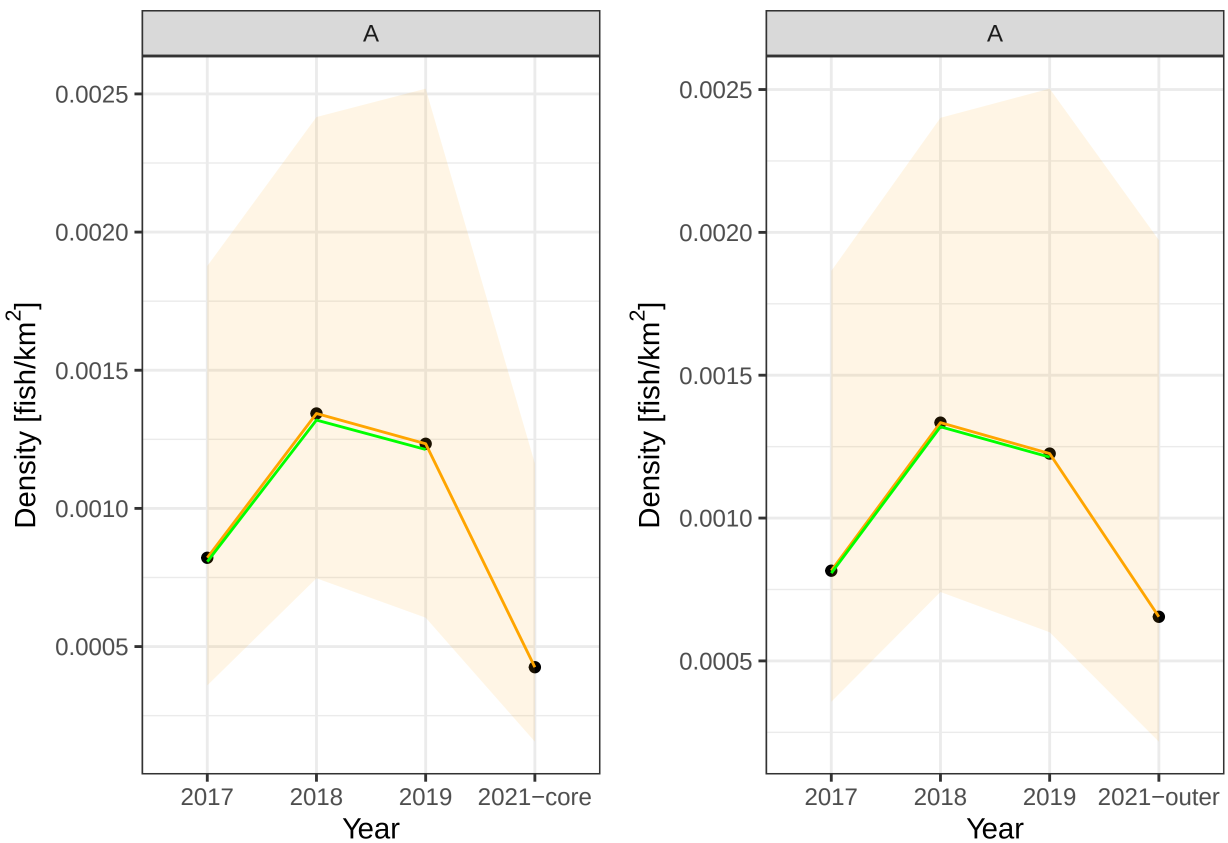


Figure 6. Estimated density (fish/km^2) of BFT for surveyed years in Region A in years 2017-2021. The left panel shows density estimates for A-core in Region A and the right panel A-outer in Region A. Orange ribbon show upper and lower confidence limits of the 95% confidence interval for the 2017-2021 models. For comparison, estimated density based on 2017-2019 data only (Chudzinska *et al.* (2021)) are shown in green.

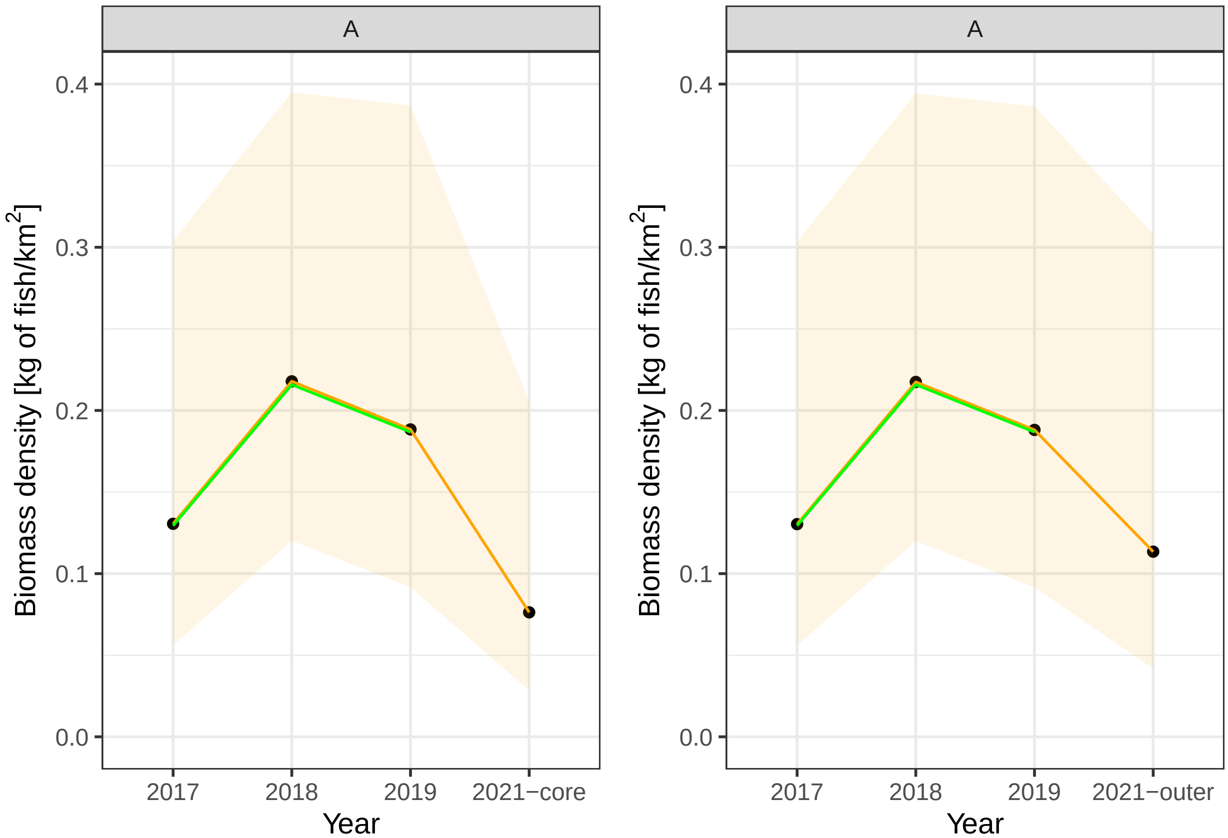


Figure 7. Estimated biomass density (kg of fish/km^2) of BFT for surveyed years in Region A in years 2017-2021. The left panel shows biomass density estimates for A-core in Region A and the right panel A-outer in Region A. Orange ribbon shows upper and lower confidence limits of the 95% confidence interval for the 2017-2021 models. For comparison, estimated biomass density based on 2017-2019 data only (Chudzinska *et al.* (2021)) are shown in green.

Model-based approach

As the same sightings were considered for the model-based approach as for the DS approach, there were not enough detections to use data from the 2021 survey only. Therefore, data from the 2017-2019 surveys in Region A were included. Hence, model based inference was, based on all sightings from 2017-2021 in A-core and from 2021 in A-outer.

Search effort and numbers of groups detected

During the four years of surveys in Region A, a total of 2.5998×10^4 km of search effort were flown and 89 groups of BFT were detected (Table 7). Two sightings from 2021 were excluded from analysis compared to design-based approach as they were detected just after search effort ended on the transect. The majority of groups were detected in the shallower waters in the region (Figure 8).

Group sizes ranged from 1 to 5000 taking into account sightings from all years (i.e. without truncation) (see below).

Table 7. Summary of survey effort and the number of groups detected and mean group size. The groups have not been truncated.

Year	Effort	Number.Groups	Mean.Group.Size
2017	4982	22	754
2018	6147	29	928
2019	5460	20	694
2021	9409	18	867

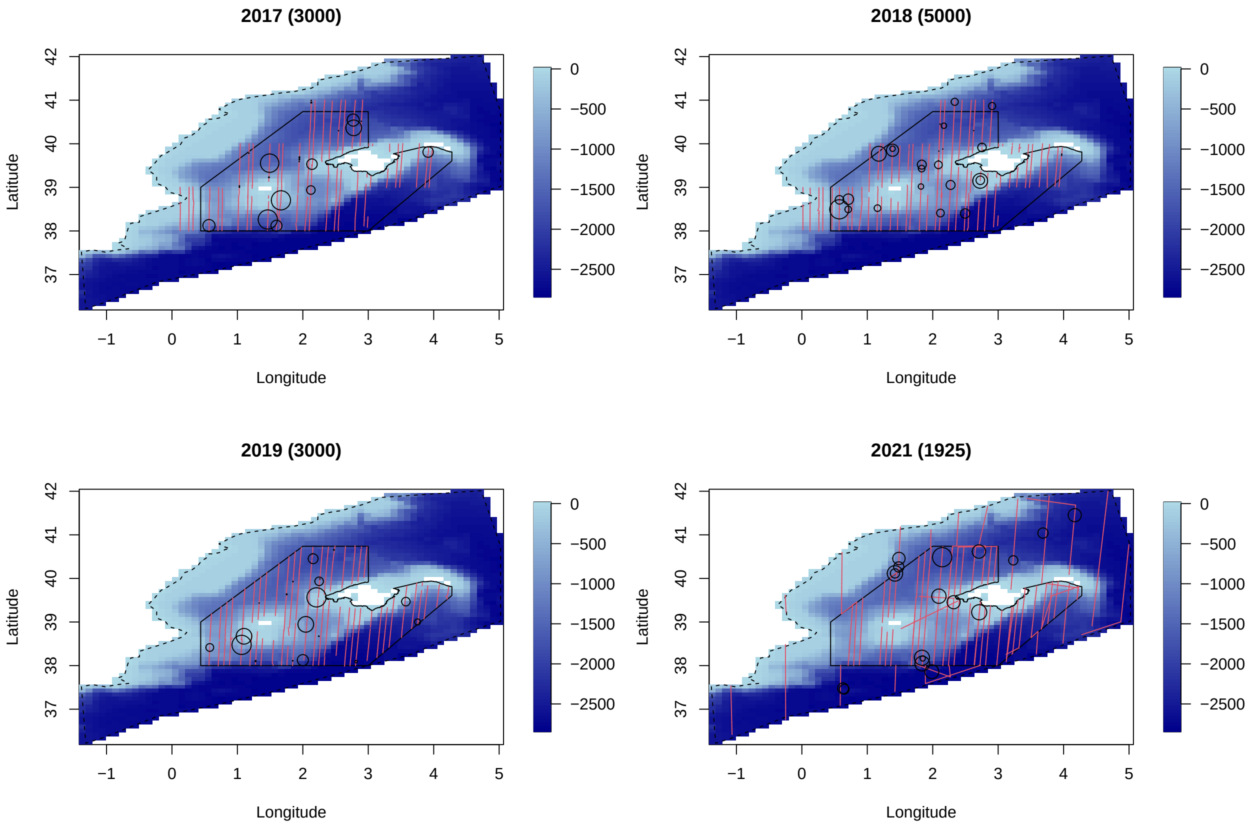


Figure 8. Location of search effort (red lines) and (truncated) sighted groups (circles) overlaid on depth (metres). The area of the circle is proportional to the size of the sighted group; maximum group size given in parentheses.

Model selection

The data consisted of 2838 segments and the average length was 9.2 km (range 0.18 - 14.9 km); 82 segments (3%) contained sightings. A few, very small segments (<0.15km) were excluded.

Given the time of year of the surveys, *SST* was slightly higher than the temperature 10 days prior to the survey day and differences were greater at higher temperatures (Figure 9).

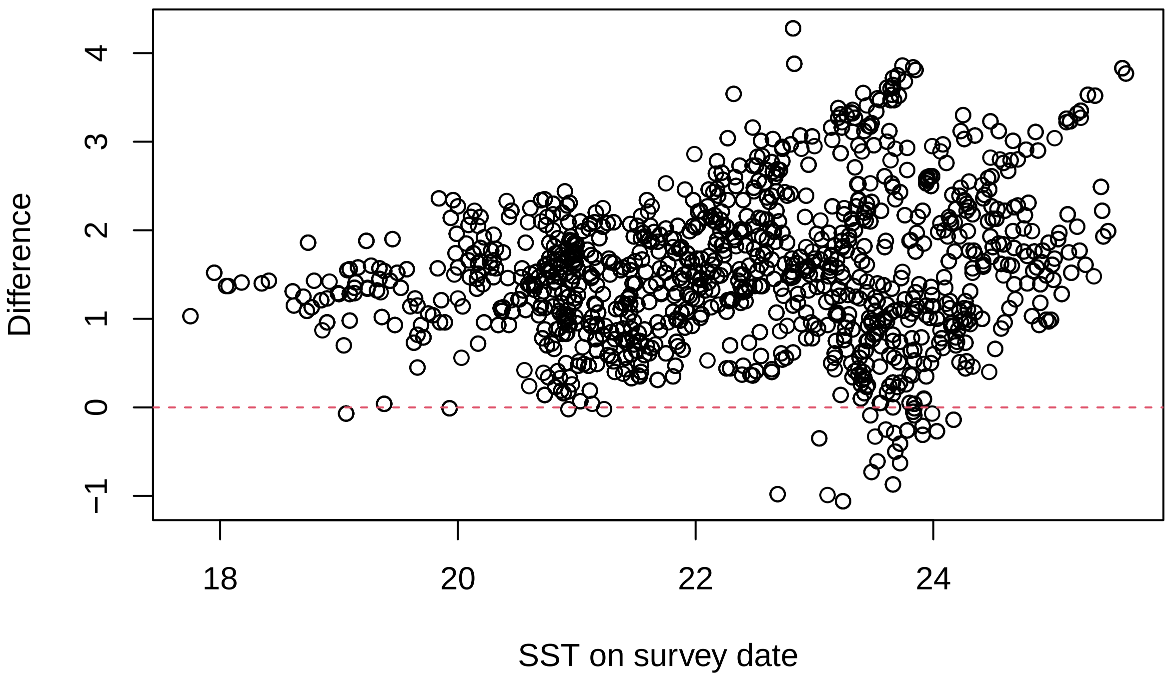


Figure 9. Difference (^{o}C) between SST and the temperature 10 days prior to the survey day. The red dashed line indicates no difference.

The variables included in both types of models (with group sizes and estimated number of groups as responses) and the deviances explained are shown in Table 8. See Appendix A for model diagnostics of the selected models (selected models are marked in bold capitals in Table 8).

Table 8. Summary of the fitted models. SST: sea surface temperature on a day of survey, SSTd10: difference between temperature on a day of survey and 10 days before, YearF: year of survey as a factor. The numbers by each term indicate effective degrees of freedom. Models marked in capitals are the final models used to estimate density and abundance.

Number	Response	Description	Smooth terms	AIC
1	Group size	‘All terms with interaction‘	s(x,y,YearF, 4.2), s(SST, 0.8), s(depth, 0.47), s(SSTd10, 1.5e-05)	633.84
2	Group size	‘All terms without interaction‘	s(x,y, 1), s(SST, 0.83), s(depth, 0.6), s(SSTd10, 0.00014)	631.39
3	GROUP SIZE	‘SST only‘	s(x,y, 1), s(SST, 0.83)	631.88
1	Group encounter	‘All terms with interaction‘	s(x,y,YearF, 11), s(SST, 0.00051), s(depth, 0.00034), s(SSTd10, 0.8)	978.95
2	Group encounter	‘All terms without interaction‘	s(x,y, 1.7), s(SST, 0.00039), s(depth, 0.00031), s(SSTd10, 0.68)	976.92
3	GROUP ENCOUNTER	‘Year-SST dif interaction‘	s(x,y, 1.7), s(SSTd10,YearF, 3.8)	978.98

Estimated density and abundance

The estimated abundances in the survey region for 10 June 2017, 2018 and 2019 are shown in Table 9.

Table 9. Model-based estimates of BFT in A-core and A-outer: number of groups and CV (CV.Group), average group size and CV (CV.Group.Size), abundance of individuals and CV and 95% CI. Note that in 2017-2019 only A-core was surveyed.

Year	Region	Groups	CV.Group	Group.Size	CV.Group.Size	Abundance	CV	Lower.CI	Upper.CI
2017	A	456	0.55	220	0.46	110281	0.78	28516	426498
2018	A	334	0.33	380	0.35	120016	0.48	49360	291813
2019	A	268	0.40	285	0.38	71985	0.56	25712	201532
2021	A-core	52	0.51	855	0.27	45605	0.55	16749	124181
2021	A-outer	145	0.55	840	0.27	114275	0.57	40658	321190

Estimated density for each year is shown in Figure 10 for A-core and A-outer but estimates for 2017-2019 in Table 9 relate only to A-core. Higher densities were estimated north of the region. The CVs of the estimates are shown in Figure 11; not surprisingly, the uncertainty associated with A-outer for 2017-2019 is high.

See Appendix B for further details of the estimated number of groups and estimated group sizes.

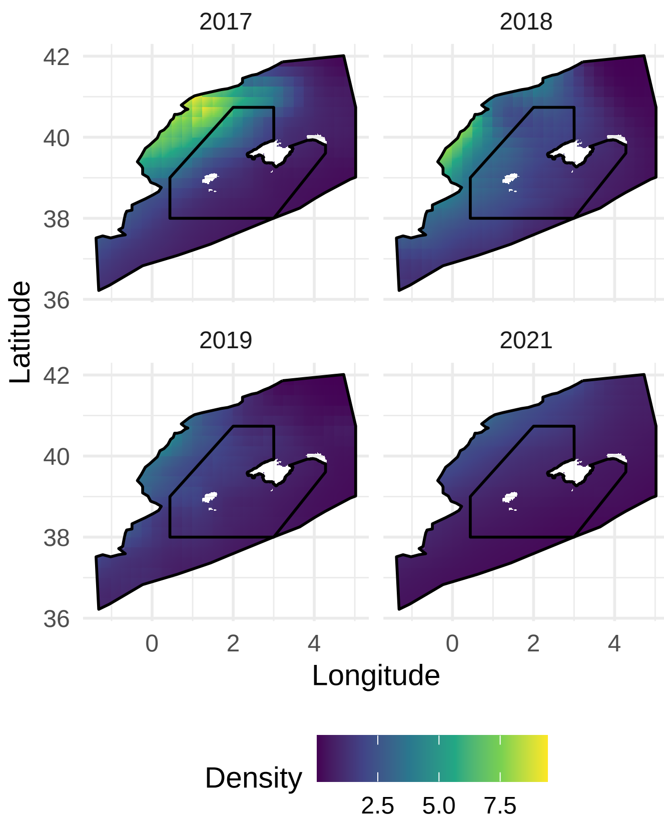


Figure 10. Estimated density of BFT (fish/km^2) for 15th June. The boundary of A-core is shown by the inner solid black line.

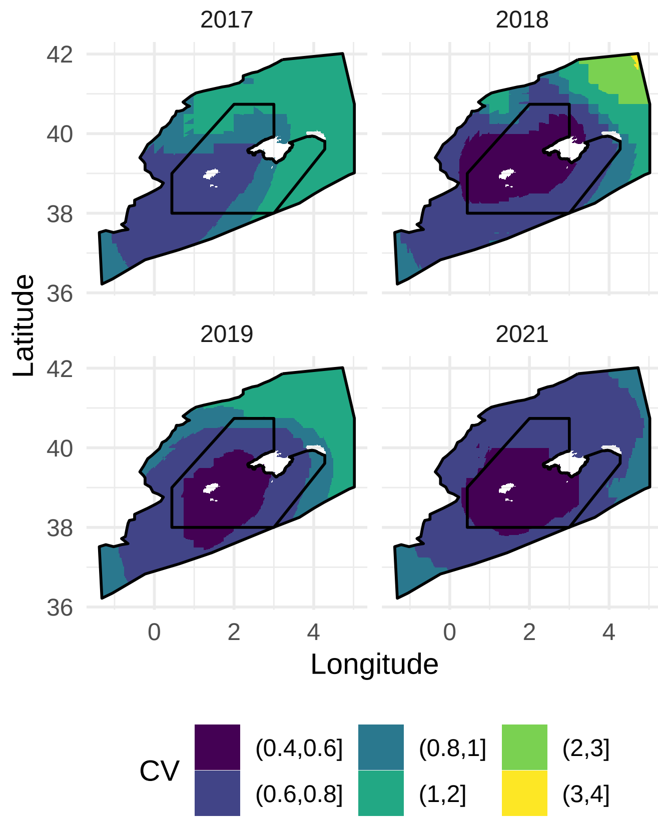


Figure 11. Coefficient of variation associated with density of BFT (not including detection function uncertainty). The boundary of A-core is shown by the inner solid black line.

Comparison of abundance estimates between design- and model-based approach are shown in Figure 12. The model-based estimates are within the 95% CI of the design-based estimates, except estimates for 2021 in the core area.

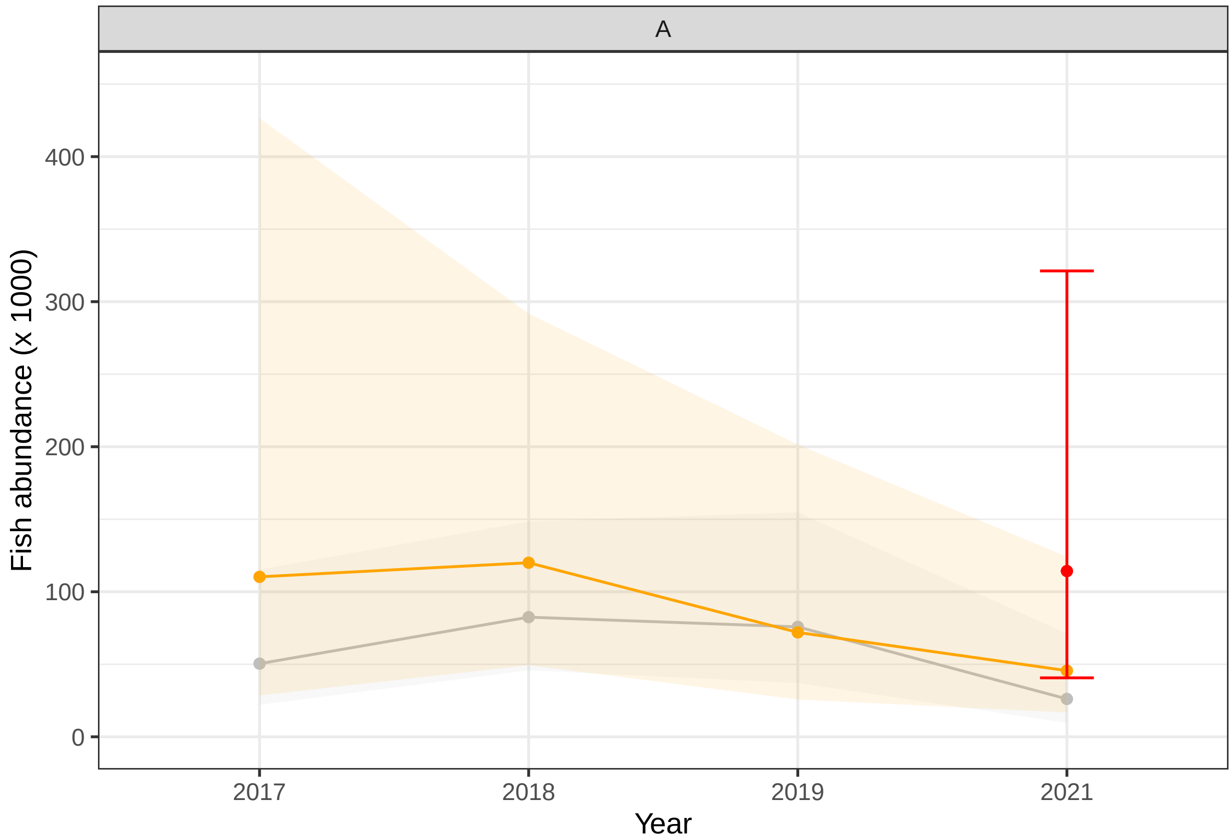


Figure 12. Comparison of model-based (orange) and design-based (grey) abundance estimates of BFT for surveyed years in Region A-core in years 2017-2021. Ribbon shows the upper and lower confidence limits of the 95% confidence interval for the 2017-2021 models. For year 2021 abundance estimates from model-based approach for A-outer is plotted in red with point being mean estimate and error bars confidence limits. Note that the area of A-core and A-outer differs and the abundance estimates are directly not comparable.

Discussion

Comparison between surveys in 2021 and previous years

The probability of detection (p, Table 6) was much higher in 2021 (regardless whether it was A-core or A-outer) than in previous years. This trend in higher probability of detection was also present when the detection function was fitted without any covariates (results available on demand). The most likely explanation for this trend is the fact that in 2021, schools were observed almost uniformly within the truncation distance (Appendix C), whereas in previous years, most schools were detected closer to the survey line. Further investigation is needed in order to understand whether these differences are driven by different way of calculating perpendicular distances between years, different practices of the survey companies, experience of the observers or fish behaviour.

Surveys in 2021 started a week later than surveys in 2017-2019. The encounter rate was much lower in 2021 (regardless of whether it was A-core or A-outer) than in previous years, which is reflected in decreasing trend in the abundance of BFT in Region A-core.

It is also worth noting that the minimum school size in 2021 was higher than in the previous years. In the previous years, numerous schools were of size = 1 fish, whereas in 2021 the minimum school size was 150 fish. These differences are reflected in larger estimated group sizes and reduced encounter rate in 2021 (regardless of whether it was A-core or A-outer) than in previous years (Table 6). Whether this is a general trend in the school behaviour of BFT or is related to a different protocol of data collection, variable experience of the observers, or some other reason needs further investigation. One possible option would be to truncate small groups from previous years to reduce the differences between 2021 and previous years.

Comparison between A-core and A-outer in 2021

The area A-core is half the size of A-outer, but had 2.5 times greater effort. Bearing in mind that the number of sightings was low in general, the encounter rate was larger in A-outer than in A-core (Table 6), which resulted in larger abundance and density estimates in A-outer. Larger schools were detected in A-core compared to A-outer. Further investigation is needed to understand what drives the differences in the behaviour of fish in these two areas.

Design- and model-based approach

The number of sightings in 2021 was too low to conduct any analysis based solely on 2021. Data from previous surveys in 2017-2019 were therefore included. A-outer in Region A was only surveyed in 2021 (out of 2017-2021 surveys). Whereas comparison between abundance, density and biomass estimates between the core area of Region A between years is straightforward, comparing the estimates in A-core and A-outer is more difficult to interpret due to differences in location, extent and coverage areas. Due to only a few additional sightings in 2021, when combining all survey years including 2021, the estimates for the previous years in A-core should be comparable to the estimates based solely on 2017-2019, and this is the case in this report.

As in Burt *et al.* (2021), the presented report is a demonstration of using model-based inference to estimate density and abundance of BFT. We limited the data to one Region and used a limited selection of potential explanatory variables. The distribution of BFT is, most likely, affected by a range of other environmental variables (e.g. distance to frontal systems, salinity) which can easily be included in the models in the future. The need to consider other variables is reflected in relatively low deviance explained by the models developed here (Table 8). (However, information on variables throughout the region of interest would be required for prediction.) Due to low number of sightings in 2021, the model including these sightings did little to improve the estimates based on only 2017-2019 sightings. The model-based estimates followed a similar pattern as the design-based estimates but were higher than the design-based estimates. Prediction in model-based methods can be affected by using values outside the range of the data on which the model was fitted but this was not the case here; predictions were only obtained for A-core for 2017-2019 (also see Appendix D). The higher abundance estimates are likely to mix of either, or both, the group encounter rate model and the group size model predicting higher values than the design-based model.

The design-based approach indicates lower abundance, density, biomass and density of biomass in 2021 than in previous years if only the A-core is taken into account. The area A-outer has comparable abundance and biomass estimated as in A-core in 2019. Given that the A-outer is twice the size of the A-core, the density of BFT also shows decrease from the previous years. The model-based approach indicates an increase in abundance in comparison to 2019. However, the uncertainties are much higher in the model-based approach and the latter estimates should be interpreted with caution. The discrepancies could also be related to the differences in school sizes between 2021 and previous years as explained above.

It also worth noting that in this report, SST was obtained from NOAA, not Copernicus as in 2017-2019 analysis (Burt *et al.* (2021)). While these look to be comparable, differences have not been studied in detail. Predictions were also done 5 days later than in Burt *et al.* (2021), to accommodate the later timing of 2021 survey.

Conclusions

- The estimates of abundance, density of biomass of BFT in the Region A show decreasing trend in comparison to previous years.

- Surveying A-outer should be considered in the future as relying solely on A-core area may result in underestimation of the abundance, density and biomass of the BFT.
- In the model-based approach, data were challenging to model due to the low number of sightings and the large variation in observed group sizes. To alleviate these problems, a multi-step modelling approach was implemented; a model was fitted to the estimated number of groups and to group sizes. Predicted values from these two models were multiplied to estimate number of individuals. Other possible approaches may consider fitting presence/absence of groups and then combining with estimated number of groups and group size.
- A limited selection of explanatory variables was used in this model-based approach but this method has the potential to include more to understand the drivers of BFT density.
- Including data from other survey Regions (C, E and G) and years would increase the number of sightings, however, the analysis is still likely to be challenging due to various issues related to data collection in the other regions, such as temporal mismatch between survey and spawning season, unstandardised survey protocols between years and regions.

Acknowledgements

This work has been carried out under the ICCAT Atlantic-Wide Research Programme for Bluefin Tuna (GBYP), which is funded by the European Union, several ICCAT CPCs, the ICCAT Secretariat, and other entities (see https://www.iccat.int/gbyp/en/overview.asp). The content of this paper does not necessarily reflect ICCAT's point of view or that of any of the other sponsors, who carry no responsibility. In addition, it does not indicate the Commission's future policy in this area.

References

Bouchet PJ, Miller DL, Roberts JJ, Mannocci L, Harris CM and Thomas L (2020) Extrapolation tools for density surface models. Methods in Ecology and Evolution. DOI: 10.1111/2041-210X.13469

Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL and Thomas L (2001) Introduction to distance sampling. Oxford University Press

Burt L, Chudzinska M, Borchers D, Miller D and Buckland S (2021) Model-based inference to estimate density and abundance of bluefin tuna. Report number CREEM-2021-04. Provided to ICCAT, August 2021 (Unpublished)

Canadas A and Vazquez A (2020) 2010-2019 Aerial Survey Data Re-Analysis. SHORT-TERM CONTRACT ICCAT GBYP 07/2019 Final report

Chudzinska M, Burt L, Borchers D, Miller D and Buckland S (2021) Design-based inference to estimate density, abundance and biomass of bluefin tuna. Reanalysis of 2010-2019 Aerial Surveys. Report number CREEM-2021-03. Provided to ICCAT, September 2021 (Unpublished).

Hedley SL and Buckland ST (2004) Spatial models for line transect sampling. Journal of Agricultural, Biological and Environmental Statistics 9:181-199

Horvitz D and Thompson D (1952) A Generalization of Sampling Without Replacement from a Finite Universe. Journal of the American Statistical Association, 47

Marques FFC and Buckland ST (2004) Covariate models for the detection function. In Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL and Thomas L (eds) Advanced distance sampling. Oxford University Press, Oxford, UK

Miller DL, Burt L, Rexstad E and Thomas L (2013) Spatial models for distance sampling data: Recent developments and future directions. Methods in Ecology and Evolution, 4:1001-1010

Miller DL, Rexstad E, Burt L, Bravington MV and Hedley S (2019a). dsm: Density Surface Modelling of Distance Sampling Data. R package version 2.2.18. http://github.com/DistanceDevelopment/dsm

Miller DL, Rexstad E, Thomas L, Marshall L and Laake JL (2019b) Distance Sampling in R. Journal of Statistical Software, 89(1), 1-28. doi:10.18637/jss.v089.i01 (URL: https://doi.org/10.18637/jss.v089.i01)

Pedersen E, Miller DL, Simpson GL and Ross N (2019) Hierarchical generalized additive models in ecology: an introduction with mgcv. PeerJ 7:e6876

Quevreus H and Quiquempois V (2021) Prospection aerienne des reproducteurs de thons rouges (*Thunnus thynnus*) Zone A - Mer des Baleares. Provided to ICCAT, July 2021 (Unpublished)

R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

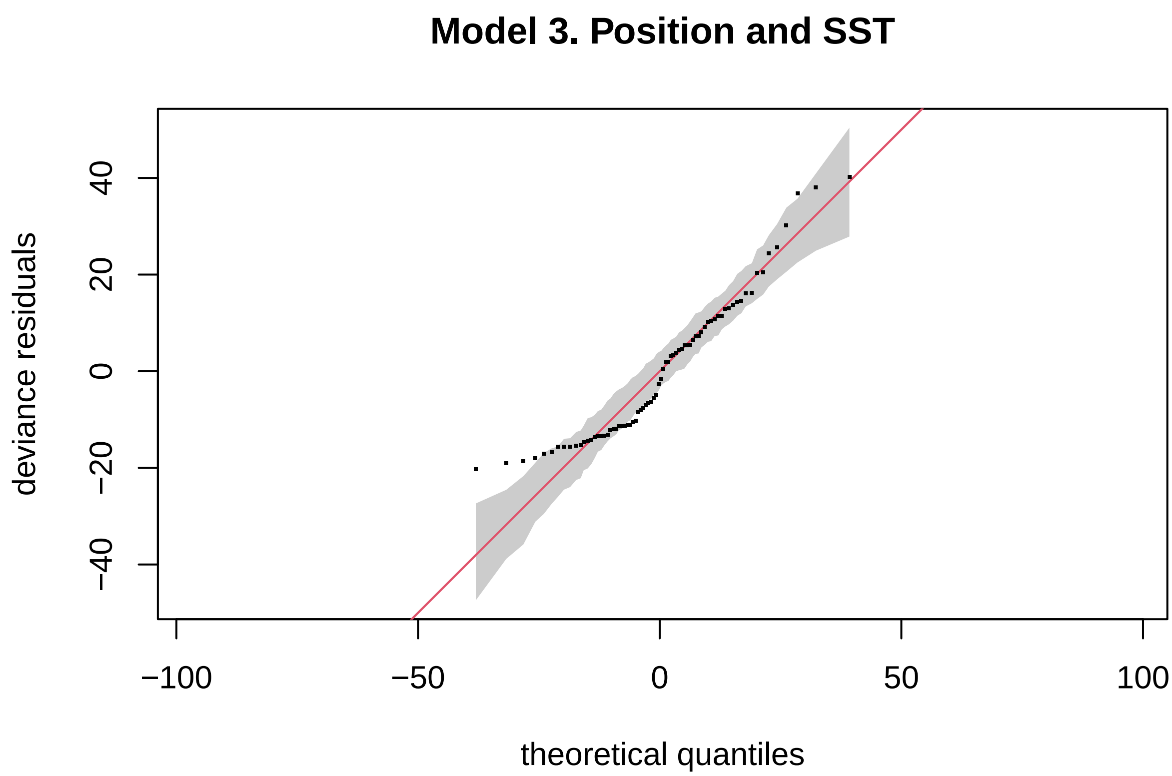
Wood SN (2017) Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall/CRC

Appendix A: Model summary and diagnostics

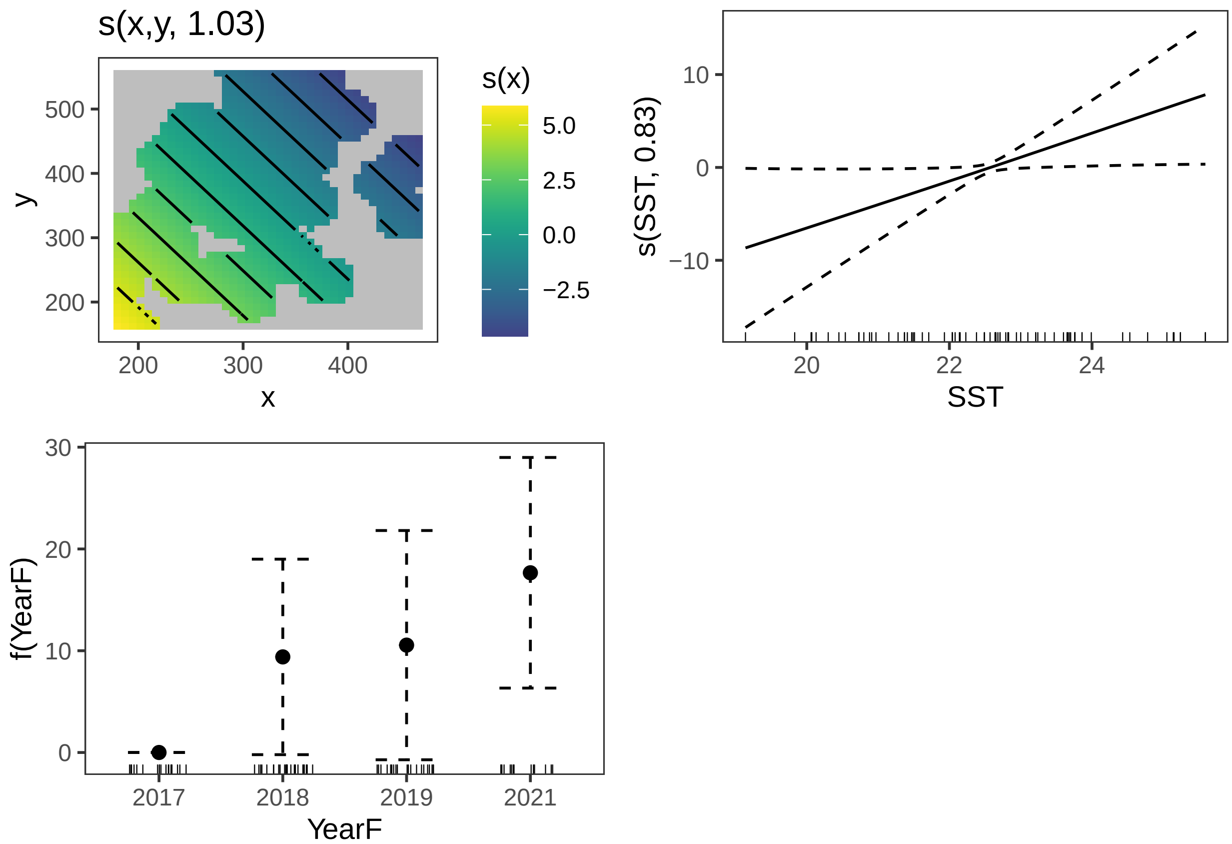
Summary output for the group size model

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## sqrtsize ~ YearF + s(x, y, bs = "ts", k = 8) + s(SST, bs = "ts")
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   10.733      3.850   2.788  0.00684 **
## YearF2018      9.392      4.901   1.916  0.05945 .
## YearF2019     10.549      5.746   1.836  0.07067 .
## YearF2021     17.659      5.781   3.055  0.00320 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##           edf Ref.df     F p-value
## s(x,y) 1.0297      7 0.373  0.0871 .
## s(SST) 0.8296      9 0.472  0.0251 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.138   Deviance explained = 19.5%
## -REML = 302.25  Scale est. = 236.98    n = 75
```

The plots below show the quantile-quantile plots (q-q plots) for all the fitted GAMs; 'Inter' indicates models where location was interacting with year. Ideally, the residual q-q plots should lie on a straight line and the figure below shows that there is some deviation. Grey shaded areas show results of 100 replicates. The number in the plot caption is the model number (see Table 8); model number 3 was selected.



The plots below show the estimated model parameters for included factors (Year; with 2017 as the reference level) and continuous variables: position and SST. The y-axis (or title in case of the plot of x and y) shows the estimated degrees of freedom. Dashed lines indicate two standard error bounds.

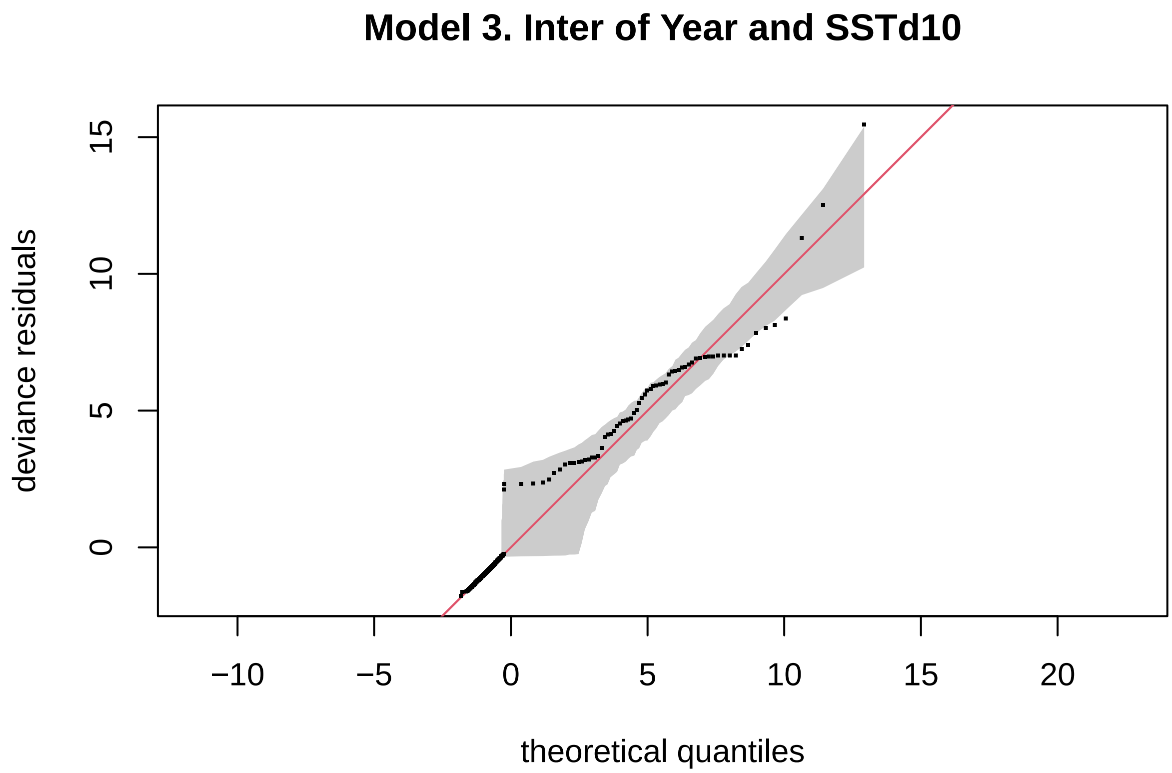


s(x,y, 1.03)
s(x)
5.0
2.5
0.0
−2.5
500
400
300
200
y
200
300
400
x
10
0
−10
s(SST, 0.83)
20
22
24
SST
30
20
10
0
f(YearF)
2017
2018
2019
2021
YearF

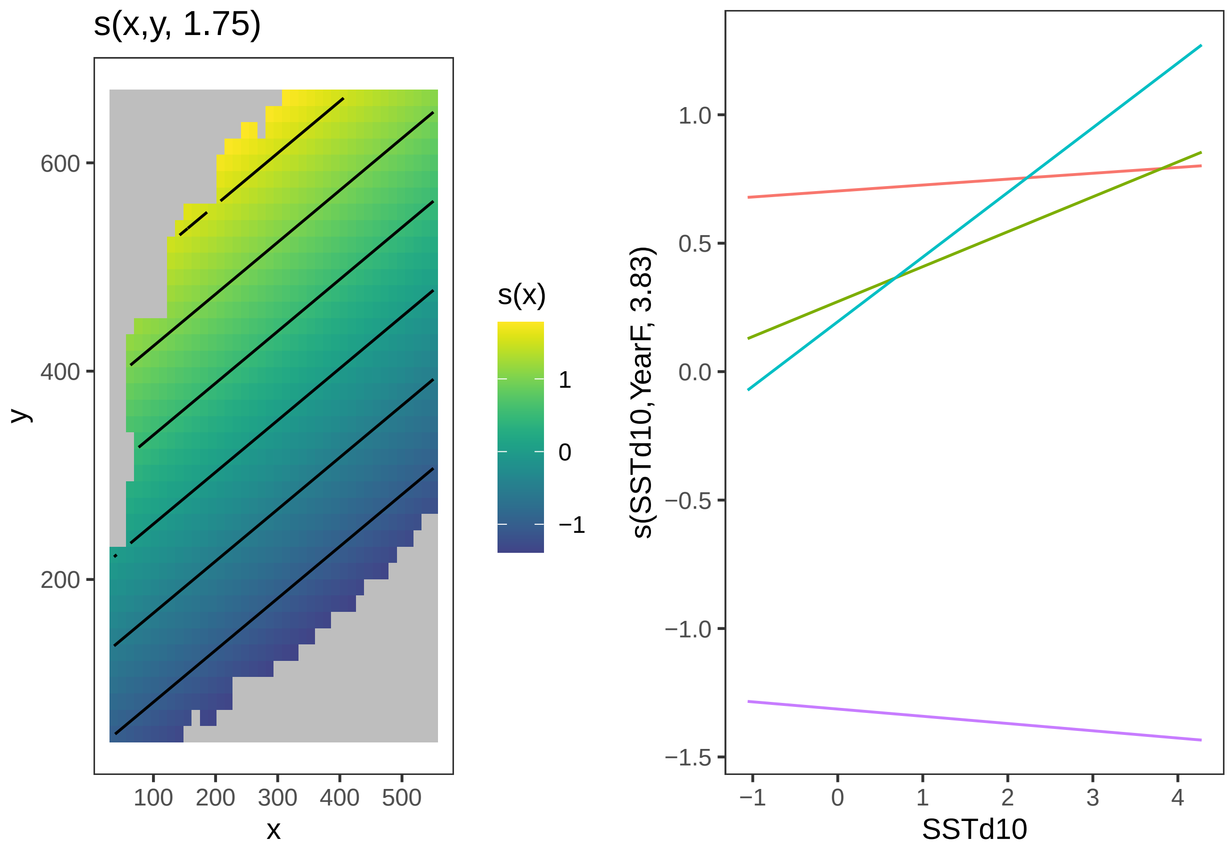
Summary output for the group encounter rate model

```
##
## Family: Tweedie(p=1.393)
## Link function: log
##
## Formula:
## abundance.est ~ s(x, y, bs = "ts") + s(SSTd10, YearF, bs = "fs",
##     k = 8) + offset(off.set)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -5.7600     0.5469  -10.53   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                   edf Ref.df     F  p-value
## s(x,y)          1.748     29 0.421 0.000908 ***
## s(SSTd10,YearF) 3.835     31 0.769 1.86e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.0101   Deviance explained = 15.1%
## -REML = 489.64  Scale est. = 17.727    n = 2838
```

Ideally, the residuals in the q-q plots should lie on a straight line and the figure below shows that there is some deviation, although less pronounced than the models for group size. Grey shaded areas show results of 100 replicates. The number shown in the individual plot caption is the model number (see Table 8); model number 3 was selected.



The plots below show the estimated model parameters for location and the interaction between Year and $SSTd10$. The y-axis (or title for the plot of x and y) shows estimated degrees of freedom.

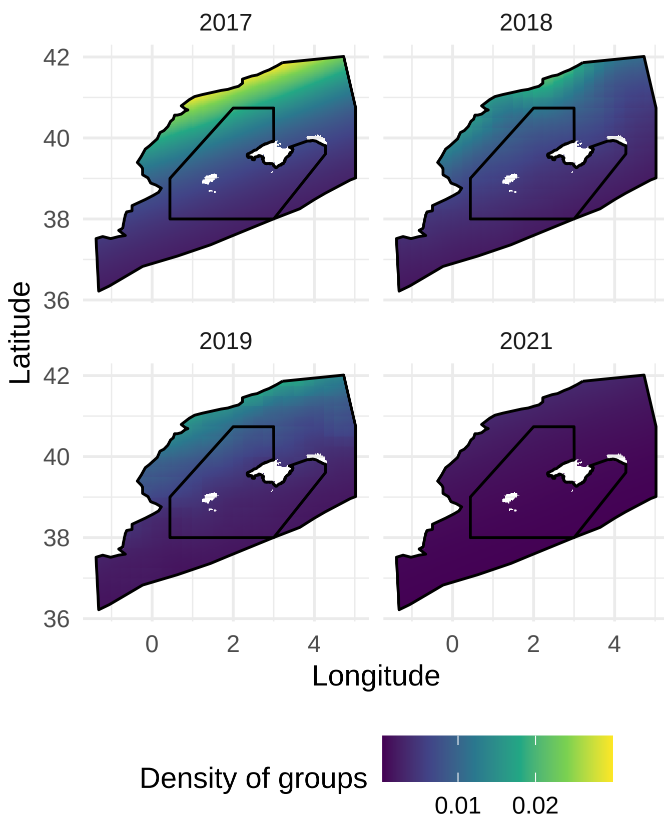


s(x,y, 1.75)
600
400
200
y
100 200 300 400 500
x
s(x)
1
0
−1
1.0
0.5
0.0
−0.5
−1.0
−1.5
s(SSTd10,YearF, 3.83)
−1 0 1 2 3 4
SSTd10

Appendix B: Estimated surfaces of BFT, number of groups and group size.

This appendix displays the surfaces of the estimated density of groups (Figure B1) and the estimated group sizes (Figure B2) and the corresponding CVs. The estimated numbers and average group sizes are shown below (and also in Table 9).

Year	Region	Groups	CV.Group	Group.Size	CV.Group.Size
2017	A	456	0.55	220	0.46
2018	A	334	0.33	380	0.35
2019	A	268	0.40	285	0.38
2021-core	A-core	52	0.51	855	0.27
2021-outer	A-outer	145	0.55	840	0.27



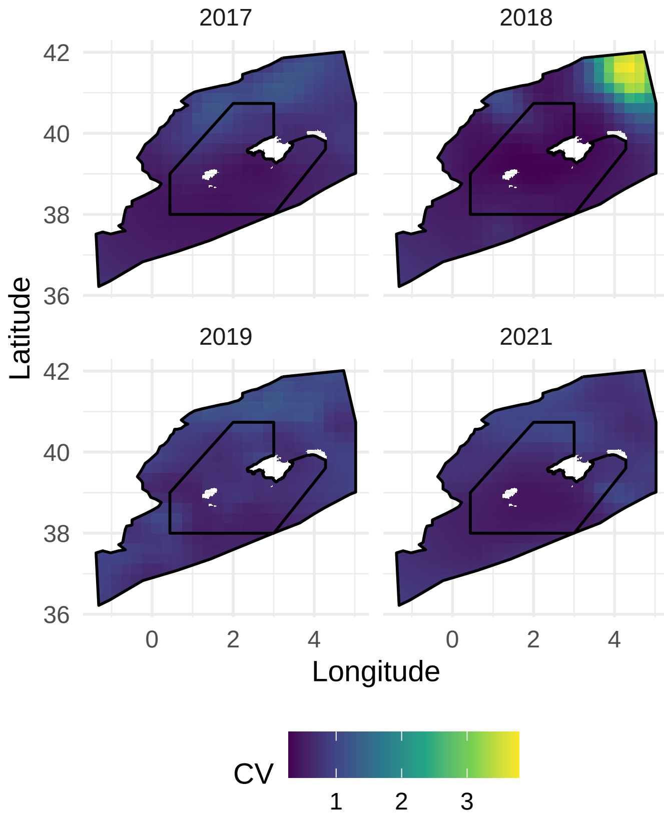
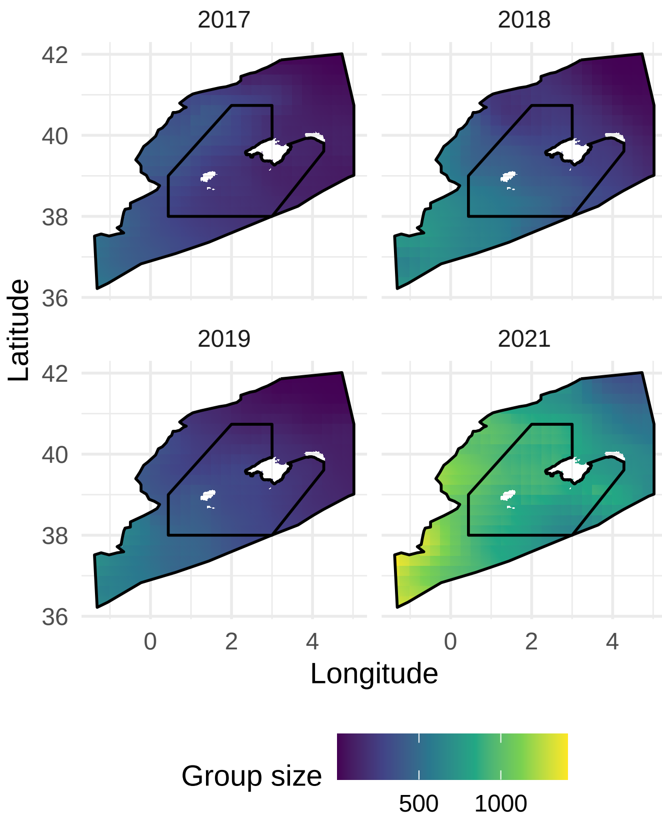


Figure B1. Estimated density of BFT groups (groups/km^2) and CV (not including detection function uncertainty) for each year using SST on 15th June.



2017
2018
2019
2021
Latitude
42
40
38
36
0
2
4
Longitude
Group size
500
1000

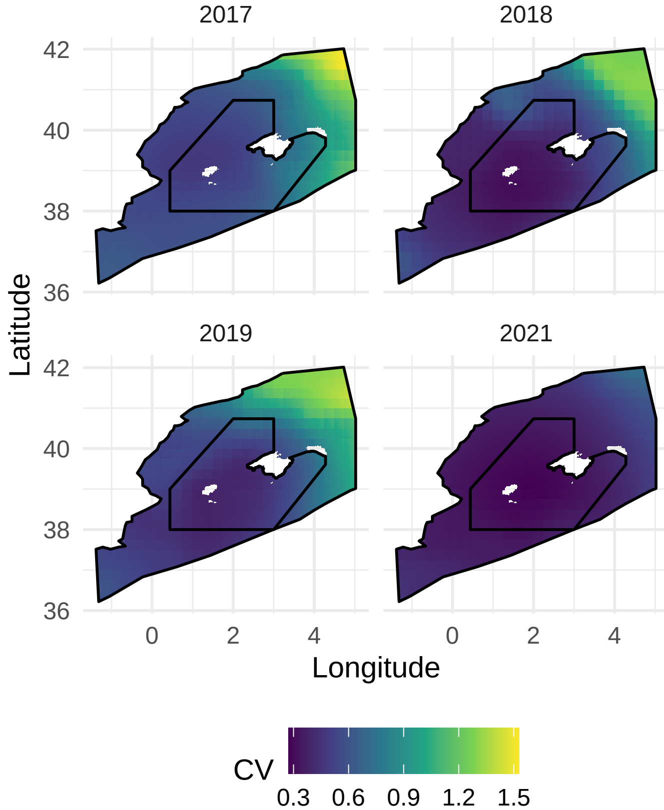
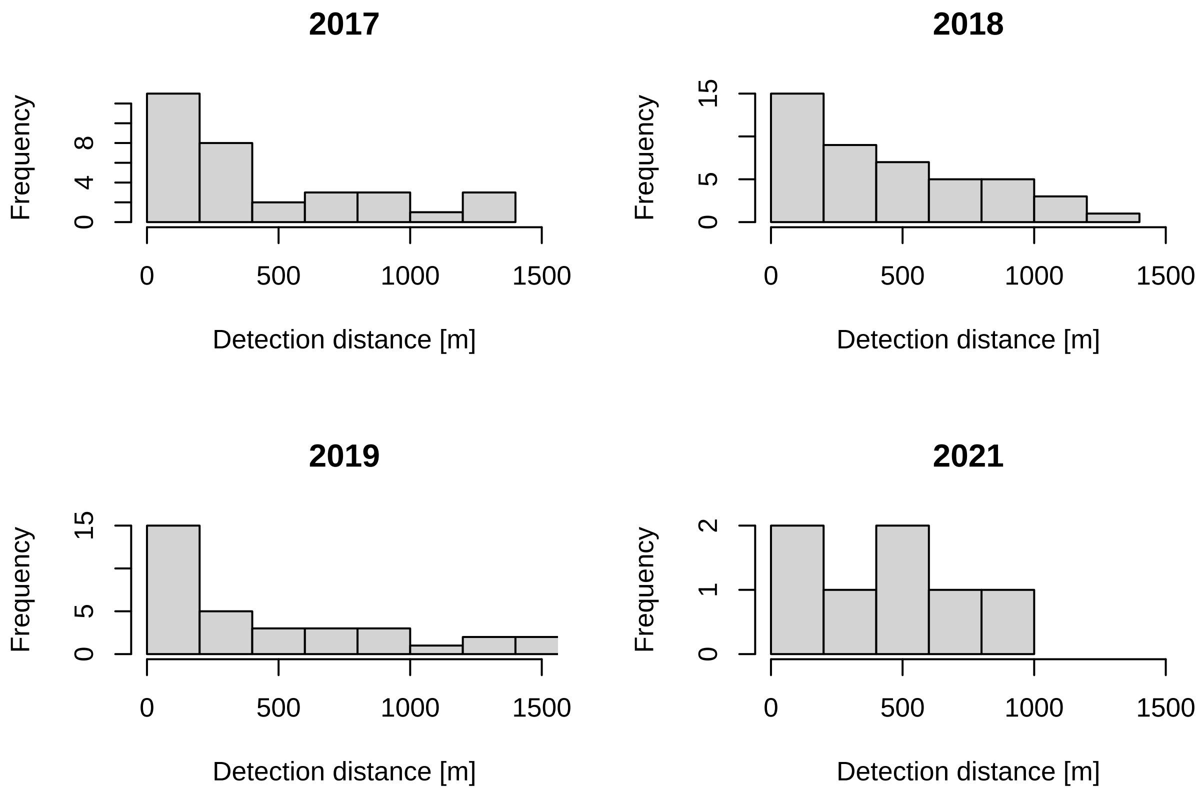


Figure B2. Estimated group sizes and CV using SST on 15th June.

Appendix C: comparison of detection distances between survey years within truncation distance of 1500 m.



Appendix D: Summary of explanatory variables associated with observed data and prediction region

To obtain density and abundance estimates from the model-based approach, predictions were required for the whole of A-core and A-outer (for 2021) regions and not just at values where data were surveyed. This appendix briefly summarises the range of values of the explanatory variables in the observed data and in the prediction region (covering both A-core and A-outer blocks).

Table D1 shows that the prediction region covers a wider range of values for the explanatory variables than associated with the segments. This is to be expected as only a small percentage of the region was covered (and A-outer only covered in 2021), however, the values are not substantially outside the segment ranges.

Table D1. Minimum and maximum values of the candidate explanatory variables associated with the observed segments and prediction grid; x (km), y (km), depth (metres), SST (^{o}C) and $SSTd10$ (^{o}C).

Variable	Segments.Min	Segments.Max	PredGrid.Min	PredGrid.Max
x	36.65	550.6	11.76	562.9
y	51.55	662.2	25.95	667.2
Depth	-2813	-2	-2930	-1
SST	17.75	25.62	17.69	26.28
SSTd10	-1.06	4.28	-2.22	4.97