ATLANTIC WIDE RESEARCH PROGRAMME FOR BLUEFIN TUNA

(ICCAT GBYP Phase 15)

SHORT-TERM CONTRACT FOR GBYP CIRCULAR # G-00109/2026

Aerial survey for Bluefin spawning aggregations: aerial survey data analysis and further development of model-based approaches for the standardization of the index of abundance - Atlantic-wide research programme for bluefin tuna (ICCAT GBYP – Phase 15 - 2026)

FINAL REPORT

September 2026

Authors:

IEO-CSIC: Alvarez-Berastegui D., Tugores M.P., Ottmann D., Ellacuría I., Reglero P.

University of St Andrews: Burt L., Paxton C.G.M., Chudzińska M.

Document control

Please consider this document as uncontrolled copy when printed.

Version	Date	Reason for issue	Prepared by	Checked by
Draft	25 August 2026	Draft	DAB, PT, LB	CP
Draft v2	27 August 2026	Draft v2	DAB, LB	LB, MC
Final draft	28 August 2026	Submission to ICCAT	DAB, LB	
Final v1	09 Sept 2026	Final submission	DAB, LB	CP

Co-funded by the European Union

This report has been prepared in response to the call for tender ICCAT GBYP circular # G-00109/2026. There were two objectives in this tender, and the report is divided into two parts, each addressing one of these objectives.

Contents

OBJECTIVE 1

Design-based inference to update the index of abundance and biomass of bluefin tuna in the Balearic Sea including the aerial visual survey for 2026 in Area A.

Authors:

CREEM, University of St Andrews: Burt L, Paxton C. G. M., Chudzińska M.

This chapter responds to the objective 1 of the SHORT-TERM CONTRACT, REFERENCE ICCAT GBYP CIRCULAR # G-00109/2026

Please cite this section of the report as: Burt *et al*. (2026) Design- based inference to update the index of abundance and biomass of bluefin tuna in the Balearic Sea including the aerial visual survey for 2026 in Area A. Report number CREEM-2026-01. Provided to ICCAT, August 2026 (Unpublished).

Summary

This section of the report has been prepared in response to objective 1, as specified in the call for tender ICCAT GBYP circular # G-00109/2026.

Aerial surveys are conducted annually in regions of the Mediterranean Sea to collect data to estimate abundance and biomass of bluefin tuna. In 2026, the region around the Balearic Islands (Block A) was surveyed. Line transect distance sampling methods were employed on the survey and 5,837 km of search effort were flown, and 8 schools of adult bluefin tuna schools were detected on search effort along the survey transects.

A detection function obtained previously using detections from surveys in 2017 to 2022 was applied to the survey data from 2017 to 2026 data to obtain a strict update of the annual indices of estimates of abundance of individuals and total biomass. Abundance for block A in 2026 was estimated to be 28,997 fish (CV=0.51) and biomass was estimated to be 3,733 tonnes (CV=0.62). These indices are likely an underestimate of the true abundance and biomass because the proportion of BFT schools at the surface to be detected has not been considered and observers may not detect all available schools on the transect.

Introduction

Aerial surveys have been undertaken in the Mediterranean Sea to detect bluefin tuna (BFT) from 2010 to 2026 under the auspices of the International Commission of the Conservation of Atlantic Tunas (ICCAT) Atlantic wide research programme for bluefin tuna (GBYP). The main objectives of this programme are to improve a) understanding of the key biological and ecological processes, b) current assessment methodology, c) management procedures and d) advice for this species. This report fulfils objective 1 as specified in the call for tender ICCAT GBYP circular # G-00109/2026 and presents the results from the 2026 GBYP Balearic Sea aerial survey data (Area A). These data were combined with data from surveys in 2017 to 2025 to update the index of BFT eastern Atlantic and Mediterranean relative spawning stock biomass and abundance, currently used within the framework of BFT stock management.

Estimates were obtained using line transect distance sampling methods (Buckland *et al*. 2001) and the biomass (total weight) and abundance (numbers of individuals) were updated as a strict update. To provide a strict update, the detection function fitted in Paxton *et al*. (2023) (i.e. the same variables and values of the model parameters) were used to obtain estimates for 2026 abundance and biomass.

Methods

Overview of the aerial surveys

In recent years, three regions, or blocks, have been surveyed in the Mediterranean Sea, denoted by A, C and E (Figure 1); block G (not shown) was last surveyed in 2019. In 2025 and 2026, only block A was surveyed. Details of the survey protocol and outcomes of the 2026 survey are provided by Pérez-Gil and Pérez-Gil (2026), and so only information relevant for the analysis is provided here.

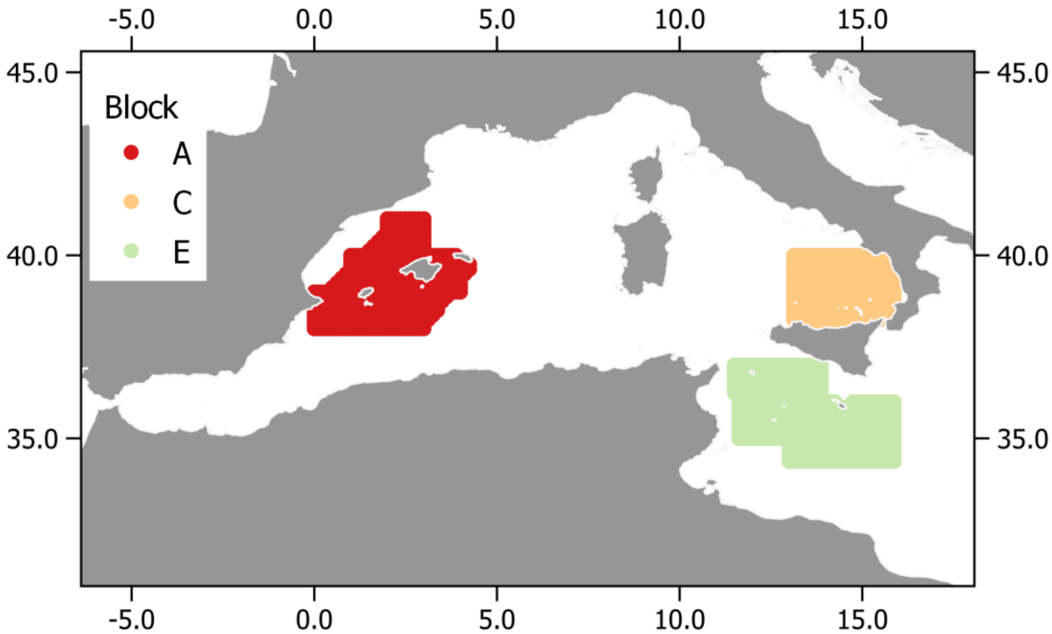


Figure 1. Depiction of three survey blocks in the Mediterranean Sea. Block G is in the Eastern Mediterranean Levantine Sea.

Table 1 summarises the timing, company and airplane type used for the survey.

Table 1. Summary of the 2026 survey including dates when search effort took place.

Block	Dates	Company	Airplane
A	01 June – 24 June	Airmed	Partenavia

Statistical methods

Although a strict update is required for this analysis, and hence the detection function has already been obtained, the analysis methods are included for context. Line transect distance sampling methods (Buckland *et al*. 2001) were used to estimate density and abundance and these methods are described below.

Data processing

Data from the 2026 survey and for previous surveys were provided by ICCAT and were processed to obtain the structure and information required for distance sampling analysis. All data from 2017 to 2025 were processed in 2025 as part of another project (Álvarez-Berastegui *et al*. 2025) to ensure it was done in a systematic and consistent way.

Estimating density and abundance

In distance sampling (DS) methodology, the perpendicular distances to detections of BFT are used to model how detectability decreases with increasing distance and hence used to estimate a probability of detection ($\hat{p}$). Using standard methodology (Buckland *et al*. 2001), the estimated density ($\widehat{D}$) and abundance ($\widehat{N}$) of fish in a survey block was obtained from

$$\widehat{D} = \frac{n}{2wL} . \frac{1}{\hat{p}} . E[s]$$

$$\widehat{N} = A . \widehat{D}$$

where for each block, A is the size of the block, n is the number of detected schools, w is the truncation distance associated with the detection function, L is the total length of transects covered on search effort and $E[s]$ is the expected school size.

Biomass of fish was also of interest and to obtain this, school size was replaced in the equation above by expected biomass (which was recorded for each detected school) and a corresponding probability of detection (see below).

Calculating perpendicular distance

The perpendicular distance from the detected school to the transect was calculated using the trigonometric relationship:

$$y_i = h_i * \tan((90 - \theta_i)$$

where y_i is the perpendicular distance between the transect and the i^{th} school, θ_i is the declination angle measured when the plane was abeam and h_i is the height of the airplane above sea level when abeam of the detected school (Figure 2). Where measurements were not taken abeam of the detected school, the measurements taken when the school was first detected were used.

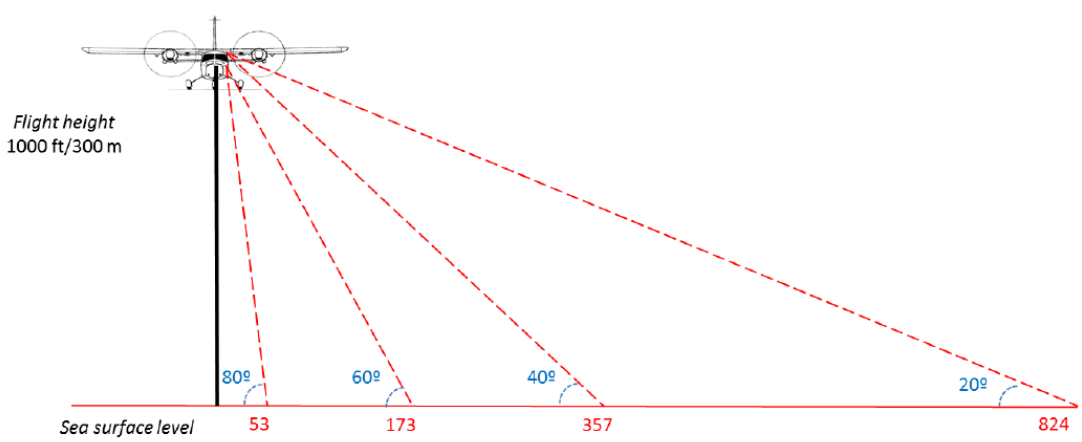


Figure 2. Example of the key declination angles and perpendicular distances at an altitude of *h* = 1000 ft = 300 m (Figure 5 from ICCAT survey protocol. Source: https://www.iccat.int/.)

Detection function

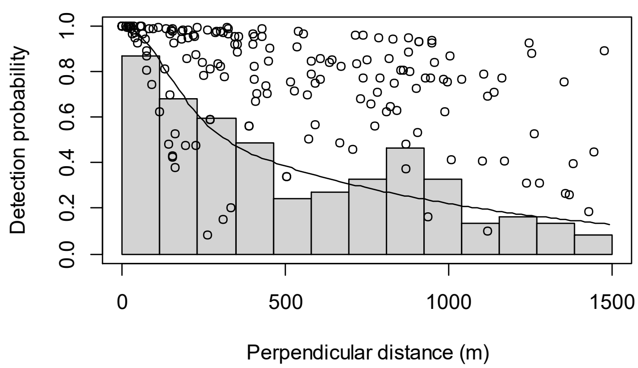
Two critical assumptions of DS methods are that all schools on the transect (i.e., at zero perpendicular distance) are detected with certainty and that distance measurements are exact (i.e., measured without error). Given these assumptions, the distribution of perpendicular distances is used to model how the probability of detection decreases with increasing distance from the transect. If detection on the transect is not certain (i.e. $g(0)<1$), the estimates will underestimate the true abundance and will represent estimates of relative numbers of animals.

In a conventional distance sampling approach, perpendicular distance is the only explanatory variable used to obtain $\hat{p}$ but the model can easily be extended to include additional explanatory variables which affect detectability, such as school size (Marques *et al*. 2007).

In Paxton *et al*. (2023), the variables selected in the detection function were *company* and *log(school size)* to estimate tuna abundance and *company* and *log(biomass)* to estimate tuna biomass. These detections functions were fitted to data from 2017 to 2022 in blocks A, C, E and G (Figure 3). These models (i.e., using the same model parameter values) were applied to obtain estimates for 2023, 2024, 2025 and 2026. This means that detections from the four most recent years were not included in the detection functions. To avoid changing the parameters, we have applied the same levels of company used by Paxton *et al*. (2023) to apply the detection function to the recent data; the levels were Action Air, Aerial Banners, Airmed, Air Perigord and Unimar.

Paxton *et al*. (2023) truncated the perpendicular distances at 1500 m, to avoid a long tail in the detection function and be consistent with previous analyses. The choice of this truncation distance was based on visual inspection of the fitted detection function. No left truncation was applied; left truncation is a common practice for aerial surveys, due to difficulties in searching directly underneath the plane. The planes used in the aerial surveys under consideration were fitted with bubble windows improving the visibility directly below the planes.

a)



b)

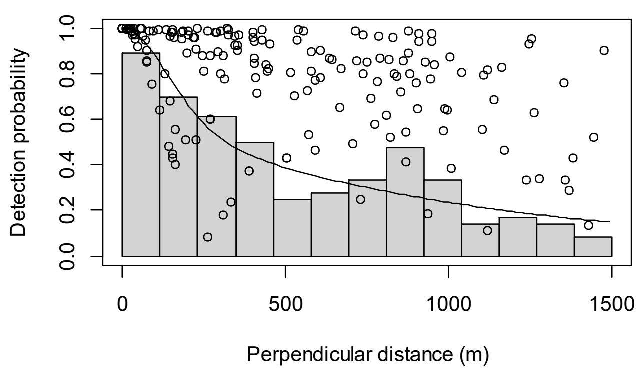


Figure 3. Detection functions (solid line) overlaid onto the scaled histogram of perpendicular distances: a) abundance and b) biomass. Dots indicate individual detections.

Estimating density and abundance and biomass

Detections and search effort were pooled within each block and year to obtain encounter rates, and hence obtain estimates of density and abundance, by year for each block surveyed in that year. The lengths of the realised transects were calculated from the recorded positions (i.e. latitude and longitude) when observers were on search effort.

This same approach was used to estimate biomass; in this case, the size of observed schools was replaced by the recorded biomass.

Sightings of BFT schools were included if detected when the observers (both professional and scientific observers) were on search effort along the transect lines. Secondary sightings (schools seen, for example, when the plane left the transect to confirm school sizes for a previous sighting) were not included.

School sizes and biomass values recorded by the professional observers were used in preference to those of the scientific observer unless these values were missing, in which case estimates from the scientific observers were used, if available (ICCAT pers. comm).

Schools that were recorded as 100% small (i.e. individual fish < 25kg, referred to as juveniles) were excluded. The remaining schools are referred to as adult schools (i.e. fish ≥ 25kg). Sightings where perpendicular distances, school sizes or biomass values were missing were also excluded from the analysis.

The analysis was performed in R version 4.4.0 (R Core Team 2024) using functions from the packages Distance (Miller *et al*. 2019) and mrds (Laake *et al*. 2025).

Results

Summary of search effort and BFT sightings for 2026 data

Table 2 summarises the search effort along the transects and number of adult BFT schools detected on search effort for the 2026 survey. Eight adult schools were detected but note that 23 schools of juvenile BFT were also detected. Figure 4 shows the locations of the search effort and detected adult schools.

Table 2. Summary of survey data for 2026. Number of transects, total length of search effort along transects, size of the area covered (i.e. length of transect x truncation distance x 2), number of adult BFT schools detected during search effort along the transect lines and adult BFT schools within 1,500m of the transects, encounter rate of schools (ER) and associated coefficient of variation (CV) of the encounter rate.

Year	Block	Number of transects	Total length (km)	Covered area (km^2)	Number of schools	Number of schools within truncation distance	ER (schools/km)	CV
2026	A	30	5,836.7	17,510	8	7	0.0012	0.42

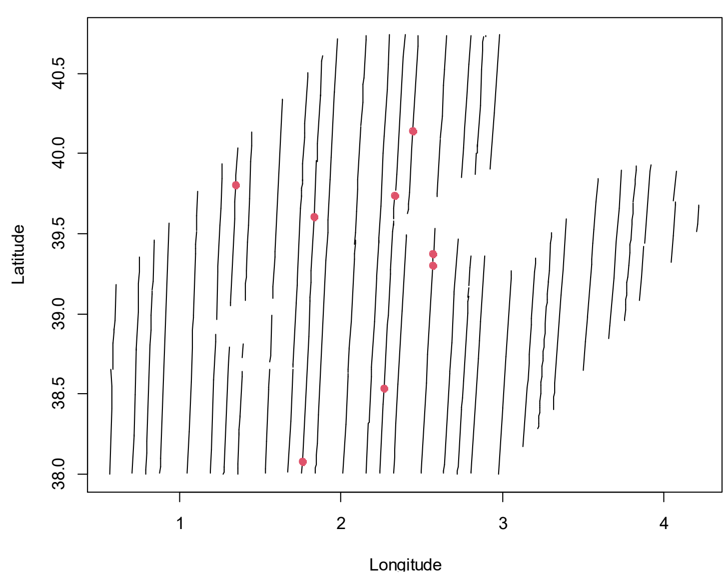


Figure 4. Location of the 2026 survey transects (lines) and detected schools (red dots) in block A. The detections represent on-effort detections of adult schools along the transects.

Estimates of abundance and biomass

The results for 2026 assuming an identical function to that fitted in Paxton *et al*. (2023) are given in Table 3.

Table 3. Estimated number of individuals (*N*, in thousands) and biomass (B, in tonnes) in block A in 2026 with lower (LCL) and upper (UCL) limits of a 95% confidence interval and associated coefficient of variation (CV).

Year	Block	*N*	CV	LCI	UCI	B	CV	LCI	UCI
		Abundance				Biomass			
2026	A	29.0	0.51	10.9	77.0	3,733	0.62	1,168	11,934

Following Burt et al. (2025), estimates for 2026 have been added to previous estimates for block A; the patterns are similar for both indices. A summary of the survey data by block and year is given in Appendix A.

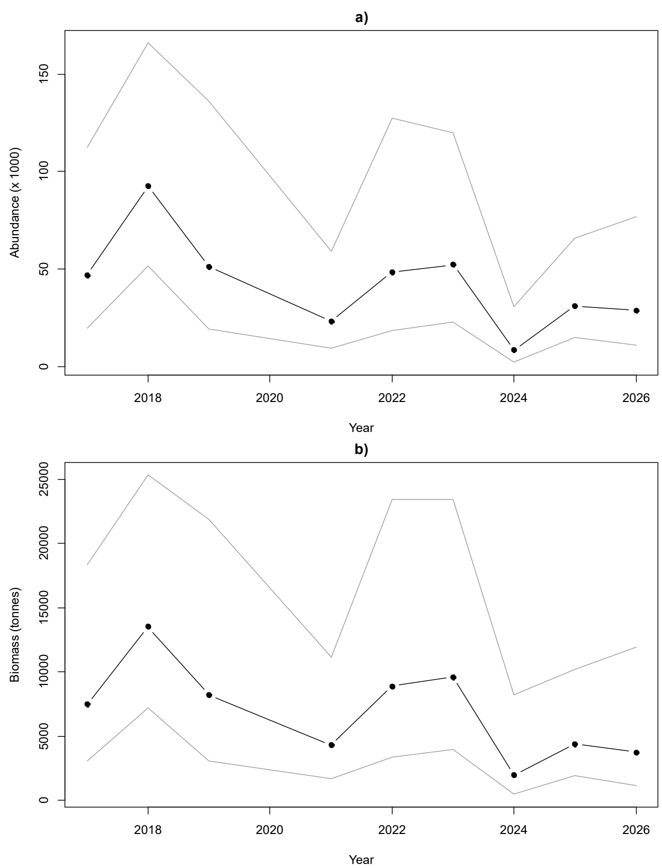


Figure 5. Estimates of a) abundance (in thousands) and b) biomass (in tonnes) of BFT for block A, using a detection function based on data from 2017 - 2022. Dots show estimated values, and grey lines show upper and lower limits of the 95% confidence interval.

The patterns for both indices are similar and the estimates for 2026 are like the estimates for 2025; the school encounter rates were very similar in these two years and although the median school size and median biomass were lower in 2026 compared to 2025, the range was larger in 2026 (Appendix B).

Discussion

It should be emphasised that the estimates given here were based on detections of schools of fish (of one or more fish) observed at the surface, or close enough to the surface to be detected. It has been assumed that surface fish available to be detected on the transect (i.e. at zero distance) were certain to be detected. It is likely that some schools may have been too deep to be detected, and available schools on the transect may have been missed although the use of bubble windows in the plane reduces this issue. Hence, these estimates provide indices of relative abundance and biomass rather than absolute abundance/biomass; neither the availability of fish to be detected nor the detection of available fish on the transect have been considered in the calculations.

The detection function used here is based on a detection function using detections from 2017 to 2022 surveys and included the logarithm of school size or biomass and company. The detection function used here is based on a detection function using detections from 2017 to 2022 surveys and included the logarithm of school size or biomass and company. It is anticipated that detection of school will decrease as the distance of the school from the transect increases, however, there is an increase in the number of schools detected between about 500 and 1000m. Numbers of schools detected are relatively small each year but surveys in 2017 and 2018 tended to have more detections than expected in this range compared to the detections closer to the transect and so had a cumulative effect.

Four more years of survey data are now available, and these additional sightings could be included to update the detection function. It is likely that model parameters have changed and could be estimated more reliably, for example, the company Airmed, who conducted the survey in 2025 and 2026 were last used in 2019 and their recent detection function may be different to that in the past. To be consistent with previous analyses, five levels for company have been used but it may be that the levels could be reduced; “Unimar” and “Unimar/Aerial Banners” seem to provide a service in collaboration and so could be combined into one level. In addition, applying a strict update only works because the companies that have undertaken the recent surveys were also used previously; if a new company is used in future, these detection function will not be applicable.

References

Álvarez-Berastegui D, Tugores MP, Burt L, Paxton CGM, Chudzinska M, Ottman D, Alvarez I, Reglero P, (2025). Preparation of aerial survey data and environmental data to estimate abundance and biomass of Bluefin Tuna in the Mediterranean Sea using model-based analysis. Provided to ICCAT, December 2025 (Unpublished).

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. (2001). Introduction to Distance Sampling. Estimating abundance of biological populations. Oxford University Press, Oxford.

Burt, L., C. Paxton, and M. Chudzińska. (2025). Design- based inference to estimate abundance, and biomass of bluefin tuna in the Mediterranean Sea: analysis including the aerial visual survey for 2025. Report number CREEM-2025-02. Provided to ICCAT, December 2025 (Unpublished).

Laake, J., D. Miller, and F. Petersma. (2025). mrds: Mark–Recapture Distance Sampling. R package version 3.0.1. https://cran.r-project.org/package=mrds.

Marques, T. A., L. Thomas, S. G. Fancy, and S. T. Buckland. (2007). Improving estimates of bird density using multiple-covariate distance sampling. The Auk 124:1229-1243.

Miller, D. L., E. Rexstad, L. Thomas, L. Marshall, and J. L. Laake. (2019). Distance Sampling in R. Journal of Statistical Software 89:1 - 28.

Paxton, C. G. M., C. S. Oedekoven, M. Chudzińska, M. Pilar Tugores Ferrá, and D. Alvarez-Berastegui. (2023). Design- and model-based inference to estimate density, abundance and distribution of BFT in the Mediterranean Sea. Report number CREEM-2023-04.

Pérez-Gil M and Pérez-Gil E (2026). Prospección aérea para el seguimiento de las concentraciones de reproductores de atún rojo en el mar Mediterráneo Mar Balear (área A)

R Core Team. (2024). R: A Language and Environment for Statisical Computing. R Foundation for Statistical Computing.

Appendix A: Summary of the survey blocks and realised survey data in each year

This appendix contains a summary of the survey blocks and survey data in each year. As part of another project (specified in the call for tender ICCAT GBYP circular # G-00510/2025), the data for the surveys listed in Table A2, were processed and formatted for distance sampling analysis in a systematic way. Thus, any decisions required regarding corrections could be applied consistently.

Table A1. Size of the survey blocks (km^2).

Block	Year	Area (km^2)
A	2017 - 2019	61,849
	2021 - 2026	61,837
C	2017 - 2025	53,868
E	2017 - 2025	93,614
G	2017 - 2025	38,788

Table A2. Company and aircraft used to undertake the surveys and data used in the analysis: k is the number of transects, L is the total length of search effort along transects, and n is the number of adult groups detected during search effort along the transects and within 1,500 m of the transects.

Year	Block	Company	Aircraft	k	L (km)	n
2017	A	Airmed	Partenavia	29	4996	18
	C	Unimar	Partenavia	25	4569	7
	E	Airmed	Partenavia	30	6478	5
	G	ActionAir	Cessna	57	4493	2
2018	A	Airmed	Partenavia	39	5996	25
	C	Unimar	Partenavia	25	4931	8
	E	Unimar	Partenavia	40	8822	7
	G	ActionAir	Cessna	56	4122	4
2019	A	Airmed	Partenavia	30	5466	12
	C	Unimar	Partenavia	23	4818	4
	E	ActionAir	Cessna	40	8567	5
	G	ActionAir	Cessna	55	4084	2
2021	A	ActionAir	Cessna	27	6317	8
	AO[1]	ActionAir	Cessna	12	2536	2
2022	A	Air Perigord	Cessna	30	5319	9
	C	Unimar/Aerial Banners	Partenavia	25	5144	11
	E	Unimar/Aerial Banners	Partenavia	30	6607	3
2023	A	Air Perigord	Cessna	30	5788	23[2]
	C	Unimar/Aerial Banners	Partenavia	25	4976	3
	E	Unimar/Aerial Banners	Partenavia	21	5165	8
2024	A	Air Perigord	Cessna	30	5679	7
	C	Unimar/Aerial Banners	Partenavia	25	4988	3
	E	Unimar/Aerial Banners	Partenavia	34	7234	8
2025	A	Airmed	Partenavia	28	5796	7
2026	A	Airmed	Partenavia	30	5837	7

[1] Data from the survey block AO (an area surrounding block A) surveyed in 2021 were not included in this report but is included in this table for completeness.

[2] One detection missing a biomass estimate was excluded from the biomass analysis.

Appendix B: School sizes and biomass by year

Figure B1 illustrates the distributions of sizes (number of fish) and biomass (total weight) for adult schools of BFT detected in block A on search effort along the transects (schools have not been truncated at 1,500m).

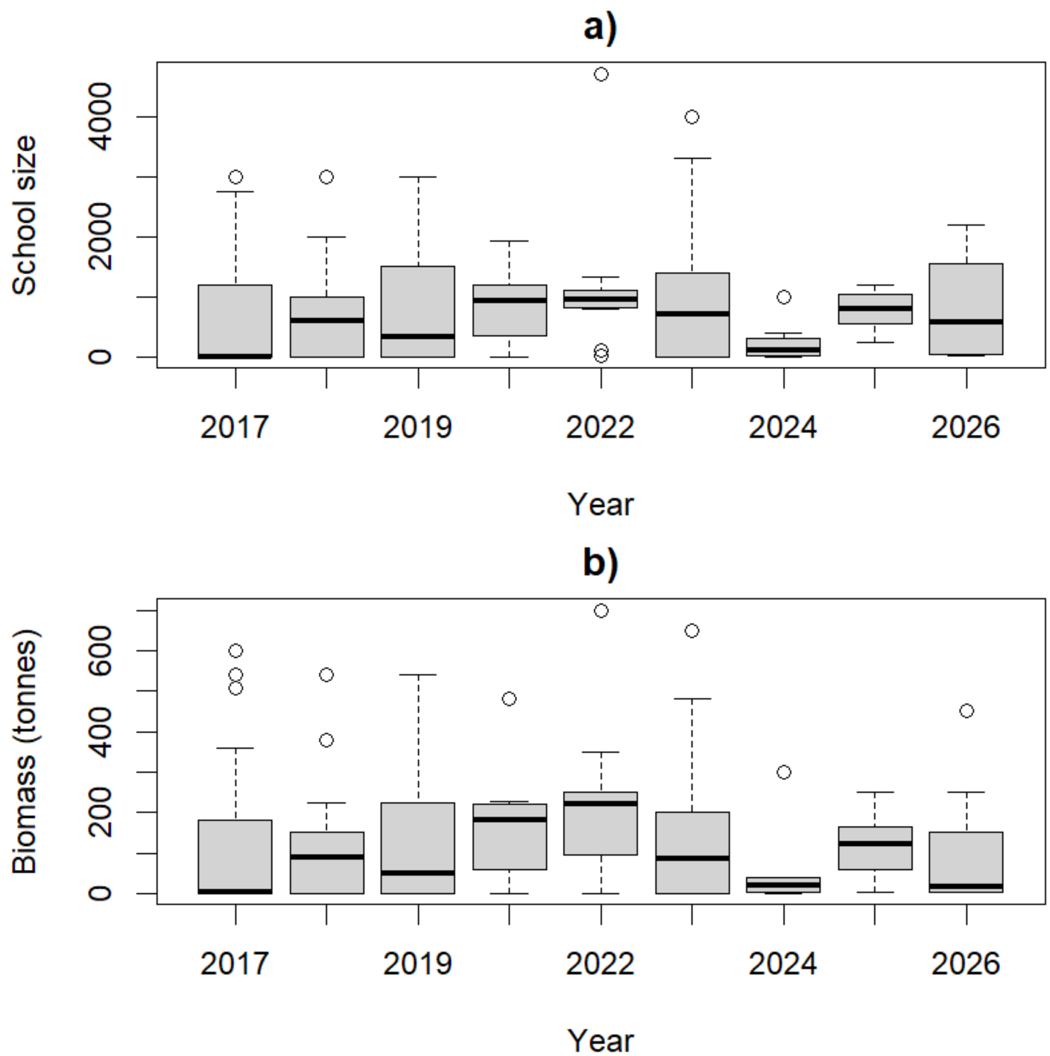


Figure B1. Distribution of a) school sizes and b) biomass for adult schools detected on search effort along transects in block A by year. The tails extend to 1.5 times the interquartile range from the box (or the maximum if smaller); points beyond that are considered unusual.

OBJECTIVE 2

Aerial survey for Bluefin spawning aggregations: aerial survey data analysis and further development of model-based approaches for the standardization of the index of abundance - Atlantic-wide research programme for bluefin tuna (ICCAT GBYP – Phase 15 - 2026)

Authors:

IEO-CSIC: Alvarez-Berastegui D., Tugores M.P., Ottmann D., Ellacuría I., Reglero P.

CREEM, University of St Andrews: Burt L., Paxton C.G.M., Chudzińska M.

This chapter responds to objective 2 of the SHORT-TERM CONTRACT, REFERENCE ICCAT GBYP CIRCULAR # G-00109/2026

Please cite this chapter as: Álvarez-Berastegui D, Tugores MP, Ottmann D, Ellacuria I, Reglero P, Burt L, Paxton CGM, Chudzinska M (2026). Aerial survey for bluefin spawning aggregations: aerial survey data analysis and further development of model-based approaches for the standardization of the index of abundance – in Atlantic-wide research programme for bluefin tuna (ICCAT GBYP – Phase 15 -2026). Provided to ICCAT August 2026 (Unpublished).

Introduction and structure

Aerial surveys have been undertaken in the Mediterranean Sea to detect Atlantic Bluefin tuna (BFT) since 2010 under the auspices of the International Commission of the Conservation of Atlantic Tunas (ICCAT), within the framework of the Atlantic-wide Research Programme for Bluefin Tuna (GBYP). The main objectives of GBYP are to improve understanding of key biological and ecological processes, enhance current stock assessment methodologies, strengthen management procedures and support scientific advice.

This section report is in response to Objective 2 in the call for tender ICCAT GBYP Circular No. G-00109/2026. It presents a detailed description of the data processing and modelling steps undertaken for analyses aimed at providing a standardized relative index of Bluefin tuna biomass in the spawning area of the Balearic archipelago, using model-based approaches endorsed by the BFT Working Group.

Objective 2 comprises three main components. These include the preparation of aerial survey data and environmental data covering the spatial and temporal extent of the surveys, the integration of both datasets into a unified analytical framework and the modelling to estimate spawning stock of BFT in the Balearic Sea.

This section of the report is structured into two sections. Section 1, developed by CREEM, University of St Andrews, includes the preparation of the aerial survey data, describes the procedures applied to the processing and standardization of the aerial survey information (for years 2010-2025). Section 2, developed by IEO-CSIC Environmental data repository and combination of aerial survey data with environmental data, addresses the development of the environmental data repository, the extraction of environmental information on the aerial survey samples and the analyses of environmental variability parameters relevant for the standardization of abundance indices.

Section 1

1.1 Methods

1.1.1. Preparation of the aerial survey data

Aerial survey data, definition of transects and segments

The aerial surveys are designed using distance sampling methodology (Buckland *et al.,* 2001). A plane flies along a series of parallel transects oriented approximately north-south (the first transect is placed with a random start point) and observers search for BFT. On detection of a school (of one or more fish), observers record key information, such as school size, total weight of the school (biomass) and angle of declination.

For the model-based analysis, the data are required to be in the form of segments of search effort along flight transects. Therefore, segments are defined as subunits of specific distance of the transects. The number of individual fish (for example) detected in a segment will form the response in the model with relevant environmental variables as explanatory variables in the model.

The survey data collected during the aerial surveys in the Balearic Sea (block A) was provided by ICCAT GBYP.

The target length of segments was 20 km, however, some segments were shorter because of natural breaks in search effort (such as going off the transect to confirm the sightings or a change in direction or conditions) and some segments were up to 30 km to avoid splitting a section into lengths much smaller than 20 km. Schools were allocated to segments on the basis of date and time.

Estimates of size (number of individual fish) and biomass (total weight) of a school recorded by the professional observers were used unless these were missing, in which case, estimates from the scientific observers were used. Where estimates from both observers were missing for a detected school (or perpendicular distance was missing), the median value recorded for that year was used (see Appendix A). These median values were calculated after removing schools recorded as juvenile (i.e. 100% small fish, see below). The median values were used rather than the mean because the distributions were right skewed and so the average can be inflated by large values.

In 2010 and 2011, the school sizes were not recorded in the same way as in later years and frequently only biomass and a general description of the weight of the fish (e.g. small, large) were recorded. In these cases, the school size was estimated from the total biomass for the school and the comments recorded by the observers.

Schools were included in the data processing if they were:

-adult schools (schools identified as 100% small schools (fish <25 kg) were excluded)

-detected when on search effort along the transect (and so schools seen when in transit or confirming a previous sighting were not included)

- within 1,500m of the transect.

The truncation distance of 1,500m has been used in previous distance sampling analyses of these data and although this results in a long tail (see results), truncating further would increase the number of segments with no detections.

Three response values were obtained for each segment:

- number of detected schools,
- total number of individual fish
- total biomass.

The detectability of schools decreases further away they are from the transect (and therefore the observers) and so these response values were adjusted for the estimated probability of detection. For example, the estimated total number of individual fish for segment i was obtained from:

$$\widehat{N}_i = \sum_{j=1}^{J} \frac{s_{ij}}{\hat{p}_{ij}}$$

where for segment i, s_{ij} is the size for school j (where J is the total number of schools in segment i) and $\hat{p}_{ij}$ is the estimated probability of detection for school j. This is to adjust for missing and underestimated schools (see below).

Biomass, adjusted for detection probability, was used as the response value in the modelling described in section 2.

1.1.2. Probability of detection

Two critical assumptions of distance sampling methods (Buckland *et al*. 2001) are that all schools on the transect (i.e., at zero perpendicular distance) are detected with certainty and that distances are measured without error. Given these assumptions, the distribution of perpendicular distances was used to model how the probability of detection decreases with increasing distance from the transect. The calculation of the perpendicular distances was described in Objective 1.

Two forms of the detection function were applied, a half-normal and hazard rate. In conventional distance sampling, perpendicular distance is the only explanatory variable, but the method can be extended to include additional explanatory variables which may affect detectability, such as school size (Marques *et al*. 2007). The variables considered here were:

- year (a factor with 12 levels),
- company undertaking the survey (a factor with 3 levels: Grup Air-Med, Air Perigord and Action Air),
- type of airplane (a factor with 2 levels: Cessna and Partenavia)
- sea state (a factor with 4 levels: 0, 1, 2 and ≥3),
- haze (a factor with 3 levels: 0, 1 and ≥2),
- logarithm of school size, and
- logarithm of biomass.

The logarithms of the latter two variables were taken to scale down the few very large values recorded (See Appendix B for Objective 1).

Detection of schools on the transect (i.e. at zero perpendicular distance) is affected by the type of windows installed on the plane. Bubble windows were installed part way through the 2011 survey (Barón et al. 2011). The report for the 2010 survey (Grup Air-Med 2010) does not mention bubble windows, however, in analyses of these data, Cañadas et al. (2010) used left-truncation to offset a lack of downward visibility. Due to differences between the 2010 and 2011 surveys and the following surveys, these two early surveys were excluded and data from 2013 to 2025 were included in the detection and hence in the following modelling.

The detection function was fitted in R (R Core Team 2024) using the mrds package (Laake *et al*., 2022). The detection function with the lowest AIC was selected and to check that the selected model was adequate, goodness-of-fit checks were performed.

Availability and perception bias

Detection on the transect may not be certain if schools are not available to be detected (i.e. are underwater) or if observers miss available schools (perception bias). The adjustment for detection probability described above does not account for availability or perception bias. Therefore, the estimates here are likely to underestimate the true values.

1.2 Results

1.2.1. Aerial survey data

Over 10 years of aerial surveys (2013 to 2025) in block A, a total of 56,063 km of search effort along the transect lines were included. The average length of the created segments was 16.5 km (with a range from 0.14 to 30 km).

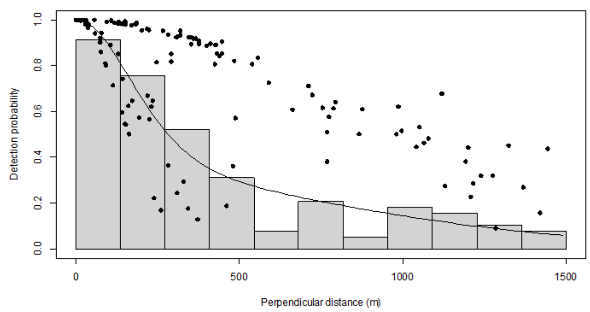
The number of detected schools included (i.e. adult schools within the truncation distance) was 129.

A half-normal detection function that included the logarithm of school size was selected (Figure 1.1 and see table below). The average probability of detection of a BFT school was 0.31 (CV=0.11). The quantile-quantile plot (Fig. 1.1) and Cramér-von Mises goodness-of-fit test did not indicate a lack of fit of the selected model (test statistic = 0.15, *p*-value=0.38).

The table below shows AIC values from fitting detection functions with different key functions and covariates. 'None' indicates that a conventional distance sampling function was fitted. Bold typeface indicates the selected model.

Covariates	Half-normal	Hazard rate
log(School size)	**1754.03**	1763.27
log(School size) + Sea state	1758.14	1767.76
log(School size) + Year	1761.99	1764.35
log(Biomass)	1774.66	1774.76
Year	1812.66	1791.39
Sea state	1817.19	1800.50
None	1823.07	1800.90

Company	1824.56	1804.20
Airplane	1822.57	1802.80
Haze	1826.80	1804.08



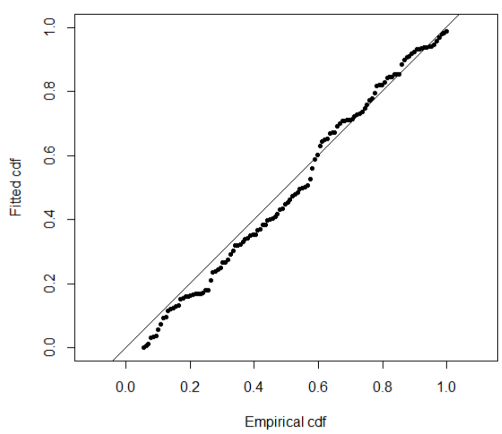
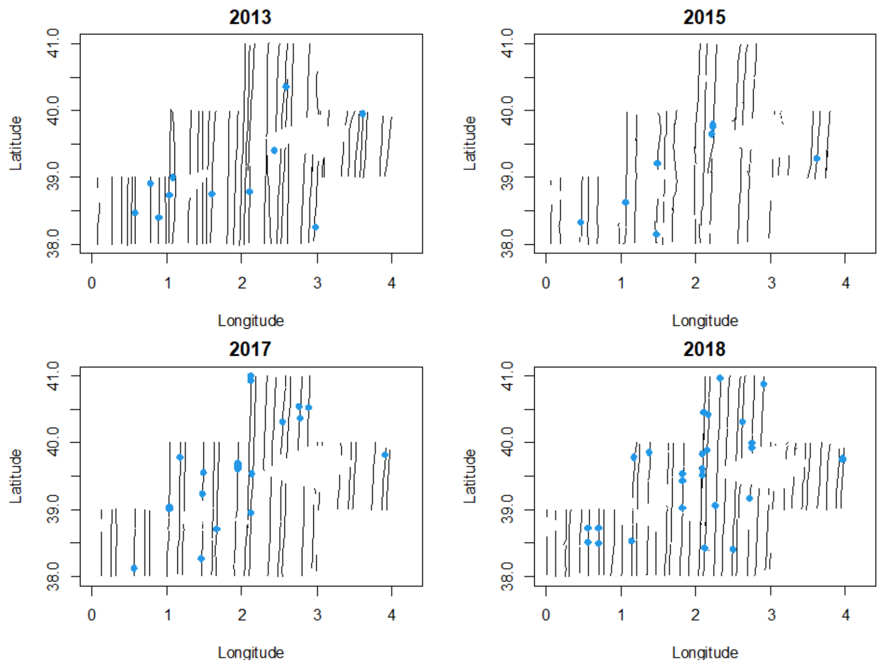
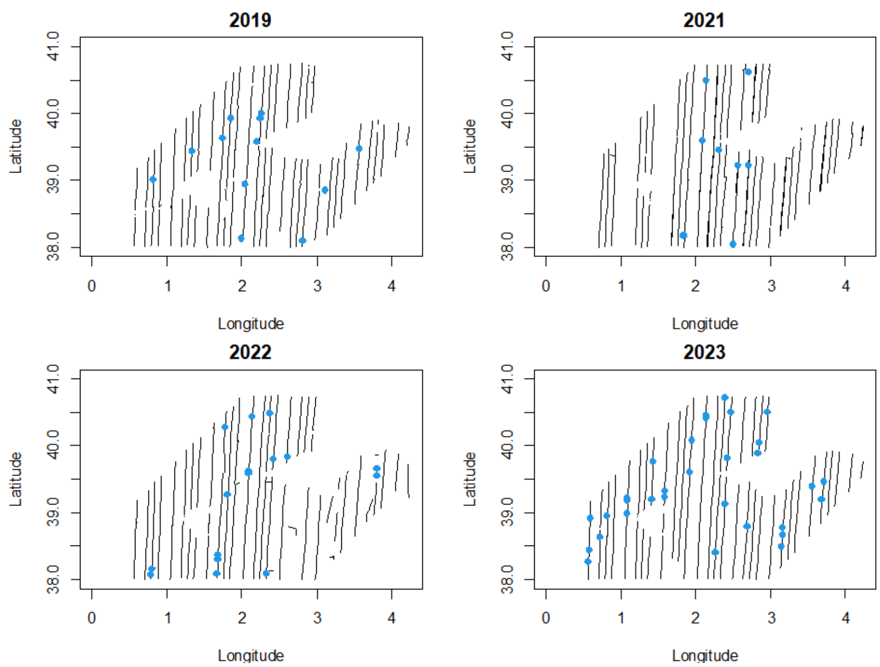


Figure 1.1. Top plot: Detection function averaged over all school sizes (line) overlaid onto a scaled histogram of perpendicular distances: the dots indicate individual schools. Bottom plot: Quantile-quantile plot for the selected detection function (see Table C1) used as a graphical check of the goodness of fit of the selected function. Points from a well-fitting model should lie along the straight line without substantial deviation.

The location of segments and detected school is shown in Figure 1.2. There were 122 segments which contained detections out of a total of 3,408 segments (3.6%).


2013
2015
2017
2018
Latitude
Longitude


2019
2021
2022
2023
Latitude
Longitude

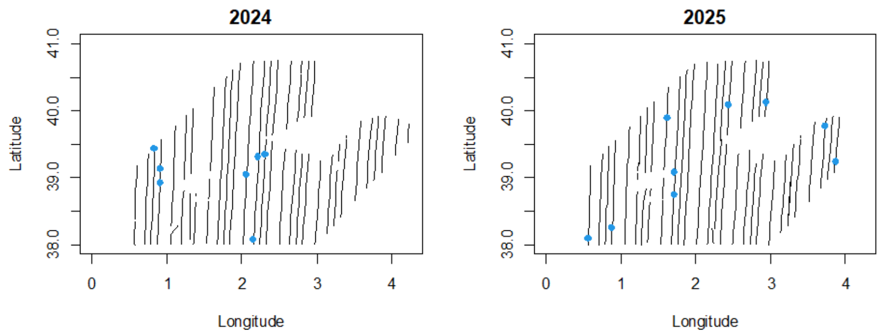


Figure 1.2. Location of transects (black lines) and midpoints of segments where schools of BFT were detected (blue dots).

1.3 Discussion

To avoid discarding detections when the information was incomplete, median values were substituted for missing values. An alternative approach may be to simply discard the detection. The sensitivity of the results to the approach taken compared to alternative approaches was not tested as part of the project due to time constraints but could be done in future.

Section 2

2.1. Background and objectives of the environmental modelling approach

The Balearic Sea is one of the principal spawning areas of the eastern Atlantic and Mediterranean stock of Atlantic bluefin tuna (Thunnus thynnus; BFT). BFT spawning distribution in this region is closely associated with sea surface temperature, salinity, water-mass boundaries and mesoscale frontal structures (Alemany et al., 2010; Reglero et al., 2012; Alvarez-Berastegui et al., 2014, 2016). Spatial and interannual variability in these conditions may therefore alter the location and concentration of spawning aggregations within the survey area.

Environmental analyses conducted during GBYP Phase 14 characterized this variability both across the wider reproductive area and within the segments sampled by the aerial surveys (Alvarez-Berastegui et al., 2025a). The analyses revealed marked interannual differences in temperature, salinity, chlorophyll-a, mixed-layer depth and frontal indicators, including changes in the position of fronts associated with recent and resident Atlantic waters and in the timing at which temperatures relevant to BFT reproduction were reached. Surveyed segments were also partially differentiated environmentally among years, indicating that successive surveys did not always sample the spawning area under equivalent conditions.

Annual variation in aerial-survey observations may consequently reflect both changes in spawning-stock abundance and environmentally driven changes in the distribution, aggregation and surface availability of BFT schools. Differences in survey timing, spatial coverage, environmental conditions and segment length may therefore affect the comparability of annual estimates (Lo et al., 1992; Thorson et al., 2015). Model-based standardization addresses this issue by estimating annual indices while accounting for observed environmental, spatial and sampling variability (Maunder and Punt, 2004; Thorson et al., 2015). Predictions under common reference conditions are intended to reduce variability attributable to the factors included in the models, although they cannot remove all environmental effects or ensure that the remaining variation is exclusively related to stock biomass. This is particularly relevant because standardized abundance indices are a primary input to the ICCAT management procedure for Atlantic bluefin tuna (ICCAT, 2022).

The objective of this section is to develop an environmentally standardized model-based index of BFT abundance and biomass for the Balearic Sea aerial surveys. Given the high proportion of segments without BFT observations, the reference approach uses a delta-lognormal formulation that models occurrence probability and positive biomass separately before combining them into an annual index (Lo et al., 1992; Stefánsson, 1996). This structure follows the approach used for the standardized BFT larval index in the Balearic Sea, based on binomial and lognormal GAM components, environmental covariates and estimated marginal means (Alvarez-Berastegui et al., 2025b). Additional Tweedie and Bayesian models were developed to compare their performance, robustness and potential advantages. The analyses also evaluate alternative treatments of segment length and survey effort, the inclusion of earlier surveys and the consistency of annual indices across modelling approaches.

2.2. Methods

2.2.1. Preparation of the integrated analytical dataset

The analysis was based on a segment-level dataset (obtained from the original transects of the aerial survey) for the Balearic Sea aerial surveys conducted between 2013 and 2025. The dataset included the detection-corrected estimates of BFT abundance and biomass produced by CREEM (see section 1 for details), survey and observation variables, segment coordinates and the environmental information assigned to each segment during this work. The distributions of segment length, sampling conditions and environmental variables were examined by year and compared between segments with and without BFT observations.

Quality-control filters were applied to restrict the analysis to segments surveyed under comparable operational and observation conditions. Segments were removed if sampling or environmental or operational covariates present extreme values that never presented positive abundances of tuna at any year. Application of these criteria reduced the dataset from 3408 to 2,612 segments. The filters applied where fly altitude >250; Turbidity >1; Segment length >=0.5 and <=29.7; glint<1 and haze <=3.

Several variables were subsequently derived for analysis. BFT presence was defined as 1 when the estimated number of individuals was greater than zero and 0 otherwise. Segment biomass density, expressed as tonnes km^{-2}, was calculated by dividing the detection-corrected biomass by the sampled area, assuming a total effective strip width of 3 km. Time of day was converted to a fraction of a 24-hour day.

The final analytical dataset contained year, sampling date and time, survey conditions, segment length and position, detection-corrected abundance and biomass, biomass density, BFT presence and the selected environmental covariates (see specific section below for details

2.2.2. Environmental and sampling covariates

Candidate covariates were selected according to their potential influence on BFT spawning distribution, surface availability or aerial-survey observations. Environmental data were obtained from the repository developed during GBYP Phase 14, using products from EMODnet and the Copernicus Marine Service. The full description of the data sources, spatial and temporal resolutions, extraction procedures and derived products is provided in Alvarez-Berastegui et al. (2025a).

The environmental covariates assigned to each segment included bathymetry, seafloor slope, sea surface temperature (SST), SST anomaly, chlorophyll-a concentration, sea surface salinity and mixed-layer depth. Bathymetry and slope describe static geographical and topographic conditions potentially related to circulation and spawning-habitat location. Temperature, salinity, chlorophyll-a and mixed-layer depth characterize the thermal, hydrographic and productivity conditions experienced during each survey.

Derived variables were included to represent dynamic oceanographic processes not fully described by instantaneous values. The spatial gradient of sea surface salinity was used as an indicator of water-mass boundaries and frontal structures, while the SST increase during the preceding 15 days represented recent seasonal warming. Because SST was strongly related to survey date, a seasonally detrended temperature variable (residual temperature: “temp_res”) was calculated from the residuals of a GAM relating SST to Julian day. This variable represents whether a segment was warmer or cooler than expected for its sampling date.

Sampling covariates included Julian day, time of day, segment position as Lon Lat, segment length. Other covariates associated with the sampling (sea state, turbidity and haze) were not used in these models as they were applied during the distance-based calibration process (see details in section 1) or because they are not compiled equally along the time series. Candidate covariates were examined for interannual coverage, missing values, pairwise relationships and biological interpretability before model formulation.

2.2.3. General model-based standardization framework

The model-based analyses were designed to estimate annual BFT indices while accounting for environmental, spatial and time differences among surveys. The primary analysis used data from 2017–2025 (excluding 2020 because no survey was conducted). This period was selected because survey protocols and data processing were sufficiently comparable over years and because it allowed direct comparison with the design-based strict-update series. Various models were subsequently fitted after adding the 2013 and 2015 surveys to evaluate the feasibility of extending the index backwards. Surveys conducted in 2010 and 2011 were excluded because substantial methodological differences prevented reliable comparison with the later period (see reports from GBYP, aerial survey documents).

Detection-corrected abundance, biomass (in tonnes) and density (tonnes per km^2) at segment level were evaluated as response variables. Because most segments contained no BFT observations and positive values were strongly right-skewed, the reference delta-lognormal approach modelled occurrence probability using a binomial GAM and positive biomass using a lognormal GAM. Predictions from both components were combined to estimate expected annual biomass. This modelling approach was the one specified in the contract but additional GAM with a Tweedie distribution and a Bayesian model were also developed.

Year was included as a categorical factor, while environmental and spatial variability was represented by selected covariates and penalized smooth functions of restricted complexity. Spatial effects were evaluated using segment coordinates. Model formulation considered biological interpretability, diagnostics, predictive performance and the stability of annual estimates.

Survey effort was represented using sampled area, calculated as segment length multiplied by a 3-km effective strip width. Alternative formulations included sampled area as an offset, segment length as a model covariate, or pre-calculated density as the response. These alternatives were compared because segment lengths were heterogeneous and could have strong effect on the standardization of biomass when calculating densities (biomass/km^2). For the delta-lognormal and Tweedie models, annual estimates were obtained using the emmeans package (Lenth, 2020), applying common environmental, spatial and effort reference conditions to all years. Predictions were calculated on the corresponding model scale and back-transformed to the response scale when transformations or non-identity link functions were used.

For the delta-lognormal model, the standardized annual index was calculated as the product of the estimated probability of presence and the back-transformed positive-biomass estimate. The latter was obtained by exponentiating the annual estimate from the log-scale positive component and therefore represents a conditional geometric mean of positive biomass. When the positive component was standardized to a reference sampled area of 1 km², the resulting combined index was expressed on a biomass-density scale (t km^{-2}) and interpreted as a standardized relative biomass index.

Standard errors for the back-transformed positive-biomass estimates were approximated using the first-order delta method. Uncertainty from the occurrence and positive-biomass components were subsequently combined using a first-order delta-method approximation. Annual standard errors, coefficients of variation and approximate confidence intervals were calculated for the combined index.

For the Bayesian model, annual posterior conditional estimates were obtained under common environmental, spatial and effort reference conditions. These estimates and their uncertainty were summarized using posterior point estimates and 95% credible intervals.

Summary of models developed

Three modelling approaches were developed and evaluated:

- **Delta-lognormal model (reference model):** Biomass (tonnes) was modelled using a two-component delta approach, combining a presence–absence component with a positive-biomass component. The area surveyed was incorporated through an offset, thereby standardizing biomass with respect to sampling effort. Three alternative formulations for the offset were tested before selecting the final reference formulation. The offset=log(area) was finally selected.
- **Tweedie model:** Biomass density was modelled directly as the response variable using a Tweedie distribution, allowing zero observations and positive continuous biomass values to be represented within a single model. Different approaches for incorporating sampling effort and standardizing the annual estimates were evaluated including weighting in some cases.
- **Bayesian model:** Biomass (tonnes) was modelled within a Bayesian framework, with surveyed area incorporated as an offset = log(area), for direct comparison with the reference delta model. Annual standardized indices and their uncertainty were obtained from the posterior predictive distributions rather than through marginal means and delta-method uncertainty propagation.

2.2.4. Reference delta-lognormal model

The “reference” index (as established in the contract) was estimated using a two-component delta-lognormal GAM fitted with the mgcv package (Wood 2017). The first component modelled BFT occurrence in each segment, defined as presence when the detection-corrected number of individuals was greater than zero. A binomial distribution with year included as a categorical factor and selected environmental, spatial and sampling variables represented by penalized smooth functions. Different link functions were tested (see table 1), the most appropriate resulted in the model with logit link and treated segment length as a flexible covariate.

The second component was fitted only to segments with BFT presence and modelled the logarithm of detection-corrected positive biomass using a Gaussian distribution. Year was again included as a categorical factor, together with the environmental and spatial covariates supported by the positive observations. The standardization to differences in sampled areas were accounted for through an effort offset formulation, but the potential bias on segment length was corrected by including that variable as smoother also.

Candidate covariates were selected according to ecological relevance, data coverage, correlation with other predictors, model diagnostics and stability of the estimated annual effects. Smooth complexity was restricted to reduce overfitting, and models were fitted using restricted maximum likelihood. The binomial component was evaluated exploring fitted-value distributions at real presences and real absences, and the area under the ROC curve. The positive component was assessed from residual distributions, residuals against fitted values, response–fitted relationships and Q-Q plots (gam.check diagnostics).

Annual marginal estimates for both components were obtained using emmeans under common environmental, spatial and effort conditions. Predictions were back-transformed to the response scale where required. The annual delta index was calculated as the product of the predicted marginal probability of presence and the predicted marginal positive biomass:

eq 2: $I_Y = \hat{p}_Y \times \hat{\mu}_Y$,

where $\hat{p}_Y$ is the standardized probability of BFT presence and $\hat{\mu}_Y$ is the standardized mean biomass conditional on presence in year *y*.

Uncertainty from both components was combined using a first-order delta-method approximation. The model was first fitted to the comparable 2017–2025 surveys and subsequently refitted after including 2013 and 2015 to evaluate the extended time series.

2.2.5. Alternative treatments of survey effort and segment length

Segment length varied among observations because of survey geometry and interruptions in search effort, including deviations from the transect to confirm BFT sightings. Consequently, shorter segments were not necessarily a random subset of the survey and could be associated with BFT presence. Several approaches were therefore evaluated to account for effort and assess the sensitivity of the annual indices.

First, the logarithm of sampled area was included as an offset in models of biomass. This formulation assumes that the expected response increases proportionally with sampled area, with the corresponding coefficient fixed at one. Predictions for a reference area can consequently be interpreted as abundance or biomass per unit area.

Second, segment length was included as a penalized smooth covariate. This approach allows a nonlinear relationship between segment length and the response, including departures from strict proportionality. However, it does not directly convert total abundance or biomass into density, therefore this option was combined with the previous use of area as offset.

Third, biomass density, calculated as detection-corrected biomass divided by sampled area, was modelled directly without an effort offset. This provides predictions in tonnes km^{-2} but may increase variability for short segments. The resulting indices were compared to determine whether their temporal patterns were robust to effort treatment. Particular attention was given to differences in absolute magnitude and units, because indices derived from offsets, reference segment lengths and density responses are not directly comparable unless expressed on a common scale.

2.2.6. Evaluation of environmental standardization

The effect of environmental standardization was evaluated by comparing the reference delta-lognormal model with a baseline model fitted without environmental, and location (lat,long) covariates. Both formulations used the same observations, annual factors, spatial structure and treatment of sampling effort, allowing differences in the resulting indices to be attributed primarily to the environmental terms.

Annual marginal estimates were obtained with emmeans and compared on both their original and normalized scales. Changes in absolute magnitude, relative temporal trends, uncertainty and the ranking of individual years were examined. The resulting series were also compared with the design-based strict-update index to assess whether environmental standardization modified or preserved the main interannual pattern.

The contribution of environmental covariates was assessed from their partial effects, uncertainty, statistical support and influence on model fit, AUC, explained deviance and residual diagnostics. Index stability was evaluated across alternative covariate combinations and from changes in annual estimates, confidence intervals and coefficients of variation. Particular attention was given to years whose estimates were especially sensitive to environmental standardization or limited positive observations.

2.2.7. Extension of the standardized time series

The primary standardized index was developed using the 2017–2025 surveys, excluding 2020 because no aerial survey was conducted. This period was considered the most homogeneous in terms of survey design, observation protocols, data recording and processing, and therefore provided the most reliable basis for model development and comparison with the strict-update series.

The possibility of extending the index backwards was evaluated by adding the 2013 and 2015 surveys and refitting the same model structures. Model diagnostics, covariate effects and annual estimates for the overlapping 2017–2025 period were compared between the primary and extended datasets. This allowed assessment of whether incorporating the earlier surveys altered model performance or the temporal pattern estimated for the methodologically comparable period.

The 2010 and 2011 surveys were evaluated but excluded from the standardized models because of substantial differences in survey protocols, aircraft observation conditions, distance measurements and the recording and estimation of school size and biomass. These differences could not be adequately represented by the available covariates and would introduce additional uncertainty into the annual index.

The extended series was therefore treated as an exploratory analysis rather than as a direct replacement for the homogeneous 2017–2025 index. Its interpretation considers the remaining differences in sampling design, data quality and observation procedures between survey periods.

2.2.8. Alternative modelling approaches

Alternative models were developed to assess whether the estimated temporal pattern depended on the response variable, probability distribution or statistical framework. These models used the same integrated dataset and were standardized under comparable environmental and sampling conditions whenever possible.

A Tweedie GAM was first fitted to detection-corrected abundance. The Tweedie distribution accommodates exact zeros and continuous positive values within a single model, avoiding the need to combine separate presence and positive components. Survey effort was represented through an offset or a common reference segment length, and annual marginal estimates were obtained using emmeans.

A second Tweedie GAM used pre-calculated density as the response. Because abundance or biomass had already been divided by sampled areas, no additional effort offset was required. This formulation provided annual estimates directly on a density scale and allowed evaluation of whether modelling density produced a more stable index than modelling total abundance with an effort correction. Various configurations were tested for the weighting of the sampled area in the different segments (see appendix SR.1)

A Bayesian model was also developed to represent parameter and prediction uncertainty through posterior distributions. Environmental, spatial, annual and effort effects were incorporated within a common probabilistic framework, and standardized annual indices were obtained from posterior predictions under common reference conditions. Posterior means and credible intervals were used to summarize annual estimates. The resulting indices were compared descriptively in terms of their temporal patterns, uncertainty and practical interpretation, without attempting a formal comparison between the frequentist and Bayesian frameworks.

2.3. Results

2.3.1. Data filtering and characteristics of the analytical dataset

The initial dataset comprised 3,408 survey segments distributed across ten aerial surveys conducted between 2013 and 2025. No surveys were undertaken in 2014, 2016 or 2020. Application of the sampling-quality and segment-length criteria retained 3,171 segments, corresponding to 93.0% of the initial records. Annual retention ranged from 84.8% in 2021 to 98.3% in 2013 (Table 1).

Examination of the sampling-variable distributions showed that observations recorded outside the selected ranges (values with no positive records any year) of altitude, segment length, glint and aircraft speed corresponded exclusively to segments without BFT detections (Figure S1). When evaluated independently, 60 segments occurred outside the selected altitude range, 47 outside the segment-length values defined in the filtering process 16 exceeded the maximum aircraft-speed threshold and two exceeded the accepted glint level. No segments exceeded the selected turbidity threshold. Because some records may have failed more than one criterion or contained missing values, these filter-specific counts should not be added to obtain the total number of excluded segments.

The filtered dataset contained 122 positive segments, representing 3.85% of the observations retained for the extended 2013–2025 analysis. The primary 2017–2025 dataset contained 2,612 segments, of which 107 were positive, corresponding to an overall occurrence rate of 4.10%. Annual occurrence rates in this primary period ranged from 2.20% in 2021 to 7.47% in 2023, with relatively high values also recorded in 2018 (7.16%). These results demonstrate the strong predominance of zero observations and support the use of modelling approaches capable of explicitly accommodating zero inflation.

2.3.2. Nominal patterns in BFT biomass density

Nominal biomass density showed substantial interannual variability over the primary 2017–2025 survey period (Figure 1; Table 2). The highest annual mean segment-level density was observed in 2018 (0.674 t km^{-2}), followed by 2023 (0.430 t km^{-2}) and 2019 (0.374 t km^{-2}). Intermediate estimates were obtained for 2022 (0.298 t km^{-2}) and 2021 (0.213 t km^{-2}), whereas comparatively low values were recorded in 2017 (0.140 t km^{-2}), 2025 (0.141 t km^{-2}) and particularly 2024 (0.044 t km^{-2}). No estimate was available for 2020 because no aerial survey was conducted that year.

Uncertainty was high in all years, with annual coefficients of variation ranging from 0.40 to 0.67. The widest relative uncertainty occurred in 2019 and 2024, while the lowest coefficient of variation was obtained in 2023. The asymmetric uncertainty intervals overlapped substantially among most years, indicating that the nominal observations provided limited support for interpreting individual annual differences independently of sampling, environmental and spatial variability. Overall, the nominal series displayed pronounced fluctuations but there was no monotonic temporal trend.

2.3.3. Reference delta-lognormal model

Occurrence component

The occurrence component was fitted to 2,612 survey segments, of which 107 contained BFT observations. Temperature residuals, segment length and bathymetry all contributed significantly to explaining the probability of BFT occurrence ($p < 0.01$; Figure 2) (See also model summary in table S1). Occurrence probability generally increased from negative to positive temperature residuals. The probability of occurrence decreased with increasing segment length up to approximately 20 km, whereas the bathymetric response indicated higher occurrence at intermediate depths and lower occurrence toward both the deepest and shallowest areas. Uncertainty increased toward the extremes of the covariate ranges, where observations were less abundant.

The occurrence model had an AUC of 0.704, indicating moderate discrimination between segments with and without BFT observations.

Standardized annual occurrence probabilities ranged from 0.024 to 0.103 (Figure 3; Table S2). The highest probability was estimated for 2018 (0.103), followed by 2023 (0.079) and 2019 (0.077). Lower probabilities were obtained for 2021 (0.033), 2022 (0.029) and 2025 (0.024). The confidence intervals overlapped for several years, indicating substantial uncertainty in the relative annual differences.

Positive-biomass component

The positive-biomass component was fitted to the 107 survey segments containing BFT observations. Surveyed area was included as a log-offset, while segment length was additionally modelled as a smooth effect to account for departures from a strictly proportional relationship between sampled area and observed biomass. Segment length, temperature residuals and salinity were all associated with BFT biomass (Figure 4; Table S3).

After accounting for the area offset, biomass generally decreased with increasing segment length up to approximately 20 km. Biomass increased from negative temperature residuals to a maximum at moderately positive values. The salinity response showed comparatively higher biomass at lower salinities and a progressive reduction toward the highest salinity values. Uncertainty increased at the extremes of the covariate ranges, where fewer observations were available.

The model explained 36.7% of the deviance, with an adjusted R^2 of 0.161. Diagnostic plots indicated some departures from the ideal Gaussian residual pattern, particularly in the tails and in the relationship between residuals and fitted values (Figure S3). These results therefore indicate that the model captured a substantial component of the variation among positive observations, while considerable unexplained variability remained.

Annual marginal estimates were obtained on the log scale under common sampling and environmental reference conditions and subsequently back-transformed to the original scale. The highest standardized positive-biomass estimate was obtained for 2022 (4.81 t km^{-2}), followed by 2018 (3.21 t km^{-2}), 2021 (2.72 t km^{-2}) and 2025 (2.46 t km^{-2}; Figure 5; Table S4). Lower estimates were obtained for 2017 (0.57 t km^{-2}) and 2023 (0.63 t km^{-2}). Uncertainty was substantial, particularly in 2021, 2022, 2024 and 2025, and the confidence intervals overlapped widely among years.

Combined delta-lognormal biomass index

The standardized biomass index showed substantial interannual variability, with a pronounced maximum in 2018 (0.329 t km^{-2}) and a decreasing trend thereafter (Figure 6;

Table 3). The next highest estimates were obtained in 2022 (0.140 t km^{-2}) and 2019 (0.138 t km^{-2}). The mean standardized biomass index was approximately 50% lower during 2023–2025 than during 2019–2022.

The lowest index value was observed in 2017 (0.025 t km^{-2}), followed by 2023 (0.050 t km^{-2}) and 2025 (0.059 t km^{-2}). In 2025, the low combined index resulted mainly from the lowest estimated occurrence probability, despite relatively high biomass in positive segments. Thus, interannual variation in the standardized index reflected differing contributions from the occurrence and positive-biomass components.

Uncertainty was considerable, with coefficients of variation ranging from 0.54 in 2018 to approximately 1.01 in 2024 and 2025. Confidence intervals overlapped widely among years, particularly for 2021, 2022, 2024 and 2025. Consequently, differences among individual annual estimates should be interpreted cautiously. No index was available for 2020 because no aerial survey was conducted that year

2.3.4. Sensitivity of the biomass index to environmental and spatial standardization

Excluding environmental and spatial covariates reduced the performance of both components of the delta-lognormal model. In the binomial component, deviance explained decreased from 6.54% to 3.57%, adjusted R^2 from 0.025 to 0.011, and AUC from 0.704 to 0.655. The reduced model also produced a narrower range of fitted probabilities and greater overlap between segments with and without observed BFT presence (Figures S2 and S4). Segment length remained significantly associated with occurrence ($p = 0.010$).

The positive-biomass component also presented a reduced performance when environmental and spatial covariates were excluded. Deviance explained decreased from 36.7% in the reference model to 23.2% in the reduced model, adjusted R^2 declined from 0.161 to 0.024, and residual variance increased from 3.634 to 4.228. Segment length remained strongly supported ($p < 0.001$), confirming the importance of retaining this sampling correction.

Environmental and spatial standardization also substantially modified the magnitude and ranking of the annual biomass estimates (Figure 7). Without these covariates, the highest index occurred in 2022 (0.123 t km^{-2}), whereas the reference model produced a pronounced maximum in 2018 (0.329 t km^{-2}). Standardization increased the 2018 and 2019 estimates by more than threefold and increased the 2024 estimate from 0.019 to 0.074 t km^{-2}. Conversely, the 2023 estimate decreased from 0.073 to 0.050 t km^{-2}. Estimates for 2017 and 2025 were comparatively stable between formulations.

Overall, the reduced model did not preserve the ranking or relative magnitude of several annual estimates and showed poorer explanatory and discriminatory performance. These results support the inclusion of environmental and spatial covariates to derive annual indices under common reference conditions.

2.3.5. Extended biomass index incorporating 2013 and 2015

Extending the delta-lognormal analysis to include the 2013 and 2015 surveys substantially changed the overall temporal interpretation of the biomass index (Figure 8). The highest estimate in the extended series occurred in 2013 (0.874 t km^{-2}), more than four times the 2018 estimate obtained from the same model (0.211 t km^{-2}). The series therefore changed from one characterized by a maximum in 2018 to one showing a very high initial estimate followed by generally lower values, although moderate increases remained evident in 2018 and 2022.

The 2013 estimate was associated with substantial model-based uncertainty (SE = 0.863; CV = 0.99; 95% CI: 0.173–4.400 t km^{-2}). This uncertainty was not caused by unusually dispersed observed positive densities: 2013 had the lowest annual CV of observed positive density and a comparatively restricted density range. Instead, its nine positive segments covered a narrow range of segment lengths (18.2–22.3 km) and occurred under strongly negative temperature residuals (−3.09 to −0.36). The common reference values used for standardization—approximately 15.4 km for segment length and 0.25 for temperature residuals—fell outside both ranges observed in 2013.

Consequently, the 2013 estimate required extrapolation toward shorter segments and warmer residual conditions, both associated with higher predicted positive biomass. This explains, at least partly, the marked difference between the observed mean positive density in 2013 (3.42 t km^{-2}) and its standardized estimate (15.61 t km^{-2}), as well as the large associated uncertainty. The 2013 estimate should therefore be interpreted cautiously because it is strongly influenced by limited covariate overlap with the other surveys.

Adding the earlier surveys also reduced the absolute estimates for all years from 2017 onwards by approximately 20–49%, although the main relative pattern within 2017–2025 was largely preserved. These results demonstrate the sensitivity of the delta-lognormal index to differences in the distribution of segment length and environmental conditions among years. They also highlight the value of developing a complementary analysis based on more homogeneous segment definitions and improved covariate overlap among surveys, thereby reducing the dependence of annual estimates on model-based extrapolation.

2.3.6. Alternative modelling approaches, Tweedie and Bayesian: common patterns and limitations

The alternative models presented in Supplementary Reports SR1 and SR2 provided complementary perspectives on the challenges involved in standardizing the aerial-survey biomass index. Because they differed in their response variables, probability distributions, treatment of zeros and representation of survey effort, their fit statistics and absolute index values were not compared directly. The comparison instead focused on temporal patterns, environmental effects and major sources of uncertainty.

For 2017–2025, the three approaches identified 2018 as a year of comparatively high biomass and 2017 as one of the lowest. In the Bayesian hurdle-lognormal model, the highest posterior estimate occurred in 2018, closely followed by 2021, with intermediate estimates for 2019, 2022 and 2023 and lower values in 2024 and 2025. However, the broad and strongly overlapping credible intervals prevented a robust ranking of most years. The hurdle component indicated a lower probability of zero biomass in 2018 and 2023 than in 2017, but a higher probability in 2025, whereas biomass conditional on detection was higher in 2021 and 2022. This demonstrates how different combinations of occurrence and positive biomass can generate similar annual index values.

The delta-lognormal and Tweedie models also supported the elevated estimate for 2018, although the ranking of other years varied among formulations. The Tweedie results were particularly sensitive to whether segment length was included as a smooth effect, an analytical weight or both. Therefore, the models supported a partially consistent broad temporal signal, while individual annual estimates remained dependent on the response distribution and treatment of survey effort.

Environmental and spatial heterogeneity contributed to all three approaches, although the variables retained were not identical. Temperature residuals, bathymetry and spatial structure were repeatedly selected, and the Bayesian model also retained the spatial salinity gradient. Mean salinity was included in some frequentist formulations but removed from the final Bayesian model during the “leave one out” based selection. Segment length was influential in the frequentist models, whereas the Bayesian model represented effort through an area offset in both hurdle components. These differences reinforce the importance of segment definition and effort treatment.

Each approach addressed different aspects of the data. The delta-lognormal model provided an ecologically interpretable separation of occurrence and positive biomass but required uncertainty to be combined across two separately fitted models. The Tweedie model accommodated zeros and positive biomass within a single distribution but remained sensitive to effort specification and the highly skewed response. The Bayesian hurdle-lognormal model jointly represented both components and propagated uncertainty through the posterior distribution. Although convergence and posterior predictive diagnostics were satisfactory, its moderate Bayesian (R^2) of approximately 0.26, high residual variation and broad credible intervals indicated that much of the variability remained unexplained.

Overall, the models supported a broad pattern of relatively high biomass in 2018 and decreased trend along following years (with no clear significant pattern along the time series). Sparse positive observations, extreme zero prevalence, heterogeneous segment lengths and limited environmental overlap among years remained the main sources of uncertainty, supporting further analyses based on more homogeneous segmentation and consistent treatment of survey effort and environmental conditions.

2.4. Discussion

This study aimed to develop a model-based framework for standardizing the GBYP aerial-survey biomass index while accounting for differences in sampling effort and environmental conditions among years. A reference delta-lognormal model was developed for the 2017–2025 period and evaluated against complementary Tweedie and Bayesian approaches. The models identified a partly consistent interannual pattern, particularly the comparatively high biomass in 2018 and low biomass in 2017, while also showing that the magnitude and ranking of some annual estimates depended on model formulation.

The data present several important analytical challenges. Positive BFT observations represented only about 4% of the retained segments, resulting in strong zero inflation and few positive observations for estimating annual biomass effects. Positive biomass was also strongly skewed, and segment lengths varied within and among surveys. Moreover, segment length may be partly influenced by survey operations because transects can be divided into shorter segments in areas with more observations. This explains why the treatment of segment length as an offset, smooth effect or weight substantially affected model results.

Environmental standardization was nevertheless essential. Removing environmental and spatial covariates reduced the performance of both delta-lognormal components and lowered the AUC of the occurrence model from 0.704 to 0.655. It also changed the magnitude and ranking of several annual estimates. Temperature residuals, salinity, bathymetry and spatial structure contributed to more than one modelling approach, demonstrating that annual differences in survey conditions can affect the observed index and potentially obscure changes in stock biomass.

Standardization becomes less reliable when the conditions observed in a particular year do not overlap with the common reference conditions. In 2013, positive observations were restricted to relatively long segments and strongly negative temperature residuals. The standardized estimate therefore required extrapolation towards shorter segments and warmer conditions that were not represented in that survey, producing a high and highly uncertain estimate despite the comparatively low dispersion of the observed densities.

These results highlight two priorities for future development. First, transects should be divided into more homogeneous and comparable segments, with spatial supports better matched to the oceanographic structures of the Balearic Sea, which may operate at scales greater than the current 20 km target length. Second, continued regional warming may generate environmental conditions without historical analogues. Future indices may therefore require environmental descriptors based on evolving thermal baselines and spatiotemporal or hierarchical models capable of representing non-stationary relationships between BFT distribution and SST.

2.5. Conclusions

The reference delta-lognormal model provided an environmentally and spatially standardized relative biomass index for 2017–2025, with a partly consistent temporal signal across alternative approaches. The biomass index shows a decreasing trend from 2018 to 2025 but with high levels of interannual variability.

Strong zero inflation, few positive segments, skewed biomass and heterogeneous segment lengths remain the principal sources of uncertainty.

Environmental and spatial covariates improved model performance and materially affected annual estimates, confirming that environmental standardization is necessary.

Extending the index to earlier surveys is limited by weak overlap in sampling protocols and environmental conditions, which can force extrapolation and produce highly uncertain estimates, as observed for 2013. Revising transect segmentation to improve comparability among years, potentially reduce the number of zero-biomass segments, and refining the selection of environmental covariates may help address this limitation.

A separate challenge is the ecological scale and temporal stability of environmental relationships. Future work should evaluate environmental descriptors at spatial scales greater than the current 20 km segments and develop modelling frameworks capable of accommodating shifting SST baselines and increasingly novel environmental conditions.

2.6 References

Alemany, F., Quintanilla, L., Vélez-Belchí, P., García, A., Cortés, D., Rodríguez, J. M., Fernández de Puelles, M. L., González-Pola, C., & López-Jurado, J. L. (2010). Characterization of the spawning habitat of Atlantic bluefin tuna and related species in the Balearic Sea (western Mediterranean). Progress in Oceanography, 86(1–2), 21–38. https://doi.org/10.1016/j.pocean.2010.04.014

Alvarez-Berastegui, D., Ciannelli, L., Aparicio-González, A., Reglero, P., Hidalgo, M., López-Jurado, J. L., Tintoré, J., & Alemany, F. (2014). Spatial scale, means and gradients of hydrographic variables define pelagic seascapes of bluefin and bullet tuna spawning distribution. PLoS ONE, 9(10), e109338. https://doi.org/10.1371/journal.pone.0109338

Alvarez-Berastegui, D., Hidalgo, M., Tugores, M. P., Reglero, P., Aparicio-González, A., Ciannelli, L., Juza, M., Mourre, B., Pascual, A., López-Jurado, J. L., García, A., Rodríguez, J. M., Tintoré, J., & Alemany, F. (2016). Pelagic seascape ecology for operational fisheries oceanography: Modelling and predicting spawning distribution of Atlantic bluefin tuna in the western Mediterranean. ICES Journal of Marine Science, 73(7), 1851–1862. https://doi.org/10.1093/icesjms/fsw041

Alvarez-Berastegui, D., Tugores Ferrá, M. P., Burt, L., Paxton, C. G. M., Chudzińska, M., Ottmann, D., Alvarez, I., & Reglero, P. (2025a). Preparation of aerial survey data and environmental data to estimate abundance and biomass of bluefin tuna in the Mediterranean Sea using model-based analysis [Final report, ICCAT GBYP Phase 14]. International Commission for the Conservation of Atlantic Tunas. https://www.iccat.int/GBYP/DOCS/Aerial_Survey_Phase_14_2025_CREEM_IEO_Data_Analysis_Task_B.pdf

Alvarez-Berastegui, D., Tugores, M. P., Torres, A. P., Martín-Quetglas, M., Balbín, R., Santiago, R., Santadreu, M., Álvarez, I., Aranda, I., Calcina, N., Carreño, A., & Reglero, P. (2025b). Bluefin tuna larval abundances in the Balearic Archipelago spawning ground (western Mediterranean): Standardized time series from 2001 to 2024. Collective Volume of Scientific Papers ICCAT, 82(4), 1–21. SCRS/2025/207. https://www.iccat.int/Documents/CVSP/CV082_2025/n_4/CV082040207.pdf

Barón, J. M. S. and M. C. Burgos. 2011. ICCAT Bluefin tuna aerial survey (2011). Informe técnico prospecciones aéreas en Baleares. Available at Aerial_Survey_Phase_2_GAM_Area_1.pdf

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. (2001). Introduction to Distance Sampling. Estimating abundance of biological populations. Oxford University Press, Oxford.

Grup Air-Med (2010). ICCAT Bluefin tuna aerial survey 2010 Technical report: Aerial surveys explorations in the reproduction area of bluefin tuna (Thunnus thynnus) in Balearics waters. Available at Aerial_Survey_Phase_1_GAM_Areas_13.pdf

International Commission for the Conservation of Atlantic Tunas. (2022). Recommendation by ICCAT establishing a management procedure for Atlantic bluefin tuna to be used for both the western Atlantic and eastern Atlantic and Mediterranean management areas (Recommendation 22-09). https://www.iccat.int/Documents/Recs/compendiopdf-e/2022-09-e.pdf

Laake, J., D. Borchers, L. Thomas, D. Miller, and J. Bishop. (2022). mrds: mark–recapture distance sampling. R package version 2.2.8. https://cran.r-project.org/package=mrds.

Lenth R.V. (2020). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version, 1.5.2-1. https://CRAN.R-project.org/package=emmeans

Lo, N. C. H., Jacobson, L. D., & Squire, J. L. (1992). Indices of relative abundance from fish spotter data based on delta-lognormal models. Canadian Journal of Fisheries and Aquatic Sciences, 49(12), 2515–2526. https://doi.org/10.1139/f92-278

Marques, T. A., L. Thomas, S. G. Fancy, and S. T. Buckland. (2007). Improving estimates of bird density using multiple-covariate distance sampling. The Auk 124:1229-1243.

Marra, G., and Wood, SN. (2011). Practical variable selection for generalized additive models. Comput Stat Data an, 55 (7) (2011), pp. 2372-2387, 10.1016/j.csda.2011.02.004

Maunder, M. N., & Punt, A. E. (2004). Standardizing catch and effort data: A review of recent approaches. Fisheries Research, 70(2–3), 141–159. https://doi.org/10.1016/j.fishres.2004.08.002

Miller, D. L., E. Rexstad, L. Thomas, L. Marshall, and J. L. Laake. (2019). Distance Sampling in R. Journal of Statistical Software 89:1 - 28.

Paxton, C. G. M., C. S. Oedekoven, M. Chudzinska, M. Pilar Tugores Ferrá, and D. Alvarez-Berastegui. (2023). Design- and model-based inference to estimate density, abundance and distribution of BFT in the Mediterranean Sea. Report number CREEM-2023-04.

R Core Team. (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.R-project.org/

Reglero, P., Ciannelli, L., Alvarez-Berastegui, D., Balbín, R., López-Jurado, J. L., & Alemany, F. (2012). Geographically and environmentally driven spawning distributions of tuna species in the western Mediterranean Sea. Marine Ecology Progress Series, 463, 273–284. https://doi.org/10.3354/meps09817

Stefánsson, G. (1996). Analysis of groundfish survey abundance data: Combining the GLM and delta approaches. ICES Journal of Marine Science, 53(3), 577–588. https://doi.org/10.1006/jmsc.1996.0079

Thorson, J. T., Shelton, A. O., Ward, E. J., & Skaug, H. J. (2015). Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES Journal of Marine Science, 72(5), 1297–1310. https://doi.org/10.1093/icesjms/fsu243

Wood, S. N. (2017). Generalized additive models: An introduction with R (2nd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315370279

List of Tables in Section 2

List of Figures in Section 2

List of Supplementary Materials

Supplementary tables and figures.

Table S1. Summary of the binomial model.

Table S2. Summary of the lognormal model with positive abundances

Table S3. Summary of the positive-biomass component of the reference delta-lognormal mode1

Table S4. Back-transformed annual marginal estimates of the positive-biomass component of the reference delta-lognormal model.

Figure S1. Distributions of the sampling variables used to define the filtering criteria, separated according to observed BFT presence or absence.

Figure S2. Receiver operating characteristic curve and distribution of fitted occurrence probabilities for segments with and without observed BFT presence.

Figure S3. Diagnostic plots for the positive-biomass component of the reference delta-lognormal model.

Figure S4. Receiver operating characteristic curve and distribution of fitted occurrence probabilities for segments with and without observed BFT presence, for the binomial model with no environmental covariates

Supplementary report SR.1: Tweedie modelling approach with tuna densities.

Supplementary report SR.2: Bayesian modelling approach with tuna biomass.

Supplementary tables

Table 1. Annual summary of aerial-survey segments before and after application of the sampling filters.

Year	Initial segments	Retained segments	Retained (%)	Positive segments	Presence (%)
2013	347	341	98.3	9	2.64
2015	252	218	86.5	6	2.75
2017	317	311	98.1	14	4.50
2018	398	363	91.2	26	7.16
2019	351	336	95.7	11	3.27
2021	428	363	84.8	8	2.20
2022	315	303	96.2	9	2.97
2023	316	308	97.5	23	7.47
2024	296	286	96.6	7	2.45
2025	388	342	88.1	9	2.63
Extended series	**3,408**	**3,171**	**93.0**	**122**	**3.85**
Primary series	**2,809**	**2,612**	**93.0**	**107**	**4.10**

Table 2. Annual arithmetic means of nominal segment-level BFT biomass density, calculated across all surveyed segments, including zero observations. SE and CV refer to the annual means. Lower and upper limits correspond to approximate 95% lognormal confidence intervals derived from the variance of the annual mean. No survey was conducted in 2020.

Year	Segments	Nominal density (t km^{-2})	SE	CV	Lower 95% interval	Upper 95% interval
2017	311	0.140	0.068	0.487	0.057	0.345
2018	363	0.674	0.286	0.425	0.303	1.498
2019	336	0.374	0.251	0.673	0.113	1.238
2021	363	0.213	0.093	0.437	0.094	0.483
2022	303	0.298	0.146	0.489	0.120	0.739
2023	308	0.430	0.173	0.403	0.201	0.919
2024	286	0.044	0.027	0.604	0.015	0.131
2025	342	0.141	0.058	0.414	0.064	0.306

Table 3: Standardized annual BFT biomass index obtained from the reference delta-lognormal model.

year	n.samples	index	se_i	UCI	LCI	mu	ppos	cv_i
2017	311	0,025	0,019	0,092	0,007	0,565	0,044	0,743
2018	363	0,329	0,177	0,884	0,123	3,211	0,103	0,537
2019	336	0,138	0,108	0,536	0,036	1,796	0,077	0,783
2021	363	0,090	0,087	0,443	0,018	2,715	0,033	0,967
2022	303	0,140	0,120	0,597	0,033	4,807	0,029	0,853
2023	308	0,050	0,030	0,147	0,017	0,634	0,079	0,594
2024	286	0,074	0,074	0,381	0,014	1,402	0,053	1,006
2025	342	0,059	0,059	0,305	0,011	2,458	0,024	1,014

Supplementary figures

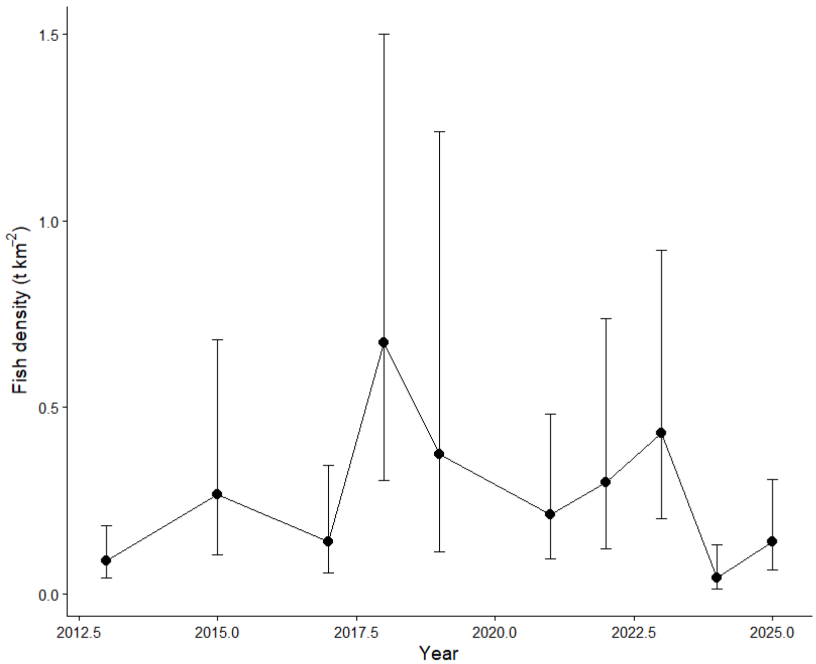


Figure 1. Annual arithmetic means of nominal segment-level BFT biomass density, including zero observations. Error bars represent approximate asymmetric 95% confidence intervals around the annual mean, calculated from its estimated standard error using a lognormal approximation.

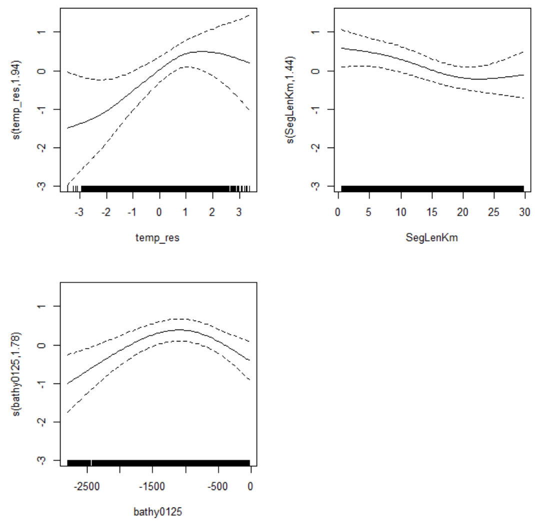


Figure 2. Partial effects of temperature residuals, segment length and bathymetry on BFT occurrence (binomial model).

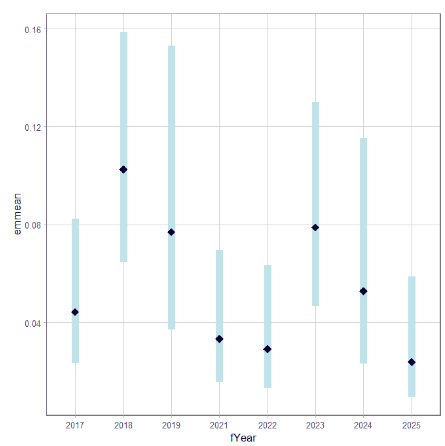


Figure 3: Standardized annual probabilities of BFT occurrence estimated by the binomial component.

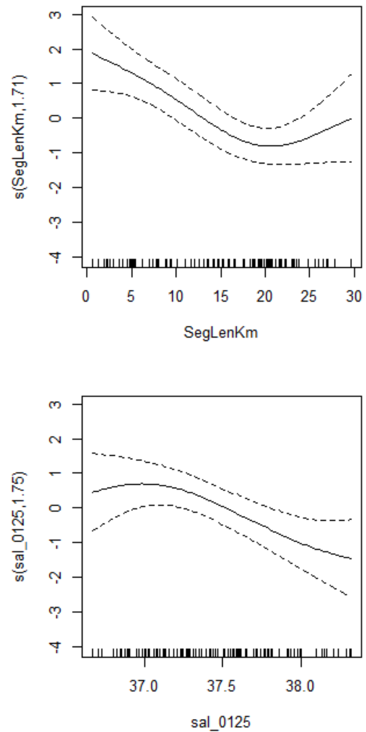


Figure 4: Partial effects of segment length, temperature residuals and salinity on positive BFT biomass.

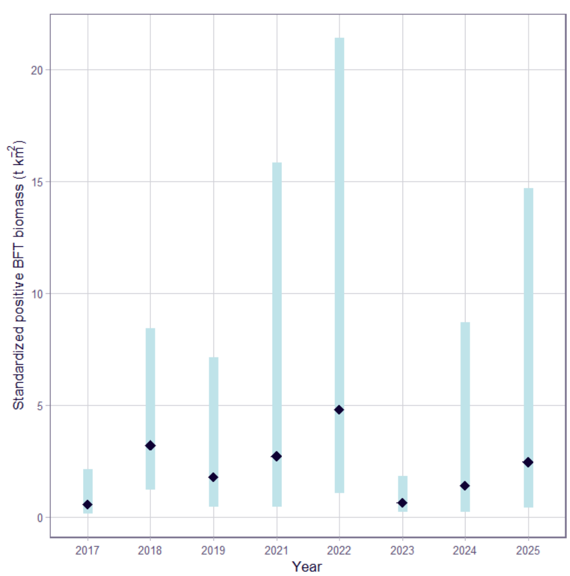


Figure 5. Standardized annual estimates of positive BFT biomass from the positive component of the reference delta-lognormal model. Points represent back-transformed annual marginal estimates conditional on BFT presence, and error bars indicate 95% confidence intervals. No estimate is shown for 2020 because no survey was conducted.

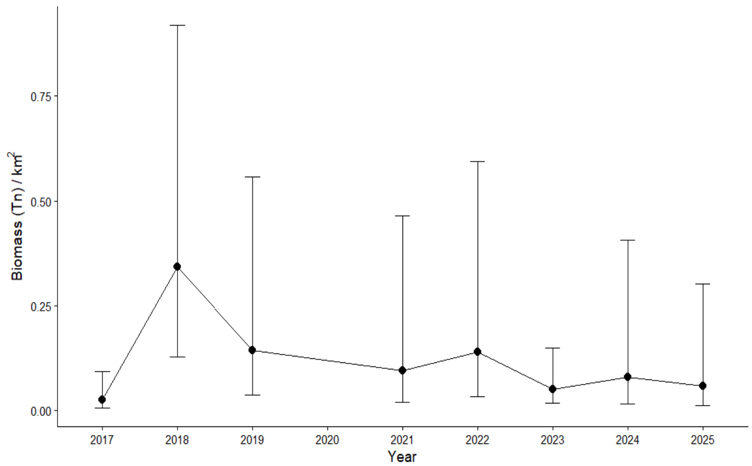


Figure 6. Standardized annual BFT biomass index estimated using the reference delta-lognormal model. Points represent the product of the annual marginal probability of occurrence and the back-transformed annual marginal positive-biomass estimate. Error bars represent asymmetric 95% confidence intervals derived from the coefficient of variation. No estimate is shown for 2020 because no aerial survey was conducted.

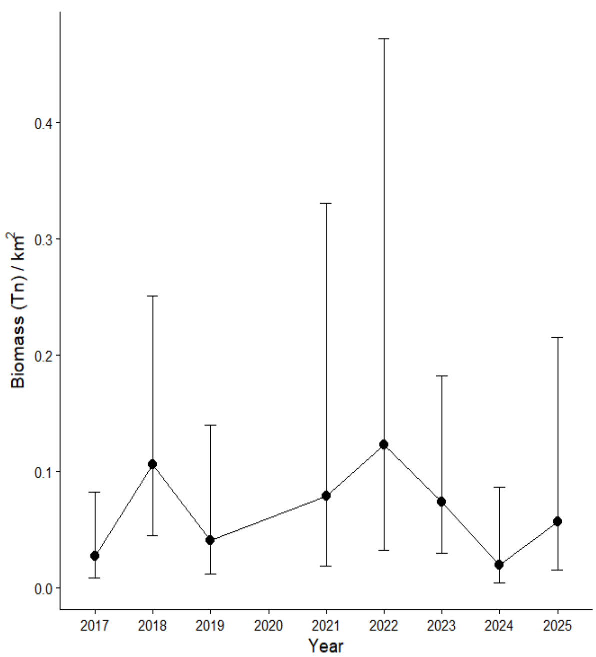


Figure 7. Comparison of standardized annual BFT biomass indices obtained from the reference delta-lognormal model and the reduced model without environmental or spatial covariates. Both formulations included year and a smooth effect of segment length, and the positive-biomass components included sampled area as a log-offset. Points represent annual estimates and error bars indicate asymmetric 95% confidence intervals. No estimate is shown for 2020 because no aerial survey was conducted.

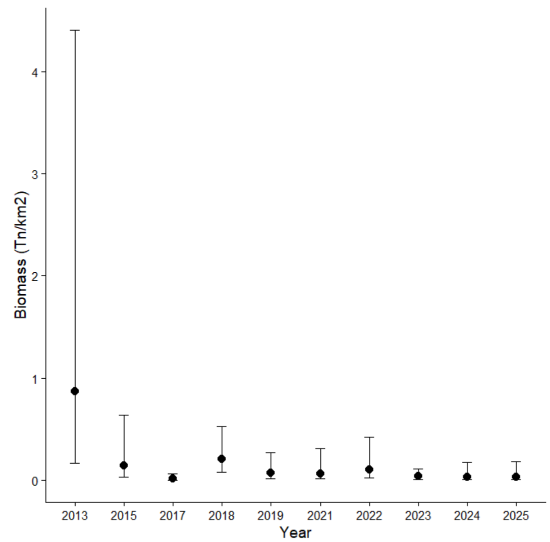


Figure 8. Extended standardized BFT biomass index obtained from the delta-lognormal model after incorporating the 2013 and 2015 surveys. Points represent annual estimates and error bars indicate asymmetric 95% confidence intervals. The high 2013 estimate was associated with substantial uncertainty and should be interpreted cautiously because of the unusually high dispersion of the observations from that survey.

Supplementary Material

Table S1. Summary of the binomial occurrence component of the reference delta-lognormal model

Model specification:

bft_pres ~ fYear + s(temp_res) + s(SegLenKm, k = 4) + s(bathy0125, k = 4)

Family: binomial; link function: logit.

A. Parametric coefficients

Term	Estimate	Standard error	z-value	p-value
Intercept	−3.501	0.303	−11.559	<0.001
Year 2018	0.924	0.387	2.389	0.017
Year 2019	0.608	0.501	1.213	0.225
Year 2021	−0.273	0.491	−0.555	0.579
Year 2022	−0.463	0.441	−1.049	0.294
Year 2023	0.612	0.354	1.732	0.083
Year 2024	0.229	0.543	0.421	0.674
Year 2025	−0.673	0.471	−1.428	0.153

B. Smooth terms

Smooth term	Effective df	Reference df	χ^2	p-value
Temperature residual (temp_res)	1.936	9	9.369	0.003
Segment length (SegLenKm)	1.438	3	8.698	0.003
Bathymetry (bathy0125)	1.775	3	10.230	0.003

Model summary

Sample size	Adjusted R^2	Deviance explained	REML criterion	AUC
2,612	0.025	6.54%	427.75	0.704

Table note: Parametric coefficients are expressed on the logit scale. Year was included as a categorical effect, with 2017 as the reference level; consequently, individual year coefficients represent contrasts with 2017. The significance tests for smooth terms are approximate. Effective degrees of freedom greater than one indicate non-linear responses.

Table S2. Standardized annual probabilities of BFT occurrence estimated by the binomial component of the reference delta-lognormal model.

Year	Probability	SE	Lower 95% CI	Upper 95% CI
2017	0.0442	0.0143	0.0233	0.0823
2018	0.1025	0.0235	0.0647	0.1587
2019	0.0769	0.0281	0.0370	0.1532
2021	0.0332	0.0127	0.0156	0.0694
2022	0.0292	0.0117	0.0131	0.0634
2023	0.0789	0.0207	0.0467	0.1301
2024	0.0528	0.0217	0.0233	0.1153
2025	0.0239	0.0111	0.0095	0.0587

Table S3. Summary of the positive-biomass component of the reference delta-lognormal model

Model specification:

log(tonnes) ~ fYear + s(SegLenKm, k = 4) + s(temp_res) + s(sal_0125) + offset(log(area))

Family: Gaussian; link function: identity.

A. Parametric coefficients

Term	Estimate	Standard error	t-value	p-value
Intercept	−0.637	0.563	−1.132	0.260
Year 2018	1.737	0.767	2.265	0.026
Year 2019	1.156	0.900	1.284	0.202
Year 2021	1.569	1.010	1.553	0.124
Year 2022	2.140	0.839	2.551	0.012
Year 2023	0.115	0.677	0.170	0.865
Year 2024	0.908	1.063	0.854	0.395
Year 2025	1.470	0.910	1.614	0.110

B. Smooth terms

Smooth term	Effective df	Reference df	F-value	p-value
Segment length (SegLenKm)	1.714	3	6.894	<0.001
Temperature residual (temp_res)	1.632	9	0.523	0.045
Salinity (sal_0125)	1.746	9	1.243	0.001

Model summary

Sample size	Adjusted R^2	Deviance explained	Residual variance	REML criterion
107	0.161	36.7%	3.6339	219.45

Table note: The response variable was the natural logarithm of positive BFT biomass in tonnes. Surveyed area was included as a log-offset but does not appear in the formula printed by summary.gam() because it was supplied separately in the model call. Parametric coefficients are expressed on the log scale. Year was included as a categorical effect, with 2017 as the reference level; individual year coefficients therefore represent contrasts with 2017. Significance tests for the smooth terms are approximate.

Table S4. Back-transformed annual marginal estimates of the positive-biomass component of the reference delta-lognormal model.

Year	Positive biomass (t km^{-2})	SE	Lower 95% CI	Upper 95% CI
2017	0.566	0.379	0.150	2.136
2018	3.212	1.561	1.223	8.432
2019	1.797	1.245	0.454	7.114
2021	2.716	2.411	0.465	15.847
2022	4.807	3.615	1.079	21.413
2023	0.634	0.338	0.220	1.829
2024	1.402	1.289	0.226	8.688
2025	2.457	2.211	0.411	14.688

Table note: Estimates were obtained on the log scale with the area offset fixed at zero, corresponding to a reference area of 1 km^2, and were back-transformed using the exponential function. Standard errors were transformed to the response scale using a first-order delta-method approximation. Confidence limits were calculated on the log scale and subsequently back-transformed.

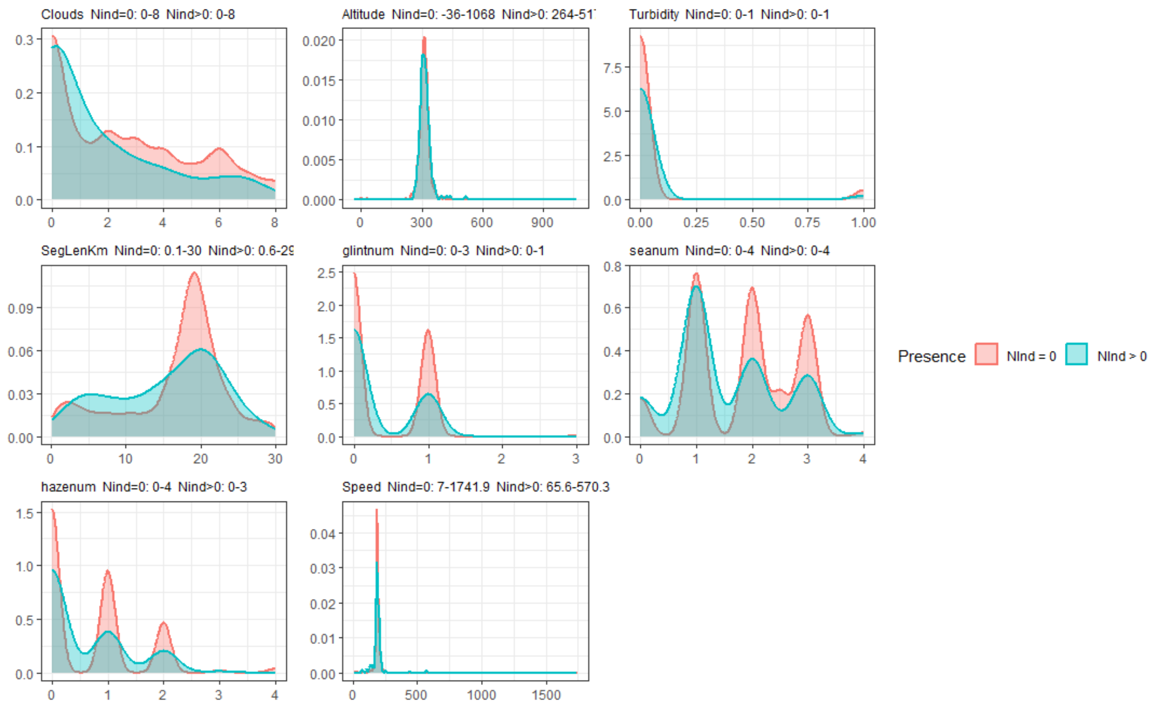


Figure S1. Distributions of the sampling variables used to define the filtering criteria, separated. Showing the range of the variable for samples with number of individuals =0 and >0.

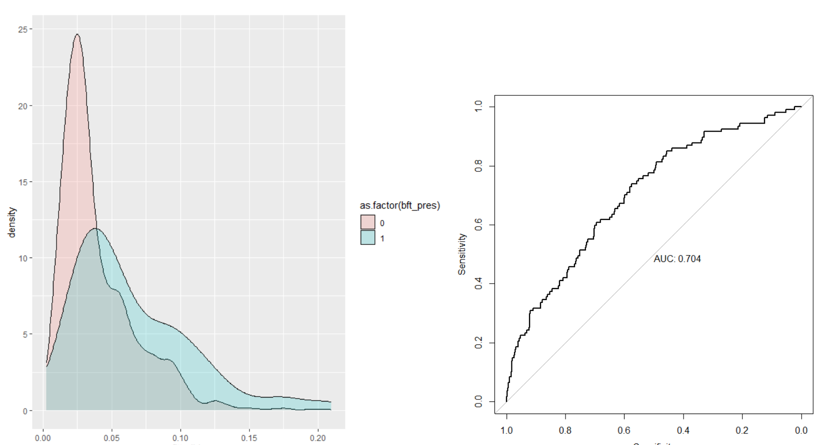


Figure S2. Receiver operating characteristic curve and distribution of fitted occurrence probabilities for segments with and without observed BFT presence.

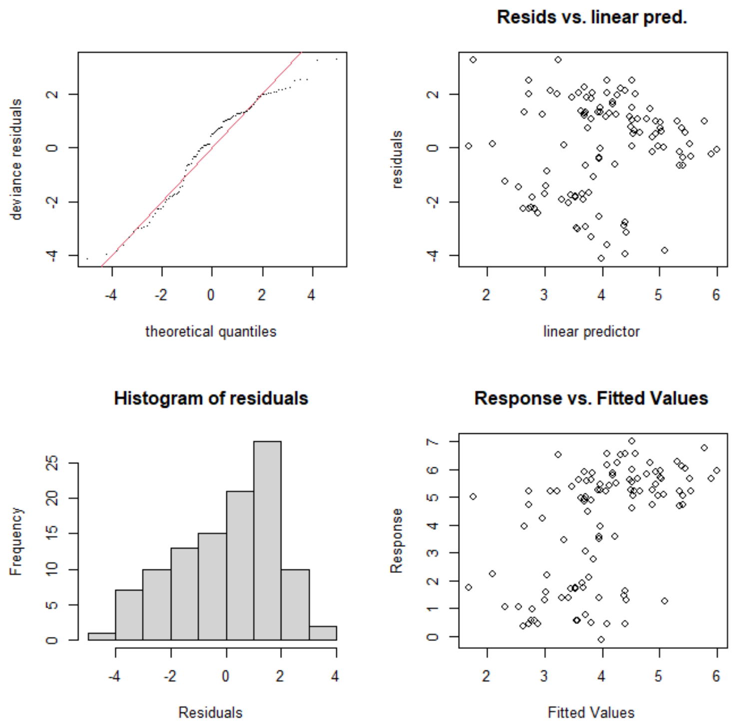


Figure S3. Diagnostic plots for the positive-biomass component of the reference delta-lognormal model.

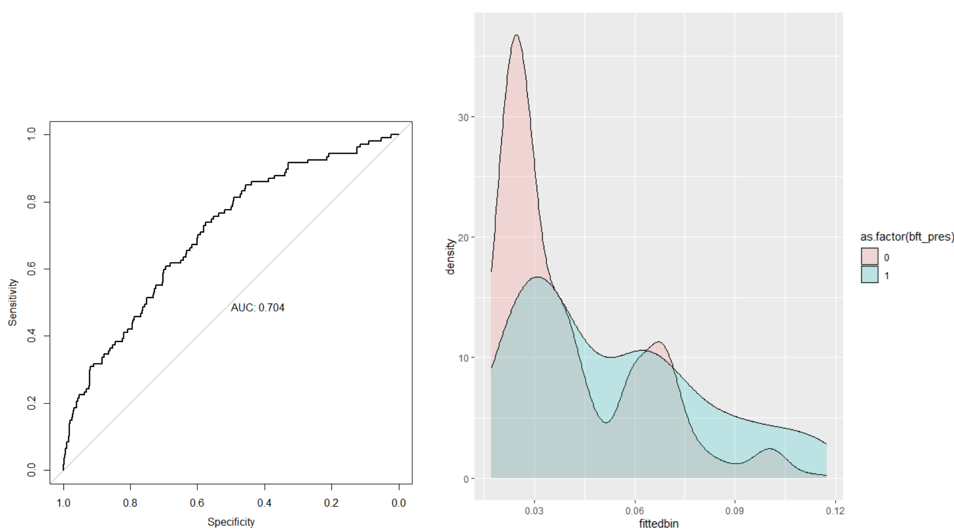


Figure S4. Receiver operating characteristic curve and distribution of fitted occurrence probabilities for the reduced binomial occurrence model without environmental or spatial covariates. Probability-density distributions are shown separately for survey segments with and without observed BFT presence.

Supplementary report SR.1: Tweedie modelling approach with tuna densities

METHODS

GAM models were fit by means of Restricted Maximum Likelihood (REML) method (Marra and Wood, 2011), using the 'mgcv' package (Wood, 2017) in R statistical software (R Core Team, 2024). Models were selected using a backward selection approach, removing the variable with the highest p-value at each step and retaining in the final model only variables with p-value < 0.05. SegLenKm and Year were mandatory in all models and were not incorporated in the backward selection approach.

Initial variables tested: MidLon, MidLat, interaction MidLon – MidLat, bathy0125, slope025, yday, fdia, mld_d0125, sal_0125, sst_anom, sst_d0125, sst15dinc, temp_res, GradSpat_SSS_d0125

For modelling tuna density, three configurations were tested: A) using SegLenKm as weights, B) using SegLenKm as smoother and C) using SegLenKm as a smoother and as weights.

For configurations A and C with SegLenKm incorporated as weights in the model, almost all explanatory variables were significant with p-value < 0.05. In this case, for each configuration three models were tested excluding highly correlated variables (see Figure S1), which were generally temperature and temporal variables. For each configuration, the three models tested included all variables that were significant at p-value and were not correlated, as well as one of the three configurations of temperature and temporal variables:

Model A: + s(yday, k=4) + s(sst15dinc, k = 4)
Model B: + s(temp_res, k=4)
Model C: + s(sst_d0125, k=4)

Within each configuration, models were compared in terms of the AIC and the deviance explained by the model and the R2 between observed and predicted values. Retained model was that with lower AIC, higher deviance explained and higher R2. In both configurations A and C, the best model resulted model 2, which included temp_res and yday.

To select between model configurations, the three final selected models were compared in terms of the highest precision of the estimated annual index, i.e. lower coefficient of variation.

RESULTS

Tweedie model with the time series 2013-2025

A. SegLenKm as weights

Table SR1.1. Model specifications for the Tweedie final selected model for Tuna density (2013-2025).

```
Family: Tweedie(p=1.648)
Link function: log

Formula:
TotWgt1325_tnkm2 ~ as.factor(Year) + s(MidLon, MidLat, k = 9) +
   +s(bathy0125, k = 4) + s(slope025, k = 4) + s(fdia, k = 4) +
   s(mld_d0125, k = 4) + s(sal_0125, k = 4) + s(temp_res, k = 4)

Parametric coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)          -1.4825     0.1543  -9.608  < 2e-16 ***
as.factor(Year)2015  -1.4232     0.2280  -6.241 4.93e-10 ***
as.factor(Year)2017  -1.4115     0.2413  -5.850 5.42e-09 ***
as.factor(Year)2018  -0.1300     0.1858  -0.699 0.484331
as.factor(Year)2019  -0.2784     0.1799  -1.548 0.121698
as.factor(Year)2021  -0.7373     0.2167  -3.402 0.000678 ***
as.factor(Year)2022  -1.6162     0.2310  -6.996 3.21e-12 ***
as.factor(Year)2023  -0.9531     0.2252  -4.233 2.37e-05 ***
as.factor(Year)2024  -1.9561     0.2249  -8.698  < 2e-16 ***
as.factor(Year)2025  -1.1228     0.2651  -4.236 2.34e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                   edf Ref.df      F  p-value
s(MidLon,MidLat) 5.432      8  4.944  < 2e-16 ***
s(bathy0125)     1.707      3  6.969 8.75e-07 ***
s(slope025)      1.940      3 13.218  < 2e-16 ***
s(fdia)          2.297      3 30.006  < 2e-16 ***
s(mld_d0125)     1.914      3 12.089  < 2e-16 ***
s(sal_0125)      2.808      3 26.365  < 2e-16 ***
s(temp_res)      2.845      3 55.231  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  -0.0014   Deviance explained = 11.2%
-REML =  11961  Scale est. = 36.139    n = 3171
```

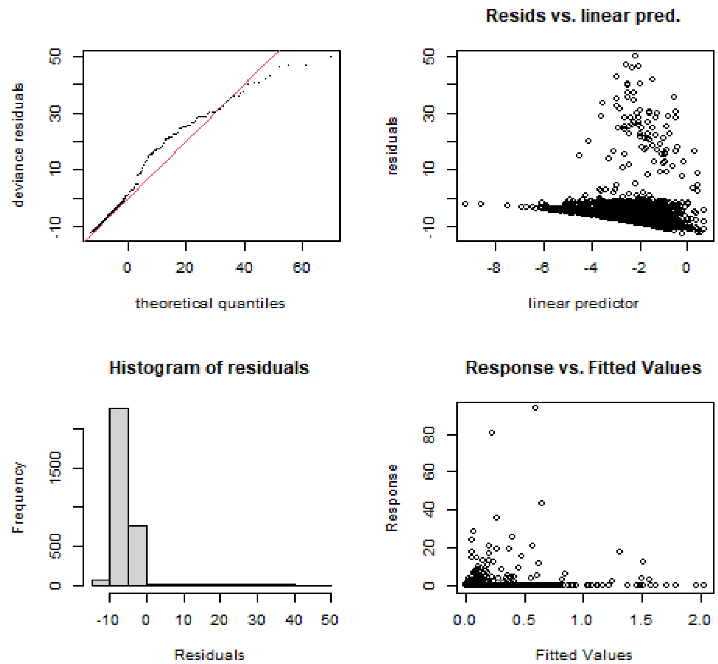


Figure SR1.1. Goodness of fit plots for the tweedie model.

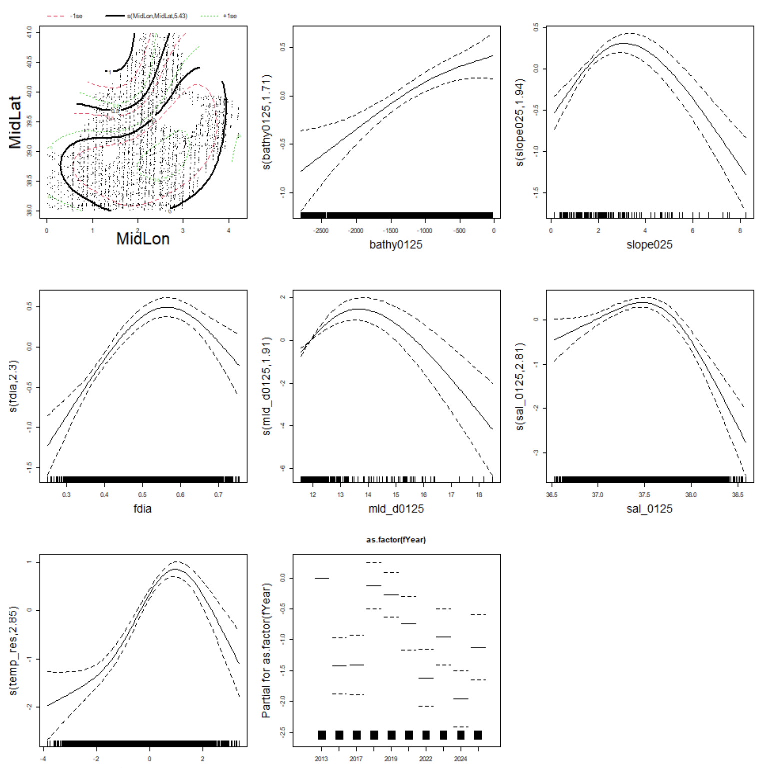


Figure SR1.2. Responses for the explanatory variables in the Tweedie model on tuna density (2013-2025).

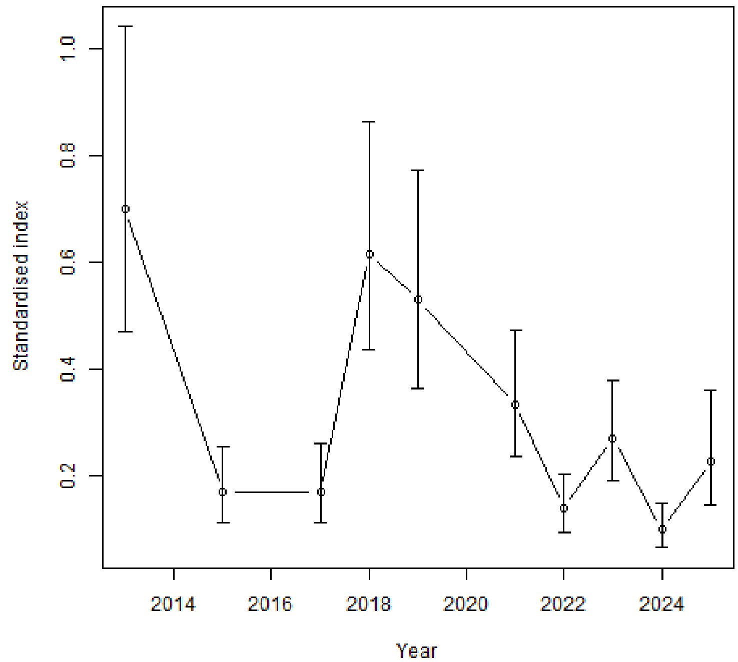


Figure SR1.3. Standardised biomass index obtained from the aerial surveys in t /Km^2.

Table SR1.2. CPUE nominal and standardised.

year	nsamples	index	UCInn	LCInn	cv_i	nominalCPUA	CPUA_LCInn	CPUA_UCInn
2013	341	0.6991	1.0419	0.4691	0.2015	0.0903	0.0439	0.1858
2015	218	0.1684	0.2540	0.1117	0.2076	0.2661	0.1020	0.6940
2017	311	0.1704	0.2601	0.1117	0.2138	0.1398	0.0556	0.3519
2018	363	0.6139	0.8647	0.4359	0.1725	0.6742	0.2985	1.5229
2019	336	0.5292	0.7716	0.3630	0.1902	0.3735	0.1100	1.2687
2021	363	0.3345	0.4722	0.2369	0.1737	0.2129	0.0922	0.4913
2022	303	0.1389	0.2037	0.0947	0.1932	0.2981	0.1181	0.7526
2023	308	0.2695	0.3789	0.1918	0.1715	0.4302	0.1982	0.9338
2024	286	0.0989	0.1493	0.0655	0.2081	0.0440	0.0144	0.1343
2025	342	0.2275	0.3589	0.1442	0.2310	0.1406	0.0635	0.3114

B. SegLenKm as smoother

Table SR1.3. Model specifications for the Tweedie final selected model for Tuna density (2013-2025).

```
Family: Tweedie(p=1.634)
Link function: log

Formula:
TotWgt1325_tnkm2 ~ as.factor(Year) + s(SegLenKm, k = 4) + s(bathy0125,
    k = 4) + s(temp_res, k = 4)

Parametric coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)          -1.2200     0.5900  -2.068   0.0387 *
as.factor(Year)2015  -1.1072     0.8469  -1.307   0.1912
as.factor(Year)2017  -1.3686     0.8545  -1.602   0.1093
as.factor(Year)2018  -0.1527     0.7109  -0.215   0.8300
as.factor(Year)2019  -0.6210     0.7143  -0.869   0.3847
as.factor(Year)2021  -1.2006     0.7589  -1.582   0.1137
as.factor(Year)2022  -1.0606     0.8672  -1.223   0.2214
as.factor(Year)2023  -0.4515     0.8334  -0.542   0.5880
as.factor(Year)2024  -1.5385     0.8428  -1.826   0.0680 .
as.factor(Year)2025  -1.4913     0.9170  -1.626   0.1040
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
              edf Ref.df      F p-value
s(SegLenKm)  1.842      3 15.176  <2e-16 ***
s(bathy0125) 1.416      3  1.586  0.0355 *
s(temp_res)  1.496      3  2.069  0.0158 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

R-sq.(adj) = 0.0331 **Deviance explained = 20.2%** REML = 9714.8 Scale est. = 35.122 n = 3171

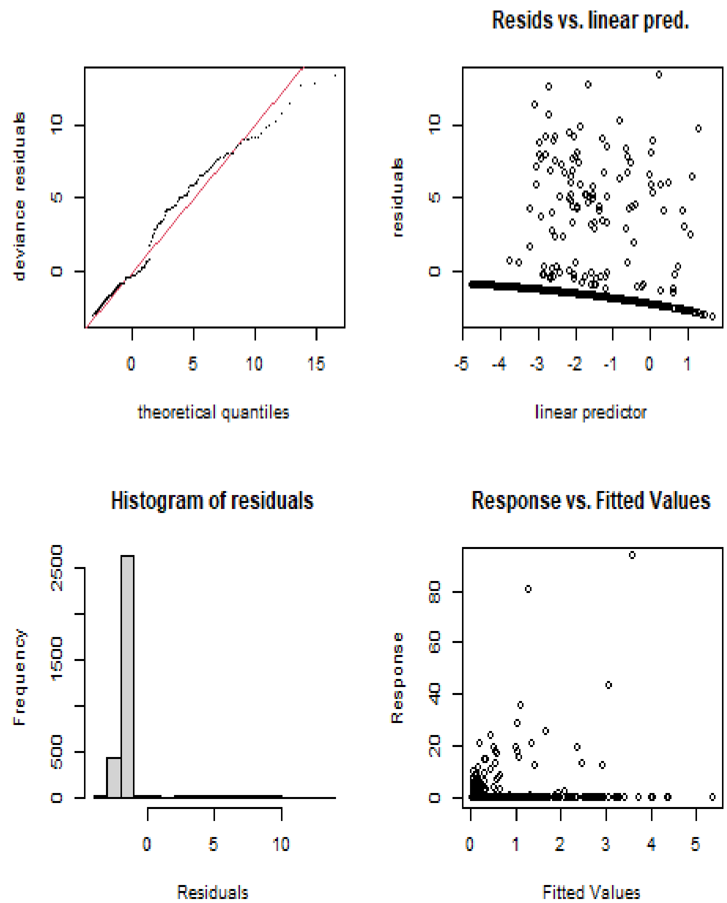


Figure SR1.4. Goodness of fit plots for the tweedie model.

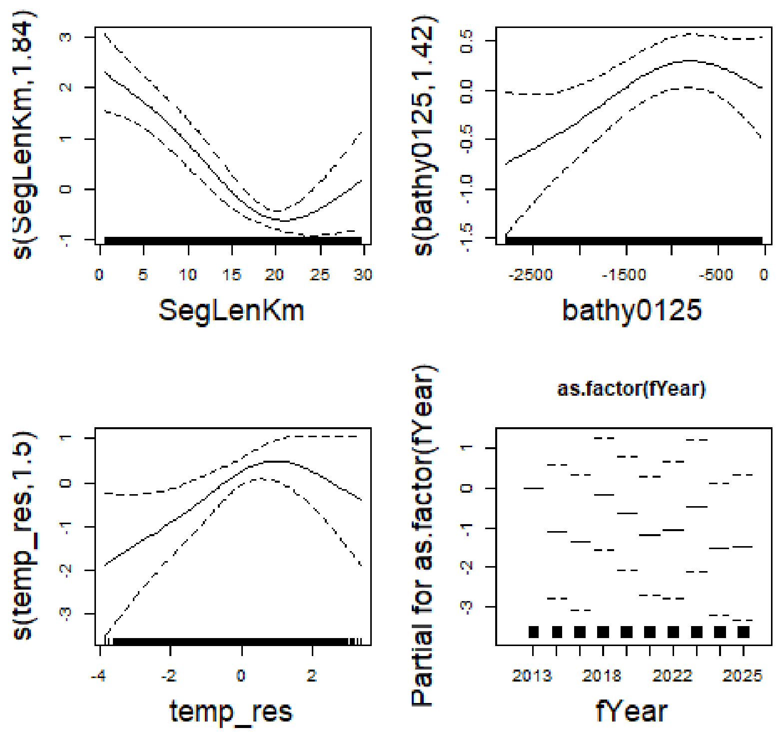


Figure SR1.5. Responses for the explanatory variables in the Tweedie model on tuna density (2013-2025).

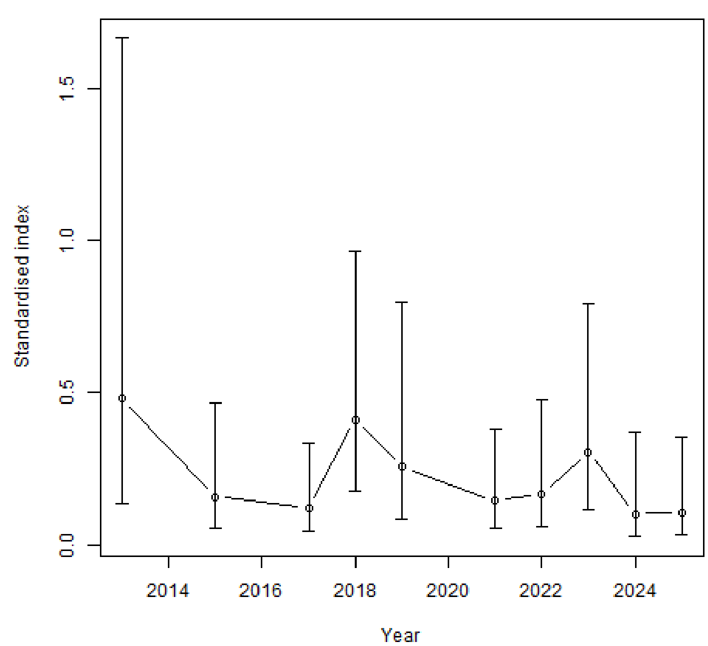


Figure SR1.6. Standardized biomass index obtained from the aerial surveys in t /Km^2.

Table SR1.4. CPUE nominal and standardized.

year	nsamples	index	UCInn	LCInn	cv_i	nominalCPUA	CPUA_LCInn	CPUA_UCInn
2013	341	0.4796	1.6635	0.1383	0.6871	0.0903	0.0439	0.1858
2015	218	0.1585	0.4677	0.0537	0.5831	0.2661	0.1020	0.6940
2017	311	0.1220	0.3361	0.0443	0.5408	0.1398	0.0556	0.3519
2018	363	0.4117	0.9627	0.1761	0.4446	0.6742	0.2985	1.5229
2019	336	0.2577	0.7987	0.0832	0.6138	0.3735	0.1100	1.2687
2021	363	0.1444	0.3817	0.0546	0.5162	0.2129	0.0922	0.4913
2022	303	0.1661	0.4770	0.0578	0.5664	0.2981	0.1181	0.7526
2023	308	0.3054	0.7896	0.1181	0.5031	0.4302	0.1982	0.9338
2024	286	0.1030	0.3709	0.0286	0.7124	0.0440	0.0144	0.1343
2025	342	0.1080	0.3542	0.0329	0.6505	0.1406	0.0635	0.3114

C. SegLenKm as smoother and weights

Table SR1.5. Model specifications for the Tweedie final selected model for Tuna density (2013-2025).

```
Family: Tweedie(p=1.627)
Link function: log

Formula:
TotWgt1325_tnkm2 ~ as.factor(Year) + s(SegLenKm, k = 4) + s(MidLon,
    MidLat, k = 9) + s(bathy0125, k = 4) + s(slope025, k = 4) +
    s(fdia, k = 4) + +s(mld_d0125, k = 4) + s(sal_0125, k = 4) +
    s(temp_res, k = 4) + s(GradSpat_SSS_d0125, k = 4)

Parametric coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)          -0.8134     0.1490  -5.459 5.17e-08 ***
as.factor(Year)2015  -2.2537     0.2175 -10.361  < 2e-16 ***
as.factor(Year)2017  -1.9549     0.2265  -8.631  < 2e-16 ***
as.factor(Year)2018  -1.1376     0.1804  -6.307 3.24e-10 ***
as.factor(Year)2019  -1.3957     0.1811  -7.709 1.69e-14 ***
as.factor(Year)2021  -1.5131     0.2145  -7.052 2.15e-12 ***
as.factor(Year)2022  -2.6230     0.2259 -11.613  < 2e-16 ***
as.factor(Year)2023  -1.6327     0.2157  -7.570 4.88e-14 ***
as.factor(Year)2024  -1.9306     0.2188  -8.823  < 2e-16 ***
as.factor(Year)2025  -1.8339     0.2534  -7.238 5.69e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Approximate significance of smooth terms:
                       edf Ref.df       F  p-value
s(SegLenKm)          2.813      3 168.370  < 2e-16 ***
s(MidLon,MidLat)     5.451      8   3.009 7.10e-05 ***
s(bathy0125)         1.842      3   5.334 4.46e-05 ***
s(slope025)          1.858      3   6.229 2.99e-05 ***
s(fdia)              1.916      3  28.544  < 2e-16 ***
s(mld_d0125)         1.915      3   9.993 2.66e-07 ***
s(sal_0125)          1.968      3  27.763  < 2e-16 ***
s(temp_res)          2.925      3  96.798  < 2e-16 ***
s(GradSpat_SSS_d0125) 1.906     3   8.541 1.52e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  -0.0181   Deviance explained =   19%
-REML =  11733  Scale est. = 30.188     n = 3171
```

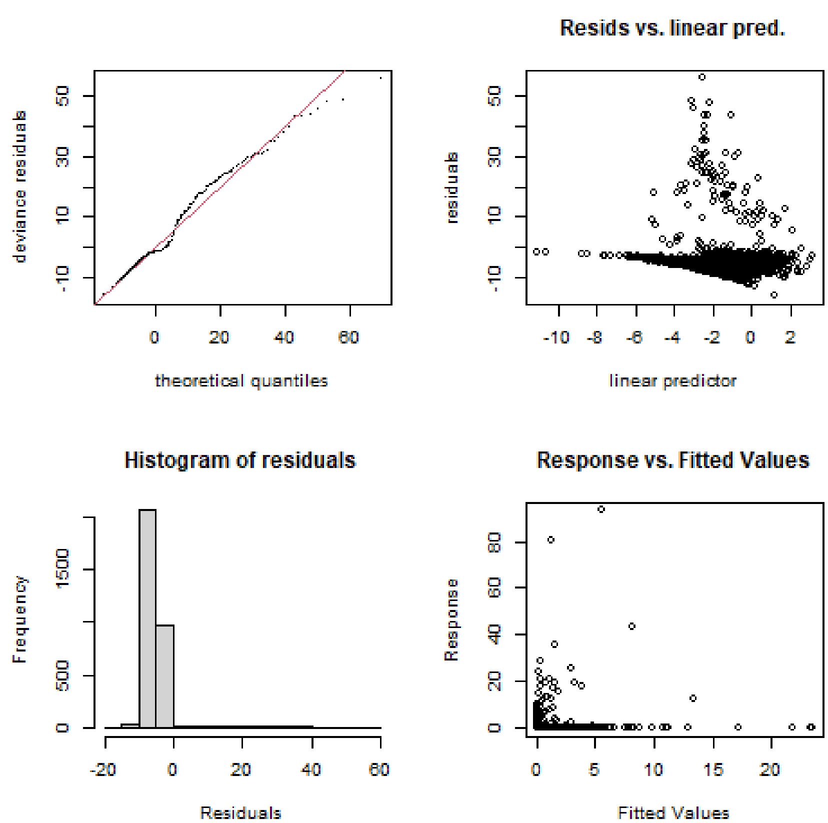


Figure SR1.7. Goodness of fit plots for the tweedie model.

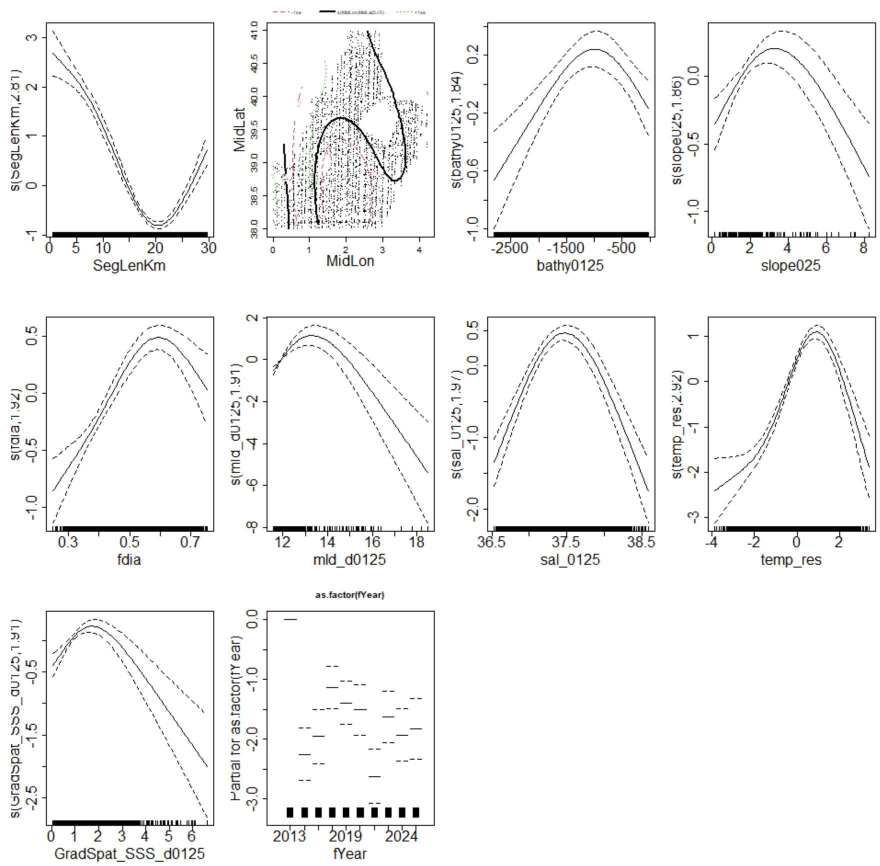


Figure SR1.8. Responses for the explanatory variables in the Tweedie model on tuna density (2013-2026).

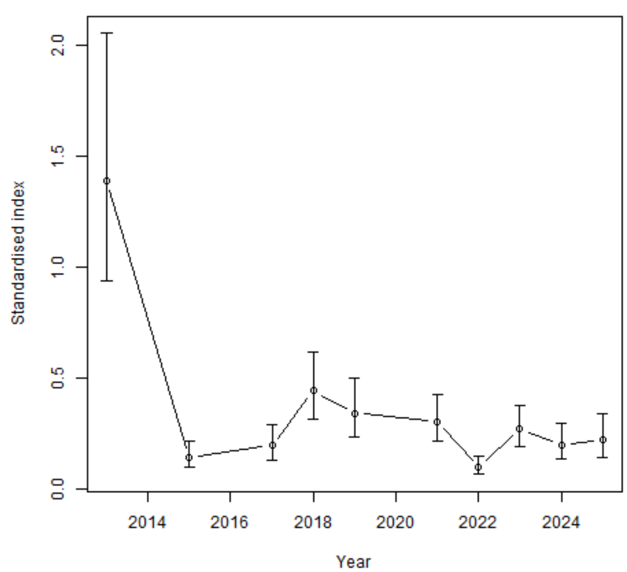


Figure SR1.9. Standardised biomass index obtained from the aerial surveys in t /Km^2.

Table SR1.6. CPUE nominal and standardised.

year	nsamples	index	UCInn	LCInn	cv_i	nominalCPUA	CPUA_LCInn	CPUA_UCInn
2013	341	1.3880	2.0540	0.9380	0.1978	0.0903	0.0439	0.1858
2015	218	0.1458	0.2153	0.0987	0.1969	0.2661	0.1020	0.6940
2017	311	0.1965	0.2908	0.1328	0.1978	0.1398	0.0556	0.3519
2018	363	0.4450	0.6198	0.3195	0.1668	0.6742	0.2985	1.5229
2019	336	0.3438	0.5033	0.2348	0.1923	0.3735	0.1100	1.2687
2021	363	0.3057	0.4274	0.2187	0.1687	0.2129	0.0922	0.4913
2022	303	0.1008	0.1482	0.0685	0.1948	0.2981	0.1181	0.7526
2023	308	0.2712	0.3767	0.1953	0.1653	0.4302	0.1982	0.9338
2024	286	0.2014	0.3005	0.1349	0.2023	0.0440	0.0144	0.1343
2025	342	0.2218	0.3433	0.1433	0.2211	0.1406	0.0635	0.3114

Tweedie model with the time series 2017-2025

TUNA DENSITY

TWEEDIE

With the model constructed with the configuration C (selected as it presented the lowest CV), we produced a new model including only data from 2017-2025. Results are as follows:

Table SR1.8. Model specifications for the Tweedie final selected model for Tuna density (2017-2025).

```
Family: Tweedie(p=1.646)
Link function: log

Formula:
TotWgt1325_tnkm2 ~ as.factor(fYear) + s(SegLenKm, k = 4) + s(MidLon,
    MidLat, k = 9) + s(bathy0125, k = 4) + s(slope025, k = 4) +
    s(fdia, k = 4) + +s(mld_d0125, k = 4) + s(sal_0125, k = 4) +
    s(temp_res, k = 4) + s(GradSpat_SSS_d0125, k = 4)


Parametric coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)          -3.2082     0.1357 -23.635  < 2e-16 ***
as.factor(fYear)2018  1.5201     0.1895   8.020 1.59e-15 ***
as.factor(fYear)2019  1.2679     0.2215   5.724 1.16e-08 ***
as.factor(fYear)2021  0.5901     0.1924   3.067  0.00218 **
as.factor(fYear)2022 -0.2225     0.1942  -1.146  0.25190
as.factor(fYear)2023  0.7178     0.1626   4.415 1.05e-05 ***
as.factor(fYear)2024  0.5168     0.2286   2.261  0.02384 *
as.factor(fYear)2025  0.5333     0.1890   2.822  0.00481 **
---
```

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:

```
                       edf Ref.df       F  p-value
s(SegLenKm)          2.918      3 205.751  < 2e-16 ***
s(MidLon,MidLat)     7.512      8  19.578  < 2e-16 ***
s(bathy0125)         1.810      3   6.220 1.96e-05 ***
s(slope025)          1.911      3   9.138 1.14e-06 ***
s(fdia)              0.997      3  24.442  < 2e-16 ***
s(mld_d0125)         1.932      3  19.256  < 2e-16 ***
s(sal_0125)          1.969      3  33.907  < 2e-16 ***
s(temp_res)          1.989      3 117.605  < 2e-16 ***
s(GradSpat_SSS_d0125) 1.564     3   4.523 0.000217 ***
---
```

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) = -0.125 **Deviance explained = 23.2%**
-REML = 9766.7 Scale est. = 30.129 n = 2612

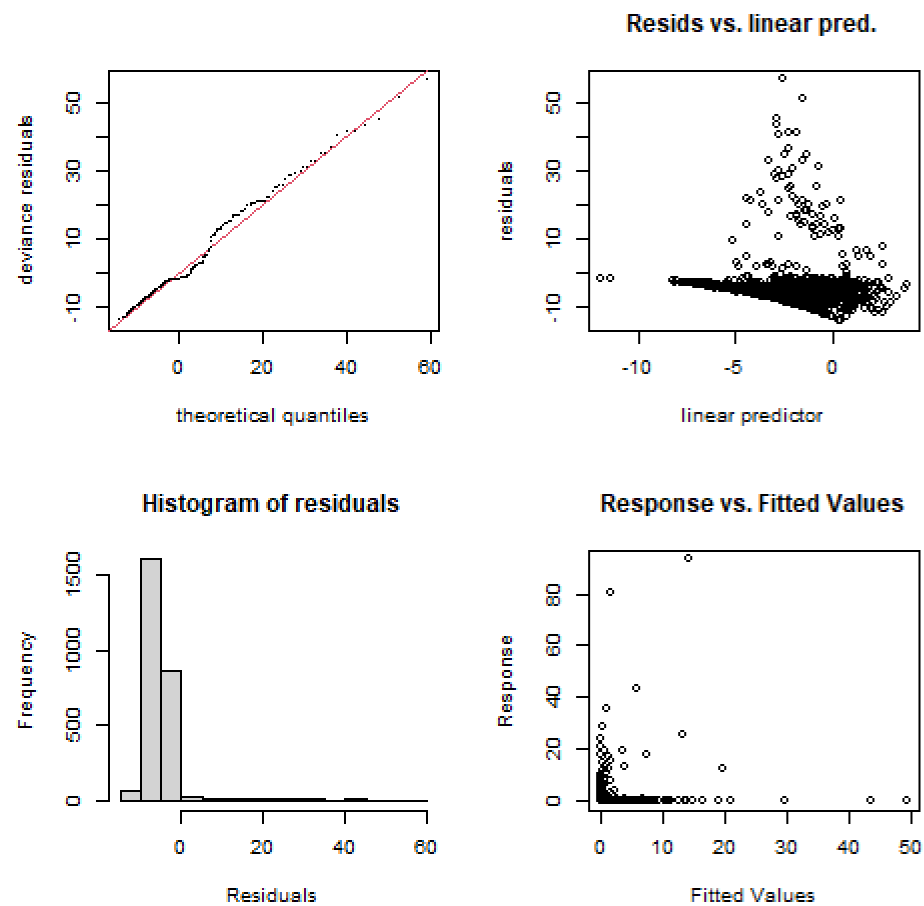


Figure SR1.10. Goodness of fit plots for the tweedie model.

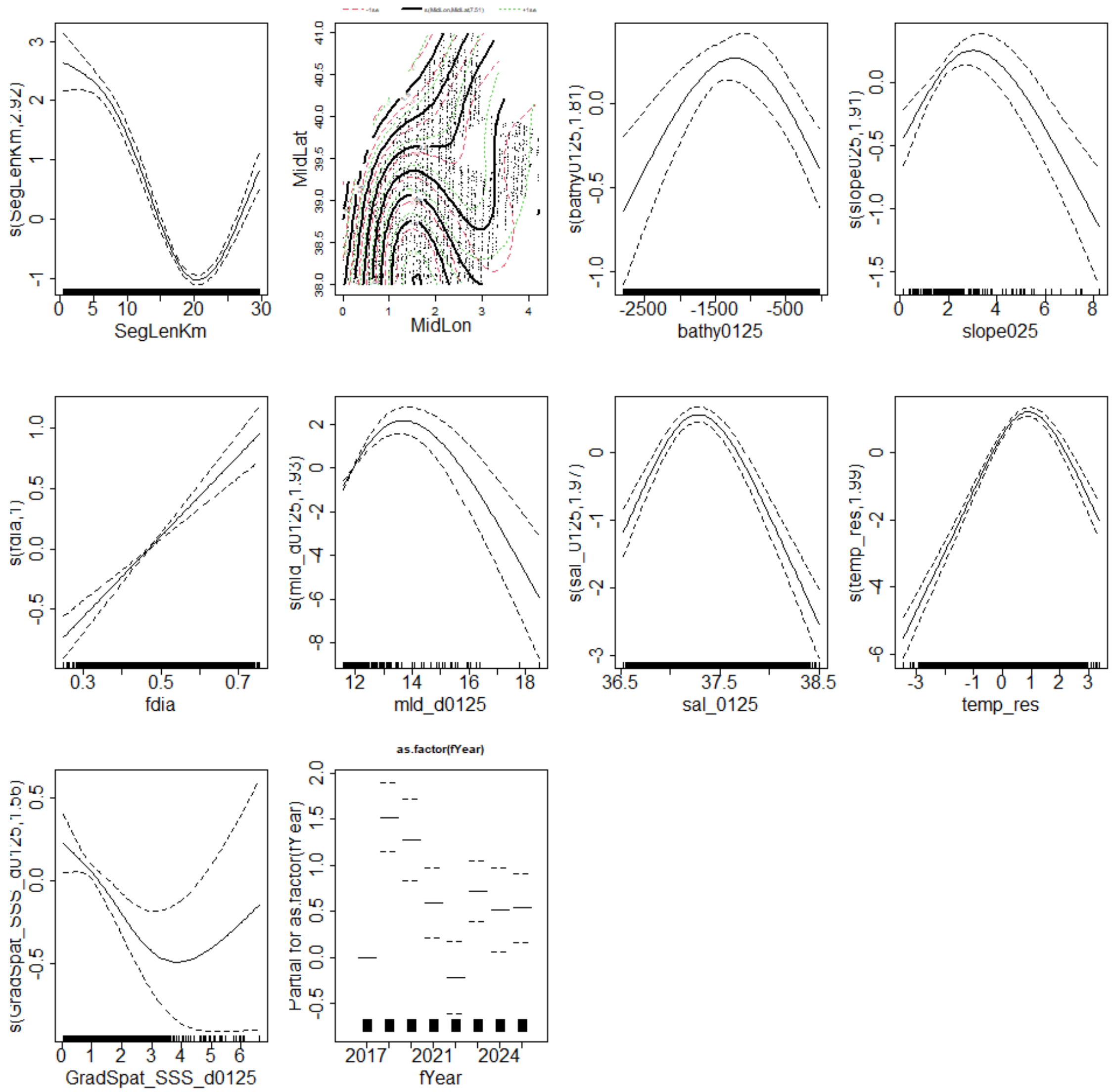


Figure SR1.11. Responses for the explanatory variables in the Tweedie model on tuna density (2013-2026).

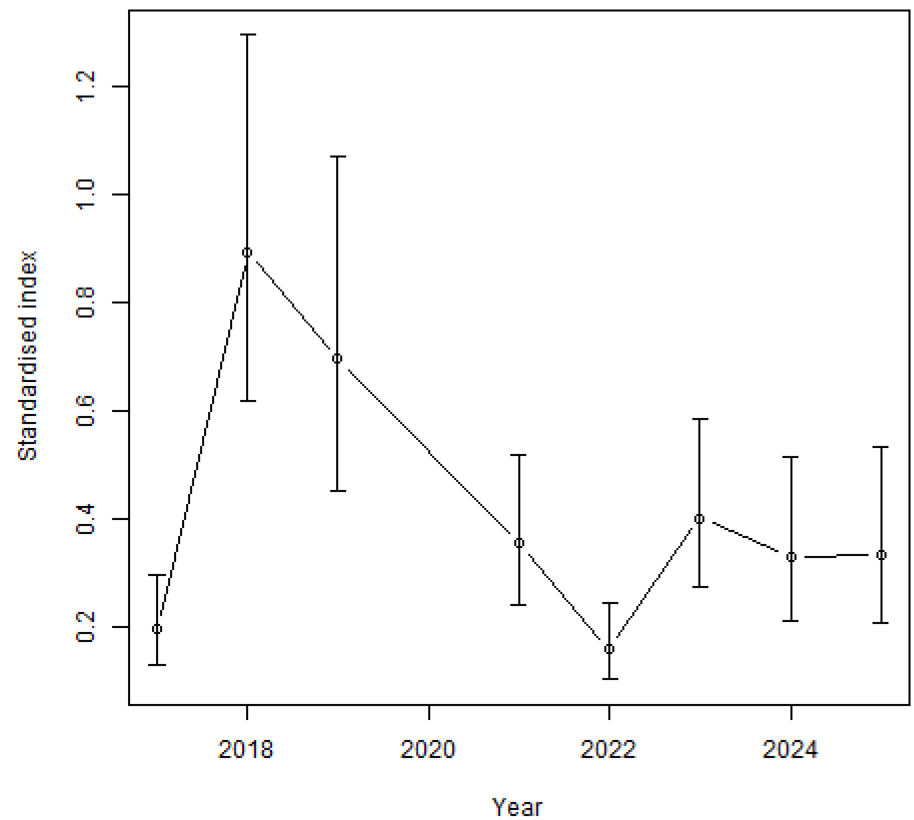


Figure SR1.12. Standardised biomass index obtained from the aerial surveys in tn/Km^2.

Table SR1.9. CPUE nominal and standardised.

year	nsamples	index	UCInn	LCInn	cv_i	nominalCPUA	CPUA_LCInn	CPUA_UCInn
2017	311	0.1953	0.2945	0.1295	0.2075	0.1398	0.0556	0.3519
2018	363	0.8932	1.2950	0.6160	0.1874	0.6742	0.2985	1.5229
2019	336	0.6941	1.0709	0.4499	0.2194	0.3735	0.1100	1.2687
2021	363	0.3524	0.5186	0.2395	0.1950	0.2129	0.0922	0.4913
2022	303	0.1564	0.2416	0.1012	0.2201	0.2981	0.1181	0.7526
2023	308	0.4004	0.5843	0.2744	0.1907	0.4302	0.1982	0.9338
2024	286	0.3275	0.5143	0.2086	0.2285	0.0440	0.0144	0.1343
2025	342	0.3330	0.5340	0.2076	0.2395	0.1406	0.0635	0.3114

Supplementary report SR.2: Bayesian modelling approach with tuna biomass

Methods

A Bayesian hurdle-lognormal generalized additive model was fitted using the brms package. The full Bayesian model was formulated as

TotWgt_tn ~ s(MidLon, MidLat, k = 9, bs = "ts") + Year + s(yday, k = 3, bs = "ts") + s(bathy0125, k = 3, bs = "ts") + s(sal_0125, k = 3, bs = "ts") + s(sst15dinc, k = 3, bs = "ts") + s(GradSpat_SSS_d0125, k = 3, bs = "ts") + s(temp_res, k = 3, bs = "ts") + offset(log(area))

hu ~ s(MidLon, MidLat, k = 9, bs = "ts") + Year + s(yday, k = 3, bs = "ts") + s(bathy0125, k = 3, bs = "ts") + s(sal_0125, k = 3, bs = "ts") + s(sst15dinc, k = 3, bs = "ts") + s(GradSpat_SSS_d0125, k = 3, bs = "ts") + s(temp_res, k = 3, bs = "ts") + offset(log(area)),

and we run 4 chains with 2000 iterations each, of which the first 500 links were discarded to make sure we only used values once the chains had converged. Chain outputs where thinned at 1/10 to reduce autocorrelation. The model was run with uninformative priors.

Model selection in Bayesian statistics cannot be done following standard frequentist procedures (p-values). Instead, Bayesian models were selected using a backward selection approach based primarily on the Leave One Out Information Criteria (LOO-IC). Variable selection was complemented with the Widely Applicable Information Criterion (WAIC), visual examination of smooth plots, and smoothing spline hyper-parameters. All the explanatory variables selected for the final model except the geographic variable longitude-latitude had a smooth parameter estimate > 1. Nevertheless, we kept longitude-latitude because it rendered a lower LOO-IC than the model without it. The selected model excluded variables yday, sst15dinc and sal_0125.

As with the variable selection, some model diagnostics for Bayesian models are different to frequentist models. For the selected Bayesian hurdle-lognormal model, we applied Pareto diagnostics (grounded on LOO cross-validations), posterior distribution of residuals, and posterior predictive checks.

Results

Bayesian hurdle-lognormal GAM

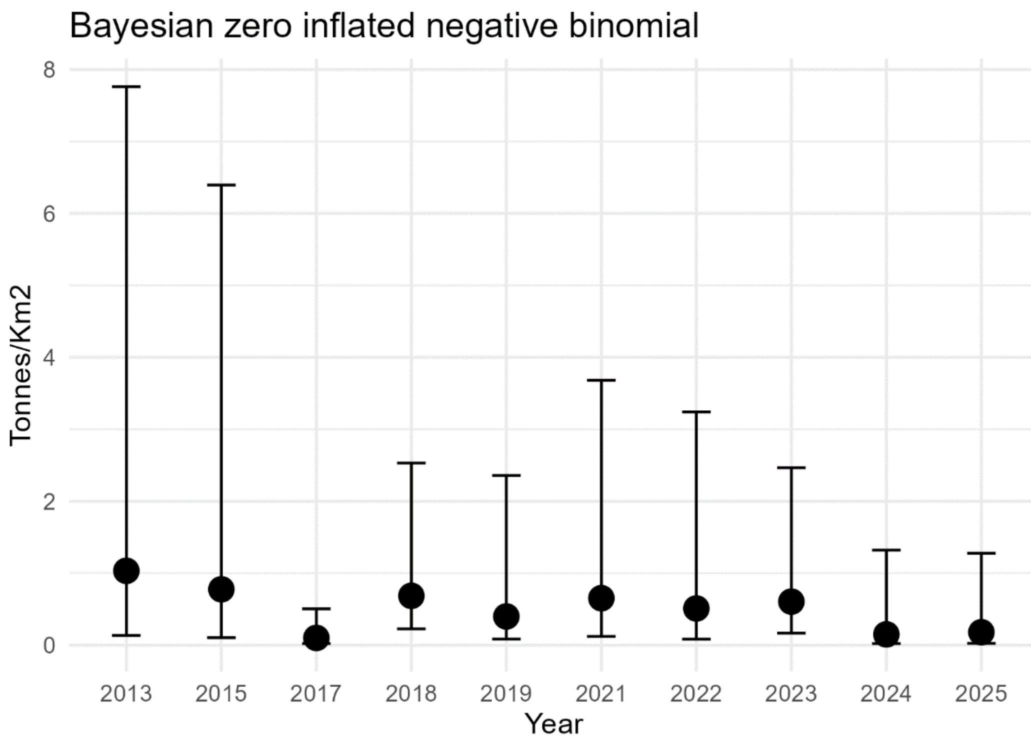


Figure SR2.1. Posterior estimates of tuna biomass with the respective credible Interval. Posterior point estimates and 95% credible intervals.

Table SR2.1. Posterior estimates of tuna biomass with the respective credible Interval. Posterior point estimates and 95% credible intervals.

Year	estimate	lower	upper	sd
2013	1.03361406	45.32998	0.13604932	7.76067213
2015	0.77678945	33.9411657	0.1051512	6.39451013
2017	0.10098053	3.60172481	0.02284347	0.50559432
2018	0.6859353	17.401617	0.22683569	2.53181134
2019	0.39956821	15.5318729	0.08391901	2.35764915
2021	0.65309495	25.7178045	0.12132331	3.68152794
2022	0.51038056	20.8969679	0.08280688	3.2426079
2023	0.6055373	17.8303839	0.16807911	2.46696366
2024	0.15012547	6.45155074	0.02233458	1.32050778
2025	0.17993567	7.99890323	0.02549327	1.27749023

Table SR2.2. Summary table of Bayesian Bayesian hurdle-lognormal model GAM distribution.

```
Smoothing Spline Hyperparameters:
          Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sds(sMidLonMidLat_1)   0.30   0.37 0.01   1.37 1.00  734   736
sds(sbathy0125_1)    1.45   1.45   0.07   5.07 1.00      770      756
sds(sGradSpat_SSS_d0125_1)  1.55  1.69  0.05   6.33 1.01   702    753
```

```
sds(stemp_res_1)      1.39   1.46 0.04     4.74 1.00      875      680
sds(hu_sMidLonMidLat_1)  0.29  0.37   0.01   1.42 1.00    722      774
sds(hu_sbathy0125_1)  2.65    2.25  0.06   8.22 1.00     568      537
sds(hu_sGradSpat_SSS_d0125_131    1.25  0.07   4.62 1.00  815      790
sds(hu_stemp_res_1)    1.72    1.39   0.15   5.31 1.00    830      746
Regression Coefficients:
           Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept    0.83      0.84    -0.70     2.47 1.00      717      735
hu_Intercept   -0.77    0.41    -1.52     0.09 1.00      791      846
Year2015      1.11      1.18    -1.22     3.43 1.00      771      772
Year2017      -1.53      1.16    -3.74     0.64 1.00      730      735
Year2018     -0.24      0.95    -2.15     1.53 1.00      788      722
Year2019      -0.37      1.01    -2.39     1.62 1.00      714      760
Year2021       0.96      1.05    -1.08     2.96 1.00      731      732
Year2022       0.46      1.19    -1.86     2.67 1.00      826      759
Year2023      -0.67      1.04    -2.74     1.21 1.00      735      814
Year2024      -1.28      1.14    -3.51     0.88 1.00      739      718
Year2025       0.11      1.30    -2.52     2.50 1.00      754      726
hu_Year2015    1.43      0.62     0.27     2.66 1.00      770      727
hu_Year2017    0.79      0.59    -0.34     1.95 1.00      736      667
hu_Year2018    0.14      0.47    -0.78     1.03 1.00      855      833
hu_Year2019    0.57      0.46    -0.36     1.45 1.00      767      666
hu_Year2021    1.45      0.56     0.38     2.61 1.00      685      735
hu_Year2022    1.18      0.65    -0.10     2.39 1.00      746      815
hu_Year2023   -0.15      0.54    -1.22     0.88 1.00      736      756
hu_Year2024    0.62      0.52    -0.43     1.66 1.00      794      775
hu_Year2025    1.87      0.68     0.57     3.10 1.00      760      862
```

Bayesian R^2= 0.25

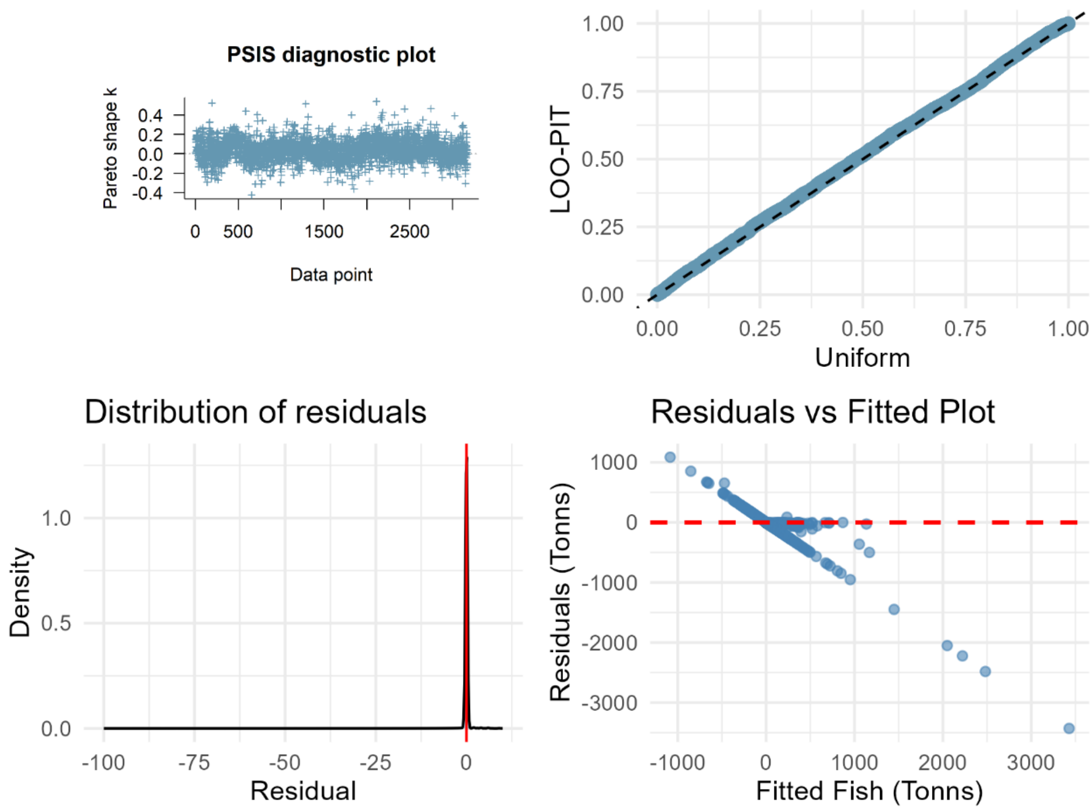


Figure SR2.2. Pareto and residual model diagnostics. A) Pareto diagnostics, all the observations have Pareto k values < 0.62. B) Empirical cumulative probabilities vs. theoretical uniform quantiles. C) Posterior distribution of residuals (negative values cropped at -100). D) Residual plot.

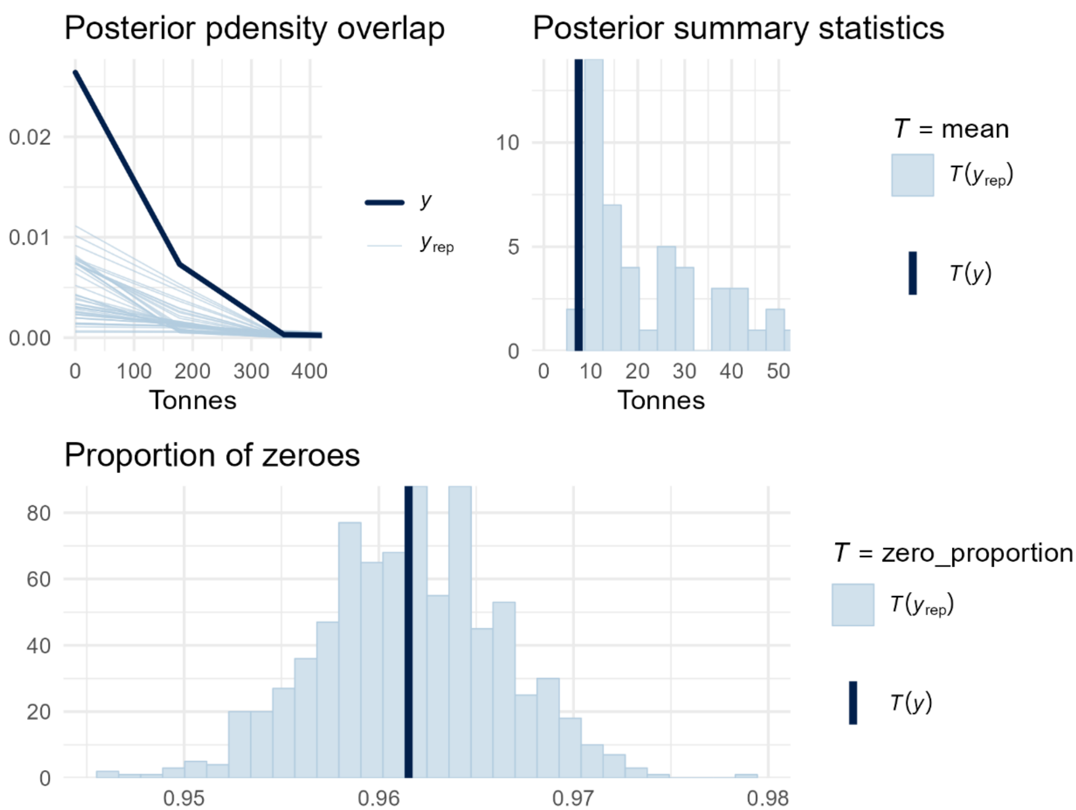


Figure SR2.3. Posterior model diagnostics. A) Posterior density overly plots (50 samples, cropped at 400 tonnes). In a good fit, the light lines should tightly cluster around the dark line, mirroring its peak, width, and tails. B) Posterior summary statistics plot (cropped at 50 tonnes). In a good fit, the dark vertical line should fall in the central region of the posterior predictive distribution. C) Posterior proportion of zeros. In a good fit, the dark vertical line should fall right in the middle of the histogram.

RESULTS
Bayesian zero-inflated negative-binomial model
Cropped data 2017-2025

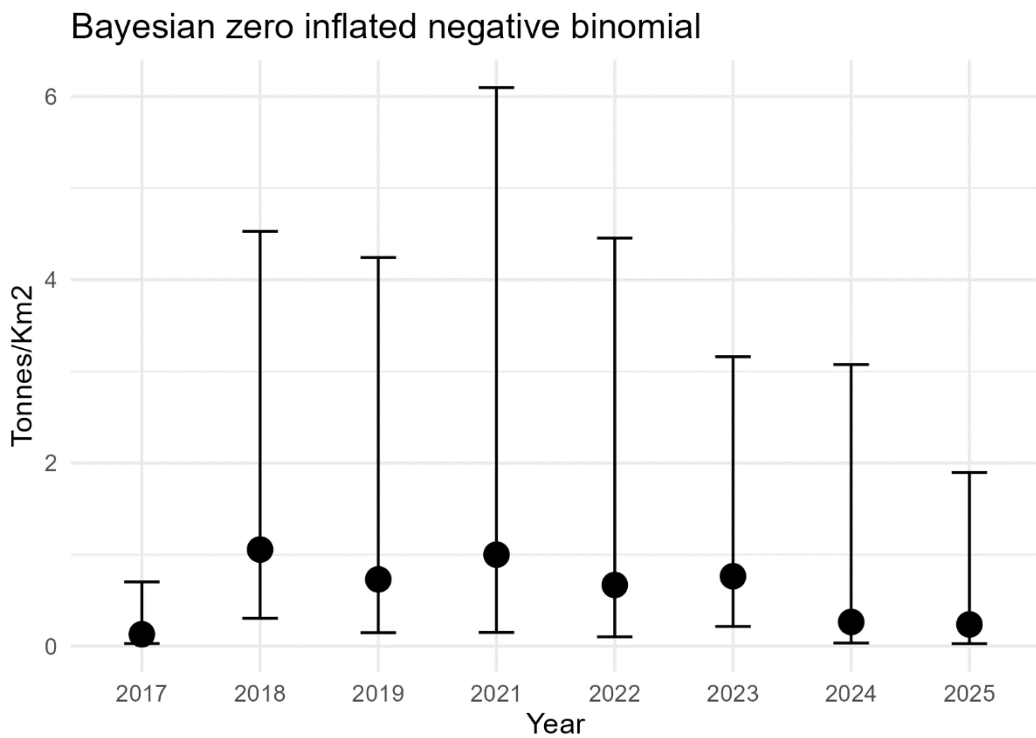


Figure SR2.4. Posterior estimates of tuna biomass with the respective credible Interval. Posterior point estimates and 95% credible intervals.

Table SR2.3. Posterior estimates of tuna biomass with the respective credible Interval. Posterior point estimates and 95% credible intervals.

Year	estimate	lower	upper	sd
2017	0.13032739	4.93940275	0.0289332	0.70164672
2018	1.05574169	33.2619558	0.30470801	4.52769206
2019	0.72947823	30.4814722	0.14785018	4.24260342
2021	0.99959781	43.0431037	0.14994386	6.09755194
2022	0.66849484	28.2784476	0.10191981	4.45511972
2023	0.76412609	24.4098157	0.21634276	3.16024942
2024	0.26467304	12.5975985	0.03529457	3.07560352
2025	0.2379791	10.4270149	0.0274323	1.89736487

Table SR2.4. Summary table of Bayesian Bayesian hurdle-lognormal modell GAM distribution

```
Smoothing Spline Hyperparameters:
                     Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sds(sMidLonMidLat_1)     0.35      0.40     0.01     1.59 1.00      779      814
```

```
sds(sbathy0125_1)    1.34    1.58   0.03   5.45 1.00      933      812
sds(sGradSpat_SSS_d0125_1)   1.68  1.86  0.05   5.69 1.00   757  773
sds(stemp_res_1)    1.52    1.52   0.04   6.09 1.00     1017      860
sds(hu_sMidLonMidLat_1)  0.45    0.46  0.01    1.71 1.00   778     732
sds(hu_sbathy0125_1)   3.14    2.74    0.10   7.76 1.00    629    596
sds(hu_sGradSpat_SSS_d0125_1)  1.15   1.26  0.03  4.56 1.00  769  681
sds(hu_stemp_res_1)     2.48    1.79   0.43   7.32 1.01    634     704

Regression Coefficients:
           Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept   -0.64    0.66   -1.94    0.65 1.00      761      681
hu_Intercept   0.13      0.32    -0.48     0.72 1.00      780      590
Year2018       1.28      0.82    -0.37     2.96 1.00      777      816
Year2019       1.13      1.00    -0.69     3.00 1.01      722      719
Year2021       2.44      1.03     0.40     4.49 1.00      802      860
Year2022       2.00      0.93     0.18     3.76 1.00      848      823
Year2023       0.83      0.75    -0.62     2.41 1.00      745      702
Year2024       0.31      1.21    -2.01     2.72 1.01      776      760
Year2025       1.62      0.97    -0.29     3.43 1.00      864      736
hu_Year2018   -0.81      0.41    -1.61    -0.02 1.00      795      799
hu_Year2019   -0.60      0.54    -1.62     0.48 1.00      857      761
hu_Year2021    0.49      0.50    -0.46     1.46 1.00      763      770
hu_Year2022    0.45      0.50    -0.50     1.41 1.00      793      621
hu_Year2023   -0.93      0.37    -1.65    -0.18 1.00      824      732
hu_Year2024   -0.44      0.58    -1.56     0.75 1.00      677      736
hu_Year2025    1.08      0.48     0.25     2.09 1.00      823      635

Further Distributional Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     2.15      0.16     1.85     2.49 1.00      723      774
```

Bayesian R^2= 0.255

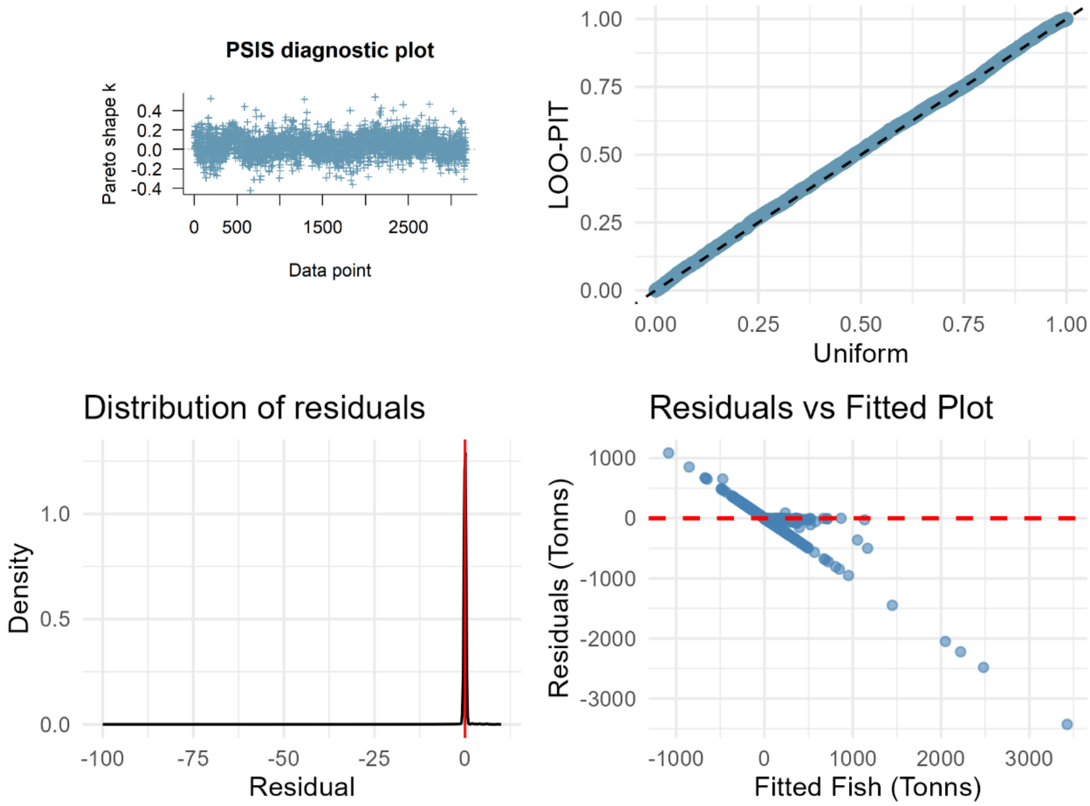


Figure SR2.5. Pareto and residual model diagnostics. A) Pareto diagnostics, all observations have Pareto k values < 0.62. B) Empirical cumulative probabilities vs. theoretical uniform quantiles. C) Posterior distribution of residuals (negative values cropped at -100). D) Residual plot.

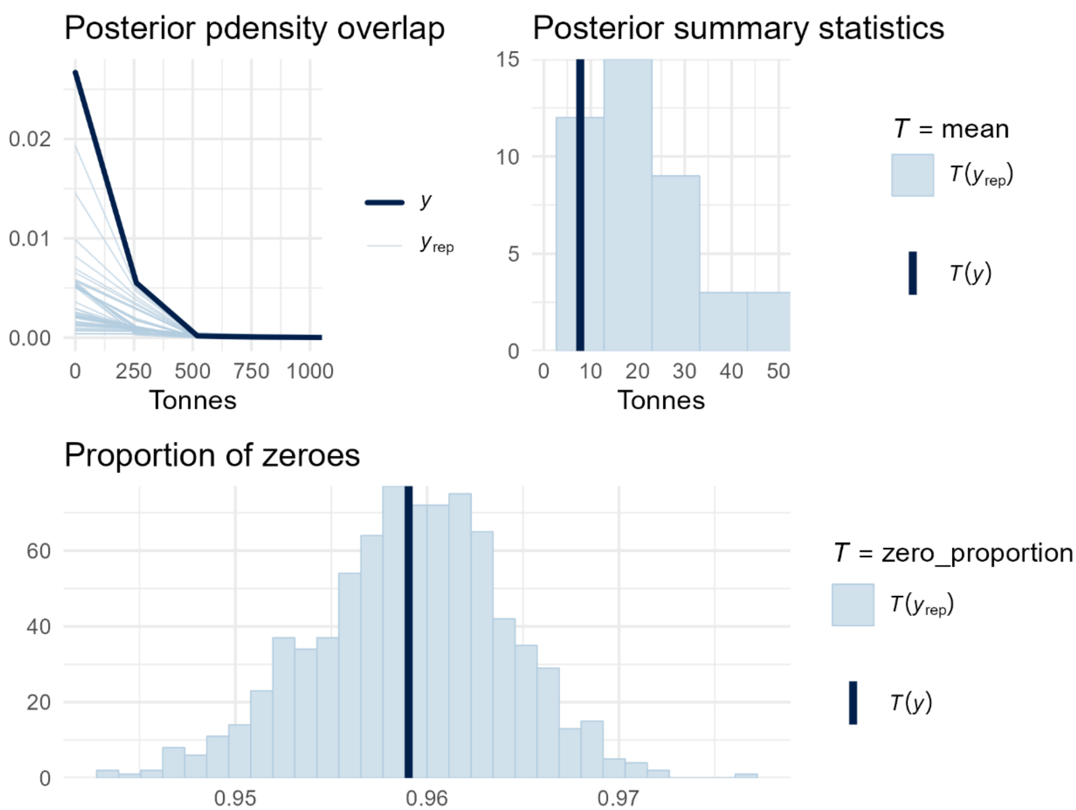


Figure SR2.6. Posterior model diagnostics. A) Posterior density overlay plots (50 samples, cropped at 400 tonnes). In a good fit, the light lines should tightly cluster around the dark line, mirroring its peak, width, and tails. B) Posterior summary statistics plot (cropped at 50 tonnes). In a good fit, the dark vertical line should fall right in the middle or meat of the light histogram. C) Posterior proportion of zeros. In a good fit, the dark vertical line should fall right in the middle of the histogram.

ACKNOWLEDGEMENTS

This work was carried out under the provision of the ICCAT Atlantic-Wide Research Programme for Bluefin Tuna (GBYP). The contents of this document do not necessarily reflect the point of view of ICCAT, which has no responsibility over them, and in no way anticipate the Commission's future policy in this area.

This work was conducted with the funding support of the ICCAT GBYP budget and was partially funded by the European Union through the EU Grant Agreement No. 101320145.