SCIENTIFIC COMMITTEE
TWENTY-SECOND REGULAR SESSION
11-19 August 2026
Apia, Samoa (Hybrid)

PROJECT 120: REVISITING THE REPRODUCTIVE BIOLOGY OF WESTERN AND CENTRAL PACIFIC OCEAN YELLOWFIN TUNA (*Thunnus albacares*)

WCPFC-SC22-2026-SA-IP17
11 July 2026

Submitted by

Sebastien Gislard[1], Joe Scutt Phillips[1], Jessica Farley[2], Marc Ghergariu[1],John Morrongiello[3] ,Joshua Barrow[1], Bianca Gardner[4] ,Roseline Perrot[5] ,Corey Wakefield[1],Joseph Widdrington[1], Alexandre Pons[1],Joanne Potts[1],Tom Peatman[6],Matthew Cunningham[1],Lui Bell[1], Francis Tofukalao[1], Caroline Sanchez[1], Vanille Barthelemy[1], Jone Amoe[1], Bradley Phillip[1], Francois Roupsard[1] & Simon J. Nicol[1]

[1] Oceanic Fisheries Program of the Pacific Community
[2] CSIRO Environment
[3] University of Melbourne
[4] MRAG Asia-Pacific
[5] Laboratoire de la Nouvelle-Calédonie
[6] Shearwater Analytics Ltd, Frome, United Kingdom

Executive summary

Reproductive parameters are fundamental inputs for stock assessment models. However, these parameters often arise from work conducted many years ago. For yellowfin tuna (*Thunnus albacares*) in the Western and Central Pacific Ocean (WCPO), the most recent WCPO-wide reproductive biology estimates are nearly 30 years old. This study presents updated maturity parameters derived from 1,634 female yellowfin tuna sampled between 2005 and 2025 across 21 EEZs. Histological analysis of gonad tissue provided estimates of an individual's maturation stage and fecundity, and when combined with fish age and length data, estimates of age- and length-at-maturity.

The WCPO-wide length at 50% maturity (L_{50}) was estimated at 90.76 cm (95% CI: 88.94 to 92.62), substantially lower than the previously accepted estimate of 104.57 cm. The 14% reduction in this estimate is largely attributable to the adoption of a more comprehensive histological classification system that improves detection of regenerating mature females. Frozen gonad preservation was found to be a suitable alternative to the more traditional formalin-based preservation methods. Only minor differences were detected between the two preservation methods, with formalin-preserved samples producing lower estimates of L_{50} and L_{95} than frozen samples (+3.0% and +7.3%, respectively, for frozen samples), supporting the consideration of frozen preservation as a practical alternative for future sampling programmes.

Age at 50% maturity (A_{50}) was estimated at 1.83 years using both growth-coefficient-derived ages and direct otolith readings, although the latter estimate was based on a smaller sample size (n = 331 vs. n= 1634).

Strong spatial variation in length at maturity was documented across stock assessment regions, with L_{50} increasing from 73.29 cm in the Archipelagic (AC) region of PNG/Solomon Islands to 94.50 cm in the Equatorial (EQ) region, and 101.96 cm in Southern (SO) region. Marked regional differences in spawning seasonality were also observed, ranging from near year-round spawning activity in equatorial regions to a clearly defined seasonal peak between December and February in southern regions.

Results on spawning fraction and spawning frequency will be presented in the final report of the P120 project. Batch fecundity will also be addressed in the final report; however, the limited number of suitable samples (n = 11) is currently insufficient to allow reliable estimation or any meaningful assessment of spatial variation in fecundity parameters. This reflects limited sampling of females in the hours preceding spawning, particularly during night-time periods when yellowfin tuna are known to spawn, as well as practical constraints in obtaining accurate whole gonad weights under field conditions.

Based on these findings, SC22 is invited to consider the following recommendations:

- Note that frozen samples present operational and safety advantages over formalin and support their adoption as the standard preservation method for future sampling programmes.

-Note that incorporating the complete set of spawning indicators in histological classifications reduces misclassification between immature and regenerating females.

-Consider adopting the WCPO-wide maturity ogive presented in this study for future stock assessments, while recognising that future refinements should better account for spatial and temporal variation in length at maturity

-Consider exploring the reported region-specific maturity parameters in management strategy evaluation frameworks for their influence on stock dynamics, given the observed spatial variability in seasonality and length at maturity.

-Endorse further investigation at finer spatiotemporal scales to better characterise regional patterns in maturity, including exploring environmental, fishing pressure, movement and genetic drivers of this observed variability.

-Endorse further investigation of otolith proteomics as a tool to reconstruct individual reproductive histories and investigate temporal variation in reproductive dynamics across the WCPO.

-Note that improved and more targeted sampling is needed across the WCPO, including better regional and size distribution coverage, dedicated night sampling to support fecundity estimation, and otolith ageing from existing paired samples to enable age at maturity analyses.

-Note that the final report of project P120 will be provided to SC23

Introduction

Reproductive parameters are critical inputs for stock assessment models, as they determine estimates of spawning potential and ultimately influence estimates of spawning biomass used in fisheries management decision-making. Uncertainty in these parameters can propagate through assessments and reduce the precision and reliability of stock status indicators for tropical tuna stocks in the Western and Central Pacific Ocean (WCPO).

Life-history traits, including reproductive parameters such as size- and age-at-maturity, fecundity, and spawning dynamics, are plastic and can thus vary across space and through time (Young et al. 2006). For example, the length-at-maturity for yellowfin tuna (*Thunnus albacares*) in the eastern Pacific Ocean (EPO) varied by 10 to 24 % across the full north-south and east-west extent of the EPO and by 8-26 % over a 20-year period (Schaefer and Fuller 2022). The most comprehensive assessment of yellowfin tuna maturity in the western and central Pacific Ocean (WCPO) was conducted nearly 30 years ago (Itano 2000). Some more recent studies of yellowfin tuna reproductive biology have been undertaken, but their utility for stock assessment models is limited because sample collection has often been restricted in spatial coverage (Shi et al. 2022). Generating these contemporary datasets at the scale required for spatially explicit stock assessment presents its own logistical challenges. Histological assessment of maturity conventionally relies on formalin-fixed tissue, which remains the gold standard for preserving the cellular structure needed for accurate reproductive staging (Musiał et al. 2016). However, formalin is not adapted to sample collection within the fisheries industry, as it requires specialised handling expertise and poses a risk in proximity to fish destined for human consumption. Frozen storage offers a more practical alternative for at-sea or port sampling, but its suitability for histological maturity assessment has not been well established. This study therefore also evaluates whether frozen preservation can serve as a viable alternative to formalin fixation, to support broader-scale sample collection in future monitoring programmes

The lack of recent broad-scale estimates of key reproductive parameters highlights the need to re-estimate these parameters using contemporary datasets, standardised ovary histology protocols, and improved analytical approaches. This study estimates maturity parameters for yellowfin tuna across the whole WCPO stock and then by stock assessment spatial strata (hereafter "regions") to support spatially explicit stock assessment frameworks.

Material and Method

Samples collection

The ovary samples analysed in this study were obtained from the Pacific Marine Specimen Bank (PMSB; Project 35b) (SPC-OFP 2019). The PMSB is a sample repository populated through coordinated sampling efforts by observers, scientists and other trained samplers operating under the Western and Central Pacific Fisheries Commission (WCPFC), national observer programmes, and fisheries agencies from Pacific Island Countries and Territories (PICTs). These samples are collected from commercial fisheries operating across the Western and Central Pacific Ocean (WCPO) and Pacific tuna research cruises implemented by SPC (SPC-OFP 2026)

Samples were collected either at sea, aboard purse seiners and longliners during commercial fishing operations, on multipurpose fishing vessels chartered for scientific cruises, or at port during the unloading of fishing vessels operating in the WCPO. When sampling occurred at sea, each fish was tagged with a cable tie labelled with a unique identifier, secured either through the gills or around the caudal peduncle, and associated metadata were recorded. When sampling occurred at landing, each fish was similarly tagged with a uniquely labelled cable tie. Fish were either identified by the onboard observer prior to unloading, with associated fishing metadata obtained from the vessel logsheet, or sampled from a designated well and subsequently backtraced to the corresponding fishing set using vessel logsheet information.

Samples originated from predefined geographic strata based on the Western and Central Pacific Ocean (WCPO) stock assessment regions (Magnusson et al. 2023). These strata are hereafter referred to as (Figure 2): Northern (Region 1; NO), Western (Region 2; W), Archipelagic (Region 3; AC), Equatorial (Region 4; EQ), and Southern (Region 5; SO). Samples from the Eastern Pacific Ocean (EPO) constituted a distinct stratum from the WCPO regions.

Fish were either processed fresh immediately on board or preserved following vessel-specific procedures prior to sampling. Whole fish preservation involved either freezing the fish at −20 °C, or placing it fresh on ice or in brine, until processing at port. Fork length (FL) was measured to the nearest centimetre using callipers. The ventral cavity was then opened by a longitudinal incision extending from the anus to the pelvic fins, and the gonads were carefully separated from surrounding tissues and removed.

Extracted gonads were preserved either in 10% buffered formalin or placed in plastic bags and frozen at −20 °C. A subset of samples was preserved using both methods to assess the potential effects of preservation technique on maturity staging and reproductive phase determination.

Histology preparation and classification

A cross section of the central part of preserved ovaries was then removed, placed in a histology cassette and fixed in 10% formalin for 24 -48h. Following fixation, samples were embedded in paraffin wax to provide structural support for sectioning. The paraffin blocks were then sectioned

using a microtome at approximately 4 µm thickness. Sections were mounted onto microscope slides and stained with haematoxylin and eosin (H&E).

Histological slides were initially examined at low magnification using a Leica M60 stereomicroscope and subsequently analysed using a Leica DMLB microscope at magnifications ranging from ×50 to ×100.

Females were classified into one of 7 reproductive phases using terminology from Brown-Peterson et al. (2011) and adapted to tropical tunas (Farley et al. 2014) (**Figure 1,Table 1**)

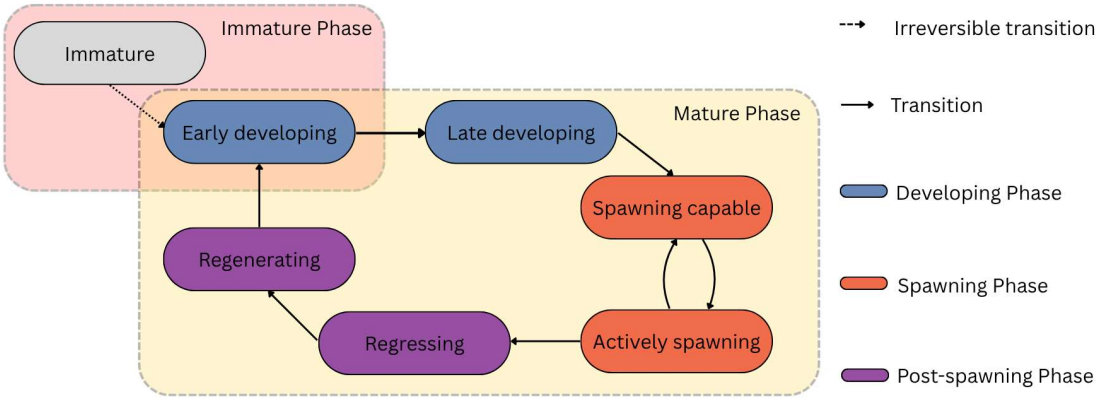

Figure 1. Reproductive maturity cycle based on histological staging following (Brown-Peterson et al. 2011) and adapted to tropical tunas (Farley et al. 2014). Stages are grouped into two broad phases: immature and mature. Arrows indicate the directionality of stage progression, with the regenerating stage looping back into early developing, reflecting the cyclic nature of fish reproduction. Colour coding represents three reproductive subgroups: developing, spawning and post-spawning.

The classification was based on three primary criteria: (1) the most advanced group of oocytes (MAGO), (2) the presence or absence of indicators of previous spawning activity (IPS), and (3) the presence and extent of atretic vitellogenic oocytes.

Females were classified as mature when the MAGO consisted of Primary, Secondary or tertiary vitellogenic oocytes (respectively Vtg1, Vtg2, Vtg3), migratory nucleus (MN), or hydrated oocytes (Hyd), or when evidence of previous spawning activity was observed, including residual hydrated oocytes (RHO), post-ovulatory follicles (POFs), muscle bundles, brown bodies, or increased ovarian wall thickness.

Both mature and immature females were classified as developing when the MAGO was at the cortical alveolar (CA) stage and in such cases, maturity status was determined based on the presence or absence of IPS. Females were classified as spawning capable when vitellogenic oocytes were present, fewer than 50% of these oocytes were atretic, and post-ovulatory follicles (POFs) were absent, whereas females were considered actively spawning when oocytes had reached the migratory nucleus or hydrated stage, or when POFs were present. Females were classified as regressing when more than 50% of vitellogenic oocytes were undergoing atresia. Finally, females with oocytes in primary growth stages were classified as regenerating if IPS were present, and as immature if IPS were absent.

Table 1 Histology-based classification of reproductive phases and maturity status in yellowfin tuna. Dominant oocyte stage, most advanced group of oocytes (MAGO), presence of post-ovulatory follicles (POF), atresia, and

indicators of prior spawning (IPS) were used to assign reproductive phases. PG = primary growth oocytes; CA = cortical alveolar oocytes; Vtg1–3 = vitellogenic oocyte stages 1–3; MN = migratory nucleus; Hyd = hydrated oocytes; POF = post-ovulatory follicles; IPS = indicators of prior spawning.

Reproductive Phase	Maturity status	Dominant oocyte	MAGO	POF	Atresia	IPS
Immature	Immature	PG	PG	Absent	Absent	Absent
Early Developing		PG	CA	Absent	Absent	Absent
Early developing	Mature	PG	CA	Absent	< 50%	Present
Late developing		PG/CA	Vtg1	Absent	< 50%	May be present
Spawning capable		Vtg3	Vtg3	Absent	< 50%	May be present
Actively spawning		Vtg3	Vtg3	Present	< 50%	May be present
		Vtg3	MN/Hyd	Present or Absent	< 50%	May be present
		MN/Hyd	MN/Hyd	Present or Absent	< 50%	May be present
Regressing		PG/CA/Vtg1,2,3	PG/CA/Vtg1,2,3	Absent	> 50%	May be present
Regenerating		PG	PG	Absent	< 50%	Present

Assessing robustness of maturity data

Comparison between readers

To assess the consistency of maturity classifications between readers, a subset of the slides were reviewed independently by two readers who were blind to each other's assessments: Reader 1 was highly experienced (>2000 previous readings), and Reader 2 was newly trained (0 previous readings).

McNemar's test was initially used to assess whether the two readers differed systematically in their binary maturity classifications, where 0 indicated immature and 1 indicated mature. The test evaluates whether the two possible types of discordant classification occurred with equal frequency: that is, whether Reader 1 classified a fish as mature and Reader 2 classified it as immature as often as the reverse. Statistical significance was assessed using a threshold of $p<0.05$.

To determine whether differences between readers varied with fish length, the probability of a mature classification was modelled using a generalised estimating equation (GEE) with a binomial error distribution and logit link. Let $Y_{ij} \in \{0,1\}$ denote the maturity classification for fish i and reader j, where $Y_{ij}=1$ indicates that the fish was classified as mature and $Y_{ij}=0$ indicates that it was classified as immature. An exchangeable working correlation structure was specified to account for the correlation between the two classifications of the same fish.

The maturity classification was assumed to follow a Bernoulli distribution:

$$Y_{ij} \sim Bernoulli(\pi_{ij})$$

where π_{ij} is the probability that fish *i* by reader *j* is classified as mature:

$$\pi_{ij} = \frac{1}{1 + \exp(-\eta_{ij})} \quad (1)$$

Here, η_{ij} is the linear predictor representing the log-odds that fish *i* is classified as mature by reader *j*, conditional on the other explanatory variables (e.g., length) included in the model.

Three model forms were considered. The first included fish length only:

$$\eta_{ij} = \beta_0 + \beta_1 Length_i \quad (2)$$

The second included additive effects of fish length and reader:

$$\eta_{ij} = \beta_0 + \beta_1 Length_i + \beta_2 Reader_j \quad (3a)$$

The third included an interaction between fish length and reader:

$$\eta_{ij} = \beta_0 + \beta_1 Length_i + \beta_2 Reader_j + \beta_3 (Length_i \times Reader_j) \quad (3b)$$

For all GEE analyses, fish length was centred on its observed mean before model fitting. This ensured that the intercept represented the expected log-odds of a mature classification for a fish of average length and improved the numerical stability of model estimation. Model comparison was based on the quasi-likelihood under the independence model criterion (QIC).

Comparison between preservation methods

To assess the consistency of maturity classifications between preservation methods, paired samples from 316 females were assessed by a single reader. The analysis followed the same structure as the reader comparison, with preservation method replacing reader as the factor of interest. The model formulations were:

$$\eta_{ij} = \beta_0 + \beta_1 Length_i + \beta_2 Preservation_m \quad (4a)$$

$$\eta_{ij} = \beta_0 + \beta_1 Length_i + \beta_2 Preservation_m + \beta_3 (Length_i \times Preservation_m) \quad (4b)$$

Length at maturity

The maturity ogive was modelled using the binomial logistic regression framework described in **Eq.1** and for which the simplest model from **Eq.2** was retained.

The length at which a given proportion of females were mature was obtained from rearranging the logistic regression equation of **Eq.1**:

$$Length_P = \frac{log\left(\frac{P}{1-P}\right) - \beta_0}{\beta_1} \quad (5)$$

Where $Length_P$ is the fork length at which a given proportion P of females is mature

From this relationship, L_{50} and L_{95} were estimated as the fork lengths at which 50% and 95% of females were predicted to be mature, respectively.

Confidence intervals (95%) for β_0, β_1 L_{50}, and L_{95} were obtained by non-parametric bootstrap resampling, where observations were resampled with replacement from the original dataset over 2,000 iterations.

Age at maturity

Age (in years) was estimated for each individual female using the following form of the Von Bertalanffy Growth Function (VBGF):

$$\text{Age} = t_0 \times \frac{1}{-K} \times \log\left(1 - \frac{\text{Length}}{L_\infty}\right) \quad (6)$$

Where *Length* is fork length (cm), t_0(-0.13), K (0.53) and L_∞(142.67) are the theoretical age at length zero, the growth coefficient, and the theoretical maximum length, respectively, and were estimated from analyses associated with the most recently adopted stock assessment model for WCPO yellowfin tuna (Magnusson et al. 2023). For females with lengths exceeding L_∞, age was assigned as the maximum age predicted by the model (i.e., the asymptotic age value).

Age was also estimated from otolith sections for a subset of females that had been assessed for maturity using gonad histology and for which paired otolith samples were also available (Wakefield et al. 2026). Two maturity ogives were therefore fitted using the same binomial logistic regression framework as described for length-at-maturity, with age substituted for fork length as the explanatory variable:

$$\eta_i = \beta_0 + \beta_1 Age_i \quad (7)$$

The first model used otolith-derived ages, while the second used ages predicted from the growth model.

Regional variation in length-at-maturity

Potential regional differences in length-at-maturity were evaluated by including Region as an explanatory term in the logistic regression model:

$$\eta_{ik} = \beta_0 + \beta_1 \text{Length}_i + \beta_2 \text{Region}_k \quad (8a)$$

where k indexes the region.

This model was compared using likelihood ratio tests and Akaike Information Criterion (AIC) with the simplest model form shown in (**Eq. 2**) and with an interactive model of the form:

$$\eta_{ik} = \beta_0 + \beta_1 Length_i + \beta_2 Region_k + \beta_3 (Length_i \times Region_k) \quad (8b)$$

Pairwise regional comparisons in length-at-maturity were assessed using model-based estimated marginal means (EMMs) derived from the full binomial logistic regression model described above.

GSI

The gonado-somatic index provides a measure of reproductive condition and is widely used as an indicator of spawning seasonality, reflecting increases in gonad weight as oocytes develop during the reproductive cycle. Only sexually mature females were included in these analyses (n = 688), as the presence of immature females, which have undeveloped gonads, may bias estimates of reproductive condition downward during the spawning period. Females weight was estimated from length using published length–weight coefficients (Macdonald et al. 2023). Gonads were weighed either fresh at port or after freezing prior to processing. The gonadosomatic index (GSI) was calculated as follows and expressed as a percentage of total body mass.

$$GSI = \frac{gonad\ weight}{fish\ weight} \times 100 \tag{9}$$

To evaluate whether the condition in which gonads were preserved prior to weighing (fresh vs. frozen) influenced gonad weight, while accounting for biological and spatio-temporal variation, a series of Gaussian generalised linear models (GLMs) were fitted.

The primary model tested the effect of preservation state (fresh or frozen) on gonad weight, with fork length, month of capture, and stock assessment region included as explanatory variables:

$$GW_i = \beta_0 + \beta_1 Length_i + \beta_2 Region_k + \beta_3 Month_l + \beta_4 Preservation_m \tag{10a}$$

where $GW_i$is the gonad weight of fish i, $\mu_i$its expected value, σ^2 the residual variance, $\beta_0$the intercept, and Length_i, Region_k, Month_l,and Preservation_m represent fork length for individual i, and the stock assessment region, the month of capture and preservation state corresponding to the respective categorical levels k, l, and m, respectively

To assess whether accounting for potential spatial variation in seasonal patterns improved model fit, an additional model including a month-by-region interaction was also considered:

$$\mathrm{GW_i} = \beta_0 + \beta_1 \mathrm{Length_i} + \beta_2 \mathrm{Region_k} + \beta_3 \mathrm{Month_l} + \beta_4 \mathrm{Preservation_m} + \beta_5 (\mathrm{Region_k} \times \mathrm{Month_l}) \tag{10b}$$

A fully interactive model was also explored as a sensitivity analysis:

$$\begin{aligned} \mathrm{GW_i} = {} & \beta_0 + \beta_1 \mathrm{Length_i} + \beta_2 \mathrm{Region_k} + \beta_3 \mathrm{Month_l} + \beta_4 \mathrm{Preservation_m} \\ & + \beta_5 (\mathrm{Preservation_m} \times \mathrm{Length_i}) + \beta_6 (\mathrm{Preservation_m} \times \mathrm{Month_l}) \\ & + \beta_7 (\mathrm{Preservation_m} \times \mathrm{Region_k}) \end{aligned} \tag{10c}$$

Model selection was based on Akaike Information Criterion (AIC) and likelihood ratio tests to compare competing candidate models. The effect of preservation method on gonad weight was then assessed by comparing a full model with a reduced model excluding the preservation term, again using AIC and a likelihood ratio test. Finally, t-tests were used to evaluate the significance of individual coefficients in the best-supported model.

To examine seasonal variation in reproductive condition independently of preservation effects, the gonadosomatic index (GSI) from all selected samples was modelled as a function of fork length and month of capture separately for each region, using the same Gaussian GLM framework and inference approach described above.

Software

All data analysis were performed in R version 4.5.0 R Core Team 2025). The 'geepack' package (v1.3.13,Højsgaard et al. 2006; Yan 2002; Yan and Fine 2004; Xu et al. 2025) was used to fit the GEEs and perform model selection.

Results

Sample collection

In this study, a total of 1634 yellowfin tuna females were sampled between 2005 and 2025 across 442 fishing trips and during 800 fishing events. Females ranged in size from 54 to 228 cm recorded fork length. Of these samples 54.4% were caught by longline fisheries and 24.5% by purse-seine fisheries. The remaining samples were associated with a variety of less common capture methods including dropline, vertical longline, pole-and-line, handline, and reel-and-rod fishing, collected either opportunistically during non-commercial fishing events or as part of tagging cruises. Sampling covered a broad geographic extent spanning 21 Exclusive Economic Zones (EEZs) and 8 high seas areas, with curved surface area of sampling locations encompassing approximately 56.5 million km^2 **(Figure 2)**. A total of 35 additional samples were collected outside the boundaries of the Western and Central Pacific Ocean (WCPO) and were excluded from this study.

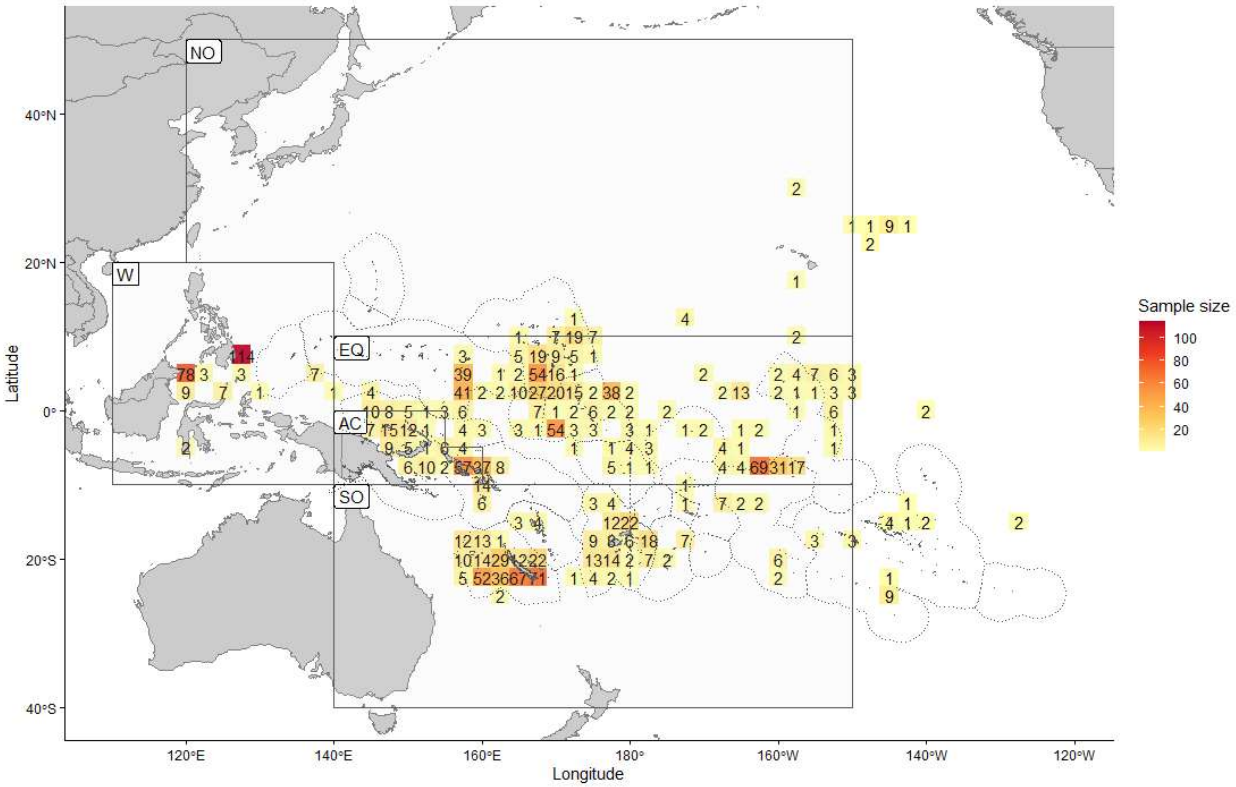

***Figure 2. Spatial distribution of female yellowfin tuna samples analysed in this study**. The heatmap is aggregated on a 1° × 1° latitude-longitude grid, with cell values representing the number of samples within each grid cell. Dotted lines indicate Exclusive Economic Zones, while solid lines represent the five stock assessment regions (NO: Northern, W: Western, EQ:Equatorial,AC:Archipelagic,SO:Southern) used for yellowfin tuna in the Western and Central Pacific Ocean.*

To assess spatial differences in sampling structure and reproductive composition, length frequency distributions and reproductive phase composition were examined across geographic strata (**Figure 3**). Sample sizes varied among regions, with limited observations in the NO and EPO regions (n = 18 and n = 35, respectively) and only a very small number of females recorded in the W region (n = 5).

Marked differences in reproductive composition were observed among strata (**Figure 3**). The Southern Ocean (SO) region was dominated by immature or regenerating females (80%), whereas the Western region was composed predominantly of spawning-capable and actively spawning females (92%). In contrast, the AC and EQ regions showed a mixed reproductive composition. These patterns suggest potential differences in reproductive timing among regions; however, they should be interpreted with caution, and a more comprehensive analysis of seasonal variation is presented later in the manuscript. However, in the AC region no length class consisted exclusively of immature females, with the 60-70 cm size class still containing a small proportion of regenerating females. This may affect subsequent length-at-maturity estimates, as limited representation of fully immature females size classes may reduce precision in characterising the lower tail of the maturity ogive.

Differences in size at maturity were also evident among regions. The smallest spawning capable or actively spawning females were 65, 62, and 105 cm for the AC, EQ, and SO regions, respectively. Similarly, the largest immature females were 97, 122, and 137 cm for the AC, EQ, and SO regions, respectively. These differences provide a preliminary indication of spatial variation in maturation patterns. However, these values represent extreme observations (smallest and largest females) and should be interpreted with caution, as they may not be representative of the overall population. More detailed analyses of spatial effects across strata are presented later in the manuscript.

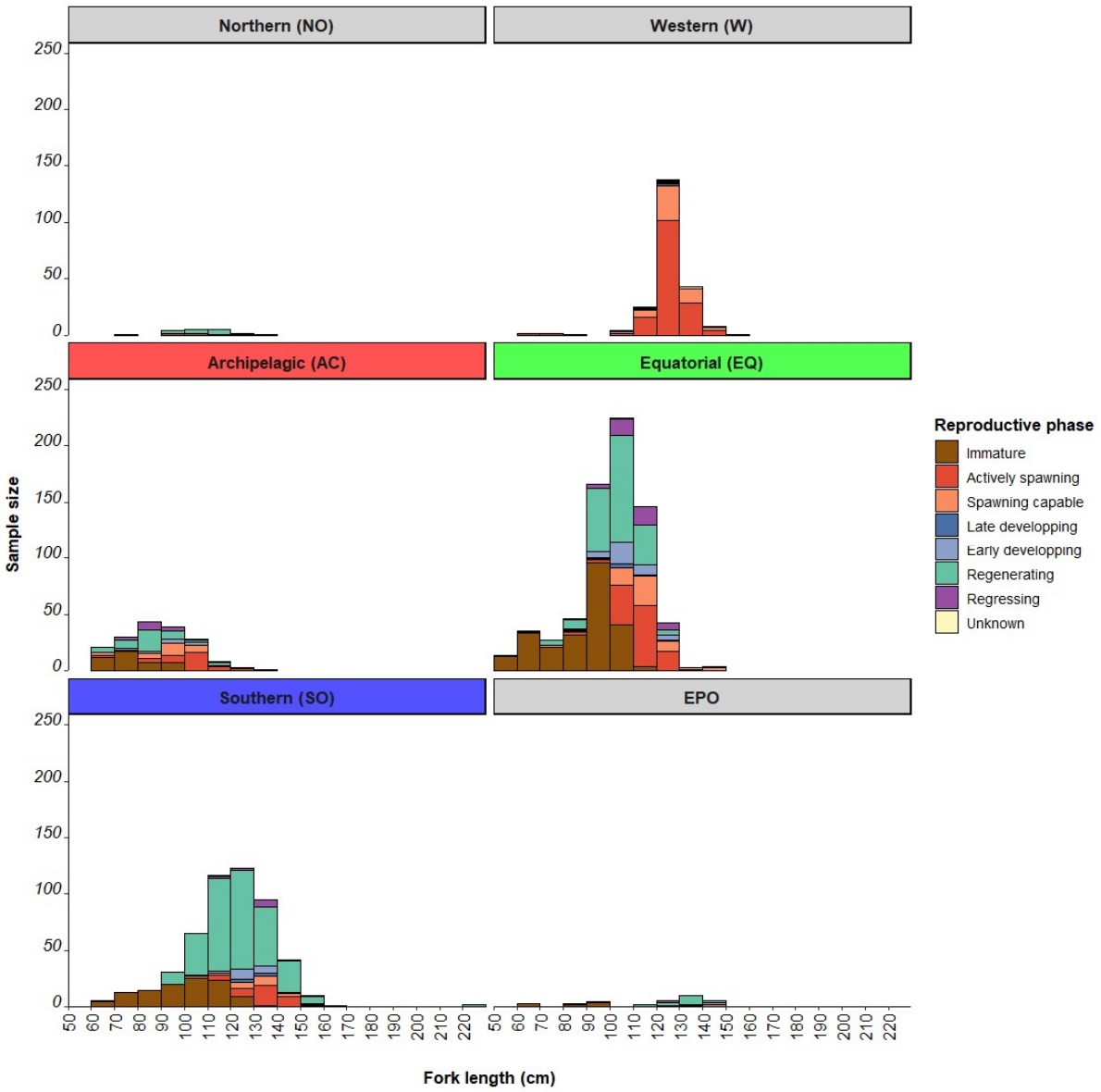

Figure 3. Length-frequency distribution by reproductive phase of female yellowfin tuna in the Western and Central Pacific Ocean (WCPO). Stacked bars represent the absolute number of female individuals in each reproductive phase within each 10 cm length class.Panels are stratified by geographic region and colours represent sampling regions, with red representing the Archipelagic region (AC), green the Equatorial region (EQ), and blue the Southern region (SO). This colour scheme is consistent across all figures in the manuscript

Method confirmation

Double blind reading assessment

In total, 9.4% of slides (n=157 slides) were reviewed independently by two readers. Reader agreement was high, with 90.5% of double-blind read slides receiving the same maturity classification (mature vs immature) from both readers, and no statistically significant difference in maturity assessments between the two readers (p-value = 0.12) **(Figure 4)**. Model comparison using the QIC quasi-likelihood under the independence model criterion (QIC) provided little support for including reader effects. The model excluding reader had the lowest QIC (214.14), while the additive and interaction models had slightly higher values (215.73 and 218.08, respectively; Δ_{QIC} = 1.58 and 3.93). Although the interaction model (**Eq.3b**) offers greater flexibility by allowing the effect of length to vary between readers, there was little evidence that this

additional complexity improved model fit. Consistent with this, Wald tests indicated that neither the reader main effect nor the length-by-reader interaction was statistically significant (respectively p =0.22 and p =0.92). Consequently, reader effects (**Eq.3a,3b**) were not retained in the final model (**Eq.2**). Nevertheless, the interaction model (**Eq.3b**) was retained for visualization purposes, as it allows reader-specific maturity ogives to be displayed and facilitates comparison of the fitted relationships between readers.

The remaining 9.5% of differently staged slides predominantly came from females in the 100-105 cm size range (50% of differing cases). This length range corresponded to a size at which maturity staging was inherently ambiguous, as reflected by 65% of assessments being classified as low confidence by both readers. Given this overall agreement, only readings from Reader 2 were retained for the remainder of the study.

In addition to the binary maturity classification, mature females received the same reproductive phase classification in 89% of cases, with all remaining discrepancies differing by only one adjacent stage. This high level of agreement indicates consistency in reproductive phase staging among mature females between readers.

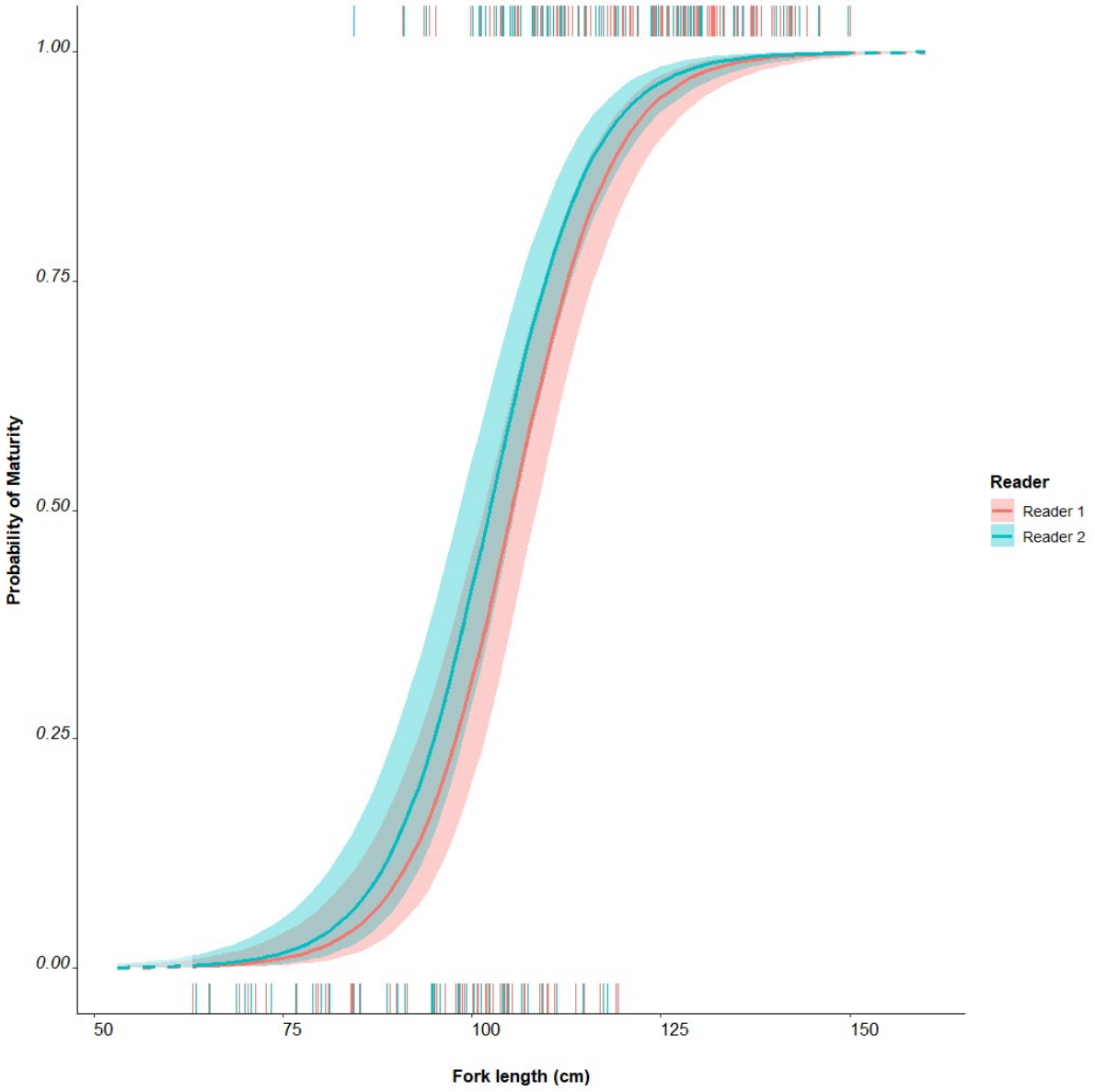

Figure 4. Maturity ogive comparison between independent readers. Solid lines show the fitted length-at-maturity relationships for each reader, and shaded ribbons represent the associated uncertainty around the fitted curves. Colours distinguish between the two independent readers. Rug ticks represent individual females, with those at the top indicating mature females and those at the bottom indicating immature females.

Assessing preservation method suitability

A total of 19.0% (n=317) of gonad samples were subsampled and preserved both in formalin (fresh-preserved) and frozen (frozen-preserved) to evaluate whether frozen samples are suitable for histological assessment. Agreement in maturity status between preservation method was high with 83.6% of samples assessed identically. The remaining 16.4% disagreement was primarily associated with fish larger than 90 cm, which accounted for 86.5% of all discordant classifications.

McNemar's test indicated a statistically significant difference in maturity assessment between preservation methods (p-value= 0.04). Models including preservation method (**Eq. 4a, Eq. 4b**) showed only marginal differences in QIC relative to the length-only model ($\Delta_{QIC} < 1.7$), indicating

similar predictive performance across models. However, the preservation term was statistically significant (Wald test, p = 0.01), suggesting a detectable effect of preservation method. Although the additive model was marginally favoured over the interaction model (Δ_{QIC} = 1.37), and the comparison between models showed no significant difference (Wald test, p = 0.25), the interaction term model (**Eq. 4b**) was retained because it allows β_1 to vary between preservation methods, consistent with the observed pattern of disagreement being concentrated among females larger than L_{50}.

Although a statistically significant difference was detected between formalin-preserved and frozen-preserved samples, the estimated L_{50} of formalin samples was only 3.0% smaller than frozen samples. There was a greater divergence at larger sizes, where the L_{95} of formalin preserved samples was 7.39% smaller than frozen samples **(Figure 5)**.

While these results suggest a small effect of preservation method on maturity assessment, particularly among larger fish, further sampling is needed to better characterise this effect and to investigate contributing factors such as the duration of frozen storage prior to histological processing.

Agreement in reproductive phase assignment for females between the two preservation methods was high, with 75.8% of samples staged identically. Of the 24.2% of samples that were assigned different reproductive phase between preservation methods, 11.3% involved discrepancies among the early developing, late developing, and regressing stages. These mismatches are not expected to affect estimates of length at maturity or fecundity, as they occur primarily within transitional stages between spawning capable and regenerating phases, which do not alter the classification of mature versus immature females or the identification of reproductively active females. These mismatches likely reflect the cyclic nature of gonadal development and the difficulty of assigning transitional individuals, where it is not always possible to determine whether an ovary is progressing toward maturation or regressing toward recovery. A further 8.1% of samples were classified differently between the spawning capable and actively spawning stages. As these stages are used to estimate spawning fraction and spawning frequency, the source of these discrepancies will be investigated further before these metrics are calculated and reported.

The remaining 4.8% of discrepancies corresponded to samples classified as late developing when preserved in formalin but as spawning capable when stored frozen. This difference is likely attributable to the difficulty of distinguishing primary from tertiary vitellogenic oocytes. The main criterion used to differentiate these stages is the distribution of vitellogenic yolk within the oocyte, which appeared less distinct in some frozen samples, potentially leading to misclassification.

Overall, most discrepancies occurred between adjacent or transitional maturity stages and are therefore unlikely to substantially influence the study outcomes. The analyses of reproductive seasonality may be the most sensitive to these differences; however, transitional stages are not central to the interpretation of the main reproductive patterns. Potential effects on spawning fraction and spawning frequency estimates will be examined further.

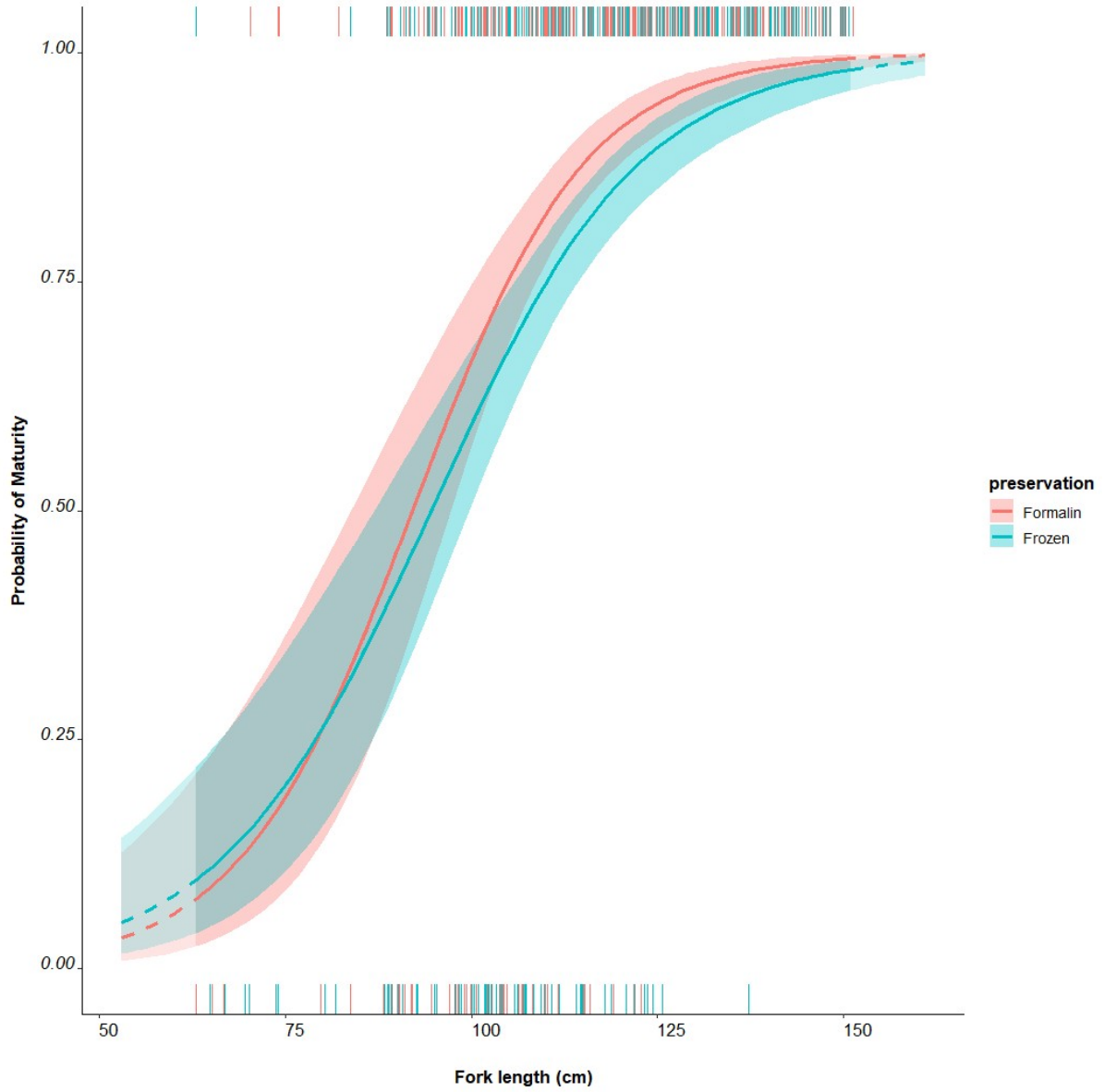

Figure 5 Maturity ogive comparison between preservation methods. Solid lines show the fitted length-at-maturity relationships for each reader, and shaded ribbons represent the associated uncertainty around the fitted curves. Dashed lines and lighter shaded areas indicate model predictions beyond the range of the observed data, representing extrapolation of the fitted relationships. Colours distinguish the two-preservation method. Rug ticks represent individual females, with those at the top indicating mature females and those at the bottom indicating immature females.

Reproductive biology parameters

General maturity ogive for the WCPO

The smallest mature female measured 62 cm FL, whereas the largest immature female measured 137 cm FL. Overall, 25.76% of sampled female were classified as immature and 74.24% as mature. The length-at-maturity relationship followed the expected sigmoidal pattern (**Figure 6)**, with an estimated L_{50} of 90.76 cm (95% CI: 88.94 to 92.62)**.** Model coefficients are provided in **Table 2**.

Age-at-maturity relationships, estimated using the established, WCPO-wide age-length relationship, showed a similar sigmoidal pattern. Using growth-coefficient-derived ages

(n=1634), the estimated A_{50} and A_{95} were 1.83 (95% CI: 1.75 to 1.91) and 3.38 (95% CI: 3.21 to 3.56) years, respectively **(Figure 7a)**.

Using otolith section age readings (n = 331), the estimates were consistent with those obtained using growth-coefficient-derived ages. The estimated A_{50} was 1.83 years (95% CI: 1.53 to 2.05), matching that obtained using growth-coefficient-derived ages, while A_{95} was higher at 4.13 years (95% CI: 3.59 to 4.77) **(Figure 7b)**.

The corresponding age at maturity coefficients are provided in **Table 2.**

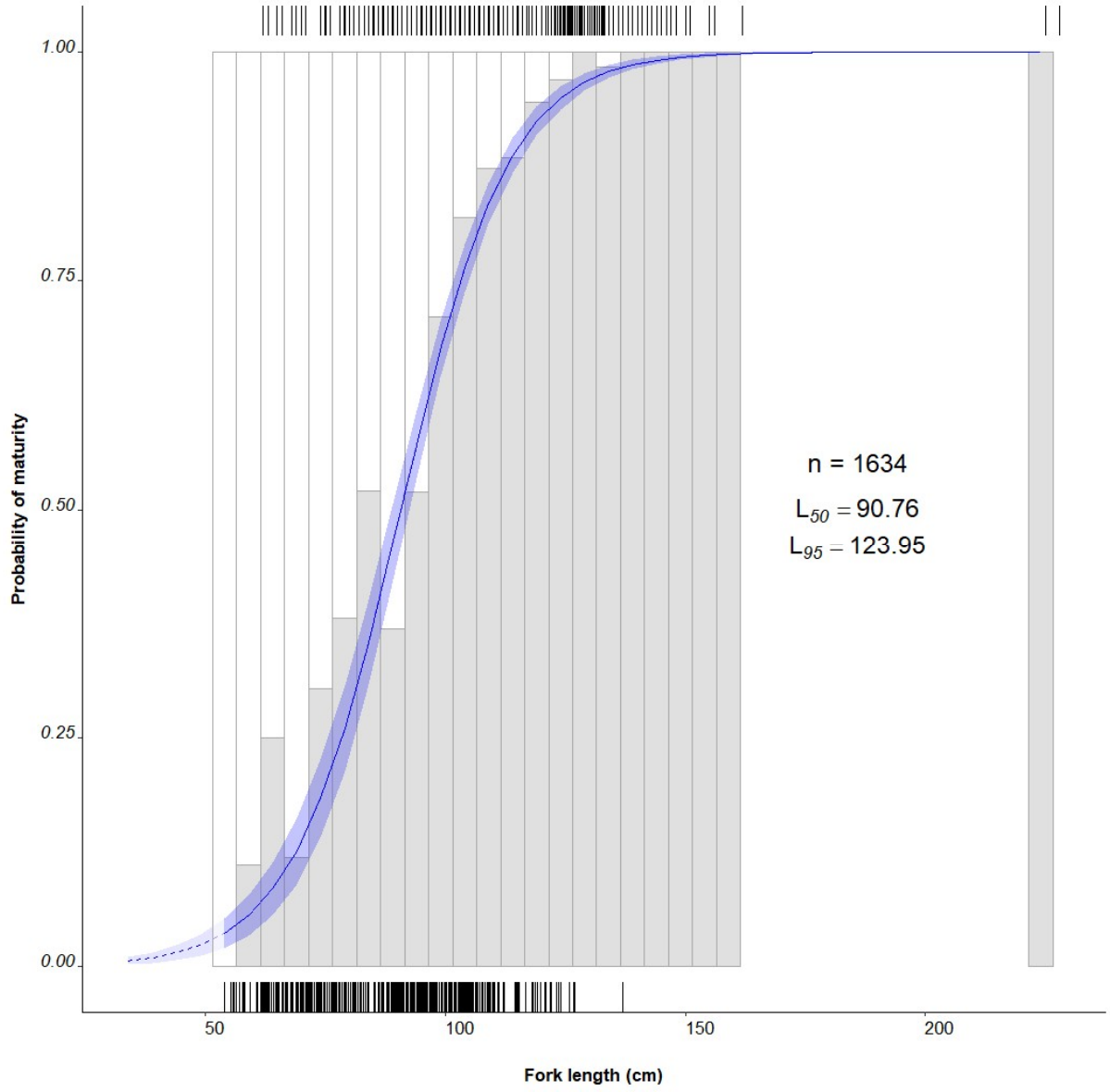

Figure 6. Length at maturity ogive for yellowfin tuna in the Western and Central Pacific Ocean. Solid blue line shows the fitted length-at-maturity relationships, and shaded blue ribbons represent the associated uncertainty around the fitted curves. Dashed blue line and lighter blue shaded areas indicate model predictions beyond the range of the observed data, representing extrapolation of the fitted relationships. Bars represent the proportion of mature (grey) and immature (white) females . The number of observations is indicated by *n*. Rug ticks represent individual females, with those at the top indicating mature females and those at the bottom indicating immature females.

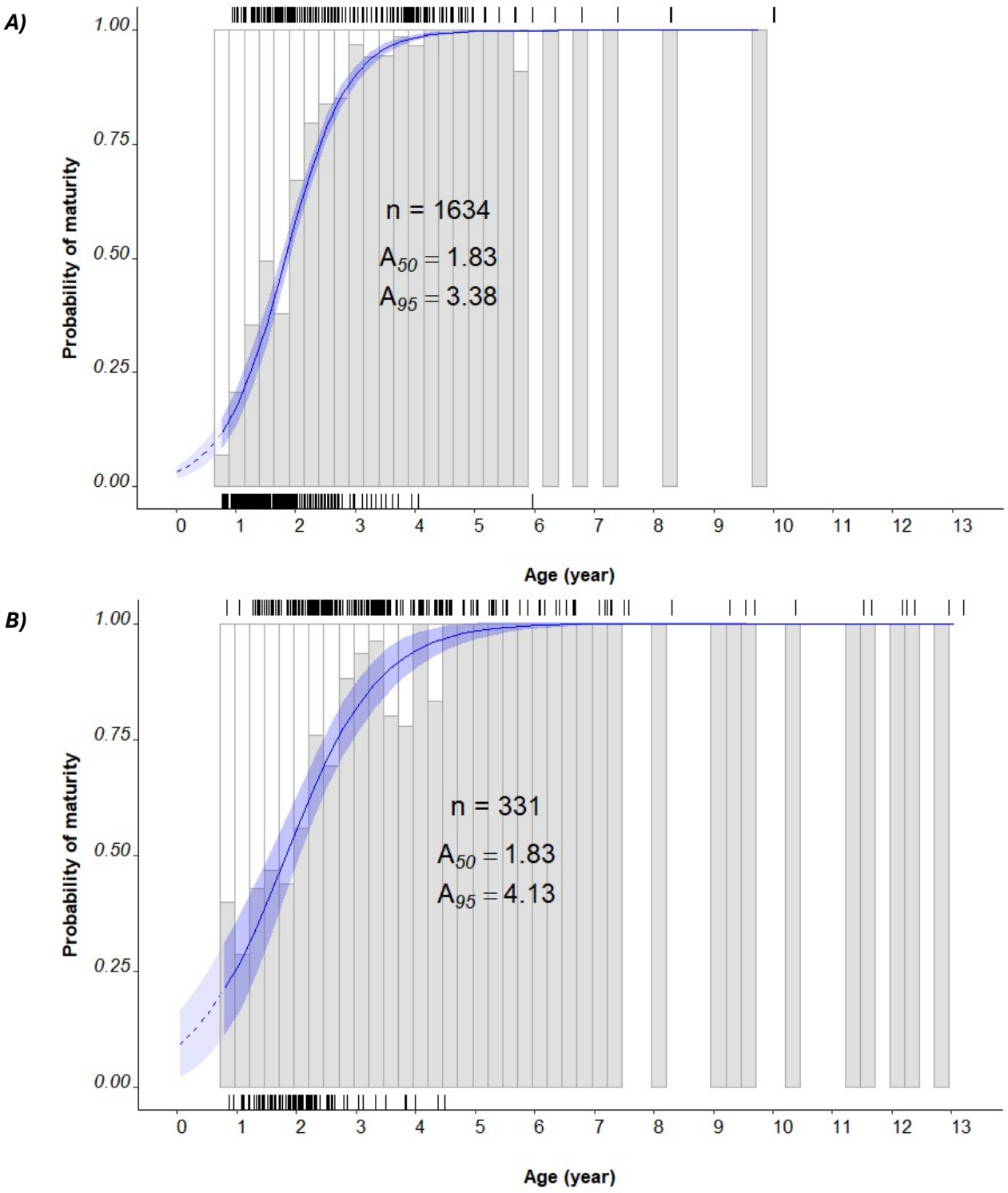

Figure 7 Age-at-maturity ogive for yellowfin tuna in the Western and Central Pacific Ocean. (*A*) Age derived from Von Bertalanffy growth function; (*B*) Age derived from otolith ageing. The number of observations is indicated by *n*. Rug ticks represent individual females, with those at the top indicating mature females and those at the bottom indicating immature females. Solid blue line shows the fitted length-at-maturity relationships, and shaded blue ribbons represent the associated uncertainty around the fitted curves. Dashed blue line and lighter blue shaded areas indicate model predictions beyond the range of the observed data, representing extrapolation of the fitted relationships. Bars represent the proportion of mature (grey) and immature (white) females within 5 cm length classes.

Maturity ogives per stock assessment region

There were clear spatial differences in length-at-maturity among Archipelagic (AC), Equatorial (EQ) and Southern (SO) regions in the WCPO **(Figure 8)**, where the majority of samples from this study were collected. This was supported by both AIC comparison and a likelihood ratio test comparing binomial logistic regression models with and without region (Δ_{AIC} = 127.34, $p < 0.001$). Although the additive model was slightly better supported than the interaction model by AIC (Δ_{AIC} = 2.281), the difference between models was not statistically significant (likelihood ratio test, $p < 0.42$). The interaction framework was nevertheless retained in the final model to allow β_1 to vary among regions, providing a more flexible representation of potential regional differences in the rate of maturation over length.

Substantial spatial variation in L_{50} was observed across regions, with estimates ranging from 73.29 cm in the AC region to 101.96 cm in the SO region, a difference of nearly 23 cm and an intermediate value of 94.50 cm in the EQ region. Model coefficients and corresponding L_{50}-L_{95} estimates per region used to construct the maturity ogives are provided in **Table 2**. The smallest difference was observed between the EQ and SO regions (7.46 cm), indicating that the contrast in size-at-maturity is less pronounced between these two regions compared to comparisons involving AC region (vs EQ region = 21.21 and vs SO region =28.67 cm).

Despite covering a smaller spatial extent, the AC region exhibited greater uncertainty in the estimated maturity ogive, likely reflecting differences in sampling structure. In particular, the lower length classes were not exclusively composed of immature females, with an absence of bins dominated solely by immature fish. Instead, these size classes already included a substantial proportion of mature females, resulting in limited resolution of the early part of the maturation curve. Alternatively, the greater uncertainty could reflect a higher heterogeneity in the sampled fish, potentially arising from females originating from different populations or experiencing different environmental conditions influencing length-at-maturity.

Length at maturity could not be modelled for three regions. Sample sizes in the NO region (n = 18) and in the EPO (n = 35) were too small to support reliable estimation of a maturity ogive. In the W region, only six immature females were available, resulting in insufficient representation of maturity states to support reliable model fitting.

Table 2 Logistic regression parameters and estimated length at maturity for yellowfin tuna by region. Estimates include sample size (n), model coefficients (β_0 and β_1), and derived length at maturity estimates (L_{50} and L_{95} expressed in cm) and age at maturity estimates (A_{50} and A_{95} expressed in year), corresponding to the fork length at which 50% and 95% of females are predicted to be mature, respectively. Values in parentheses represent 95% confidence intervals. Regions with insufficient sample size or inadequate representation of immature or mature females were not fitted.

Region	Length at maturity					Age at maturity from otoliths ageing					Age at maturity derived from length				
	n	β_1	β_0	L_{50}	L_{95}	n	β_1	β_0	A_{50}	A_{95}	n	β_1	β_0	A_{50}	A_{95}
EPO	35	Not included in present study				Not determined					Not determined				
Northern (NO)	18	Require more females													
Western (W)	222	Require more immature females													
Archipelagic (AC)	172	0.10 (0.07 to 0.14)	-7.32 (-10.82 to -4.76)	73.29 (67.17 to 77.56)	102.76 (95.26 to 111.58)										
Equatorial (EQ)	707	0.13 (0.05 to 0.21)	-12.13 (-19.40 to -5.49)	94.50 (92.50 to 96.19)	117.45 (113.99 to 121.12)										
Southern (SO)	515	0.12 (0.04 to 0.20)	-12.21 (-19.74 to -5.54)	101.96 (99.12 to 104.56)	126.55 (122.53 to 130.43)										
WCPO	1634	0.09 (0.08 to 0.10)	-8.05 (-09.10 to -7.18)	90.76 (88.94 to 92.62)	123.95 (121.13 to 126.60)	331	1.28 (0.92 to 1.64)	-2.34 (-3.21 to -1.47)	1.83 (1.75 to 1.91)	4.13 (3.59 to 4.77)	1634	1.90 (1.69 to 2.18)	-3.48 (-3.21 to -3.56)	1.83 (1.53 to 2.05)	3.38 (3.21 to 3.56)

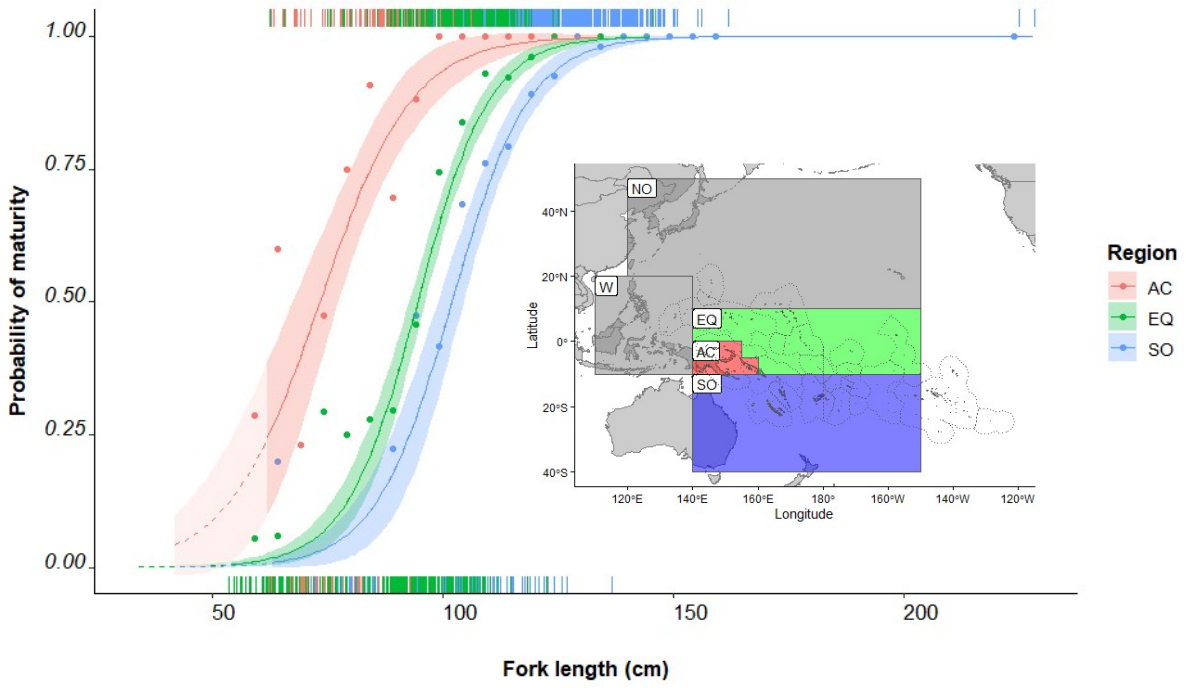

Figure 8 Length at maturity ogive for yellowfin tuna per WCPO stock assessment region. Solid lines show the fitted length a maturity relationships, and shaded ribbons represent the associated uncertainty around the fitted curves. Dashed line and lighter shaded areas indicate model predictions beyond the range of the observed data, representing extrapolation of the fitted relationships. Points represent the proportion of mature females in 5 cm length classes. Rug ticks represent individual females, with those at the top indicating mature females and those at the bottom indicating immature females. Colours correspond to sampling regions, with red representing the Archipelagic region (AC), green the Equatorial region (EQ), and blue the Southern region (SO). An inset map is included to show the geographic stratum of the study area, with regions highlighted using the same colour scheme as in the main figure. This colour scheme is used consistently throughout subsequent figures to facilitate comparisons among regions.

Seasonal dynamics of reproductive activity in the WCPO

There was no evidence that gonad preservation condition affected gonad weight. The best-supported model **(Eq.10b)** included a region-by-month interaction (Δ_{AIC} = 22.61; p-value = 6.96 × 10^{-6} ***), and adding preservation condition did not improve model fit Δ_{AIC} = 0.27; p-value = 0.14). Consistent with this result, the preservation condition coefficient was not significant in the GLM (t test, p-value= 0.14). Therefore, frozen and fresh gonad samples were pooled for subsequent GSI analyses. A total of 1,203 mature female yellowfin tuna were included in the seasonality analysis (**Figure 9**). Strong regional differences in spawning seasonality were observed, with some regions exhibiting continuous year-round spawning activity, while others showed marked seasonal peaks in reproductive activity.

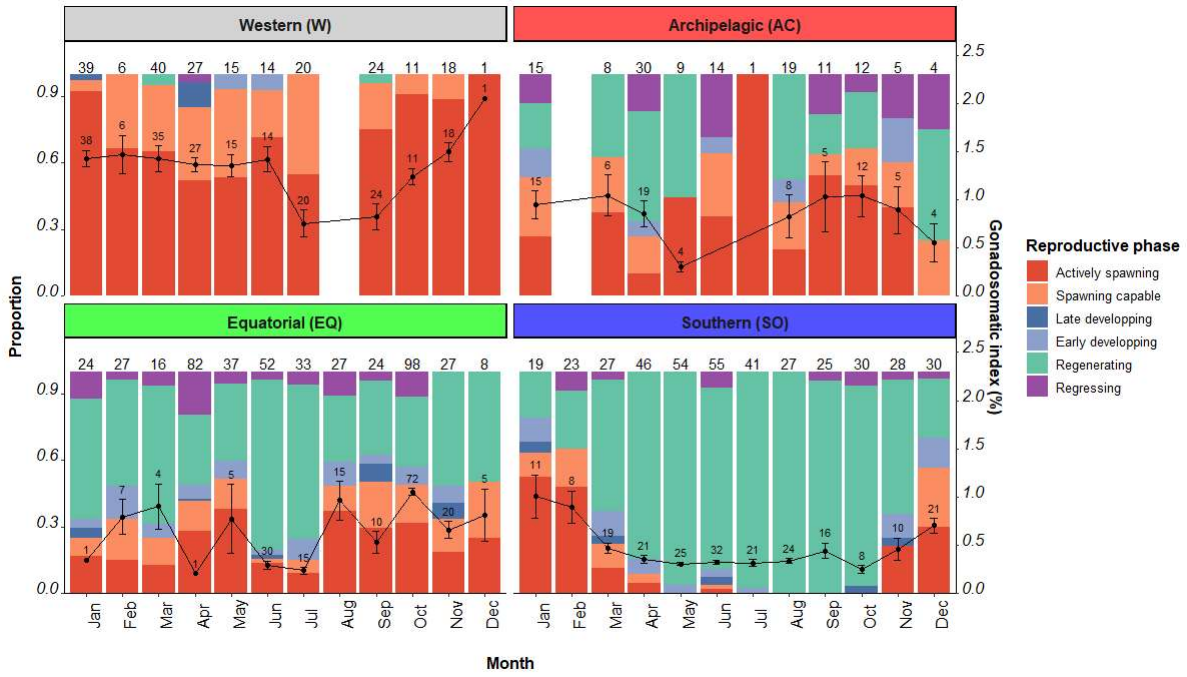

Figure 9 Seasonality of reproductive activity and gonadosomatic index (GSI) of mature female yellowfin tuna across Western and Central Pacific Ocean (WCPO) regions. Stacked bars show the monthly proportion of females assigned to each reproductive phase based on histological classification. The black line and points represent the monthly mean GSI (%), referenced to the secondary y-axis on the right-hand side, and errors bars represent the standard error around the mean. Numbers above stacked bars indicate the monthly sample size used for reproductive phase classification while number above error bars indicate the sample size used for GSI estimation in each month. Strip colours represent sampling regions, with red representing the Archipelagic region (AC), green the Equatorial region (EQ), and blue the Southern region (SO). This colour scheme is consistent across all figures in the manuscript

In region W, spawning activity appeared to occur throughout the year, with consistently high (85-100%) proportions of females classified as either spawning-capable or actively spawning throughout the year. This was corroborated with a generally elevated and stable GSI. In the W region, gonads represented on average 1.24-2.05% of a female's body mass from October to June. There was a statistically significant decrease in GSI to 0.75% and 0.82% in July and September ($p = 3.57\ e^{-05}$*** and $p = 8.69e^{-05}$ ***, respectively) relative to March as the reference month. No data were available for August. However, these patterns should be interpreted with caution for months with less than 10 samples (February, August, December), as estimated proportions of reproductive phases and GSI are highly sensitive to individual observations, a single female can shift a proportion by 10% or more.

In contrast, there was strong evidence of seasonal spawning in the SO region (**Figure 9**), with peak spawning occurring between December and February. During this period, a relatively high proportion of females were classified in spawning capable and actively spawning stages (56.67–65.22%), accompanied by significantly elevated GSI values ranging from 0.70% to 1.00% (p-values = $1.87e^{-05}$ ***, $3.99e^{-10}$ *** and $1.70e^{-06}$ *** respectively) relative to June as the reference month. In contrast, an apparent non-spawning season was observed from May to October, during which no spawning capable and actively spawning females were recorded. The exception was two actively spawning fish (out of a sample of 55) collected in Fijian waters in June. November and March, and to a lesser extent April, appear to represent transitional periods, with 8.70% to 22.22% of females classified in spawning capable and actively spawning stages.

The EQ region exhibited a more mixed pattern of spawning activity throughout the year, with weaker seasonal structuring, although reduced reproductive activity was still evident in June and

July. Relatively high proportions of females in actively spawning and spawning capable stages were observed from August to May, ranging from 25.00% to 51.35%, while lower proportions were recorded in June and July (approximately 15.38% and 15.15%, respectively). Limited GSI data for December to May restrict detailed seasonal comparisons; however, the reduced reproductive activity observed in June and July is supported by significantly lower GSI values (0.23% and 0.29%, respectively) compared to August, October, and November (p-values = $6.59e^{-07}$***, $2.60e^{-13}$*** and $4.72e^{-04}$***). However, GSI in June and July did not differ significantly from September, although interpretation is limited by the low sample size in that month.

In the Archipelagic region, actively spawning and spawning capable females were observed year-round, and regenerating females were also present in almost every month (10 out of 12). However, these patterns should be interpreted with caution given the relatively low sample size, with nearly half of the months (5 out of 12) having fewer than 10 females sampled, and given the high variability observed in GSI.

This lack of a pronounced pattern of spawning activity throughout the year in the EQ region may reflect the broad spatial extent of the sampling, with females collected across a wide longitudinal and latitudinal range, potentially integrating fish from different subpopulations and environmental conditions within the EQ region. Indeed, the three regions differed markedly in the spatial extent represented by sampling. The AC region represented the smallest spatial extent (~2 million km^2; approximately 13.98% and 9.20% of the EQ and SO regions, respectively), while sampling in the SO region was largely concentrated around New Caledonia and Fiji within a relatively narrow latitudinal band, although spanning approximately 20° of longitude. In contrast, the EQ region exhibited the broadest spatial distribution, with samples dispersed across the Federated States of Micronesia, Marshall Islands, Kiribati (Gilbert Islands), and Cook Islands. This resulted in multiple longitudinal clusters separated by approximately 10-40°, while latitude ranged between ~5° and 10°.

Discussion

Implications of preservation methods for reproductive sampling

Formalin is widely regarded as the gold standard for histological preservation because it maintains tissue integrity and cellular structure. However, its hazardous nature makes its use challenging in commercial fishing operations, where handling formalin at sea increases the risk of spills and contamination in environments associated with fish for human consumption (Musiał et al. 2016). While frozen samples overestimated L_{50} and L_{95} by 3.0% and 7.39%, respectively compared to formalin preserved samples, this shift appears small and its potential implications for stock assessment outputs remain uncertain. Further investigation could include sensitivity analyses to evaluate the extent to which preservation of gonads may influence key stock assessment parameters and management advice. Nevertheless, considering the substantial operational and safety benefits, frozen preservation should be adopted for future reproductive sampling programmes, particularly where the use of formalin is impractical or poses logistical constraints. Moreover, reliance on formalin would likely restrict sampling to highly controlled environments, limiting achievable sample sizes and reducing the spatial and seasonal representativeness of sampling.

Revisiting WCPO length-at-maturity estimates: the role of reproductive classification criteria

The WCPO L_{50} estimated in this study (90.76 cm) is substantially lower than that reported (Itano 2000) for the same basin-wide area (WCPO;104.57 cm), which remains the most recent WCPO-wide assessment conducted nearly 30 years ago. However, these estimates are not directly

comparable due to differences in the classification of regenerating females. Itano (2000) assigned females to the regenerating stage based solely on the absence of developed oocytes (primary growth or cortical alveolar stage) combined with evidence of beta-atresia. The present study applied a broader set of criteria to estimate reproductive status, following (Farley et al. 2014). Gonad wall thickness, residual hydrated oocytes, and the presence of muscle bundles and late stages of atresia (referred to as brown bodies) were included as indicators of previous spawning activity. Our approach extends the detectable window of post-spawning evidence beyond atresia alone, improving the identification of regenerating mature females that might otherwise be classified as immature. The effect of classification criteria on the length at maturity ogive was not uniform across the WCPO, varying with spawning seasonality: in highly seasonal regions, mature fish spend an extended period in a regenerating phase between spawning events, during which atresia of vitellogenic oocytes, the sole criterion used by Itano (2000), is no longer detectable several months after spawning (Miranda et al. 1999). A classification scheme relying solely on this transient marker is likely to systematically misclassify a proportion of regenerating females as immature in regions where spawning seasonality is strongest. In contrast, in regions where spawning is more continuous, females may spend less time in a non-spawning (regenerating) condition reducing the likelihood of misclassification. This pattern is reflected in the region-specific L_{50} estimates under the current study's classification criteria versus Itano's (2000) classification criteria applied to the same dataset, with Itano's criteria resulting in larger L_{50} estimates by 14.21, 13.19, and 46.99 cm for regions AC, EQ, and SO respectively, with the disproportionately large shift in SO consistent with its more pronounced seasonality in these fish at high latitudes. As a result, any WCPO-wide maturity estimate based solely on atresia in non-spawning females will be biased toward larger length-at-maturity estimates, with the degree of inflation determined by the relative contribution of highly seasonal regions to the overall dataset.

Despite methodological and spatial differences between Itano's (2000) L_{50} estimates and those from the present study, a partial comparison was possible. Pooling the Archipelagic (AC) and Equatorial (EQ) regions from the present study using Itano's (2000) maturity criteria yielded an L_{50} of 105.29 cm based on frozen-preserved samples, which is comparable to the 107.90 cm reported by Itano (2000) for the equivalent pooled areas (B, C, D) based on formalin-preserved samples, representing a difference of approximately 2.5%.

Spatial and temporal variability in yellowfin tuna maturity and implications for stock assessment

Nonetheless, the regional differences in length at maturity observed in this study suggest that the interpretation of a single WCPO-wide maturity ogive may not adequately capture the variability in length at maturity across the WCPO. Because sampling effort was not evenly distributed among regions, the resulting WCPO-wide ogive is more strongly influenced by regions contributing the greatest number of samples and may therefore not best represent the length at maturity of yellowfin tuna across the WCPO weighted by catch. Likewise, the dataset combines observations collected over more than 20 years (2005-2025) without accounting for temporal stratification, despite the possibility that length at maturity may have changed through time. In comparison, the effect of preservation method on maturity classification was relatively small, suggesting that future efforts should focus primarily on better accounting for spatial, and where possible temporal, variation in length at maturity. If region-specific maturity ogives cannot be integrated into the current stock assessment framework, alternative approaches, such as the multivariate response Bayesian regression model described by Cousido-Rocha et al. (2024), should be explored to account for spatial and temporal variation while retaining a single WCPO-wide maturity ogive for use in stock assessment and monitoring.

Similar work has been conducted on yellowfin tuna in the Eastern Pacific Ocean (EPO) using a comparable classification system, although those studies relied solely on gonad wall thickness rather than the complete set of spawning indicators applied in the present study (Schaefer 1998; Schaefer and Fuller 2022). The spatial coverage of the EPO study differs markedly from that of the present study. The EPO dataset included substantial sampling north of 10°N, whereas the present study had only limited representation in that region (n = 34; 2.04% of total samples). In contrast, no samples were collected between 5°S and 5°N in the EPO study, while considerable equatorial coverage was achieved in the present study, with 35.23% of all samples (n = 588) falling within those latitudes. These differences should be considered when interpreting cross-basin comparisons of maturity parameters.

The EPO-wide L_{50} was reported at 81.56 cm, which is 10.68% lower than the WCPO-wide estimate presented here. Despite this basin-level difference, the spatial patterns in L_{50} within the EPO mirror those observed in the WCPO. In the WCPO, L_{50} values differed across spatial strata, with the highest estimates observed in the southern region (SO: 101.96 cm), followed by lower values in the equatorial zone (EQ: 94.50 cm), and the lowest in the Archipelagic area (AC:73.29 cm). A similar spatial pattern has been reported in the EPO, where L_{50} was highest in the southern region (5°S-20°S; ~90.6 cm), and increased from west to east, rising from 70.9 cm in the westernmost area (85°W-100°W) to 78.6 cm further east (105°W–135°W). A decrease in length at maturity over a 30-year period was documented in the EPO, with fishing pressure identified as a primary driver (Schaefer 1998; Schaefer and Fuller 2022).

Improving age at maturity estimates

Age at maturity can be estimated using two approaches: fitting the maturity ogive to age data derived from age-length conversion or directly reading otolith sections from females with paired gonad maturity assessments. The first approach carries uncertainty associated with both age–length modelling and inter-individual variability in growth, which propagates into age estimates. The second approach avoids age-length conversion but carries uncertainty from otolith ageing error and is typically more labour-intensive, resulting in smaller sample sizes. These smaller sample sizes further limit the ability to explore spatial and temporal patterns in age at maturity. In this study, age at maturity was not estimated at the regional level, as the available paired age–histology data (n = 331) were insufficient to be divided among regions while preserving sufficient sample sizes for region-specific estimates. This is a notable gap given that, as shown above, maturity parameters vary across regions. Growth rates are also expected to vary spatially due to environmental differences such as temperature and food availability and therefore cannot be reliably inferred from WCPO-wide length-at-age relationships alone. Moreover, growth is closely linked to maturity through energy allocation trade-offs between somatic growth and reproductive investment under differing ecological conditions (Williams et al. 2012; Chapman et al. 2011). As a result, spatial heterogeneity in both environmental conditions and reproductive dynamics limits robust inference of age at maturity from a single WCPO-wide length-based relationship. Regional estimates of age at maturity would therefore require either a substantially larger set of aged otolith sections with paired gonad assessments, or the development of region-specific growth curves to support robust age-length conversion and, in turn, region specific estimates of age at maturity. Similarly, the role of individual movement histories should not be overlooked when considering apparent spatial variation in length-at-maturity, or any life history parameters, from sampled fish.

Reproductive seasonality and environmental drivers

Estimates of gonadosomatic index rely partly on body weights derived from length–weight relationships because accurately weighed females are limited. This reflects logistical constraints

associated with at-sea sampling and the condition of landed fish, which are often partially processed. While this introduces uncertainty in absolute GSI values, the strong consistency between GSI patterns and histologically derived reproductive phases suggests that this approach provides a reliable proxy for seasonal reproductive dynamics.

The combined analyses of GSI and reproductive phases shows a clear latitudinal pattern in reproductive seasonality. Higher-latitude regions exhibit more pronounced seasonal spawning, with activity concentrated in a defined summer period, whereas equatorial regions show more continuous reproductive activity throughout the year.

The broader spatial extent of sampling in equatorial regions may also contribute to the apparent reduction in seasonality, as it likely integrates individuals exposed to heterogeneous environmental conditions and potentially distinct subpopulations with differing reproductive schedules.

Overall, these findings support the hypothesis that reproductive seasonality in yellowfin tuna is spatially structured and likely influenced by environmental gradients, particularly temperature. Further work incorporating high-resolution oceanographic data and movement ecology would help to better resolve the mechanisms driving this variability.

Challenges in estimating fecundity parameters

Results on spawning frequency and spawning fraction will be presented in the final report of the P120 project. Batch fecundity is discussed below; however, the limited number of suitable samples (n = 11) prevents reliable fecundity estimation and any meaningful assessment of the spatial variation in maturity parameters described above.

The limited number of samples suitable for spawning frequency and batch fecundity estimation likely results from two main factors. First, sampling was conducted during fishing operations, which are predominantly daytime activities, resulting in few fish caught and sampled between 20:00 and 04:00. As yellowfin tuna have been observed to spawn predominantly during the night (Schaefer 1996), and batch fecundity estimates require examination of samples from females in hours preceding spawning, this sampling bias explains both the scarcity of post-ovulatory follicles in early stages and the low number of hydrated oocytes observed in this study (n = 16) (; Schaefer et Fuller 2022; Itano 2000). Hydrated oocyte counts are generally considered to be the most accurate data for batch fecundity estimation, as they represent the final pre-release stage occurring four to six hours before spawning (Schaefer 1998). While batch fecundity can alternatively be estimated from gonads containing migratory nucleus oocytes by applying a size-class hiatus to isolate oocytes destined for imminent hydration, this approach requires careful validation and introduces additional uncertainty.

The second limitation was accurate whole gonad weight as a prerequisite for batch fecundity estimation. At-sea sampling conditions make precise weighing inherently difficult due to vessel movement and limited onboard storage space for biological samples during commercial fishing operations, meant that only partial gonads were retained in many cases, preventing subsequent weighing of the whole gonad.

These limitations highlight the need for a dedicated sampling campaign targeting mature females during peak spawning activity, with sampling conducted at night between 20:00 and 04:00 to ensure adequate representation of hydrated oocytes and reliable batch fecundity estimation.

Towards an integrated understanding of yellowfin tuna reproductive dynamics

Histological assessment of gonad tissue provides a snapshot of the reproductive condition of a fish at the time of sampling, and population-level patterns are inferred by aggregating these observations across fish, seasons, and regions. While this approach has proven valuable in characterising spawning seasonality and maturity parameters, it remains inherently limited to the moment of capture and cannot capture the life history of individual fish. Consideration of these observed dynamics in fully, spatially explicit modelling frameworks is likely the only way to fully explain the apparent spatial variation reported here. While consideration of discrete, large and fixed regions is a pragmatic choice given the current stock assessments for these species, these artificial boundaries may not reflect biological reality, and their use is likely constrained by inherent model structure and assumptions. Otolith proteomics, which may allow the identification of maturity-related biomarkers encapsulated in otolith material and their linkage to the age of the fish, represents a potentially powerful approach to reconstruct individual spawning histories from archived material (Thomas and Swearer 2019). Such an approach could provide a powerful tool to investigate temporal variation in reproductive biology and better characterise the drivers of reproductive dynamics across the WCPO. In parallel, an improved understanding of growth and bioenergetic processes in this species would enable a more mechanistic approach to describing life-history dynamics, and their likely impact on stock status, robustness to management procedures, and future projections under climate change.

Recommendations

SC22 is invited to:

- Note that frozen samples present operational and safety advantages over formalin and support their adoption as the standard preservation method for future sampling programmes.

-Note that incorporating the complete set of spawning indicators in histological classifications reduces misclassification between immature and regenerating females.

-Consider adopting the WCPO-wide maturity ogive presented in this study for future stock assessments, while recognising that future refinements should better account for spatial and temporal variation in length at maturity

-Consider exploring the observed region-specific maturity parameters in management strategy evaluation frameworks for their influence on stock dynamics, given the observed spatial variability in seasonality and length at maturity.

-Endorse further investigation at finer spatiotemporal scales to better characterise regional patterns in maturity, including exploring environmental, fishing pressure, movement and genetic drivers of this observed variability.

- Endorse further investigation of otolith proteomics as a tool to reconstruct individual reproductive histories and investigate temporal variation in reproductive dynamics across the WCPO.

-Note that improved and more targeted sampling is needed across the WCPO, including better regional and size distribution coverage, dedicated night sampling to support fecundity estimation, and otolith ageing from existing paired samples to enable age at maturity analyses.

-Note that the final report of project P120 will be provided to SC23

Acknowledgment

We gratefully acknowledge the contributions of the observers, port samplers, coordinators, fishing operators, technicians, transporters, and scientific staff involved in the Pacific Marine Specimen Bank (PMSB) sampling program, whose efforts made this study possible. Their dedication and continued support for biological sample collection across the WCPO provided the foundation for this work.

We also thank the staff of Laboratoire de la Nouvelle-Calédonie for their professionalism and high-quality histological processing of samples conducted throughout this study. Additional histological processing support was provided by the Veterinary Laboratory Services at the University of Queensland and the John A. Burns School of Medicine (JABSOM) of the University of Hawaii at Manoa.

We thank NOAA Fisheries in Honolulu for providing access to laboratory facilities that supported histological sample processing.

This project is co-funded by the European Union (EU) under the European Maritime, Fisheries and Aquaculture Fund (EMFAF) implemented through the Western and Central Pacific Fisheries Commission (WCPFC) under the project titled: "Updated reproductive biology of tropical tunas for WCPFC stock assessment and enhancement of the western and central Pacific biodegradable FADs" (WCPFC REPRO FAD). Project ID: 101146206

Declaration of generative AI and AI-assisted technologies in the SC paper preparation process.

Large Language model (ChatGPT, OpenAI and Claude,Anthropic) was used to assist with code development and debugging and drafting and editing portions of the manuscript for clarity and language. All study design, analytical decisions, model specification, data analysis, interpretation of results, and scientific conclusions were made by the authors. All AI-assisted text and code were reviewed, verified, and edited by the authors, who take full responsibility for the content of this manuscript.

Bibliography

Brown-Peterson, Nancy J., David M. Wyanski, Fran Saborido-Rey, Beverly J. Macewicz, and Susan K. Lowerre-Barbieri. 2011. “A Standardized Terminology for Describing Reproductive Development in Fishes.” *Marine and Coastal Fisheries* 3 (1): 52–70. https://doi.org/10.1080/19425120.2011.555724.

Chapman, Erik W., Christian Jørgensen, and Molly E. Lutcavage. 2011. “Atlantic Bluefin Tuna (*Thunnus Thynnus*): A State-Dependent Energy Allocation Model for Growth, Maturation, and Reproductive Investment.” *Canadian Journal of Fisheries and Aquatic Sciences* 68 (11): 1934–51. https://doi.org/10.1139/f2011-109.

Christensen, Rune H. B. 2025. *Ordinal—Regression Models for Ordinal Data*. https://CRAN.R-project.org/package=ordinal.

Cousido-Rocha, Marta, Francisco Izquierdo, Joaquín Martínez-Minaya, et al. 2024. “A Novel Statistical Approach to Deal with Spatial Bias in Maturity Ogive Estimation.” *Canadian*

Journal of Fisheries and Aquatic Sciences 81 (4): 497–507. https://doi.org/10.1139/cjfas-2023-0219.

Farley, Jessica H., Simon D. Hoyle, J. Paige Eveson, Ashley J. Williams, Campbell R. Davies, and Simon J. Nicol. 2014. "Maturity Ogives for South Pacific Albacore Tuna (Thunnus Alalunga) That Account for Spatial and Seasonal Variation in the Distributions of Mature and Immature Fish." *PLoS ONE* 9 (1): e83017. https://doi.org/10.1371/journal.pone.0083017.

Hadley Wickham. 2016. "Ggplot2: Elegant Graphics for Data Analysis." Springer-Verlag New York. https://ggplot2.tidyverse.org.

Højsgaard, Søren, Ulrich Halekoh, and Jun Yan. 2006. "The R Package Geepack for Generalized Estimating Equations." *Journal of Statistical Software* 15/2: 1–11.

Itano, David. 2000. *The Reproductive Biology of Yellowfin Tuna (Thunnus Albacares) in Hawaiian Waters and the Western Tropical Pacific Ocean*.

Lenth, Russell V., and Julia Piaskowski. 2026. *Emmeans: Estimated Marginal Means, Aka Least-Squares Means*. https://doi.org/10.32614/CRAN.package.emmeans.

Macdonald, Jed, Peter Williams, Francois Roupsard, et al. 2023. *Project 90 Update: Better Data on Fish Weights and Lengths for Scientific Analyses*. Nos. SC19-ST-IP-04. 19th Regular Session of the Scientific Committee. https://meetings.wcpfc.int/node/19346.

Magnusson, A., J. Day, T. Tears, et al. 2023. *Stock Assessment of Yellowfin Tuna in the Western and Central Pacific Ocean: 2023 (7Aug2023) - Final (Rev.02 on 15Sep2023) | WCPFC Meetings*. WCPFC-SC19-2023/SA-WP-04 (Rev. 2). https://meetings.wcpfc.int/node/19352.

Miranda, A. C. L., N. Bazzoli, E. Rizzo, and Y. Sato. 1999. "Ovarian Follicular Atresia in Two Teleost Species: A Histological and Ultrastructural Study." *Tissue and Cell*, ahead of print. https://doi.org/10.1054/tice.1999.0045.

Musiał, Agata, Ryszard W. Gryglewski, Stanislas Kielczewski, Marios Loukas, and Justyna Wajda. 2016. *Formalin Use in Anatomical and Histological Science in the 19th and 20th Centuries*.

R Core Team. 2025. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Schaefer, Kurt M. 1998. *Reproductive Biology of Yellowfin Tuna (Thunnus Albacares) in the Eastern Pacific Ocean*. Bulletin of the Inter-American Tropical Tuna Commission 21, no. 94:98-112.

Schaefer, Kurt M., and Daniel W. Fuller. 2022. "Spatiotemporal Variability in the Reproductive Biology of Yellowfin Tuna (*Thunnus Albacares*) in the Eastern Pacific Ocean." *Fisheries Research* 248 (April): 106225. https://doi.org/10.1016/j.fishres.2022.106225.

Shi, Xiaofei, Jian Zhang, Xiao Wang, Yixi Wang, Cheng Li, and Jiangao Shi. 2022. "Reproductive Biology of Yellowfin Tuna (Thunnus Albacares) in Tropical Western and Central Pacific Ocean." *Fishes* 7 (4): 4. https://doi.org/10.3390/fishes7040162.

SPC-OFP. 2019. *Project 35b: WCPFC Tuna Tissue Bank (25July) - Rev.1 | WCPFC Meetings*. 15th Regular Session of the Scientific Committee Nos. SC15-RP-P35b-01. https://meetings.wcpfc.int/node/11272.

SPC-OFP. 2026. *PROJECT 42: PACIFIC TUNA TAGGING PROJECT REPORT AND WORK-PLAN FOR 2026-2029*. Nos. SC22-PTTP-RP1. https://meetings.wcpfc.int/node/26451.

Thomas, Oliver R. B., and Stephen E. Swearer. 2019. "Otolith Biochemistry—A Review." *Reviews in Fisheries Science & Aquaculture* 27 (4): 458–89. https://doi.org/10.1080/23308249.2019.1627285.

Wakefield, Corey, Joseph Widdrington, Alexandre Pons, Simon Nicol, and Paul Hamer. 2026. "Project 129 Update: Building an Age–Length Data Stream for Tuna Assessments." *WCPFC-SC22-2026-SA-WP10*.

Wickham, Hadley, Mara Averick, Jennifer Bryan, et al. 2019. "Welcome to the Tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Williams, Ashley J., Jessica H. Farley, Simon D. Hoyle, Campbell R. Davies, and Simon J. Nicol. 2012. "Spatial and Sex-Specific Variation in Growth of Albacore Tuna (Thunnus Alalunga) across the South Pacific Ocean." *PLOS ONE* 7 (6): e39318. https://doi.org/10.1371/journal.pone.0039318.

Xu, Z., Fine, et al. 2025. "On GEE for Mean-Variance-Correlation Models: Variance Estimation and Model Selection." *Statistics in Medicine* 44: 1–2.

Yan, Jun. 2002. "Geepack: Yet Another Package for Generalized Estimating Equations." *R-News* 2/3: 12–14.

Yan, Jun, and Jason P. Fine. 2004. "Estimating Equations for Association Structures." *Statistics in Medicine* 23: 859–80.

Young, Jeffery L., Zosia B. Bornik, Michelle L. Marcotte, et al. 2006. "Integrating Physiology and Life History to Improve Fisheries Management and Conservation." *Fish and Fisheries* 7 (4): 262–83. https://doi.org/10.1111/j.1467-2979.2006.00225.x.