Stock assessment of yellowfin tuna in the western and central Pacific Ocean

John Hampton

Oceanic Fisheries Programme
Secretariat of the Pacific Community
Noumea, New Caledonia
1 Background

1.1 Biology

Yellowfin tuna are distributed throughout the tropical and sub-tropical waters of the Pacific Ocean. However, there is some indication of restricted mixing between the western and eastern Pacific based on analysis of genetic samples (Ward et al. 1992) and tagging data (Figure 1). Adults (larger than about 100 cm) spawn, probably opportunistically, in waters >26°C (Itano 2000), while juvenile yellowfin are first encountered in commercial fisheries (mainly surface fisheries in Philippines and eastern Indonesia) at several months of age.

Yellowfin tuna are relatively fast growing, and have a maximum fork length (FL) of about 180 cm. The growth of juveniles departs from von Bertalanffy type growth with the growth rate slowing between about 40 and 70 cm FL (Lehodey and Leroy 1999). The natural mortality rate is strongly variable with size, with the lowest rate of around 0.6–0.8 yr \(^{-1}\) being for pre-adult yellowfin 50–80 cm FL (Hampton 2000). Tag recapture data indicate that significant numbers of yellowfin reach four years of age. The longest period at liberty for a recaptured yellowfin tagged in the western Pacific at about 1 year of age is currently 6 years.

1.2 Fisheries

Yellowfin tuna, an important component of tuna fisheries throughout the western and central Pacific Ocean (WCPO, east of 150\(^\circ\)W), are harvested with a diverse variety of gear types, from small-scale artisanal fisheries in Pacific Island and southeast Asian waters to large, distant-water longliners and purse seiners that operate widely in equatorial and tropical waters. Purse seiners catch a wide size range of yellowfin tuna, whereas the longline fishery takes mostly adult fish.

Since 1990, the yellowfin tuna catch in the WCPO has varied between 320,000 and 485,000 t (Figure 2). Purse seiners harvest the majority of the yellowfin tuna catch (49% by weight in 1996–2000), with the longline and pole-and-line fisheries comprising 15% and 3% of the total catch, respectively. Yellowfin tuna usually represent approximately 20–25% of the overall purse-seine catch and may contribute higher percentages of the catch in individual sets. Yellowfin tuna are often directly targeted by purse seiners, especially as unassociated schools.

Longline catches in recent years (56,000–73,000 t) are well below catches in the late 1970s to early 1980s (which peaked at 117,000 t), presumably related to changes in targeting practices by some of the larger fleets. Catches in the ‘Other’ category in Figure 2 are largely composed of yellowfin tuna from the Philippines and eastern Indonesia. These catches come from a variety of gear types (e.g. ringnet, gillnet, handline and seine net) and have increased steadily over the past decade. Based on catch data provided by those countries, recent catches represent approximately 35% of total WCPO yellowfin tuna catches.

Figure 3 shows the spatial distribution of yellowfin tuna catch in the WCPO for the past ten years. The majority of the catch is taken in equatorial areas, with declines in both purse-seine and longline catch towards the east. Also, the east-west distribution of catch is strongly influenced by ENSO events, with larger catches taken east of 160\(^\circ\)E during El Niño episodes.

2 Data compilation

The data used in the yellowfin tuna assessment consist of catch, effort, length-frequency and weight-frequency data for the fisheries defined in the analysis, and tag release-recapture data. The details of these data and their stratification are described below.
2.1 Spatial stratification

The geographic area considered in the assessment is the WCPO, defined by the coordinates 40°N–35°S, 120°E–150°W. Within this overall area, a five-region spatial stratification was adopted for the assessment (Figure 3). This change from the previous seven-region configuration was adopted to somewhat simplify the spatial configuration and fisheries definitions.

2.2 Temporal stratification

The time period covered by the assessment is 1962–2001. Within this period, data were compiled into quarters (Jan–Mar, Apr–Jun, Jul–Sep, Oct–Dec).

2.3 Definition of fisheries

MULTIFAN-CL requires the definition of “fisheries” that consist of relatively homogeneous fishing units. Ideally, the fisheries so defined will have selectivity and catchability characteristics that do not vary greatly over time (although in the case of catchability, some allowance can be made for time-series variation). For most pelagic fisheries assessments, fisheries defined according to gear type, fishing method and region will usually suffice. Fifteen fisheries have been defined for this analysis (Table 1).

2.4 Catch and effort data

Catch and effort data were compiled according to the fisheries defined above. Catches by the longline fisheries were expressed in numbers of fish, and catches for all other fisheries expressed in weight. Effort data for the Philippines and Indonesian fisheries were unavailable and defined as missing. Effort data units for purse seine fisheries are defined as days fishing and/or searching, allocated to set types based on the proportion of total sets attributed to a specified set type (log, FAD or school sets) in logbook data. For the longline fisheries, we used estimates of effective effort derived in a separate study (Bigelow et al. 1999, updated to include 2000 data). Essentially, effective effort is an estimate of the numbers of longline hooks fishing in the mixed layer above the thermocline, which is believed to define yellowfin tuna habitat. The estimates take into account the time and spatial variability in the depth of the mixed layer (using oceanographic databases) and variation in the fishing depth of longliners as indicated by distributions of the numbers of hooks between floats. The effective effort estimates were derived at 5°-month resolution for the Japanese distant-water longline fleet and raised to represent the total longline catch by 5°-month before aggregating into the five-area-quarterly stratification used in the model. Longline effort data were not available for 2001 – we assumed the same levels of effective effort as the corresponding quarters of 2000 and declared the catches for 2001 to be missing.

Within the model, effort for each fishery was normalized to an average of 1.0 to assist numerical stability. In the case of the longline fisheries, the normalization occurred over the five “all nationalities” fisheries rather than individually. Also, effort for these fisheries was divided by the relative size of the respective region. The application of these procedures allowed longline CPUE to index exploitable abundance in each region (rather than density), which in turn allowed simplifying assumptions to be made regarding the spatial and temporal constancy of catchability for the longline fisheries.

2.5 Length-frequency data

Available length-frequency data for each of the defined fisheries were compiled into 100 2-cm size classes (10–12 cm to 208–210 cm). Each length-frequency observation consisted of the actual number of yellowfin tuna measured. The data were collected from a variety of sampling programmes, which can be summarized as follows:
Philippines: Size composition data for the Philippines domestic fisheries derived from a sampling programme conducted in the Philippines in 1993–94 have recently been augmented with data from the 1980s and for 1995.

Indonesia: Limited size data were obtained for the Indonesian domestic fisheries from the former IPTP database. Under the assumption that most of the catch is by pole-and-line gear, catches by the SPC tagging vessels operating in Indonesia in 1980 and 1991–93 have also been used to represent the size composition of domestic fishery catches.

Purse seine: Length-frequency samples from purse seiners have been collected from a variety of port sampling programmes since the mid-1980s. Most of the early data is sourced from the U.S. National Marine Fisheries Service (NMFS) port sampling programme for U.S. purse seiners in Pago Pago, American Samoa and an observer programme conducted for the same fleet. Since the early 1990s, port sampling and observer programmes on other purse seine fleets have provided additional data. Only data that could be classified by set type were included in the final data set.

Longline: The majority of the historical data were collected by port sampling programmes for Japanese longliners unloading in Japan and from sampling aboard Japanese research and training vessels. It is assumed that these data are representative of the sizes of longline-caught yellowfin generally in the WCPO. In recent years, data have also been collected by OFP and national port sampling and observer programmes in the WCPO.

2.6 Weight-frequency data

Individual weight data are available for several longline fisheries. In some cases (e.g. the Australian longline fishery), the weight data represent a large proportion of the total catch. In addition to the Australian fishery, weight data are also available from vessels unloading in Guam and from various unloading ports around the region where tuna are exported. Data were compiled by 1 kg weight intervals over a range of 1–100 kg. As the weights were generally gilled-and-gutted weights, the frequency intervals were adjusted by a gilled-and-gutted to whole weight conversion factor for yellowfin (1.1561).

2.7 Tagging data

A large amount of tagging data was available for incorporation into the MULTIFAN-CL analysis. The data used consisted of yellowfin tuna tag releases and returns from the OFP’s Regional Tuna Tagging Project conducted during 1989–1992. Tags were released using standard tuna tagging equipment and techniques by trained scientists and technicians. The tag release effort was spread throughout the tropical western Pacific, between approximately 120°E and 170°W (see Kaltongga 1998 for further details).

For incorporation into the MULTIFAN-CL analysis, tag releases are stratified by release region (all yellowfin tuna releases occurred in regions 2, 3, 4 and 5), time period of release (quarter) and the same length classes used to stratify the length-frequency data. A total of 39,424 releases were classified into 27 tag release groups in this way. Of the 4,952 tag returns in total, 4,098 could be assigned to the fisheries included in the model. Tag returns that could not be so assigned were included in the non-reported category and appropriate adjustments made to the tag-reporting rate parameters. The returns from each size class of each tag release group were then classified by recapture fishery and recapture time period (quarter). Because tag returns by purse seiners were often not accompanied by information concerning the set type, tag-return data were aggregated across set types for the purse seine fisheries in each region. The population dynamics model was in turn configured to predict equivalent estimated tag recaptures by these grouped fisheries.

3 Structural assumptions of the model

As with any model, various structural assumptions have been made in the yellowfin model. Such assumptions are always a trade-off to some extent between the need, on the one hand, to keep
the parameterization as simple as possible, and on the other, to allow sufficient flexibility so that important characteristics of the fisheries and population are captured in the model. The mathematical specification of structural assumptions is given in Hampton and Fournier (2001). The main structural assumptions used in the yellowfin model are discussed below and are summarized in Table 2.

3.1 Observation models for the data

There are four data components that contribute to the log-likelihood function – the total catch data, the length-frequency data, the weight-frequency data and the tagging data. The observed total catch data are assumed to be unbiased and relatively precise, with the SD of residuals on the log scale being 0.07.

The probability distributions for the length-frequency proportions are assumed to be approximated by robust normal distributions, with the variance determined by the effective sample size and the observed length-frequency proportion. Effective sample size is assumed to be 0.1 times the actual sample size for non-longline fisheries and 0.2 times the actual sample size for longline fisheries, with a maximum effective sample size for all fisheries of 1000. Reduction of the effective sample size recognises that length-frequency samples are not truly random and would have higher variance as a result. The differential treatment of longline and purse seine fisheries occurs because sampling coverage of purse seine catches is generally lower than longline, and the purse seine samples tend to be clumped by wells.

A similar likelihood function was used for the weight-frequency data. The only exception was that the effective sample size for the Australian longline fishery was assumed to be equal to the actual sample size because the sample coverage represented a high proportion of the catch thus increasing the likelihood that the samples were random.

A log-likelihood component for the tag data was computed using a negative binomial distribution in which fishery-specific variance parameters were estimated from the data. The negative binomial is preferred over the more commonly used Poisson distribution because tagging data often exhibit more variability than can be attributed by the Poisson. We have employed a parameterization of the variance parameters such that as they approach infinity, the negative binomial approaches the Poisson. Therefore, if the tag return data show high variability (for example, due to contagion or non-independence of tags), then the negative binomial is able to recognise this. This would then provide a more realistic weighting of the tag return data in the overall log-likelihood and allow the variability to impact the confidence intervals of estimated parameters. A complete derivation and description of the negative binomial likelihood function for tagging data is provided in Hampton and Fournier (2001) (Appendix C).

3.2 Tag reporting

While the model has the capacity to estimate tag-reporting rates, we provided Bayesian priors for fishery-specific reporting rates. Relatively informative priors were provided for reporting rates for the Philippines and Indonesian domestic fisheries and the purse seine fisheries, as independent estimates of reporting rates for these fisheries were available from tag seeding experiments and other information (Hampton 1997). For the longline fisheries L1–L5, we have no auxiliary information with which to estimate reporting rates, so a relatively uninformative prior was used and a common reporting rate estimated. All reporting rates were assumed to be stable over time. The proportions of tag returns rejected from the analysis because of insufficient data were incorporated into the reporting rate priors.

3.3 Tag mixing

We assume that tagged yellowfin gradually mix with the untagged population at the region level and that this mixing process is complete by the second quarter after release.
3.4 Recruitment

“Recruitment” in terms of the MULTIFAN-CL model is the appearance of age-class 1 fish in the population. Earlier attempts to fit the model to yellowfin data using a standard annual recruitment formulation proved unsuccessful. The model could not find a significant growth signal in most of the size data and the von Bertalanffy parameter $K$ typically converged to zero. While there are clear length modes in the size data and these can be followed in some cases for a year or more, the appearance of modes is somewhat erratic and is certainly not with a consistent annual spacing for all fisheries. This situation is to be expected given that yellowfin spawning does not follow a clear seasonal pattern in the tropics but occurs sporadically when food supplies are plentiful (Itano 2000). To solve this problem, we introduced additional structure into the model to allow multiple recruitments per year. The results presented in this report were derived using four recruitments per year, which are assumed to occur at the start of each quarter.

The distribution of recruitment among the five model regions was estimated and allowed to vary over time in an unconstrained fashion. The time-series variation in spatially-aggregated recruitment was somewhat constrained by a lognormal prior. The variance of the prior was set such that recruitments of about three times and one third of the average recruitment would occur about once every 25 years on average.

Spatially-aggregated recruitment was assumed to have a weak relationship with the parental biomass via a Beverton and Holt stock-recruitment relationship (SRR). The SRR was incorporated mainly so that a yield analysis could be undertaken for stock assessment purposes. We therefore opted to apply a relatively weak penalty for deviation from the SRR so that it would have only a slight effect on the recruitment and other model estimates (see Hampton and Fournier 2001, Appendix D).

Typically, fisheries data are very uninformative about SRR parameters and it is generally necessary to constrain the parameterisation in order to have stable model behaviour. We have incorporated a beta-distributed prior on the “steepness” ($S$) of the SRR, with $S$ defined as the ratio of the equilibrium recruitment produced by 20% of the equilibrium unexploited spawning biomass to that produced by the equilibrium unexploited spawning biomass (Francis 1992; Maunder and Watters 2001). The prior was specified by mode = 0.9 and SD = 0.04 ($a = 46, b = 6$). In other words, our prior belief is that the reduction in equilibrium recruitment when the equilibrium spawning biomass is reduced to 20% of its unexploited level would be fairly small (a decline of 10%).

3.5 Age and growth

The standard assumptions made concerning age and growth in the MULTIFAN-CL model are (i) the lengths-at-age are assumed to be normally distributed for each age-class; (ii) the mean lengths-at-age are assumed to follow a von Bertalanffy growth curve; (iii) the standard deviations of length for each age-class are assumed to be a linear function of the mean length-at-age; and (iv) the probability distributions of weights-at-age are a deterministic function of the lengths-at-age and a specified weight-length relationship (see Table 2).

For any specific model, it is necessary to assume the number of significant age-classes in the exploited population, with the last age-class being defined as a “plus group”, i.e. all fish of the designated age and older. This is a common assumption for any age-structured model. For the results presented here, 20 quarterly age-classes have been assumed.

Previous analyses assuming a standard von Bertalanffy growth pattern indicated that there was substantial departure from the model, particularly for sizes up to about 80 cm. Similar observations have been made on yellowfin growth patterns determined from daily otolith increments and tagging data (Lehodey and Leroy 1999). We therefore modelled growth by allowing the mean lengths of the first eight quarterly age-classes to be independent parameters, with the last twelve mean lengths following a von Bertalanffy growth curve.
3.6 Selectivity

Selectivity is fishery-specific and was assumed to be time-invariant. Selectivity coefficients have a range of 0–1, and for the longline fisheries (which catch mainly adult yellowfin) were assumed to increase with age and to remain at the maximum once attained. Selectivities for longline fisheries L1–L5 were constrained to be equal.

The selectivity coefficients are expressed as age-specific parameters, but were smoothed according to the degree of length overlap between adjacent age-classes. This is appropriate where selectivity is thought to be a fundamentally length-based process (Fournier et al. 1998). The coefficients for the last four age-classes, for which the mean lengths are very similar, are constrained to be equal for all fisheries.

3.7 Catchability

Catchability was allowed to vary slowly over time (akin to a random walk) for all non-longline fisheries and the Australian longline fishery using a structural time-series approach, and seasonally for all fisheries apart from the Philippines and Indonesian fisheries (in which the data were based on annual estimates). Random walk steps were taken every two years, and the deviations were constrained by a prior distributions of mean zero and variance specified for the different fisheries according to our prior belief regarding the extent to which catchability may have changed. For the Philippines and Indonesian fisheries (1–3), no effort estimates were available. We made the prior assumption that effort for these fisheries was proportional to catch, but set the variance of the priors to be high (equivalent to a CV of about 0.7 on the log scale), thus allowing catchability changes to compensate for failure of this assumption. For the purse seine fisheries, the catchability deviation priors were assigned a variance equivalent to a CV of 0.10 on the log scale. We assumed that catchability for the non-Australian longline fisheries varied seasonally, but that its average annual value was both constant over time and among the different regions (fisheries). This assumption seemed reasonable given that the estimation of “effective” fishing effort was designed to remove spatial and temporal variability in CPUE due to targeting changes and variation in the depth of optimal yellowfin habitat. In essence, we allowed longline CPUE constructed using effective effort to index the exploitable abundance both among areas and over time.

3.8 Effort variability

Effort deviations, constrained by prior distributions of zero mean, were used to model the random variation in the effort – fishing mortality relationship. For the Philippines and Indonesian fisheries, for which reliable effort data were unavailable, we set the prior variance at a high level (equivalent to a CV of about 0.7 on the log scale), to allow the effort deviations to account for fluctuations in the catch caused by variation in real effort. For the purse seine fisheries and the Australian longline fishery, the variance was set at a moderate level (CV of about 0.2). For the L1–L5 longline fisheries, the variance was set at a low level (CV of about 0.1) to reflect our assumption that longline CPUE (using the effective effort estimates) provides a relatively good index of abundance.

3.9 Movement

Movement was assumed to be time invariant and to occur instantaneously at the beginning of each quarter. For age-independent movement, there would be two non-zero transfer coefficients for each region boundary, i.e. 12 transfer coefficients. We allowed each of these coefficients to be age-dependent in a simple linear fashion, enabling the rate of movement across each region boundary to increase or decrease linearly with age.
3.10 Natural mortality

Natural mortality was assumed to be age-specific, but invariant over time and region. Penalties on the first difference, second difference and deviations from the mean were applied to restrict the age-specific variability to a certain extent.

3.11 Initial population

The population age structure in the initial time period in each region is determined as a function of the average total mortality during the first 20 quarters and the average recruitment in quarters 2–20 in each region. This assumption avoids having to treat the initial age structure, which is generally poorly determined, as independent parameters in the model.

4 Results of the base-case analysis

4.1 Fit of the model to the data

The fit to the total catch data by fishery is very good (Figure 5), which reflects our assumption that observation errors in the total catch estimates are relatively small.

The fit to the length data is displayed in Figure 6 for length samples aggregated over time for each fishery. This provides a convenient means of assessing the overall fit of the model to the length data for each fishery. On the whole, the model appears to have captured the main features of the data, particularly for the larger, more heavily sampled fisheries.

The fit to the weight-frequency data aggregated over time is shown in Figure 7a and an example of the fits to time-series data (for the Australian longline fishery) is shown in Figure 7b. There is some systematic lack of fit to the data for fisheries LL 2 and LL 3. Length-frequency data are also available for these fisheries and these data are not always consistent with the weight-frequency data. This is probably because the weight-frequency data come mainly from Taiwanese longliners in the case of LL 2 and from Hawaii and Pacific Island longliners in the case of LL 3. In contrast, the length-frequency data are sourced mainly from Japanese longliners. It is possible that these fleets have different selectivity characteristics, which might cause the inconsistencies in the data sets. This could be dealt with in future analyses by defining separate fisheries for the Japanese and other fleets mentioned.

The fits of the model to the tagging data compiled by calendar time and by time at liberty are shown in Figure 8a and b. The model results are fairly consistent with the aggregate tag-return data. The fits for individual fishery groups are shown in Figure 8c. Note the discrepancy in observed and predicted returns for the Australian longline fishery. The higher numbers of observed tag returns indicates an inadequacy of the spatial structure of the model and a failure in this instance of the tagged fish mixing assumption. It seems that at least some of the large numbers of tags released in the vicinity of this fishery (off north-eastern Australia at about 16°S) in 1991 and 1992 remained in this area and were hence vulnerable to the Australian longline fishery for an extended period. This problem could be remedied by defining an additional model region that more closely approximates the area of operation of that component of the Australian longline fleet that returned these tags.

4.2 Tag-reporting rates

Estimated tag-reporting rates by fishery are shown in Figure 9. The estimates for the purse seine fisheries are all near the modes of their prior distributions. The estimates for the Philippines and Indonesian fisheries are significantly below its prior mode, indicating that the model has used information contained in the data to estimate this reporting rate. The estimates for the longline fisheries are highly variable, ranging from near zero to the upper limit allowed (0.9).
4.3 Age and growth

Using the four-recruitment-per-year formulation, the model was able to detect a coherent growth signal in the size data. The estimated growth curve is shown in Figure 10. The non-von Bertalanffy growth of juvenile yellowfin is clearly evident, with a pronounced reduction in growth rate in the 40–70 cm size range. This growth pattern is similar to that observed in both otolith and tagging length-increment data (Lehodey and Leroy 1999).

4.4 Selectivity

Estimated selectivity coefficients are generally consistent with expectation (Figure 11). For the purse seine fisheries, the selectivities are very similar for log and FAD sets. However, school sets are focused more towards larger yellowfin. There are also similarities among regions in the selectivity estimates for the purse seine fisheries, suggesting that some aggregation of parameters may be possible in future analyses. Longline selectivity coefficients increase to full recruitment at 2–3 years of age.

4.5 Catchability

Time-series changes in catchability are evident for several fisheries (Figure 12). There is evidence of strongly increasing catchability in all of the purse seine fisheries.

In this analysis, catchability in the longline fisheries has been assumed to be constant over time, with the exception of seasonal variation, which is apparent mainly in the North Pacific longline fishery (LL 1).

The overall consistency of the model with the observed effort data can be examined in plots of effort deviations against time for each fishery (Figure 13). If the model is coherent with the effort data, we would expect an even scatter of effort deviations about zero. Some outliers would also be expected, which prompted the use of robust estimation techniques. On the other hand, if there was an obvious trend in the effort deviations with time, this may indicate that a trend in catchability had occurred and that this had not been sufficiently captured by the model. No unusual variability in the residuals is apparent in Figure 13, suggesting that the model has extracted all the information present in the data regarding catchability variation. In particular, there is no evidence that the assumption of no time-series or regional variation in catchability for the longline fisheries is inappropriate.

Figure 13 also provides an indication of the relative variability of the effort data with respect to the model. The plots suggest that the data for the LL 2 and LL 3 fisheries (the longline fisheries in the tropics) provide the best information on the stock dynamics. Of the purse seine fisheries, the PS/LOG fisheries have less variable effort deviations than the school and FAD-set fisheries.

For the LL 1 – LL 5 longline fisheries, we assumed catchability to be constant among regions, as well as over time (with the exception of seasonal variation). This assumption was considered appropriate because of the use of standardized effort for these fisheries, i.e. the numbers of longline hooks fishing in yellowfin tuna habitat (the upper mixed layer) in a standardized area. Given this assumption and treatment of the longline effort data, we would expect that longline CPUE would provide an index of exploitable abundance (population-at-age multiplied by age-specific selectivity and summed across age-classes) in each region. Figure 14 compares longline exploitable abundance and CPUE for each region. There is generally good, though not perfect agreement between the time-series patterns of both variables in each region. This indicates that the model results reflect longline CPUE as intended by the catchability assumptions.

4.6 Natural mortality

Natural mortality shows considerable variation with size and age-class (Figure 15). For the mid-sizes of ~50-100 cm, the estimates are in the range 0.1-0.2 qtr⁻¹. Estimates for both smaller and larger yellowfin are higher, giving the M-at-age curve a U-shaped appearance. The right hand end of
the curve begins its upward movement at around the size at first maturity, peaking at about 140 cm in size. This has been postulated to be due to higher female mortality associated with spawning.

4.7 Movement

A representation of the dispersal patterns resulting from the estimated movement parameters is shown in Figure 16, which shows the changes in the relative distributions over time of cohorts originating in each region. Movement patterns in the tropics (regions 2 and 3) imply that net movement is strongly west to east.

It is also possible to use the movement coefficients, the average proportions of the total recruitment occurring in each region and the age-specific natural mortality rates to estimate the equilibrium stock composition in each region in the absence of fishing (Figure 17). The model results imply that 31% of the equilibrium biomass in region 1 would be composed of fish recruited in that region. The contributions of local recruitment to equilibrium biomass in the other regions is 74% (region 2), 59% (region 3), 27% (region 4) and 79% (region 5).

4.8 Recruitment

The total recruitment estimates (Figure 18a) are characterised by a strong increase during the late 1970s and 1980s followed by a slight decline in the 1990s. The most recent recruitment estimates are considerably lower than average. The precision of the total recruitment estimates is indicated by the approximate 95% confidence intervals (Figure 18b). For the whole period considered by the model, the average recruitment CV is about 0.15. However, the CV is higher towards the end of the time series and approaches about 0.4 for the last several quarters. This degradation in performance of recruitment estimates is expected for cohorts that have experienced relatively little fishing.

4.9 Biomass

Time series of total and adult relative biomass, by region are shown in Figure 19. Initially, most of the biomass is estimated to have occurred in the tropical regions 2 and 3. Recently, biomass has increased in both the northern (region 1) and southern (region 5) regions. Overall, biomass increased strongly in the late 1970s and 1980s driven by the increased recruitment, but has been in decline since the mid-1990s. The recent decline has been particularly marked in regions 2 and 3. Despite these recent declines, current total biomass is estimated to be at similar levels to the 1960s and 1970s. The biomass of adult yellowfin tuna (assumed to be 25% of age-class 7, 50% of age class 8, 75% of age class 9 and 100% of older age-classes) shows similar trends and spatial distribution to the total biomass.

4.10 Fishing mortality and the impact of fishing

Average fishing mortality rates for juvenile and adult age groups are shown in Figure 20 for the total model area. Fishing mortality for adults has increased steadily over the time series. In contrast, juvenile fishing mortality has increased rapidly particularly since the early 1990s. A major factor in this increase has been the increase in reported catches in Indonesia. Catches reported by the national authorities in Indonesia have been used in this assessment, but further validation is required. Increased purse seine catches using FAD sets in region 3 have also contributed to the increase.

For a complex model such as this, it is difficult to readily interpret fishing mortality rates and other parameters to obtain a clear picture of the estimated impact of fishing on the stock. To facilitate this, we have computed total biomass trajectories for the population in each region using the estimated recruitment, natural mortality and movement parameters, but assuming that the fishing mortality was zero throughout the time series. Comparison of these biomass trajectories with those incorporating the actual levels of observed historical fishing provides a concise, integrated picture of the impacts of the total fishery on the stock. Biomass trajectories for each region and for the WCPO in total are shown in Figure 21. The greatest impacts have occurred in regions 2 and 3, where the “actual” biomass is about
50% of the “unfished” biomass. This result would suggest that there has been a significant depletion of the sub-populations in these regions, primarily by the domestic fisheries of the Philippines and Indonesia and the combined purse seine fishery. For the WCPO in total, the current biomass is estimated to be around 37% less than that which would have occurred in the absence of fishing.

4.11 Yield and reference point analysis

The use of reference points provides a framework for quantitatively determining the status of the stock and its exploitation level. Two types of reference points are often now required for fisheries management: the fishing mortality at maximum sustainable yield (F_{MSY}) is used as an indicator of overfishing; and the biomass at MSY (B_{MSY}) is used as an indicator of an overfished state. It is possible for overfishing to be occurring, but for the stock to not yet be in an overfished state. Conversely, it is possible for the stock to be in an overfished state but for the current level of fishing to be within the overfishing reference point. In this case, the stock has presumably been depressed by past overfishing and would recover to a non-overfished state if the current level of fishing was maintained. It is likely that these reference points, or something similar, will be used for stock status determinations in the new WCPO tuna commission. We have therefore developed a reference point analysis within the MULTIFAN-CL model framework as an example of how this might be applied in WCPO tuna fisheries.

The reference point analysis has been carried out as follows:

1. Estimate population model parameters, including the parameters of a Beverton and Holt stock-recruitment relationship (SRR).
2. Estimate a “base” age-specific fishing mortality vector, F_{age}, various multiples of which are assumed to maintained into the future; for the yellowfin tuna assessment, the average F_{age} over 1996–2000 was used.
3. For various multiples of F_{age} compute the equilibrium population-at-age, and equilibrium yield using the estimated SRR, natural mortality and other parameters.
4. Compute the equilibrium total biomass, equilibrium adult biomass and equilibrium fishing mortality (averaged over age classes) at MSY. These equilibrium quantities are the reference points.
5. Compare the actual estimated biomass and fishing mortality levels at time t with these reference points. This is done by computing the ratios B_{t}^{total} / B_{MSY}^{total}, B_{t}^{adult} / B_{MSY}^{adult}, F_{t} / F_{MSY} and their 95% confidence intervals and comparing them with 1.0. Values of F_{t} / F_{MSY} significantly greater than 1.0 would indicate overfishing, while values of B_{t}^{total} / B_{MSY}^{total} and/or B_{t}^{adult} / B_{MSY}^{adult} or less than 1.0 would indicate an overfished state.

Note that these somewhat simplistic notions make assumptions about equilibrium behaviour of the populations. This aspect of reference points and in particular those based on equilibrium models has been roundly criticised (with some justification) in some fisheries circles. One criticism is that long-term changes in recruitment might occur through environmental or ecosystem changes that have little or nothing to do with the fisheries. More generally, it is not unreasonable to view many fish populations as being in a continual state of flux with an equilibrium condition never being reached or maintained for any length of time. In reality, therefore, MSY, F_{MSY} and B_{MSY} are “moving targets” and not static quantities. At best, they should be considered as averages over time, and additional analyses undertaken in cases where it is suspected that important non-fishery-induced changes in productivity may have occurred.

The estimated SRR used in the yield and reference point analyses for yellowfin tuna is shown in Figure 22. The scatter of recruitment-biomass points is fairly typical of most fisheries data sets – there is very little information on how recruitment might respond to very low biomass levels. For this reason, it is necessary to constrain the behaviour of the curve in the region towards the origin by the
prior assumption for “steepness”. To recap, the assumption was that significant (>10%) recruitment decline occurs only at adult biomass of <20% of virgin levels, i.e. that average recruitment is quite robust to adult biomass decline.

The estimated equilibrium yield using a base \( F \)-at-age given by the 1996–2000 average is shown in Figure 23. This analysis indicates that, at the 1996–2000 average \( F \)-at-age (i.e. a fishing mortality multiplier of 1.0), the equilibrium yield is approximately 275,000 t per year. The maximum equilibrium yield (equivalent to MSY) of about 312,000 t is achieved at a \( F \)-multiplier of 1.7. These equilibrium yields are considerably lower than the actual catches that occurred during 1996–2000, which averaged about 425,000 t per year. This is because the yield analysis is based on an equilibrium model in which equilibrium recruitment is predicted on the basis of the SRR shown in Figure 22. However, recruitment during much of the 1990s was at a relatively high level which enabled the recent high catches to occur. To demonstrate this effect, we repeated the yield analysis using the estimated average levels of recruitment for 1962–1980 (low-recruitment period) and 1981–2001 (high-recruitment period) instead of using a single SRR. These results are shown in Figure 24. For the high-recruitment period, the yield at an \( F \)-multiplier of 1.0 is 400,000 t, which is close to the 1996–2000 average catch, and the MSY is approximately 500,000 t at an \( F \)-multiplier of about 2. However, if recruitment were to return to pre-1981 levels, the MSY would then be only about 290,000 t. Note that these estimates of MSY based on average recruitment may be optimistic as they ignore any possibility of recruitment decline at low stock sizes.

Recruitment effects such as noted above can obviously bias status determinations based on a comparison of catch and MSY. A comparison of \( F \) with \( F_{MSY} \) is better from this point of view as the effects of recruitment are removed. The ratios of \( F_i/F_{MSY} \) and \( B_i^{\text{adult}}/B_{MSY}^{\text{adult}} \) are shown in Figure 25. \( F_i/F_{MSY} \) has been beneath the overfishing reference point throughout the time series. Also, while adult biomass has fallen recently, \( B_i^{\text{adult}}/B_{MSY}^{\text{adult}} \) has remained above 1.0, indicating that the population has yet to reach an overfished state under the definition used here.

5 Retrospective analysis

Retrospective analysis is undertaken by repeating the model fitting procedure on data sets in which some number of recent years’ data have been removed. This objective is to see whether or not there are consistent biases in population parameters for the most recent years. It also demonstrates the impact of adding new data to the model. The analysis undertaken here consisted of repeating the base-case analysis on data sets with the most recent 1 year (retro-1), 2 years (retro-2), 5 (retro-5) and 10 (retro-10) years of data removed. The results obtained for total recruitment, biomass, juvenile and adult fishing mortality, fishery impact, yield and reference points are shown in Figure 26. The results for the retro-1 and retro-2 analyses are very similar to those for the base-case analysis. The retro-10 results are also similar, at least up until about three years before the end of the time series. At this point, the recruitment and biomass decline and fishing mortality increases. These patterns were not observed in the longer time series. The retro-5 results show similar patterns to the longer time series, but absolute biomass is lower and juvenile fishing mortality is higher. There is also a tendency for divergence during the last three years of the retro-5 time series. These different recruitment, biomass and fishing mortality results cause substantial variation in the corresponding yield curves, although a lot of this would be due simply to the differences in base \( F \)-at-age (which in each was the average for the last-but-one five years of the analysis). The \( F \)- and \( B \)-ratio reference points show much less variability, which suggests that they may be more robust measures of stock status than the yield curves themselves.

6 Conclusions

The yellowfin tuna model has integrated catch, effort, length-frequency, weight-frequency and tagging data into a coherent analysis that is broadly consistent with other information on the biology and fisheries. The major conclusions of the analysis with respect to stock assessment are:
1. Recruitment shows considerable variation at several different time scales, but the main feature of the estimates is an increase in total recruitment during the late 1970s to early 1990s, followed by some decline.

2. The main feature of both the total and adult biomass estimates is a strong increase in the late 1970s and a decline since the mid-1990s. The decline would appear to be driven by the lower recruitments that have occurred in recent years, although higher fishing mortality has also contributed.

3. Fishing mortality for juvenile yellowfin tuna has increased strongly since about 1992, partly as a result of catchability increases in the purse seine fisheries. But a significant component of the increase is attributable to the Philippines and Indonesian fisheries, which have the weakest catch, effort and size data. This is of continuing concern. There has been recent progress made in the acquisition of a large amount of historical length frequency data from the Philippines and regular sampling operations are now in place there. However, uncertainty with the total catch and size composition data for the Indonesian fishery continues to be a problem.

4. The overall impact of fishing on stock biomass is estimated to be in the vicinity of 35% in recent years. The impact is differentially high in the tropical regions (around 50%) compared to the subtropical regions.

5. Further work in this year’s assessment has been carried out on the application of reference points. The estimates of $F_i/F_{MSY}$ and $B_i/B_{MSY}^{adult}$ suggest that the stock is neither being overfished nor is it in an overfished state. We note, however, that the current trends in both ratios are towards the reference points. Continued monitoring and improvement in fisheries statistics is required.

6. Recommended research and monitoring required to improve the yellowfin tuna assessment include the following:
   - Continued monitoring and improvement in fisheries statistics is required. In particular, better data generally are required for the Philippines and Indonesian fisheries.
   - New conventional tagging experiments, undertaken regularly, would provide additional information on recent levels of fishing mortality, refine estimates of natural mortality and possibly allow some time-series behaviour in movement to be incorporated into the model.
   - In view of the importance placed on longline effort data by this model, additional archival tagging is required to characterise the depth distribution of yellowfin tuna and its environmental correlates across the stock range to enable better estimation of effective longline effort.

7 **Acknowledgements**

We are grateful to Naozumi Miyabe of the Japan National Research Institute of Far Seas Fisheries for his assistance in compiling data for the Japanese fisheries.

8 **References**


<table>
<thead>
<tr>
<th>Fishery</th>
<th>Nationality</th>
<th>Gear</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>LL 1</td>
<td>All</td>
<td>Longline</td>
<td>1</td>
</tr>
<tr>
<td>LL 2</td>
<td>All</td>
<td>Longline</td>
<td>2</td>
</tr>
<tr>
<td>LL 3</td>
<td>All</td>
<td>Longline</td>
<td>3</td>
</tr>
<tr>
<td>LL 4</td>
<td>All excl. Australia</td>
<td>Longline</td>
<td>4</td>
</tr>
<tr>
<td>LL 4A</td>
<td>Australia</td>
<td>Longline</td>
<td>4</td>
</tr>
<tr>
<td>LL 5</td>
<td>All</td>
<td>Longline</td>
<td>5</td>
</tr>
<tr>
<td>PS/LOG 2</td>
<td>All</td>
<td>Purse seine, log sets</td>
<td>2</td>
</tr>
<tr>
<td>PS/FAD 2</td>
<td>All</td>
<td>Purse seine, FAD sets</td>
<td>2</td>
</tr>
<tr>
<td>PS/SCH 2</td>
<td>All</td>
<td>Purse seine, school sets</td>
<td>2</td>
</tr>
<tr>
<td>PS/LOG 3</td>
<td>All</td>
<td>Purse seine, log sets</td>
<td>3</td>
</tr>
<tr>
<td>PS/FAD 3</td>
<td>All</td>
<td>Purse seine, FAD sets</td>
<td>3</td>
</tr>
<tr>
<td>PS/SCH 3</td>
<td>All</td>
<td>Purse seine, school sets</td>
<td>3</td>
</tr>
<tr>
<td>PH/R 3</td>
<td>Philippines</td>
<td>Ringnet</td>
<td>2</td>
</tr>
<tr>
<td>PH/H 3</td>
<td>Philippines</td>
<td>Handline</td>
<td>2</td>
</tr>
<tr>
<td>ID 3</td>
<td>Indonesia</td>
<td>Various</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 2. Main structural assumptions used in the yellowfin tuna analysis.

<table>
<thead>
<tr>
<th>Category</th>
<th>Assumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation model for total catch data</td>
<td>Observation errors small, equivalent to a residual SD on the log scale of 0.07.</td>
</tr>
<tr>
<td>Observation model for length-frequency data</td>
<td>Normal probability distribution of frequencies with variance determined by effective sample size and observed frequency. Effective sample size assumed to be 0.1 times actual sample size for non-longline fisheries and 0.2 times for longline fisheries with a maximum effective sample size of 100.</td>
</tr>
<tr>
<td>Observation model for weight-frequency data</td>
<td>Normal probability distribution of frequencies, variance determined by effective sample size and observed frequency. Effective sample size assumed to be equal to the actual sample size for the Australian longline fishery, 0.2 times the actual sample size for other longline fisheries with a maximum effective sample size of 1000.</td>
</tr>
<tr>
<td>Observation model for tagging data</td>
<td>Tag numbers in a stratum have negative binomial probability distribution, with fishery-specific variance parameter.</td>
</tr>
<tr>
<td>Tag reporting</td>
<td>Informative priors for purse seine fisheries (based on tag seeding), moderately informative priors for Philippines and Indonesian fisheries, relatively uninformative prior for longline fisheries. Longline reporting rates constrained to be equal. All reporting rates constant over time.</td>
</tr>
<tr>
<td>Tag mixing</td>
<td>Tags assumed to be randomly mixed at the model region level from the quarter following the quarter of release.</td>
</tr>
<tr>
<td>Recruitment</td>
<td>Occurs as discrete events at the start of each quarter. Spatially-aggregated recruitment is weakly related to spawning biomass in the prior quarter via a Beverton-Holt SRR (beta prior for steepness with mode at 0.9 and SD of 0.04). The spatial distribution of recruitment in each quarter is allowed to vary in an unconstrained fashion.</td>
</tr>
<tr>
<td>Initial population</td>
<td>A function of the equilibrium age structure in each region, which is assumed to arise from the total mortality and movement rates estimated for the initial 20 quarters of the analysis.</td>
</tr>
<tr>
<td>Age and growth</td>
<td>20 quarterly age-classes, with the last representing a plus group. Juvenile age-classes 1–8 have independent mean lengths; adult age-class mean lengths constrained by von Bertalanffy growth curve. Mean weights ($W_f$) computed internally by estimating the distribution of weight-at-age from the distribution of length-at-age and applying the weight-length relationship $W = aL^b$ ($a=0.00002784$, $b=2.8992$ estimated from available length-weight data).</td>
</tr>
<tr>
<td>Selectivity</td>
<td>Constant over time. Various smoothing penalties applied. Coefficients for the last 4 age-classes are constrained to be equal. Longline selectivities are non-decreasing with increasing age. Non-Australian longline fisheries share selectivity parameters.</td>
</tr>
<tr>
<td>Catchability</td>
<td>Constant over years and among regions for longline fisheries (effort data are scaled to reflect different region sizes). Seasonal variation for all fisheries apart from Philippines and Indonesian fisheries. Non-longline fisheries and the Australian longline fishery have structural time-series variation, with random steps (catchability deviations) taken every 2 years. Catchability deviations constrained by a prior distribution with (on the log scale) mean 0 and SD 0.1 (SD is 0.7 for Philippines and Indonesian fisheries with missing effort data).</td>
</tr>
<tr>
<td>Fishing effort</td>
<td>Variability of effort deviations constrained by a prior distribution with (on the log scale) mean 0 and SD 0.22 (SD is 0.7 for Philippines and Indonesian fisheries with missing effort data).</td>
</tr>
<tr>
<td>Natural mortality</td>
<td>Age-dependent but constant over time and among regions. Smoothing penalties constrain the age-dependency.</td>
</tr>
<tr>
<td>Movement</td>
<td>Age-dependent but constant over time and among regions. Age-dependency for each coefficient (2 per region boundary) is linear.</td>
</tr>
</tbody>
</table>
Figure 1. Long-distance (>1,000 nmi) movements of tagged yellowfin tuna.

Figure 2. WCPO yellowfin tuna catch, by gear.
Figure 3. Distribution of yellowfin tuna catch, 1991–2000. The heavy lines indicate the spatial stratification used in the MULTIFAN-CL model.
Figure 4. Catch per unit effort by fishery. Units are catch number per 100 hooks fishing in yellowfin tuna habitat (fisheries LL1, LL2, LL3, LL4, LL5), catch number per 100 nominal hooks (LL4A) and catch (t) per day fished/searched (all PS fisheries).
Figure 5. Observed (circles) and predicted (lines) total catches by fishery and quarter. Catches are in tonnes for all fisheries except the longline (LL) fisheries, where the catches are in number of fish.
Figure 6. Observed (histograms) and predicted (line) length frequencies for each fishery aggregated over time.
Figure 7a. Observed (histograms) and predicted (line) weight frequencies for each fishery aggregated over time.
Figure 7b. Time-series fit (lines) to weight-frequency data (histograms) for the Australian longline fishery. Vertical bars indicate the estimated mean weights at age.
Figure 8. Observed (circles) and predicted (lines) tag returns (a) totals by recapture period, (b) totals by time at liberty, and (c) by various fishery groupings.
Figure 9. Estimated tag-reporting rates by fishery (histograms). The black squares indicate the modes of the priors for each reporting rate and the small bars indicate a range of ±1 SD.

Figure 10. Estimated mean lengths-at-age (heavy line) and the variability of length-at-age (dotted lines represent ± 2 SD).
Figure 11. Selectivity coefficients, by fishery.
**Figure 12.** Estimated catchability trends for each fishery.
Figure 13. Effort deviations by time period for each fishery. The X-axis scale is 1962–2002. The Y-axis scale is in SD units.
Figure 14. Estimates of exploitable abundance (solid lines) and CPUE (dashed lines) for the longline fisheries in each region. Both variables have been smoothed to remove seasonal variation and are scaled relative to their averages.
Figure 15. Estimated natural mortality rate by age-class plotted against mean length-at-age with 95% confidence intervals.
Figure 16. Relative distributions over time of cohorts recruited in each region.
Figure 17. The estimated origin of recruits (different patterns within bars) for yellowfin sub-populations in each region (different bars). Sub-populations are computed at equilibrium and in the absence of fishing.
Figure 18. Time-series of estimated yellowfin tuna recruitment for the base-case model: (a) quarterly recruitment estimates (circles) and moving 4-quarter average spatially-aggregated recruitment (heavy line); (b) 95% confidence intervals for quarterly recruitment; and (c) quarterly recruitment estimates by region (circles) and 4-quarter moving averages (heavy lines).
Figure 19. Estimated relative total (a) and adult (b) yellowfin tuna biomass by region, and total biomass with 95% confidence intervals (c) for the base-case model.

Figure 20. Estimated annual average fishing mortality rates for the WCPO.
Figure 21. Comparison of the estimated biomass trajectories (lower heavy lines) with biomass trajectories that would have occurred in the absence of fishing (upper thin lines) for the base-case model. (a) For the WCPO (with the percentage difference in the trajectories also plotted in grey), and (b) for each region.
Figure 22. Spawning biomass – recruitment estimates (on a relative scale) and the fitted Beverton and Holt stock-recruitment relationship (SRR) incorporating a prior on steepness of 0.9. The dashed lines are the 95% confidence intervals on the SRR.

Figure 23. Predicted equilibrium yield and 95% confidence intervals as a function of fishing mortality (relative to the average fishing mortality-at-age during 1996-2000).
Figure 24. Predicted equilibrium yield assuming constant equilibrium recruitment, as a function of fishing mortality (relative to the average fishing mortality-at-age during 1996-2000). The two curves represent recruitment at two average levels corresponding to early (1962–1980) and recent (1981–2001) estimated recruitment.
Figure 25. Ratios of (a) $F_t/F_{MSY}$ and (b) $B_{adult, t}/B_{adult, MSY}$ with 95% confidence intervals. The horizontal lines at 1.0 in each case indicate the overfishing (a) and overfished state (b) reference points.
Figure 26. The base-case analysis compared to analyses with the last 1, 2, 5 and 10 years of data removed.