SCIENTIFIC COMMITTEE

FIFTH REGULAR SESSION

10-21 August 2009

Port Vila, Vanuatu

CHARACTERISTICS OF POTENTIAL REFERENCE POINTS FOR USE IN WCPFC TUNA STOCK ASSESSMENTS

WCPFC-SC5-2009/ME-WP-02

S. J. Harley[1], S. D. Hoyle[1], J. Hampton[1], and P. Kleiber[2]

[1] Oceanic Fisheries Programme, Secretariat of the Pacific Community, Noumea, New Caledonia
[2] Pacific Islands Fisheries Science Center, National Marine Fisheries Service, Honolulu, Hawaii, USA.

Characteristics of potential reference points for use in WCPFC tuna stock assessments

Shelton Harley, Simon Hoyle, John Hampton and Pierre Kleiber

1. Introduction

This report describes analyses undertaken to address the following request for material to support the workshop on reference points to be held during the Methods Specialist Working Group at SC-5. SPC-OFP were required to provide advice on the following:

1. Suitability of MSY-based reference points as default limit reference points and how they may be operationalized, as described in SC4 Summary Report paragraph 202;
2. Evaluate stock status for key stocks against a variety of reference points;
3. Consider alternative methods for assessing stock status against reference points; and
4. Evaluate the sensitivity of reference points to alternative model structures and stability over time with respect to the incorporation of new data.

In this paper we:

- discuss the suitability of *MSY*-based reference points as default limit (or conservation) reference points and how they could be integrated into the decision making process, e.g. operationalized;
- introduce a variety of reference points that can be evaluated from stock assessment results and discuss the types of information that they can provide to fishery decision makers;
- consider approaches to addressing the technical issue of how status is evaluated against the reference points, namely how can we incorporate uncertainty and risk; and
- illustrate sensitivity of reference points to particular model assumptions, which is an important consideration for choosing reference points that are robust, and explaining to managers why results sometimes change from year to year.

We have not undertaken any specific modelling to address point[3] (1), instead we have consulted one of the key references on the subject, namely Annex II of the UNFSA (Annon. 1995) and used this as the basis for answering the two questions (1) Are MSY-based reference points suitable as default limit reference points, and (2) how can they be operationalized. It it is not entirely clear, how Article 5(b) of the WCPF Convention relates to this question, as it appears to relate to target reference points rather than limit reference points. We recognize from the outset that this is more of a technical policy / legal question than one for scientists to consider alone. Nevertheless, we have attempted to do what has been asked, but note the need for further legal advice on this important question.

[3] Nevertheless, some of the analysis and discussions provided in the responses to 2-4 will be relevant.

In order to answer the other questions we will use a range of stock assessment modelling results. To address (2) we will introduce and describe various reference points, including the types of information that they can provide for managers. The assessments for SP-ALB, BET, and YFT conducted this year will include evaluation against many of these reference points, but we have not included these results here.

Point (3) addresses how we account for uncertainty when determining whether or not we have exceeded a reference point. We will discuss some of the approaches that have been used for this and introduce some alternative approaches for consideration. Some of these alternative options are in their early development. This work will be applied to the model runs included in the YFT structural sensitivity analysis (Harley et al. 2009a) that is based on the 2007 YFT assessment model (Langley et al. 2007).

Point (4) addresses the issue of why the estimates of particular reference points and stock status can change between assessments and among sensitivity analyses. We will show how these estimates are sensitive to changes in various model assumptions and inputs, in particular, the assumptions made about the relationship between the numbers of spawning fish and the number of juveniles they produce, the addition of new data, and changes in the mix of fishing gears used to exploit the stock. These analyses will be based on some model runs undertaken with the 2007 YFT assessment.

The rest of the paper is structured as follows: we address each question in turn, describing the approaches used and key results; then there is a single section at the end which includes the key conclusions and recommendations for future work that arose from each.

2. MSY-quantities as default limit reference points

In addressing this first question we refer to Annex II of the UN Fish Stocks Agreement (Annon. 1995), namely the "GUIDELINES FOR APPLICATION OF PRECAUTIONARY REFERENCE POINTS IN CONSERVATION AND MANAGEMENT OF STRADDLING FISH STOCKS AND HIGHLY MIGRATORY FISH STOCKS". What we have interpreted as the relevant parts of this text to our task are listed below. They have been numbered by us to allow us to refer back to these (emphasis added):

a) *Two types of precautionary reference points should be used: conservation, or limit, reference points and management, or target, reference points. Limit reference points set boundaries which are intended to constrain harvesting within safe biological limits within which the stocks can produce maximum sustainable yield*

b) *Precautionary reference points should be stock-specific to account, inter alia, for the reproductive capacity, the resilience of each stock and the characteristics of fisheries exploiting the stock, as well as other sources of mortality and major sources of uncertainty*

c) *Fishery management strategies shall ensure that the risk of exceeding limit reference points is very low. If a stock falls below a limit reference point or is at risk of falling below such a reference point, conservation and management action should be initiated to facilitate stock recovery*

d) *When information for determining reference points for a fishery is poor or absent, provisional reference points shall be set. Provisional reference points may be established by analogy to similar and better-known stocks*

e) *The fishing mortality rate which generates maximum sustainable yield should be regarded as a minimum standard for limit reference points*

It is also useful to define what is meant by default, and for the purposes of question (1) we have defined default as "*an option that is selected automatically unless an alternative is specified*".

Point (e) obviously seems very important, but even among the authors of this paper there were different interpretations, in particular with respect to what is meant by the phrase "a minimum standard". Two potential interpretations were:

- If F_{MSY} = 0.2, we can either use 0.2 or a **lower value** of fishing mortality as a *F*-based limit reference because F_{MSY} is a minimum standard and we intake this interpretation literally; or
- The phrase "minimum standard" would be interpreted as "default" which can be defined as "an option that is selected automatically unless an alternative is specified". Therefore, you use F_{MSY} as a limit reference point unless you have specified an alternative which is still consistent with the other guidelines noted in Annex II.

The first definition is rather troublesome as it leaves very little flexibility and might in fact provide a constraint for achieving *MSY*, whilst the second provides some flexibility, and there are other provisions within Annex II providing guidance on what alternative limit reference points need to achieve (a and b). It is possible that Article 5(b) of the WCPF-Convention might provide some of the basis for choosing an alternative, but this is not entirely clear.

Are MSY-based reference points suitable as default limit reference points?

Considering the criteria above we ask two questions in relation to MSY-based reference points:

- Are they stock specific, i.e. do they incorporate the information described in (b)?
- Do you want to have a very low risk of exceeding them for biological reasons (i.e. to keep within safe biological limits) as recommended in (a)?

Within the MULTIFAN-CL framework (and most other age-structured models), estimation of MSY-based quantities incorporates many sources of information including the biological parameters for the species (e.g. growth, natural mortality, and maturity), and the recent fishery conditions through the estimate of age-specific fishing mortality. Therefore, one might conclude that MSY-based reference points in the context of the MULTIFAN-CL assessments, meet the criteria of stock-specific reference points that consider the important information. The one caveat on this statement is uncertainty in the steepness of the spawner-recruitment relationship, which is a key determinant of resilience to overfishing. This topic comes up throughout this paper, and approaches for incorporating this uncertainty into estimates of reference points and stock status will be an important consideration for the use of MSY-based reference points. Notwithstanding this we would conclude that F_{MSY}, and SB_{MSY} or SB_{MSY} as estimated from MULTIFAN-CL meet the criteria for stock-specific limit reference points.

Having management strategies that have a very low risk of exceeding these reference points (c) should stop the stock from being reduced below safe biological limits (a). We note that this would be a very precautionary approach as requiring fishing strategies to have a very low probability of biomass falling below B_{MSY} or fishing mortality exceeding F_{MSY}, would result in average biomass levels well above B_{MSY} and yields would be lower than MSY [4].

Note: The above is based on assuming that SB_{MSY} or F_{MSY} are safe biological limits. However, whilst this is often assumed, is may not always be the case for all stocks. Adoption of reference points that satisfies criteria (a), i.e. constrains harvesting within safe biological limits, is ultimately premised on an understanding of what a safe biological limit is for a given stock.

Based on this understanding, the guidance provided through Annex II of the UNFSA (Annon. 1995), and taking our second interpretation of (e), we would conclude that F_{MSY} and SB_{MSY} are appropriate as default limit reference points.

Notwithstanding this conclusion, there are alternatives that could be used as a limit reference points without compromising (a) and (e) above. In the case of the four key tuna stocks assessed by the WCPFC, we would suggest that it is not necessary to use the default limit reference points and alternative limit reference points could be developed. It is possible that an alternative could still be based on the concept of MSY, for example $x\%SB_{MSY}$, or it might be a *MSY*-proxy such as $x\%SB_0$ [5], or a quantity that incorporates non-equilibrium population dynamics such as $x\%SB_{current\ F=0}$. The value for *x* would be determined on a careful examination of relevant data (e.g. spawner recruitment data for that stock, other stocks of the same species, and stocks of similar species) to ensure appropriate characterization of resilience, and through the use of simulation modelling. The robustness of the reference point to factors not directly related to biological risk should be considered, e.g. how do the reference points respond to changes in the mix of fishing gears.

A final comment on the development of limit reference points relates to the relevance of multi-species/fisheries interactions. We would suggest that this is an issue for target reference points, not limit reference points. It seems sensible that, in multi-species fisheries, target reference points should be developed taking into account the interactions among species. It is important to consider the impact of fishing strategies, developed to achieve a target reference point (on average) for one stock, on the ability to achieve target and limit reference points for other stocks. Trade-offs between stocks should be considered when incompatible target reference points are encountered in mixed fisheries, but we would conclude that this does not apply to limit reference points which are developed to keep individual stocks at safe biological levels.

Operationalising limit reference points

The second part of the question considers how these reference points could be operationalized. We have interpreted this to mean how they could be incorporated into a fishery decision making process. This leads to the topics of decision rules, and management strategy evaluation (MSE) which have been covered in previous papers presented at the WCPFC-SC (Davies and Basson 2008).

---

[4] Simulation studies should be used to examine the expected biomass levels and yields that result from strategies that look to avoid exceeding MSY -based limit reference points

[5] For some New Zealand fish stocks, management strategies are evaluated against a criterion that spawning biomass cannot be predicted to drop below $20\%SB_0$ more than 10% of time (Francis 1992).

Decision rules are simply a set of rules that describe the management response that results from a particular stock assessment outcome (e.g. estimates of stock status relative to one or more reference points). A very simple example, similar to what is currently done in SC advice to the Commission:

> *$F_{current}$ /F_{MSY} = 1.3, therefore fishing mortality must be reduced by 23% (calculated* $1 - (1/(F_{current}\ /F_{MSY}))$

However, it is possible to imagine more complex decision rules that take into account multiple variables and uncertainty, e.g.:

> *$F_{current}$ /F_{MSY} = 1.3, and $SB_{current}$ /SB_{MSY} = 0.8 and there is a 90% probability that $SB_{current}$ / SB_{MSY} will continue to decline at current levels of fishing effort, therefore fishing mortality /catch /effort must be reduced by xx%*

The identification of appropriate limit reference points and associated decision rules is often undertaken via detailed simulation studies. These simulations studies should consider uncertainty in future conditions (e.g. recruitment variation and autocorrelation) and uncertainty in the current stock status (e.g. parameter and structural uncertainty) to allow accurate estimation of probabilities of exceeding limit reference points under a given fishing strategy. Modification to MULTIFCAN-CL software would be necessary to allow recruitment variability to be included in projections and there are questions regarding the best approach for describing uncertainty in current status (see response to question 3). The delays in data provision (e.g. the one-two year lag in receiving DWFN longline catch and effort data) reduces our ability to react quickly to changes in stock status. This will need to be factored into the simulation studies, and would be expected to result in more conservative management strategies.

The biggest problem that will be encountered in this process is likely to be the lack of a clear and simple management strategy to evaluate. Francis (1992) examined constant catch and constant fishing mortality as two alternative strategies. When this type of analysis was undertaken for the Commission for the Conservation of Southern Bluefin Tuna (CCSBT), the scientists had the benefit that the Commission managed the fishery with a single total allowable catch. Therefore, the MSE involved adjusting the catch limit up and down in response to changing stock status to ensure that management objectives were met (e.g. rebuilding targets achieved).

The current management strategy for bigeye and yellowfin tuna for WCPFC, CMM2008-01, includes catch limits, effort limits, area closures, gear restrictions, capacity limits, and unqualified exemptions (i.e. some fleets or geographical areas have no catch or effort limits associated with them). Further there is no formula for how these various 'limits' would be modified up or down in response to a change in stock status. For this reason, Hampton and Harley (2009) were only able to predict the likely consequence of the provisions of CMM2008-01, but to fully develop limit reference points and decision rules it will be necessary to incorporate the basis for a management response to changing stock conditions.

Other complexities will include the consideration of the regional patterns in exploitation, for example in the yellowfin tuna assessment the tropical regions are most heavily exploited while higher latitudes are only lightly exploited. So far the WCPFC has considered this, to an extent, by focusing many measures in the region between 20 N and 20 S.

3. Potential reference points for WCPFC assessments

We consider a reference point to be a level of a particular indicator that we are interested in addressing performance against. There are several indicators that can potentially be used to construct reference points that are useful for fishery management purposes. Some examples of the types of indicators that can be used to construct reference points and the types of information that these can provide to managers are provided in Table 1.

From Table 1 it is clear that reference points can be developed to meet different information needs and that not all reference points provide the same types of information. Specific reference points and indicators considered in this paper are provided in Table 2 and some of these are described in further detail below. This is not meant to be an exhaustive list, and, as dialogue with fishery managers continues, we would expect other reference points to be developed to help meet their information needs. In particular, we have not defined any reference points that might be derived through economic analysis, e.g. the biomass that might occur on average from a fishery strategy that maximizes the economic outputs from the fishery, a.k.a the biomass that supports the maximum economic yield (B_{MEY}). Once defined these can also be evaluated, but they are dependent on knowing the parameter values which characterize the economic or social system being managed.

Whilst most readers will be familiar with the standard *MSY*-based reference points (e.g. $B_{current}$ $/B_{MSY}$), there are three other types of reference points that might require further explanation, namely the, $B_{current\ F=0}$reference points, otherwise known as the 'no fishing' reference points; the $B_{F_{current}}$, or equilibrium biomass reference points; and the yield per recruit related reference points near the bottom of Table 2.

No-fishing reference points

Once a stock assessment run is completed, and the time series of recruitment and the starting biomass have been estimated, the stock assessment model is projected forward from the start of the model period and no catches are removed from the population and recruitments (estimated previously) are added into the model[6]. Such reference points have the advantage of not relying on equilibrium (on average) assumptions and can take account of periods of above or below average recruitment. These are typically evaluated by comparing a recent estimate of biomass from the model that includes fishing to the estimate of biomass from the model that didn't. The resulting output is an estimate of the level of depletion. If recent recruitment has been close to the average, and fishing mortality held more or less constant, then $B_{current}$ $/B_0$ should be similar to $B_{current}$ $/B_{current\ F=0}$. Figure 1 shows the regional fished and 'un-fished' biomass trajectories based on run 10 from the 2009 bigeye tuna assessment.

[6] The implementation of this also takes into account the spawner recruitment relationship, e.g. if we believe levels of fishing have reduced spawning biomass such that recruitment will also be reduced, then we need to correct for this in our simulations where no fishing has occurred.

Equilibrium biomass reference points

Typically we compare biomass over a recent period to the MSY -levels, e.g. $B_{current}$ $/B_{MSY}$, but it is also possible to make some 'equilibrium-based' comparisons to the MSY -level as well. Simply speaking, we calculate what level of biomass would we expect to get if the current pattern of fishing mortality was maintained 'forever' and recruitment was as predicted by the spawner recruitment relationship. It is essentially the same as doing a very long term projection with current fishing mortality. For example $B_{F_{current}}$ $/B_{MSY}$ is the ratio of the biomass that is predicted to occur if current fishing mortality was continued, relative to the B_{MSY}. It provides further information on the long-term sustainability of current fishing mortality rates.

Yield-per-recruit based reference points

The estimates of MSY and related quantities are strongly related to the mix of fishing gears (Maunder 2002), i.e. you can get a different MSY depending on how you harvest the stock with respect to ages and sizes of fish that you catch. A simple way to examine this is by looking at the relative biomass of a cohort of fish projected through time in the absence of fishing. As the fish age, they increase in weight due to growth, but decrease in number due to natural mortality. In the example of bigeye tuna (Figure 2) the biomass of the cohort is maximized at an age of 15 quarters which, based on the estimated growth curve, equates to a length of about 125 cm. If it was possible to harvest all of the fish at that age (and ignoring the need for spawning), then yield would be maximized. Catching the fish at any other age would result in a loss of potential yield. Of course it is not possible limit catch to a single precise age (as all fishing gears select fish over a range of sizes and ages and fishing is not a single instantaneous event). Also the ability to modify the age of capture depends on whether the species is taken as a target or bycatch species.

Considering that, and noting that typically we only calculate a MSY that is obtained based on the mix of fishing gears over some recent period, it is also possible to see how MSY could be changed by changing the mix of fishing gears and how much potential yield is being forgone due to the pattern of fishing. Figure 3 (top) shows the annual estimates of MSY, based on the mix of fishing gears in that year, and Figure 3 (bottom) indicates the mean age of capture in the different fisheries estimated in the bigeye tuna assessment.

Estimation of many of the reference points listed in Table 2 and evaluation of current conditions against them will be undertaken for the 2009 assessments for south Pacific albacore, bigeye and yellowfin tuna and these results are not replicated here.

4. Alternative methods for assessing stock status against reference points

Modern fisheries stock assessment and fisheries management is seldom based around a single number describing the status of the stock against a reference point. More commonly both a 'best' estimate and some quantification of the uncertainty, e.g. probability of being above or below the reference point of interest, is used in formulating management advice.

For assessments presented to the WCPFC-SC a variety of approaches for describing uncertainty have been considered for formally describing stock status:

- Likelihood profiles on stock status relative to MSY -related reference points (e.g. $B_{current} / B_{MSY}$ and $F_{current} / F_{MSY}$);
- Point estimates of status relative to reference points for a small number of one-change sensitivity analyses); and
- Point estimates of status relative to reference points for a large number of model runs), e.g. southwest Pacific swordfish assessment (Kolody et al. 2009).

These approaches can be categorized into those that attempt to consider statistical or parameter uncertainty within a single model run (e.g. likelihood profiles), versus those that attempt to incorporate structural uncertainty across alternative model structures (e.g. point estimates for a large number of model runs). Computing constraints make it difficult to consider both types of uncertainty with the types of models currently used to assess the main tuna stocks.

In addition, the structural sensitivity approach based on partially confounded factorial designs was applied by Hoyle et al. (2008) to assess the key sources of uncertainty in the 2008 bigeye tuna assessment, Also, an alternative stock assessment approach, Stock Synthesis, has been used to help verify / test the robustness of the results from MULTIFAN-CL assessments and examine potential areas of uncertainty (e.g. robustness of conclusions to consideration of length-specific selectivity) (Langley and Methot 2008). However, in neither case were the results used directly in the provision of advice on stock status.

Methods

Here we consider three approaches for determining stock status against reference points:

1. Base case model results with uncertainty based on likelihood profiles (i.e. statistical uncertainty resulting from an assumed model structure);
2. Base case model plus one-change sensitivity analyses with uncertainty based on combined likelihood profiles (a limited combination of statistical and structural uncertainty);
3. A full cross grid of 128 point estimates based on all possible combinations of seven key model assumptions (each with two options) (see Harley et al. (2009a) for further details of the options) (i.e. a more complete treatment of structural uncertainty, but without statistical uncertainty).

Results and discussion

Figure 4 and Figure 5 provide the likelihood profiles for $B_{current} / B_{MSY}$ and $F_{current} / F_{MSY}$ respectively for the base case model run for YFT, the one change sensitivities undertaken as part of the YFT SSA, and a overall profile that combines the individual profiles. Unfortunately the base case profiles in these examples are obscured by the profile for model 07 (size frequency weighting). In the development of the SSA it was considered that all of the runs were approximately equally likely. The first thing to note is that there is little overlap between many of the scenarios, and some scenarios have no overlap. Some models have a high likelihood of exceeding a reference point, whilst others have next to zero likelihood; this is clearly shown in the third column of Table 3.

It is therefore clear that an individual likelihood profile, which attempts to characterize within model statistical uncertainty, is insufficient to adequately describe the [more] ‘real’ uncertainty in stock status in relation to these reference points which can be seen from the one change sensitivity results[7]. If you combine the individual likelihood profiles, and assume equal weight, 13% of the 'mass' of the combined profile for $B_{current}\ /B_{MSY}$ is less than one and 34% of the mass of the combined profile for $F_{current}\ /F_{MSY}$ is greater than one. Whilst there is likely to be some theoretical statistical discussions as to the merit of this type of approach – the latter type of estimate does seem to more accurately account for the uncertainty in the stock assessment.

The one-change sensitivity analyses do not allow for potential interactions among the different factors. Given the seven factors and 2 options for each, this provides 128 possible configurations[8]. Whilst it is not currently feasible to undertake these number of runs calculating likelihood profiles, it is possible to get point estimates for the key reference points (takes about six days using around 20-30 machines). Figure 6 uses a Kobe-style plot of biomass versus fishing mortality to display the point estimate of current conditions from the base case, the one-change model runs, and the runs in the full 'grid'. This type of information provides the basis for another approach based on the proportion of the individual model runs that exceed a given reference point. For $F_{current}\ /F_{MSY}$, these values are provided in Table 3. These numbers are quite different to the likelihood profile values as they integrate uncertainty in the other factors, but the overall probabilities are comparable between the two approaches.

It is also possible to estimate uncertainty based on a normal approximation of the variance covariance matrix and application of the delta method (and this is commonly done to produce confidence limits of recruitment and biomass trajectories), and run a subset of the 128 runs based on alternative partially confounded factorial designs (Hoyle et al. 2008). This latter approach was tested for the pre-assessment workshop (Harley et al. 2009b) and shows promise, but rather than describe these other approaches here, they should be included in any future work programe in this area.

The results presented here cast strong doubts as to the validity of simply using a single base case model to characterize the uncertainty in stock status and there are at least two potential approaches that offer improvements, but there is further work needed to address statistical issues, and those of the selection of models to include, e.g. representativeness.

5. Sensitivity of reference points

Using the model runs undertaken using the 2007 YFT stock assessment data set we illustrate how sensitive various reference points are to alternative data/model structures and biological assumptions. This will also include consideration of stock status relative to these reference points under these alternatives. Sources of uncertainty that we consider include the steepness of the spawner recruitment relationship, changes in catch histories, and changes in CPUE series (Figure 7). In addition we will

[7] One reason for the narrowness of the profiles is likely due to steepness being fixed rather than estimated, but we note the difficulties in estimating steepness elsewhere in this paper so are not discussed further here.

[8] We call this a Structural Uncertainty Analysis (SUA), which is different from a Structural Sensitivity Analysis (SSA) in that the former is specifically aimed to characterize uncertainty in an assessment while the later is used in an explorative way to better understand what the model is sensitive to (e.g.. a tool that can also be used to help develop research needs).

provide an example of a retrospective analyses where we will use a single model and run the assessment model several times, each time excluding another year of data.

Results and discussion

The assumed value for the steepness of the spawner recruitment had a very substantial impact on many of our reference points (Figure 8 and Figure 9). MSY, F_{MSY}, $F_{current}/F_{MSY}$, were all very sensitive to the assumed value of steepness (83 - 138% difference over the range compared to that estimated from a steepness of 0.62). These quantities increased in an almost linear fashion. B_{MSY} and B_{MSY}/B_0 were less sensitive (39% and 30% changes, respectively), but this still represents considerable variation. Interestingly the impact on the estimated B_0 was only 10% and $B_{current}$ varied only 1%. Therefore, in this example the assumed value of steepness had little impact on absolute biomass levels either in the past or present, but the impact on stock status relative to the reference points was considerable. However, interestingly the alternative biomass based reference point, $B_{current}/B_{current\ F=0}$, was only slightly less sensitive than $B_{current}/B_{MSY}$ (40% versus 44%) to uncertainty in steepness. This is because steepness is still factored in to the calculation of $B_{current\ F=0}$ in increased scaling of recruitment for lower values of steepness.

In this example the objective function value did not vary significantly over most of the range of the steepness values considered. This is not surprising as it is typically very difficult to estimate steepness from a single data set, particularly if the stock has not been reduced to low levels to allow better characterization of the left-hand side of the spawner-recruitment curve.

For the current WCPO tuna stocks, the assessment models do not predict that biomass has declined below B_{MSY} levels and therefore the expected declines in average recruitment would be very small. To get good estimates of steepness you need several observations of recruitment levels at low stocks sizes – we do not have these for our stocks. Given that recruitment variation is very high and stock reductions are moderate, unless steepness is very low (in which case we would have already detected a decline in average recruitment) any estimation of steepness involves extrapolation well outside our observations in the face of very noisy spawner recruitment data. Therefore, it is extremely doubtful that steepness can be estimated within our individual stock assessment models. In the face of such variation, steepness estimates can be driven by structural artifacts in the models or data.

So steepness is important, but it is unlikely that we can tell what it is from the data we have before us. We provide some recommendations for future work to help either reduce the uncertainty relating to steepness or better incorporate it into the management advice process later in the paper.

Table 4 and Table 5 show the impact on key reference points from two sets of sensitivity analyses. In the first example two alternative catch histories were compared and steepness was allowed to be estimated. The higher catches were associated with a higher MSY and higher estimated steepness. The increase in steepness was likely driven by increased catches of small fish in recent years which would result in higher recent recruitment over a time when spawning biomass was at its lowest.

In the second example, the inclusion of effort creep in longline fisheries led to a higher MSY, B_0, and B_{MSY}, but had little impact on $B_{current}$, therefore stock status was more pessimistic. What appears (to the untrained eye) to be only a small change in the CPUE series (Figure 7) can take the stock from a state where overfishing is not occurring and the stock is not in an overfished state to a situation where the stock is overfished and overfishing was occurring.

The final series of model runs used to assess the sensitivity of the reference points was the retrospective analyses. Estimates of some of the key reference points across the model runs are provided in Figure 10. Not all of these quantities are comparable over the model runs, e.g. none of the quantities that include either $B_{current}$ or $F_{current}$, but they are included here to show that the updated data each year has led to large changes in the stock status over the time period. Quantities that are comparable are the equilibrium quantities MSY, B_0, B_{MSY}, and B_{MSY}/B_0 and it is seen that there is relatively little change in any of these quantities over the time period shown. The main changes that are apparent in MSY (i.e. decline) are in fact being driven by the changes in the size at harvest as proportionally more of the catch is taken in fisheries taking small fish. In this example, the equilibrium quantities are relatively robust in their estimation.

Figure 11 and Figure 12 show the trajectories for key model outputs from the retrospective model runs. These plots provide some insights into how reference points are calculated and the reliability placed on recent estimates of some model quantities. It should be noted that most recent (or final) recruitment estimates are typically highly uncertain and are usually 'corrected' once more data is available (e.g. the end of the series 'sticks out') and there are some extreme values such as for the model that uses data up to 2004. Therefore recent recruitment estimates should always be viewed with caution. That same extreme recruitment estimate had an impact on fishing mortality and there is certainly clear evidence that, in general, the most recent estimates of fishing mortality are 'corrected' with the addition of new data. Some of the end year patterns in recruitment are carried through to the estimates of total biomass, but it appears that spawning biomass is far more stable. This is because it is comprised more of older fish in the population (for which the model has 'converged') and less on the younger cohorts which are more uncertain.

There is a general question about how many of the most recent periods (quarters) should be excluded due to their uncertainty. Willingness to accept this 'uncertainty' depends on whether there is also a trend in stock status indicators, e.g. if things were relatively stable then you might be willing to reject more recent estimates, but if there was a strong increase in fishing mortality then you might be willing to accept some of the more recent uncertainty to keep the assessment conclusions as relevant as possible to recent conditions. Presently reference points are developed based on either a definition of 'current', which involves excluding the most recent year and then taking the average over the proceeding four years, or 'latest', which takes the value for the last time period included in the model. Fishing mortality based reference points are only evaluated with respect to their 'current' values, while the biomass based reference points are evaluated at both 'current' and 'latest' levels. As B_{MSY} and MSY are calculated using the fishing mortality at age profiles, they are based on the same period as the F-based reference points, but both current and latest biomass levels can be evaluated against them.

Based on the results of this retrospective analysis the decision to base fishing mortality reference points only on their 'current' values seems appropriate given the uncertainty displayed in the most recent estimates. Similarly, comparisons based on B_{latest} are probably too uncertain, but comparisons against SB_{latest} are more appropriate and should be given more weight than the $SB_{surrent}$ comparisons.

One type of robustness that we have not explicitly considered here is the stability of reference points from one assessment to the next. There are continual improvements in the methodologies used in the assessments (e.g. introduction of length-specific selectivity) and biological studies are improving our understanding of key biological processes. Further there are uncertainties in some key catch statistics (e.g. IDPH catches and PS catches of juvenile BET/YFT) which we are seeking to reduce over time and this often results in changes to the historical data. All of these developments can and do result in changes in the absolute estimation of reference points and stock status. We have illustrated the types of changes that can occur from different estimates / assumptions about steepness and historical catches, and how new data can change our view of the world, which can provide some insights into changes in assessments over time. The one key issue to recognize is that the stock assessment models are not so complex that changes between assessments cannot be explained and understood. One of the key roles in the assessments is to carefully examine and explain any such changes that occur. Nevertheless, this is a key area that managers must understand, as significant changes in fishing strategies could be driven by changes in stock assessment methodologies or data that lead to a different picture of the state of the stock against previously agreed reference points.

6. Conclusions and recommendations for future work

MSY-quantities as default limit reference points

Conclusions:

- This is more of a technical policy / legal issue than one for scientists to consider alone and further legal advice should be sought, including the potential for Article 5(b) to have a bearing on the selection of limit reference points;
- Conservation-based limit reference points have an explicit biological purpose and are they are developed to constrain harvesting within safe biological limits[9]. Consequently spawning biomass will be a better biomass indicator than total biomass for the development of limit reference points relating to biological risk;
- Based on our understanding Annex II of the UNFSA, F_{MSY} and SB_{MSY} should be considered default limit reference points and their use would be consistent with precautionary principles;
- Alternative limit reference could be developed for the four main tuna stocks and possible candidates include: $x\%SB_{MSY}$, $x\%SB_0$, and $x\%SB_{current\ F=0}$ with the value for *x* being determined from the analysis of relevant data and simulation studies. These alternative reference points would need to be developed to ensure that principals of limit reference points are still achieved, e.g. to maintain harvesting within safe biological limits (e.g. steepness will still be an issue here, even for non-MSY based reference points).

[9] Notwithstanding this, fishery managers might still seek to have alternative limit reference points (e.g. reference points that they wish to avoid with high probability), that are above those required for biological purposes, for other reasons.

- Reference points need to be associated with decisions about what happens if they are triggered (or to avoid them being triggered) and the triggering should consider uncertainty / risk. The identification and development of decision rules typically requires involvement from other stakeholders in the fishery (not just scientists).

Future work / recommendations:

- A literature review / meta-analysis (e.g. Myers et al. 1999) should be undertaken to provide insights into levels of depletion that may serve as appropriate limit reference points for Pacific tuna stocks and provide more information on plausible values for steepness. Such studies should also help identify what is meant by a "safe-biological limit" for a given stock;
- Management strategy evaluation (MSE) is a framework that should be used to further develop limit reference points (and targets as well). Key activities will include:
    - Simulation studies to assess the implications of alternative limit reference points. Such studies should include uncertainty in future conditions (e.g. stochasticity and autocorrelation in recruitment) and uncertainty in the current status (e.g. parameter and structural uncertainty). The potential impact of delays in data provision should also be included;
    - Scientists working with fishery stakeholders to allow the design of management strategies that can be tested through simulation studies.

Potential reference points for WCPFC assessments

Conclusions:

- A range of reference points can be developed that provide different information for decision makers and it is useful to consider these when formulating management advice;

Future work / recommendations

- Scientists should be engaged in discussions with fishery decision makers over reference points, particularly with respect to getting insights into the types of information that fishery managers find useful. Such work should include bioeconomic modelling and other approaches to help construct some target reference points that take into account economic considerations (e.g. B_{MEY}).

Alternative methods for assessing stock status against reference points

Conclusions

- There is a real need to incorporate uncertainty into the calculation of stock status and the provision of management advice;
- Current approaches rely heavily on a single model run (e.g. base case) and so-called parameter uncertainty (uncertainty within one model run). However, it is now clear that structural uncertainty is often larger and alternative plausible model runs can often give quite different results; and

- The Structural Uncertainty Analysis (SUA) provides a potential approach for incorporating many of the major sources of uncertainty, but is only at an early stage of its development.

Future work / recommendations

- Stock assessment methods research should examine alternative approaches for describing uncertainty. In particular, efforts should be directed at further developing the SUA approach. Areas to focus on include:
  - Development of diagnostics for individual model runs
  - Consideration of model selection and weighting schemes
  - Integration of parameter uncertainty

Sensitivity of reference points

Conclusions

- There are many possible changes to stock assessment models that can change stock status outcomes, e.g. the assumed steepness of the spawner-recruitment relationship, trends in longline CPUE (considered to reflect abundance) and absolute levels of catches. Such differences are typically greater than those seen by simply adding another year of data to the same model.
- Higher levels of steepness lead to higher MSY and lower B_{MSY}, e.g. the biomass that supports the MSY is a much lower proportion of the overall biomass;
- Steepness is therefore very important, but reliable estimates of steepness are not possible from the data we have for our stocks at the moment;
- Retrospective analyses suggest the following:
  - Some MSY related quantities are quite robust to the addition of new data (e.g. MSY and B_0);
  - Recent recruitment and fishing mortality estimates are uncertain and this carries through into estimates of total biomass;
  - Spawning biomass estimates are far more stable and it should be possible to construct reasonably robust reference points using more recent estimates than is advisable for fishing mortality or total biomass; and
  - Stock projections should consider the uncertainty in the most recent estimates of recruitment.

Future work / recommendations

- Sensitivity analyses should play an important role in the development of robust stock status advice;
- When stock assessments are updated any material changes from the previous assessment should be clearly outlined to allow for the impacts of individual changes to be evaluated;
- It is not recommended that model runs where steepness is estimated should form the basis of management advice, but uncertainty in steepness must be incorporated into management advice;

- Retrospective analyses should be conducted as regular part of a stock assessment;
- Reference points based on fishing mortality and total biomass should not include the most recent year as problems of uncertainty likely outweigh the benefits of being 'current'; and
- Reference points based on spawning biomass reference should include more recent time periods.

7. Acknowledgements

The authors thank Wez Norris, Rob Campbell, Adam Langley, and Professor Martin Tsamenyi for their very useful comments on the paper and some of the concepts contained within it.

8. References

Anon. 1995 The UN Fish Stocks Agreement, Annex II.

Davies, C., and Basson, M. 2008. Approaches for identification of appropriate reference points and implementation of MSE within the WCPO. WCPFC-SC4-2008/GN-WP-10, Port Moresby, Papua New Guinea, 11-22 August 2008.

Francis, R. I. C. C. 1992. Use of risk analysis to assess fishery management strategies: a case study using orange roughy (*Hoplostethus atlanticus*) on the Chatham Rise, New Zealand. Can. J. Fish. Aquat. Sci. 49: 922-930.

Harley, S. J., Hoyle, S. D., and Bouyé, F. 2009a. General structural sensitivity analysis for the yellowfin tuna stock assessment. WCPFC-SC5-2009/SA-IP-03, Port Vila, Vanuatu, 10-21 August 2009.

Harley, S. J., Davies, N., Hoyle, S. D. 2009b. Report from the SPC pre-assessment workshop, Noumea, April 2009. WCPFC-SC5-2009/SA-IP-01, Port Vila, Vanuatu, 10-21 August 2009.

Harley, S. J., Hoyle, S. D., Langley, A., Hampton, J., and Kleiber, P. 2009c. Stock assessment of bigeye tuna in the western and central Pacific Ocean. WCPFC-SC5-2009/SA-WP-04, Port Vila, Vanuatu, 10-21 August 2009.

Hoyle, S. D., Langley, A., and Hampton, J. 2008. General structural sensitivity analysis for the bigeye tuna stock assessment. WCPFC-SC4-2008/SA-WP-03, Port Moresby, Papua New Guinea, 11-22 August 2008.

Kolody, D., Campbell, R., and Davies, N. 2009. Final report on south-west Pacific swordfish (*Xiphias gladius*) stock assessment 1952-2007. WCPFC-SC5-2009/GN-IP-02, Port Vila, Vanuatu, 10-21 August 2009.

Langley, A., Hampton, J., Kleiber, P., and Hoyle, S. 2007. Stock assessment of yellowfin tuna in the western and central Pacific Ocean, including an analysis of management options. WCPFC SC4 SA WP-1, Honolulu, Hawaii, 13-24 August 2007.

Langley, A., and Methot, R. 2008. A preliminary stock assessment of bigeye tuna in the western and central Pacific Ocean using stock synthesis 3 (SS3); a comparison with MULTIFAN-CL. WCPFC-SC4-2008/SA-WP-02, Port Moresby, Papua New Guinea, 11-22 August 2008.

Maunder, M. N. 2002. The relationships between fishing methods, fisheries management and the estimation of MSY. Fish and Fisheries 3: 251-260

Myers, R. A., Bowen, K. G., and Barrowman, N. J. 1999. The maximum reproductive rate of fish at low population sizes. Canadian Journal of Fisheries and Aquatic Sciences 56: 2404-2419.

Table 1: Types of indicators that can be used to construct reference points

Indicator / Reference points	Example of the type of information provided
Biomass	Are current levels of biomass sufficient to allow the species to fulfill its role in the ecosystem?
Spawning biomass	Are levels of spawning biomass sufficient to ensure that recruitment is not compromised?
Fishing mortality	Are the current levels of fishing mortality sustainable in the long terms?
Catch	Are current catch levels sustainable in the long terms?
Average size of fish caught	Are we harvesting the population in such a way as to maximize potential yields?
CPUE / vulnerable biomass	Are catch rates (or the biomass of the sizes of fish targeted) at levels that will allow for profitable fisheries?

Table 2: Description of various reference points and indicators considered in this study.

Symbol	Description
$C_{current}$	Average annual catch over a recent period[10]
C_{latest}	Catch in the most recent year
$F_{current}$	Average fishing mortality-at-age[11] for a recent period
F_{MSY}	Fishing mortality-at-age producing the maximum sustainable yield (*MSY*[12])
$Y_{F_{current}}$	Equilibrium yield at $F_{current}$
$Y_{F_{MSY}}$	Equilibrium yield at F_{MSY}. Better known as *MSY*
$C_{current}\ /MSY$	Average annual catch over a recent period relative to MSY
$C_{latest}\ /MSY$	Catch in the most recent year relative to *MSY*
F_{mult}	The amount that $F_{current}$ needs to be scaled to obtain F_{MSY}
$F_{current}\ /F_{MSY}$	Average fishing mortality-at-age for a recent period relative to F_{MSY}
B_0	Equilibrium unexploited total biomass
B_{MSY}	Equilibrium total biomass that results from fishing at F_{MSY}
B_{MSY}/B_0	Equilibrium total biomass that results from fishing at F_{MSY} relative to B_0
$B_{current}$	Average total biomass over a recent period
B_{latest}	Total biomass in the most recent year
$B_{F_{current}}$	Equilibrium total biomass that results from fishing at $F_{current}$
$B_{current\ F=0}$	Average total biomass over a recent period in the absence of fishing
$B_{latest\ F=0}$	Total biomass predicted to exist in the absence of fishing
SB_0	Equilibrium unexploited total biomass[13].
$B_{current}\ /B_0$	Average total biomass over a recent period relative to B_0
$B_{latest}\ /B_0$	Total biomass in the most recent year relative to B_0
$B_{F_{current}}\ /B_0$	Equilibrium total biomass that results from fishing at $F_{current}$ relative to B_0
$B_{current}\ /B_{MSY}$	Average total biomass over a recent period relative to B_{MSY}
$B_{latest}\ /B_{MSY}$	Total biomass in the most recent year relative to B_{MSY}
$B_{F_{current}}\ /B_{MSY}$	Equilibrium total biomass that results from fishing at $F_{current}$ relative to B_{MSY}
$B_{current}\ /B_{current\ F=0}$	Average total biomass over a recent period / the biomass in the absence of fishing
$B_{latest}\ /B_{latest\ F=0}$	Total biomass in the most recent year / the biomass in the absence of fishing
$Crit_{age}$	The age at which harvest would maximize the yield per recruit
$Crit_{length}$	The length at which harvest would maximize the yield per recruit
$Mean_{age}$	The mean age of the catch over a recent period
$Mean_{length}$	The mean length of the catch over a recent period
Y_{lost}	The proportion of the maximum yield per recruit lost by the mean age at harvest

[10] Some recent period used for the purpose of averaging fishing mortality or other quantities. Typically excludes the most recent year due to uncertainty, but covers the preceding four years, e.g. 2003-2006.

[11] This age-specific pattern is dependent on both the amount of fishing and the mix of fishing gears, e.g. relative catches of small and large fish

[12] MSY and other MSY-related quantities are linked to a particular fishing pattern and the MSY will change, for example, based on changes in the relative catches of small and large fish

[13] Similar quantities as above for total biomass can also be calculated for spawning biomass and are not repeated here

Table 3: Point estimates of F_{CURR}/F_{MSY} , and probability of that $F_{CURR} > F_{MSY}$ based on a likelihood profile and the grid of model runs (i.e. the proportion of runs that included the particular model option where $F_{CURR} > F_{MSY}$) for the base model and the one-change sensitivity analyses. Note that the grid probability for the base model is based on all model runs with steepness = 0.62.

		$p(F_{CURR}>F_{MSY})$	
	F_{CURR}/F_{MSY}	Likeprof	Grid
All runs		0.343	0.180
Base model	0.939	0.051	0.359
Steepness alt.	0.567	0.000	0.000
Mortality alt.	1.016	0.698	0.188
Growth alt.	0.619	0.000	0.000
CPUE CV alt.	1.097	0.990	0.203
Size N alt.	1.021	0.053	0.188
Effort creep alt.	1.065	0.953	0.234
SB calculation alt.	0.847	0.000	0.125

Table 4: Estimates of key reference points for a yellowfin tuna stock assessment with two alternative time series of catch from the fisheries of Indonesia and the Philippines. Note that steepness was estimated in each case.

Reference point	**Low catch**	**High catch**
MSY (mt)	376,760	452,400
$F_{current}/F_{MSY}$	1.00	0.91
B_0 (mt)	3,420,000	4,002,000
B_{MSY} (mt)	1,536,000	1,626,000
B_{MSY}/B_0	0.45	0.41
$B_{current}/B_0$	0.50	0.50
$B_{current}/B_{MSY}$	1.12	1.23
$B_{current}/B_{current\ F=0}$	0.63	0.57
steepness	0.60	0.71

Table 5: Estimates of key reference points for a yellowfin tuna stock assessment with two alternative assumptions about effort creep. Note that steepness was estimated fixed at 0.62 in each case. Harley et al. (2009a) provides further details of the scenarios considered.

Reference point	No creep	Creep
MSY (mt)	412,800	484,000
$F_{current} / F_{MSY}$	0.94	1.07
B_0 (mt)	3,898,000	4,738,000
B_{MSY} (mt)	1,654,000	1,997,000
B_{MSY} / B_0	0.42	0.42
$B_{current} / B_0$	0.49	0.37
$B_{current} / B_{MSY}$	1.16	0.88
$B_{current} / B_{current\ F=0}$	0.52	0.47
steepness	0.62	0.62

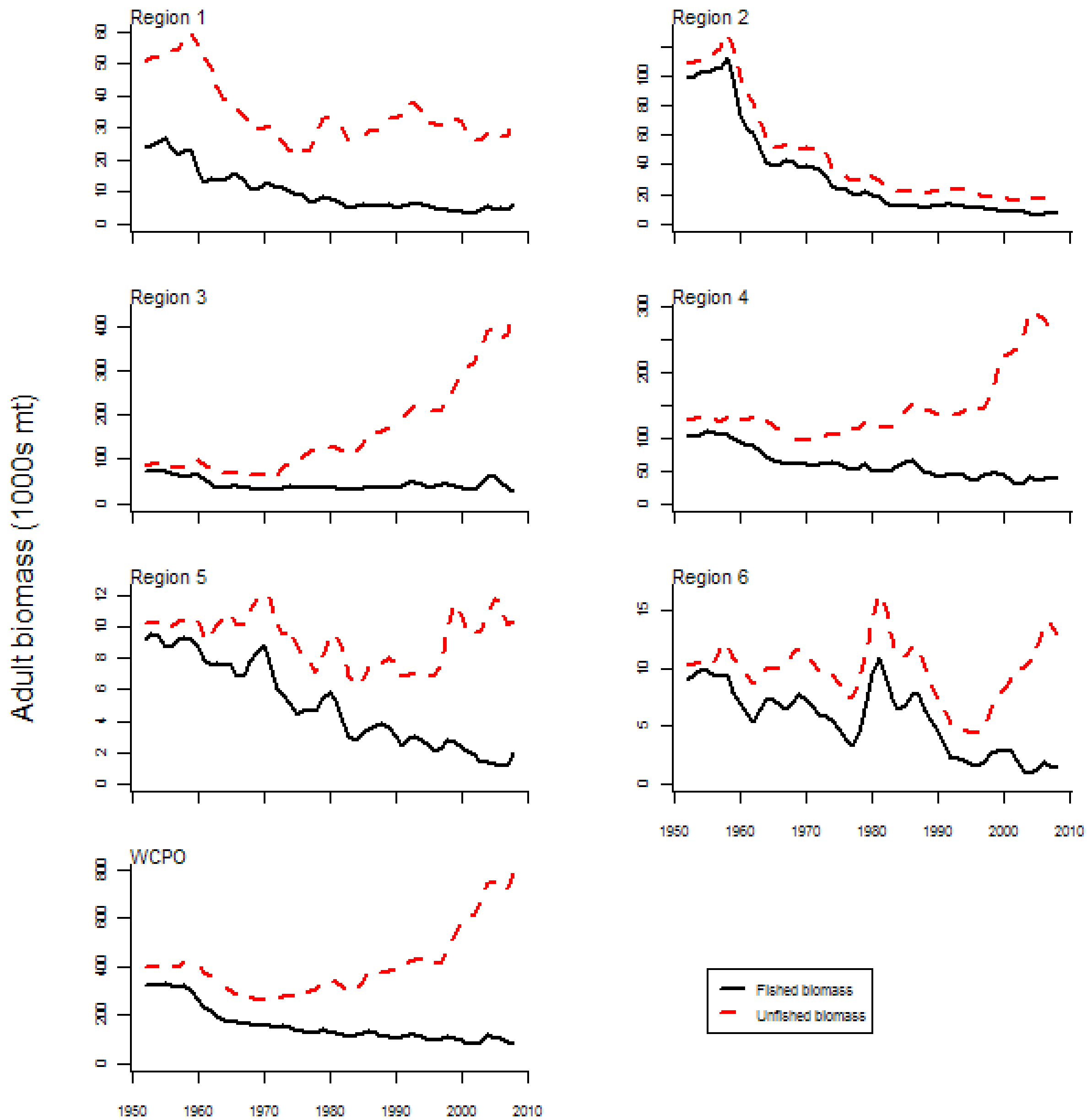


Figure 1: Estimated fished and unfished biomass trajectories by region and entire WCPO for bigeye tuna based on run 10 from the 2009 assessment (Harley et al. 2009c).

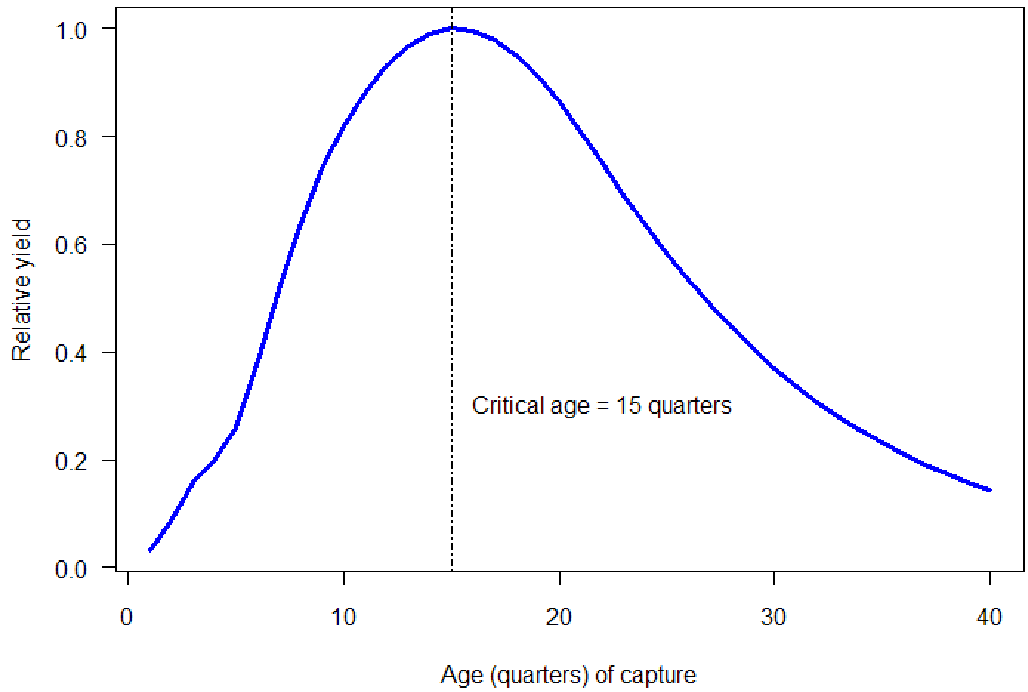
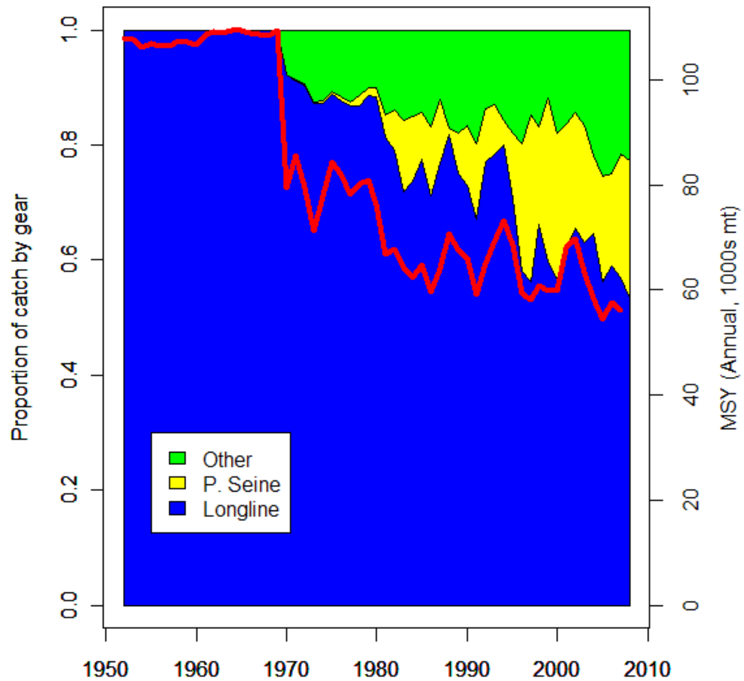


Figure 2: Estimates of the relative yield that would be taken if all fish were harvested instantaneously at that age (which is equal to the biomass at age of a cohort). The maximum, age 15 quarters, is referred to as the critical age.



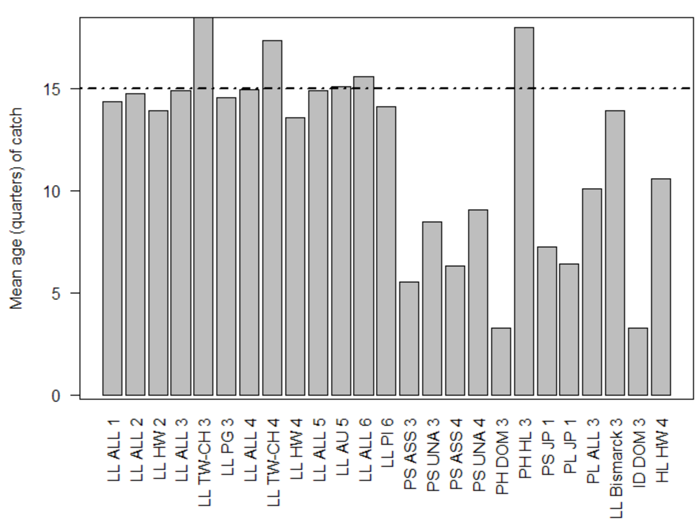


Figure 3: MSY calculated for each year (based on the mix of gears in that year) versus the proportion of catch taken by the main gear types (top) and the average age of fish taken by each of the fisheries defined in the 2009 bigeye tuna stock assessment (bottom).

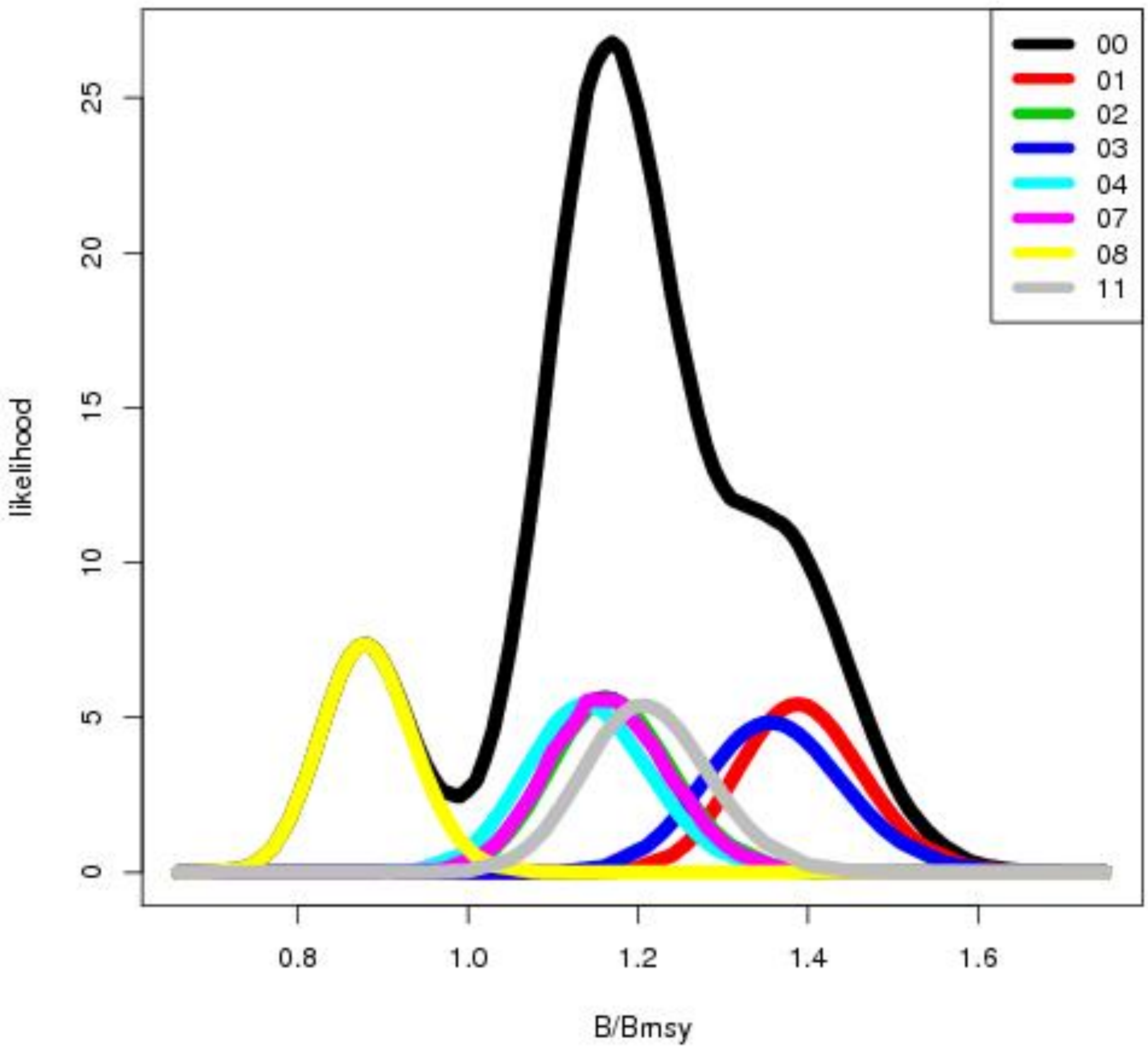


Figure 4: Likelihood profiles for yellowfin tuna B_{CURR}/B_{MSY} for the base case (2007 assessment) and selected one change sensitivity analyses.00-base;01-steepness; 02-mortality;03-growth; 04-CPUE CV;07 size frequency weighting;08-effort creep;11-SB calculation method. The heavy black line is a combined profile.

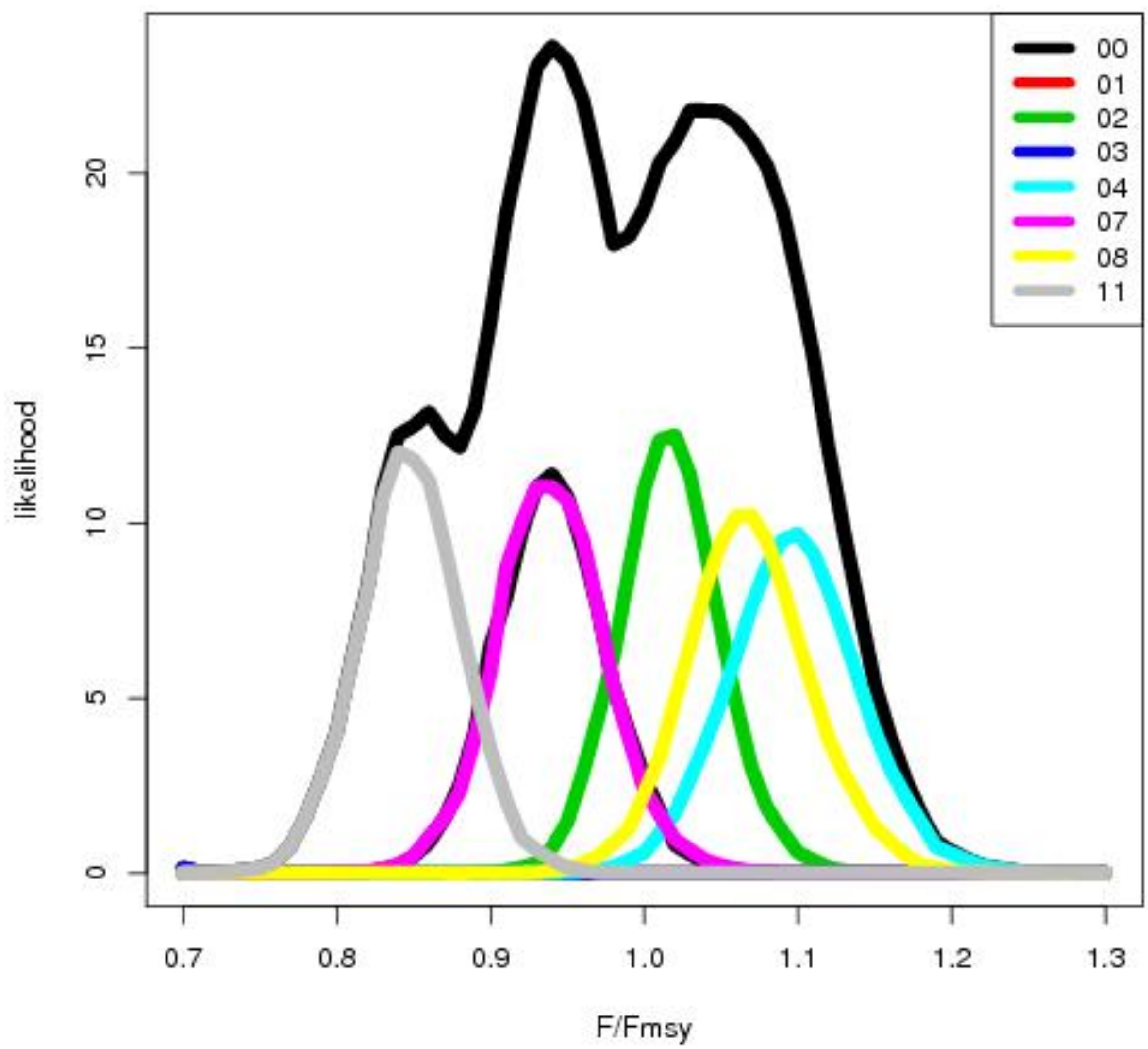


Figure 5: Likelihood profiles for yellowfin tuna F_{CURR}/F_{MSY} for the base case (2007 assessment) and selected one change sensitivity analyses.00-base;01-steepness; 02-mortality;03-growth; 04-CPUE CV;07 size frequency weighting;08-effort creep;11-SB calculation method. The heavy black line is a combined profile.

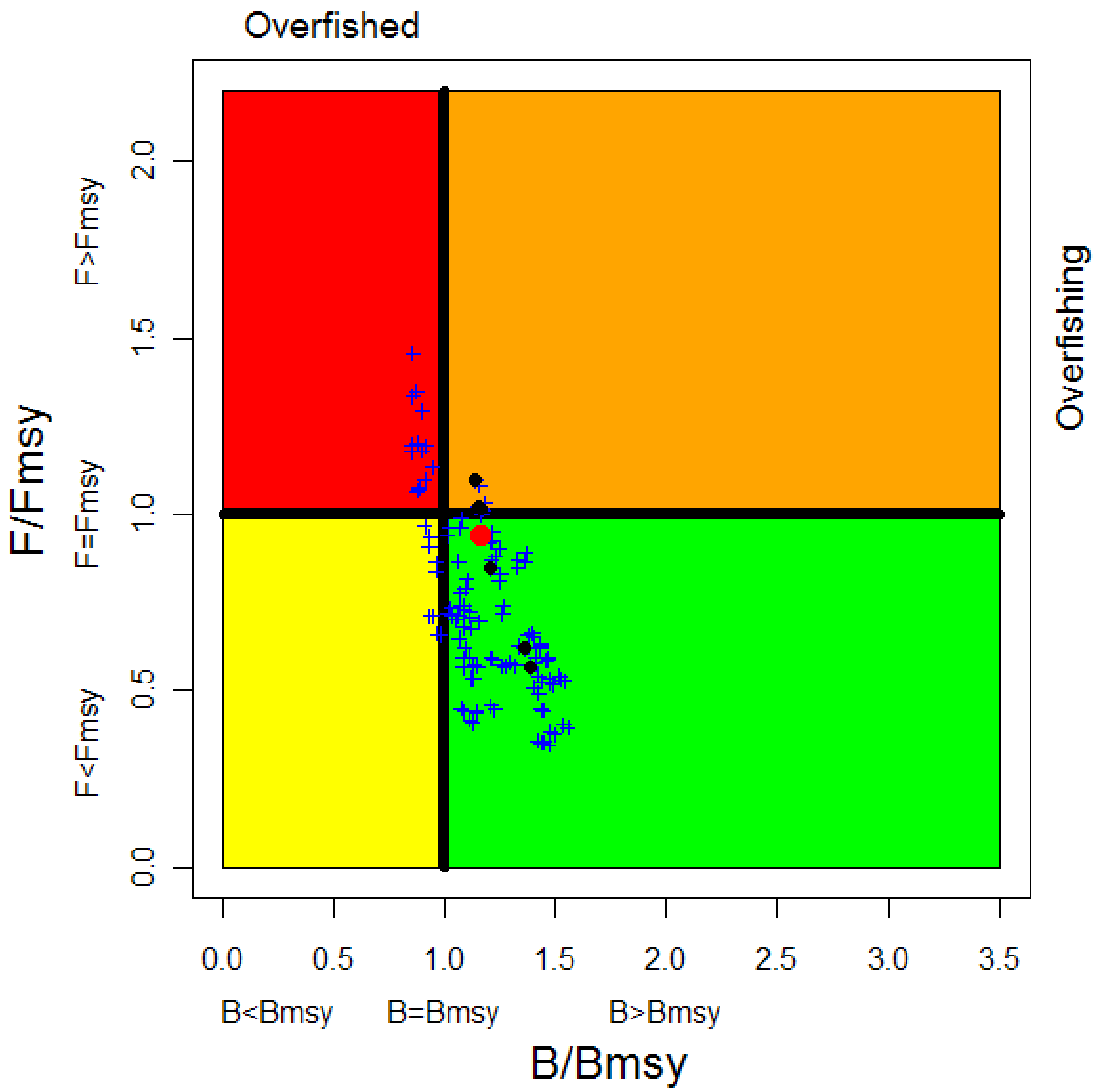


Figure 6: A Kobe-style plot showing $B_{current}/B_{MSY}$ versus $F_{current}/F_{MSY}$ for the yellowfin tuna base case model (2007 assessment) of the structural sensitivity analysis (large red circle), the one change sensitivity analyses (smaller black circles), and all model runs in the full grid (blue pluses +).

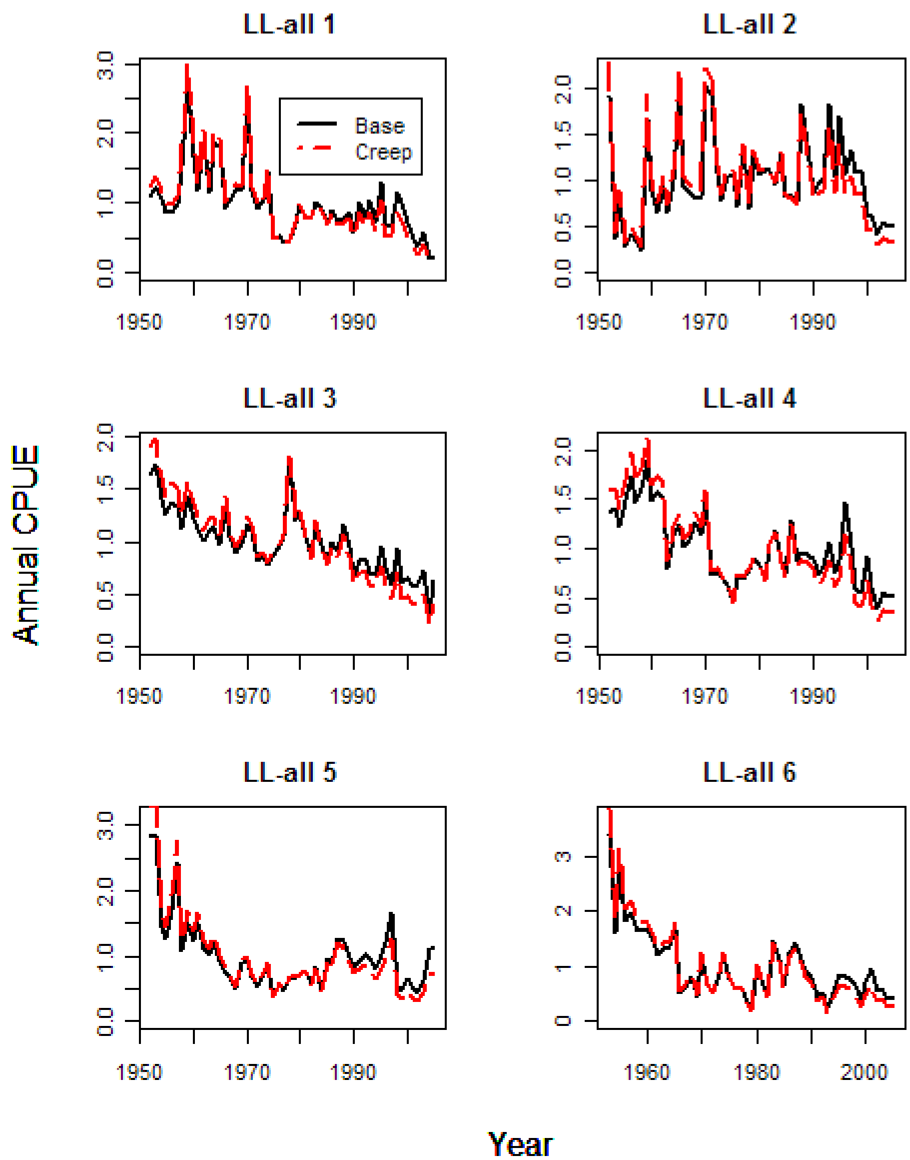


Figure 7: Alternative CPUE series used for the main longline fisheries in each region factoring in potential effort creep.

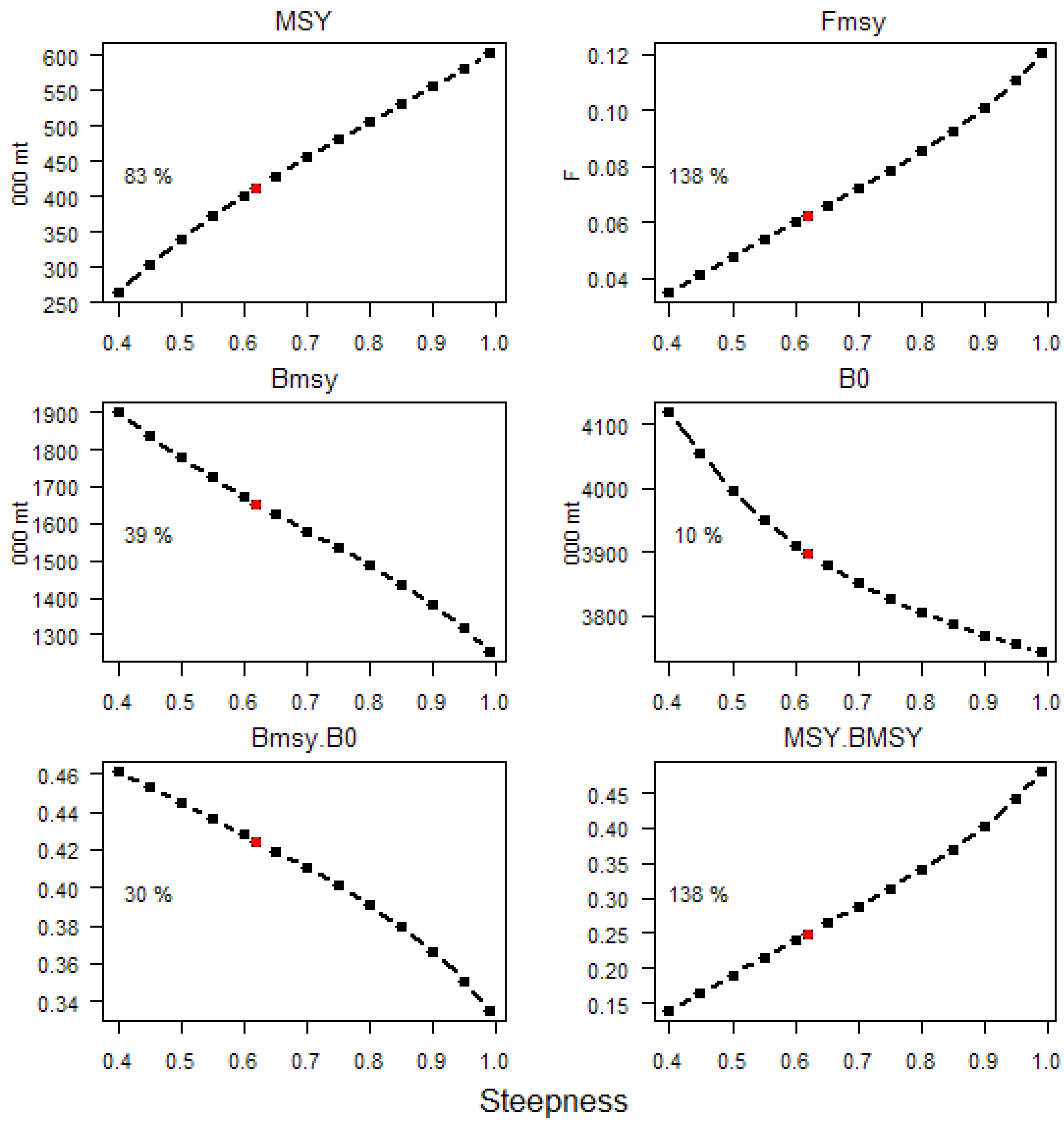


Figure 8: Impact of the assumed value of steepness on important assessment and management quantities. The red circle indicates the value for the base case model of the YFT SSA (Harley et al. 2009a) (steepness = 0.62). The percentage change in each quantity over the range of steepness values relative to the base case value is provided to allow evaluation of the relative impact.

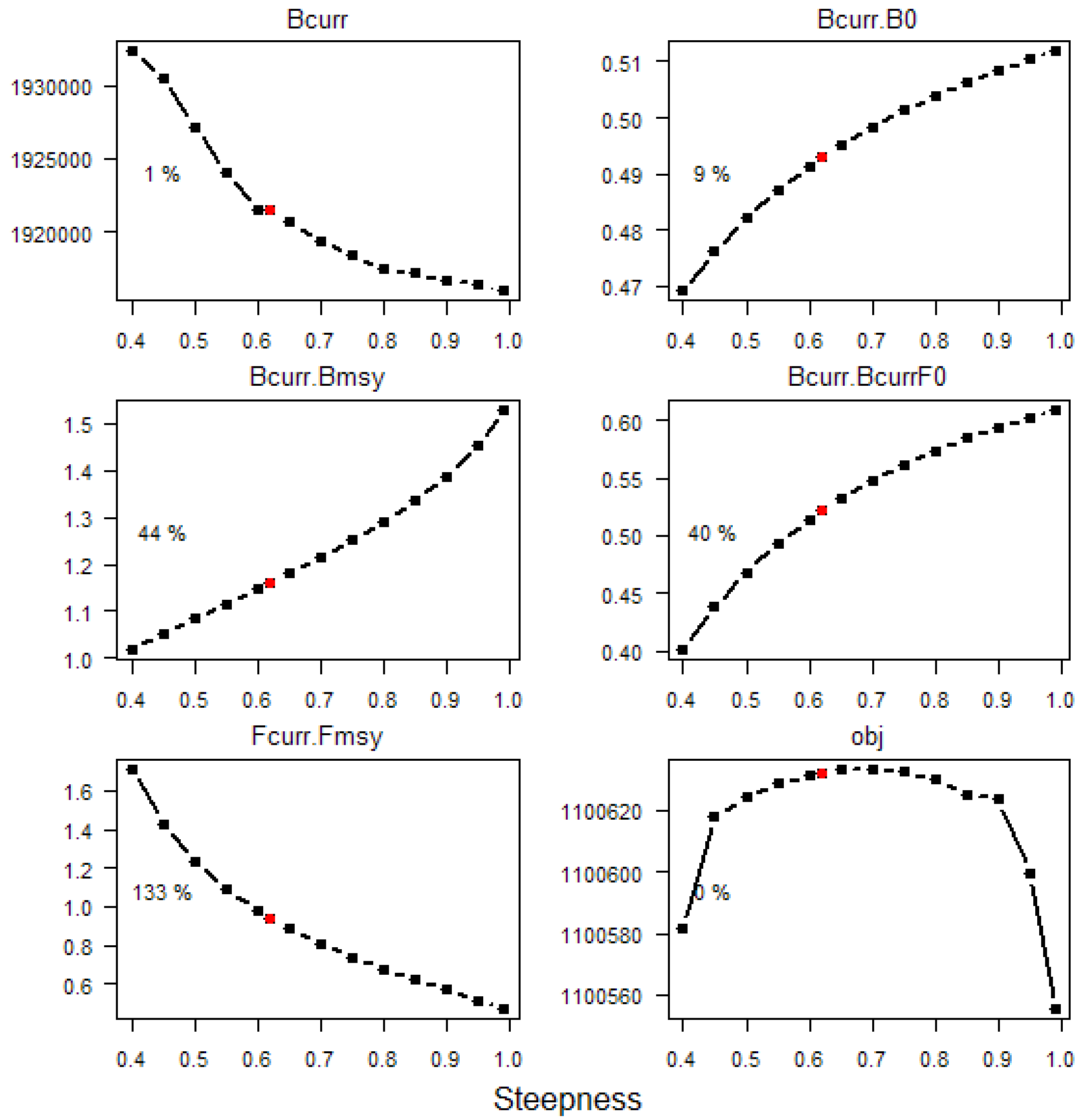


Figure 9: Impact of the assumed value of steepness on important assessment and management quantities. The red circle indicates the value for the base case model of the YFT SSA (Harley et al. 2009a) (steepness = 0.62). The percentage change in each quantity over the range of steepness values relative to the base case value is provided to allow evaluation of the relative impact. 'obj' is the objection function or measure of fit of the model run.

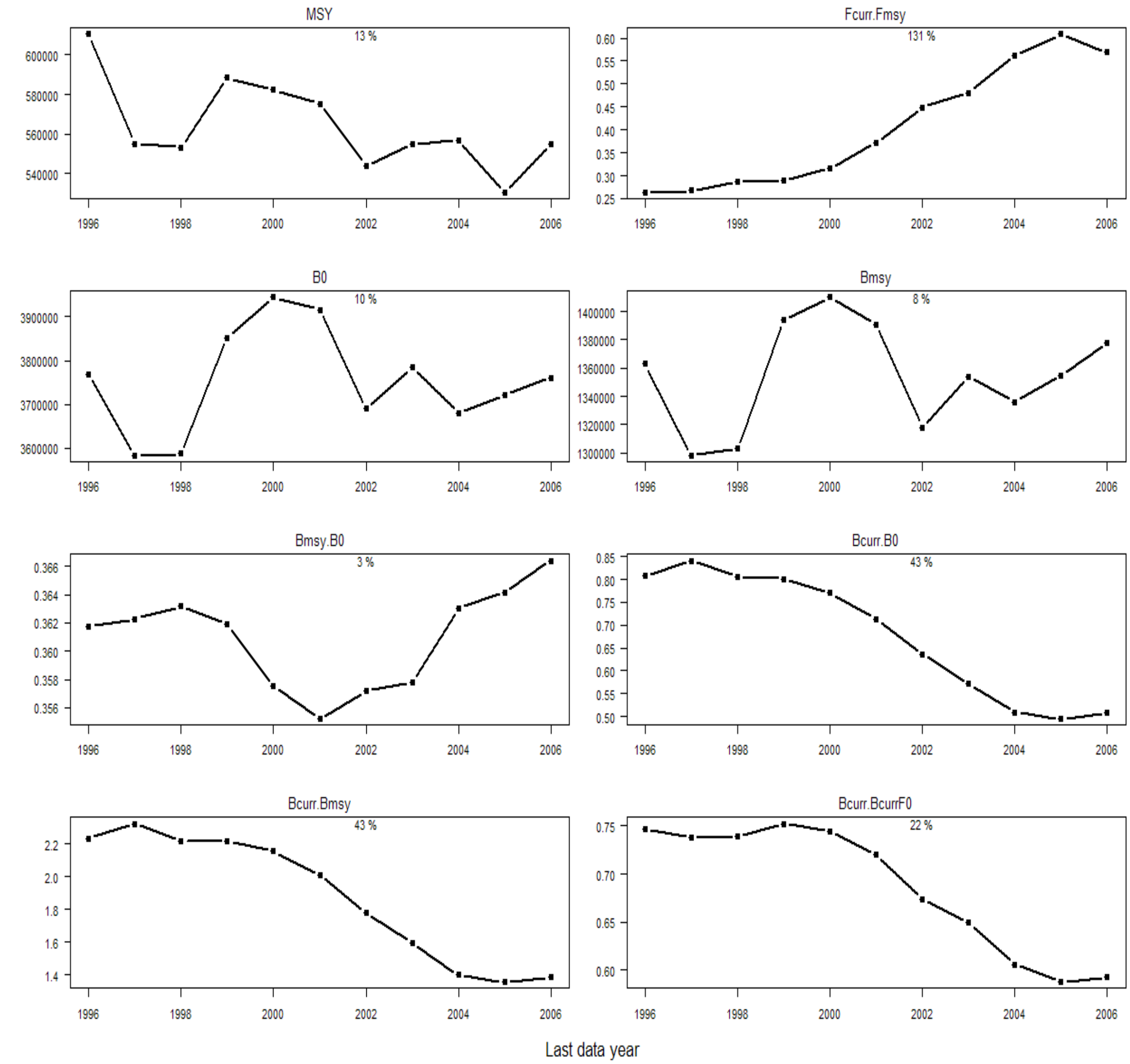


Figure 10: Estimates of important assessment and management quantities from the retrospective analysis. The percentage change in each quantity over the model runs relative to the value estimated when all data is used is provided to allow evaluation of the relative change.

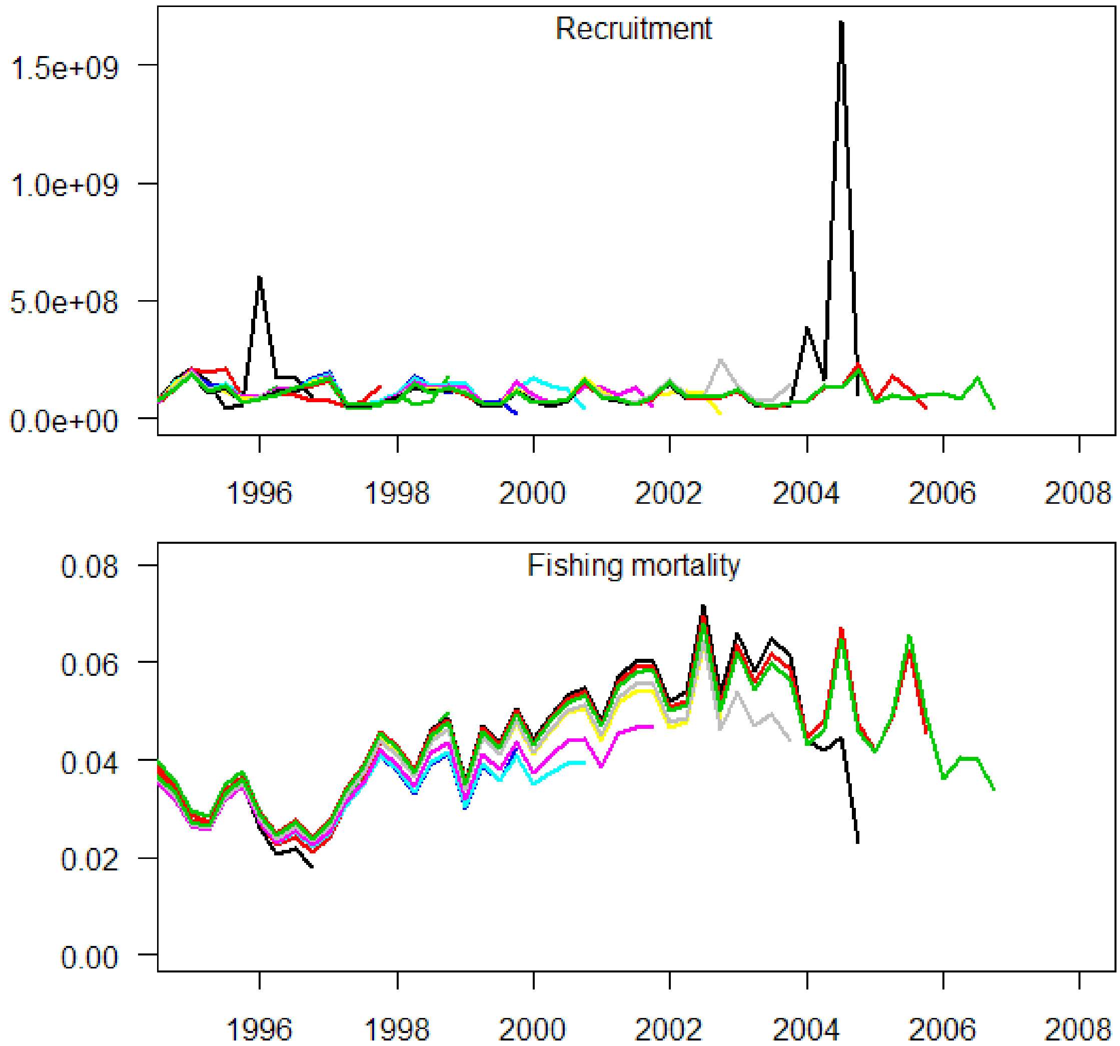


Figure 11: Trajectories for recruitment (top) and fishing mortality (bottom) for the retrospective analysis. Each line represents a models with different last data years.

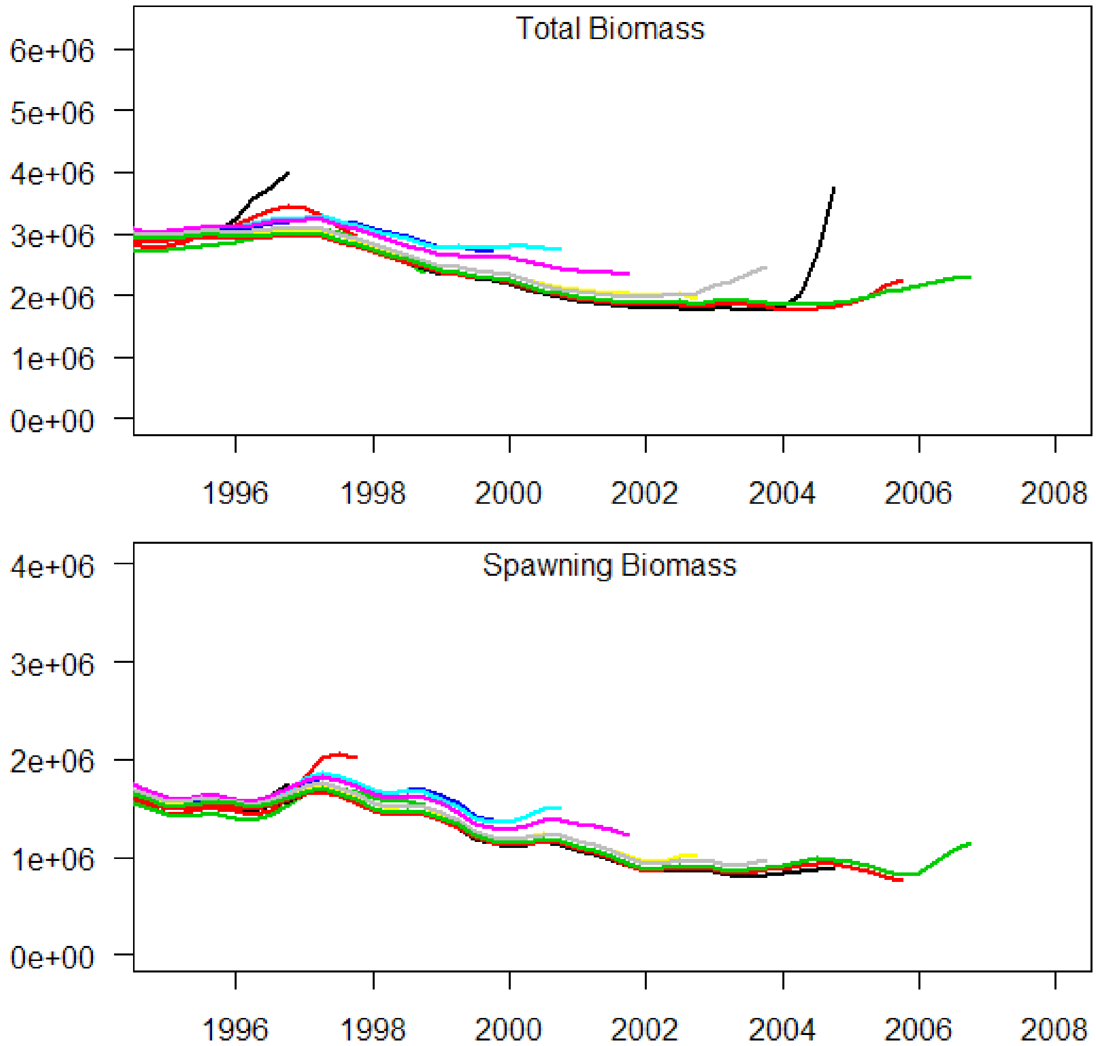


Figure 12: Trajectories for total biomass (top) and spawning biomass (bottom) for the retrospective analysis. Each line represents a models with different last data years.