

Exploration of Indian Ocean Bigeye Tuna Stock Assessment Sensitivities 1952-2008 using Stock Synthesis (Updated to include 2009)

Dale Kolody (Dale.Kolody@IOTC.org)

Miguel Herrera

Julien Million

IOTC Secretariat
PO Box 1011
Victoria, Seychelles

Summary

Note: the main text of this document describes preliminary work that was distributed prior to the IOTC WPTT 2010. Attachment 1 and 2 describe work that was undertaken and presented during the WPTT 2010. Methods are essentially the same, but the attachment results supersede the main text.

A stock assessment of the Indian Ocean bigeye tuna (*Thunnus obesus*, BET) population 1952-2008 is presented using Stock Synthesis (SS3) software. BET was judged to be the lowest priority of the main species of the IOTC Working Party on Tropical Tunas (WPTT) for 2010, and due to overlapping time commitments, this analysis was conducted prior to the usual data exchange process, and does not include the most recent data. This work extends the analysis described in Shono *et al.* 2009 (IOTC-2009-WPTT-20) in three substantial ways: *i*) the non-longline fisheries have been disaggregated into 3 fleets, *ii*) the Regional Tuna Tagging Programme – Indian Ocean (RTTP-IO) data are included, and *iii*) there is a fairly extensive exploration of alternative model assumptions. The model is presented primarily as an exploratory tool with which the WPTT can debate the best approach for future BET assessments and priorities among species. Core assumptions in all models included:

- Spatially-aggregated, age-structured, sex-aggregated population, iterated on a quarterly time-step 1952-2008
 - Four fishing fleets:
 - LL - combined longline (primarily Japan and Taiwan)
 - PSFS - unassociated Purse Seine (PS) sets in the western equatorial region
 - PSLS - FAD/log associated PS sets in the western equatorial
 - Other - includes PS outside the core area plus all other non-longline fleets.
-

- Beverton-Holt stock-recruitment dynamics, with fixed steepness and spawning biomass proportional to the total mass of mature fish. Annual recruitment deviates were estimated from 1985-2007 (in some cases).
- Mean length-at-age and weight-at-age relationships were adopted from IOTC-2009-WPTT-20 (constant over time)
- Maturity was adopted from IOTC-2009-WPTT-20 (constant over time with 50% mature at ~age 2).
- Age-based selectivity was estimated for each fleet independently, with independent parameters for each age (or group of consecutive ages) to admit the possibility of logistic, dome-shaped or polymodal functions.
- The RTTP-IO data have been included, with recaptures up to the end of 2008, including point estimates for the reporting rates derived from a tag seeding experiment on-board the European/Seychelles purse-seine fleet. A number of shortcuts were taken with the initial treatment of the tagging data. These shortcuts were expected to have a trivial effect, and to be resolved when adopting the new data. However, it is apparent that the errors are not trivial and are biased toward an underestimation of fishing mortality. (These issues are addressed in the revised document attachment 1).
- Objective function terms included the likelihoods for the fit to the standardized CPUE from the Japanese LL fleet (1960-2008), Catch-at-Length from all fleets, tag recoveries from the PSLs fleet, and priors on all estimated parameters. Estimated parameters included: virgin recruitment, CPUE catchability, independent selectivity by fleet (with temporal variability in some cases), and recruitment deviates (in some cases).

Additional important assumptions are listed in the model grid definitions below. Initial modelling efforts indicated that the size composition in the non-longline fleets could not be well fit, and there were potentially important stock status implications associated with different, somewhat arbitrary, assumptions. To better understand these sensitivities, a systematic exploration of the interactions among 8 different sets of assumptions was undertaken. Grid B included 54 models to look at the interactions among length-at-age, selectivity, size composition sampling and the influence of the tag data:

- 2 Length-at-age variance options (CV low, high)
- 3 Tag recovery negative binomial overdispersion options ($\tau = 2, 20, 200$)
- 3 PSLs selectivity assumptions (stationary, annual deviates for tagging years 2005-9, or annual deviates 1985-2008)
- 1 recruitment deviate option, $sd(\log(\text{devs})) = 0.6$
- 1 CPUE series option (no catchability trend)
- 1 steepness option ($h = 0.75$)
- 1 M vector $M(\text{age } 2y+) = 0.4$
- 3 Catch-at-Length likelihood weights (LL high+PSLs high, LL high+PSLs low, LL low+PSLs low)

Most of these options could be dismissed as either *i*) not making much difference to the stock status inferences, or *ii*) representing a strong and influential assumption that could not be justified from the available data. Only multiple levels of the tag recovery overdispersion options were carried into

the next level of the analysis. Grid C, with 108 models, examined the interactions among a number of other fundamental life history and data assumptions:

- 1 Length-at-age option (CV high)
- 3 Tag recovery negative binomial overdispersion options (2, 20, 200)
- 1 PLS selectivity assumptions (stationary)
- 2 recruitment deviate option ($sd(\log(\text{devs})) = 0.6, 0$)
- 2 CPUE options (no catchability trend, catchability trend of 0.47%/y)
- 3 steepness option ($h = 0.55, 0.75, 0.95$)
- 3 M vectors: $M(\text{age } 2y+) = 0.32, 0.4, 0.48$
- 1 Catch-at-Length likelihood weight (LL low+PSLS low)

The maximum posterior Density (MPD) estimates from these models indicate a broad range of uncertainty, including many plausible scenarios in which the B_{MSY} and F_{MSY} reference points were exceeded. The more pessimistic interpretations were generally associated with deterministic recruitment, effort creep, lower stock-recruit steepness, and reduced influence of the tagging data. In this case, the data were not expected to be very informative with respect to the key model selection decisions required to reduce the uncertainty among models. A somewhat subjective (but transparent) scheme was devised with which to weight the 108 models and provide stock status summaries based on the aggregate results. It is recognized that other members of the WPTT may have additional insight into the plausibility of the different scenarios, and this decision matrix can be easily updated at the discretion of the WPTT. For 2008, $B/B_{MSY} = 1.05$ (weighted mean of MPD estimates), with a range of 0.64 - 1.69 (most extreme MPD estimates). Similarly $F_{2008}/F_{MSY} = 0.91$ (0.40, 1.78). The central tendency of these results is similar to those presented in the WPTT 2009 report. The weighting scheme was also applied to deterministic constant catch projection results (catches at 60%, 80%, 100%, 120% and 140% of 2008 levels) in the format of a Kobe-2 Strategy Matrix (for the projected stock status in years 2011 and 2018). A number of recommendations for future assessments are discussed.

TABLE OF CONTENTS

SUMMARY	1
1. INTRODUCTION	5
<i>Fishery History</i>	5
2. METHODS.....	5
<i>Temporal units</i>	6
<i>Spatial Structure</i>	6
<i>Fleet definitions</i>	6
<i>Total catch data</i>	7
<i>Japanese CPUE as a relative abundance index and catchability assumptions</i>	7
<i>Size Composition Data</i>	8
<i>Selectivity</i>	9
<i>Stock Recruitment</i>	10
<i>Tag Releases, Recaptures and Reporting Rates</i>	10
<i>Growth, Natural Mortality and Maturity</i>	12
<i>Software</i>	13
<i>Model Specifications</i>	13
<i>Uncertainty Quantification</i>	14
<i>Projections</i>	15
3. RESULTS AND DISCUSSION.....	15
<i>Reference Case Model</i>	15
<i>Uncertainty Grid C – What are the implications of the key data and life history assumptions and how do they interact?</i>	19
4. CONCLUSIONS.....	23
5. ACKNOWLEDGEMENTS.....	24
6. REFERENCES	25
ATTACHMENT 1: UPDATE OF THE INDIAN OCEAN BIGEYE TUNA STOCK ASSESSMENT, INCLUDING THE 2009 DATA, AND REVISED GROWTH AND TAG RECOVERY ASSUMPTIONS	59
EXTENDED SUMMARY	59
<i>Revision of age-length relationship</i>	60
<i>Revision of tag data</i>	60
<i>Revisiting the interactions among the size sampling, selectivity and tag overdispersion assumptions</i>	61
<i>The Updated Stock Status Grid</i>	63
<i>Example models</i>	64
<i>Priorities for the future</i>	65

1. Introduction

Recent Indian Ocean bigeye tuna (*Thunnus obesus*, BET) assessments have suggested that the stock is not likely to be in an overfished state ($B > B_{MSY}$), and overfishing is probably not occurring ($F < F_{MSY}$) at this time (WPTT 2009). However, in 2009, the two assessments which provided confidence intervals recognized that there was some probability that the stock could be in an overfished state, and overfishing could be occurring. Having been assessed last year, and in the absence of strong stock status concerns, the Working Party on Tropical Tuna (WPTT) considered BET to be a lower priority than YFT and SKJ for stock assessment in 2010. Accordingly, this assessment was conducted in advance of the normal data exchange process, and the most recent data were not included. However, this analysis does extend the 2009 analysis in a number of directions:

- The WPTT tagging data was included (this was reportedly attempted previously, but the specific problems encountered were not described in Shono *et al.*, 2009),
- The effects of potentially increasing efficiency in Japanese longliners was considered, based on estimates from the operational data in the Pacific fleet (Hoyle, 2009),
- Sensitivity to uncertainty in several poorly quantified assumptions (and interactions among assumptions) is illustrated.

While the results can be interpreted as a stock assessment, the analysis is presented primarily to help the WPTT prioritize future research and assessment needs for the BET stock relative to the other species. The stock status inferences are broadly consistent with those described in the WPTT 2009. However, a key feature of this work is the illustration that the stock status inferences are sensitive to several subjective model assumptions and it is expected that the broader experience of the WPTT might help to constrain this uncertainty. However, the more important result might be the initiation of a discussion about how best to *i)* reduce the uncertainty in future assessments, and *ii)* provide useful management advice given that a large degree of uncertainty is probably inevitable.

Fishery History

Figure 1 illustrates the spatial structure that has been discussed in relation to BET fisheries in recent years. The catch histories are broken up by area and gear in Figure 2 to Figure 5. Artisanal fisheries (*e.g.* Maldives) have caught small numbers of BET for centuries, while large catches have only been recorded since the industrial Japanese longline fishery began operating in the Indian Ocean in the early 1950s. The Japanese catches have fluctuated considerably, though recent catches are similar to those observed in the 1950s. Longliners have consistently represented the dominant gear, with the Taiwanese fleet taking the largest catches (in mass) since greatly expanding operations in the 1980s and 1990s. The bulk of the catches have been taken from the western equatorial region, though the longline fishery has a broad distribution throughout most of the Indian Ocean. Purse Seiners started operating in the 1980s (primarily in the western equatorial region), however, the reported catches from this fleet, combined with the artisanal fleets, represent less than 20% of the total catch in mass. Total catch has declined by 30-40% since the peak observed in the late 1990s.

2. Methods

Fisheries data, research data, and model assumptions are described under the following headings. For continuity of the arguments, related data and model assumptions are described together. The SS3 control file for the reference case model is appended to resolve any incomplete or ambiguous descriptions.

Note: the original Control.SS file was replaced by attachment 3, which applies to the updated analysis in attachment 2.

Temporal units

Data were disaggregated by quarter (quarter 1 = Jan-Mar), and the model was iterated on quarterly time-steps, to represent potentially important seasonal dynamics, over the period 1952-2008 (plus 10 years of projections). However, the model may not have been configured to most appropriately describe seasonal processes. The fish population was given a plus-group accumulator age of 15 years.

As implemented in these models, SS3 has limited freedom to resolve variable recruitment by quarter (*i.e.* seasonal recruitment is possible, but year|quarter interactions cannot be estimated). This limitation could be overcome by restructuring the model (and various time-dependent functions) with quarters defined as years.

The importance of seasonally variable selectivity has not been examined.

Spatial Structure

The population dynamics are spatially aggregated, however, the 10 areas shown in Figure 1, have been used to define fisheries with consistent operational characteristics (*e.g.* as discussed further in relation to fleet definitions, CPUE, selectivity assumptions and tag releases/recaptures).

There remains an open question of the appropriate spatial structure to use for this tuna population (and most others). Qualitatively, the tagging data suggest that BET migrate reasonably quickly, and indicate at least some basin-scale movements (*e.g.* Figure 7). Unfortunately, the limited distribution of tag releases, and small number of returns (and absence of tag reporting rate estimates) outside of the European/Seychelles purse seine fleets, mainly operating in the western equatorial Indian Ocean, makes it difficult to quantify large-scale movements at this time. While there may be some relatively discreet sub-populations, or slow mixing rates among sub-regions, there is no evidence that this is the case in the core area where most of the catch is taken, and presumably where the bulk of the population is located.

It is noted that recent yellowfin (YFT) tuna assessments in the Indian Ocean have adopted a 5 region spatial structure (Langley *et al.*, 2009). It is not clear that this has resulted in an improved understanding of the YFT stock dynamics, as the estimated movement rates show some counterintuitive patterns that likely reflect the unbalanced distribution of tag releases and recoveries (and inability of the software to resolve inter-annual movement variability). Comparison of the disaggregated YFT model with a similarly parameterized aggregated model did not reveal substantive differences in the overall stock status estimates (WPTT 2009).

The BET spatial structure should be revisited in the future, in relation to more detailed analyses of the tagging data, and space/time patterns in the size composition and CPUE data.

Fleet definitions

Four fleets were defined on the basis of gear type and area, with data from all nationalities combined:

- 1) LL –longline fleets (primarily Japanese and Taiwanese) aggregated across all areas

- 2) PSFS – Purse Seine Free School sets (not associated with drifting objects), only in area 3.
- 3) PSLS – PS sets associated with FADs or logs, only in areas 1+3.
- 4) Other – all other gear (*i.e.* artisanal gear and PS sets of unknown type or location outside the areas defined for PSFS and PSLS).

Shono *et al.* (2009) used 2 fleets, with 2-4 above aggregated into a single fleet. We opted to split the non-longline fleets in an attempt to identify purse seine fisheries with homogeneous characteristics. The *a priori* expectation was that this would improve the validity of the stationary selectivity assumption (for the PSFS and PSLS fleets), which would in turn improve the estimation of relative year class strength, improve the consistency between the catch-at-size and tag recoveries, and allow the external tag reporting rate estimates to be applied directly to these fleets. It was expected that fishery 4 would represent a relatively small heterogeneous mix of fisheries, such that the size composition data could be considered uninformative and greatly down-weighted. As discussed subsequently, these efforts have not fully achieved the desired results.

Total catch data

The total catches were calculated by the Secretariat (Herrera, 2009). This is a complicated process that requires a number of approximations and substitutions for fleets with poor data (including those discussed below under size composition data). The catch time series for the 4 fleet structure is shown in Figure 6.

Catch in mass was used in the model for all fleets, and was assumed to be known essentially without error ($SD(\log(\text{catch in mass error})) = 0.01$; this is not actually used in the SS3 likelihood, but might reflect a tolerance in the iterative solving of the catch equations).

Japanese CPUE as a relative abundance index and catchability assumptions

Okamoto *et al.* (2009) describe the standardized Japanese CPUE series adopted as the relative abundance index for the period 1960-2008. Several differences from Shono *et al.* (2009) were applied here:

- A single annual CPUE index based on the mean of the 4 quarters was used (assigned to quarter 1 in the model dynamics). This was done purely to simplify the data handling and interpretation of graphical results (qualitatively, the 4 quarters appeared to be consistently proportional throughout the time series)
- A constant CV of 5% was adopted. Shono *et al.* (2009) used a unique CV for each data point derived from the CPUE analyses. These estimates were generally less than a CV of 5%, and considering that there were 4 (very similar) CPUE series, this effectively corresponds to a much higher CPUE weighting than this analysis.
- An additional CPUE scenario was admitted in which it was assumed that catchability increases by 0.47% / year (arithmetically, *i.e.* this corresponds to a total efficiency increase from 1952-2008 of 26.3%). This value is adopted from Hoyle *et al.* (2009), which estimated a vessel-technology effect using operational level data from the Pacific, which could not be identified using the 5°x5° aggregated CPUE data. It is noted that this increasing efficiency trend only accounts for one possible source of increasing efficiency, and it would not be surprising if the real effort creep was higher (*e.g.* sensitivity analyses in the Pacific regularly include catchability trends of 2%/year).

The two CPUE series are compared in Figure 10. These two catchability assumptions are referred to in the model grids as:

- EC0 = constant catchability
- EC47 = increasing catchability.

We do not actually believe that the CV of 5% is realistic. We are essentially saying that the model is conditioned to the relative abundance index. Preliminary exploration with a relaxed CPUE CV (10-20%) resulted in temporal trends in abundance that were dubious if one believes the CPUE series (*i.e.* the model often estimated a very steep CPUE decline over the last few years, while the observations are stable). Examining the way in which the other data sources conflict with the CPUE series is useful for understanding the model, however, in general, we would not have much confidence in any stock status inferences from models that fail to fit the core features of the relative abundance index.

Catchability for the LL CPUE was assumed to be stationary in the models. Note that, in general, it is not possible to meaningfully compare likelihood terms for models that use different sources of data. However, in this case, one might consider the two CPUE time series to be the same data, with an additional effort creep parameter imposed, and the likelihood does potentially provide a means for considering which CPUE option is more consistent with the model assumptions and other data sources. However, as discussed elsewhere, there are still reasons why this comparison is probably not a good idea.

Size Composition Data

The catch-at-length data were compiled by the secretariat (Herrera, 2009). This process involves a number of approximations, substitutions and unit conversions because some fleets have very poor data, some fleets do not report data at the appropriate resolution, and different measurement procedures are often used (*e.g.* mass vs: length). These data were generated as a 'best guess' to produce a complete time series for all fleets (*e.g.* which might be a requirement for data intensive VPA-type assessments). There are many extrapolations in time and space (*i.e.* assuming that unsampled strata are the same as the nearest strata in space or time that was sampled). Actual sample sizes are unknown in many cases, and the representativeness of sampling across vast areas is often questionable. The secretariat produced another type of size composition file for BET in 2010, which provides an estimate of the sample sizes (and omits strata without samples), but this was not available in time for this analysis.

Catch-at-length distributions aggregated over time, and time series of mean length are shown by fleet in Figure 11. The PSFS and Other fleets have bimodal distributions. In the PSFS fleet in particular, the mode shifts between consecutive quarters, in a pattern which is not obviously seasonal, nor evident in the other fleets. This is believed to be largely due to limitations in the definition and recording of the set-type, as there are reportedly two very different types of free school sets (mixed schools with small BET, YFT and SKJ, and mixed schools of large YFT, BET and ALB). At this time, it is not possible to further partition the fisheries by free-school set-type, but the model can attempt to accommodate this bimodality as either a sampling problem (*i.e.* down-weight the size data) or temporally variable selectivity.

The model is potentially sensitive to the size composition data assumptions. In this assessment, all length composition strata from all fleets were assigned a default sample size of 200. It is not a good idea to place too much confidence in these data because of the problems mentioned above (sampling problems and non-stationary selectivity), and also because the model does not resolve seasonal recruitment variability, so length-at-age of younger cohorts can be poorly represented. The sample size assumption was further reduced via a downweighting co-efficient (λ) of 0.1 in the objective function at all times for PSFS and Other fleets, and in some cases for LL and PSLs. The downweighting is similar to assuming a sample size of 20. Different input sample size assumptions are examined in the model grids:

- CL1010 : $\lambda_{LL} = 1.0, \lambda_{PSLS} = 1.0, \lambda_{PSFS} = 0.1, \lambda_{Other} = 0.1$
- CL1001 : $\lambda_{LL} = 1.0, \lambda_{PSLS} = 0.1, \lambda_{PSFS} = 0.1, \lambda_{Other} = 0.1$
- CL1010 : $\lambda_{LL} = 0.1, \lambda_{PSLS} = 0.1, \lambda_{PSFS} = 0.1, \lambda_{Other} = 0.1$

The catch-at-size distributions are represented in 103 bins of length 2cm. The multinomial likelihood was used in the model, with an additional 0.1% added to each length bin (predicted and observed) to make the term more robust to outliers.

Selectivity

A number of different approaches were investigated to address the problems with the PS selectivity stationarity and size composition sampling assumptions. Initially, a double-normal pseudo-length-based function was estimated for each non-longline fleet. The estimated fishing mortality (*i.e.* quarterly value of the most highly selected age classes) in the non-longline fisheries was unbelievably high for all of the models that used the pseudo-length-based selectivity (this problem is also evident in Shono *et al.* 2009). This likely happened because either *i*) the model attempts to remove a large number of weakly selected cohorts or *ii*) the model attempts to remove a very large number of small fish while the size composition data suggests that the actual catch consisted of a much smaller number of larger fish. Despite reasonable convergence (as defined by the maximum gradient criteria), it also appeared the length-based selectivity models were prone to identify local minima (in the cases examined).

Age-based selectivity resolved the high F and convergence problems (though did not resolve the poor fit to the non-longline size composition data). One parameter for each age (or group of consecutive ages) was estimated as a free parameter (7 for the LL fleet and 6 for the others).

Different options for temporal variability in selectivity were also explored and there were convergence problems with this approach unless informative priors were used.

Three age-based selectivity options were considered in the model uncertainty grids:

- sc: stationary selectivity for all fleets
- st3: annual selectivity deviates estimated for PSLs years 2005-2008, the years with tagging data (normal prior, mode 0, CV 0.6)
- st23: annual selectivity deviates estimated for PSLs 1985-2008 (normal prior, mode 0, CV 0.6) .

Stock Recruitment

A Beverton-Holt stock recruitment relationship was assumed, with fixed steepness (0.55, 0.75 or 0.95). It was assumed that spawning biomass is equal to the mass of the mature females (*i.e.* no disproportionate adjustment for size-dependent fecundity). The different steepness assumptions are defined in the model grids:

- h55: $h = 0.55$
- h75: $h = 0.75$
- h95: $h = 0.95$

Deviations from the stock-recruitment relationship were assumed to follow a lognormal distribution, and were only estimated from 1985-2007 (the period in which there may be informative PS size composition data). Deviations prior to 1985 could be used to explain CPUE variability, but there was no strong requirement for additional freedom to explain the data during the early period of relatively light exploitation. The only exception might be the few extreme CPUE observations in the late 1970s, but, assuming that they are credible, these reflect a transient event that should not have much influence on the long term dynamics.

There was a declining trend in the recruitment deviations over time for all models examined, such that the models interpret declining CPUE as a mix of depletion and recruitment trends (also illustrated in Shono *et al.* 2009). If the recruitment trend is actually an artefact, then the model will be over-optimistic with respect to MSY-related reference points. Given the problems with the fit to the size composition data, it would be difficult to have much confidence in the relative cohort strength estimates derived from these data. However, down-weighting the size composition data, and replacing all of the non-longline size composition data by the long-term mean size distributions also resulted in recruitment deviation trends. This suggests that at least part of the problem is derived from the conflict between the stock recruitment relationship and the CPUE series. The alternative approach was explored in which models were run without the estimation of any recruitment deviations. This is obviously unrealistic, but may provide a more robust estimator of productivity in this case (*i.e.* this formulation resembles a surplus production model that ignores the potentially misleading vagaries of the size composition data). The recruitment deviation CV assumptions are defined in the model grids as:

- r0 : $SD(\log(\text{dev})) = 0$
- r6 : $SD(\log(\text{dev})) = 0.6$

The $-0.5\sigma^2$ 'bias correction' to the stock recruitment relationship was differentially applied, such that two time periods, one with estimated deviations and one without, have the same expected mean recruitment. This was not ideal in this case, because the majority of models with estimated recruitment had a RMSE lower than the assumed value of 0.6. Hence the mean recruitment during the estimation period was always lower than expectation. To remove this effect, one can estimate σ , however, in the cases in which this was done, σ approached 0 (due to the low weighting on the size composition data), and essentially this was redundant with the r0 option.

Tag Releases, Recaptures and Reporting Rates

The RTTP-IO tag release (2005-7) and recapture (2005-8) data are described in Hallier and Million (2009). Figure 7 shows the spatial distribution of all releases and recoveries up to 2008. The only

reporting rate estimates available are derived from tag seeding experiments in the European/Seychelles purse seine fisheries (Gaertner and Hallier 2009), from which the large majority of the recoveries were reported, so these fleets should be the most useful for interpreting the tagging data. Data were processed by the IOTC Secretariat.

In the models reported here, only the PSLS fleet recoveries were included in the model. As noted previously, the PSFS fishery seemed to be a heterogeneous mix of sets with different selectivity. This renders the stationary selectivity assumption dubious for this fleet (and it only accounts for a small amount of catch as shown in Figure 6), and was expected to cause problems with the tag recoveries. The LL and Other fleets have very few tag returns, and poor estimates of reporting rates, such that they were not expected to be very informative (recoveries and reporting rates set to 0). Initial explorations that included estimation of stationary reporting rates for LL and Other fleets were examined, but not pursued in detail.

In the population model, tagged fish are assumed to have identical dynamics to the general population. For this to be true, the tags should first fully mix with the general population. It was assumed that full mixing was achieved 4 quarters after release. However, for various reasons, mixing and reporting assumptions tend to not conform to the ideal situation and tag recoveries tend to contain greater error than expected. The negative binomial distribution allows for overdispersion relative to the ideal (*e.g.* Poisson distribution). Three options were explored for the overdispersion parameter τ (applied equally across all tag groups):

- t2 : Overdispersion parameter = 2 (close to ideal Poisson tag recovery assumptions)
- t20 : Overdispersion = 20 (intermediate value adopted in Pacific BET assessments)
- t200 : Overdispersion = 200 (tags highly downweighted).

The probability density functions for the 3 options are illustrated in Figure 12 for a mean of 100 recoveries. The ADMB `log_negbinomial_density` function is parameterized in terms of $\tau = \sigma^2/\mu$; $\tau > 1$, and this is equivalent to the R function `dnbinom(x, size, mu)` where $\sigma^2 = \mu + \mu^2/\text{size}$, $pr = \text{size}/(\text{size} + \mu)$, such that for $\mu=100$, $\tau = (2,20,200) \sim \text{size} = (100,5,0.5)$.

The length of release of each tag is recorded in the database, but the model dynamics are based on ages. The (annual) age of each individual tag was estimated based on the mean growth curve, assuming a 1 January birthdate, and an independent tag release event was tracked for each age/year/quarter release strata (24 release events in total). This age/length processing is external to the model, and the discrete age assignment means that growth should not be estimated within the model when tags are used.

Recovered tags that were not assigned a release length were removed from the release and recovery datasets. A total of 34565 tag releases were included.

- Recovered tags which could not be assigned to a fleet were removed from the recovery dataset. These were primarily PS tags from which it was impossible to determine if they were PSFS or PSLS sets (*i.e.* set-type unknown and/or date of recovery unknown and/or area unknown). It was initially assumed that this was a small number, but actually accounts for over 30% of recoveries.

- Tag shedding was assumed to be 0. However, the instantaneous (annual) shedding rate estimated by Gaertner and Hallier (2008) was 0.024 (0.014 – 0.30 95% CI).
- The tag reporting rate estimates for the European/Seychelles PS fleet were applied to the PSFS and PSLs fleets, however, some small portion (expected to be <10%) of the PSFS and PSLs fleet catch composition came from the non-European/Seychelles PS fleets, *i.e.* PS fleets from Thailand and Iran, (for which there is no reporting rate estimate, and it is presumably much lower than the European/Seychelles fleet).
- For convenience of interpretation of graphical software, the recoveries of tags from the non-mixing period were removed from the recovery dataset. This introduces an additional error to the treatment of the tags because the recovered tags from the non-mixing period are assumed not to be caught and therefore not removed from the population. For a particular release event, this has no effect on the fishing mortality estimated within the non-mixing period, but the magnitude of the error in subsequent time periods will depend on the number of recovered tags in the non-mixing period relative to the total number of tags at liberty. For the largest release event (7619 tags) the number of recovered tags in the (4 quarter) pre-mixed period in the PSLs fishery is 974 (13%) that were not removed from the tag population (after including adjustment for reporting rates).

These errors are not trivial, and they are all biased in the direction that will lead to an underestimate of fishing mortality. This needs to be addressed before the tagging results can be considered meaningful.

SS3 cannot admit temporal variability in tag reporting rates at this time. Instead, the number of reported tags is inflated by a factor of $1/(\text{reporting rate at time } t)$, and the reporting rate is fixed at 100% for the informative fleet(s).

For the uninformative fleets (everything other than PSLs for the results presented here in detail), tag recoveries and reporting rates were set to 0.

Note that the tag implementation, especially with seasonal dynamics, is a relatively new and poorly documented feature within SS3. It is possible that there may be errors in the software, or current implementation, that adds to the exploratory nature of this assessment. However, the initial results seemed to be largely consistent with expectations (this downplayed the urgency in resolving the remaining the tag implementation problems).

Growth, Natural Mortality and Maturity

Mean length-at-age was adopted from Shono *et al.* (2009) and assumed to be constant over time. The adoption of a von Bertalanffy growth curve does not adequately describe the sigmoid shape that has been estimated for this species in the Indian Ocean (*e.g.* Eveson 2008). This is currently a limitation in SS3 software, but probably not overly influential given the other problems with the juvenile size composition. Two different options for the variance of the length-at-age distribution were explored in the model uncertainty grids (shown in Figure 8):

- gr05: lower variance option (adopted from Shono *et al.* 2009)
- gr20: higher variance option.

The lower variance was judged to be less attractive than the higher variance, because it is potentially very restrictive in its assumptions about recruitment timing and within season growth (both of which would inflate the variance, particularly for the youngest fish). The high variance option is not derived from any specific analysis, and the gr20 option is obviously questionable in so far as the variance in length at age actually declines between intermediate and older ages. The uncertainty exploration suggests that there can be non-trivial implications for this assumption, but they are probably not important if one assumes that juvenile fishery size composition is only weakly informative.

The baseline M vector and maturity schedule were derived from Shono *et al.* (2009), both constant over time (Figure 9). Three alternative M vectors were explored in the uncertainty grid in recognition that this term is often poorly quantified.

- M08: M for all ages 20% lower than Figure 9
- M10: M as shown in Figure 9
- M12: M for all ages 20% higher than Figure 9.

Exploratory attempts to estimate M were not very believable, in that the model consistently preferred lower M for younger ages.

Software

The model was implemented with Stock synthesis SS V3.10B. This is a powerful and flexible stock assessment package with efficient function minimization, implemented with AD Model Builder (<http://admb-project.org/>). Technical details are (mostly) described in Methot (2000, 2009). Typical function minimization required about 1 minute (up to 5 min. with non-stationary selectivity), on a 3.0 GHz PC (not including inverse Hessian calculations).

Model Specifications

The assessment is described in 4 stages:

- 1) Reference Case - Identify a baseline 'reference case' model that is plausibly consistent with the data (*i.e.* reasonable fit to the CPUE series, gross features of the size composition data and tagging data, and stationary recruitment assumptions). This involved an initial specification, detailed analysis of diagnostics, and single dimension deviations from the previous specification, until a satisfactory model could be identified.
- 2) Explore options for addressing the non-longline size composition problems - Using a haphazard exploration of models, we were not able to identify any specification with a good fit to the size composition data from the non-longline fleets. A systematic search was subsequently initiated, in which 54 models were defined to examine the implication of various assumptions related to size composition sampling assumptions, selectivity, length-at-age variance and the influence of the tag data (Grid B in Table 1). On the basis of these results, a minimum set of alternative specifications were sought which could be described as
 - Plausible (or at least no worse than the alternatives), and
 - Sufficiently representative of the uncertainty to be carried into stage 3.

- 3) Sensitivity analyses - a grid of 108 models was defined that considers the interactions among 4 different sets of assumptions (grid C in Table 1). We attempt to quantify the core assumptions and interactions that influence the model by examining the marginal effect of each assumption option on the fit to the data, and key stock status reference points.
- 4) Stock status estimates - The final stock status estimates are derived from a synthesis of the 108 sensitivity results of gridC. Key summary diagnostics are considered, and a somewhat subjective (but hopefully transparent) weighting scheme is adopted. The stock status summary consists of the weighted estimates of B_{2008}/B_{MSY} , F_{2008}/F_{MSY} , presented in a Kobe plot, and 3 and 10 year constant catch projections presented in a Kobe 2 Strategy Matrix (See *Uncertainty Quantification* and *Projections* below).

In general, models were compared on the basis of:

- CPUE RSME – describes the fit to the CPUE series. Ideally, this value should be very similar to the assumed SD of the CPUE observation errors. In most of the models discussed here, the CPUE RMSE in the period of interest is <0.1 , which indicates a good fit, and no justification for rejecting models. This is of course because the observation error was assumed to be very low in the first place, because models with a poor CPUE fit were not worth considering.
- ESS (Effective Sample Size) – describes the fit to the size composition data for each fleet (averaged over all observations). ESS indicates how well the predicted size composition data fits the observations (irrespective of the assumed weighting for that data). An ESS of 200 means that on average, the fit is as good as would be expected for a true random sample of 200 (regardless of what the actual sample size was). The ESS does not explicitly distinguish between random noise and systematic lack of fit (and it is the latter quality that we are usually most interested in). However, when used as a relative index to compare models fit to the same data set, lower ESS is usually associated with a higher systematic lack of fit.
- Recruitment trend – this measure describes the systematic lack of fit that arises when the recruitment deviates are estimated for this specific situation (the RMSE and auto-correlation would be of more general interest in most applications). The recruitment trend is defined here as the drop in recruitment (deviation) between years 1985 and 2007, as estimated from the slope of a linear regression through the recruitment dev time series. The units were chosen for ease of interpretation (*i.e.* to consider the question – should we really believe a model which estimates the average recruitment dev has had a 30% decline between 1985 and 2007, or is this indicative of a systematic problem in the model?).
- Likelihood terms –The likelihoods are useful for qualitative discussions of which options appear to be more compatible, *etc.*, but literal interpretation of likelihoods in these models will generally lead to some counter-intuitive results, and over-optimistic perceptions of precision (*e.g.* see below).

Uncertainty Quantification

The stock assessment process often appears to involve a haphazard search for one or a very few model specifications which appear to be plausibly consistent with the data, and a priori expectations. Most commonly, some form of statistical method is used to describe uncertainty distributions (*e.g.* likelihood profiles or Bayesian posteriors) for the quantities of interest under the

assumption that a particular model is ‘correct’. However, in this case, there are some fundamental problems with interpreting the likelihoods literally: *i)* the data are not the same (*i.e.* changing the input sample size between two models invalidates a direct comparison of the likelihoods), *ii)* these are complicated highly parameterized models with many assumptions that are poorly justified, usually with evidence of systematic failures in the model fit, that probably mean a strict interpretation of the likelihood is not justified, and *iii)* it is known from simulation studies, that some parameters cannot be estimated reliably with the type of data and observational contrast that are typically available (*e.g.* M, steepness).

The process used here takes the alternative approach of focussing on the model selection uncertainty, which is usually much greater than the statistical uncertainty conditional on any individual model. We only consider the Maximum Posterior Density (MPD) estimates, and stock status estimates are derived from a weighted average of the MPD estimates. This is similar to the approach used by the CCSBT (originally in the context of stock assessment, and subsequently in the development of operating models for Management Strategy Evaluation). Figure 43 illustrates the idea conceptually – once it is admitted that there is a range of plausible models (*e.g.* the two normal distributions in Figure 43), if the statistical uncertainty bounds are narrow relative to the difference in the central tendency, it seems sensible to admit that the real uncertainty spans the broader range between the models (*e.g.* the red box, or better yet, the red box plus the tails of the black distributions).

A comparison of the two approaches might be considered by an analogy of observing a large street mural at night. The first case is analogous to observing the part that happens to fall under the streetlamp. The second case is like walking around with a little flashlight. The view is never as impressive in the second case, but you are less likely to miss something important.

Projections

Projections were conducted from the MPD estimates of all models at catch levels of 60%, 80%, 100%, 120% and 140% of 2008 levels (assuming 2008 selectivity and catch allocations among fleets). The projections used deterministic recruitment from the stock recruitment relationship (starting in 2007). This approach ignores two important sources of uncertainty: statistical uncertainty in the parameter estimates, and recruitment variability (the latter of which cannot currently be accommodated within the SS3 software). However, as in the previous section, the approach does incorporate the model selection uncertainty, which is probably greater than both of these sources of uncertainty in most cases. Three and Ten year projection results are summarized in a management decision table (Kobe 2 Strategy Matrix).

3. Results and Discussion

Reference Case Model

A number of initial model specifications were explored to identify a sensible baseline reference case for the assessment. The unsuccessful attempts are not presented in detail, but a number of general problems are noted for the record in the methods and following discussion. The core assumptions from the selected ‘ref’ model are listed in Table 1, along with two similar models that illustrate alternative assumptions related to tag recoveries (reft002 and reft200). There is no preferential

status assigned to any of these models. Key features of the behaviour of these models are illustrated in Figure 13 - Figure 19.

Figure 13 illustrates the fit to the CPUE series (excellent for the period 1985-2008), and the PSLS tag recovery data. The ref model is contrasted with models that increase (reft002) and decrease (reft200) the influence of the tags. In all 3 cases, the CPUE provides an excellent fit. The seasonal pattern of the tag recovery data does not fit particularly well in any case, and, qualitatively, the different tag overdispersion assumptions do not seem to make a big difference to the quality of fit to the CPUE or tags.

Figure 17 illustrates the fit to the temporally-aggregated size distributions and time series of mean size for each fleet ref. In general, the LL fit was always good, and the non-longline fleets were always poor. The difference in fit between ref, reft002 and reft200 was almost imperceptible (not shown).

Figure 16 illustrates the 2 recurring troubles with the recruitment time series when the annual deviates are estimated. First, there is a downward trend from 1985-2007. The most recent recruitment (which is really not influenced by the size composition data) is always very low. The model uses a dev_vector assumption, in which the deviates sum to a mean of 0 and the flexibility of the weakly constrained cohorts is often exploited to balance large deviations in other years that are tightly constrained. SS3 provides the flexibility to anchor an arbitrary number of recent recruits to the stock recruitment function. If 1-2 additional years are constrained in this way, the exceptional deviate was simply shifted back 1-2 years (not shown). The second problem is the low variance estimated on the recruitment deviates. When applied with the bias correction on the mean, the recent recruits fall below the expected value. Both of these problems tend to suggest that recent productivity has been lower than average due to anomalous recruitment, and that in general the stock should be more productive than recent years. This interpretation might be overoptimistic (not estimating the recruitment deviations removes both of these problems, but it comes at a cost in terms of the fit to the CPUE series as discussed under *Uncertainty grid C*).

Figure 18 illustrates the selectivity estimates by fleet. It seems unlikely that the LL selectivity should have such a strong dome-shape, with very low selectivity after age 3. Perhaps there is a biological justification on the basis of spatial partitioning by age, but it is probably a more fundamental problem with the growth or M assumptions. All of the non-longline fleets show some form of bimodal selectivity. Given the obvious bimodal size composition distributions in PSFS and Other fleets, this is not really surprising, but the specific functions estimated are questionable. It is not clear that any single compromise selectivity function would be more appropriate. The question is really how much the selectivity functions affect the stock status inferences. There are two main effects of the selectivity: *i*) ensuring that the correct size/age composition is removed from the population, and *ii*) estimating relative year-class strength through the separable assumption. For the non-longline fisheries the first problem should not be an issue because these fisheries are small relative to the LL fishery. The second issue might cause poor estimation of some individual year classes (particularly the more recent years) if the selectivity changes are random, and potentially more serious long term biases if there is a trend in the selectivity changes. The interaction between the stationary selectivity assumption and the tagging data may also have important consequences.

Time series of spawning biomass and fishing mortality for ref, reft002 and reft200 are shown in Figure 19. Increasing influence of the tagging data tends to support a more optimistic interpretation of the stock status.

Uncertainty Grid B: How to handle the PS size composition data?

The reference case (and associated models), strongly suggested that there are considerable problems with the non-longline size composition data, and potentially the fit to the tagging data, that are worth further attention. Several options were considered to reduce the problem:

1. Improve the selectivity assumptions. Selectivity in these models reflects the combination of two very different concepts, gear selectivity and availability. Gear selectivity is a relatively straightforward concept that describes which of the fish exposed to the gear are captured (*e.g.* selective net mesh sizes). Availability is the much more complicated question of the spatial and temporal distribution of the fish and fleet (*i.e.* understanding gear selectivity is not very helpful if you do not actually know which fish are exposed to the gear). In most fisheries models the two concepts are inextricably confounded. For Indian Ocean BET, the assumption of stationary selectivity, particularly in the non-longline fleets is questionable for a number of reasons: *e.g.* *i)* some of the fleets do not operate consistently in space and time (*i.e.* in particular the 'Other' fleet is a heterogeneous mix of fisheries), and *ii)* PS fleets (in particular) target schools, and different schools have different size composition characteristics (*i.e.* if you always successfully target schools with a unique size/age composition, it will not tell you much about the relative abundance of fish of different sizes/ages in other schools). The PSFS fleet seems to have three different catch-at-size distributions: small fish, large fish and a bimodal mixture of the two (the latter probably represents the aggregation of the two set types). The observed size composition of the PSLS fleet seems to be more consistent than the other two, which presumably indicates the consistency of the size composition of BET found around floating objects. But it may be optimistic to expect that this results in stationary selectivity even for this most-homogenous of fleets. If stationary selectivity is incorrectly assumed, it could have undesired consequences on year-class strength, and may also constrain the fit to the tagging data. Two possibilities were considered:
 - a. Revise the fleet definitions into more heterogeneous units. This was the reasoning behind the disaggregation of the non-longline fleet in the first place, and it is not clear that further disaggregation will help the problem.
 - b. Explicitly model the temporal variability in selectivity. This is the most satisfactory approach if the data are reliable. The main downside is the added complexity and computation time for the model.
2. Improve the representation of the size sampling. Perhaps the selectivity is reasonably stationary, but the size sampling is poor. The sampling procedures are poorly quantified in these fisheries, such that it is currently unclear what the sample sizes are and whether they are random in space and time. Downweighting the size composition data to the point that it has a weak influence relative to the other data sources should remove some of the misleading artifacts of the size composition data. This might still result in unbelievable time

series of year-class strength, but the size composition data should not dominate when there are conflicts with the other data. This will also result in fishery removals that are not very consistent with the observed catch-at-size samples, but this should not make a big difference in this case, since most of the catch is taken by the longline fisheries.

3. Do not attempt to estimate the recruitment deviations. This is an extreme way of reducing the size composition influence on the year-class strength. This approach essentially reduces the model to a form of production model. The flexibility is greatly reduced, and the influence of some data sources is minimized, but may be more robust to some questionable data. The downside is that useful information may be lost from other sources (*i.e.* tags and CPUE). Option 3 was explored in the following section.

In an attempt to find the best approach for dealing with these potentially inter-related problems, Grid B was defined (Table 1), which consists of 54 models with different options for 4 sets of assumptions related to:

- 2 X variance on length-at-age,
- 3 X selectivity (stationary or temporally variable),
- 3 X tag recovery options (overdispersion 2, 20, 200)
- 3 X different weightings for the size composition data.
- Other assumptions as defined in Table 1 and the Methods section.

None of the 54 model specifications seemed to experience outright convergence failures (but some of the models with time varying selectivity might have been marginal). Plots of key summary statistics are shown in Figure 20 - Figure 28. These figures suggest that:

- All of the models had an excellent fit to the CPUE series (RMSE (1985-2008) \sim 0.05 – 0.1, Figure 20). The quality of fit is slightly worse with the low variance growth curve (gr05) and higher CL data weighting (CL1010, CL1001).
- All of the options produced a satisfactory fit to the longline size data (ESS \sim 900 – 1600), and none of the options produced a very good fit to the non-longline size data (ESS \sim 35-113) (Figure 21 - Figure 24). On the basis of the ESS, the PSFS fleet (with the most irregular bimodal catch pattern) was the best fit of the non-longline fleets, and least sensitive to the alternative assumptions. Lower variance on the length-at-age function (gr05) was the most influential factor affecting the fit to the size composition data. This is what might be expected given the rapid in-season growth of the juveniles and possibility of continuous recruitment. The model is clearly doing a poor job of fitting the very small size classes, and this should really not be an important issue because we are reasonably comfortable assigning an age to these lengths. This probably should have been handled by using fewer, aggregated length bins for the first age class.
- The biomass and fishing mortality estimates among the different models show considerable variability (Figure 25 - Figure 27). None of the models estimate B_{MSY} to be exceeded (prior to the projection years), while 5 models suggest that $F_{2008} > F_{MSY}$. The tag recovery assumptions are the most influential with respect to the stock status estimates (higher overdispersion results in more pessimistic stock status interpretation).

- The slope of the least squares linear trend in the recruitment deviations over the period 1985-2007 was calculated as an additional index of interest that reflects the systematic lack of fit to the stock recruitment relationship. This reported value was converted to the percentage change from the beginning to the end of the time period, *i.e.* such that a value of -25 might be loosely interpreted as a systematic decline in recruitment of 25% between 1985 and 2007, after accounting for the effect of changing spawning stock size. Figure 28 indicates that the distribution of this index closely resemble the biomass stock status estimates (*i.e.* the factors that drive the recruitment trend also increase B_{2008}/B_{MSY}).

On the basis of these results, the decision was made to maintain the uncertainty in the tag overdispersion for the next stage of uncertainty quantification. The other dimensions were reduced to a single option. The broader length-at-age option, gr20, was maintained because, relative to gr05, it does a better job of admitting continuous recruitment and in-season growth, provides a better fit to the size composition data, and represents less of a constraint to the stock status indices. The stationary selectivity assumption was adopted because the temporal variability did not seem to resolve any problems and was computationally expensive (with potentially unreliable convergence). The size composition down-weighting option, CL0101, was adopted because the concerns about over-influential size data could not be reduced in any other way tested.

Uncertainty Grid C – What are the implications of the key data and life history assumptions and how do they interact?

A grid that encompasses the major elements of uncertainty for this stock was applied with a balanced design of 108 model specifications (Table 1), including:

- 3 X Tag recovery options (negative binomial overdispersion 2, 20, 200)
- 2 X Recruitment deviate options ($sd(\log(\text{devs})) = 0.6, 0$)
- 2 X CPUE series (no catchability trend, catchability trend of 0.47%/y)
- 3 X steepness option ($h = 0.55, 0.75, 0.95$)
- 3 X M vectors ($M(\text{age } 2y+) = 0.32, 0.4, 0.48$)
- Other assumptions as defined in Table 1 and the Methods section

The model summary diagnostics are disaggregated by the individual assumption options in Figure 29 - Figure 37, from which we note the following points:

- Neither the fit to the size composition nor the fit to the CPUE data provides much justification for selecting preferable scenarios among the grid C models:
 - The fit to the CPUE series is reasonable in all cases (Figure 29). As would be expected, the models with estimated recruitment deviates (r6, RMSE $\sim 0.04-0.07$) fit somewhat better than the deterministic recruitment model (r0, RMSE $\sim 0.08-0.14$).
 - The fit to the LL size composition data is much better than the non-longline fleets, however, the difference in fit among model specifications is trivial for all of the model options (Figure 30-Figure 33).
- Slightly more than half of the models suggest that $B(2008) < B_{MSY}$, and slightly less than half the models suggest that $F_{2008} > F_{MSY}$ (Figure 35 - Figure 34). The more pessimistic scenarios are associated with higher tag overdispersion (t200), deterministic recruitment (r0), Effort

creep (EC47), lower steepness (h55) and lower M (M08). Not surprisingly, the model options have similar relative effects on the MSY estimates (Figure 37).

- With stochastic recruitment (r6), the models always estimate a downward trend in the deviations over time (sometimes exceeding 50%, Figure 39). In all cases examined, the trend is anchored by a very large negative deviate in the last year (which cannot be justified on the basis of the size composition data). In general, the same model options that lead to a more optimistic interpretation of stock status are also associated with larger trends in the recruitment deviations.
- While one has to be careful interpreting the likelihoods in this context, there are some trends worth noting (Figure 40). The objective function generally suggests a better fit for the more optimistic steepness (h95) and M (M12) options. The catchability options are mixed (lowest and highest values are associated with effort creep EC47). The intermediate overdispersion value was preferred for the tagging data (t20). The comparison between the stochastic and deterministic recruitment options (r6, r0) is not very useful.

To provide a further reference for assessing the plausibility of the different models, detailed results from two of the most productive ('maxMSY'; MSY = 183000 t) and unproductive ('minMSY'; MSY = 89000 t) models were selected (assumptions defined in Table 1). Figure 41 and Figure 42 illustrate the fit to the CPUE and tagging data for these two models. As would be expected, the model with the estimated recruitment deviates (maxMSY) provides the better fit to the CPUE data; RMSE CPUE(1985-2008) = 0.03 vs 0.14. But the more important question is whether the fit to the CPUE of minMSY is plausible? The CPUE RMSE for minMSY is considerably higher than the assumed value in the model (0.05), but well within the bounds of what we would expect is reasonable for a relative abundance index derived from commercial CPUE. minMSY predicts a long-term decline in abundance which is consistent with the CPUE, but there is a systematic failure to fit the final few years of relatively stable CPUE. It would be difficult to reject outright minMSY on the basis of the CPUE, because this lack of fit is comfortably within the magnitude that can be attributed to observation error. The tagging data also appear to be better fit for maxMSY than minMSY, which would be expected given the different assumed overdispersion values (combined with the difference in recruitment freedom). However, none of the models provided a good fit to the seasonal variability in the tagging data, and it is not clear that we have enough confidence in the tagging data implementation to discriminate on this basis at this time. The fit to the size composition data for these two models is even less informative (ESS values differ by less than 2 for all fisheries). These plots are not shown, because they are almost identical to each other and the reference case (Figure 17). We would be inclined to conclude that the model fits to the data do not provide a sufficiently compelling criteria to choose between these models. If we are prepared to accept that these two most extreme models are plausibly consistent with the data, then we can expect that other models can also be identified which span the range between the two.

Stock Status

The range of models represented in Grid C recognizes a considerable amount of uncertainty in the stock status. The question is how much uncertainty is appropriate?

Obviously some of the models defined in Grid C are closer to reality than others, but are we able to objectively choose among them (and have other important dimensions been left out?) As in a classical Bayesian analysis, there are two basic ways to assign credibility to different options:

- 1) Prior weighting – experience from other fisheries systems has obviously helped to formulate the alternative models in the first place. Can the different scenarios be weighted to reflect the prior beliefs of the analysts or broader WPTT? There is obviously an element of subjectivity in this process. But this is true in any model formulation and selection. At least in this case the weightings are transparent and open to criticism.

- 2) Likelihoods – what do the data tell us about the plausibility of the different models? To weight the parameter estimates by the likelihood, we are assuming that the model is correct, and the data were generated according to a well-defined statistical process. In this case that is exactly what we are doing within each model for some parameters (*i.e.* virgin recruitment, selectivity, LL catchability). However, we tend to be sceptical about the capacity of the model to estimate many key parameters (*e.g.* steepness, M). Sometimes a particular parameter (*e.g.* M) is freely estimated, and if the result is within prior expectations people tend to accept the result. Otherwise they revert to the prior (which is equivalent to saying that they are always largely constrained by the prior). Given the oversimplifications and dubious statistical assumptions in these models, some parameter estimates are always questionable. Often the data appear to be very unlikely within the constraints of the model and this will pull a parameter estimate toward a very precise value (or bound), while common sense might suggest that, qualitatively, the fit to the important data sources does not seem all that different with alternative parameter values. There is a further problem with the likelihood weighting in this case, in that it is simply not meaningful to compare all of these models (*i.e.* if we have fundamental doubts about the validity of the model to estimate year-class strength, the better likelihood for the stochastic recruitment model is not a reason for rejecting the deterministic recruitment model).

We developed the current stock status estimates on the basis of only a prior weighting on the different model options. The justification follows from consideration of the following points with respect to the five sets of assumption options in grid C:

- Estimation of the stock recruitment deviations is problematic. The models never fit the non-longline fleet size composition data very well, the longline fleet is probably not very informative about year-class strength, selectivity is likely changing over time for some or all fleets, and the model does not seem to resolve the early growth pattern and length-at-age distribution very well. Thus the prime justification for the stochastic recruitment comes from the CPUE series (with possibly a minor influence from the tags for the most recent recruits). We weighted the two scenarios equally at this time, recognizing that the deterministic recruitment may lead to a more robust estimator given the size composition problems, but also admitting that the stochastic recruitment trends are possible, and required to explain the CPUE pattern. Recruitment Variance summary weight:
 - $r_6 = 0.5$
 - $r_0 = 0.5$

- We do not have a strong view of the validity of the effort creep scenario. It is generally assumed that fleets would become more efficient over time, and Hoyle *et al.* (2009) provide evidence for this in the Pacific. However, further analyses (Hoyle *et al.*, 2010) were

inconclusive. There are mechanisms that could undermine this assumption (*e.g.* changing target species, replacement of experienced crews and skippers during economic downturns, *etc.*). The goodness of fit indices do not suggest that either catchability scenario is more consistent with the data. In the absence of compelling evidence for effort creep, we would tend to favour the stable catchability scenario, but admit that effort creep remains a possibility. Catchability option summary weight:

- $EC0 = 0.8$
- $EC47 = 0.2$
- Stock recruitment curve steepness is generally difficult to estimate, especially if there is poor contrast in stock size. There may also be physical or ecological factors driving long-term trends in recruitment productivity. There is a certain self-reinforcing circularity in adopting values from other oceans (*i.e.* especially if the data are no better in these other fisheries). We would tend to give higher weight to the central steepness value, but not rule out the possibility of the higher or lower options (and note that the likelihood values tend to support the higher steepness). Steepness summary weight:
 - $h55 = 0.1$
 - $h75 = 0.7$
 - $h95 = 0.2$
- Subsequent to the initial analysis, definite biases in the processing of the tag data were identified, such that the lower tag weight is clearly preferable at this time (and removing the tags would be better yet). Tag overdispersion summary weight:
 - $t002 = 0.1$
 - $t020 = 0.3$
 - $t200 = 0.6$
- The different M options were all weighted equally.

For this analysis, the best point estimates for biomass and fishing mortality reference points are presented as the mean of all 108 models in gridC, with each model weighted according to its specific combination of assumptions. The uncertainty distribution is defined by the minimum and maximum of the MPD estimates from gridC (*i.e.* irrespective of the weighting). In one sense it might be argued that this distribution overstates the uncertainty, because some of the models at the extreme may have a very low weighting. Conversely, it might also be argued that the uncertainty distribution is under-estimated, because the most extreme models are actually MPD estimates, and the statistical uncertainty conditional on these models is ignored. Using this method:

- $B_{2008}/B_{MSY} = 1.05 (0.64, 1.69)$
- $F_{2008}/F_{MSY} = 0.91 (0.40, 1.78)$

The mean option-weighted time series for the biomass and fishing mortality reference points are shown in Figure 44, with the corresponding Kobe plot in Figure 45. The Kobe-2 strategy matrix for the weighted results of grid C is presented in Table 3.

Other members of the WPTT may be in a better position to comment on the proposed weighting scheme adopted, and the summary plots and matrices can easily be updated to reflect the deliberations of the WPTT as required. However, it is not worth revisiting this issue unless/until the analysis can be updated with the new data and revised tagging assumptions.

4. Conclusions

1. This analysis represents an extension of the 2009 BET stock assessments, including new fleet definitions, inclusion of the RTTP-IO, and an extensive exploration of model assumptions (and interactions among assumptions). Unfortunately, the most recent year of data has not been included, and the implementation of the tagging data is biased in a known direction (but unknown magnitude). If time permits, these issues will be addressed in a revised document before/during the WPTT.
2. All of the models presented provided a reasonable fit to the CPUE series and the longline size composition data. However, despite attempts to define homogenous fleets with very flexible selectivity functions (including temporal variability), none of the models provided a satisfactory fit to the non-longline size data or the seasonal patterns in the tag data. This raises a fundamental question of whether stochastic recruitment can be usefully estimated at this time.
3. The relative credibility of the most influential assumptions in the model cannot really be quantified on the basis of the model fit to the data. A subjective weighting scheme is proposed for synthesizing the results into stock status estimates and uncertainty distributions. The central tendency of the results suggest that the stock is fully, but not over-exploited (near F_{MSY} and B_{MSY}), but the full range of results suggests that the stock status could be considerably more optimistic or pessimistic than the mean. Furthermore, these conclusions are expected to be optimistic, given the known direction of the biases in the current tag implementation (though the magnitude of the bias and degree of influence of the tags is not known).
4. A number of issues are identified for future assessments:
 - The use of the commercial longline CPUE series as a relative abundance index is critical to the assessment. Any analyses that can help quantify the effects of changing catchability over time, or constructing realistic estimates of uncertainty around these series are very important. It is noted that work undertaken in 2010 in Japan between the scientists of the Secretariat of the Pacific Community and the Far Seas Fishery laboratory has been fruitful for Pacific stocks (*e.g.* Hoyle *et al.*, 2010). Similar collaborative work is encouraged for the Indian Ocean.
 - Total fishery removals are probably the most important data in the stock assessment process. Any efforts that improve the estimates (and uncertainty) of the removal data is valuable, and obvious points of concern include errors in the artisanal fleet catches and Purse Seine species composition sampling.
 - There are several avenues related to fitting the size composition data that can be attempted. However, given the doubts about the stationary selectivity assumption in some fleets, it is not clear that these efforts will lead to obvious improvements in the estimates of year class strength in the assessment:

- Investigate the size composition data to better understand the statistical properties of the samples within and among years and fleets. Combined with a spatial analysis, this may help improve the homogeneity of fleet definitions.
- Revisit the mean growth equation in relation to the two stage growth curve estimated from tagging studies (this deviation from traditional growth curves cannot currently be represented in SS3 software).
- Revisit the growth equation variance so that the predicted length-at-age distributions better represent the recruitment seasonality and within year growth, particularly for the youngest ages.
- Depending on these analyses, it may prove useful to reformulate the seasonal representation of the SS3 model (*e.g.* to define model years as quarters to accommodate continuous recruitment, or to partition fleets by season if there is evidence for important seasonal changes in selectivity).
- It is expected that ongoing analyses of the tagging data will be useful in multiple ways:
 - Estimates of F and M from independent tagging analyses (*e.g.* Brownie-Petersen estimators) should be compared with those derived from the integrated analysis to see if major discrepancies can be identified and resolved.
 - Estimates of movement from fine-scale analyses (*e.g.* advection-diffusion models) might help to determine the appropriate spatial structure of the assessment, and appropriate mixing times for tag releases.
- Meta-analysis of other BET populations might help to constrain some of the model assumptions (recognizing that there is a risk of circular arguments).
- While the emphasis on model selection uncertainty undertaken here probably provides a more realistic representation of the real assessment uncertainty, it is not clear how best to include this information in the management advice provided to the Commission. Management Strategy Evaluation is one possible avenue that can more fully exploit this representation of uncertainty. In this context, this model, or something similar could be used as an operating model.

5. Acknowledgements

We are grateful for the contributions of various parties toward the compilation of data and preparatory analyses, particularly Hiroaki Okamoto and Hiroshi Shono for assembling the longline CPUE and various life history parameters used in the analysis. The SS3 software and supporting graphical routines produced by Rick Methot, Ian Taylor (and colleagues) were very useful and much appreciated.

6. References

- Gaertner, D. and J.P. Hallier. 2009. An updated analysis of tag-shedding by tropical tunas in the Indian Ocean. IOTC-2009-WPTT-34.
- Herrera, M. Estimation of catch-at-size, catch-at-age and total catch per area. IOTC-2009-WPTT-12.
- Hoyle, S. 2009. CPUE Standardization for bigeye and yellowfin tuna in the western and central Pacific Ocean. WCPFC-SC5-2009/SA-WP-1.
- Hoyle, S., H. Shono, H. Okamoto, and A.D. Langley. 2010. Analyses of Japanese longline catch and effort for bigeye tuna in the WCPO. WCPFC-SC6-2010/SA-WP-2.
- Langley, A., M. Herrera, J.P. Hallier and J. Million. 2009. Stock assessment of yellowfin tuna in the Indian Ocean using MULTIFAN-CL. IOTC-2009-WPTT-10.
- Methot, R.D. 2000. Technical description of the stock synthesis assessment program. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-43, 46 p.
- Methot, R.D. 2009. User Manual for Stock Synthesis. NOAA Fisheries, Seattle, USA.
- Okamoto, H., K. Satoh, and H. Shono. Japanese longline CPUE for bigeye tuna in the Indian Ocean up to 2008 standardized by GLM. IOTC-2009-WPTT-05.
- Shono, H., K. Satoh, H. Okamoto, and T. Nishida. 2009. Updated stock assessment for bigeye tuna in the Indian Ocean up to 2008 using Stock Synthesis III (SS3). IOTC-2009-WPTT-20.
- WPB 2009. Report of the eleventh session of the IOTC working party on tropical tunas. Mombasa, Kenya, 15-23 Oct 2009.

Table 1. Model assumption combinations used in selected models and grids (where a grid represents the list of models with all possible combinations of the indicated options)

Model Option	Model or Grid						
	ref	reft002	reft200	opt	pes	Grid B (54 models)	Grid C (108 models)
Length-at-age Variance	gr20	gr20	gr20	gr20	gr20	gr05 gr20	gr20
Tag Recovery Overdispersion	t020	t002	t200	t002	t200	t002 t020 t200	t002 t020 t200
Selectivity	sc	sc	sc	sc	sc	sc st3 st23	sc
Recruitment Deviations	r6	r6	r6	r6	r0	r6	r0 r6
Catchability	EC0	EC0	EC0	EC0	EC47	EC0	EC0 EC47
Stock-Recruit Steepness	h75	h75	h75	h95	h55	h75	h55 h75 h95
Natural Mortality	M10	M10	M10	M08	M08	M10	M08 M10 M12
Catch-at- Length weighting	CL0101	CL0101	CL0101	CL0101	CL0101	CL1010 CL1001 CL0101	CL0101

Table 2. Weighting scheme adopted for the 108 model scenarios represented in the Kobe-2 Strategy Matrix (Table 3).

	Option Weighting		
Growth Curve Variance	gr05 = 1.0		
Tag Overdispersion	t002 = 0.1	t020 = 0.3	t200 = 0.6
Selectivity	sc = 1.0		
Recruitment sd(log(dev))	r0 = 0.5	r6 = 0.5	
Catchability	EC0 = 0.8	EC47 = 0.2	
Steepness	h55 = 0.1	h75 = 0.7	h95 = 0.2
M	M08 = 0.33	M10 = 0.33	M12 = 0.33
Catch-at-Length weighting	CL0101 = 1.0		

Table 3. Kobe-2 Strategy Matrix indicating the estimated stock status implications of different constant catch strategies, with the assumptions described in grid C (Table 2).

Stock status Reference Point	Projection Time frame	Weighted percentage of scenarios that violate the Reference Point				
		C(2008) -40%	C(2008) -20%	C(2008)	C(2008)+20%	C(2008)+40%
P(Bt<Bmsy)	In 3 years	11	18	33	55	74
	In 10 years	11	18	33	69	80
P(Ft>Fmsy)	In 3 years	<1	5	23	59	80
	In 10 years	<1	5	23	62	80

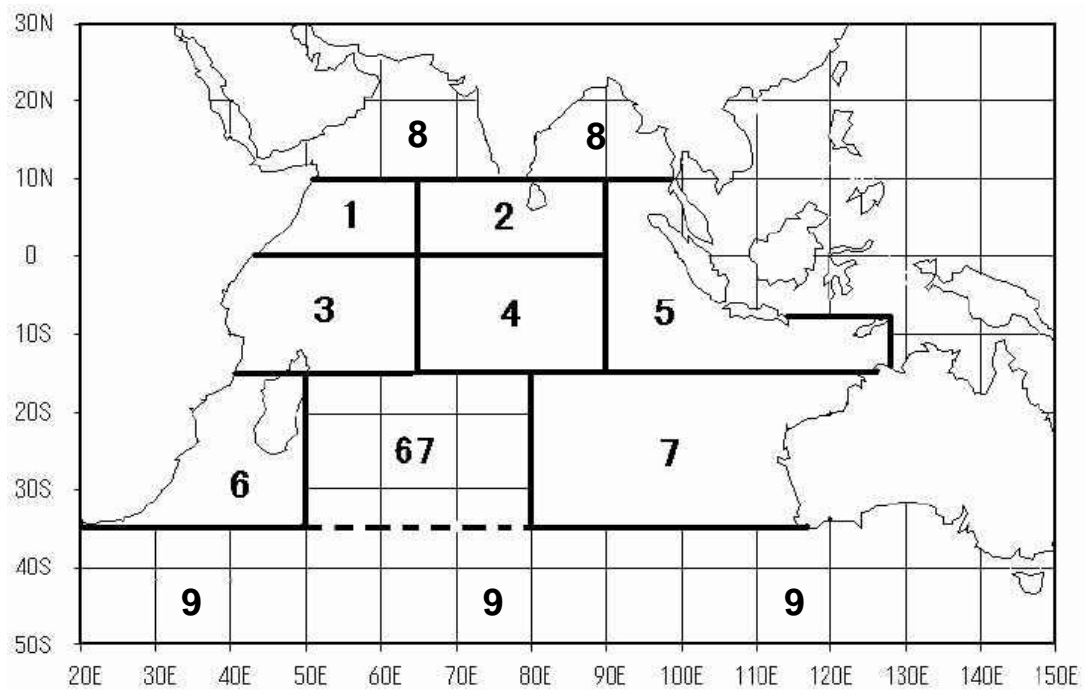


Figure 1. Regions used in the BET assessment for the definition of fleets.

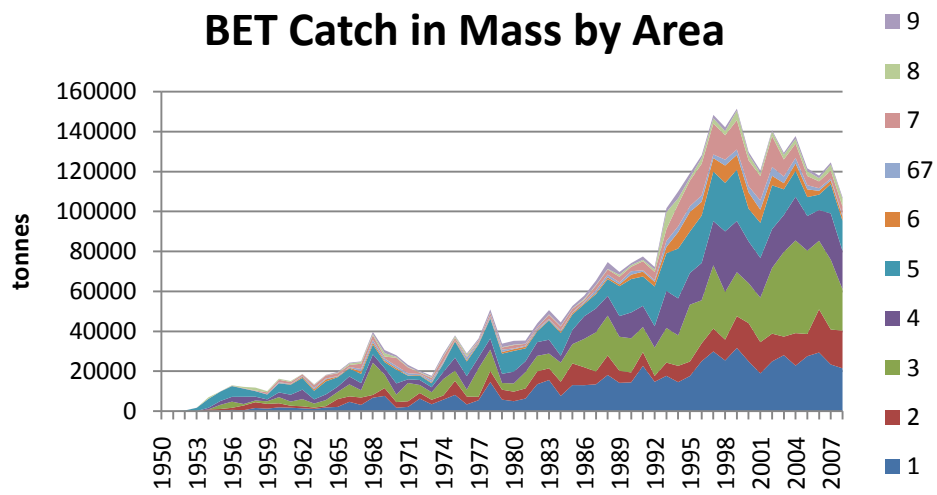
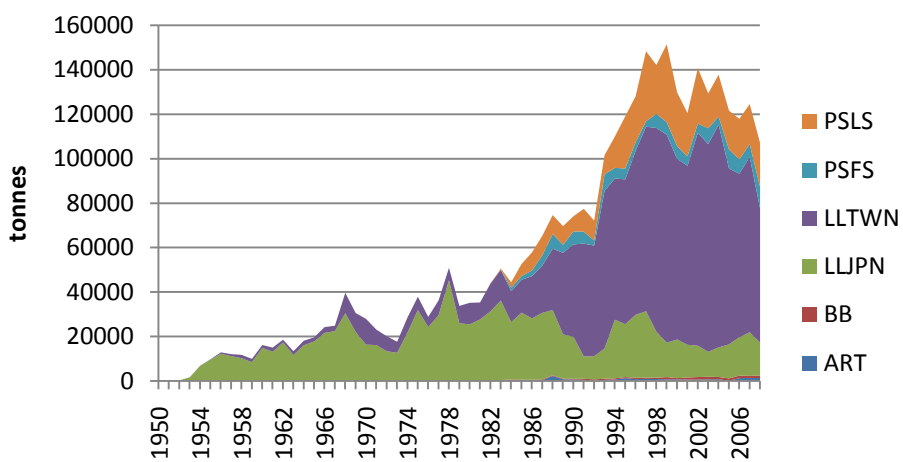


Figure 2. Bigeye tuna catch in mass disaggregated by the areas shown in Figure 1.

BET Catch in Mass by Gear



BET Catch in Numbers by Gear

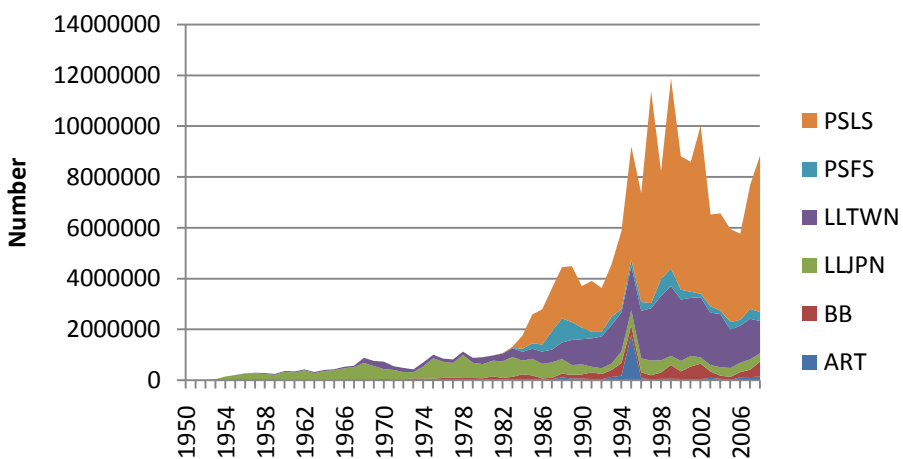


Figure 3. Bigeye tuna catch in mass (top) and numbers(bottom) disaggregated by fleet (PS=Purse Seine, LS=FAD-associated, FS= Free School, LL=Longline, TWN=Taiwan, JPN=Japan, BB = bait boat, ART=Artisanal).

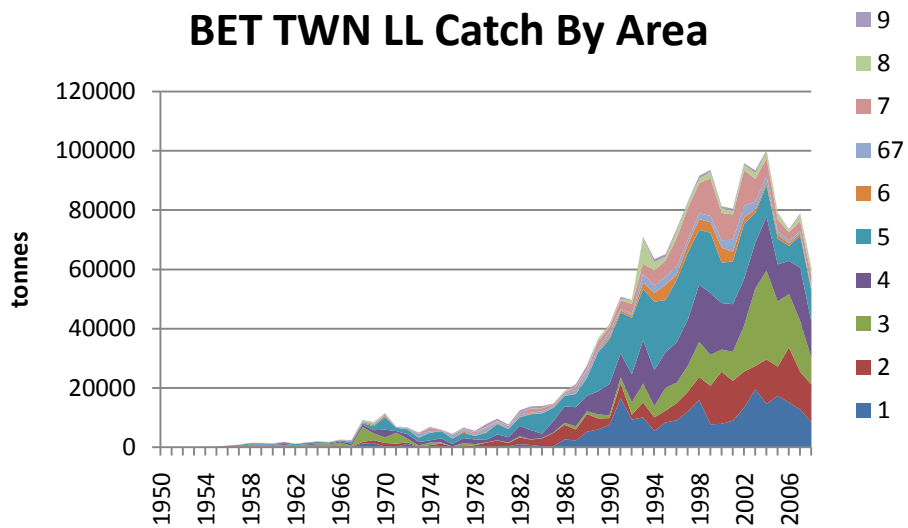
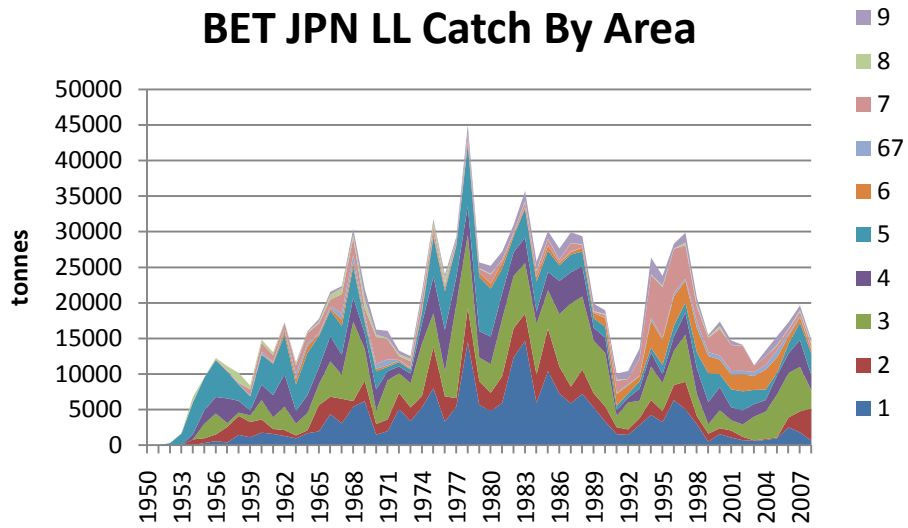


Figure 4. Bigeye tuna longline catch in mass over time for the Japanese (top) and Taiwanese (bottom) fleets, disaggregated by area.

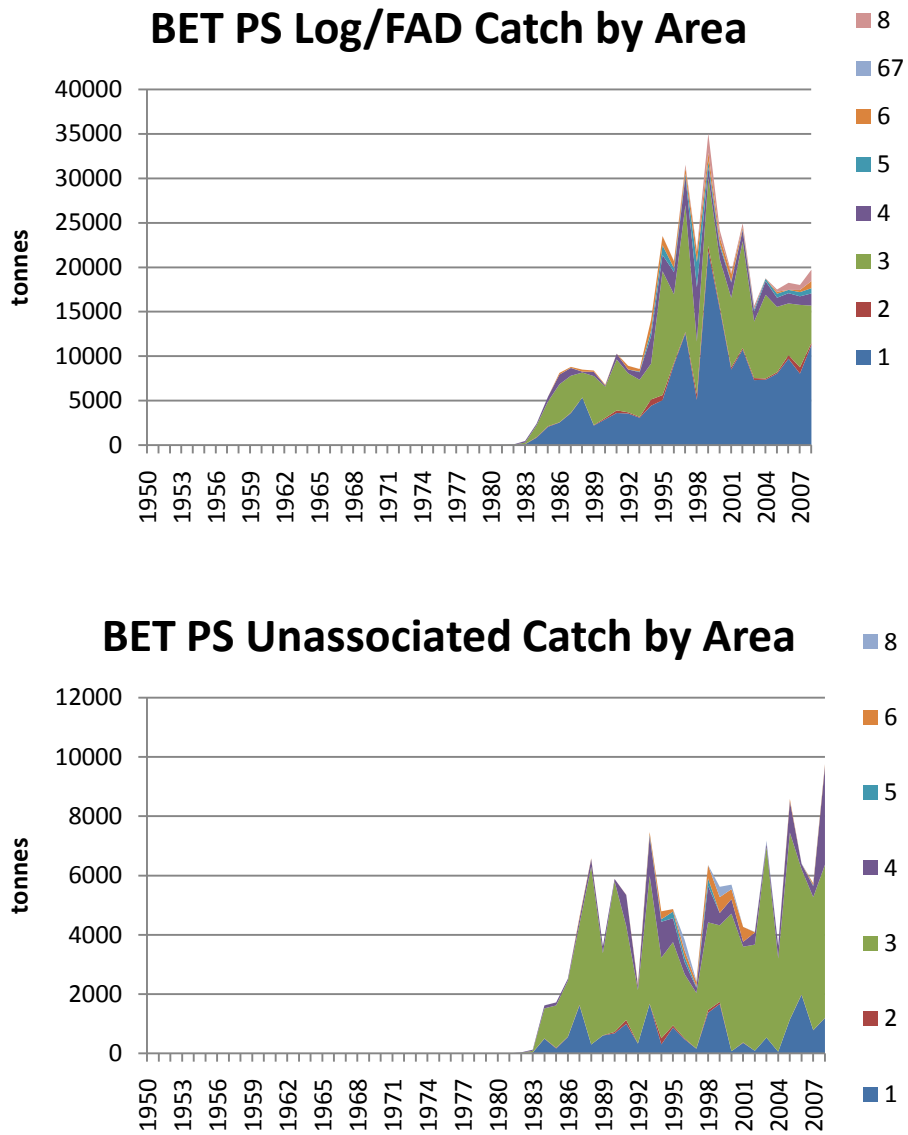


Figure 5. Bigeye tuna Purse Seine catch in mass over time for FAD-associated (top) and free school (bottom) sets, disaggregated by area.

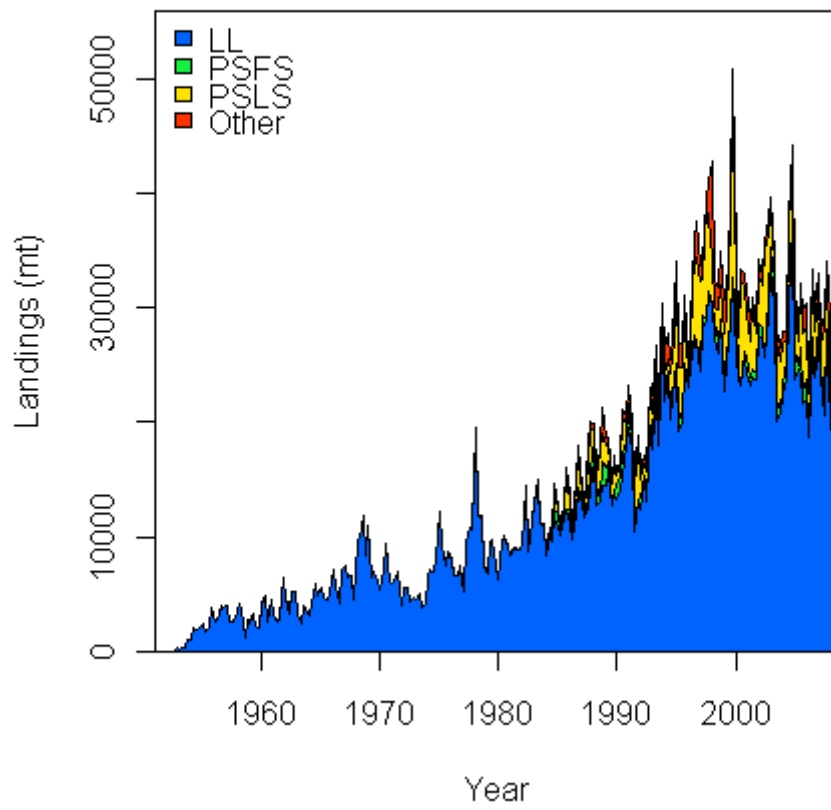


Figure 6. Catch partitioned by the 4 fleets adopted in this analysis.

BET Tag Movements

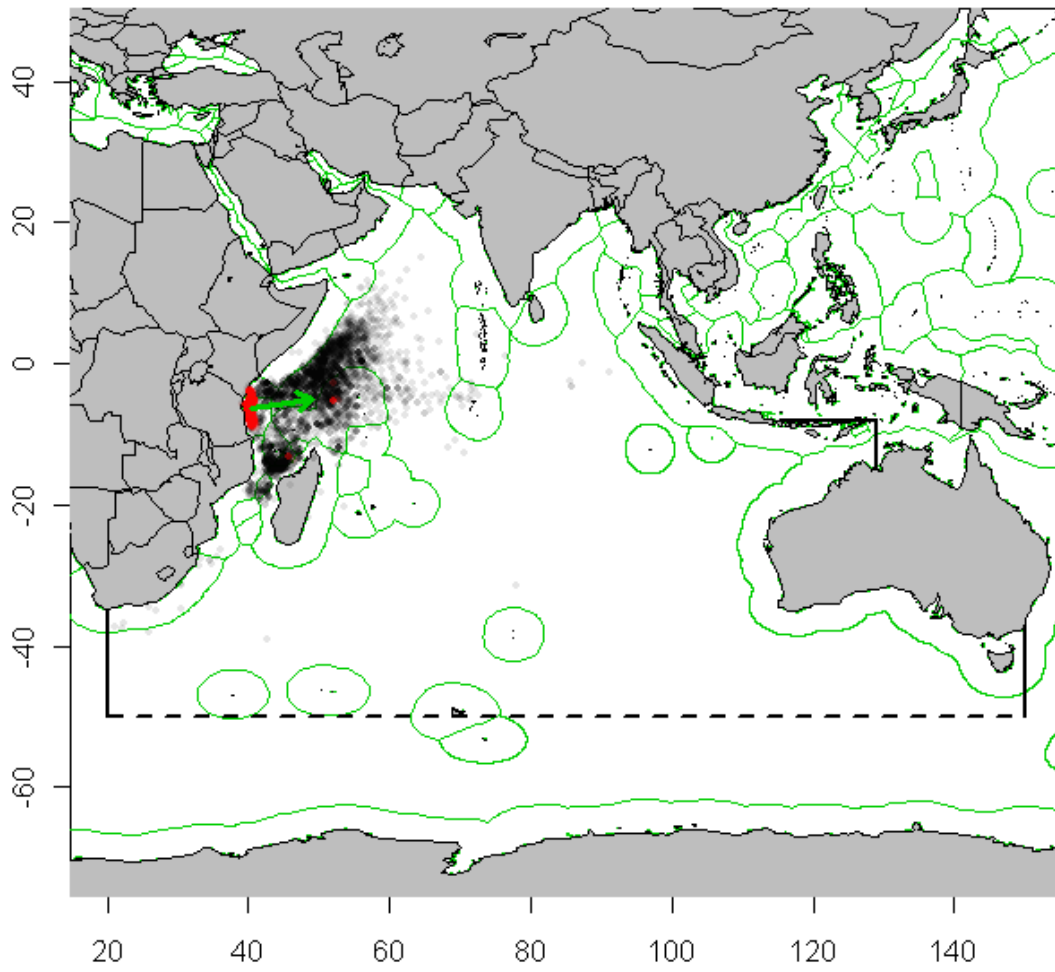


Figure 7. Summary of RTP BET tag releases (red circles) and recoveries (black circles). Full colour saturation indicates 20+ tags. The green arrow indicated the mean displacement of all recovered tags.

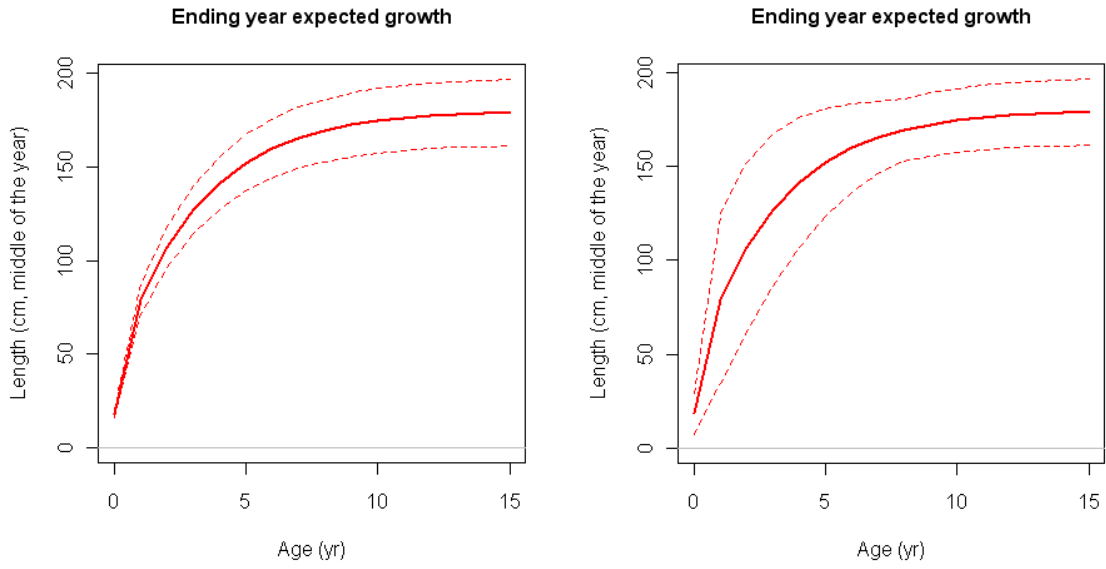


Figure 8. Comparison of length-at-age relationships used in the models (left panel = gr05, right panel=gr20).

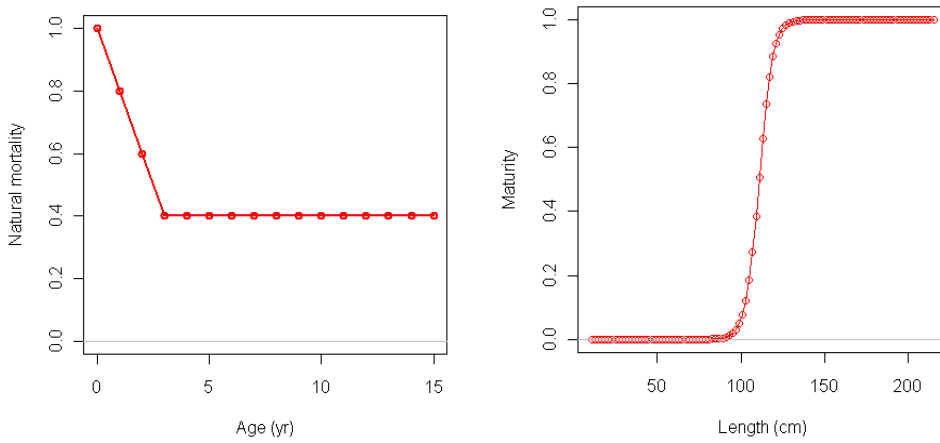


Figure 9. Assumed BET mortality (left) and maturity (right) schedules.

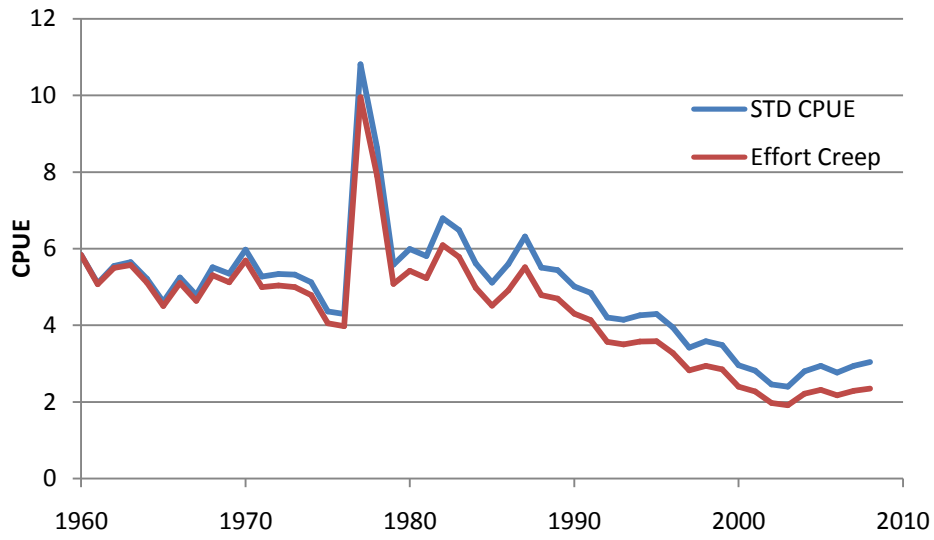


Figure 10. Standardized JPN longline CPUE, and the effort creep scenario (linear 0.47%/y).

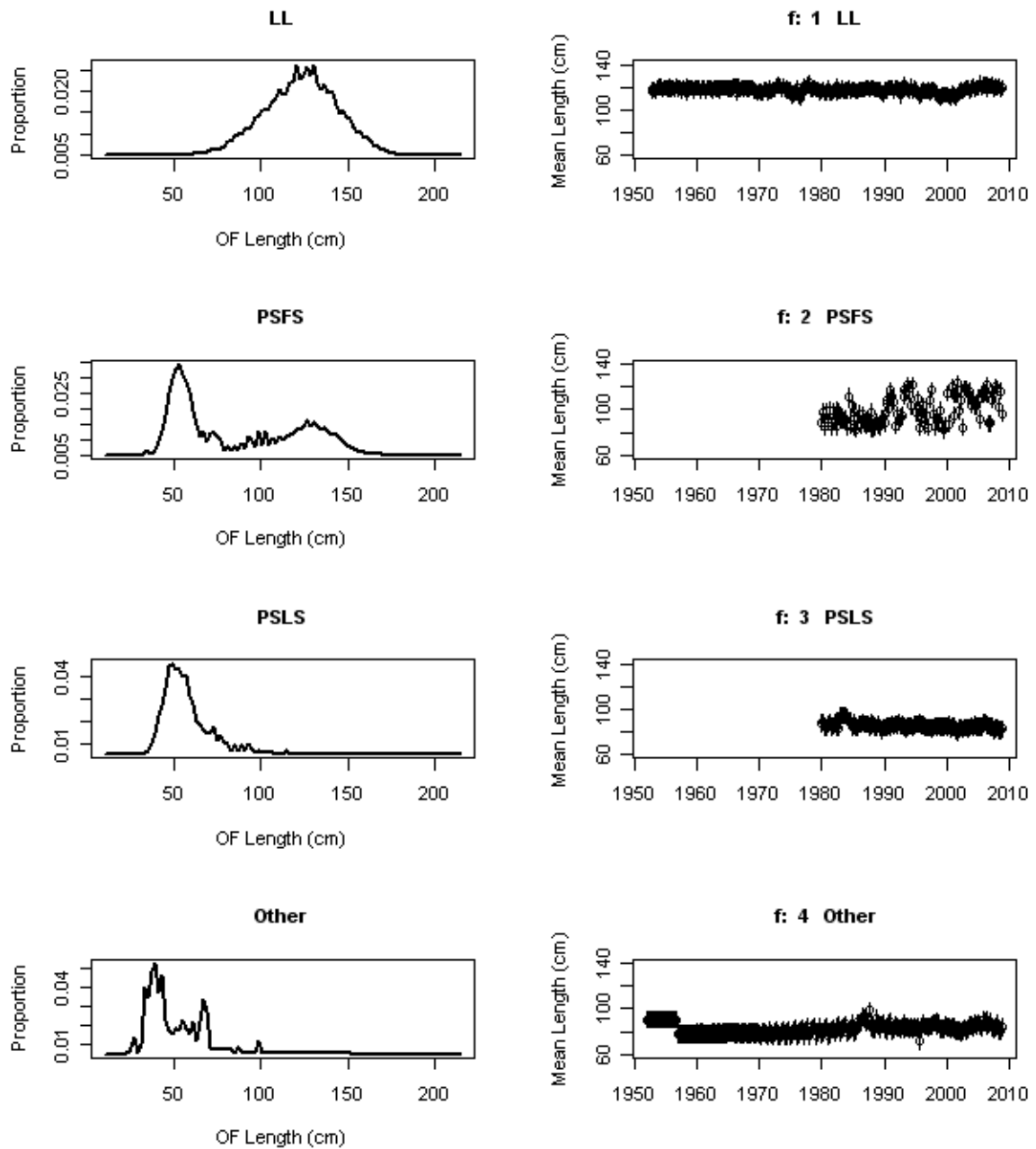


Figure 11. Size composition data by fleet. The left column represents the sum over time with all years weighted equally, right column represents the quarterly time series of mean length (with 95% CI for an assumed sample size of 200).

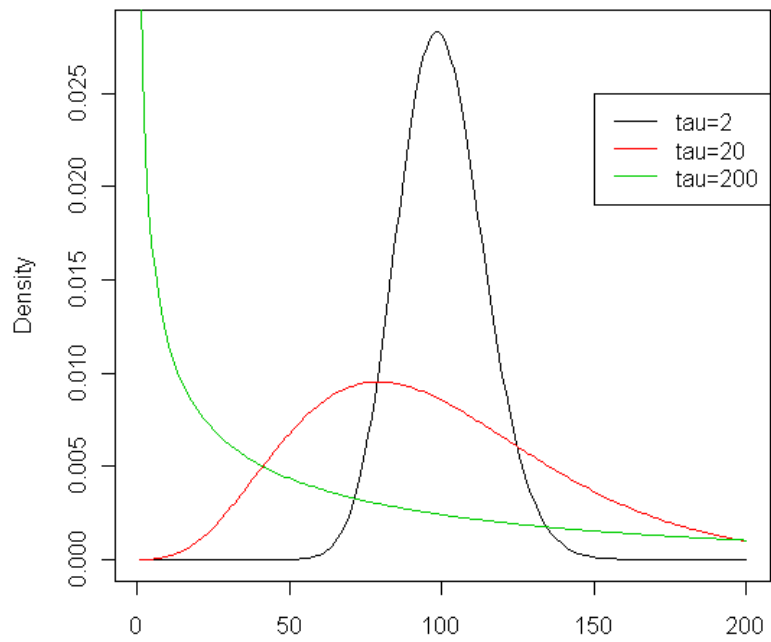


Figure 12. Negative binomial probability density functions for 100 expected tag recoveries and range of overdispersion (τ) parameters (2, 20, 200).

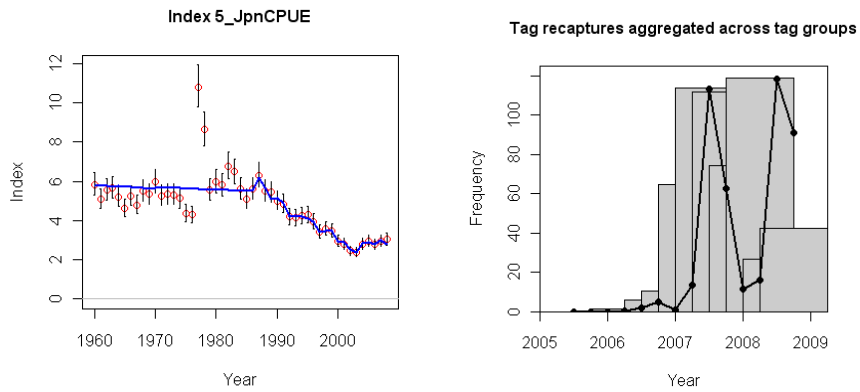


Figure 13. Left panel: Predicted (line) and observed (points) CPUE for the reference case. Right panel: Predicted (line) and observed (bars) tag recoveries for the PSL fishery.

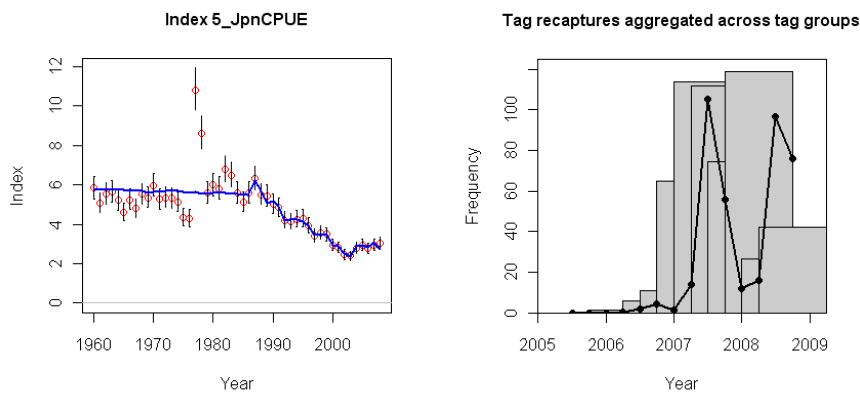


Figure 14. Left panel: Predicted (line) and observed (points) CPUE for ref002 (reference case with increased tag influence). Right panel: Predicted (line) and observed (bars) tag recoveries for the PSL fishery.

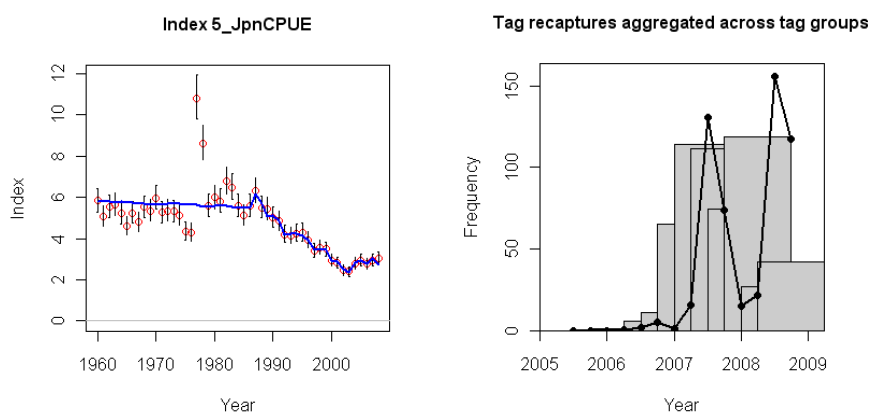


Figure 15. Left panel: Predicted (line) and observed (points) CPUE for ref200 (reference case with decreased tag influence). Right panel: Predicted (line) and observed (bars) tag recoveries for the PSL fishery.

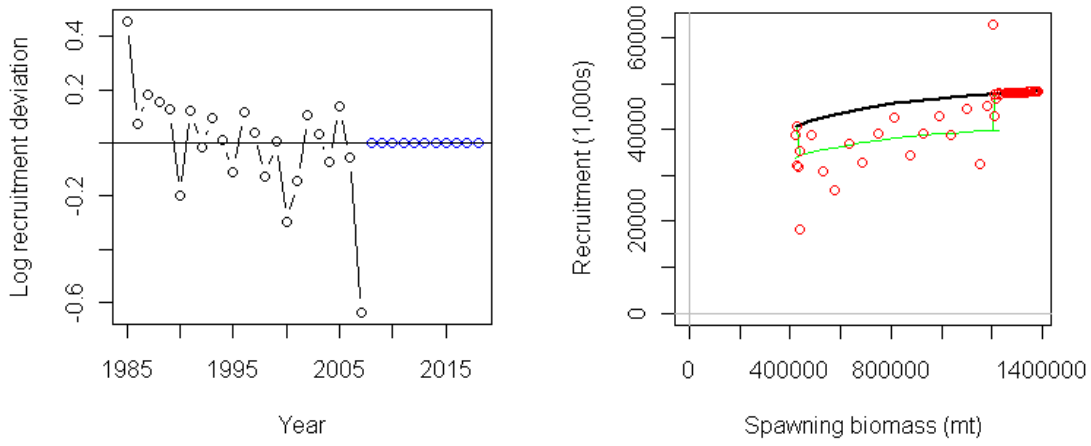


Figure 16. Estimated recruitment deviations (left panel), and stock recruitment relationship (right panel) for the reference case.

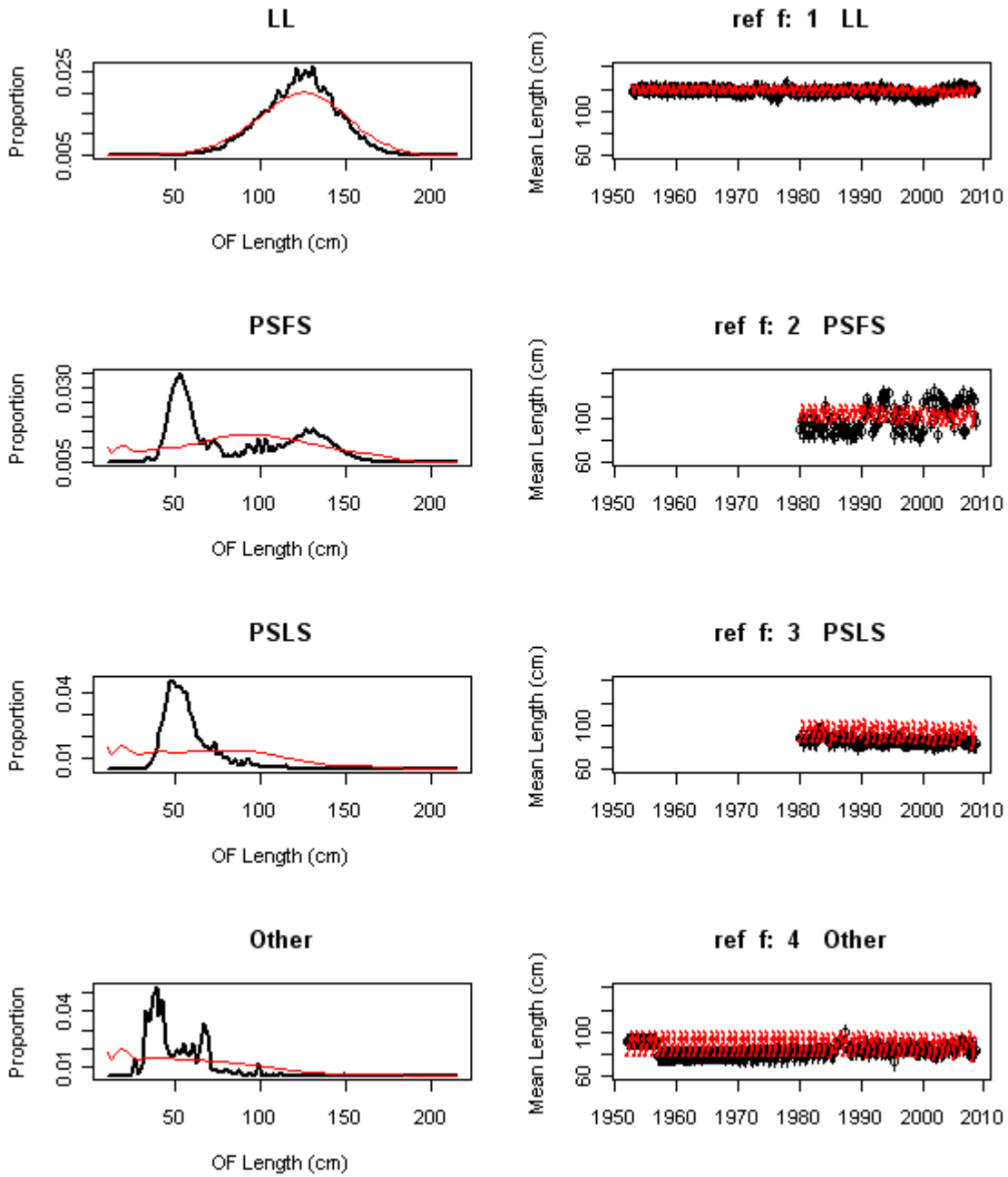


Figure 17. Predicted (red) and observed (black) catch size composition for the reference model (left column is the aggregated distribution over time, right column is the time series of mean catch).

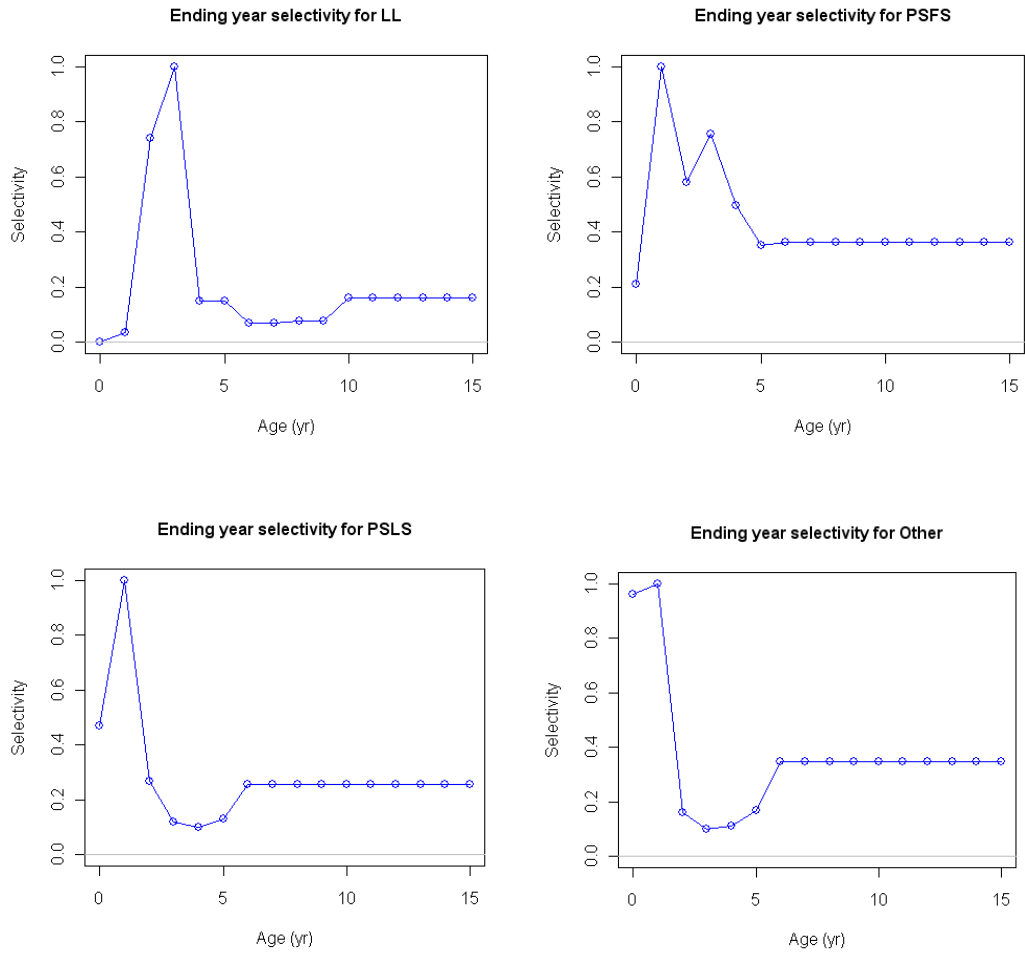


Figure 18. Selectivity estimated for the reference case model.

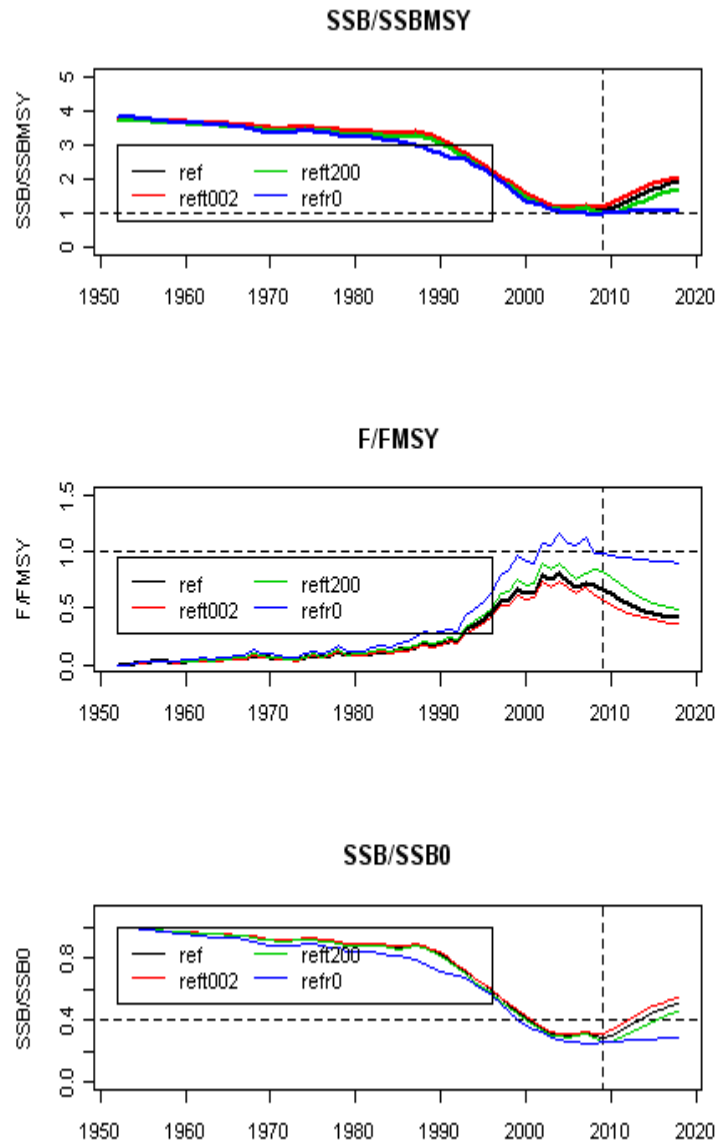


Figure 19. MPD Biomass and Fishing mortality time series (including 10 years of constant catch projections) for the reference case and models with increased (ref002) and decreased (ref200) influence of the tagging data. ref0 illustrates the reference case model without the estimation of recruitment deviations.

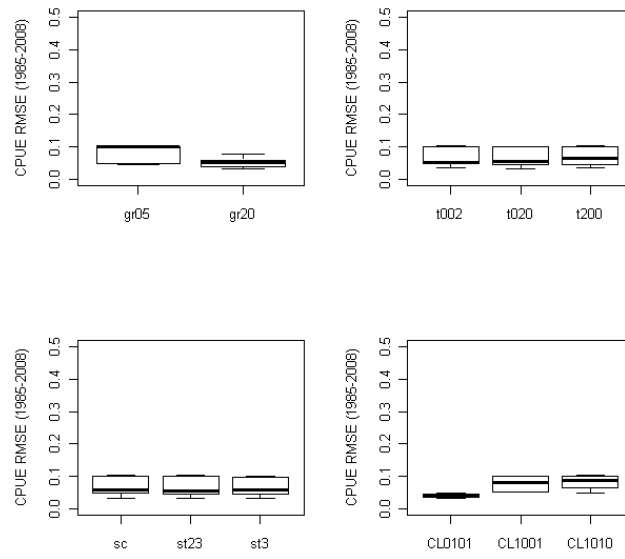


Figure 20. Summary of model fit to the CPUE series for the 54 models from grid B, partitioned by the different assumptions.

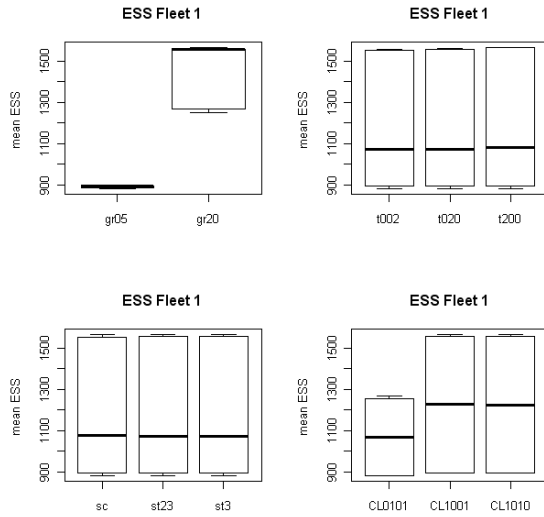


Figure 21. Summary of model fit to the LL fleet size composition data for the 54 models from grid B, partitioned by the different assumptions.

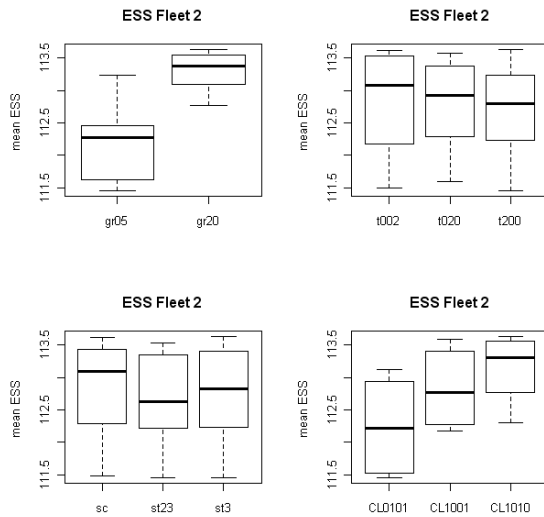


Figure 22. Summary of model fit to the PSFS fleet size composition data for the 54 models from grid B, partitioned by the different assumptions.

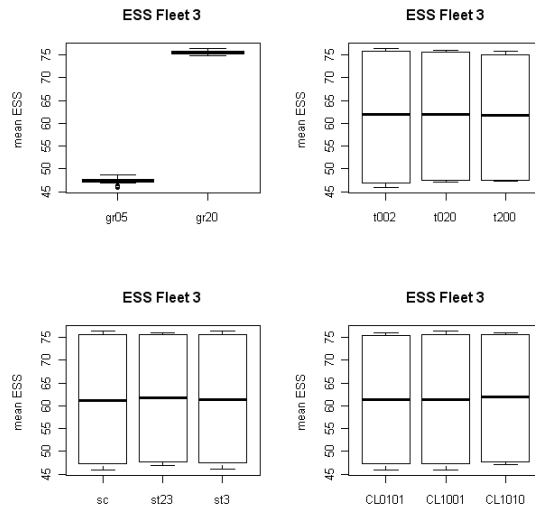


Figure 23. Summary of model fit to the PLS size composition data for the 54 models from grid B, partitioned by the different assumptions.

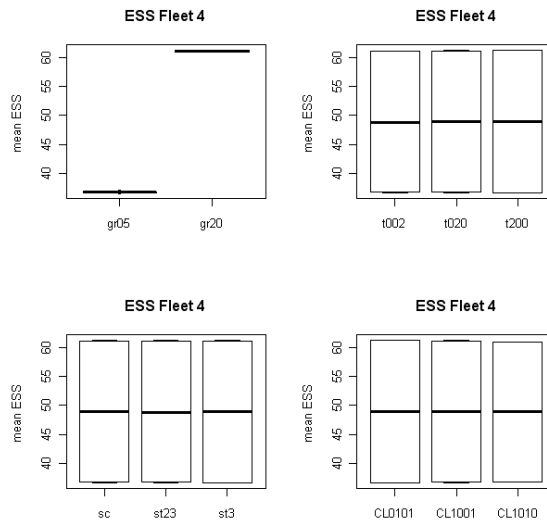


Figure 24. Summary of model fit to the Other fleet size composition data for the 54 models from grid B, partitioned by the different assumptions.

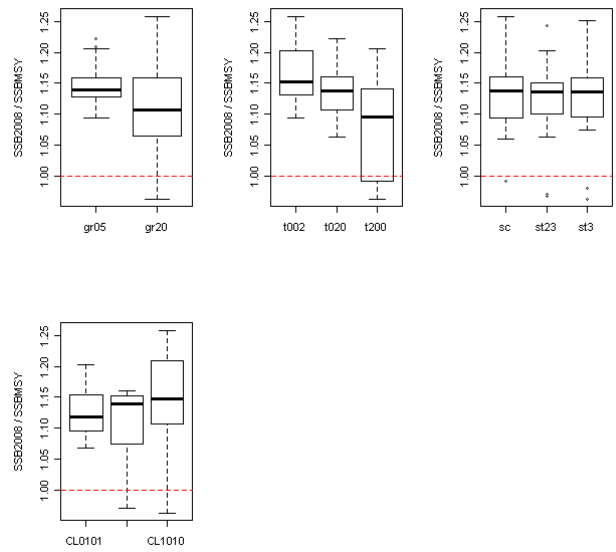


Figure 25. Summary of current biomass status estimates from the 54 models from grid B, partitioned by the different assumptions.

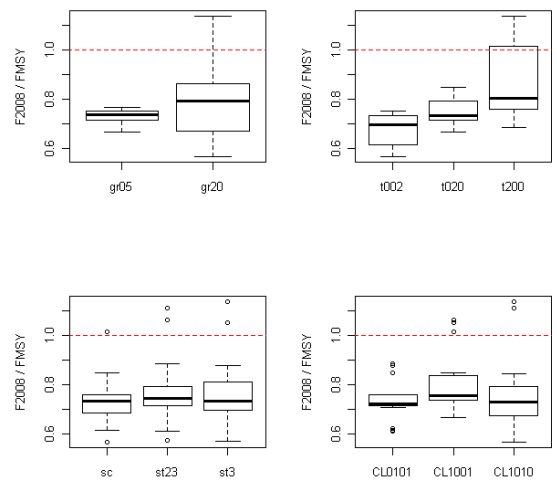


Figure 26. Summary of current fishing mortality estimates from the 54 models from grid B, partitioned by the different assumptions.

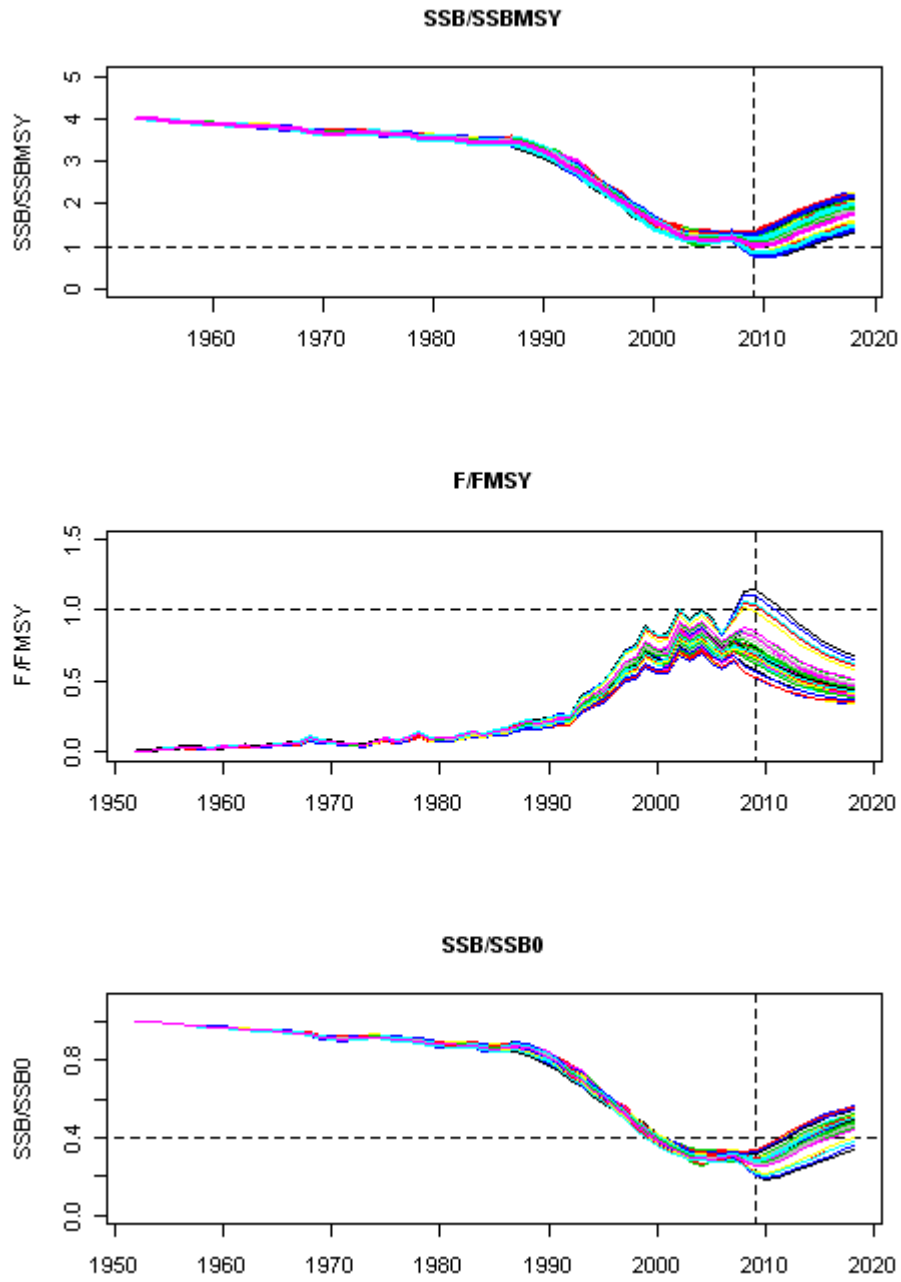


Figure 27. MPD biomass and fishing mortality time series estimates for the 54 models defined in grid B (including 10 years of constant catch projections at C(2008)).

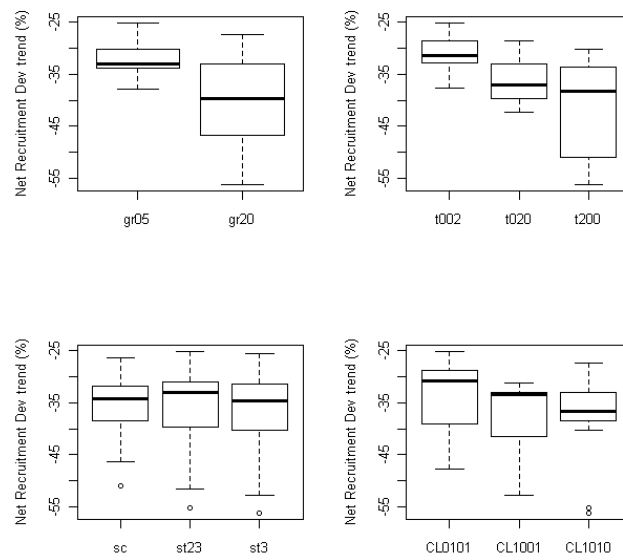


Figure 28. Summary of the recruitment deviation trend from the 54 models from gridB, partitioned by the different assumptions.

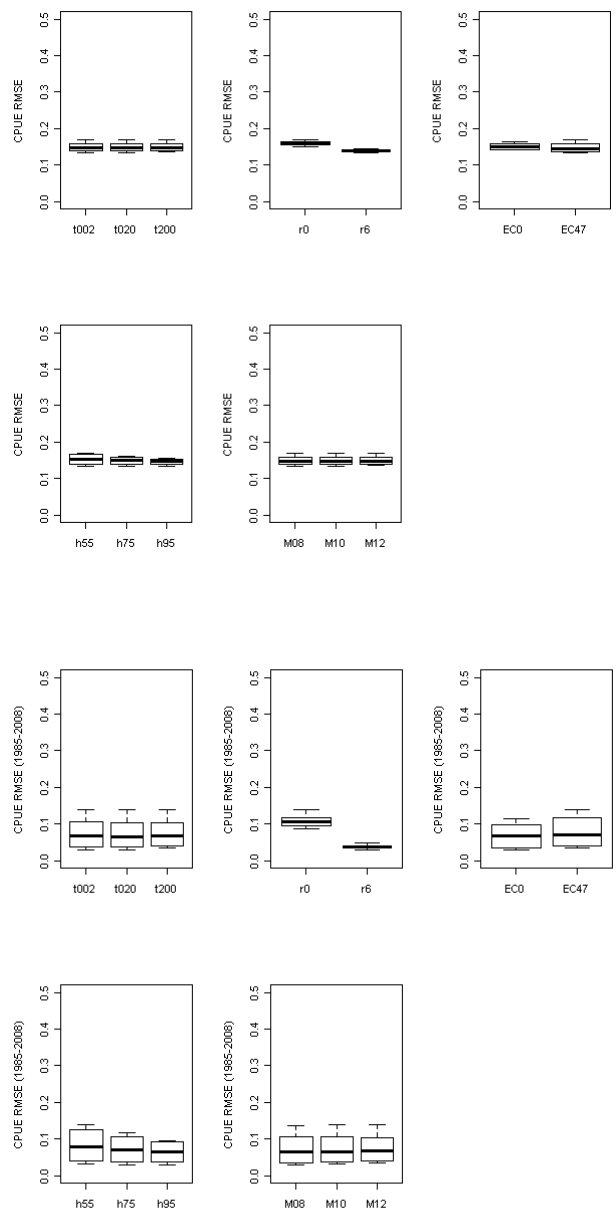


Figure 29. Summary of model fit to the CPUE series for the 108 models from grid C, partitioned by the different assumptions. Top 5 panels are the whole time series (1960-2008), bottom 5 panels are only the year with recruitment deviations estimated (1985-2008).

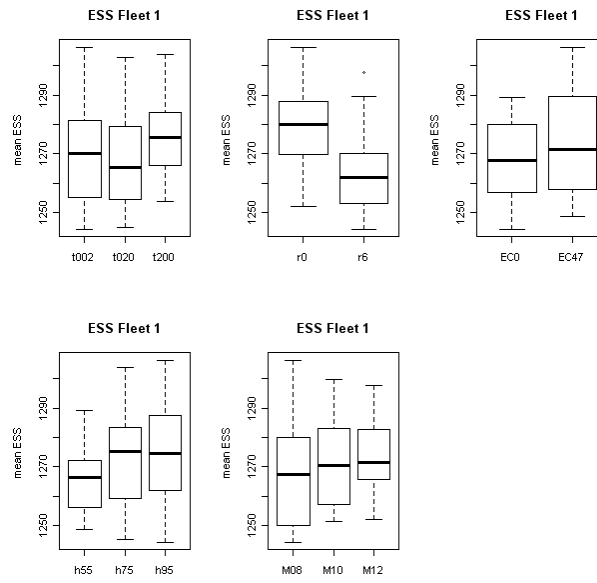


Figure 30. Summary of model fit to the LL fleet size composition data for the 108 models from grid C, partitioned by the different assumptions.

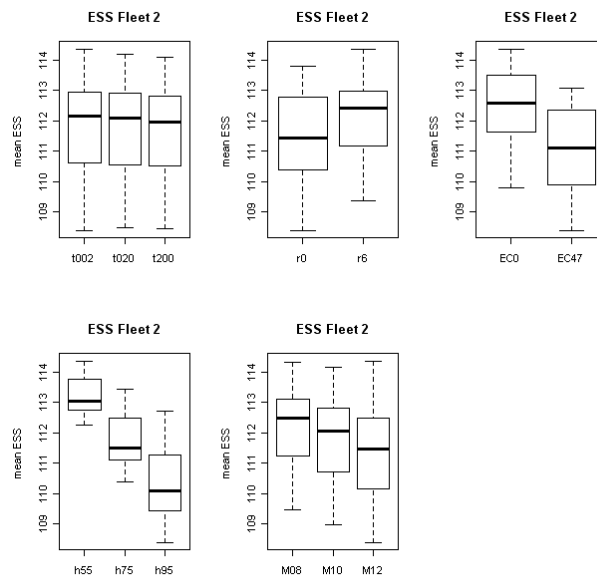


Figure 31. Summary of model fit to the PSFS fleet size composition data for the 108 models from grid C, partitioned by the different assumptions.

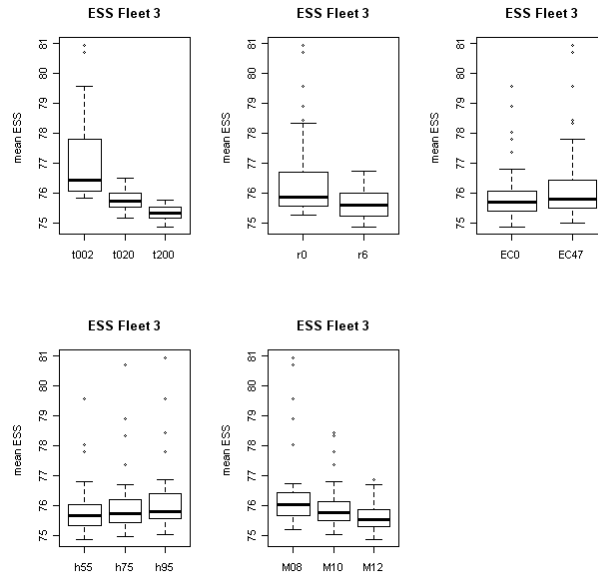


Figure 32. Summary of model fit to the PSLs fleet size composition data for the 108 models from grid C, partitioned by the different assumptions.

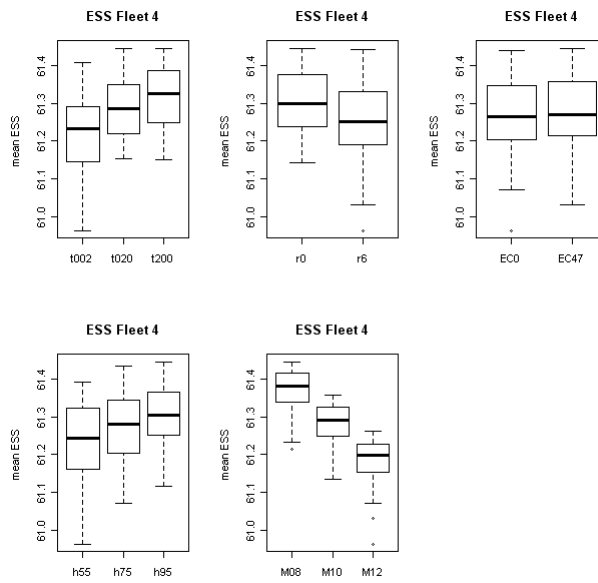


Figure 33. Summary of model fit to the 'Other' fleet size composition data for the 108 models from grid C, partitioned by the different assumptions.

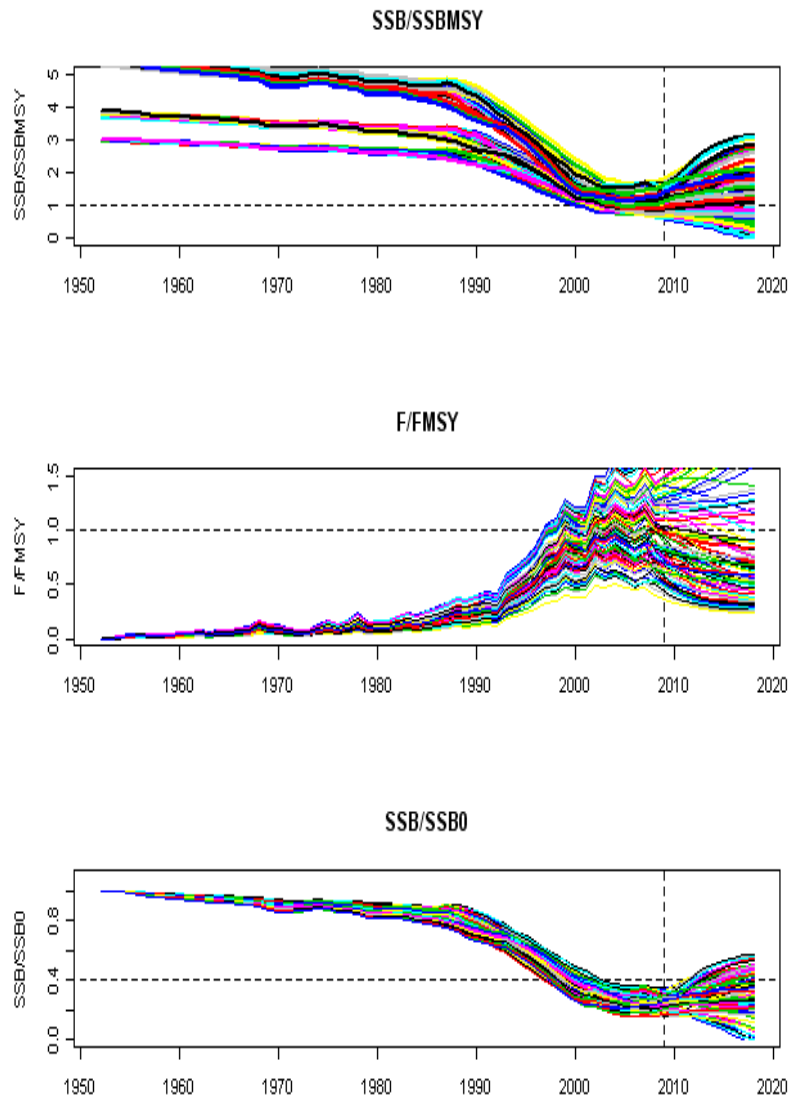


Figure 34. MPD time series of biomass and fishing mortality from the 108 models from grid C, including 10 years of constant catch, $C(2008)$, projections.

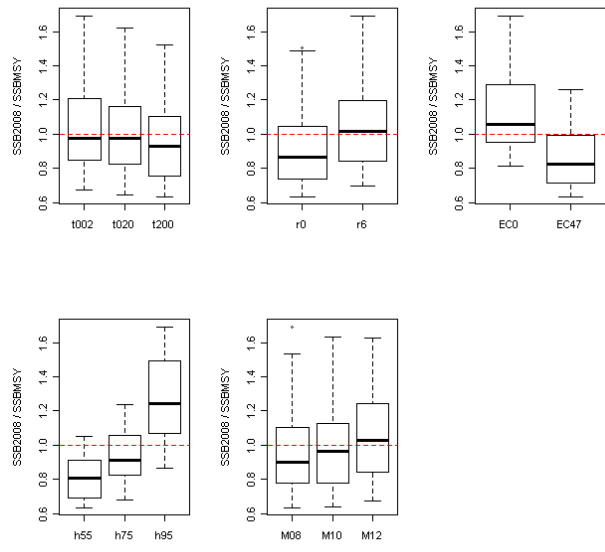


Figure 35. Summary of current (2008) spawning stock biomass relative to BMSY from the 108 models from grid C, partitioned by the different assumptions.

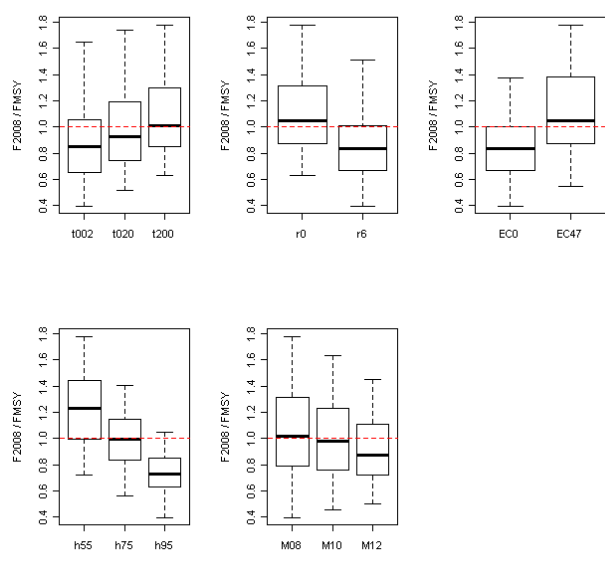


Figure 36. Summary of the fishing mortality relative to FMSY from the 108 models from grid C, partitioned by the different assumptions.

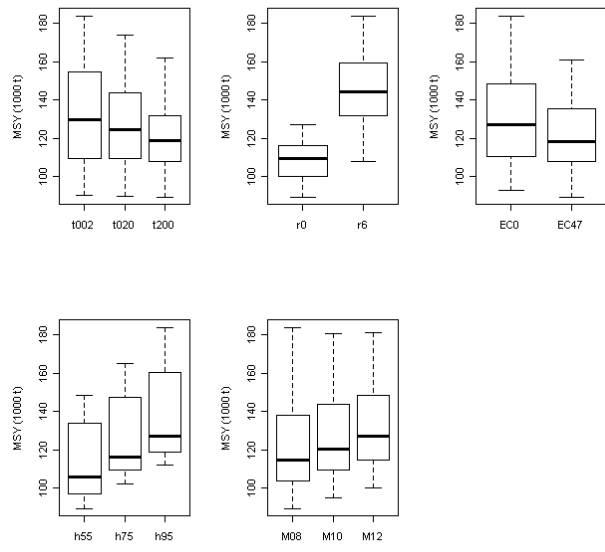


Figure 37. Summary of MSY estimates from the 108 models from grid C, partitioned by the different assumptions.

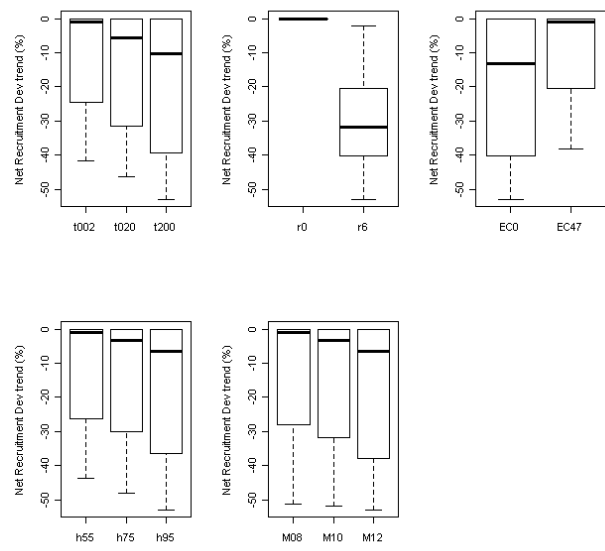


Figure 38. Summary of the recruitment deviation trend from the 108 models from grid C, partitioned by the different assumptions.

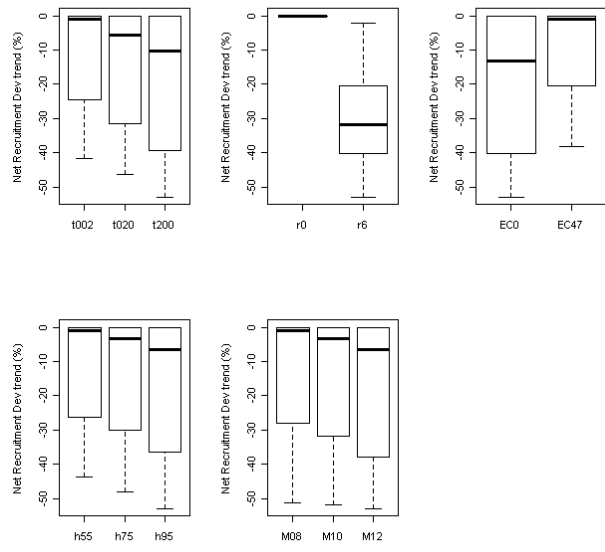


Figure 39. Summary of the recruitment deviation trend from the 108 models from grid C, partitioned by the different assumptions.

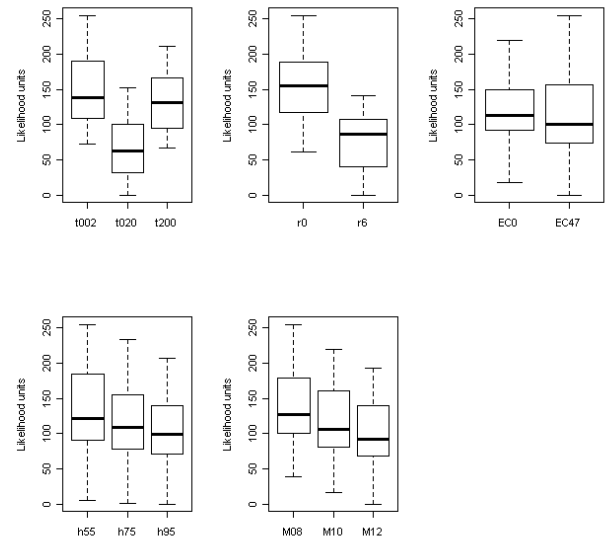


Figure 40. Summary of the relative likelihood values from the 108 models from grid C, partitioned by the different assumptions. Note that not all likelihoods are directly comparable.

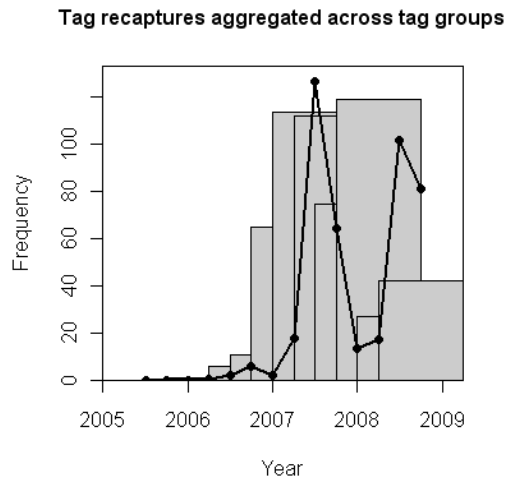
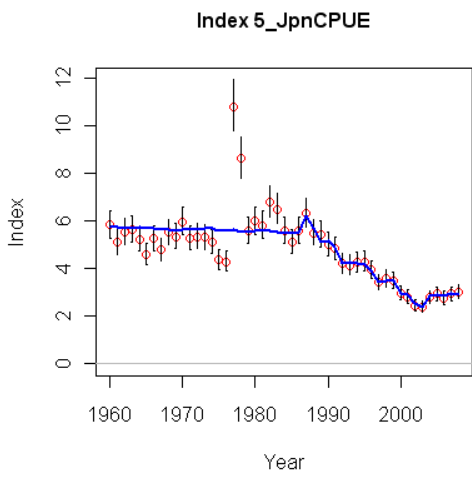


Figure 41. Left panel: Predicted (line) and observed (points) CPUE for the optimistic model maxMSY. Right panel: Predicted (line) and observed (bars) tag recoveries for the PSLF fishery.

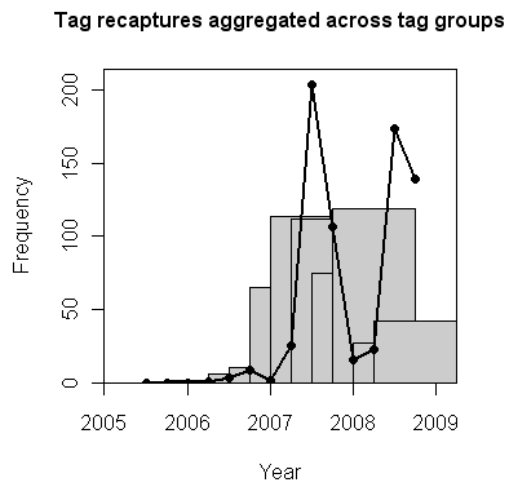
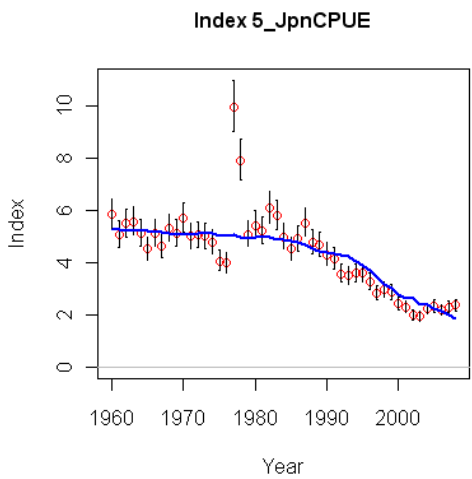


Figure 42. Left panel: Predicted (line) and observed (points) CPUE for the pessimistic model minMSY. Right panel: Predicted (line) and observed (bars) tag recoveries for the PSLF fishery.

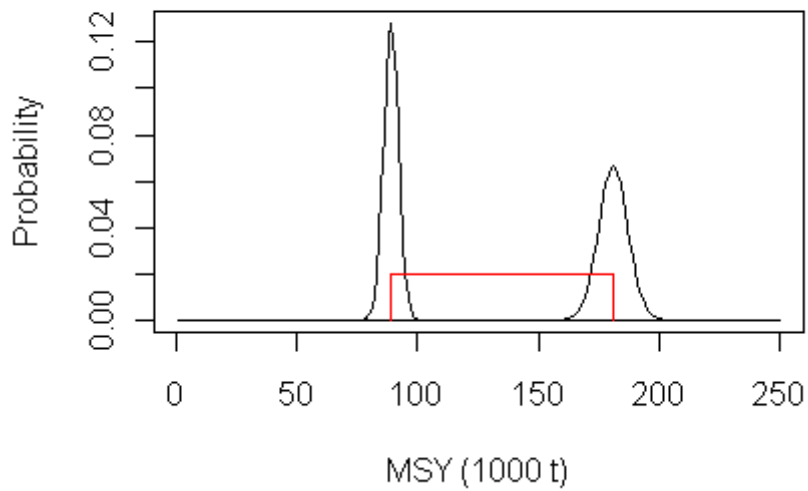


Figure 43. Conceptual illustration comparing statistical uncertainty (black lines, e.g. roughly corresponding to the Inverse Hessian multivariate normal estimates for models pes and opt) and the model selection uncertainty based on MPD modes (red line).

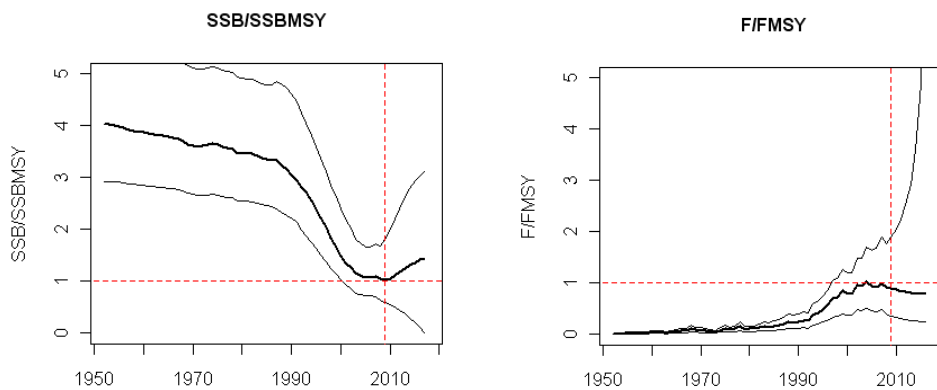


Figure 44. Time series of MSY reference point estimates from grid C, including 10 years of catch projections at 2008 levels. Thick black lines represent the weighted mean value across all 108 models. Thin lines represent the most extreme MPD value from the 108 models.

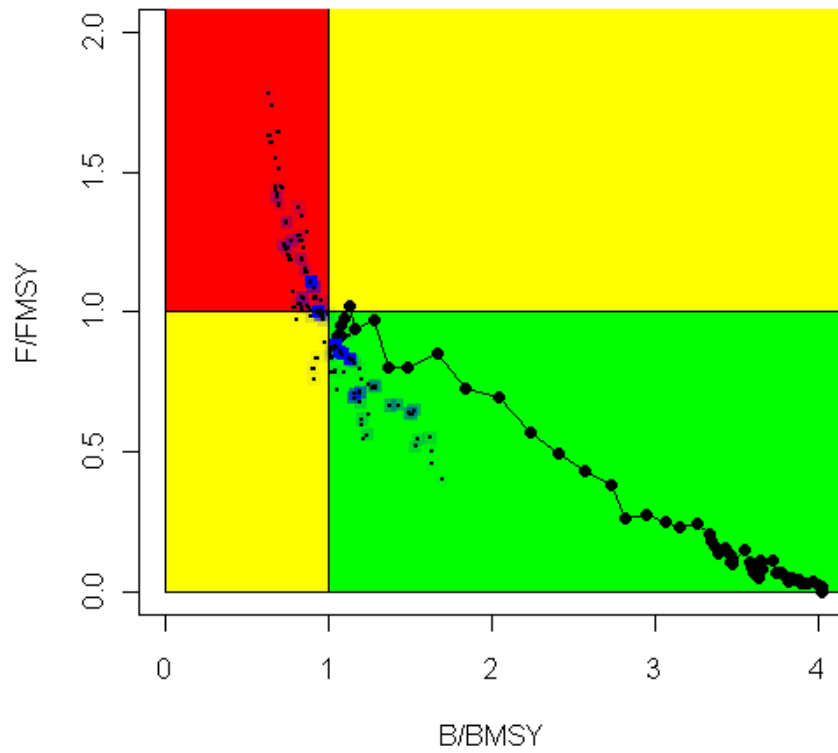


Figure 45. Kobe plot for grid C, using the authors proposed weighting scheme for model assumptions. Black circles represent the weighted mean values for each year. Blue squares indicate the MPD estimates for 2008 corresponding to each individual grid C model, with colour density proportional to the weighting (small black points indicate individual models which might not otherwise be visible).

Attachment 1: Update of the Indian Ocean Bigeye Tuna Stock Assessment, including the 2009 Data, and Revised Growth and Tag Recovery Assumptions

Extended Summary

This attachment briefly summarizes the update to the Indian Ocean BET assessment as conducted during the WPTT 2010. The methods are essentially the same as in the main text, however, these results supersede the main text due to the following revisions:

- New catch data series have been employed (1952-2009), largely unchanged except for the addition of 2009.
- New size composition data series have been employed (1952-2009). Attempts were made to derive a time series of effective sample size for the PS fleets, based on some assumptions of the sampling procedure (20% of fish sampled in a well, of which 5% are BET). When this was done (including a further arbitrary downweighting by a factor of 10 for all strata), the sample size estimates were consistently extremely large (or 0) for all fleets. Thus the upper limit of 200 was retained as described in the main text (and subjected to further downweighting in many cases). However, about 30% of the original observations were removed, as they were the result of space/time substitutions in the 2009 data series.
- The age-length relationship was reviewed and revised to match the Eveson and Million (2008) curve derived from the tagging data. However, SS3 has limited capacity to describe the inflection around age 2 in this curve. This was partially accounted for in the model by aggregating the smallest size composition (<50cm) observations and predictions into a single bin. This removes some of the information from the size composition data, but this is preferable to using a growth curve that is known to be grossly incorrect.
- The length-based maturity vector was retained from the original text, but given the revision to the growth curve, the age of 50% maturity is older (~4 years).
- New standardized CPUE series have been employed (1970-2009, Figure A1), largely unchanged except for the addition of 2009 (IOTC-2009-WPTT-29). Notably, the trend over the final 5 years has actually reversed from upward to downward, but the distinction is probably not meaningful relative to the general level of noise in the series.
- The treatment of the tag recovery data (including recoveries from 2009) has been revised to account for several biases in the estimated number of recaptures (and the revised growth equation).
- Revised management advice is provided in a similar format to the main text. The WPTT 2010 was encouraged to review the proposed plausibility weightings for the different assessment options, but no specific changes were suggested.
- A template of the SS3 Control.SS file is included in Attachment 3 to resolve any ambiguities in model specification. Data and control files are archived with the IOTC secretariat. Table 4 summarizes the main assumption interactions explored.

Revision of age-length relationship

The von Bertalanffy growth equation used in the main text was compared with the function derived from the 2008 analysis of tagging data (Eveson and Million 2008, IOTC-2008-WPTT-09), and found to be considerably different (Figure 47). The Richards model was adopted to achieve a considerably better, though still compromised fit (SS3 has limited growth options based on simple parameteric growth equations). To minimize the effects of the growth curve deviation for ages younger than 2, all length observations smaller than 50cm were aggregated in a single bin. The Eveson-Million growth curve was used for estimating ages of tagged fish. There remains an inconsistency in the mass-at-age for the youngest ages that could not be addressed, but this is expected to have a trivial influence on the model overall.

Revision of tag data

The following points were addressed in the revised treatment of the tagging data:

- Tag recovery data were included up to 2009
- The tag release ages were re-estimated on the basis of the Eveson-Million (2008) growth curve.
- The (instantaneous annual) tag shedding rate was assumed to be 0.024, with no initial shedding, as estimated by Gaertner and Hallier (2008).
- The EU PS tag recoveries of unknown set type were re-distributed according to the total proportion of known FS and LS set types in the PSFS and PSLs fisheries (by quarter).
- A correction term was applied to account for recaptures in the PSLs fleet from EU vessels that did not land in the Seychelles (the dominant port for landing and the only port with reliable tag reporting estimates). The correction was derived from the aggregate across 2005-2009.
- A correction term was applied to account for recaptures in the PSLs fleet from non-EU vessels (these vessels do not land in the Seychelles). The correction was derived from the aggregate across 2005-2009.
- The small number of tags with unknown release length were removed from the analysis.
- Tag recoveries from the assessment model non-mixing period were retained in the recovery dataset (and inflated according to the assumptions above). These tag recoveries are removed from the tagged population, but are not fit in the objective function.

The estimate for the total PSLs recaptures (where PSLs has the spatial and nation definition used in the assessment model) was calculated:

$$\hat{R} = \frac{1}{\hat{P}_{EU}} \left[\frac{1}{\hat{r}^{sea}} (R_{LS}^{sea} + \hat{P}_{LS}^{sea} R_{unk}^{sea}) + \frac{1}{\hat{P}_{EU}^{SEZ}} \left(\frac{1}{\hat{r}^{SEZ}} (R_{LS}^{SEZ} + \hat{P}_{LS}^{SEZ} R_{unk}^{SEZ}) \right) \right]$$

where subscripts indicate fishery types (EU = European Union/Seychelles, LS=log set-type, unk=unknown set-type), and superscripts indicate recovery locations (sea = aboard the fishing vessel, SEZ = port of Seychelles), (time subscripts are omitted for readability):

\hat{R} =estimated number of recaptures for all components of the PSLs fishery (for a particular time period),

$\frac{1}{\hat{P}_{EU}}$ accounts for the proportion of the PSLs recaptures that are not from the EU fleet, where $P=0.86$, was estimated as the proportion of PSLs catch from EU relative to PSLs catch from all nations (2005-9). The average from individual quarters is similar at 0.87, with a range from 0.57 – 1.04 (the value can exceed 1.0 because of errors in conversion factors required to estimate total catches by nation)

$\frac{1}{\hat{r}^{sea}}$ accounts for the proportion of tags removed at sea, but not reported; r is assumed to be 1.0.

R_{LS}^{sea} represents the number of recoveries at sea of known set-type LS from the recovery database.

\hat{P}_{LS}^{sea} is the proportion of tags recovered at sea of unknown set-type which are actually of set-type LS, estimated as the proportion of tags of known set-type LS recoveries at sea of all known set-type recoveries at sea summed over all years (=0.94).

R_{unk}^{sea} is the number of recoveries from unknown set-types at sea from the recovery database.

$\frac{1}{\hat{P}_{EU}^{SEZ}}$ is the scaling factor to account for the EU PSLs recaptures not moving through the port of Seychelles, estimated by the mean of the quarterly proportions of EU PSLs catch landed in the Seychelles relative to the total EU PSLs catch (2005-2008) = 0.85 (quarterly range of 0.57 – 0.97).

$\frac{1}{\hat{r}^{SEZ}}$ accounts for the unreported tags from the EU PS landings in the Seychelles, where r is the annual reporting rate estimated from tag seeding experiments (with 2009 assumed to be equal to the 2008 provisional estimate of 0.9).

R_{LS}^{SEZ} represents the EU PS recoveries in the Seychelles of known set-type LS from the recovery database.

\hat{P}_{unk}^{SEZ} is the proportion of tags recovered in the port of Seychelles of unknown set-type which are actually type LS, estimated from the proportion of PSLs tags of known set-type recovered in the port of Seychelles = 0.91.

R_{unk}^{SEZ} represents the EU PS recoveries in the port of Seychelles of unknown set-type from the recovery database.

The coastal fleets on the east coast of Africa, *i.e.* in Kenya and Zanzibar, have intercepted some tags near the release point, before they were fully mixed with the broader population, however reporting rates are largely unknown for these fleets. This is an unknown, but probably small number of tags that also potentially bias the tag inferences (and would likely lead to an underestimate of the fishing mortality).

Revisiting the interactions among the size sampling, selectivity and tag overdispersion assumptions

Similar to the main text, an expanded grid of 144 models was fit to examine the interactions among length-at-age variance, selectivity, size composition sample assumptions, and the influence of the tag data:

- 2 Length-at-age variance options (CV low, high)
- 8 Tag recovery options related to the incomplete mixing period and the negative binomial overdispersion parameter (τ). Note that $\tau = 70$ is consistent with recent WCPFC BET assessments (in which the input MULTIFAN-CL τ parameter is internally transformed by the addition of 50):
 - nt: tags not used
 - t002: $\tau = 2$, 4 quarters of incomplete mixing
 - t020: $\tau = 20$, 4 quarters of incomplete mixing
 - t070: $\tau = 70$, 4 quarters of incomplete mixing
 - t200: $\tau = 200$, 4 quarters of incomplete mixing
 - tODE: τ estimated ($\sim 140-150$ in cases examined), 4 quarters of incomplete mixing
 - M2t20: $\tau = 20$, 2 quarters of incomplete mixing
 - M2t70: $\tau = 70$, 2 quarters of incomplete mixing
- 3 PLS selectivity assumptions (stationary, annual deviates for tagging years 2005-8, or annual deviates 1985-2008)
- 3 Catch-at-Length likelihood weights (approximate input sample size) (LL 200, PLS 200), (LL 200, PLS 20), (LL 20, PLS 20); in all cases PSFS and Other fisheries sample size is 20.
- 1 recruitment deviate option, $SD(\log(\text{devs})) = 0.6$ (annual recruitment deviates were estimated from 1985-2008)
- 1 CPUE series option (no catchability trend), $sd(\log(\text{devs})) = 0.05$
- 1 stock-recruitment steepness option ($h = 0.75$)
- 1 natural mortality vector $M(\text{age } 2y+) = 0.4$.

Figure 48 - Figure 51 show the summary diagnostics of the fit to the CPUE, LL size composition and PLS size composition, and the MSY estimates. Conclusions are similar to those in the main text. Different combinations of these assumptions did not result in substantial differences in the fit to the CPUE or size composition data, except that the broad CV on length-at-age was probably preferable to the narrow CV (a large increase in the input effective sample sizes had only a trivial improvement in the quality of fit to the size composition data). Assumptions related to the tagging data had the greatest influence on the stock productivity (MSY) estimates. Notably (and perhaps counter-intuitively) removal of the tagging data and the assumption of low overdispersion ($\tau = 2$), resulted in similar, optimistic stock status interpretations, while intermediate-high overdispersion ($\tau = 20-200$) resulted in more pessimistic stock status estimates. This emphasizes that there is a fundamental qualitative difference between down-weighting the tagging data (*i.e.* assuming a reduced number of effective releases, or adding a co-efficient to the tag likelihood term) and increasing the overdispersion parameter. The highest overdispersion values were dropped from further consideration (*i.e.* $\tau > 70$), because these values assign a very high probability density to recaptures of 0. The lowest rate of overdispersion ($\tau = 2$) was also dropped because the quality of fit to the tags was never sufficient to suggest that the process approximates the ideal Poisson process. Additional consideration is warranted on how to fit and evaluate the fit to the tagging data, and the appropriate mixing period. There was also some variability in the fit to the PLS size composition data in the most recent years. This also seems to be related to conflicts with the tagging data, and merits further attention.

A representative range of these assumptions (most importantly spanning the range of plausible tag options) was carried forward into the stock status grid (below).

The Updated Stock Status Grid

As in the main text, the stock status grid of 288 models was intended to quantify the effects and interactions among a number of other fundamental life history assumptions that are known to be important and difficult to estimate. The grid was similar to the main text, except that more options were included with respect to the tagging and size composition assumptions:

- 1 Length-at-age option (CV high)
- 4 Tag recovery options related to the incomplete mixing period and the negative binomial overdispersion parameter (τ):
 - nt: tags not used
 - t020: $\tau = 20$, 4 quarters of incomplete mixing
 - t070: $\tau = 70$, 4 quarters of incomplete mixing
 - t70M2: $\tau = 70$, 2 quarters of incomplete mixing
- 1 PLS selectivity assumption (stationary)
- 2 recruitment deviate option ($SD(\log(\text{devs})) = 0.6, 0$). The latter was intended primarily to test whether year class strength was being strongly influenced by stationary selectivity assumptions, and to see if the model was using the flexibility in the poorly estimated recent recruits to reduce some other conflict.
- 2 CPUE options (no catchability trend, catchability trend of 0.47%/y)
- 3 steepness option ($h = 0.55, 0.75, 0.95$)
- 3 M vectors: $M(\text{age } 2y+) = 0.32, 0.4, 0.48$
- 2 Catch-at-Length input sample size assumptions (LL 200, PLS 200) (LL 20, PLS 20). PSFS and Other fleets 20.

The maximum posterior Density (MPD) estimates from these models indicate a broad range of uncertainty, including many plausible scenarios in which the B_{MSY} and F_{MSY} reference points were exceeded. The more pessimistic interpretations were generally associated with the CPUE catchability trend, low stock-recruit steepness, low M, and full tag mixing after 2 quarters. The likelihood of some of these assumptions cannot be compared within the context of the model (*e.g.* catchability trends, CL weighting, tag mixing periods). Some parameters might be estimated in principle (*e.g.* M, steepness, growth), however experience with other stocks and simulations suggests that these estimates are frequently very poor. For example, in these models, M estimates were doubtful because the youngest ages tended to have the lowest M; estimation of the variance on the length-at-age relationship resulted in the inference that a substantial portion of the fish shrink with age.

Summary diagnostics from the stock status grid are presented in Figure 52 - Figure 56. The fit to the CPUE and size composition data is similar to that described previously. Figure 55 describes the slope of the recruitment deviate trend for the range of models, and suggests that i) as in the main text, there remains a trend in the recruitment deviates (if they are estimated), which may suggest a systematic lack of fit to the stock recruitment relationship, and ii) the magnitude of the trend is most

influenced by the tagging and CPUE catchability assumptions. Qualitatively, it is evident that most of this trend is driven by the anomalously low deviates in the most recent (poorly estimated) years (e.g. Figure 63), and the recruitment behaviour is generally better than the original analysis described in the main text.

MSY-related reference points are partitioned by model assumption in Figure 56 - Figure 58. Most of the assumptions have a non-trivial influence on the stock status estimates, with stock-recruit steepness being the most influential. Table 5 indicates the subjective, but transparent, scheme that was used to weight the results of the 288 models in the stock status grid (i.e. Bayesian posteriors equal to the priors), to provide a synthesis of results. WPTT participants were encouraged to provide additional insight to revise the weightings, but no specific changes were proposed at the meeting. The median, 5th and 95th percentiles of the biomass and fishing mortality time series are shown in Figure 59, along with the most extreme of the Maximum Posterior Density (MPD) estimates. Key reference points are summarized in Table 6, and a Kobe plot is included in Figure 60.

A Kobe-2 Strategy Matrix (management options decision table) is presented in Table 7, based on the following assumptions:

- 10 years with deterministic recruitment (from the stock recruitment relationship).
- Constant catch with proportions distributed among fleets as in 2009.
- 5 catch levels at Catch(2009) -40%, -20%, +0%, +20%, +40%.

There is a large degree of uncertainty in the stock status derived from this analysis, but this is not surprising given the uninformative 'one way trip' nature of the fishery to date, general uncertainty of life history parameters, and questionable validity of the PS stationary selectivity assumption.

Example models

Similar to the results in the main text, it was found that the different models fit the core features of most of the important data series reasonably well, as indexed by the summary diagnostics. However, since these diagnostics are not especially meaningful to those that are unfamiliar with them, or outside of the context of a specific set of models, more traditional, model-specific diagnostics are also provided to illustrate some key points (selected models are defined in Table 8). Fits to the CPUE series (Figure 61), and size composition data (Figure 62) are shown for two models with MSY estimates near the extremes of the observed range (biomass and fishing mortality time series are shown in Figure 65, illustrating that the management implications of the two models are very different). The fits to the data are very similar, typical of the other models in the grid, and qualitatively illustrate that the data seem to be almost equally consistent with very different stock status interpretations. The actual likelihood values may strongly differ for a visually similar quality of fit, however, as noted previously, not all models are directly comparable on the basis of the likelihoods. For the models that are comparable in principle, a comparison of the likelihoods is useful, but we do not have enough confidence in the statistical properties of these sorts of fisheries models to accept a literal interpretation of the likelihood values. The stock recruitment relationships are compared in Figure 63, and the problems noted in the original analysis seem to be reduced (though the low recent recruitment events are still questionable). Figure 64 illustrates that the selectivity functions estimated for the optimistic and pessimistic models are very similar to each other and those from the main text.

The fit to the tagging data was consistently disappointing across a broad range of models. Tag recovery fits are shown for the pessimistic model in Figure 66 (tags were not fit in the optimistic model). Tag fits under a range of other assumptions are compared in Figure 67. While the fit to the tagging data is qualitatively disappointing, it is not immediately obvious that the fit is worse than observed in other tuna fisheries. We note that there are two important considerations in comparing the fit to the tags between these models and typical MULTIFAN-CL results (*e.g.* Langley *et al.* 2010, IOTC-2010-WPTT-23). First, there are a large number of predicted and observed tags during the pre-mixing period that are uninformative in the models, but are fit perfectly, and included in the MULTIFAN-CL summary graphics. The effect of including the unmixed tags is shown for the pessimistic model in Figure 66 (and it is not obvious that it makes much difference in this case). Plotting only the fully mixed tags is clearly a more useful representation for the purposes of evaluating the quality of fit (though this may not be the only reason for plotting tags), and this is the convention adopted in Figure 67. Second, most MULTIFAN-CL tuna applications estimate catches with an assumed error distribution. This provides an extra degree of freedom that allows the model to partition the lack of fit to the tags between the tag recovery error and the catch prediction error (*e.g.* IOTC-2010-WPTT-23 illustrates YFT catch deviations of up to 40%, coinciding with individual tag recovery events in the PS fisheries). While it is certainly possible that there could be large catch errors, it seems more likely that the bulk of the problem is attributable to problems with the tagging assumptions, at least for the Indian Ocean PS fleets (note that catch deviations in later versions of the YFT assessment were more tightly constrained).

Given the sensitivity of the stock status to the tag results, it is worth investing some effort to ensure that the tag data are treated properly (and the uncertainty is illustrated) *i.e.* *i)* noting the distinct difference between downweighting the likelihood and increasing overdispersion, *ii)* investigating the appropriate lag for the mixing period (and evidence for incomplete mixing across the whole Indian Ocean), and *iii)* properly accounting for seasonal variability in some of the tag recapture pre-processing calculations.

Priorities for the future

One can easily improve the fit to the different data series by increasing the degrees of freedom in the model (*e.g.* adding non-stationary selectivity, spatial structure, *etc.*), however before doing this it is worth considering what one hopes to achieve. Is more flexibility going to result in the extraction of useful information from the available data, in a way that makes any difference to the management advice, or is it simply going to introduce more uncertain parameters that obfuscate the important issues and make it more difficult to represent the uncertainty? At this time, further tinkering with model structural assumptions is expected to be less productive than re-examining some of the fundamental inputs. The following is a proposed list for improvements, roughly in order of perceived priority:

1. The assessment is highly dependent on the standardized Japanese CPUE series as a relative abundance index. The Japanese LL fleet has had dramatic changes in spatial distribution of effort over time (in addition to other substantive operational changes), and it is doubtful that these effects are adequately accounted for in the standardization. There is a large literature on this problem in tuna fisheries, and a number of approaches that have been

employed elsewhere should be used in the Indian Ocean as well. It may not be possible to prove that a better CPUE series has been obtained, but it is easy to demonstrate a range of uncertainty associated with series derived from different plausible assumptions. This key source of uncertainty needs to be reflected in the assessment.

2. Fine-scale spatial analyses of the tagging, CPUE and size composition data should be undertaken to address a number of issues: *i*) homogeneity of fishery operations, (*i.e.* consistent operations are more likely to have stationary catchability/selectivity), *ii*) tag mixing rate assumptions (the assessment is sensitive to the choice of a 2 or 4 quarter mixing period – is there independent support for a better choice?), *iii*) homogeneity of the fish population (Is there any benefit to resolving the spatial sub-structure of the population? Do we actually have enough data and the right modelling tools to resolve the spatial processes?).
3. It might be possible to improve the tag recapture estimates by *i*) moving the non-EU catch from the PSLs (and PSFS) fishery to the ‘Other’ fishery, and *ii*) applying season-specific (rather than multiyear mean) correction factors for EU landings outside of the Seychelles.
4. Improved resolution of the growth curve for fish of age less than 2 years. This seems to require a modification to the SS3 software.
5. Increased resolution of temporal processes (*e.g.* treating quarters as years) might be useful for representing seasonal spawning and higher resolution selectivity issues (*e.g.* the very sharp mode of age 1 fish in the PSLs selectivity). However, this is most relevant for younger ages and probably not worth pursuing until the length-at-age representation can be improved for younger fish. Furthermore, the main advantage that might be gained from this improvement is only realized in the context of a separable selectivity assumption in the PSLs fleet. If the stationary selectivity assumption breaks down under closer scrutiny, then this approach is likely to introduce a heavy computational burden for no real benefit.
6. It may be possible to re-estimate the growth curve parameters, particularly $L(\infty)$, now that the RTTP tags have been at liberty for several years.
7. The assessment is conditional on the assumption that the catch data are known without error. This time series is obviously not perfect, and it is doubtful that one can estimate an improved catch series within the context of the model. However, there may be reason to believe that the species composition estimates from the artisanal fleets are biased toward under-reporting of BET. Alternative plausible estimates of the BET surface fishery catch should be explored to evaluate the total catch uncertainty. If these alternative series are substantially different, then this uncertainty should be represented in the assessment.

Table 4. Summary of assumption option abbreviations.

Assumption	Option
Tag dynamics p=mixing period τ=overdispersion	<i>nt</i> ; tags are not fit <i>t002</i> ; p=4, τ=2 <i>t020</i> ; p=4, τ=20 <i>t070</i> ; p=4, τ=70 <i>t200</i> ; p=4, τ=200 <i>t20M2</i> ; p=2, τ=20 <i>t70M2</i> ; p=2, τ=70 <i>tODE</i> ; p=4, τ=estimated (~140-150 in results examined)
Recruitment σ=sd(log(dev))	<i>R6</i> ; σ=0.6 <i>R0</i> ; σ=0 (deviates not estimated)
LL Catchability	<i>EC0</i> ; no trend <i>EC47</i> ; linear catchability increase at 0.47% / year
Beverton-Holt SR Steepness (h)	<i>h55</i> ; h=0.55 <i>h75</i> ; h=0.75 <i>h95</i> ; h=0.95
Natural Mortality Vector	<i>M08</i> ; 80% of Figure 9 ($M(a_{2+})=0.32$) <i>M10</i> ; Figure 9 $M(a_{2+})=0.40$ <i>M12</i> ; 120% of Figure 9 ($M(a_{2+})=0.48$)
Catch-at-Length maximum input N (by fleet)	<i>CL1010</i> ; LL N=200, PSLS N=200; (PSFS N=20, Other N=20) <i>CL1001</i> ; LL N=200, PSLS N=20; (PSFS N=20, Other N=20) <i>CL0101</i> ; LL N=20, PSLS N=20; (PSFS N=20, Other N=20)

Table 5. Weighting scheme for integrating the 288 models in the stock status grid.

Assumption	Option Weighting			
Tags (p =mixing period, τ =overdispersion)	<i>nt</i> no tags 0.5	<i>t020</i> $p=4, \tau=20$ 0.167	<i>t070</i> $p=4, \tau=70$ 0.167	<i>M2t70</i> $p=2, \tau=70$ 0.167
Recruitment $sd(\log(\text{dev}))$	<i>R0</i> $\sigma=0$ 0.5	<i>R6</i> $\sigma=0.6$ 0.5		
Catchability	<i>EC0</i> no trend 0.8	<i>EC47</i> q increasing 0.2		
SR Steepness	<i>h55</i> $h=0.55$ 0.1	<i>h75</i> $h=0.75$ 0.7	<i>h95</i> $h=0.95$ 0.2	
Natural Mortality	<i>M08</i> $M(a=2+)=0.32$ 0.33	<i>M10</i> $M(a=2+)=0.40$ 0.33	<i>M12</i> $M(a=2+)=0.48$ 0.33	
Catch-at-Length input N (LL, PSLs)	<i>CL1010</i> $N=200, 200$ 0.5	<i>CL0101</i> $N=20, 20$ 0.5		

Table 6. Stock status summary table. Percentiles are drawn from a cumulative frequency distribution of MPD values with models weighted as in Table 5.

Reference Point	median	5 th and 95 th percentiles	MPD range
SSB_{2009}/SSB_{MSY}	1.20	0.88 – 1.68	0.67 - 1.91
F_{2009}/F_{MSY}	0.79	0.50 – 1.22	0.40 - 1.79
MSY (1000 t)	114	95 – 183	81 - 214
SSB_{2009}/SSB_0	0.34	0.26 – 0.40	0.22 – 0.42
SSB_{2009} (1000 t)	381	236 - 762	184 - 1150

Table 7. Kobe 2 Strategy matrix derived from the the stock status grid with models weighted as in Table 5.

Stock status Reference Point	Projection Time frame	Weighted proportion of scenarios that violate the Reference Point				
		C(2009) -40%	C(2009) -20%	C(2009)	C(2009)+20%	C(2009)+40%
$P(B_t < B_{MSY})$	In 3 years	0.19	0.24	0.28	0.40	0.50
	In 10 years	0.19	0.24	0.30	0.55	0.73
$P(F_t > F_{MSY})$	In 3 years	<0.01	0.06	0.22	0.50	0.68
	In 10 years	<0.01	0.06	0.24	0.58	0.73

Table 8. Grid assumptions from the example models (other assumptions are identical to each other and the other models described in the text).

Assumption	Example Model		
	Optimistic	Pessimistic	Tag Fit Comparisons
Tags (p =mixing period, τ =overdispersion)	<i>nt</i> no tags	<i>t70M2</i> $p=2, \tau=70$	<i>t002</i> ; $p=4, \tau=2$ <i>t020</i> ; $p=4, \tau=20$ <i>t200</i> ; $p=4, \tau=200$ <i>t20M2</i> ; $p=2, \tau=20$ <i>tODE</i> ; $p=4, \tau$ =est. <i>nt</i> (no tags)
Recruitment $sd(\log(\text{dev}))$	<i>R6</i> $\sigma=0.6$	<i>R0</i> $\sigma=0$	<i>R6</i> $\sigma=0.6$
Catchability	<i>EC0</i> no trend	<i>EC47</i> q increasing	<i>EC0</i> no trend
SR Steepness	<i>h95</i> $h=0.95$	<i>h55</i> $h=0.55$	<i>h75</i> $h=0.75$
Natural Mortality	<i>M12</i> $M(a=2+)=0.48$	<i>M08</i> $M(a=2+)=0.32$	<i>M10</i> $M(a=2+)=0.40$
Catch-at-Length input N (LL, PSLS) (PSFS=20, Other=20)	<i>CL1010</i> $N=200, 200$	<i>CL0101</i> $N=20, 20$	<i>CL1010</i> $N=200, 200$

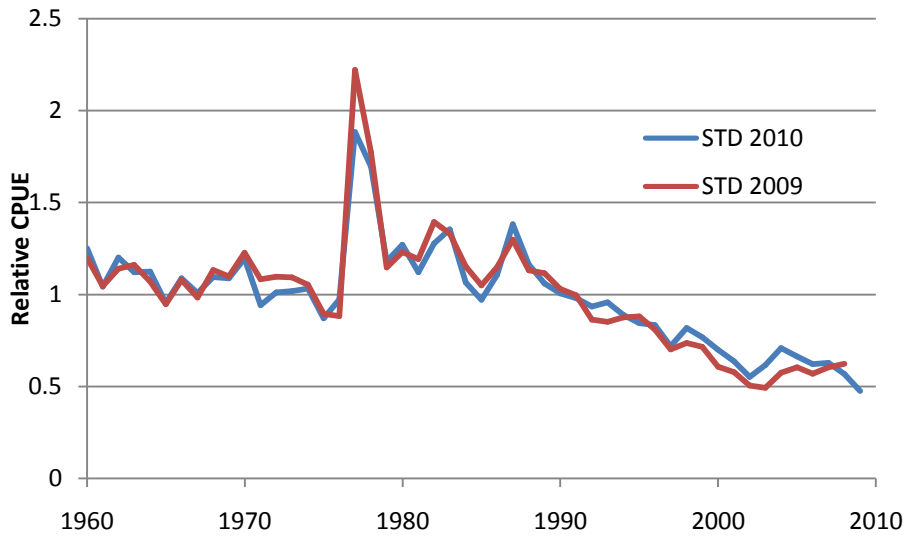


Figure 46. Comparison of standardized BET CPUE series calculated in 2009 and 2010.

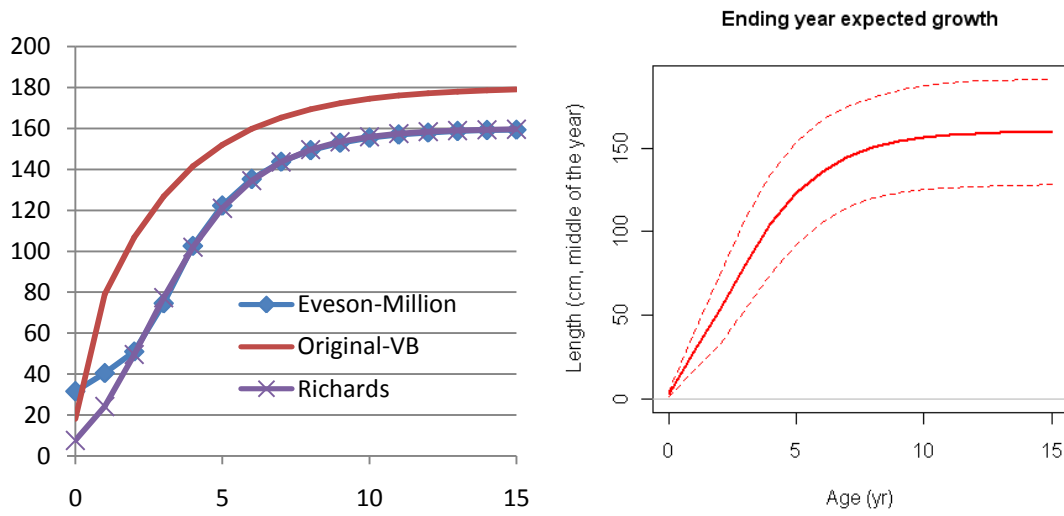


Figure 47. Comparison of length-at-age relationships. The Richards version replaces the Original-VB used in the main text. Right panel indicates the 95 %iles of the preferred (gr20) length-at-age distribution.

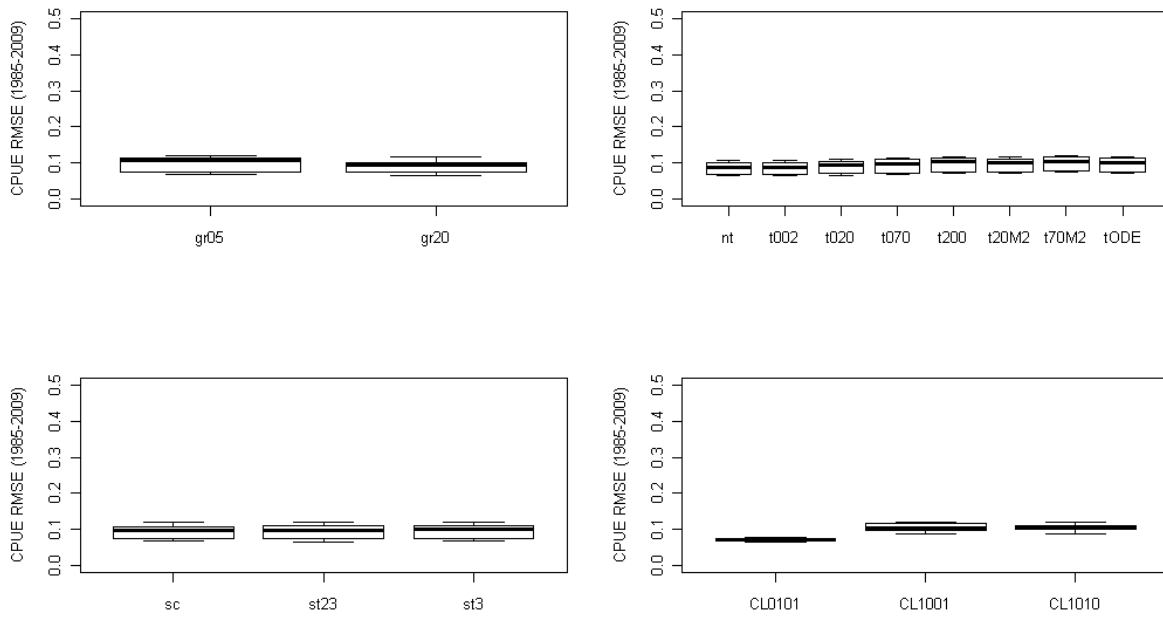


Figure 48. Summary of model fit to the CPUE series for the 144 models from the size/selectivity/tag interaction grid, partitioned by the different assumptions. Model abbreviations in Table 4.

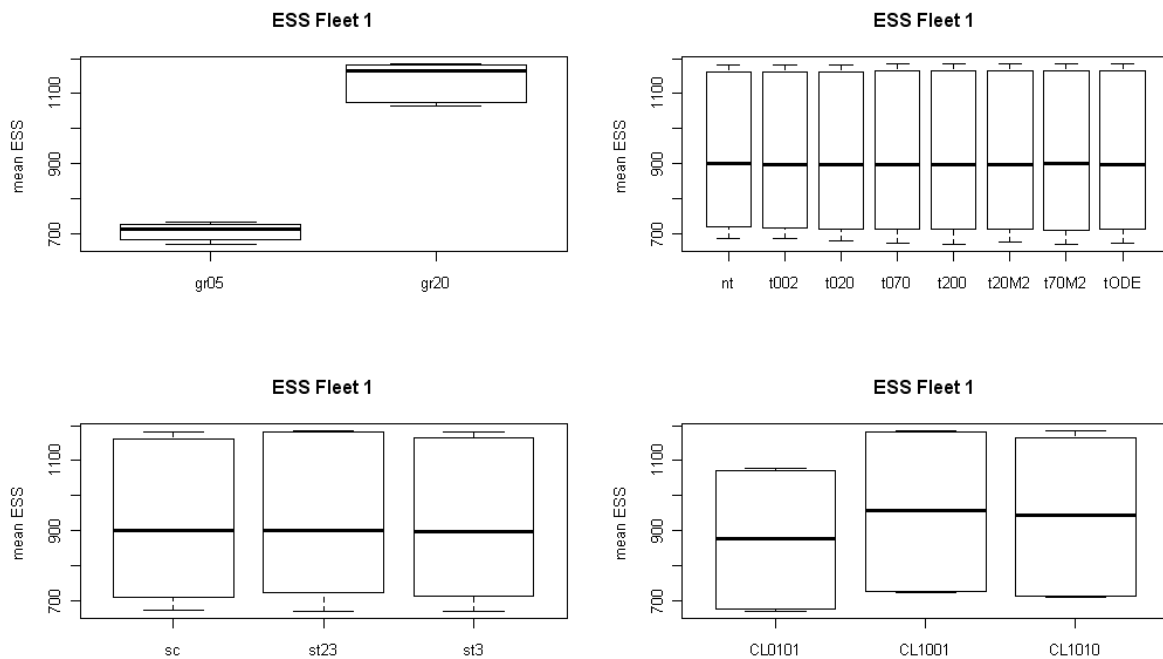


Figure 49. Summary of model fit to the LL size composition data for the 144 models from the size/selectivity/tag interaction grid, partitioned by the different assumptions.

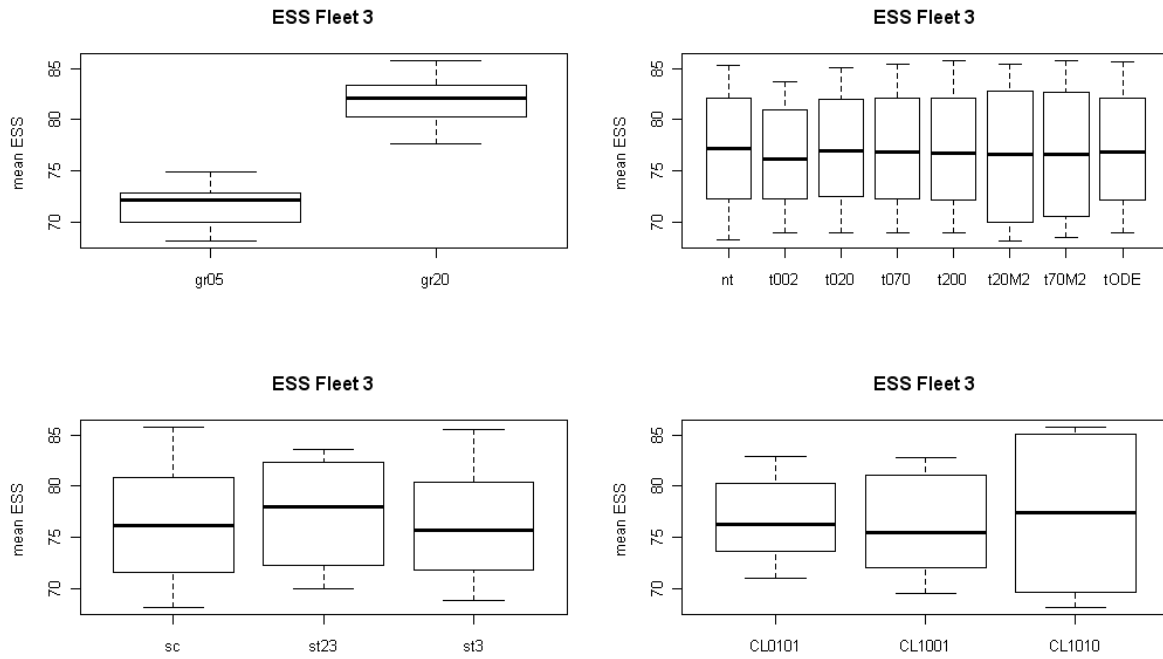


Figure 50. Summary of model fit to the PLS size composition data for the 144 models from the size/selectivity/tag interaction grid, partitioned by the different assumptions. Model abbreviations in Table 4.

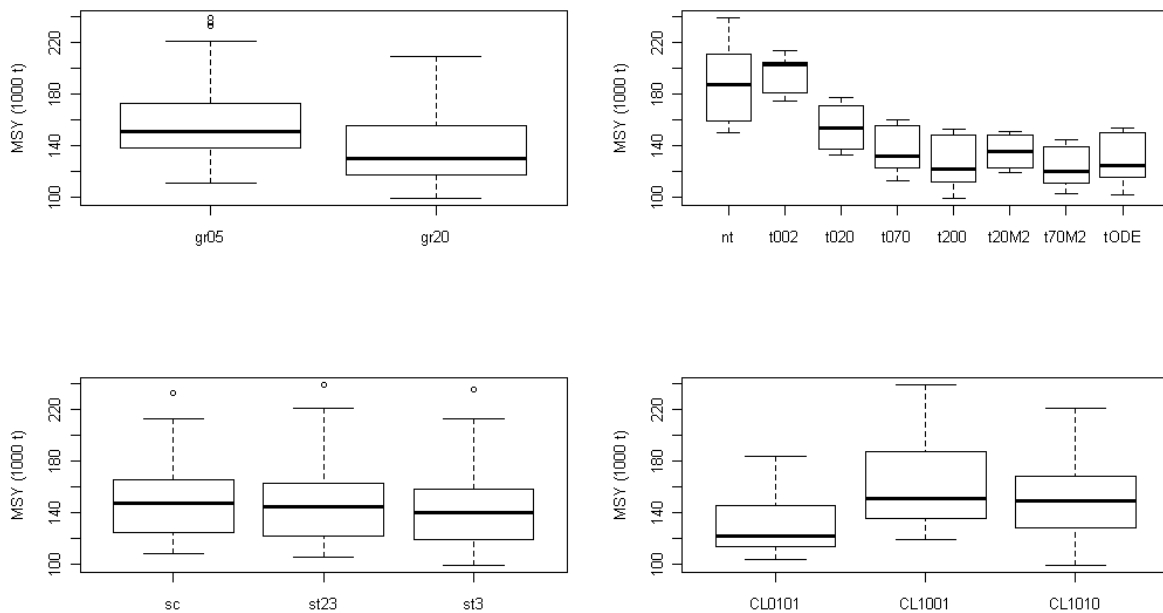


Figure 51. Summary of the MSY estimates for the 144 models from the size/selectivity/tag interaction grid, partitioned by the different assumptions. Model abbreviations in Table 4.

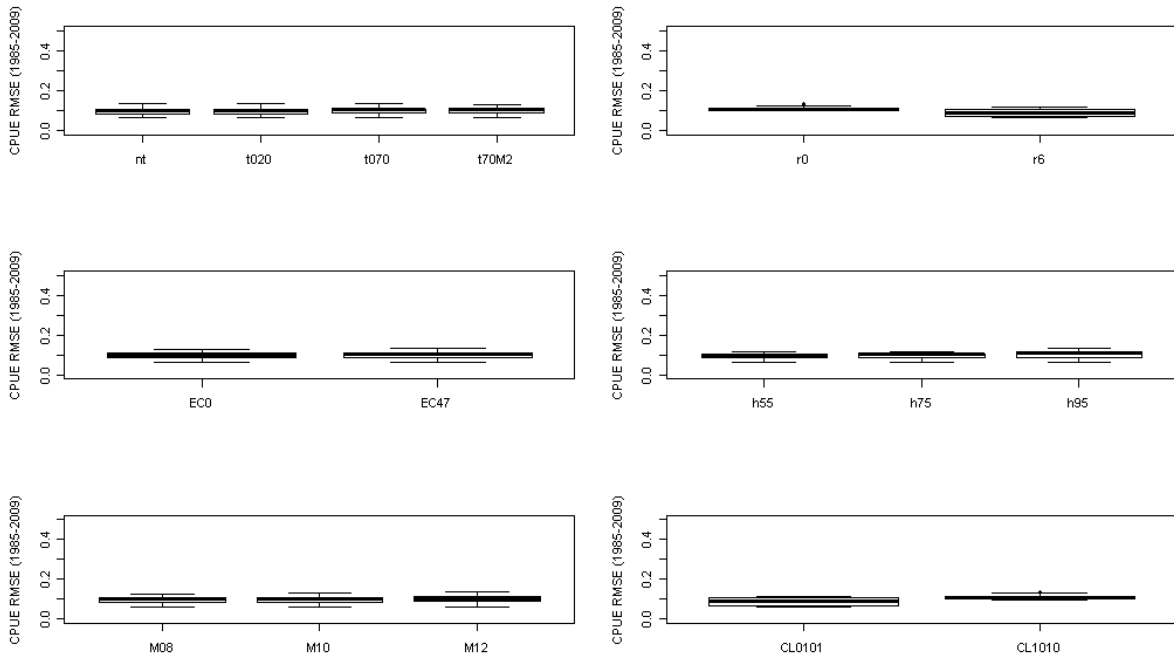


Figure 52. Summary of model fit to the CPUE series for the 288 models from the stock status grid, partitioned by the different assumptions. Model abbreviations in Table 4.

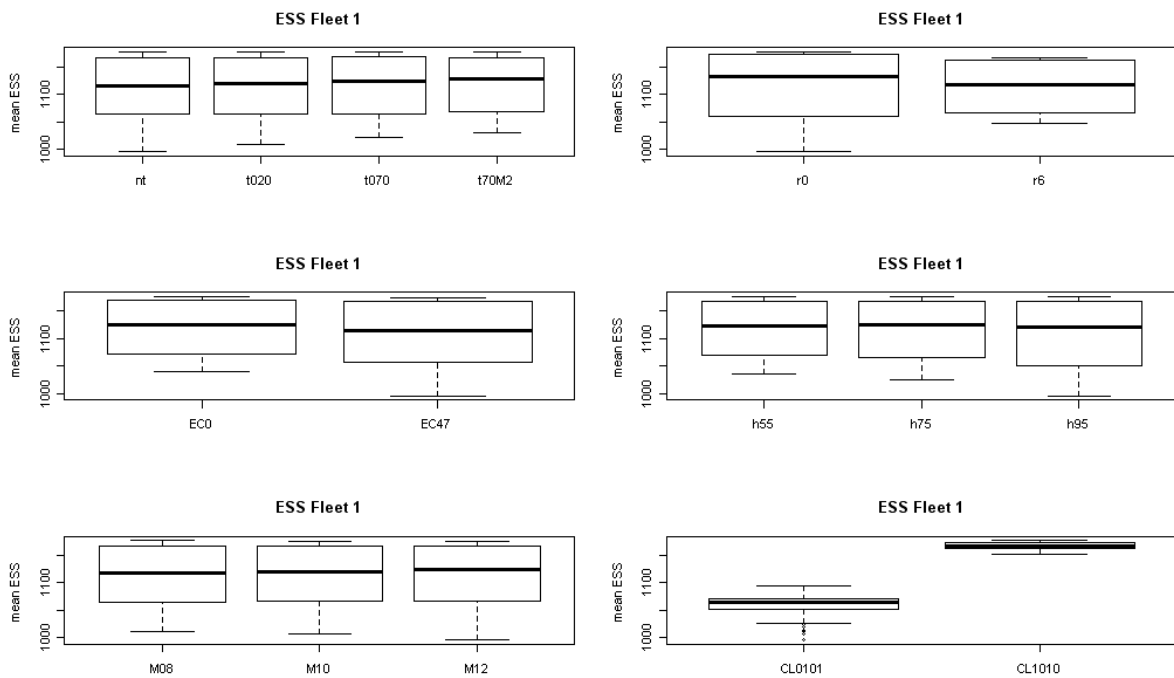


Figure 53. Summary of model fit to the LL size composition data for the 288 models from the stock status grid, partitioned by the different assumptions. Model abbreviations in Table 4.

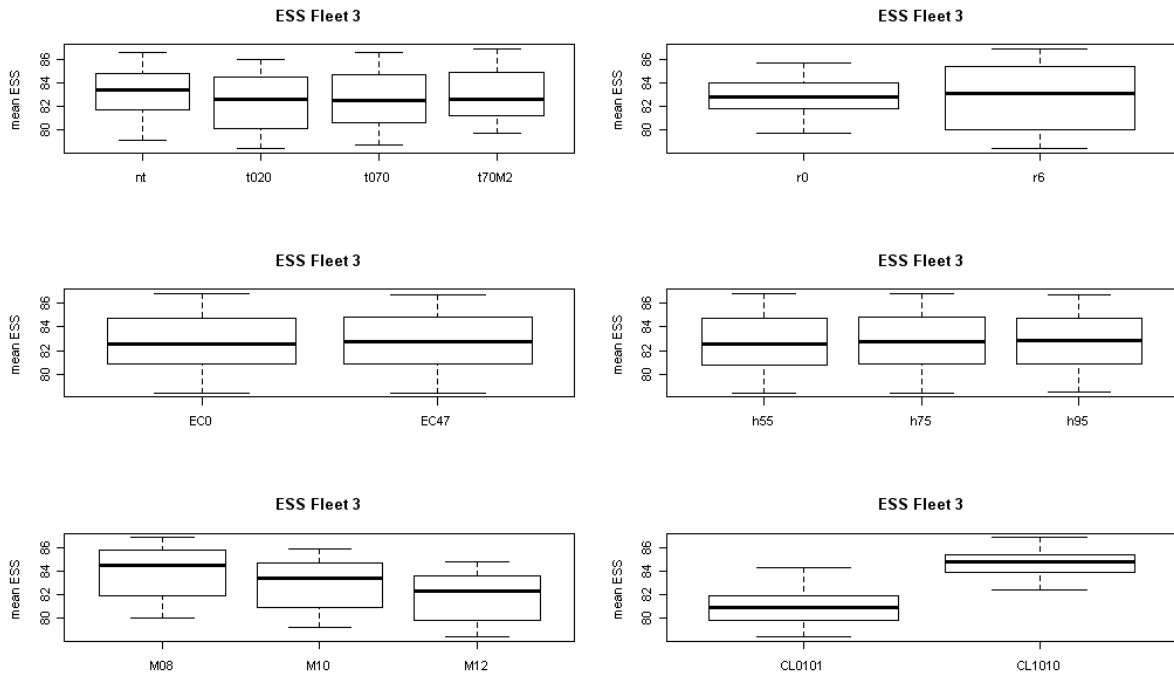


Figure 54. Summary of model fit to the PLS size composition data for the 288 models from the stock status grid, partitioned by the different assumptions. Model abbreviations in Table 4.

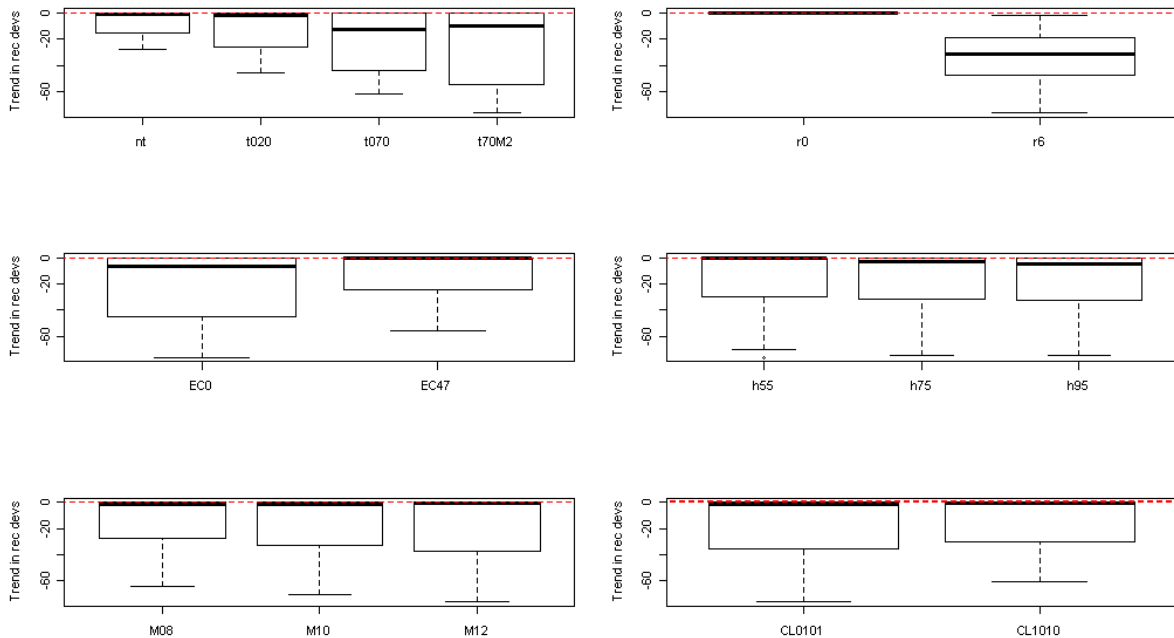


Figure 55. Summary of the linear trend (slope in %) in MPD recruitment deviates (1985-2009) for the 288 models from the stock status grid, partitioned by the different assumptions. Model abbreviations in Table 4.

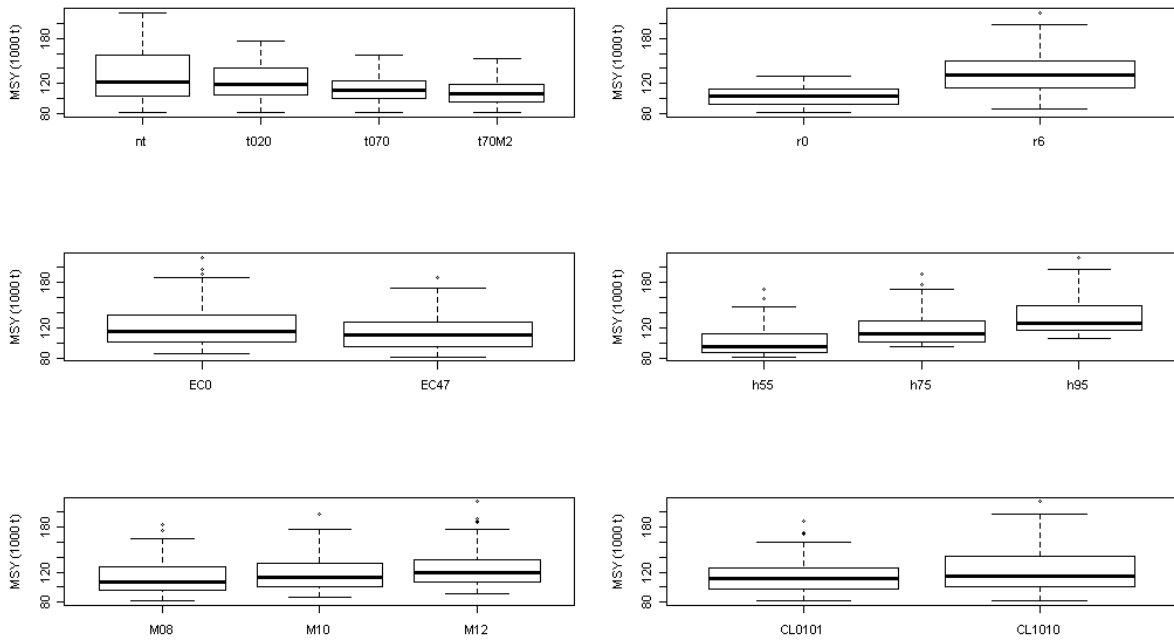


Figure 56. Summary of MSY MPD estimates for the 288 models from the stock status grid, partitioned by the different assumptions. Model abbreviations in Table 4.

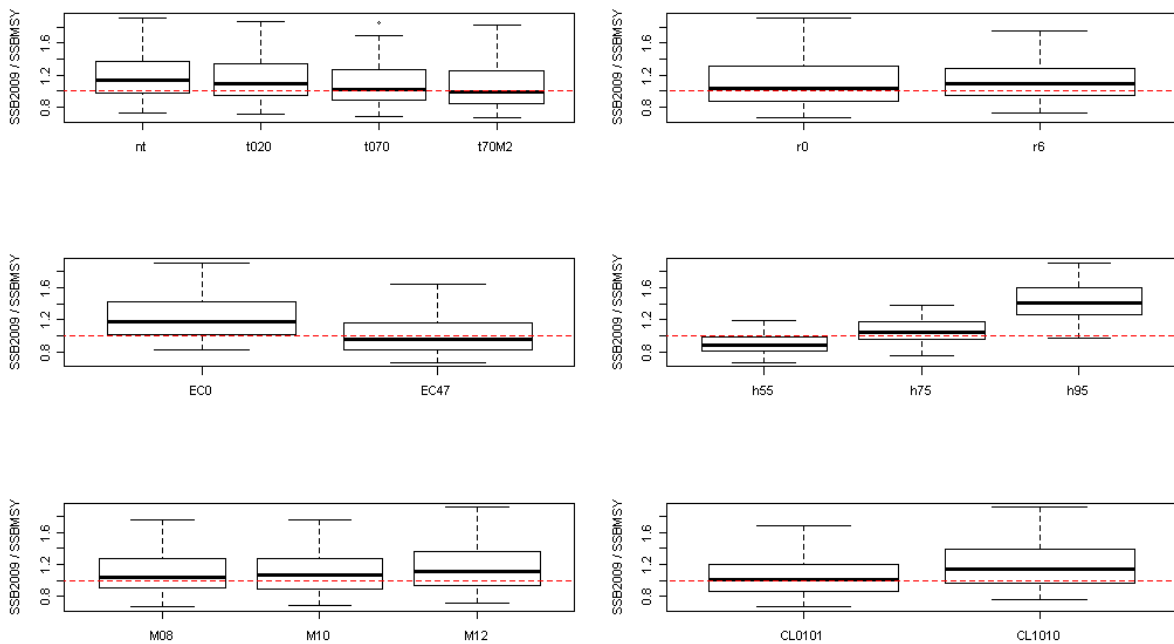


Figure 57. Summary of SSB(2009)/SSB(MSY) MPD estimates for the 288 models from the stock status grid, partitioned by the different assumptions. Model abbreviations in Table 4.

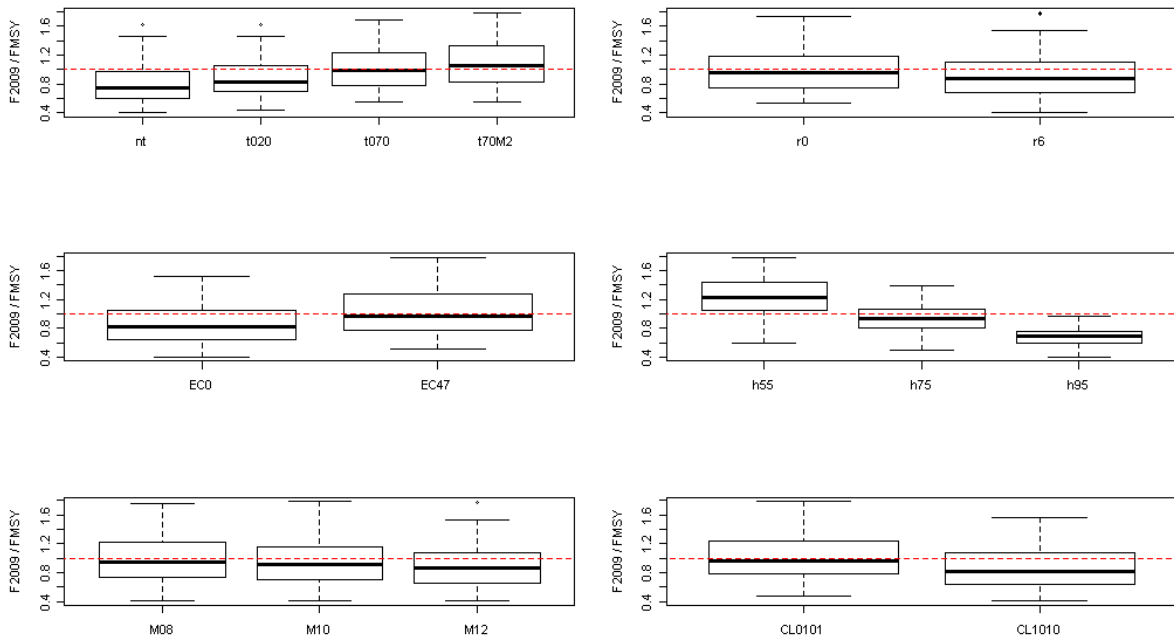


Figure 58. Summary of $F(2009)/F(MSY)$ MPD estimates for the 288 models from the stock status grid, partitioned by the different assumptions. Model abbreviations in Table 4.

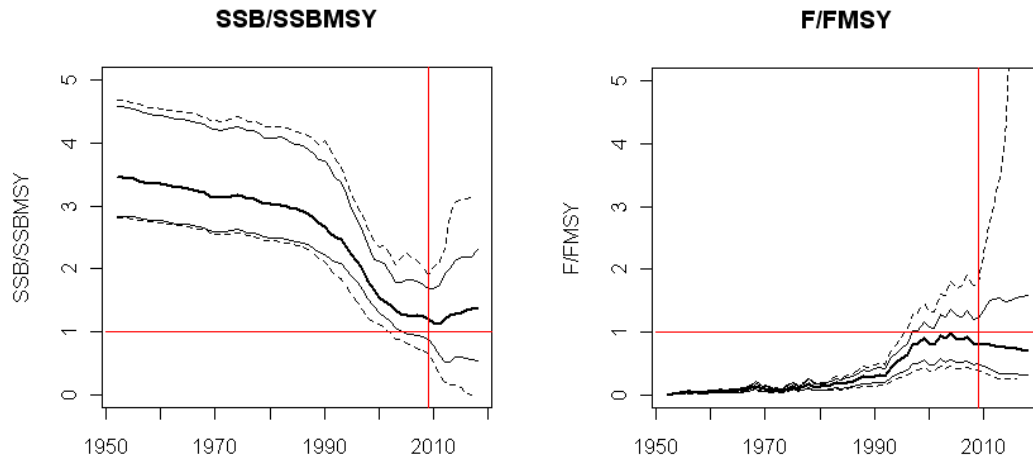


Figure 59. Time series of MSY reference point estimates from the stock status grid, including 10 years of catch projections at 2009 levels. Thick black lines represent the median MPD value from the weighted (Table 5) aggregate of 288 models. Thin black lines represent the 5th and 95th percentiles, and broken lines indicate the full MPD range.

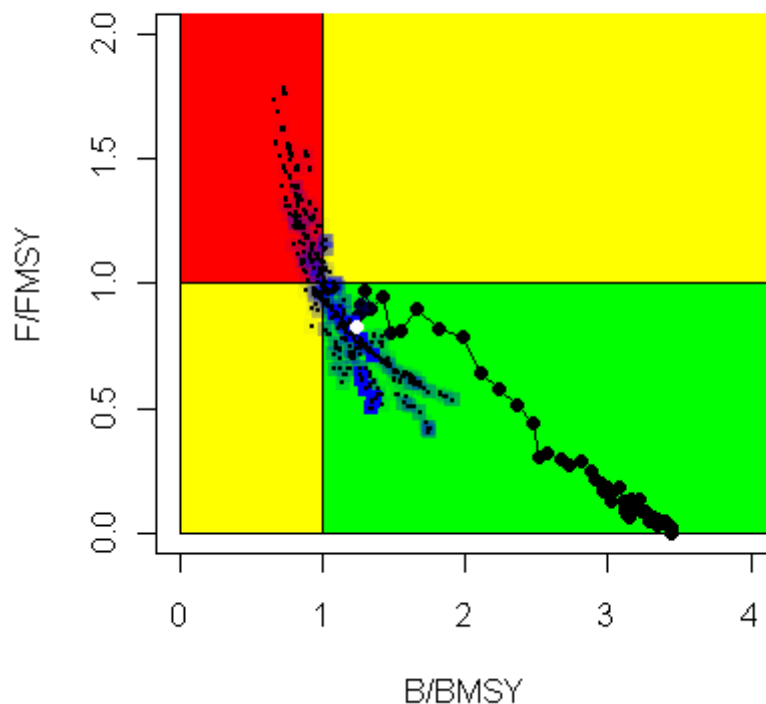


Figure 60. Kobe plot for the 288 model stock status grid. Black circles represent the time series of annual median values from the weighted (Table 5) stock status grid (white circle is 2009). Blue squares indicate the MPD estimates for 2009 corresponding to each individual grid C model, with colour density proportional to the weighting (each model is also indicated by a small black point, as the squares from highly downweighted models are not otherwise visible).

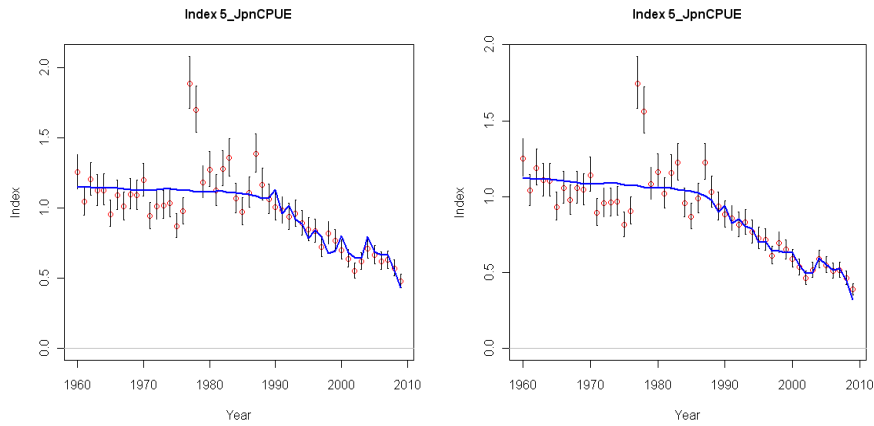


Figure 61. Predicted (line) and observed (points) fit to the LL CPUE series for two of the most optimistic (left) and pessimistic (right) models from the stock status grid.

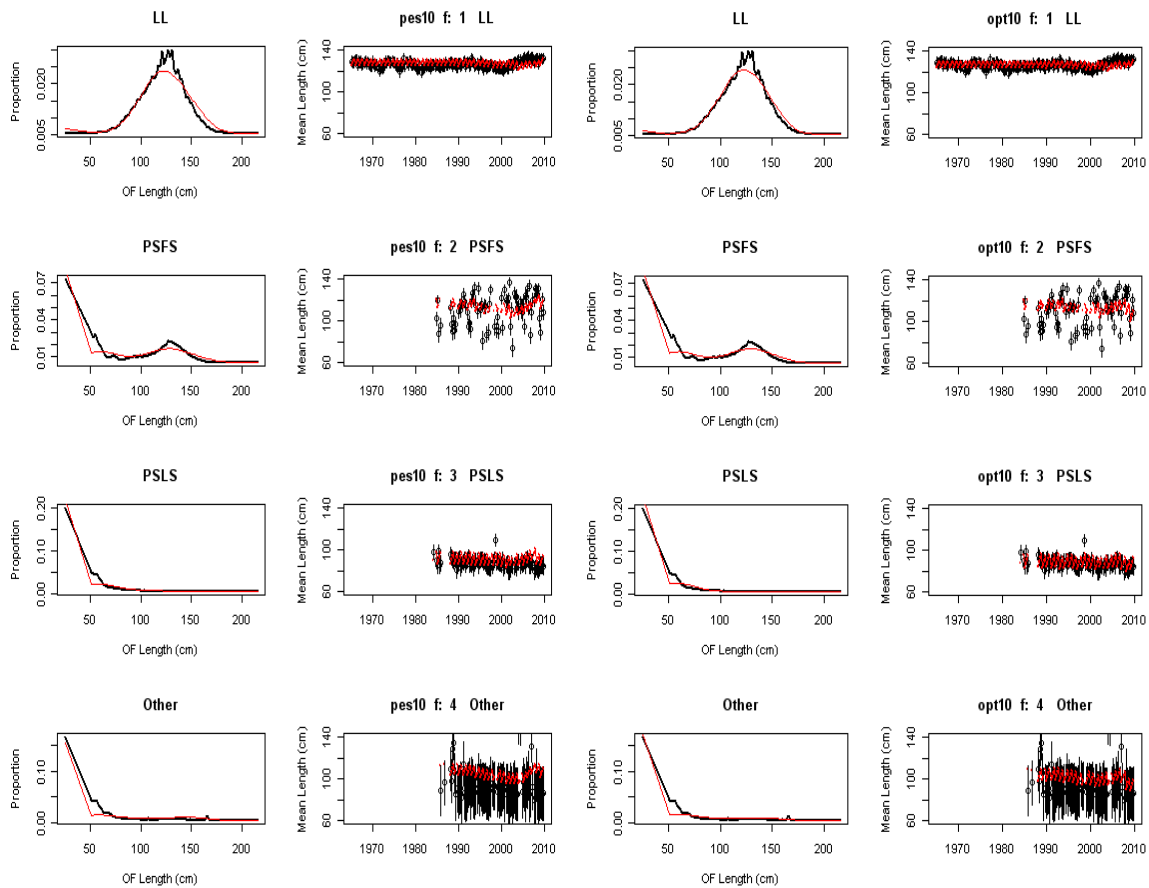


Figure 62. Predicted (red) and observed (black) fit to the size composition data for two of the most optimistic (left 2 columns) and pessimistic (right 2 columns) models from the stock status grid. Columns 1 and 3 represent the distributions aggregated over time, columns 2 and 4 represent the time series of mean size.

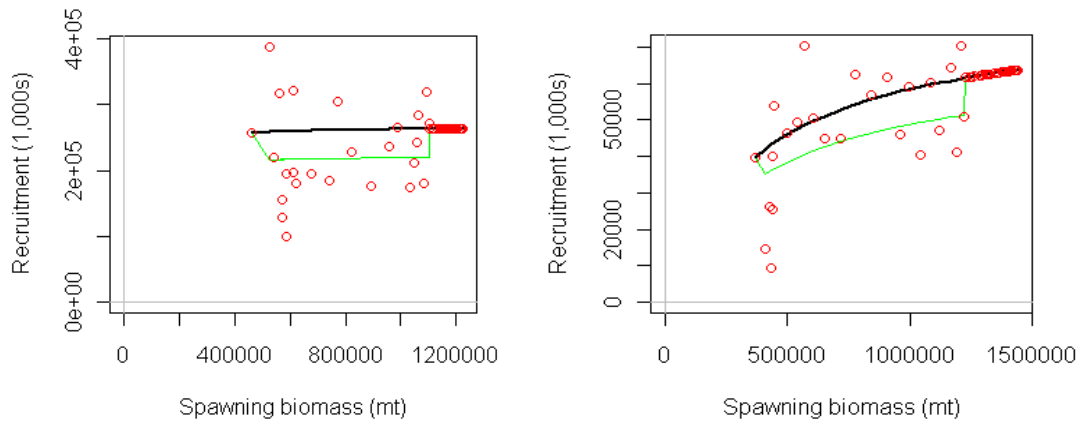


Figure 63. Estimated stock-recruitment relationships for the most optimistic (left) and pessimistic (right) models from the stock status grid.

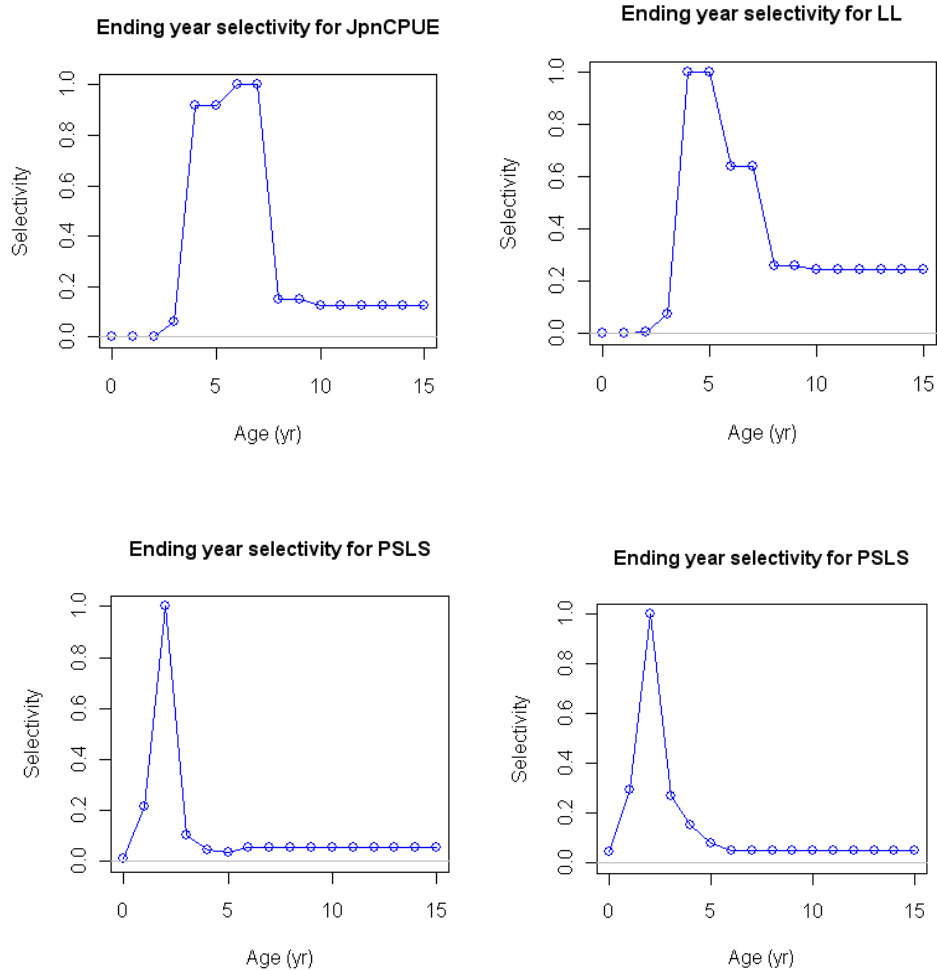


Figure 64. Comparison of the estimated selectivity functions for the optimistic (left) and pessimistic (right) example models for the LL (top) and PSLs (bottom) fleets.

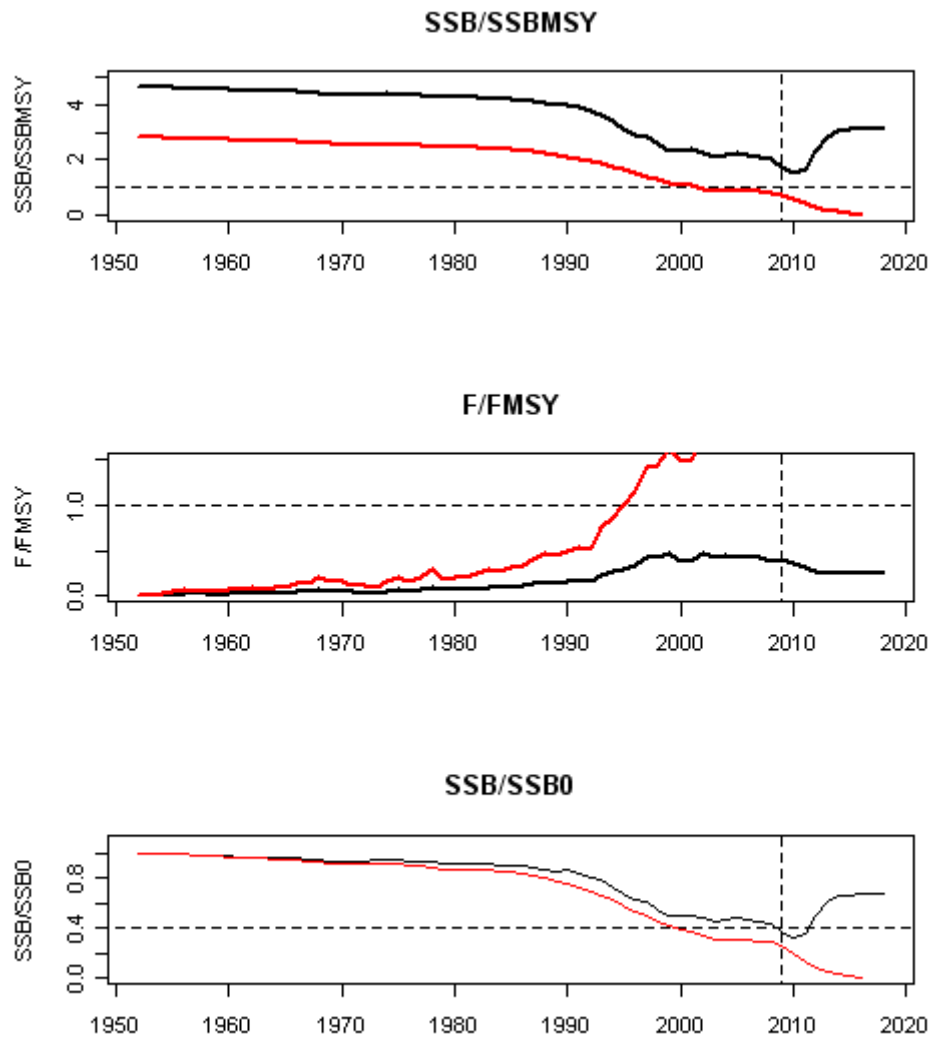


Figure 65. Biomass and fishing mortality time series from two of the most optimistic (black) and pessimistic (red) models from the stock status grid.

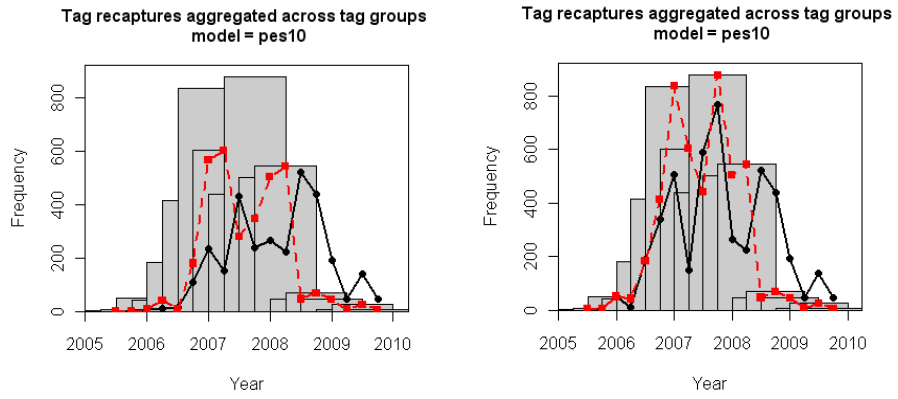


Figure 66. Predicted (red broken lines) and observed (black solid lines) tag recoveries from the pessimistic example model. Columns indicate tag recoveries including the period of incomplete mixing (2 quarters in this case) and are not informative in the model. In the left panel, the un-mixed tags are removed from the predictions and observations, in the right panel the unmixed tags are included.

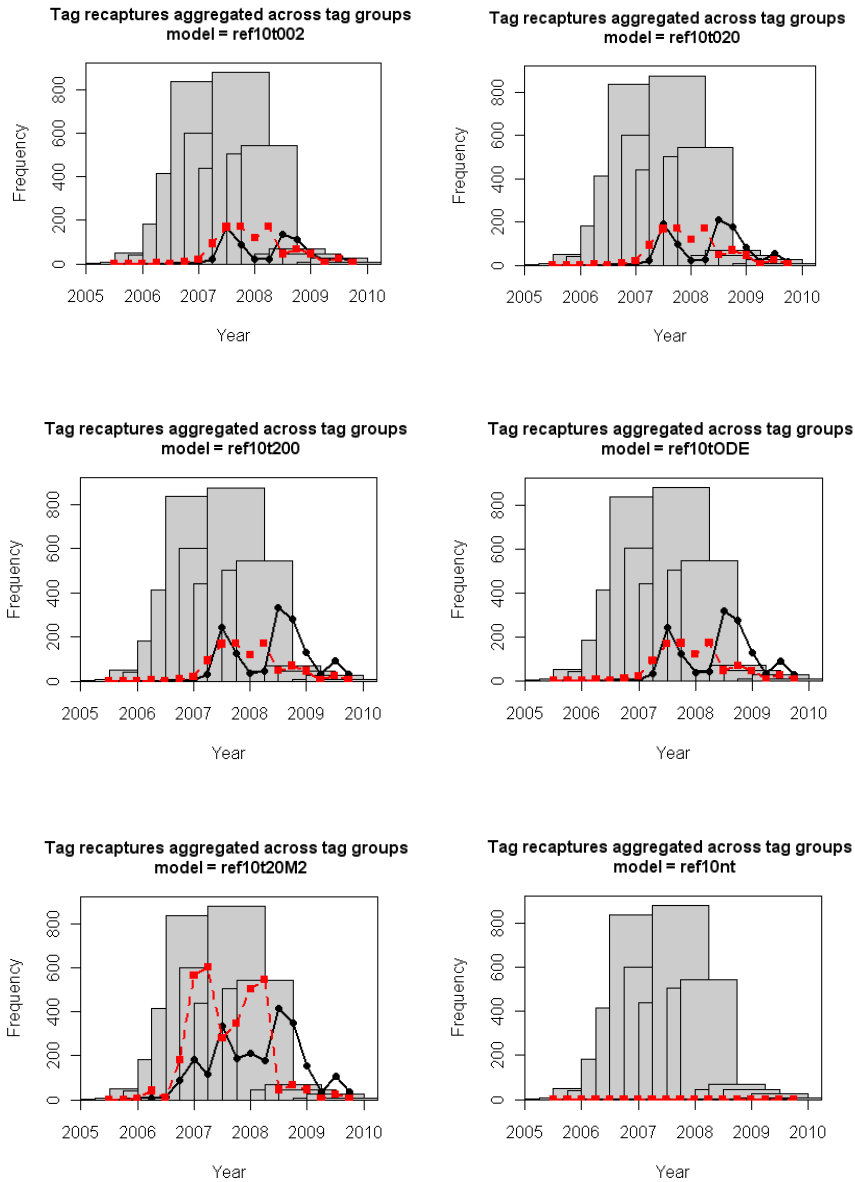


Figure 67. Comparison of the predicted (red broken lines) and observed (black solid lines) fully mixed tag recoveries for a range of example models defined in Table 8. Identical assumptions are used in all models except for the tag overdispersion parameter and the assumed period of complete mixing (options identified by the digits in the model name following 'ref10').


```

-4 4 0 -1 99 -3 0 0 0 0 0.5 0 0 # RecrDist_Area_1_
-4 4 0 -1 99 -3 0 0 0 0 0.5 0 0 # RecrDist_Seas_1_
-4 4 -4 0 -1 99 -3 0 0 0 0 0.5 0 0 # RecrDist_Seas_2_
-4 4 -4 0 -1 99 -3 0 0 0 0 0.5 0 0 # RecrDist_Seas_3_
-4 4 -4 0 -1 99 -3 0 0 0 0 0.5 0 0 # RecrDist_Seas_4_
1 1 1 1 -1 99 -3 0 0 0 0 0.5 0 0 # CohortGrowDev
# 0 # custom_MG-env_setup (0/1)
# -2 2 0 0 -1 99 -2 #_placeholder for no MG-environ parameters
# 0 # custom_MG-block_setup (0/1)
# -2 2 0 0 -1 99 -2 #_placeholder for no MG-block parameters
#_seasonal_effects_on_biology_parms
0 0 0 0 0 0 0 0 #_femwtlen1,femwtlen2,mat1,mat2,fec1,fec2,Malewtlen1,malewtlen2,L1,K
# -2 2 0 0 -1 99 -2 #_placeholder for no seasonal MG parameters
# -2 2 0 0 -1 99 -2 #_placeholder for no MG dev parameters
# -4 # placeholder for #_MGparm_Dev_Phase
#_Spawner-Recruitment
1 #_SR_function
#_LO HI INIT PRIOR PR_type SD PHASE
0 35 14 14 0 10 1 # SR_R0 ## changed from 15 to 14 ##
# xxx h55 0.201 0.99 0.55 0.55 0 10 -2 # SR_steepness
# xxx h75 0.201 0.99 0.75 0.75 0 10 -2 # SR_steepness
# xxx h95 0.201 0.99 0.95 0.95 0 10 -2 # SR_steepness
0 10 0.6 0.6 0 10 -6 # SR_sigmaR
-5 5 0 0 0 1 -3 # SR_envlink
-5 5 0 0 0 1 -4 # SR_R1_offset ## changed from -4 (fixed) to 1 (estimated) ##
0 0.5 0 0 -1 99 -2 # SR_autocorr
0 #_SR_env_link
0 #_SR_env_target_0=none;1=devs;2=R0;3=steepness
# xxx r0 0 #do_recdev: 0=none;1=devvector;2=simple deviations
# xxx r6 1 #do_recdev: 0=none;1=devvector;2=simple deviations
1985 # first year of main recr_devs; early devs can precede this era
2008 # last year of main recr_devs; forecast devs start in following year (from 2008 to 2007)
4 #_recdev phase
1 #0 # (0/1) to read 11 advanced options
0 #_recdev_early_start (0=none; neg value makes relative to recdev_start)
-4 #_recdev_early_phase
-10 #_forecast_recruitment phase (incl. late recr) (0 value resets to maxphase+1)
1 #_lambda for prior_fore_rec occurring before endyr+1
960 #_last_early_yr_nobias_adj_in_MPD
1492 #_first_yr_fullbias_adj_in_MPD
2008 #_last_yr_fullbias_adj_in_MPD
2009 #_first_recent_yr_nobias_adj_in_MPD
1 #_max_bias_adj_in_MPD
0 # period of cycle in recruitment
-15 #min rec_dev
15 #max rec_dev
0 #_read_recdevs
#_end of advanced SR options
#Fishing Mortality info
0.2 # F ballpark for tuning early phases ## changed from 0.1 to 0.4 ##
2000 # F ballpark year(neg value to disable)
3 # F_Method: 1=Pope; 2=instan. F; 3=hybrid (hybrid is recommended) ## changed from 1 to 3 ##
8 # max F or harvest rate, depends on F_Method ## We can change from 0.99 to 4 if F_method is hybrid(3) ##
# no additional F input needed for Fmethod 1
# read overall start F value; overall phase; N detailed inputs to read for Fmethod 2
9 # read N iterations for tuning for Fmethod 3 (recommend 3 to 7)
# Fleet Year Seas F_value se phase (for detailed setup of F_Method=2)
#_initial_F_parms
#_LO HI INIT PRIOR PR_type SD PHASE ## changed the following maximum values from 0.9 to 3.99 ##
# Shono's initial parms and warning...
# 0 3.99 0.76543 0.76543 0 100 -1 # InitF_1_LL (longline) ## changed initial value from 0.07654 to 0.76543 ##
# 0 3.99 0.00012 0.00012 0 100 -1 # InitF_2_PS (purse seine)
# Please be careful about the above values for initial F.
# alternates...why did Shono not free it up???
0 3.99 0.0 0.0 0 100 1 # InitF_1_LL (longline) ## changed initial value from 0.07654 to 0.76543 ##
0 3.99 0.0 0.0 0 100 1 # InitF_2_PFS
0 3.99 0.0 0.0 0 100 1 # InitF_3_PSL
0 3.99 0.0 0.0 0 100 1 # InitF_4_Other

#_Q_setup
# A=do power, B=env-var, C=extra SD, D=devtype(<0=mirror, 0/1=none, 2=cons, 3=rand, 4=randwalk); E=0=num/1=bio, F=err_type
#_A B C D E F ## change the following values of error-type from 0 to 30 for the future ##
0 0 0 0 1 0
0 0 0 0 1 0
0 0 0 0 1 0
0 0 0 0 1 0
0 0 0 0 0
# 0 #_0=read one parm for each fleet with random q; 1=read a parm for each year of index
#_Q_parms(if_any)
## Double normal size selectivity option
## Start Size Sel Block
##_size_selex_types
##_Pattern Discard Male Special
# 24 0 0 0 # 1
# 24 0 0 0 # 2
# 24 0 0 0 # 3
# 24 0 0 0 # 4
# 5 0 0 1 # 1
##_age_selex_types
##_Pattern Discard Male Special
# 10 0 0 0 # 1
# 10 0 0 0 # 2
# 10 0 0 0 # 3
# 10 0 0 0 # 4
# 10 0 0 0 # 5

```

```

## LO HI INIT PRIOR PR_type SD PHASE env-var use_dev dev_minyr dev_maxyr dev_stddev Block Block_Fxn
## 1. LL (longline)
##Begin double normal time sel time series with replacment block
##double normal for fishery 1
## LO HI INIT PRIOR PR_type SD PHASE
##50 200 100 100 199 3 0 0 0 0.5 2 2 # SizeSel_1P_f1 and JPN CPUE
##-6 4 -3 -3 199 3 0 0 0 0.5 2 2 # SizeSel_1P_2_
##-19 8.3 8.3 199 3 0 0 0 0.5 2 2 # SizeSel_1P_3_
##-1 9 4 4 199 3 0 0 0 0.5 2 2 # SizeSel_1P_4_
##-15 -5 -10 -1 199 -3 0 0 0 0.5 2 2 # SizeSel_1P_5_
##-5 9 1.7 -1 199 3 0 0 0 0.5 2 2 # SizeSel_1P_6_
###-15 -5 -999 -1 199 -3 0 0 0 0.5 2 2 # SizeSel_1P_5_
###-5 9 -999 -1 199 -3 0 0 0 0.5 2 2 # SizeSel_1P_6_
##
###double normal for fishery 2 time block 1 (element 14: 0=multiplicative, 1=additive, 2=replace)
##15 200 40 40 199 3 0 0 0 0.5 1 2 # SizeSel_1P_1_f2
##-6 4 -5 -5 199 3 0 0 0 0.5 1 2 # SizeSel_1P_2_
##-19 8.3 8.3 199 3 0 0 0 0.5 1 2 # SizeSel_1P_3_
##-1 9 4 4 199 3 0 0 0 0.5 1 2 # SizeSel_1P_4_
##-15 5 -10 -1 199 3 0 0 0 0.5 1 2 # SizeSel_1P_5_
##-15 9 -10 -1 199 3 0 0 0 0.5 1 2 # SizeSel_1P_6_
##
###double normal for survey 1
##-5 3 1 -4 1 0.05 -3 0 0 0 0.5 0 0 # size sel mirror p1 f3
##-5 3 -1 -4 1 0.05 -3 0 0 0 0.5 0 0 # size sel mirror p2 f3
##
##1 #custom block set up
###double normal for fishery 1 time blocks with block Fxn=replacement
##15 200 100 100 199 3 # SizeSel_1P_1_f2
##15 200 100 100 199 -3 # SizeSel_1P_1_f2
##15 200 100 100 199 -3 # SizeSel_1P_1_f2
##15 200 100 100 199 -3 # SizeSel_1P_1_f2
##-6 4 -3 -3 199 3 # SizeSel_1P_2_
##-6 4 -3 -3 199 -3 # SizeSel_1P_2_
##-6 4 -3 -3 199 -3 # SizeSel_1P_2_
##-6 4 -3 -3 199 -3 # SizeSel_1P_2_
##-6 4 -3 -3 199 -3 # SizeSel_1P_2_
##-1 9 8.3 8.3 199 3 # SizeSel_1P_3_
##-1 9 8.3 8.3 199 -3 # SizeSel_1P_3_
##-1 9 8.3 8.3 199 -3 # SizeSel_1P_3_
##-1 9 8.3 8.3 199 -3 # SizeSel_1P_3_
##-1 9 8.3 8.3 199 -3 # SizeSel_1P_3_
##-1 9 4 4 199 3 # SizeSel_1P_4_
##-1 9 4 4 199 -3 # SizeSel_1P_4_
##-1 9 4 4 199 -3 # SizeSel_1P_4_
##-1 9 4 4 199 -3 # SizeSel_1P_4_
##-1 9 4 4 199 -3 # SizeSel_1P_4_
##-15 -5 -10 -1 199 -3 # SizeSel_1P_5_
##-15 -5 -10 -1 199 -3 # SizeSel_1P_5_
##-15 -5 -10 -1 199 -3 # SizeSel_1P_5_
##-15 -5 -10 -1 199 -3 # SizeSel_1P_5_
##-15 -5 -10 -1 199 -3 # SizeSel_1P_5_
##-5 9 1.7 -1 199 3 # SizeSel_1P_6_
##-5 9 1.7 -1 199 -3 # SizeSel_1P_6_
##-5 9 1.7 -1 199 -3 # SizeSel_1P_6_
##-5 9 1.7 -1 199 -3 # SizeSel_1P_6_
##-5 9 1.7 -1 199 -3 # SizeSel_1P_6_
##
###double normal for fishery 2 time blocks with block Fxn=replacement
##15 200 40 40 199 3 # SizeSel_1P_1_f2
##15 200 40 40 199 -5 # SizeSel_1P_1_f2
##15 200 40 40 199 -5 # SizeSel_1P_1_f2
##15 200 40 40 199 -5 # SizeSel_1P_1_f2
##15 200 40 40 199 -5 # SizeSel_1P_1_f2
##-6 4 -5 -5 199 3 # SizeSel_1P_2_
##-6 4 -5 -5 199 -5 # SizeSel_1P_2_
##-6 4 -5 -5 199 -5 # SizeSel_1P_2_
##-6 4 -5 -5 199 -5 # SizeSel_1P_2_
##-6 4 -5 -5 199 -5 # SizeSel_1P_2_
##-1 9 8.3 8.3 199 3 # SizeSel_1P_3_
##-1 9 8.3 8.3 199 -5 # SizeSel_1P_3_
##-1 9 8.3 8.3 199 -5 # SizeSel_1P_3_
##-1 9 8.3 8.3 199 -5 # SizeSel_1P_3_
##-1 9 8.3 8.3 199 -5 # SizeSel_1P_3_
##-1 9 4 4 199 3 # SizeSel_1P_4_
##-1 9 4 4 199 -5 # SizeSel_1P_4_
##-1 9 4 4 199 -5 # SizeSel_1P_4_
##-1 9 4 4 199 -5 # SizeSel_1P_4_
##-1 9 4 4 199 -5 # SizeSel_1P_4_
##-15 5 -10 -1 199 3 # SizeSel_1P_5_
##-15 5 -10 -1 199 -5 # SizeSel_1P_5_
##-15 5 -10 -1 199 -5 # SizeSel_1P_5_
##-15 5 -10 -1 199 -5 # SizeSel_1P_5_
##-15 5 -10 -1 199 -5 # SizeSel_1P_5_
##-15 9 -10 -1 199 3 # SizeSel_1P_6_
##-15 9 -10 -1 199 -5 # SizeSel_1P_6_
##-15 9 -10 -1 199 -5 # SizeSel_1P_6_
##-15 9 -10 -1 199 -5 # SizeSel_1P_6_
##-15 9 -10 -1 199 -5 # SizeSel_1P_6_
##-15 9 -10 -1 199 -5 # SizeSel_1P_6_
##2 # selparm_adjst_method 1=direct, 2=logistic transform...seems to be ignored for replacement blocks
###End double normal time sel time series with replacment block
##Begin double normal time sel time series with multiplicative devs block
##double normal for fishery 1
## LO HI INIT PRIOR PR_type SD PHASE

```



```

# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_-20 10 -2. -2. 1 2. 2 0 0 0 0 0 0 # TG_report_fleet: 1_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_-20 10 20. 20. 1 0.2 -4 0 0 0 0 0 0 # TG_report_fleet: 2_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_ # Initial tag reporting rate for each fleet, transformation = rep rate = exp(p)/(1+exp(p)) (apparently!)
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_1_ #PS recoveries already inflated by RR (PSSL and PSFS)
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_-20 10 -2. -2. 1 2. 2 0 0 0 0 0 0 # TG_report_fleet: 1_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_-20 10 20. 20. 1 0.2 -4 0 0 0 0 0 0 # TG_report_fleet: 2_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_1_-20 10 20 1 0.2 -4 0 0 0 0 0 0 # TG_report_fleet: 2_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_-20 10 -2. -2. 1 2. 2 0 0 0 0 0 0 # TG_report_fleet: 1_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_ #PS recoveries already inflated by RR
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_-20 10 -10. -10. 1 2. 2 0 0 0 0 0 0 # TG_report_fleet: 1_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_1_-20 10 -10. -10. 1 0.2 2 0 0 0 0 0 0 # TG_report_fleet: 2_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_-20 10 20 1 0.2 -4 0 0 0 0 0 0 # TG_report_fleet: 2_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_-20 10 -10. -10. 1 2. 2 0 0 0 0 0 0 # TG_report_fleet: 1_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_1_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_ # LO HI INIT PRIOR PR_type SD PHASE
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_1_ # Exponential decay rate in reporting rate for each fleet (default=0, negative value to get decay)
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_-4 0 0 0 2 -4 0 0 0 0 0 0 # TG_rpt_decay_fleet: 1_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_-4 0 0 0 2 -4 0 0 0 0 0 0 # TG_rpt_decay_fleet: 2_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_-4 0 0 0 2 -4 0 0 0 0 0 0 # TG_rpt_decay_fleet: 1_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_1_-4 0 0 0 2 -4 0 0 0 0 0 0 # TG_rpt_decay_fleet: 2_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_1_ # Variance_adjustments_to_input_values
# xxx t70M2 150 70 70 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_1_ #_1_2_3

# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_ # Initial tag reporting rate for each fleet, transformation = rep rate = exp(p)/(1+exp(p)) (apparently!)
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_1_ #PS recoveries already inflated by RR (PSSL and PSFS)
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_-20 10 -2. -2. 1 2. 2 0 0 0 0 0 0 # TG_report_fleet: 1_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_-20 10 20. 20. 1 0.2 -4 0 0 0 0 0 0 # TG_report_fleet: 2_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_ # Initial tag reporting rate for each fleet, transformation = rep rate = exp(p)/(1+exp(p)) (apparently!)
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_1_ #PS recoveries already inflated by RR (PSSL and PSFS)
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_-20 10 -2. -2. 1 2. 2 0 0 0 0 0 0 # TG_report_fleet: 1_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_-20 10 20. 20. 1 0.2 -4 0 0 0 0 0 0 # TG_report_fleet: 2_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_1_-20 10 20 1 0.2 -4 0 0 0 0 0 0 # TG_report_fleet: 2_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_-20 10 -2. -2. 1 2. 2 0 0 0 0 0 0 # TG_report_fleet: 1_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_ #PS recoveries already inflated by RR
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_-20 10 -10. -10. 1 2. 2 0 0 0 0 0 0 # TG_report_fleet: 1_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_1_-20 10 -10. -10. 1 0.2 2 0 0 0 0 0 0 # TG_report_fleet: 2_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_-20 10 20 1 0.2 -4 0 0 0 0 0 0 # TG_report_fleet: 2_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_-20 10 -10. -10. 1 2. 2 0 0 0 0 0 0 # TG_report_fleet: 1_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_1_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_ # LO HI INIT PRIOR PR_type SD PHASE
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_1_ # Exponential decay rate in reporting rate for each fleet (default=0, negative value to get decay)
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_-4 0 0 0 2 -4 0 0 0 0 0 0 # TG_rpt_decay_fleet: 1_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_-4 0 0 0 2 -4 0 0 0 0 0 0 # TG_rpt_decay_fleet: 2_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_-4 0 0 0 2 -4 0 0 0 0 0 0 # TG_rpt_decay_fleet: 1_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_1_-4 0 0 0 2 -4 0 0 0 0 0 0 # TG_rpt_decay_fleet: 2_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_2_
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_3_1_ # Variance_adjustments_to_input_values
# xxx t200 150 200 200 1 0.001 -4 0 0 0 0 0 0 # TG_overdispersion_1_ #_1_2_3
# Initial tag reporting rate for each fleet, transformation = rep rate = exp(p)/(1+exp(p)) (apparently!)
#PS recoveries already inflated by RR (PSSL and PSFS)
#-20 10 -2. -2. 1 2. 2 0 0 0 0 0 0 # TG_report_fleet: 1_
#-20 10 20. 20. 1 0.2 -4 0 0 0 0 0 0 # TG_report_fleet: 2_
#-20 10 20 20 1 0.2 -4 0 0 0 0 0 0 # TG_report_fleet: 2_
#-20 10 -2. -2. 1 2. 2 0 0 0 0 0 0 # TG_report_fleet: 1_
#PS recoveries already inflated by RR
-20 10 -10. -10. 1 2. 2 0 0 0 0 0 0 # TG_report_fleet: 1_
-20 10 -10. -10. 1 0.2 2 0 0 0 0 0 0 # TG_report_fleet: 2_
-20 10 20 20 1 0.2 -4 0 0 0 0 0 0 # TG_report_fleet: 2_
-20 10 -10. -10. 1 2. 2 0 0 0 0 0 0 # TG_report_fleet: 1_
# LO HI INIT PRIOR PR_type SD PHASE
# Exponential decay rate in reporting rate for each fleet (default=0, negative value to get decay)
-4 0 0 0 2 -4 0 0 0 0 0 0 # TG_rpt_decay_fleet: 1_
-4 0 0 0 2 -4 0 0 0 0 0 0 # TG_rpt_decay_fleet: 2_
-4 0 0 0 2 -4 0 0 0 0 0 0 # TG_rpt_decay_fleet: 1_
-4 0 0 0 2 -4 0 0 0 0 0 0 # TG_rpt_decay_fleet: 2_
1# Variance_adjustments_to_input_values
#_1_2_3
0 0 0 0 #_add_to_survey_CV
0 0 0 0 #_add_to_discard_CV
0 0 0 0 #_add_to_bodywt_CV
1 1 1 1 #_mult_by_lencomp_N
1 1 1 1 #_mult_by_agecomp_N
1 1 1 1 #_mult_by_size-at-age_N
30 #_DF_for_discard_like
30 #_DF_for_meanbodywt_like
4 #_maxlambdaphase
1 #_sd_offset
13 # number of changes to make to default Lambdas (default value is 1.0)
# Like_comp codes: 1=survey; 2=disc; 3=mnwt; 4=length; 5=age; 6=SizeFreq; 7=sizeage; 8=catch;
# 9=init_equ_catch; 10=recrdev; 11=parm_prior; 12=parm_dev; 13=CrashPen; 14=Morphcomp; 15=Tag-comp; 16=Tag-negbin
#like_comp_fleet/survey phase value sizefreq_method

#CPUE
1 3 1 1 . 1
1 3 2 1 . 1
1 3 3 1 . 1

```

```

#Size
# xxx CL1050 4 1 1 1. 1
# xxx CL1050 4 2 1 0.1 1
# xxx CL1050 4 3 1 5. 1
# xxx CL1050 4 4 1 0.1 1

# xxx CL1010 4 1 1 1. 1
# xxx CL1010 4 2 1 0.1 1
# xxx CL1010 4 3 1 1. 1
# xxx CL1010 4 4 1 0.1 1

# xxx CL1001 4 1 1 1. 1
# xxx CL1001 4 2 1 0.1 1
# xxx CL1001 4 3 1 0.1 1
# xxx CL1001 4 4 1 0.1 1

# xxx CL0110 4 1 1 0.1 1
# xxx CL0110 4 2 1 0.1 1
# xxx CL0110 4 3 1 1. 1
# xxx CL0110 4 4 1 0.1 1

# xxx CL0101 4 1 1 0.1 1
# xxx CL0101 4 2 1 0.1 1
# xxx CL0101 4 3 1 0.1 1
# xxx CL0101 4 4 1 0.1 1

# tags...not clear on assignment definitions
# 15 tag-comp does not seem to show up in report file, but weightings do change result trivially?
15 1 2 1. 1
15 2 2 1. 1
15 1 3 1. 1
15 2 3 1. 1
15 1 4 1. 1
15 2 4 1. 1
# lambdas (for info only; columns are phases)
# 0 #_CPUE/survey:_1
# 0 #_CPUE/survey:_2
# 1 #_CPUE/survey:_3
# 1 #_lencomp:_1
# 1 #_lencomp:_2
# 0 #_lencomp:_3
# 1 #_init_equ_catch
# 1 #_recruitments
# 1 #_parameter-priors
# 0 #_parameter-dev-vectors
# 100 #_crashPenLambda
0 # (0/1) read specs for extra stddev reporting
# 0 1 -1 5 1 5 1 -1 5 # placeholder for selex type, len/age, year, N selex bins, Growth pattern, N growth ages, NatAge_area(-1 for all), NatAge_yr, N Natages
# -1 1 1 1 1 # placeholder for vector of selex bins to be reported
# -1 1 1 1 1 # placeholder for vector of growth ages to be reported
# -1 1 1 1 1 # placeholder for vector of NatAges ages to be reported
999

```