Exploratory Modelling of the Indian Ocean Swordfish Fishery, using an age-structured, sex-structured and spatially-disaggregated implementation of Stock Synthesis software

Dale Kolody
(Dale.Kolody@csiro.au)

CSIRO Marine and Atmospheric Research, Hobart, Australia
May 2009

Abstract

This document briefly describes some preliminary work toward an Indian Ocean swordfish stock assessment using Stock Synthesis 3 software, and discusses options that might be pursued prior to the IOTC Billfish Working Party in July 2009. Two example models are presented, that fit the total catch in mass (1950-2007) and size composition data (1971-2007) from 24 fleets defined by the IOTC Secretariat, and the standardized Japanese longline CPUE series (1980-2006) from Semba and Nishida (IOTC-2008-WPB6-Info1). The model population is age-structured, sex-disaggregated, and spatially disaggregated into 4 regions (NW = North-West, NE, SW, SE). The preliminary explorations suggest that this freely distributed NMFS software is reasonably easy to use, very flexible, numericallyefficient and supported with convenient R graphical utilities. This application does not constitute an assessment for several reasons, and it is hoped that three important areas can be advanced before July:

1. Incorporation of additional data:
    - Updated, standardized Japanese and Taiwanese longline CPUE series
    - Standardized CPUE from La Reunion and the Seychelles
    - Catch by sex from the Spanish fleet
2. Exploration/quantification of key uncertainties, including:
    - Stock structure, migration patterns and mixing rates
    - Natural mortality
    - Growth rates (uncertainty due to age estimation methods)
    - Process and observation error characteristics of the CPUE and size composition data.
3. A spatially restricted assessment focussing on the SW Indian Ocean, assuming that it is a somewhat discrete sub-population, and possibly more depleted than the other regions.

It is hoped that this document will help to illustrate what is feasible, and stimulate discussion in advance of the July WPB particularly in relation to biological assumptions, data interpretation, and the objectives of the assessment. However, it is also important to emphasize that this approach remains experimental. At this time it should be seen as complementary to, but not a replacement for, the other methods that have been used in recent years.

Introduction

The 2008 WPB report describes three different modelling approaches that were used for the assessment of the Indian Ocean swordfish fishery:

1. Spatially-aggregated surplus production model (age-aggregated, deterministic recruitment) – this model was used to estimate the depletion and production characteristics of the whole Indian Ocean, and provided continuity with the methods applied in 2006.
2. Spatially disaggregated Pella-Tomlinson production model (age-aggregated, deterministic recruitment) – this 4 area model attempted to describe the apparent differential depletion across different regions in the Indian Ocean (and used life history considerations to bound the surplus production 'shape' parameters). The better fit models were consistent with lower migration rates and higher depletion in the SW, but the data were also reasonably consistent with a highly mixed population. Numerical issues left some question about the general applicability of the model.
3. Spatially-aggregated, age- and sex-structured integrated analysis – this model represented the first attempt to include the swordfish size composition data from the Indian Ocean into an assessment. Results were preliminary, and generally suggestive of a much less productive stock than the other two methods. The use of size composition data potentially adds another means of estimating fishing mortality effects on the population, and assists in the estimation of year-class strength. However, it also includes strong assumptions about temporally stable selectivity and the randomness of sampling. The spatial aggregation prevents examination of area-specific effects.

The models described in this paper attempt to combine the best features from the models above to i) estimate depletion in different sub-areas, ii) use the size composition data to potentially extract additional information about recruitment variability and fishing mortality. Additional data may also be used if possible (standardized CPUE from La Reunion and the Seychelles; Spanish sex composition data).

There are some potential downsides to adopting the more complicated assessment framework, but the most obvious ones do not seem to be debilitating:

- Reliance on third party software – this is probably not a problem in this case because the developer (Rick Methot, NMFS) actively supports the Stock Synthesis software, it is widely used, very flexible and core features have presumably become more reliable over time as bugs have been identified and fixed. Source code is also distributed.

- Additional overheads for the user – the software is reasonably well documented, with many example applications available, and to date has proved easier to use than other similar products.

- Increased computation time – the function minimization uses the highly efficient AD Model Builder software. Initial applications of the highly disaggregated swo assessment suggest minimization in 10 minutes or less (2.4 GHz processor).

- Over-parameterization – there is not enough data to estimate all of the potentially important processes. This is inevitable with virtually any complicated stock assessment, and swordfish data is not as good as many other fisheries (i.e. no tags, stock structure poorly understood, growth rates uncertain, M unknown). With a sufficiently complicated model, one can represent important details, examine the sensitivity of the model to the inestimable quantities, prioritize further research for reducing these uncertainties, and develop management strategies that are robust to these uncertainties to the extent possible. However depending on the purpose of the assessment, complicated models may not be superior to simpler methods (noting that the simple models have their own problems related to restrictive assumptions and serious biases).

This paper attempts to summarize the exploratory modelling undertaken with SS3 to date and some plans for what might be achieved before the July WPB. It is hoped that interested parties will provide constructive feedback on important issues that have been missed, alternative interpretations of the system that might be represented, priorities of interest, and likely timelines for the provision of additional data before the WPB.

Why are we doing this assessment?

It is worth emphasizing that this exploration was undertaken in part for the author to gain familiarity with the Indian Ocean swordfish fishery and SS3 software. However, depending on what questions the assessment is intended to answer, it is not clear that more complicated models are always necessary, or more effective than simpler models. With this in mind, I pose a number of questions that I am assuming are relevant, but look for broader clarification:

1. Are the IOTC management objectives being met for Indian Ocean swordfish?
    a. Are management objectives being met in some regions but not others?
2. How confident are we in the assessment conclusions?
    a. Are there additional analyses or data requirements that could be used to improve our confidence in these conclusions?
3. Can we quantify the implications of alternative management options that will allow managers to choose actions that will increase the likelihood that the IOTC will meet its management objectives?
    a. Are there fundamentally different data requirements for some of these options? Which options will provide the best management return from the investment in monitoring and analysis?

At the 2009 WPB, I would hope to make progress toward addressing questions 1 and 2, while question 3 might provide a focus for longer term strategic planning.

Software

The current model was implemented with Stock synthesis SS V3.03 (Rick Methot, NOAA). Graphics are mostly from the R functions SSv3_plots (Google code: BETA May 13, 2009, Ian Stewart and Ian Taylor, NWFSC).

Data

Total catch in mass, catch length frequency, and catch length frequency by sex was provided by Miguel Herrera (IOTC secretariat).

Standardized CPUE series were provided by Semba and Nishida (2008, IOTC-2008-WPB6-Info1) from 1980-2006, and are subject to revision pending new analyses in 2009. The current data are annual, but quarterly indices could be used as well.

Other sources of biological data are described under model assumptions below.

Model Assumptions

Spatial Structure

The model is disaggregated into 4 areas corresponding to those used in the catch rate standardization analysis of Semba, Nishida and Wang (2008, IOTC-2008-WPB6-Info1) (Figure 1). Some evidence suggests that there may be genetic distinction within the IO, but this is the subject of ongoing investigation. Given the vast size of the Indian Ocean, and the migration rate inferences that have been made from tagging studies (particularly in other oceans), it seems unlikely that there would be rapid mixing processes across the whole basin, even if the population was genetically homogeneous. As such, localized overfishing could result in negative localized consequences even if the overall stock is not overfished, or at serious risk of declining genetic diversity. The 4 area structure seems reasonable for resolving the disproportionately high depletion observed in the SW region. The 4 areas also correspond to convenient partitions for many of the national fleets.

Migration Dynamics

There are very few direct observations of swordfish migration in the Indian Ocean. I am aware of a few conventional tag releases and recaptures near the Australian coast, with no indication of large scale movements. There are plans to deploy a number of electronic tags on swordfish in the Indian Ocean, however, I am not aware of any recoveries.

We can indirectly infer that there are probably some relatively large seasonal migrations. Swordfish are caught in the temperate waters south of 35S, however, the spawning regions (and larval distributions) tend to be in the tropical regions. At least in the southern hemisphere this suggests directed seasonal migrations.

- Is there any evidence for seasonal migration in the northern part of the IO? Is it out of phase with the southern hemisphere?

At this stage, migration can probably only be considered in terms of the rate of random diffusive mixing (i.e. are the 4 populations rapidly mixing, or relatively isolated from each other?)

Fishery Definitions

In the example model, 24 fleets were defined, corresponding to the data aggregation units of the catch data as supplied by the IOTC Secretariat. Each 'fleet' resides in a single area only. If the same nation operates in more than one region, these operations are described here as a separate fleet. Many of the small fleets were aggregated in the data as provided by the IOTC secretariat. It would probably be sensible to further reduce the number of fleets, but given that there did not seem to be any real computational burden, this was not yet pursued.

Time Period

The model was run from 1950-2007 using a quarterly timestep. This model could probably be run on an annual timestep, but the quarterly timestep does allow the model to potentially resolve seasonal migration characteristics.

- check for seasonal cycles in CPUE (are they out of phase in the northern and southern regions?).
- check for seasonality in size composition (is it attributable to growth or migration?)

Age and Sex Structure

The model population is age- and sex-structured with cohorts of 0-40+ years, for each of two sexes. Sex-specific characteristics potentially include:

- growth curves
- natural mortality
- age-based selectivity (when derived from a length-based function that is the same for both sexes)
- SS3 supports the fitting to catch size composition disaggregated by sex (but this was not yet implemented)

Age and Size

There is strong evidence for sex dimorphism in swordfish, and it is conceivable that aggregating data across heterogeneous units can lead to statistical biases in these sorts of models. However, it is not clear that this is a high priority for the assessment because there is currently considerable uncertainty about all swordfish age estimation methods, and a number of other related (e.g. mortality) model assumptions.

Two sets of growth curves were explored to bracket two relatively extreme cases that might be plausible given the scarcity of age validation data for swordfish (Figure 2)

- CSIRO curve, derived from South-East Indian Ocean fin rays samples.
- NMFS curve, derived from Hawaiian samples.

The biology of the Hawaiian swordfish may differ considerably from the Indian Ocean, however, if the NMFS age estimation method is more accurate than the CSIRO method, then the Hawaiian growth curve is probably preferable. The

Australian growth curves should also be compared with others derived for the Indian Ocean.

Maturity and Spawning Stock Biomass

Stock Synthesis can use age-specific vectors of female maturity or fecundity for biomass spawning calculations. While a number of studies quantify the relationship between size and maturity, the uncertainty of age estimation that undermines the growth relationships also undermines the maturity/fecundity by age relationship. Two relatioships were explored, roughly corresponding to extremes associated with the growth curves:

- 50% maturity ~age 10, corresponding to the CSIRO study (mostly based on Western Pacific samples).
- 50% maturity ~age 4, corresponding with one of the youngest age at maturity used in swordfish assessment.

The fecundity relationship was not yet examined.

Selectivity

All fleets seem to suggest rather similar size composition, with none of them indicating strong modal structure that can be informative for resolving year class strength in juveniles. SS3 supports length- and age-based selectivity with numerous functional forms. Three options were examined initially

- age-based, one parameter per age - results in polymodal selectivity unless strong smoothing constraints are used; strong peak around age 5
- length-based, double normal – results in dome shape with peak around 170-200 cm
- length-based logistic – plateau starts around 180 cm

The usual application of selectivity in these sorts of models involves the assumption that it is constant over time for a given gear type and area. Fleets should be disaggregated whenever there is reason to suspect that they are harvesting a different proportion of the population. However, for fleets with no data, selectivity should probably be assumed to resemble the most similar fleet.

- if there are real changes in selectivity over time, it is likely that this approach can lead to misleading recruitment time series.
- a related problem can arise with respect to non-random sampling of the size composition data, which results in misleading inferences about size composition and/or selectivity

In the exploration, two options were tested: all fleets share 1 selectivity, and each fleet with size composition data represents a unique selectivity (18 total).

Catchability

Catchability is assumed to be constant over time for the fleets with standardized CPUE. In this initial application, only the Japanese fleet was used. It is further assumed that catchability is equivalent among areas, such that cpue is a measure of density, and relative abundance by area is the product of area and density.

- The area-density assumption is commonly used to share information among areas, however, without more information about the standardization process, it is not clear that this is the most appropriate way to use this data.

- When available, the Taiwanese fleet should provide an additional perspective on these area assumptions.

Catch in mass observation errors

For the purposes of the modelling, total catches are assumed to be known without error at this time.

Catch-at-Size sampling characteristics

Some of the sample sizes are very large for some fleets. In the context of the current model, this might cause a misleading overfitting to the size composition data for a number of reasons, including: i) sampling is probably not truly random for all fleets, ii) selectivity is probably not truly constant for any fleet, iii) interactions with other model assumptions (e.g. M) might create troublesome model size estimates. Sample sizes have been treated:

- all sample sizes downweighted by a factor of 10.
- a somewhat arbitrary constant (1%) has been added to each of the predicted and observed length bins to reduce the influence of outliers on the multinomial catch-at-size likelihood term.

CPUE characteristics

The standardized CPUE was assumed to be proportional to selected abundance (numbers) and highly (unrealistically) informative with a CV of 5-10%. The annual indices were assumed to sorrespond to abundance in quarter 1. A number of additional considerations:

- Are the standardized cpue trends from the different nations consistent with each other within each area?
- Are the relative weightings among areas in the Japanese and Taiwanese surveys consistent?
- Are there operational reasons why we might believe one series more than another?

Stock Recruitment Relationship

A Beverton-Holt stock recruitment relationship was assumed, with a fixed steepness of 0.8. In SS3 there is a single annual spawning biomass calculation, but the recruitment can be partitioned in various ways. For this application:

- recruitment was assumed to occur once annually (quarter 1),
- area-specific parameters were estimated to distribute the recruitment (constant over time),
- annual deviations from the stock recruitment curve were estimated in some cases. The deviations were highly constrained SD(log(devs)) = 0.1.

Fishing Mortality

The “hybrid” fishing mortality parameterization was used, in which SS3 starts with Pope’s approximation and then conducts a fixed number of iterations to approximate instantaneous F from the Baranov catch equation.

Initial Population

The population was assumed to be in unfished equilibrium in 1950, the start of the catch data series.

Model fitting

The models fit to date involved minimization of an objective function with the following terms:

Likelihoods:

- CPUE – lognormal observation errors
- Length frequencies – multinomial distribution (downweighted sample sizes with 1% added to each bin to reduce the influence of rare observations)
- lognormal annual recruitment deviates

Prior distributions:

- parameters are either fixed or extremely diffuse priors were used

Penalties:

- smooth penalties as parameters approach bounds (bounds were rarely hit in the explorations)

Estimated parameters:

- q for each CPUE series (shared across Japanese longline fleets)
- mean virgin recruitment
- selectivity (6 parameter double normal)
- recruitment proportion by area
- annual recruitment deviations from the stock recruitment relationship

Additional estimable parameters that are usually fixed:

- natural mortality
- migration rates - two sets of two age-specific parameters (with linear interpolation between ages) for every combination of two adjacent areas
- stock recruitment steepness

Results and Discussion

Some results from two example models are illustrated below. Both models included the following assumptions:

- Beverton Holt stock-recruitment relationship (steepness = 0.8)
- Japanese CPUE observation error SD(log(observation error)) = 0.05
- Catchability constant across all 4 regions
- Size samples downweighted by factor of 10
- 24 fleets, with 18 different (double-normal, size-based) selectivities (constant over time)
- Migration rates negligible (fixed at <1%/year)

The models differ as indicated in the table below

	Example model 1 (IOswo39p4)	**Example model 2 (IOswo39p4)**
Growth	slow (CSIRO-Eastern IO)	fast (NMFS-Hawai'ian)
Maturity	50% age 10	50% age 4
Mortality (per y)	0.2	0.4
Recruitment	Stochastic CV = 0.1	Deterministic

Figure 3 and Figure 4 show the fits to the CPUE data for the two models. Given the unrealistically low observation errors assumed for these fleets, it is not surprising that the fit is reasonable. However, the good CPUE fit comes at a cost in the poor fit to the size composition data. Figure 5 and Figure 6 show the fit for the Japanese SW fishery only (similar plots for each of 18 fleets could be shown), and indicate that the mode of the size distribution is missed quite badly in some years.

Among the models explored to date, the option that seems to have the biggest impact on the assessment seems to be whether or not the recruitment deviations are estimated. Not surprisingly, the model fits the data better when recruitment deviations are estimated. Perhaps more surprising is that the model fits as well as it does when the deviations are not estimated. The main (and fairly consistent) effect of the recruitment deviations is to cause an increase in biomass around the beginning of the CPUE time series (extra recruits around 1975-85), with a subsequent recruitment decline immediately afterward which suggests that part of the steep CPUE decline post-1990 is recruitment driven, not fishery driven (Figure 7, Figure 8). Further work would need to be done to determine if this explanation is plausibly consistent with the other CPUE series and the size composition data (or whether it might reflect a catchability issue in the Japanese CPUE series or an artefact of some other assumption).

All of the models examined to date (that fit the Japanese CPUE trends reasonably well by qualitative inspection), suggested that the overall stock is substantially depleted (e.g. Figure 9). However, one should not read too much into these preliminary results for a number of reasons.

As a proof of concept, it seems reasonable to proceed with SS3 as an appropriate tool for this assessment. This is a fairly complicated implementation, primarily designed to push the limits of the software in the current context. Given the available data and understanding of swordfish, it would be unwise to have much confidence in many of the values that could be estimated (e.g. movement rates, steepness, natural mortality). However, it seems clear that the software was able to efficiently generate these estimates (though global minimization issues were not checked), so it should prove reliable for a more structured exploration of a subset of these features.

A couple downsides to the SS3 software have been identified which are relevant for this assessment:

- SS3 cannot resolve time-area interactions in recruitment variability. This is probably not serious in so far as the size data does not seem to be very informative about year class strength anyway.
- Many of the common stock status reference points generated by SS3 cannot be partitioned by area, so separate models would need to be formulated to examine individual areas under the assumption of discrete populations.

The following is a proposed list of priorities that should be addressed in moving toward an actual assessment:

- Clarify the questions that we are seeking to answer in undertaking this assessment, particularly with respect to the sub-populations.
- Obtain updated standardized CPUE series
    - Japan
    - Taiwan
    - La Reunion
    - Seychelles
- Revisit fleet disaggregation with respect to size composition data and shared selectivities
- Structured examination of the biological uncertainties and interactions among conflicting data.
- Incorporate size by sex data.

Conclusions:

1. Stock Synthesis 3 software provides a powerful, flexible and numerically efficient framework for integrating a diverse range of structural features and statistical assumptions for the assessment of Indian Ocean swordfish (though this does not mean that everything that can be included in a model should be included in a model).

2. Before a formal assessment can be undertaken, several additional data sources need to be provided, including updated standardized CPUE from Japan and Taiwan, and ideally La Reunion and the Seychelles.

3. The implications of many uncertain assumptions need to be examined. Hopefully this paper might encourage discussion of obvious problems before the 2009 WPB. However, given time constraints, it seems likely that many avenues for analysis and improvement will need to be pursued in subsequent years.

4. The sex-specific catch data from some fleets might be useful as well, either integrated directly into the model, or qualitatively in relation to the formulation of

biological assumptions. This is probably a low priority in terms of assessment inferences, but might help structure the debate about migration dynamics.

5. The specifics of the approach pursued should be considered in relation to the questions that the assessment is designed to answer, and clarification on this point should be sought from the broader WPB/SC when planning the next phase of assessment work.

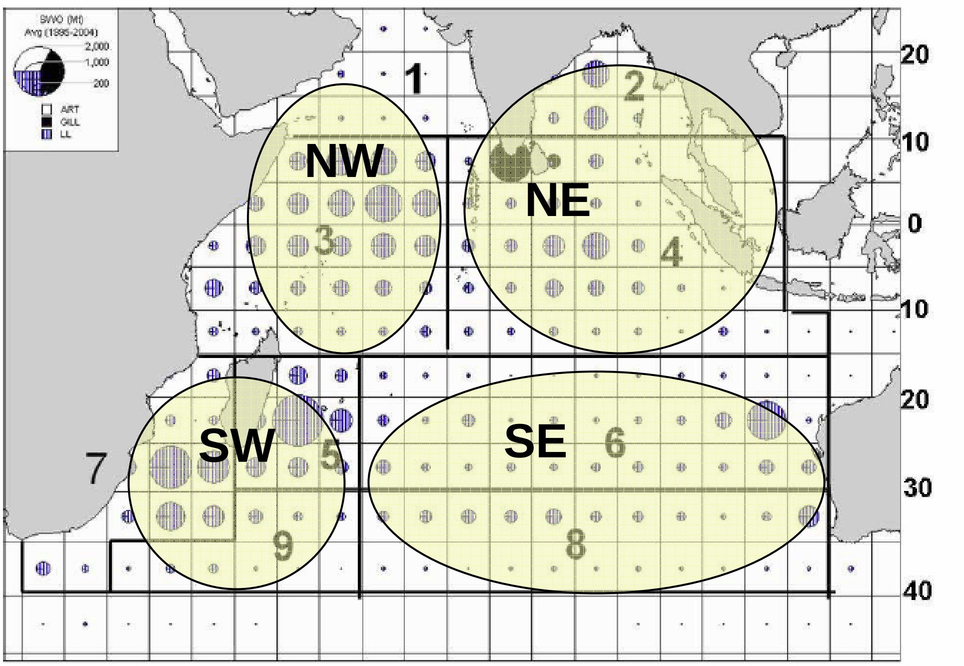

Figure 1. Spatial structure showing the 4 areas used for the exploratory model, superimposed on the IOTC statistical areas and historical swordfish catch distribution.

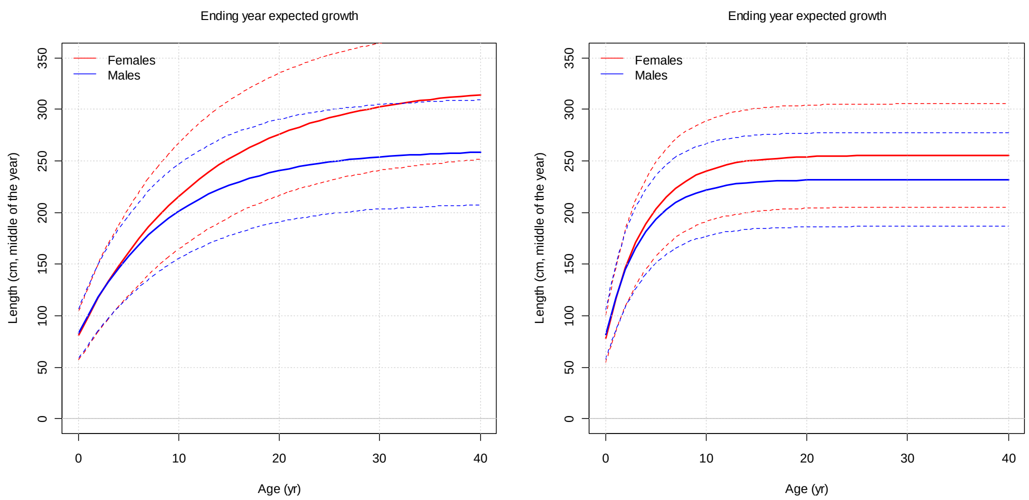

Figure 2. Growth curves (Lower jaw fork lengths) estimated for the Eastern Indian Ocean (CSIRO, top panel, corresponding to example model 1), and north-central Pacific (NMFS, bottom panel, corresponding to example model 2). Age estimation comparisons revealed differences of fin ray annulus interpretation when readers from both labs read the same fin rays. Real differences in biology might account for some of the estimated growth rate discrepancy, but most of it is seems to be attributable to the unresolved problem of age estimation (Young et al. 2008, WCPFC-SC3-BI SWG/WP-1**).**

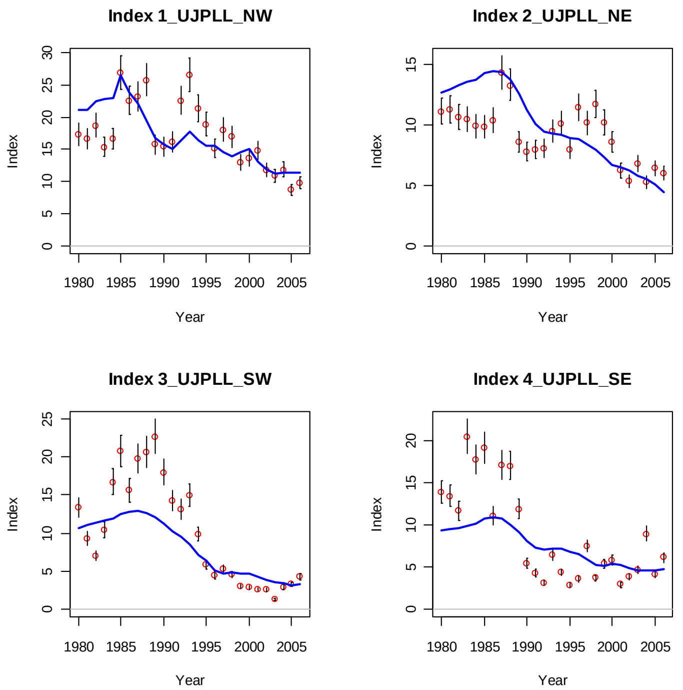

Figure 3. Predicted (lines) and observed (circles with 95% error bars) standardized Japanese longline CPUE for example model 1 (4 panels represent the 4 regions).

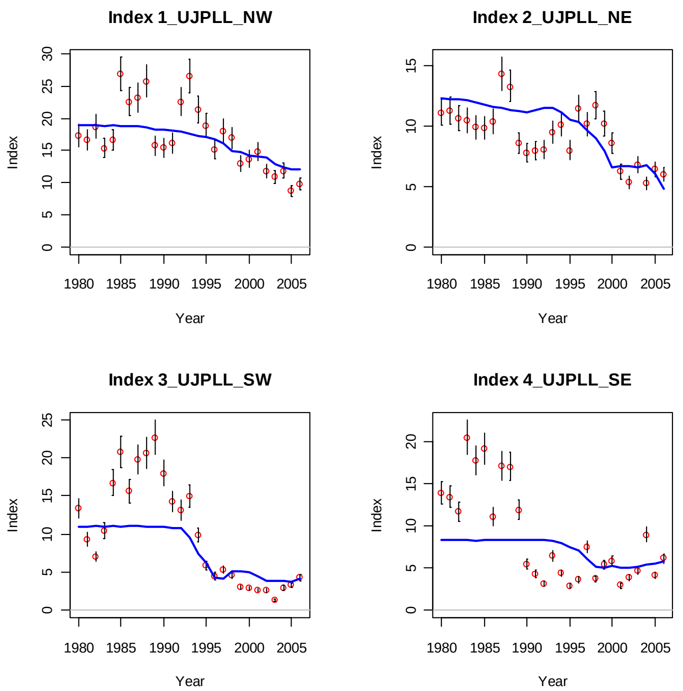

Figure 4. Predicted (lines) and observed (circles with 95% error bars) standardized Japanese longline CPUE for example model 2 (4 panels represent the 4 regions).

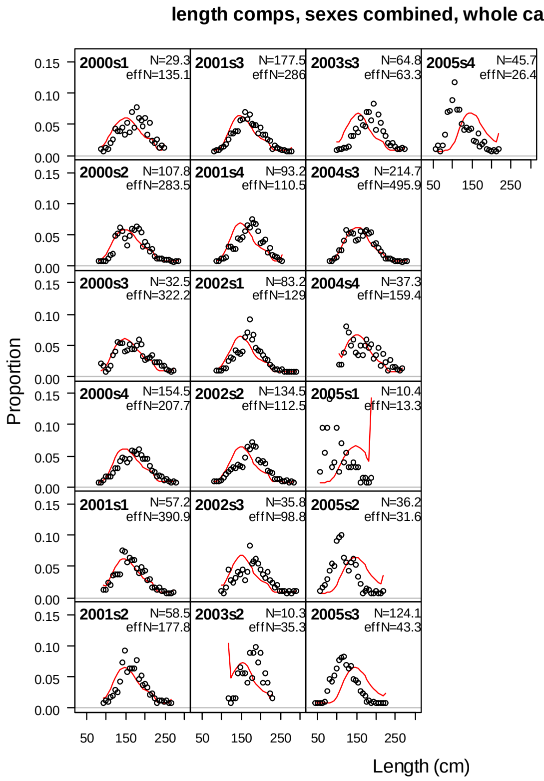

Figure 5. Predicted (lines) and observed (circles) fit to the Japanese size composition data from the SW Indian Ocean from example model 1.

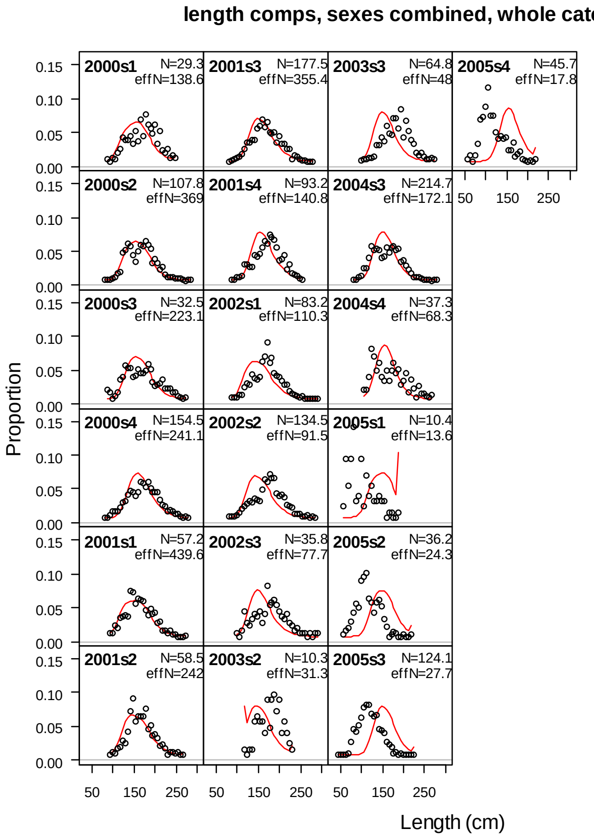

Figure 6. Predicted (lines) and observed (circles) fit to the Japanese size composition data from the SW Indian Ocean from example model 2.

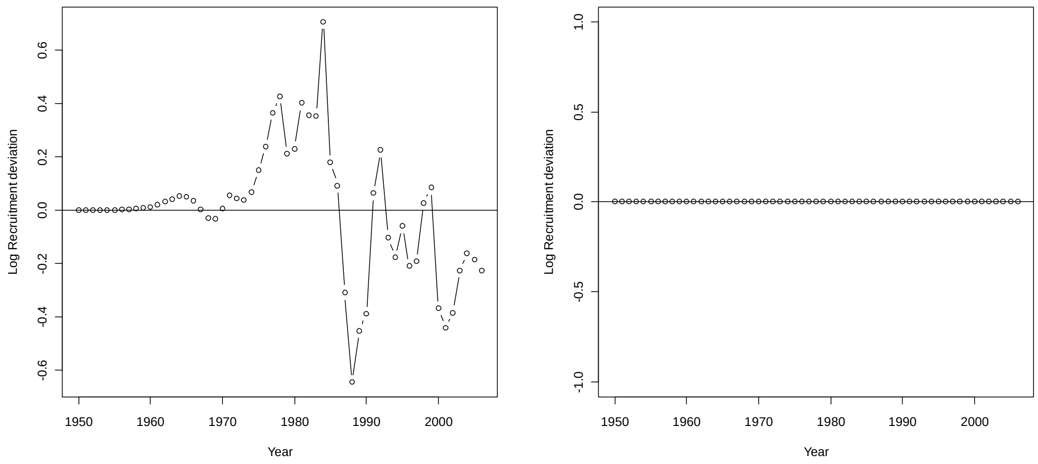

Figure 7. Deviations from the stock recruitment relationship estimated for example model 1 and example model 2 (i.e. the relationship is deterministic in model 2).

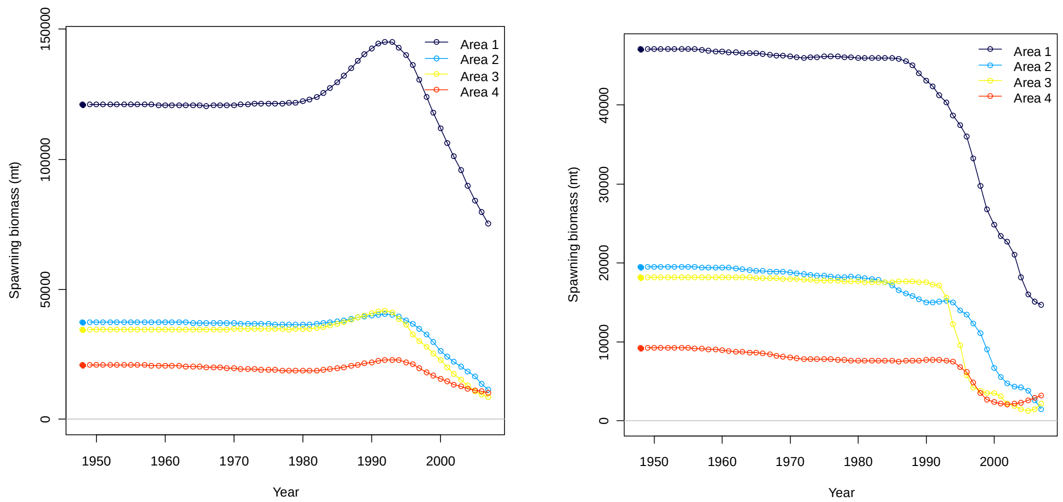

Figure 8. Estimated biomass time series for each of the 4 regions in the Indian Ocean swordfish for example model 1 (left) and 2 (right).

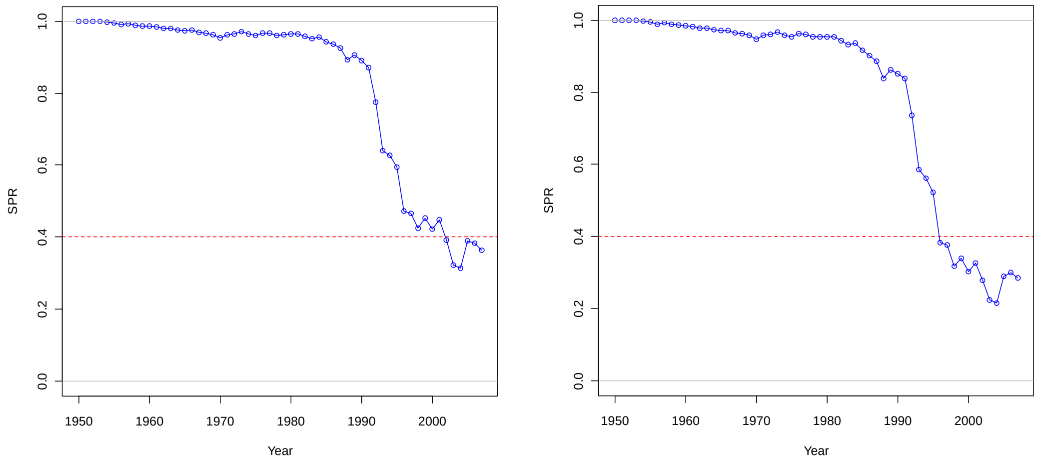

Figure 9. Estimated Spawning Potential Ratio (egg production (t) / virgin egg production) for example model 1 (left) and 2 (right).