SCIENTIFIC COMMITTEE
SECOND REGULAR SESSION

7-18 August 2006
Manila, Philippines

SOUTH-WEST PACIFIC SWORDFISH STOCK STATUS SUMMARY FROM MULTIPLE APPROACHES

WCPFC-SC2-2006/SA WP-7

Paper prepared by

Dale Kolody [1]*
Nick Davies [2]
Robert Campbell [3]

[1] CSIRO Marine and Atmospheric Research, Hobart, Australia
[2] NIWA Ruakaka, New Zealand
[3] CSIRO Marine and Atmospheric Research, Aspendale, Melbourne, Australia
* for correspondence email dale.kolody@csiro.au

WCPFC-SC2 Working Paper
WCPFC-SC2 SA-WP-7

South-West Pacific Swordfish Stock Status Summary from Multiple Approaches

Dale Kolody[1*]
Nick Davies[2]
Robert Campbell[3]

[1] CSIRO Marine and Atmospheric Research, Hobart, Australia
[2] NIWA Ruakaka, New Zealand
[3] CSIRO Marine and Atmospheric Research, Aspendale, Melbourne, Australia
* for correspondence email dale.kolody@csiro.au

Executive Summary

This paper presents a brief overview of assessment approaches applied to SW Pacific swordfish (*Xiphias gladius*) this year, with a primary emphasis on comparing results from Multifan-CL and CASAL models. Each of these approaches is described in considerably more detail in separate WCPFC-SC2 working papers (WCPFC-SC2 ME-WP-3 and WCPFC-SC2 ME-WP-4).

The stock assessment covers the South-West Pacific Ocean (0-50S, 140E-175W) for the period 1952-2004. Swordfish have been exploited in this region primarily as by-catch in the Japanese longline tuna fisheries since the 1950s. Total catches and catch rates remained fairly consistent from about 1970-1996, after which the Japanese fleets were no longer able to access Australian and New Zealand fishing zones, and catches from this fleet have declined steadily since then. Australian and New Zealand catches increased dramatically in the mid-1990s, such that total annual catches in 1997-2004 were roughly double the levels in the preceding period. Pacific Island, Korean, and Taiwanese catches also increased during this period, but remain a small proportion of the total. In the mid-1990s, the Australian fleet gradually expanded offshore with some of the fleet specifically targeting swordfish. Declining catch rates and declining sizes in core areas of the fishery since 1997 have raised concerns about the biological and economic sustainability of the fishery. This assessment attempts to integrate the available fisheries data on total catch, catch rates, and size composition with biological studies on age, growth, reproductive dynamics and stock structure, to provide a summary of the current stock status.

The inferences of most fisheries stock assessment models are sensitive to the subjective assumptions that are required to formulate tractable estimators, and one can generally be more confident in results that are robust to alternative plausible assumptions. In recognition of this fact, we used the swordfish situation as an opportunity for exploring and comparing different models, with primary emphasis on two assessment packages, Multifan-CL and CASAL. Both packages are flexible, generic modelling tools that have rich and overlapping feature sets relevant for describing pelagic fisheries. Different spatial structures were explored, including single area and 5 area models (including single stock with homogenous mixing among areas, and 5 stocks with shared spawning grounds but discrete foraging areas). Additional assumptions related to process and observation variances, and structural constraints among fleets and areas were also tested. In total, several hundred Multifan-CL models were fit, compared with a handful of CASAL models (4 of which are reported here). We qualitatively describe the different models, (including a brief exploration of age-aggregated production models), and review performance to date in consideration of future assessment work.

Assuming that the catch data are reliable, the strongest signals for estimating the impact of the fishery on the swordfish population relate to the declining catch rates in the core areas of the fishery, and the declining sizes in the Australian fishery. The majority of models seem to be able to fit the main data reasonably well. To some extent, the models can interpret the gross features of these data either as a direct impact of the fishery, or via trends in recruitment that are largely independent of the fishery (generally increasing recruitment through the 1970-80s, and declining in the

1990s). Most models suggest that both mechanisms are occurring, with the relative importance of the two driven by sensitivity to the structural and statistical assumptions of the models, and not easily distinguished by the available data.

At this time we are not able to conclude that either Multifan-CL or CASAL are preferable tools for conducting an assessment of this sort, however, we feel that the Multifan-CL assessment had the benefit of an extensive exploration of model uncertainty. The stock status summary is lifted directly out of ME-WP-4, which should be consulted for details. The status summary represents a synthesis of the Bayesian Maximum Posterior Density (MPD, or best point estimate) results from a subset of 10 models (the most plausible ensemble), selected from several hundred results. In the following conclusions, the estimates represent the median (and range) of the MPD results from the plausible model ensemble, such that if one of the models at the extreme end of the range were actually a perfect unbiased estimator, there would be a 50% chance of the true value being more extreme than the uncertainty bound indicates:

1. We consider the relative Total Stock Biomass (TSB) estimates for recent years to be the most reliable reference points, because they are the most closely linked to the highest quality data, and are reasonably robust to the alternative model assumptions explored. The MPD results from the plausible model ensemble indicate:
    - TSB(2004)/TSB(1995) median = 0.70, range = (0.56 – 0.74).

2. All of the Spawning Stock Biomass (SSB – roughly corresponding to age 10+ fish) reference points are much more uncertain than TSB because SSB represents a small portion of the catch, and may be badly biased by natural mortality assumptions, and the model aggregation of sex-specific characteristics of growth, mortality and migration. Furthermore, the southern range of the stock seems to consist predominantly of mature females, but this region is poorly sampled by the fishery and it is difficult to relate abundance in this southern part of the population to the core population.
    - SSB(2004)/SSB(1995) = 0.75 (0.51 – 0.86).

3. The ratio of current biomass over the estimated biomass that would have been observed in the absence of fishing (NF) provides a measure of the fishery impact on the population that might be more meaningful than the biomass ratio at two points in time if the population experiences non-stationary production dynamics (which these assessments tend to suggest).
    - TSB(2004) / TSBNF(2004) = 0.59 (0.31 – 0.69)
    - SSB(2004) / SSBNF(2004) = 0.49 (0.15 – 0.65).

4. The data are not sufficient to estimate a stock recruitment relationship reliably, and most or all models explored suggest some form of non-stationary (or at least highly variable) recruitment dynamics. This seriously undermines the usefulness of the MSY-related reference points. However, in so far as these reference points have been calculated, the majority of MPD estimates from the plausible model ensemble suggest that biomass (total and spawning) are probably above levels that would sustain MSY and fishing mortality is probably below F(MSY).

- TSB(2004)/TSB(MSY) = 1.7 (0.87 – 3.0)
- SSB(2004)/SSB(MSY) = 3.4 (0.75 – 6.4)
- F(2004)/F(MSY) = 0.70 (0.33 – 2.2).

5. The apparent optimism of the MSY-related reference points is countered by the stock projections (assuming constant future recruitment according to the estimated stock recruitment relationships, and constant effort at 2004 levels), which suggest biomass declines over the short term:
    - TSB(2009) / TSB(2004) = 0.88 (0.78 – 1.00)
    - SSB(2009) / SSB(2004) = 0.84 (0.71 – 0.86)

Despite the emphasis on model uncertainty, there remain a number of assumptions which probably influence these conclusions and remain largely beyond the scope of this assessment, including: 1) catchability of the fleets may be changing in ways that cannot be reliably estimated through the catch rate standardization methods employed, 2) the link between our operational definition of the SW Pacific model domain, and the broader Pacific (and possibly Indian Oceans) is unclear, and 3) all of these models ignore sex-specific population characteristics (natural mortality, growth and migration), which may contribute to potential biases in estimators.

Introduction

From 1997-2004, the annual reported broadbill swordfish (*Xiphias gladius*) catch in the SW Pacific (Figure 1, areas 1-5) has been roughly double that of the preceding period 1971-1996 (Figure 2). During this recent period, catch rates and mean size composition have declined substantially in core areas of the fishery, providing the impetus for a formal model-based assessment. This year, two approaches were pursued simultaneously, partly as a test case for comparing assessment software, but also in recognition of the value of exploring model uncertainty. The main details of the assessments, including a description of the data, the rationale for the assumed spatial structure, and specific model assumptions are provided in separate WCPFC SC2 Methods working papers: Kolody et al (2006) describes the Multifan-CL approach and Davies et al (2006) describes the CASAL implementation. This document is a largely qualitative paper intended to draw out the key results from both approaches, and describe how the essential features of the assessment relate to the general impressions that have arisen from the data-based indicators.

Assessment Approaches

A useful assessment model needs to be able to explain the basic trends that are evident in the data. If the results deviate from the expectations that are intuitively formulated when we examine the indicators, then we need to re-examine the plausibility of the models and/or our intuition. Unfortunately, we often have no way of knowing if our assumptions are reasonable, but at least within the context of a model, the assumptions are concisely articulated and open to scrutiny. The modelling process is a useful tool for exploring whether or not our assumptions are consistent with the data and our understanding of the system, and models can be iteratively improved as our understanding of the system improve and additional data are collected. Unfortunately, multiple models with considerably different management implications might be equally consistent with the data but there is no way to ensure that these alternatives will be identified.

In the initial phases of formulating the assessment, we found it useful to think about alternative migration patterns. We refer to two general categories as homogenous mixing and foraging grounds site fidelity (Figure 3). Early perceptions of the stock based on preliminary catch rate standardization suggested that there was a large differential rate of decline among regions, and this warranted some form of spatial dis-aggregation to describe properly. Furthermore, the strong seasonality in CPUE suggested annual migration patterns in relation to spawning (or other seasonal events). However, it proved difficult to reconcile both large seasonal migration and localized depletions within the context of a homogenously mixing population. This was a large part of the impetus behind the development of foraging grounds site fidelity models within CASAL and the production models. However, in the updated iteration of the assessment, the need for this alternative migration representation was greatly reduced.

The intended scope of the different approaches is outlined in Table 1.

Data-Based Indicators

Data-based indicators have been used to monitor the Australian East-Coast longline fishery for several years (e.g. Campbell 2005), and have prompted concerns about the status of the swordfish population. Coincident with the substantially increased regional catches around 1997 (Figure 2), there have been declining catch rates in the Australian and New Zealand domestic fisheries (Figure 4). Catch rates in the Japanese fleet (Figure 5) are much more variable than the Australian and NZ fleets, however, the standardized core area indices for Australia and Japan quite clearly show similar declining trends .

The core Australian fishery has the best size sampling data of all the fleets, and demonstrates a clear decrease in mean size over this period (Figure 7). Size trends are not obvious in the other fisheries (but sample sizes are much smaller).

Multifan-CL

Multifan-CL (e.g. Kleiber et al 2003) is a flexible integrative assessment modeling framework initially developed and routinely applied to the assessment of tuna species of the Western and Central Pacific Ocean (Kleiber and Yokawa 2002 describe a North Pacific swordfish assessment). Multifan-CL attempts to integrate all of the available catch, effort and size data in an arbitrary spatial arrangement, with seasonal dynamics.

The assessment on which this paper is based involved a primary emphasis on model uncertainty, such that several hundred individual models were fit with different combinations of assumptions, and a plausible ensemble of 10 models was defined on the basis of the following criteria (evaluated at the Maximum Posterior Density):

1) adequate numerical convergence
2) minimum quality of fit to core CPUE series
3) minimum quality of fit to size data
4) fishing mortality sensible (i.e. in the southern part of the range where swordfish is caught as by-catch in the SBT fishery, we would not expect high fishing mortality
5) moderate – high stock recruitment curve steepness

CASAL

CASAL (e.g. Bull et al 2003) is also a flexible, generic software package with many of the same features as Multifan-CL. Notable differences include: CASAL has the capacity to disaggregate populations by sex, and to represent multiple stocks. There are many additional subtle differences that can be important in specific situations. For most applications we would generally expect similar advantages and disadvantages between the two packages, but we are not aware of any other direct comparisons between the two.

Spatially-Disaggregated Pella Tomlinson (SDPT) production model

In recognition of the poor size sampling for most of the swordfish fishery, and the notorious difficulty in estimating natural mortality and stock recruitment productivity, simple age-aggregated production models were briefly explored as an alternative to the fully integrated models. The SDPT also provided a means of exploring alternative migration formulations, including different specifications for homogenous mixing and foraging grounds site fidelity. The models used deterministic production dynamics and fit area-specific CPUE as a relative abundance index. Both versions of migration seemed to provide very good agreement with the observed CPUE trends. We do not provide detailed results here, except to note that when the underlying productivity dynamics were constrained in a manner resembling the Multifan-CL production dynamics, the stock status estimates were generally similar. However, given the declining size trends in the Australian fishery, and the potential for long term shifts in recruitment regimes in this stock, it would be difficult to argue that the production models should constitute a primary basis for the assessment.

Model Assumptions

Core assumptions in all of the Multifan-CL and CASAL models included:

Key model assumptions include:

- 10 fisheries (Table 2)
- iterated on a quarterly timestep 1952-2004
- age-structured populations ages 0–19+ (in years)
- age-length-mass relationships are assumed known based on biological studies and cross-calibration of fisheries observations
- populations are sex-aggregated with age-length relationships an average of males and females
- maturity schedule approximates the female maturity
- Fishery selectivity is assumed constant over time (with different assumptions about variation among fleets)
- CPUE is assumed to be highly informative for the 3 Australian fleets (1997-2004) and the NZ domestic fleet (1998-2004), and weakly informative for the Japanese fleets (1971-2004).
- Catch-at-size effective sample sizes are down-weighted (to compensate for non-random sampling, potential selectivity shifts in relation to species targeting changes, and the inadequate representation of sex dimorphism in the model)
- Age-specific natural mortality vectors were fixed input

Contrast in the key assumptions among different versions of Multifan-CL and CASAL as applied in the swordfish assessment are summarized in Table 3.

Comparison of Model Results and Performance

Table 4 describes some of the general characteristics of the performance of the various models as applied to SW Pacific swordfish. Most models fit most of the data reasonably well. But they all had some unresolved problems in relation to some elements of the fit to the data, our prior expectations of stock dynamics, or the numerical performance of the estimators. There was considerable sensitivity to structural and statistical assumptions among the formulations tested and it is unlikely that the available data is sufficiently informative to resolve these issues. Table 5 describes the available stock status estimates resulting from the different models. With respect to the different modelling approaches, we note the following:

- The simplest interpretation of the data-based indicators might lead to an interpretation something like: catch increases since 1997 (catch history in Figure 2) have led to a biomass in 2004 of about 40% of what it was in 1997 (core area CPUE from Figure 6). And since the stock has been exploited for several decades prior to 1997, the stock must have already been below virgin levels in 1997. Interpreted in the context of an assessment model, we tend to conclude total biomass has not declined this much in the last decade because of the non-vulnerable component of the stock, while spawning biomass could be more or less depleted than 40%. Obviously one cannot infer too much from the CPUE trend alone, but the magnitude of the decline does make it difficult to embrace some of the model inferences that suggest the fishery has a negligible impact on the stock.

- We spent the most time on the Multifan-CL 5 Area models, and through the exploration of several hundred alternative specifications, we feel that a subset of the best fitting models provides the best representation of the stock status. While we are reasonably comfortable with the aggregate estimates of the stock status, we do not have much confidence in the migration estimates. This led to brief exploration of single area Multifan-CL models, but the results to date did not achieve a satisfactory fit to the catch-at-size data, and as such these stock status estimates are not reported.

- The spatially-disaggregated CASAL estimators generally fit the core area data reasonably well (except for the size composition in the southern SBT-targeting fisheries). The single area model had trouble fitting the size data. We do not feel that sufficient exploration of the uncertainty was conducted for the CASAL results to constitute an assessment. Figure 8 illustrates the relationship in two reference points for the 4 CASAL models reported in Table 4, and 144 models from the Multifan-CL model uncertainty exploration. This plot suggests that the CASAL MPD estimates tend to clump within a small subset of the space explored by the Multifan-CL uncertainty grid. The CASAL estimates are more optimistic than the 10 model plausible Multifan-CL ensemble, with only a small degree of overlap between the two. However, Table 5 suggests greater overlap for other reference points (and the incorporation of parameter uncertainty would increase the overlap further).

- The SDPT models were computationally convenient, and provided stock status results within the bounds of the Multifan-CL (5 area) estimates. While it is comforting that the results were in line with the more complicated models, we would not consider these models useful as the prime assessment tool for this stock, because: 1) they cannot admit non-stationarity in production dynamics (at least in the deterministic case), 2) they ignore the signal in the size frequency data, and 3) as implemented, there was a circularity in parameterization, in that bounds on production characteristics were approximated in relation to the age-structured characteristics of Multifan-CL.

This assessment has provided a useful opportunity to compare the Multifan-CL and CASAL assessment packages in a challenging context. But at this time, it is not obvious that either package has superior numerical performance or a preferable feature set for this type of assessment. We would encourage continued use of both models.

Stock Status Summary

Given the previous considerations regarding exploration of model uncertainty, this stock status summary is adopted straight from the Multifan-CL assessment (Kolody et al 2006), and we encourage reference to that document.

In the following, we report against reference points with an ad hoc definition of uncertainty bounds based on the outcome of the plausible model ensemble. We report the median and the range of the MPD estimates from the plausible models. This will probably result in a practical interpretation similar to a Bayesian posterior or frequentist confidence interval. However, it is important to emphasize that each MPD result is based on an individual model fitting. If the model at the lower bound happened to have the best assumptions and constituted a perfect unbiased estimator for the quantity of interest, then there would actually be a 50% probability that the true value was outside of the stated uncertainty bounds. This approach to uncertainty quantification has less of a theoretical basis than the usual approach, but we think that it will usually lead to a more reasonable estimate of the real uncertainty in most cases. The following stock status conclusions are presented roughly in order of perceived reliability:

1) We consider the relative Total Stock Biomass (TSB) estimates for recent years to be the most reliable reference points, because they are the most closely linked to the highest quality data, and are reasonably robust to the alternative model assumptions explored. The MPD results from the plausible model ensemble indicate:
    - TSB(2004)/TSB(1995) median 0.70, range (0.56 – 0.74).

2) All of the Spawning Stock Biomass (SSB – roughly corresponding to age 10+ fish) reference points are more uncertain than TSB because SSB represents a small portion of the catch, and may be badly biased by natural mortality assumptions, and the model aggregation of sex-specific characteristics of growth, mortality and migration. Furthermore, the southern range of the stock

seems to consist predominantly of mature females, but this region is poorly sampled by the fishery and it is difficult to relate abundance in this southern part of the population to the core population.

- SSB(2004)/SSB(1995) = 0.75 (0.51 – 0.86).

3) The ratio of TSB relative to the biomass estimated to have occurred in the absence of fishing (TSBNF) provides a measure of the fishery impact on the population that might be more meaningful than the biomass ratio at two points in time if the population experiences non-stationary production dynamics (which these assessments tend to produce).
    - TSB(2004) / TSBNF(2004) = 0.59 (0.31 – 0.69)
    - SSB(2004) / SSBNF(2004) = 0.49 (0.15 – 0.65).

4) The data are not sufficient to estimate a stock recruitment relationship reliably, and most or all models explored suggest some form of non-stationary (or at least highly variable) recruitment dynamics. This seriously undermines the usefulness of the MSY-related reference points. However, in so far as these reference points have been calculated, the majority of MPD estimates from the plausible model ensemble suggest that biomass (total and spawning) are probably above levels that would sustain MSY and fishing mortality is probably below F(MSY).
    - TSB(2004)/TSB(MSY) = 1.7 (0.87 – 3.0)
    - SSB(2004)/SSB(MSY) = 3.4 (0.75 – 6.4)
    - F(2004)/F(MSY) = 0.70 (0.33 – 2.2).

5) The apparent optimism of the MSY-related reference points is countered by the stock projections (assuming constant future recruitment according to the estimated stock recruitment relationships, and constant catches at 2004 levels), which suggest biomass declines over the short term:
    - TSB(2009) / TSB(2004) 0.88 (0.78 – 1.00)
    - SSB(2009) / SSB(2004) 0.84 (0.71 – 0.86)

Assessment Issues for Future Consideration

We refer to the individual papers for discussion on how the assessment might be improved in the future, through the revision of models, and probably more importantly, the collection of additional data. However, we consider that the most valuable contribution from these models might be realized through the development of operating models that reflect the assessment uncertainty, and which might be used to identify robust harvest strategies that use feedback decision rules to manage the uncertainty.

References

Bull, B., R.I.C.C. Francis, A. Dunn, A. McKenzie, D.J. Gilbert and M.H. Smith. 2003. CASAL (C++ algorithmic stock assessment laboratory): CASAL User Manual v2.01-2003/08/01. NIWA Technical Report 124. 223 p.

Campbell, R.A. 2005. Annual indices of swordfish availability in the South West Pacific. Working Paper SA-WP-6 presented to the 1st Meeting of the Scientific Committee of the Western and Central Pacific Fisheries Commission, held 8–19 August, Noumea, New Caledonia.

Davies, N., D.Kolody, and R.Campbell. (2006). CASAL Stock Assessment for South-West Pacific Broadbill Swordfish 1952-2004. Methods Specialist Working Group paper presented at the 2nd meeting of the Scientific Committee of the Western and Central Pacific Fisheries Commission, 7-16 August 2006. WCPFC-SC2 ME-WP-4.

Kleiber, P., Hampton, J. and D. Fournier. 2003. MULTIFAN-CL User's Guide. http://www.multifan-cl.org/usersguide.pdf.

Kleiber, P. and Yokawa, K. 2002. Stock assessment of swordfish in the North Pacific using MULTIFAN-CL. Standing Committee on Tuna and Billfish. SCTB15 Working Paper BBRG-3, 15 p.

Kolody, D.; Campbell, R.; Davies, N. (2006). Multifan-CL Stock Assessment for South-West Pacific Swordfish 1952-2004. Methods Specialist Working Group paper presented at the 2nd meeting of the Scientific Committee of the Western and Central Pacific Fisheries Commission, 7-16 August 2006. WCPFC-SC2 ME-WP-3.

Table 1. Qualitative comparison of assessment approaches used for SW Pacific swordfish.

Assessment Approach	**Data-based Indicators**	**Multifan-CL**	**CASAL**	**Spatially-Disaggregated Pella Tomlinson**
Intended Purpose	Monitoring fishery trends	Monitoring Fishery Trends Quantified predictions of fishery impact on population and potential effects of future management Quantifying Uncertainty for potential use in development of robust Management Procedures for Australian East-coast fishery	Monitoring Fishery Trends Quantified predictions of fishery impact on population	test whether simple models yield similar results to fully integrated models

Table 2. SW Pacific swordfish assessment fishery definitions.

Fishery Number	Area	Fishing Nation(s)
1	1	Japan (plus other DWF and PIN)
2	2	Japan (plus other DWF and PIN)
3	3	Japan (plus other DWF and PIN)
4	4	Japan (plus other DWF and PIN)
5	5	Japan (plus other DWF and PIN)
6	2	Aus
7	3	Aus
8	5	Aus
9	4	NZ Domestic
10	5	NZ Charter

Table 3. Qualitative comparison of key assumptions in the different modelling approaches. Multiple entries on the table indicate alternative assumptions. For Multifan-CL, many interactions among assumptions were explored in the context of model uncertainty (factorial combinations), while CASAL results represented one dimensional sensitivity trials.

Assumptions	MFCL 5 Area	MFCL 1 Area	CASAL 5 Area	CASAL 1 Area	SDPT 5 Area
spatial structure	homogenous mixing	none	homogenous mixing spawning grounds mixing 5 stocks with unique foraging grounds but shared spawning	none	homogenous mixing single stock with foraging grounds site fidelity and shared spawning grounds
number of fisheries	10	10	10	10	5
Recruitment constraints	Beverton-Holt steepness: 0.4, 0.65, 0.9 CV: 0.1, 0.4	Beverton-Holt steepness: 0.65, 0.9 CV: 0.1, 0.4	Beverton-Holt steepness: 0.9 CV: 0.2 Mean YCS = 1.0	Beverton-Holt steepness: 0.9 CV: 0.2 Mean YCS = 1.0	deterministic approximation to age-structured equilibrium steepness = 0.5, 0.9
Mortality (mean over ages)	0.16 0.28 0.24 0.41	0.16 0.28 0.24	0.16	0.16	deterministic approximation to age-structured equilibrium M = 0.2-0.4
CPUE (CV)	Jpn: 0.7 Aus: 0.15 NZ dom: 0.15 NZ charter: >0.7	single aggregate series CV = 0.15	Jpn: 0.4 Aus: 0.2 NZ dom: 0.2 NZ charter: 0.4	single aggregate series CV = 0.15	core areas CV=0.15 peripheries CV=0.7
size sample downweighting	5, 10, 20, 100	5 10	10	10, 50, 75, 100	n/a
catchability	2 groupings: (1) Aus+NZ, (2) all Jpn	only a single aggregate CPUE series was used	3 groupings: (1) Aus (2) NZ (3) all Jpn	only a single aggregate CPUE series was used	all fisheries independent but constrained by priors
effective area of fisheries	1)geographical 2)fished areas 3)non-0 SWO	non-0 SWO used to create aggregate CPUE series	Not used (implicitly assumes all areas equal)	non-0 SWO used to create aggregate CPUE series	geographical
selectivity	2 groupings: SBT-targeting (f=5,10) all others	10 groupings: all fisheries independent	3 groupings: Jpn Aus NZ	3 groupings: Jpn Aus NZ	n/a
sex-structure	aggregated	aggregated	aggregated	aggregated	n/a

Table 4. Qualitative comments on model performance. The 5 area MFCL results are based on the 10 model plausible ensemble, CASAL results on individual models.

Model characteristics	MFCL 5 Area	MFCL 1 Area	CASAL 5 stock	CASAL 1 stock 5 area mixed	CASAL 1 area	SDPT homogenous mixing and foraging site fidelity
fit to CPUE	excellent except Jpn seasonality and Jpn Area 1 increasing trend		Good overall trends, & for Aus seasonality	Poor seasonality	Poor seasonality	excellent except area 1 (but pre-1997 period mostly not included)
fit to Catch-at-Length/Mass	reasonable fit to all series, including Aus declining trend	not satisfactory in initial trials	Excellent fit to Aus series, poor fit to Jpn series in area 5, excellent fit to Aus declining trend,	Reasonable fit to all series areas 1 - 4, poor fit to Jpn series in area 5	Reasonable fit to Aus series, poor fit to other series	n/a
recruitment dynamics	generally estimates some trends in unfished biomass, cannot resolve steepness		often suggests CPUE trend completely driven by recruitment trend	often suggests CPUE trend completely driven by recruitment trend	often suggests CPUE trend completely driven by recruitment trend	(determinsitic productivity)
selectivity	dome-shaped preferred, but non-decreasing plausible		dome-shaped preferred	dome-shaped preferred	dome-shaped preferred; implausible when size data is downweighted	n/a
Migration estimation	do not consider estimates reliable, but stock status inferences are reasonably robust to migration rate priors		Appear to be plausible	Implausible estimates and not able to estimated for the homogeneous mixing option	Not relevant	seasonality predictions excellent but estimates not considered reliable
effective area assumptions in CPUE	inferences reasonably robust to alternatives		Not used implicitly assumed all areas equal	Not used implicitly assumed all areas equal	Implicit in the single CPUE time series	priors largely over-ridden by catchability estimation
Fishing Mortality	often estimated to be dubiously high in the SBT bycatch fisheries for some age classes		Estimated to be low	Estimated to be low, but high for models assuming homogeneous mixing	Estimated to be low but high for models when size data is downweighted	not examined
Numerical convergence issues	generally reliable (<60m per MPD)		Generally fast (<60 m)	Migration parameter estimation unreliable for models assuming homogeneous mixing	Fast (<15 m)	seconds

Table 5. Key stock status reference points for the various SW Pacific swordfish assessment approaches explored. The 5 area MFCL results are based on the 10 model plausible ensemble, and indicate the median and range of the best point estimates (Maximum Posterior Densities). Each CASAL result represents a single model. SDPT results are the range from point estimates of 6 models. TSB = Total Stock Biomass, SSB = Spawning Stock Biomass, NF = No Fishing, F = aggregate fishing mortality (catch/population)

Management Quantity	**CPUE trend (fig. 6)	MFCL 5 Area	CASAL 5 stock	CASAL 1 stock 5 area mixed	CASAL 1 area	CASAL 1 area - size data downweighted	SDPT homogenous mixing and site fidelity
TSB(2004)/TSB(1995)	~0.4	0.696 (0.563-0.74)	0.84	0.77	0.85	0.73	
SSB(2004)/SSB(1995)	~0.4	0.753 (0.509-0.862)	0.93	0.82	0.95	0.86	
2004 TSB/TSB(NF)	<0.4	0.586 (0.312-0.694)	0.78	0.63	0.94	0.62	
2004 SSB/SSB(NF)	<0.4	0.487 (0.148-0.654)	0.75	0.58	0.94	0.58	
TSB(2004)/TSB(MSY)		1.72 (0.873-2.97)	n.a.	n.a.	n.a.	n.a.	1.03 – 1.61
SSB(2004)/SSB(MSY)		3.35 (0.749-6.43)	n.a.	n.a.	n.a.	n.a.	1.03 – 1.61
TSB(2009)/TSB(2004)		0.876 (0.782-0.998)	n.a.	n.a.	n.a.	n.a.	
SSB(2009)/SSB(2004)		0.802 (0.714-0.861)	n.a.	n.a.	n.a.	n.a.	
***F(2004)**		0.0569 (0.032-0.159)	0.023	0.042	0.006	0.051	
***F(2004) /F(MSY)**		0.7 (0.326-2.24)	0.232	0.218	0.052	0.513	0.52 – 1.41

* note that Multifan-CL F(2004) fishing mortality results are presented in biomass, CASAL in numbers

** CPUE trend interpretation using the implicit assumptions: CPUE proportional to biomass, stable unfished equilibrium at B(0) and the stock is not in a virgin state in 1995

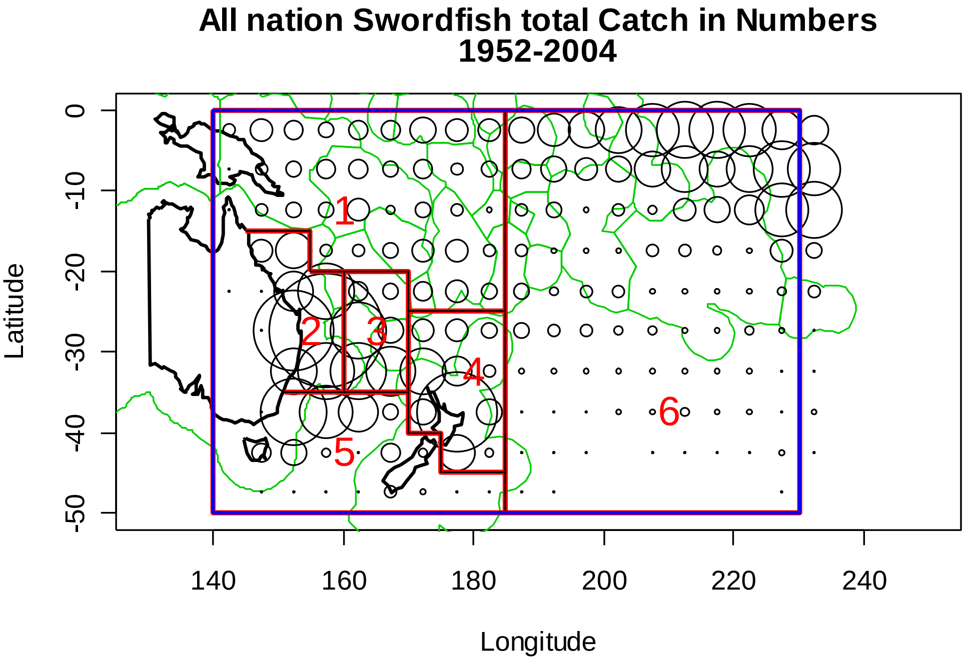

Figure 1. Spatial considerations in the development of the SW Pacific swordfish assessment. Regions 1-5 correspond to the core assessment area, where we have the best understanding of the fisheries data and biology. Area 6 was initially defined for sensitivity trials but this was not pursued. The area of the circles represents the relative catch (numbers) in each 5x5 degree square summed over 1952-2004.

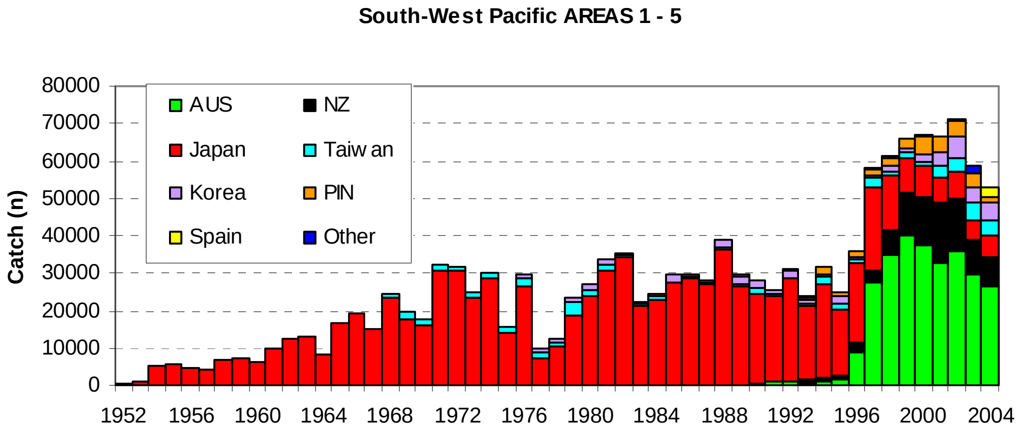

Figure 2. Total swordfish catch history in the SW Pacific defined as areas 1-5 in Figure 1.

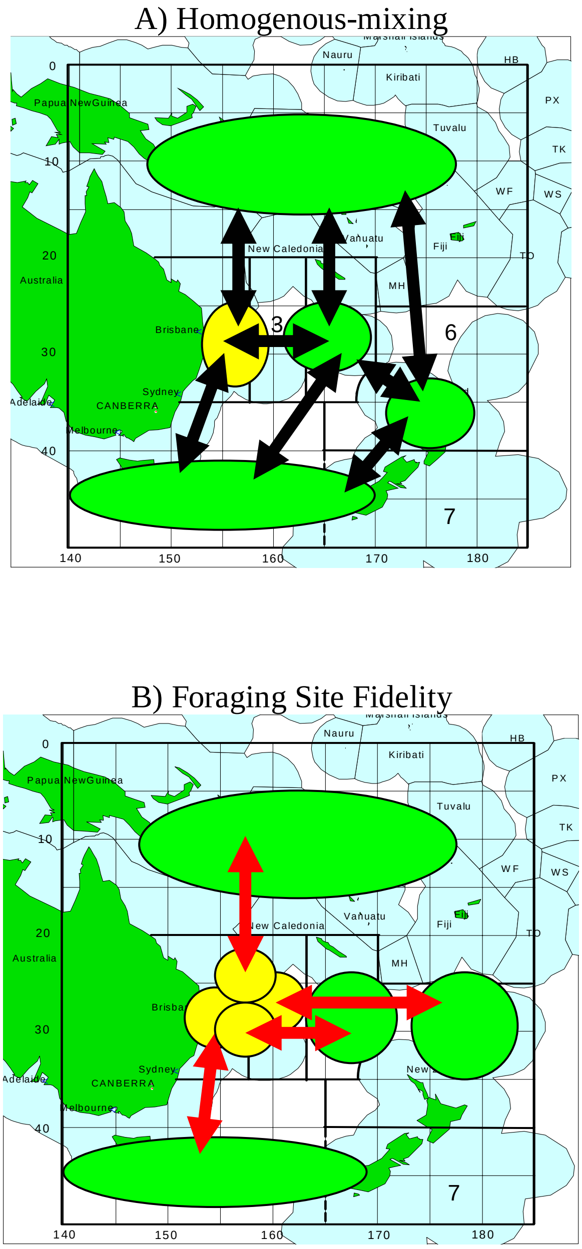

Figure 3. Schematic representation of the different migration models discussed. Arrows indicate possible movement links, ovals indicate sub-populations (green indicates foraging grounds; yellow indicates spawning grounds). Panel B represents foraging grounds site fidelity in which mature individuals always return to the same foraging areas (It could be a single stock or multiple stock situation, depending on whether the spawners and larvae mix).

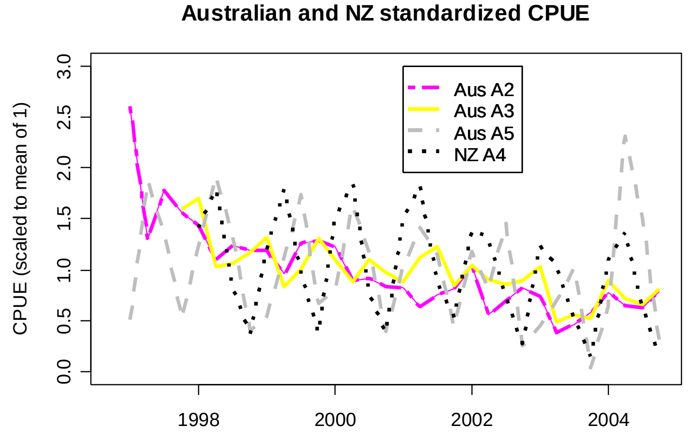

Figure 4. Standardized Catch rates (normalized to a mean of unity) for the Australian and New Zealand domestic fleets over the time period that we assume CPUE provides a useful relative abundance index.

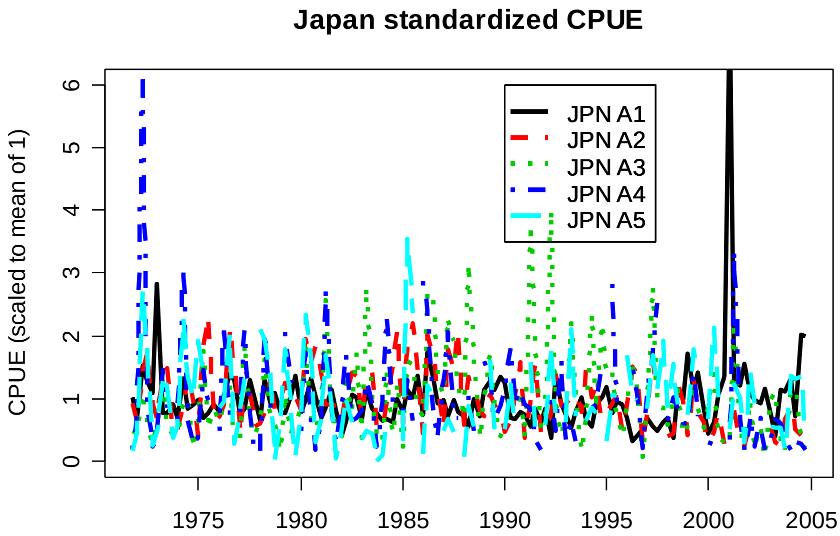

Figure 5. Standardized catch rates (normalized to a mean of unity) for the Japanese fleet over the time period considered to be informative for the assessment.

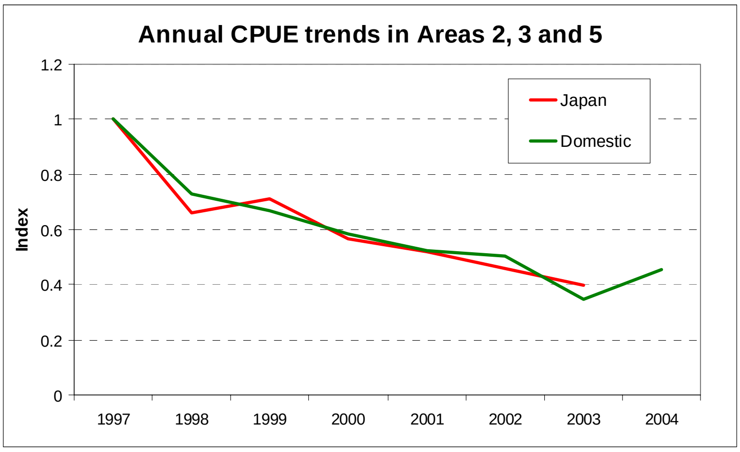

Figure 6. Comparison of the annual standardized catch rates of the Japanese and Australian domestic fleets in the regions where operations overlap in the SW Pacific.

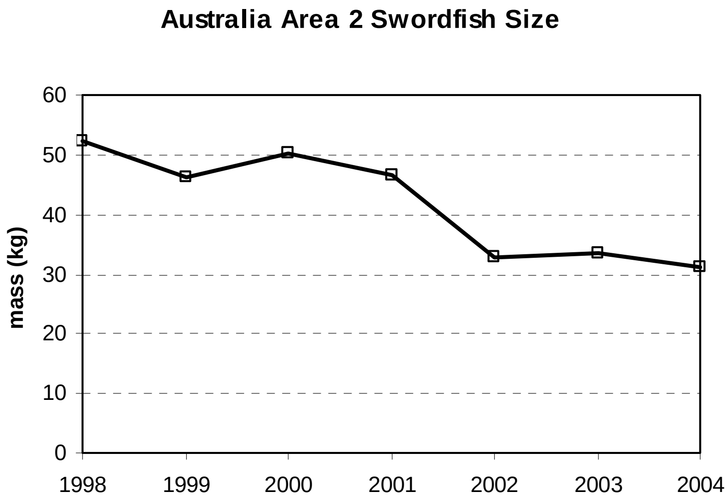

Figure 7. Declining size (trunked mass) trend in the core Australian swordfish fishery.

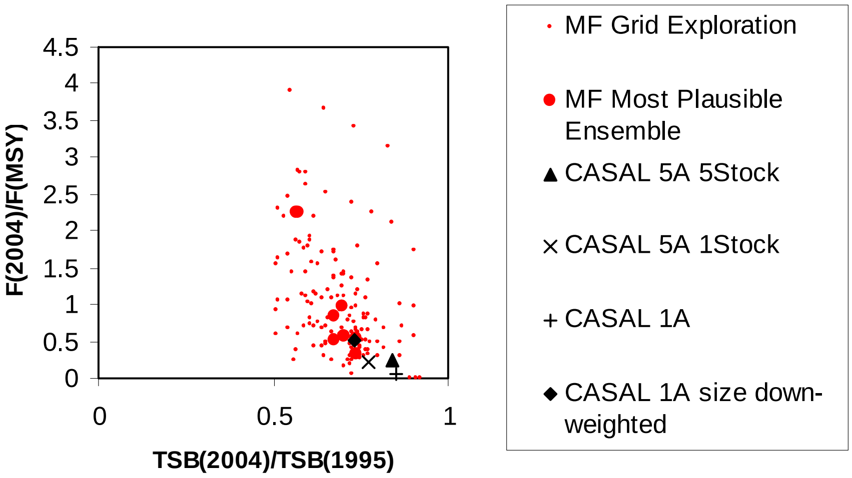

Figure 8. Joint uncertainty in fishing mortality and total stock biomass reference points illustrating the degree of overlap among the MPD estimates from a selection of CASAL and Multifan-CL models,