CCSBT-MP/0304/09

ADDITIONAL COMMENTS ON OPERATING MODEL SPECIFICATIONS FOR EVALUATION OF SBT MANAGEMENT PROCEDURES

Dale Kolody
Tom Polacheck

Prepared for the CCSBT 2nd Management Procedure Workshop
7-9 and 13-14 April 2003, Queenstown, New Zealand, CCSBT-MP/0304/09

Table of Contents

Abstract

The development of Management Procedures (MPs) for Southern Bluefin Tuna (SBT) requires an operating model that can simulate a range of population dynamics, fishery behaviour, and data collection methods that represent a diverse range of plausible scenarios (including some that are difficult to quantify with existing stock assessment methods). In this document, we make specific proposals and qualitative comments about several features of the SBT operating model (historical conditioning and future projections) that will need to be considered for MP robustness trials, particularly in reference to issues flagged at the CCSBT 2002 SAG/SC. In addition to these features, we suggest that there may be merit in having an operating model uncertainty hierarchy that provides perfect data, so that we can distinguish the degree to which candidate decision rules are succeeding on the basis of estimator quality, or rebuilding strategy.

Introduction

The Commission for the Conservation of Southern Bluefin Tuna (CCSBT) has agreed to the development of Management Procedures (MPs) for the SBT fishery (Anon. 2002). The success of the MP evaluation procedure is dependent on the simulated fish and fishery dynamics in the operating model, so it is important that this model represents a reasonable range of plausible dynamics. We need to express both the uncertainty in the current state of the stock, and the uncertainty about what may happen in the future. The more thoroughly that this is achieved, the more confidence we have that the MPs will be robust.

We have suggested functional relationships and/or parameter values below as a starting point for further debate, however, we have not made suggestions about how to deal with the potentially large number of dimensions that will arise if variability in each factor is considered independently in all possible combinations with every other factor.

The numbered headings below follow the numbering scheme of the CCSBT SC Report 2002, Attachment 4 (except for the duplicate numbering errors in that report).

The case for an operating model with perfect data

We have found it useful to think about the operating model in two more or less distinct parts: 1) fundamental stock characteristics that determine the ultimate limits to the catch and stock size, (eg the stock and recruitment relationship, growth rates, natural mortality, and current state of the stock), and 2) data characteristics that limit the accuracy and precision with which one can make inferences about the stock characteristics (eg relationship between CPUE and abundance). Perfect knowledge about the stock characteristics would allow one to develop truly optimal management strategies (provided that a unique performance indicator could be agreed). However, the data characteristics limit the knowledge of stock characteristics, and management strategies require lower catches to maintain the same level of biological risk when inferences are uncertain. By providing an (unrealistic) operating model with perfect data, we may be able to partition the search for decision rules, with different emphasis on rebuilding strategies and efficient estimators. This may assist in designing rules with improved overall performance. In addition, these perfect data cases could also provide a reference point against which realistic decision rules could be compared. As such, we think it would be worthwhile to have an additional uncertainty hierarchy in which perfect information was provided to the management procedure.

1) CPUE-abundance relationship

A generalized relationship between longline CPUE and abundance was proposed in the CCSBT-SC Attachment 4, with 7 user-defined parameters. Given the complexity of the functional relationship (e.g. it is a seven parameter function), there is little basis for an *a priori* selection of an appropriate range of parameter values. Polacheck and Kolody (2003) present results from conditioning the oprerating model over a range of values for the various parameter values. We propose that those conditioning results plus other ones performed by other members and any additional ones run at the April workshop be used as a basis for selecting a set of functional relationship to use in the next stage of testing candidate management procedures.

2) Selectivity

It is difficult to know how a fishery would change its targeting practices in response to economic conditions, or changing age composition. However, given that we believe it has happened in the past, and seems to create problems for CPUE interpretation, we believe that this is something that we should be admitting. The 2002 SAG/SC agreed that "...random selection of historic (estimated) selectivities in the projections phase is not desireable..." The most appropriate solution might be to use a selectivity function that reflects the economic utility of different age classes as proposed in Kolody et al (2002), however, given the difficulty in defining future economics, we propose the following combination of random and systematic selectivity selection for future projections.

All fisheries: Random noise is introduced into the selectivity by simulating a binomial sampling process with probabilities determined by the age-specific fishing mortality probability, and (in a fashion analogous to the catch-at-size sampling), an imposed Effective Sample Size (ESS) that is artificially low (eg ~50-200).

ie (depending on the fishing mortality parameterization), if F(a) is the deterministic multiplicative fishing mortality (eg C(a) = F(a)N(a)), then F'(a) is the stochastic fishing mortality:

F'(a) = Random_Binomial_Sample(N = ESS, P = F(a)) / ESS

This will produce deviations from the deterministic selectivity which are qualitatively similar to what one might expect with a patchy age-structured spatial distribution. The CV will be greater for the less selected age classes. It might also be argued that less abundant age classes might have a higher CV than more abundant ages, which could easily be implemented by linking ESS to the absolute numbers-at-age, eg

ESS' = ESS(N(a)/1000000)

In addition to this random noise we propose the following selectivity deviations which have time series trends linked to the age structure:

Longline fisheries: We propose a simple function that causes selectivity to deviate from a baseline according to the age-structure, such that a large cohort will be disproportionately targeted, and "followed" to some extent as it ages.

H'(a) = H(a)*(P(a)/P*(a))^k
where
H(a) is a baseline (equilibrium age-structure) selectivity
P(a) is a baseline (equilibrium age-structure) proportion-at-age, and
k >1 (k= 0 is the standard selectivity parameterization)
H'(a) would probably be re-scaled to a mean of one in the actual implementation

One obvious drawback of this parameterization is the decrease in selectivity of cohorts adjacent to a large one, which may be unrealistic if the cohorts mix (a size or age-based smoothing of adjacent selectivities may be simultaneously desired)

For the juvenile fishery, we propose that the selectivity be given considerable freedom to approach a constant catch composition, for example:

H(a) = (beta)H'(a) +(1-beta)H*(a)
where
H* is a baseline selectivity (eg mean over the last 10 years historical)
H' is a flexible selectivity that that changes so that the age composition of the catch is identical over time.

eg assuming a fishing mortality relationship C=FN:

F'(a) = P(C'(a))/P(N(a)),

C'(a) = the desired catch age composition (eg mean over last 10 years), and

beta = the relative weight between the baseline selectivity and desired selectivity

H'(a) = F'(a) / (Effective Effort); ie assuming F = H*Effort

We would not expect these parameterizations to be used in historical conditioning.

3) Regime Shifts / stock recruitment auto-correlation

We think a reasonable approach for approximating the effects of recruitment regime shifts (within a 20 year time horizon) is to impose a high (>0.8) auto-correlation in the deviations from the stock recruitment relationship. In doing this, there needs to be a sensible relationship between the degree of auto-correlation and variance around the stock recruitment curve – e.g. high correlation with low variance does not result in dynamics which approximate regime shifts.

In this regard, we note that the degree of auto-correlation and variance in stock recruitment in the conditioning results to date indicates that these are related to the level that steepness is fixed at in the stock/recruitment relationship. In particular, fixing steepness values at 0.90 results in both high variances and auto-correlations in the residuals around the stock/recruitment curve. This suggests that a steepness value of 0.90 is only supportable by the data if there are large auto-correlations and a relatively high CV (or equivalently a regime shift). As such, if there is agreement that we should impose stock recruitment steepness that is not strongly supported by the data (eg h = 0.9), then we think it is also necessary to use the resultant CV and auto-correlation (empirical calculation) from these scenarios in the projections. Conditioning explorations indicate that low steepness scenarios tend to have lower auto-correlation (and a lower CV). Thus we feel that projections in low productivity scenarios are consistent with or without strong auto-correlation (we would consider the low auto-correlation scenario to be a higher priority, but both are supportable).

Invoking high auto-correlation to simulate a recruitment regime shift seems to be plausibly consistent with past observations of SBT dynamics, but we should note that this is not as potentially difficult situation for managers as an abrupt regime shift (ie an instantaneous and sustained shift in mean productivity).

4) Variability in Recruitment

The fit of the operating model to the data, and inspection of the available Indonesian direct-ageing data do not suggest very high recruitment variation among consecutive cohorts. However, we feel that sigma = 0.4 should be maintained as a minimum value, and higher values should be used, if indicated from the conditioning results. See the section above for discussions of the relationship between steepness, CV, auto-correlation and regime shifts.

5) Depensation

Simple exploration of SR depensation parameter values indicated that the data were broadly consistent with depensation (Polacheck and Kolody 2003), and we suggest that this should be considered: (phi = 0.2 recommended; values from 0.1- 0.2 yielded very similar results in conditioning trials).

6) By-Catch

In the current projection component of the operating model, it is assumed that Indonessian SBT catches will be regulated in a similar manner to other SBT fisheries. However, Indonesia is not currently a member of the CCSBT and a large component of the SBT catch is a by-catch of their tropical longline fishery for yellowfin and bigeye. As such, provisions need to be made in the testing of candidate management procedures and in any full specification of a management when it is implemented for the possibility that Indonesia's SBT catches will not be regulated. We suggest that scenarios be tested in which the TAC set by a candidate management procedure is assumed to apply only to CCSBT members and that the Indonesian catches are determined by a constant fishing mortality rate equal to the average of the last three years. In doing this, the estimates of the catch of SBT by Indonesia in the current year would be available to a candidate management procedure. This is based on the expectation that the current longline catch monitoring programs in Indonesia will continue and provide reasonably accurate estimates of its catch (and its age distribution).

7) Availability of tagging data

In the initial development and testing of candidate management procedures, it was agreed that no tagging data would be available from current and future SBT tagging programs. As this point, we do not think that it would be productive to attempt to simulate future tag releases and also the tag returns from current and future tagging. This is because of the large uncertainty associated with the potential success of the tagging programs related both to the likely levels of commitment to future releases and more importantly the uncertainty about whether reporting rates can be meaningfully estimated given current levels of observer coverage in longline fisheries. This is unfortunate in that we think that results from successful tagging programs potentially could be highly informative in improving the performance of candidate management procedures. Thus, while attempting to base a management procedure on tagging data would appear to be too speculative at this time and could delay the management development procedure process, a decision not to incorporate tagging data should not preclude consideration of its use in any future revision to an agreed management procedure.

8) Sample Sizes

The effective sample sizes for the historical catch-at-length or catch-at-age data are parameter values which need to be specified in conditioning the operating model. Conditioning results can be sensitive to the specified effective sample sizes (Polacheck and Kolody 2003). In addition, sampling effort and its spatial/temporal coverage has varied greatly over time (Eveson and Polacheck, 2002). However, the effective sample sizes used in conditioning do not reflect this variability. Effective sample sizes are assumed to be constant except for their pre-1965 level in two of the longline fisheries. Eveson and Polacheck (2002) developed a method to assign relative weights or effective sample sizes among years within a specific fishery based on information on the sampling fraction, coverage and whether weight samples were used to estimate length distributions. They applied their approach to longline fishery 1 and the surface fishery. The approach can also be extended to longline fisheries 3 and 4. Application of this approach can not fully solve the question of what effective sample sizes should be used in conditioning as it only provides an estimate of the relative effective sample size between years and an absolute value needs to be specified in one year to scale the time series of relative estimates. In addition, the approach cannot be applied to longline fishery 2 or to pre-1965 longline data.

Because there are a large number of possible combinations of effective sample sizes that could be considered plausible for the historical data, and because the results from conditioning can be sensitive to the values selected, determination of an appropriate set to represent the underlying uncertainty is complex. Polacheck and Kolody (2003) present some results of conditioning the operating model for different effective sample sizes. These results along with others tabled or produced at the upcoming workshop should be used as a basis for determine what range of effective sizes should be used in final conditioning trials.

In addition, there is a need to consider what range of uncertainty should be included in the catch at age data generated by the projection software and made available to candidate management procedures. There is also the question of whether information on the precision of the catch at age data are to be provided for management. We think that these specific issues are probably best left to the next SAG meeting since at that point it should be clear whether any of the likely candidate management procedures are actually using this information.

9) Growth Models

A number of studies have suggested that SBT assessment models are sensitive to assumptions about the quality of the early catch-at-length data (eg Kolody and Polacheck 2001, Polacheck and Kolody 2003). This may in part be due to unknown sampling characteristics prior to 1965, or alternatively, it may reflect good sampling, but a change (possibly density dependent) in the length-at-age over time (Polacheck et al 2002). Simple explorations of the limited available data suggest that a smaller maximum length for cohorts born prior to 1960 is a plausible hypothesis and would be consistent with a density dependent response to the large reduction in the spawning stock in the 1960s. It may be worth explicitly adding an alternative growth curve that admits this possibility (conditioned with some simple assumptions about early spawning ground selectivity and the CL frequency distribution to constrain the L(infinity)). Consideration of alternative growth curves may provide insights to the relative plausibility of different steepness and regime shift scenarios as assumptions about early length frequency data can have a substantial effect on the stock/recruitment relationship estimated in the conditioning process. Suggestions for possible parameterisations for growth curves for cohorts born prior to 1960 are presented in Polacheck et al (2002).

10) Spawning Biomass

Preliminary attempts to estimate spawning potential (Bravington, 2003) suggest that spawning potential varies considerably with age, however, relatively minor changes in relative effective SSB time trends seemed to result (relative to the case with spawning potential directly proportional to fish mass) when compared with the SSB estimates from Kolody and Polacheck (2001). There may however be an important difference, in that the direction of the trend in (effective) SSB in the most recent years changes from increasing to decreasing in some scenarios. Attempts to explore this explicitly in the CCSBT operating model need to be undertaken (but there seems to be a problem in the current version of conditioning software).

11) Errors in Catches

The current operating models assume that the absolute level of historic catch (either number or weights are known without error). This assumption is made both in the conditioning the operating model to historic data and in future projections used to test the management procedures. Uncertainty exists about the historical catches, particularly in the period since the introduction of restrictive quotas in the 1980s. Alternative hypotheses to represent this uncertainty have been used in past CCSBT assessments. These or alternative ones could be developed by the workshop and considered for use in conditioning the operating model. However, past assessment results have not been highly sensitive to this uncertainty. As such, this could be considered a lower priority than a number of the other issues discussed in this paper. However, the errors in catch can have a large effect on predicted performance of candidate management procedures if there are substantial implementation errors in future catch (i.e. differences between specified TACs and realized catch), particularly if there is a discontinuity in the error in catches between the historic time series used in conditioning and in future projections. In this regard, scenarios which assume an implementation error in future catch would be worth considering. Assuming that future catches are in fact 10 or 20% greater then the implemented TAC would provide an indication of the robustness of candidate management procedures to such errors.

12) Further Coding Issues

Proposals for alternate fishing mortality specifications. It is a concern that selectivity patterns become severely altered when fishing mortality is very high. If we have found that this is a problem in the first stage of testing, we support Vivian Haist's suggestion of using the instantaneous fishing mortality parameterization.

Operating Model catch mass calculations

In the current operating model, inconsistencies and problems in interpretation can arise between the specified TAC, the actual catch mass, and the observed catch mass that need to be recognized.

The difference between the actual catch mass and observed catch mass is probably best considered in the context of TAC implementation errors. As the operating model is currently defined, even though the catch mass is removed exactly, there are observation errors in the catch-length sampling, such that scaling the observed length frequencies by the mass-at-length, will have a random deviation from the actual TAC depending on the sampling characteristics. It would be more realistic to have TAC implementation errors that are actually somewhat greater than this error, since it does not really make sense to be able to always achieve the TAC more accurately than it can actually be calculated from the observed catch.

The difference between the historical (2001) TAC and the 2001 catch is a more difficult issue. The operating model is conditioned on LL1 and LL3 catch in numbers. However, the selectivities do not result in a perfect match between the predicted and observed catch length frequency distributions. This translates into discrepancies between (1) the historic total catch mass in any given operating model scenario, (2) the total catch mass calculated by multiplying the "observed" historical catch at length data times the weight/length relationship and (3) the "official" historical catch statistics in weight. These discrepancies may create potential problems and confusion in either interpretation of trial results or implications for actual implementation. These included:

1. The fact that the actual most recent catches (i.e. 2001) used in the simulations differ between operating scenarios confounds comparison of the catch levels realized by candidate management procedures relative to "current" catch levels.
2. The fact that the actual most recent catches (i.e. 2001) differ between operating scenarios means that an initial fixed specified TAC level by a candidate management procedure may result in an implicit cut in catch in some scenarios and an increase in others.
3. For all of the eight initial operating model scenarios, the "true" catch in 2001 appears to be on the order of 5% greater then the estimated from the observed data going into the operating model. However, the quota value provided to a candidate management procedure for the 2001 quota (catch) to be used in setting the 2002 TAC is the "true" catch for that scenario. This implies that there is a consistent "implementation" bias built into the historical data in conditioning that is not carried into projections – i.e. that there will be no discrepancy in the future between "true" catches and the "observed" catch.

4. The "official" catch statistics for the most recent years differ from the catch weights being used in the operating model. This discrepancy means that the TAC "currency" in the projections is not equivalent to the "currency" being used in TAC setting discussions in the CCSBT. This means that in any actual implementation that there will be a need to determine an appropriate "conversion" factor in order to achieve the intended performance of a candidate management procedure. In addition, there needs to be a degree of caution in presenting trial results so as not to create incorrect expectations about the implications for future catch levels.

References

Anon. 2002. Report of the first management procedure workshop. Tokyo, Japan, 3-4,6-8 March, 2002

Bravington, M., T. Davis and J. Farley. 2003. Estimating relative per capita SBT egg production as a replacement for SSB. CCSBT-MP/info/1

Eveson, P. and T. Polacheck. 2002. A Method for Determining Relative Weighting Factors for Length-Frequency Data. CCSBT-ESC/0209/33

Kolody, D., T. Polacheck, M. Basson and A. Preece. 2002. Issues on process and observation models to be considered for the SBT fishery operating model used to evaluate management procedures. CCSBT-SC/09/34

Kolody, D. and T. Polacheck. 2001. Application of a statistical catch-at-age and –length integrated analysis model for the assessment of southern bluefin tuna stock dynamics 1951-2000. CCSBT-SC/0108/13

Polacheck, T. and D. Kolody. 2002. The behaviour and fit of alternative operating model specification for testing the performance of southern bluefin tuna candidate management procedures. CCSBT-MP/0304/7

Polacheck, T., G.M. Laslett and J.P. Eveson. 2002. An integrated analysis of the growth rates of southern bluefin tuna for use in estimating catch at age in stock assessments. CCSBT-SC/0209/35