CCSBT-ESC/0809/30

`

The potential use of indicators as a basis for management advice in the short term

Marinelle Basson
Campbell Davies

Prepared for the CCSBT 5th Management Procedure Workshop 2-7 September and the 13th Meeting of the Extended Scientific Committee 8-12 September 2008
Rotorua, New Zealand

Table of contents

Abstract 1
1. Introduction 1
2. Indicators as a basis for advice in the Short term 3
    2.1 Suggested indicators 3
    2.2 The decision rule 6
    2.3 Exceptional circumstances 10
    2.4 A role for 'scenario modelling'? 10
    2.5 Data requirements for a decision rule 11
3. A basis for advice in the Medium and Longer term 11
References 13

Abstract

We suggest the use of indicators in a simple decision rule as an alternative to using scenario modelling as a basis for advice on short term future catches. One of the main reasons for this suggestion is the authors' concerns about the scientific credibility of work based on the catch (and associated CPUE) scenarios which have not been arrived at through a scientific process. We describe the steps involved, and provide examples for approaches to take, in developing a decision rule based on indicators considered to be unaffected by the historic unreported catches. Specifying an indicator-based decision rule is not much different from formalising the underlying decision-making process that has been used by the SAG and SC over several years in the recent past to provide advice on catches to the Commission. It is appropriate that the details of such a decision rule be defined within the context of the Management Procedure Workshop and/or Scientific Committee meetings of the CCSBT. The intention and hope is that this paper will assist in that process.

The last section of the paper describes an approach for developing a management procedure, with associated testing framework, based on data that are less fishery dependent than longline CPUE which was used in the first MP development process. The approach is sketched in conceptual terms at this stage, but with the intention of developing a more detailed outline of the approach and work required over the coming year.

1. Introduction

The Work plan of the SAG/SC includes the development of methods for constant catch projections with a view to providing advice to the Commission in 2009 on stock status and the short-term risks of different TAC levels (CCSBT-SC12 Report, Attachment 10). This schedule was determined at the 2007 (SC12) meeting. At that point the notion was that scenario modelling could be used as a basis for advice, but there was an expectation that:

1. more information on the uncertainties in the unreported catch would be gathered
2. more information about whether/how CPUE might be affected would come to light

This expectation was explicit in the sense that an item to that effect is also included in the work plan ("Further examination of issues related to the Japan Market Review and Australian Farm Report").

No further information has been made available to the SAG and SC regarding either of these important points. Much work has been done on CPUE standardisation over the past year, but again, this is has been based on historic data with no additional information or data on whether unreported catches have affected CPUE or not.

In our view, there are three substantive concerns about current attempts to re-condition the operating model using the various catch and CPUE 'data' scenarios:
1. there is little basis for confidence in the input catch and CPUE data to the model, particularly for the recent past (mid to late 1980's to at least 2005)
2. the uncertainty included in the catch scenarios, and therefore also the CPUE scenarios, were not arrived at through a scientific process, and it is unknown whether the true catches (i.e. actual removals) lie within the range of scenarios considered
3. CPUE is still being used as the primary index of abundance in the model and there are several problems associated with this,

4. including the problem of linking CPUE in different time-periods together, or alternatively, dealing with CPUE indices for several unconnected shorter time series (see below)

The scenarios previously considered are still the only ones available to the SAG, with no information about whether these scenarios really do span the range of plausibility. They may, but the SAG/ and SC have no way of knowing or judging this. In our view, the most serious problem for the scientific credibility of work based on these scenarios lies in the fact that the scenarios have not been arrived at through a scientific process. Even given the re-analyses and re-standardisation of CPUE, the series is still based on the same historic data. Again, these data (or parts of the data) may be reliable, but it remains unknown whether these data have been affected by unreported catches or not. (Admittedly, this is assumed to be taken care of through consideration of the various scenarios, but since these are based on the catch scenarios, the points raised above also apply to the CPUE scenarios).

Historically, there have always been concerns about the reliability of longline CPUE as an index of abundance. These concerns include the general (i.e. not just relevant to SBT) issues of the relationship between CPUE and abundance (is it linear or a power function?) and issues of changes in 'q' (catchability) that may not be detected by standardisation. In addition, there is the potential implications of the spatial contraction of the fleet over time. In the past, we have 'lived with' these issues, and in the MP testing process we tried to take some account of them by including a range of assumptions (about the CPUE-abundance relationship) and several series (of CPUE indices). Arguably more important, however, are concerns about the continuity between three (possibly four) parts of the full CPUE time-series:
a) the period prior to unreported catches being identified as an issue (in the recent work of the CPUE WG this has been taken as pre-1989)
b) the period during which unreported catches may be an issue (from 1989 – at least 2005)
c) the most recent period following changes in the domestic management of the Japanese longline fleet (ITQ system), as well as, the reduction in total catch quota (which may even mean that this component should be treated as two rather than one period in the time series)

Continuity between the first two components (a) and (b) may not be a major issue, but still has to be dealt with in a somewhat ad-hoc way. Bridging the gap between the last two components could be much trickier and could have a strong effect on the short term predictions from scenario modelling. The issue of consistency in the recent CPUE series was specifically identified in paragraph 172 of CCSBT-SAG7 report, 2006.

We emphasise that we are aware of the need to be pragmatic under these circumstances (of uncertain historic data) and that it would not be sensible to simply 'throw out' all the historic data. We also note that some quantities and results from the scenario modelling may still be of some use. For example, estimated selectivity curves may be usable (provided assumptions about the size distribution of the unreported catches are reasonable); patterns of relative recruitment may not be too unreasonable and estimates of relative depletion may be reasonable, at least for each individual scenarios considered[1]. Nonetheless, using scenario modelling which contains the flaws and concerns outlined above as a basis for advice –

[1] Serious difficulties arise if attempts are made to combine results because there is no basis for choosing weights for the different scenarios, and because there is little knowledge about the representativeness of the various scenarios.

particularly as the only basis for advice - on future catch levels is, in our view, both risky and difficult to justify on a scientific basis.

We therefore suggest an alternative approach, namely using indicators in a pre-agreed decision rule. Such an approach can either be used alone or, with appropriate caution, in combination with scenario modelling to quantify the relative risk. The suggestions made in this document are meant as a starting point for discussion during this year's (2008) Management Procedure Workshop and Scientific Committee meetings – the appropriate forum in which to develop the detailed specification of such a decision rule. In many ways, the approach we are suggesting is really just formalising much of the underlying decision-making process that has been used by the SAG and SC over several years in the recent past to provide advice to the Commission.

Being aware of the need for advice in the short term, as well as a need to develop a reliable and (hopefully) rigorous and robust approach in the medium and longer term, we discuss alternatives under two sections:

Section 2. Short term (1- 3 or 5 years for example)
Section 3. Medium and Longer term (4+ or 6+ years for example)

2. Indicators as a basis for advice in the short term

An important part of the SAG/SC 'assessment' of SBT stock status has, in the recent past and at least since 2002, included a review of indicators. Management advice from the SC has been strongly informed by the review of indicators (see 2002-2007 Reports of the SC). In addition, the meta-rules and procedure for dealing with exceptional circumstances developed by the SAG and SC as part of the full management procedure, are also largely based on a review of indicators (e.g. Report of MPWS4, Attachment 8). Since the revelation of unreported catches, a subset of indicators considered to be unaffected, or possibly only slightly affected, by the unreported catches have been identified (Report of SC11, Table 1; reproduced as Table 1 below). In our view, these can – and should – form the basis for advice in the short term. Admittedly, the indicators do not provide absolute indices of fishable or spawning biomass, but this is not essential. Besides, the scenario modelling provides such a wide range of absolute biomass levels and the SAG/SC has historically cautioned against over-interpreting these estimates, and considered the relative depletion levels in preference.

In addition to annual 'point estimates', an indicator should ideally have some estimate of variability (e.g. a CV for each point estimate), to establish whether the indicator has changed or not. Without this, an indicator can still be used, but it should be recognised that a decision rule may respond primarily to 'noise' rather than a 'signal' in the indicators. It is, therefore, worth bearing this in mind when selecting a set of indicators to use in a decision rule.

2.1 Suggested indicators

We suggest that a decision rule be constructed, based on an agreed list of indicators which are either fishery independent, or unlikely to have been affected by the historic unreported catches. A starting point for such a list of indicators, and for discussion at this meeting, can be identified from the section on 'Synthesis of indicators' in the Report of SAG8 (2007), and from Table 1 in the Report of SC11 (2004). It is informative to retain the grouping into three categories of the population:

Indicators of recruitment
Size distribution in the New Zealand longline fishery (charter and domestic)
Aerial survey in Great Australian Bight
Estimates of F for age 3 and 4 from conventional tagging data

Indicators of spawning stock biomass
size/age distribution of Indonesian catch

Exploitable biomass in the longline fishery
none

The above list does not include all the indicators listed in Table A1 as 'unaffected'. In a simulation study of the use of multiple indicators in decision rules, Basson and Dowling (2008) found that there was little to be gained in performance by using more than 4 or 5 indicators together, and often 2 to 3 indicators were sufficient. It is arguably more important to have some indicators in each of the population categories identified above rather than have a large number in any one category.

It is immediately obvious that most of the indicators reflect recruitment, and there are none which reflect the fishable biomass in the longline fishery. This situation could be reasonable for a short-lived or early maturing species, but for SBT, which matures at around 10-12 years, this is of concern. In the very short term, the recruitment indices can still be used to inform management advice, but ultimately it will be crucial to develop indicators[2] for the older part of the stock. We discuss this further below.

The recruitment indices are all potentially affected by substantial changes in the movement dynamics of SBT if this implies that a very different proportion of the total juvenile population visits the GAB and/or New Zealand waters each year. This has always been known and noted as a caveat to the interpretation of these indices. Nonetheless, under the notion of management using a precautionary approach it implies that signals, particularly negative signals, in these indices need to be seriously considered rather than ignored if there is uncertainty about their interpretation.

Regarding the New Zealand longline size frequency data, it is not obvious how a full size frequency distribution would be used as an indicator in a decision rule. It is, however, easy to see how one might be able to construct a simple index based on the size frequency. For example, the proportion of fish below a certain size, and/or the mean size of fish in the size frequency could be used. The 'tails' of the distributions are more sensitive to sampling noise and to natural variability than quantities such as the mean or median (Basson and Dowling, 2008). Also, a change in the proportion of small fish can be due to an a change in the number of small fish, or in the number of large fish. Given the concerns about

[2] By 'develop indicators' we include the notion of using an assessment model to integrate different data components and estimate quantities, such as age-specific harvest rates. The important point is that the relevant data (i.e. fishery independent data and data from all fishing fleets) for such a stock assessment need to be collected, and through a process which ensures acceptable levels of confidence and consistency.

absence/low numbers of small fish in the New Zealand fishery in the recent past, however, an index based on the proportion small fish may well be entirely appropriate, in spite of its potential weakness.

The commercial spotting index (SAPUE) has not been included in the list above, since it is based on data gathered during commercial fishing operations and standardisation of this series therefore suffers from many of the problems associated with commercial CPUE data. In contrast, the Aerial survey index is based on a scientific line transect design which is used each year together with a consistent set of protocols. This provides underlying concistency in the way in which the data are collected. The coefficients of variation (CVs) for the aerial survey index are therefore also more likely to be reliable than those of the SAPUE index.

Although tagging data can, under some circumstances, provide estimates of abundance, the conventional tagging data from the SRP have been used to estimate harvest rates in the surface fishery. This is a useful additional piece of information on this component of the stock, since in some sense it reflects ‘escapement’ into the older (5+) age classes which form part of the longline catch. The tagging data can, however, also be used to monitor changes in movement patterns which may alert us to possible changes in the interpretation of other indices. For example, the observed changes in the patterns of returns from fish tagged as 1-yr olds in western Australia versus ones tagged as 1-yr olds in southern Australia (Polacheck and Eveson, 2007) is of some concern, and continued monitoring of this would be prudent. We note that tag return data could potentially also shed some light on the absence/reduction of small fish in New Zealand size frequency data through the information on movement patterns.

Regarding spawning biomass and the size frequency distribution of the Indonesian catch, the comments made with respect to the New Zealand fishery size frequency distribution are again relevant. The Indonesian catch (i.e. its magnitude) has been used as an indicator in the past because fishing took place on the spawning ground and there was effectively no quota limitation. Now that Indonesia has become a member of the CCSBT, this quantity will no longer be suitable as an indicator. The fishing patterns of the Indonesian fleet have also changed (see Basson et al., 2007) with some fishing taking place off the spawning grounds. This change may also be of relevance to the interpretation of size frequency.

Although medium to long term considerations are discussed below, we note that work is underway to develop an estimator of absolute spawning biomass based on the close-kin genetics approach (Bravington and Grewe, 2008 (CCSBT-ESC/0809/29)). If this approach proves to be successful, a time series of estimates can be developed. This would be a very valuable time-series which could be used directly in a decision rule in future. In light of this initiative, we do not currently consider it useful or appropriate to pursue survey-based approaches for the spawning component of the stock.

Regarding the exploitable biomass in the longline fishery, in our view tagging is still one of the most promising approaches in our view. The difficulties encountered with tag returns from longline vessels during the SRP tagging program have been discussed in the past (see e.g. Basson et al., 2007; Report of SC12, 2007), and potential solutions to this will be considered at this meeting (see e.g. Davies et al. 2008 (CCSBT-ESC/0809/13) and Harley et al. 2008 (CCSBT-ESC/0809/14)). The sooner an index that reflects this part of the population can be added to the suite of indices the better.

No CPUE-based quantities have been included in our suggested list. This is deliberate and consistent with the comments above. However, we would not rule out the potential inclusion of a CPUE index if the catch and effort data are collected with high observer coverage, and if the index can be considered to be informative in the context it is being used.

2.2 The decision rule

It is important to recall that all the scenarios considered at SAG7 (2006) indicate a stock that is still at a historic low relative abundance: between about 8% and 17% of unexploited, on average (Table 7 of the SAG7 report). Many other indicators also suggest that it is unlikely that the stock has increased substantially following the high harvest rates in the 1980s (for example, the aerial survey suggests recent recruitment below that in the 1990s; the Indonesian catch age frequency shows much smaller proportions of old, large fish; see Indicator analyses in the recent SC reports). It is against this background, and with the Commissions stated objective of rebuilding of the stock to 1980 levels (Report of CCSBT 13) as a key objective for management that we consider a decision rule. The intention is not for this type of decision rule to be used for a long period, but rather to act as an interim rule.

We suggest, as a starting point for discussion, the following (first given in summary, followed by the reasoning behind the suggestion):

Catch in year (t+1) = Catch in year (t), (where Catch is assumed to be equal to the TAC)

UNLESS:

i) one or more indicator shows a decline[3] or has a value below a pre-agreed level (‘reference point’) in which case the recommended catch should be decreased, OR;

ii) if all indicators have values above some pre-agreed level (or ‘reference point’), or show increases then an increase in catch can be considered; though unless a rebuilding target has been reached, this will jeopardise the speed of rebuilding and imply a higher risk to the stock

The suggestion above obviously requires several detailed definitions:

i) how often the decision rule would be applied (we have written t+1, implying annual application, but the catch could be set for, say 3 years - as agreed for the simulation tested MP - and only adjusted according to the rule every 3 years);

ii) if trends in indicators are used, the number of years over which to estimate the slope, and criteria for deciding whether the slope is ‘significant’ need to be defined;

iv) if pre-agreed reference values are used, these need to be defined;

v) the way in which an indicator (with or without associated CV) is compared to the reference level should be defined, e.g. the probability level that would signify being ‘above’ or ‘below’ the reference level;

[3] For some indicators, e.g. estimate of harvest rate (F), this relationship should obviously be reversed, so that for example, the catch is decreased if the harvest rate is increasing, or is above some predefined level.

iii) the extent of the decrease in the catch as a function of the indicator values or trends, and;

iv) the extent of any increase in the catch as a function of the indicator values or trends.

The third and fourth points are deliberately listed separately because the decision rule need not be symmetrical. The advantages of an asymmetric rule are discussed below.

<u>'Current catch' as a starting point</u>

The decision rule has been written in terms of Catch (rather than TAC) in recognition of the fact that it is the Commission which agrees and sets a TAC. In this sense, the implicit assumption is that the catch will be equal to, or very close to, the agreed TAC.

The suggested decision rule implicitly assumes that the starting point (i.e. at first implementation) is the current catch, or more specifically the current TAC. Justification for basing the decision rule on the current TAC lies in the following reasoning. Absolute abundance (biomass) could be greater than previously thought (e.g. Table 7 and Attachment 6, SAG7 report). If under-reporting has ceased, then the current catch will be lower than before. The probability of the current catch being too large and leading to further declines should therefore be lower than before. This statement is subject to recruitment being average and not unusually low. On the other hand, if absolute biomass is not greater than previously thought (i.e. similar to; Table 7 and Attachment 6, SAG7 report), but current catch is lower than before (again, if under-reporting has ceased), then the probability of further population declines should be lower than before. This is again subject to recruitment being average and not abnormally low.

For both above cases we do note, however, that if recruitment has been as low relative to historic levels as shown by indicators (and estimated in the operating model) for 1999-2001 there will inevitably be a decrease (or 'dip') in the SSB time-series approximately 10-12 years hence, i.e. around 2011-2013 (see e.g. SSB projections in Attachment 6 of SAG 7 Report). This situation alone suggests that any near future increases in catches would increase the risk of low relative spawning biomass for the cohorts born in those years of low recruitment.

We are, of course, aware that the justification given here is based on results from the scenario modelling which we criticise above. But is important to note HOW these results are being used here; they are being used in a qualitative sense rather than a quantitative sense.

<u>Adjusting the current catch</u>

The notion of monitoring indicators for trends or for particularly low, or high, values and adjusting catches in response is commonplace and requires little explanation. It is, however, worth considering two aspects: how to use indicators and how to adjust the catch.

Regarding the use of reference levels or trends in indicators, simulation studies and theoretical considerations have shown that a decision rule which simply responds to increases or decreases in an indicator is very unlikely to lead to rebuilding of a stock if the stock is already over-depleted (Basson and Dowling, 2008). In the case of SBT, use of historic (or biological) reference levels may be more successful and appropriate than trends in indicators. (The word ‘historic’ is used in the sense of defining an empirical reference level based on the long-term time series and broad knowledge of the stock and fishery history.) It is obvious that an asymmetric decision rule which only allows decreases, but not increases in catch (or much smaller increases than decreases) will have a higher probability of rebuilding than a decision rule which is symmetrical in terms of increases and decreases. If trends in indicators are used, but with an asymmetric decision rule, there should be some probability of rebuilding. A combination of reference levels and trends can, of course, be used; this is discussed further below. We repeat that the intention is not to use this rule in the long term and in any case, once rebuilding targets have been achieved there should no longer be a need for a strongly asymmetric decision rule.

Discussions about the acceptable magnitude of change in the catch, when a change is required, took place during the development of the simulation tested MP. The outcomes and agreements from these discussions can be used directly in this context to define the adjustments to make to current catch if the decision rule indicates the need for a change.

Reference levels

The task of defining the details in a decision rule, along the lines suggested above, will inevitably have to be approached with pragmatism. Given that the decision rule is not meant to be used for a long time period, a common sense approach to this task should be acceptable. In this section we suggest some ways in which reference levels for indicators could be defined. It is, however, important to maintain a view of the whole decision rule; it will ultimately be several components and characteristics of the rule working together to provide a framework for maintaining or adjusting the catch level.

Take the aerial survey as an example. In light of our understanding of relative spawning biomass level (which is still low) and the very low recruitments (between 1999-2003) shown by several indicators (e.g. the SAPUE indicator, several catch size frequency distributions) and estimated in the operating and scenario modelling, it would be prudent to define a lower reference level which would trigger action – a reduction in catch – if the aerial survey index falls below that level. The definition of whether ‘the index is below the reference level’ needs to be clear. For example, is it the point estimate that is to be compared with the reference level, or should the confidence interval around the point estimate be considered, in the sense of “if there is an x% probability that the index is below the reference level, then...”. This second approach of agreeing on a probability that the indicator is below the reference level is, in our view, the more appropriate one. Regarding a reference level for use in the short term, one suggestion would be to use the lower quantile of the series (i.e. the 25th percentile), a value of 0.55 (the series has been scaled so that the mean is 1). An associated probability of the index being below that value should ideally be relatively small (say less than 10%, for example). A more informed probability can be arrived at by considering the series and its precision in a bit more detail, as well as the SAPUE index values in years representing the very low recruitments (primarily 2003 and 2004 when those cohorts were 2-4 years old). Recall that the aerial survey was not conducted between 2001 and 2004. Figure 1 shows the aerial survey index point estimates with the quantiles of the series. The standardised commercial spotting index, rescaled to have the same mean over the years of overlap (2005-2008) is also plotted. For technical reasons confidence intervals are not shown here; see figures and tables in Eveson et al. (2008, CCSBT-ESC/0809/24) and in Farley and Basson (2008, CCSBT-ESC/0809/) for this information.

It is not possible to comment on the absolute level of risk to the stock implied by such a reference level and in the long term, such a level may be too risky. It can, for example, be argued that recent recruitments, since about 1999, have been substantially below levels in the 1990s which in turn are substantially below average recruitment levels in the 1960s and 1970s[4] (see e.g. CCSBT-ESC/0609/25, CCSBT-ESC/0409/23). In the short term, however, we consider that the logic set out above can be justified as being unlikely to have unacceptably high risk. It is also worth noting that it is currently particularly difficult to estimate any absolute risk given the historic low relative abundance (i.e. uncertainty about stock-recruit dynamics at such a low relative spawning biomass) and given the uncertainties about the historic input data.

From the point of view of rebuilding the SBT stock, defining upper reference levels (above which increases in the catch can be considered) are arguably less critical, particularly in the short term. Such reference levels can, however, be useful as a kind of 'performance' indicator and they are understandably of interest to the fishing industry. In the context of the aerial survey, it would not be unreasonable to suggest the upper quantile (at a value of 1.34) as an upper reference level, but here it may be prudent to also require that the indicator has a positive slope over the previous several (say, 4 or 5) years before an increase in the catch is considered, to provide additional confidence it represents a real increasing trend in recruitment.

Similar considerations can be used for other indices, though some will be more difficult to specify than others. For example, in the case of tag-based estimates of age-specific (and year-specific) harvest rates (F) it is not obvious how best to distil and simplify the information without losing important signals. The first step would be to decide on a set of assumptions to use (i.e. mortality vector and reporting rate assumptions[5]), and the second step would be to decide how to condense the estimates into a more manageable index. For example, it would be convenient not to have to separately deal with 'age 4 mortality in 2006 for individuals tagged at age 2', and 'age 4 mortality in 2006 for individuals tagged at age 3'. It may be reasonable to combine estimates for animals tagged at ages 2 and above, but animals tagged at age 1 should probably not be included in such a combined index, given the recent (unusual[6]) results for this age group (Polacheck and Eveson, 2007; Eveson and Polacheck, 2008). Careful consideration is required if setting reference levels for these estimates given that they may (it is unsure whether they do or do not) reflect the harvest rate on the whole juvenile population, or on just a part of the juvenile population which visits the GAB in summer. A reference point based on an absolute level of F obtained from one of the common approaches (e.g. F0.1, Fspr35% - based on relative spawner per recruit considerations) which assume that the harvest rate applies to the total stock could be unrealistic. Consideration of trends in current F instead of comparison with a reference point, and/or empirically-determined reference point (e.g. based on historic estimates from tagging in the 1990s and broader consideration of historic stock status and overall exploitation) may be more realistic and appropriate.

[4] Although model-based estimates of absolute recruitment are affected by the unreported historic catches, all the scenario modelling results suggest some consistency in relative recruitment patterns over time.

[5] This choice will matter more or less depending on how the F estimates are used.

[6] F estimates for a given age and year are rather different when based on those tagged at age 1 compared to those tagged at age 2 or older.

2.3 Exceptional circumstances

It would be sensible and prudent to again include a process for dealing with 'exceptional circumstances' though the details of such a process would depend on the design of the decision rule, in other words, on which indicators are included directly in the decision rule and which are not. The 'exceptional circumstances' should essentially relate to indicators that are not included in the decision rule. It should, possibly, only relate to indicators in the decision rule if a value is well outside the historically observed range of values.

2.4 A role for 'scenario modelling'?

We are suggesting that a decision rule, developed along the lines illustrated above, be used as the basis for management advice on total catch in the short term. We are, however, aware that there may be advantages in also being able to evaluate the implications of the current catch level, or alternatives derived from a decision rule, in the scenario modelling framework. An important point would be to decide beforehand how such a joint approach would work and what role each of the two approaches would play. It is in this context that the range of scenarios and their relative weightings (if combined) become important. The scenario modelling conducted in 2006 (SAG7 report) were presented as separate scenarios and not combined for this very reason. It would, for example, be appropriate to do forward projections in the modelling framework using the decision rule derived catch to quantify risks for the range of scenarios (e.g. the previous set of scenarios used in SAG7, or a new set which may be defined at this year's meetings (i.e. 2008)).

The suggestion was made at SAG7 that indicators (particularly the aerial survey index) and/or recent tagging data could be included in the operating model used for scenario modelling. We do not consider the inclusion of indicators in the scenario modelling as an alternative to the approach suggested here. The main reason is that, in our experience, past assessment analyses have shown that the catch and CPUE data dominate assessment results. In our view, it is unlikely that this would change unless the model can be made to run without the CPUE series, or with only those sections of the CPUE time-series that can be considered unaffected by the unreported catches. Even in this case, the catch series will have a strong effect. This approach would therefore still suffer from the problems (catch and CPUE data issues) that motivated this paper. Having said that, if the indicators and recent tagging data are included in the operating model, some limited simulation testing of the suggested indicator-based decision rule could arguably be attempted. The word 'limited' is intended to reflect both the fact that testing would not be extensive, as well as, the fact that the results from the testing would be limited by the shortcomings of the existing scenario modelling already mentioned. Our suggestion to use an indicator-based decision rule as a basis for advice in the short term is in fact made without the notion that its performance should first be formally tested. The reason being that we see little advantage in using the existing model to test such a decision rule, and testing using an alternative model (not formally conditioned to historic catch and CPUE data) is not feasible in the short term. Furthermore, the proposed short term use of the indicator-based decision rule reduces the need for testing, particularly if the rule is explicitly constructed to be precautionary.

2.5 Data requirements for a decision rule

In this section we only comment on particular or additional data requirements for a decision rule; all existing data requirements for assessment purposes remain (e.g. Anon 2007; Davies et al. 2007). First, there will be a continued need to ensure that reliable catch data are collected. The decision rule is predicated on the assumption that the actual removals are equal to (or at least very close to) the TAC. Second, there will be a need to ensure that those indicators included in the decision rule will be continued (at least in the short term) and the relevant data collected and analyses conducted. In this regard we note that the SRP tagging program has been temporarily suspended, in the hope that the problems of reporting rates in the longline fisheries can be resolved. The potential to use estimates of juvenile fishing mortality in a decision rule means that the possibility of restarting tagging should be considered. It would be ideal if this could be done in such a way that returns and reliable reporting rates can be obtained from the longline fishery. However, even if the tag-return problems associated with the longline fleets cannot be resolved this year, serious considerations should still be given to restarting some form of tagging program. We consider this to be particularly relevant in the light of the apparent changes in movement dynamics of 1-year old fish off the west and south coast of Australia, and noting that undetected substantial changes in spatial dynamics of 2-4 year olds could have an impact on our interpretation of the other indices of recruitment.

3. A basis for advice in the Medium and Longer term

In this section we consider ways of developing a management procedure, with associated testing framework, based on fishery independent data. Although data that are entirely fishery independent may be hard to obtain, the aim is to consider approaches that are less fishery dependent than longline CPUE which was used in the first MP development process.

A well-designed and reliable tagging program which can provide estimates of harvest rates, and possibly estimates of abundance and natural mortality, is still one of the most promising monitoring and management tools for large pelagics. There is no equivalent of ‘groundfish surveys’, or trawl surveys for large pelagics, and tagging is one of the few relatively fishery-independent approaches. We do not discount the aerial survey, but note that it only surveys juveniles (2-4 year olds) in one of its summer habitats, the Great Australian Bight (GAB). Tagging could potentially provide information on the full range of fished ages, and if so, this information could be used in a management procedure.

There may, of course, be problems in terms of the logistics and in terms of underlying assumptions (in the methods which are then used to estimate relevant quantities). For example, problems with estimating reliable reporting rates and violation of the assumption of full mixing spring to mind. In the context of management strategy evaluation (MSE), however, the impacts of these issues on an MP can be evaluated with a view to designing a robust system. This is discussed further below, but first we suggest how such an approach could form the basis for advice even in the medium term.

Assuming that relatively reliable estimates of annual harvest rates (Fs) can be obtained from all fisheries, and hence for a broad age range, it should be possible to construct a decision rule which uses these harvest rates instead of CPUE. Such a decision rule could be a relatively simple ‘indicator-based’ type of rule which responds to changes in Fs. This can be done on the basis of overall Fs or on the basis of age-specific Fs. An alternative would be build the decision rule into a simple model/estimation framework (in concept somewhat like the production model used with the CPUE-based MP, but obviously very different in detail). We emphasise that, at this stage, the 'sketch' given here is mostly conceptual, and there is much detail which would require more thought and attention.

One potential advantage of such a harvest-rate based approach is that reference points can relatively easily be defined on the basis of spawner per recruit calculations and used in the construction of the decision rule. These calculations only require estimates of mortality-at-age, maturity-at-age (and size at age if done in terms of biomass), and selectivity curves for each fishery. The first three quantities are all assumptions within the models (scenario and previously assessment and operating model). The selectivity curves estimated in such a model (even given the uncertainty in catches) should be adequate for this purpose. Obviously, uncertainties in these quantities can be explored when calculating candidate reference points.

Other fishery independent indicators can, of course, be incorporated into a decision rule, or used to modify an initial 'recommended catch' if its value is particularly low or particularly high (or if it has a downward or upward trend).

While estimates of harvest rate provide information on whether a stock is being overfished or not, estimates of abundance are required to evaluate whether a stock is overfished or not. Of particular relevance in this regard is the spawning stock, and whether rebuilding of the spawning biomass is taking place. The main fisheries are now taking primarily immature SBT, so that even if abundance estimates could be obtained from tagging, they would primarily relate to the fishable stock. The hope is that the close-kin approach to estimating spawning biomass (Bravington and Grewe, 2008, i.e. CCSBT-ESC/0809/29) will be successful and that it can, in future, also be used in an MP together with tag-based harvest rates.

An important part of the development of a tag-based MP would be the simulation testing of the performance of such an MP. In our view this would require a spatial model since one important aspect would be the interaction between the design of the tagging program, the stock (and fishery) movement dynamics and the estimated harvest rates. We also envisage taking a rather different approach to conditioning. In particular, we intend to reduce the reliance on historic catch and CPUE data in any conditioning, while still ensuring that the broad features of the stock history are adequately reflected. This is again currently only a conceptual 'sketch', but we note, for example, that estimates of relative depletion from the scenario modelling cover a narrower range than absolute estimates of biomass; it may be possible to use these types of characteristics in conditioning.

There are, as always, caveats and potential problems associated with tagging. Fortunately, some issues can be explored in the context of the MSE process. The interaction between the design of the tagging program, the stock (and fishery) movement dynamics and the estimated harvest rates has already been mentioned. The relevance of this has recently been highlighted by the observation of different fishing mortality estimates from fish tagged at age 1 versus those tagged at older ages and relative to the same age classes in the 1990 tagging program (again, see Eveson and Polacheck, 2008, i.e. CCSBT-ESC/0809/22); the reason for this is still unclear, but it could be associated with changes in movement dynamics and mixing. Fortunately, there is a growing dataset of archival and other electronic tags, and a large project (see Polacheck et al. 2008, i.e. CCSBT-ESC/0809/23) to analyse and model the movement dynamics of juvenile SBT. The data analyses and modelling work which has already started, but which will increase substantially over the next two years, is likely to play a very important role in constructing plausible scenarios of spatial dynamics for the MSE work and for designing a tagging program that is robust to alternative movement hypotheses. We also note that a significant information gaps on the spatial dynamics of the population is similar information for the spawning (12-20+) component of the population, although there have been initial and ongoing attempts to address this (Evans et al 2007).

Logistic issues, such as those associated with the return of recaptured tags and the estimation of reporting rates from all the fishery components obviously need to be overcome. Some simulation testing can be conducted to explore the effects of, for example, poorly estimated reporting rates and to establish (or review) acceptable levels of precision required. But, these issues also need to be overcome in practice if such a tag-base MP system is to work in reality. In our view, the reporting rate issues previously raised should not be insurmountable. The discussions about whether how they can be overcome, e.g. with genetic tagging or PIT tagging, or some other novel tag-recovery technique, will form part of this year's SC meeting. (see Harley et al. 2008 and Davies et al. 2008).

For completeness we note that we implicitly assume, and strongly support, the continued collection of data for stock assessment purposes, including reliable catch and effort data, size frequency data, direct ageing and basic biological information. Although reliable CPUE information can still play a relevant role as a 'local' (where the fishery takes place) indicator of density, we firmly believe that the time has come to find alternative ways to monitor and manage pelagic stocks.

This paper is a first (small) step in this direction. Our intention is to develop a more detailed framework and outline for the proposed approach and for the work required. Although aspects of this work will take several years (3-5 years), we note that a start can be made to using tag-based estimates from all fishing components in an indicator-based decision rule as soon as the practical tagging problems can be resolved.

Acknowledgements

The ideas presented here have benefited from discussions with colleagues in the Pelagic Fisheries and Ecosystems group but the views expressed are those of the authors. This paper is a contribution from the CSIRO Wealth from Oceans Flagship and was partly supported by funding from the Australian Fisheries Management Authority and the Department of Agriculture, Fisheries and Forestry.

References

Anon. 2007. CCSBT Report of the Twelfth Meeting of the Scientific Committee, 12-18 September 2007. Hobart, Australia.

Basson, M. and N.A. Dowling. 2008. Development of a robust suite of stock status indicators for the Southern and Western and the Eastern Tuna and Billfish fisheries. Final Report to the FRDC, Project 2003/042. 350p.

Basson, M., Andamari, R., Proctor, C., and Sadiyah, L. 2007. An update on the use of the Indonesian Fishery school dataset to obtain a standardised CPUE series for SBT on the spawning grounds. CCSBT-ESC/0709/15.

Bravington, M., and Grewe, P. 2008. Update on the close-kin genetics project for estimating the absolute spawning stock size of SBT. CCSBT-ESC/0809/29.

Davies, C., Preece, A., Basson, M., 2007. A review of the Commission's Scientific Research Program, and considerations of current priorities and ways forward. CCSBT-ESC/0709/16.

Davies, C., Moore, A., Grewe, P., and Bradford, R. 2008. Report on the potential and feasibility of genetic tagging of SBT. CCSBT-ESC/0809/13.

Eveson, P., and Polacheck, T. 2008. Updated analyses of tag return data from the CCSBT SRP tagging program. CCSBT-ESC/0809/22

Eveson, P., Bravington, M., and Farley, J. 2008. Aerial survey: updated index of abundance and preliminary results from calibration experiment. CCSBT-ESC/0809/24.

Farley, J., and Basson, M. 2008. Commercial spotting in the Australian surface fishery, updated to include the 2007/8 fishing season. CCSBT-ESC/0809/25.

Harley, S., Bradford, R., and Davies, C. 2008. Using passive integrated transponder (PIT) technology to improve performance of CCSBT's conventional tagging program. CCSBT-ESC/0809/14.

Polacheck, T., Chang, K.S., Hobday, A., West, G., Eveson, P., and Chung, K.N. 2008. Update on the Global Spatial Dynamics archival tagging project - 2008. CCSBT-ESC/0809/23.

Polacheck, T. and Eveson, P. 2007. Updated analyses of tag return data from the CCSBT SRP tagging program. CCSBT-ESC/0709/19.

Table 1. A copy of Table 1 from the Report of the 11th meeting of the Scientific Committee, 2004, (page 6,7).

Table 1: Potential influence of catch anomalies (affected, potentially affected, or unaffected) and funding limitations for the SBT fisheries indicators.

Indicator	Influence of catch anomalies	Information to determine extent of effect	Security of future provision of data
CPUE trends in Japanese LL fishery	Affected	Independent verification using fine-scale data. Further analysis of observer data*.	Ongoing, but expect coverage and continuity to change
CPUE by year/age class in Japanese LL fishery	CPUE affected, proportions by age potentially affected	Independent verification of fine-scale data. Further analysis of observer data*.	Ongoing, but expect coverage and continuity to change
Length frequency in Japanese LL Fishery	Potentially Affected	Independent verification of fine-scale data. Further analysis of observer data*.	Ongoing, but expect coverage and continuity to change
CPUE and length frequency for New Zealand LL charter fishery	Unaffected		Ongoing
CPUE and length frequency for New Zealand LL domestic fishery	Unaffected		Ongoing
Indonesian catch, age composition, and CPUE	Unaffected		No funding beyond May 2007+
Estimates of past total SBT catch	Affected	Resolve uncertainties in farm and market anomaly reviews	Will require increased resources for a comprehensive catch verification system
Acoustic index	Unaffected		Funding ceased in 2006
Troll survey	Unaffected		At feasibility stage
Fishery independent aerial survey	Unaffected		Funding secure for 06/07 season only
Commercial spotting index	Unaffected		Ongoing
Conventional tagging	Potentially affected	Reporting rates from LL fisheries	Funding ceased 2006, future funding awaits Commission decision
Growth Rates	Unaffected		Partially dependent on tagging program

* Availability of detailed data may be limited for reasons of confidentiality

\+ Indonesian fishery indicators may be affected by shifts in catch locations and targeting

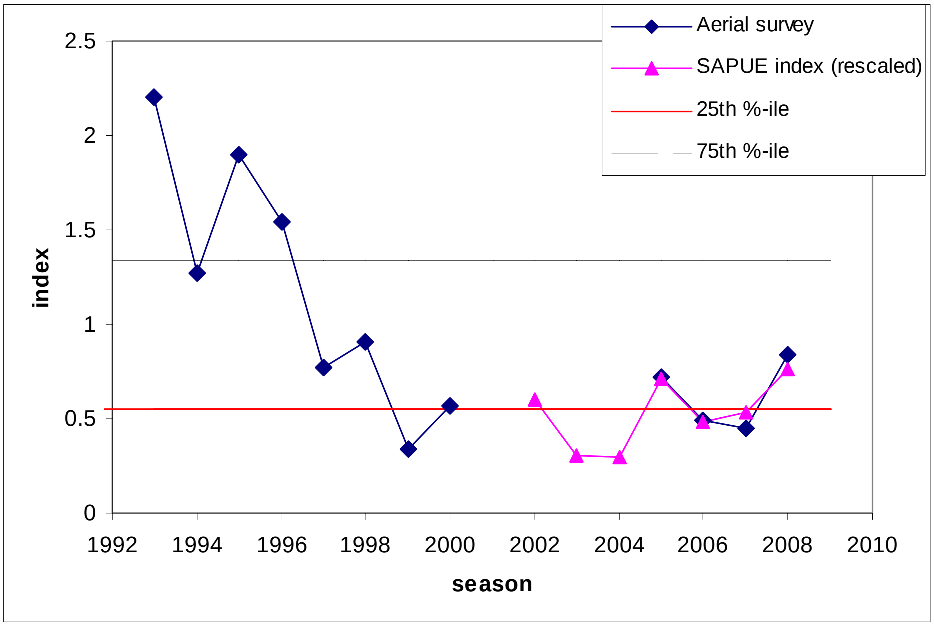

Figure 1. Point estimates of the aerial survey index (up to 2008) with horizontal lines at the 25th and 75th percentiles. The standardised commercial spotting index, rescaled to have the same mean as the aerial survey for the years of overlap (2005-2008) is also shown for information.