Fisheries Research 94 (2008) 339–350

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

Salvaged pearls: lessons learned from a floundering attempt to develop a management procedure for Southern Bluefin Tuna

D. Kolody*, T. Polacheck, M. Basson, C. Davies

CSIRO Marine and Atmospheric Research, P.O. Box 1538, Hobart, Tasmania 7001, Australia

ARTICLE INFO

Article history:
Received 30 November 2007
Received in revised form 21 August 2008
Accepted 22 August 2008

Keywords:
Harvest strategy
Management procedure
Southern Bluefin Tuna
Fisheries management
Management strategy evaluation

ABSTRACT

During 2002–2006, the Commission for the Conservation of Southern Bluefin Tuna (CCSBT) engaged in a multilateral process to develop and simulation test a Management Procedure (MP) for the international SBT (*Thunnus maccoyi*) fishery. The CCSBT Scientific Committee reached consensus in 2005, and recommended an MP to the Commission (including both the pre-specified data collection methods and decision rule for recommending a Total Allowable Catch (TAC)). The MP was adopted, in principle, by the Commission. However, revelations of substantial, long-term under-reported catches undermined confidence that the MP was likely to achieve the intended management objectives. Consequently, it has not been implemented, pending further work to determine the implications of the historical data problems, and the progression of compliance measures to improve future data collection. This is a discouraging outcome of a lengthy and resource intensive process that had been recognized as a promising solution to a difficult management impasse. However, the CCSBT did become aware of the serious catch under-reporting problem, and reached a consensus agreement on the first substantive TAC change since 1989, and the MP process may have contributed to this progress. We outline a range of lessons from the CCSBT MP experience that we would expect to be relevant to other fisheries engaging in a similar process. Foremost among the lessons: formal Management Procedures cannot be expected to resolve all of the hard problems faced by fisheries managers, and agreement on data monitoring, sharing and verification standards should be established before MP development is pursued.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

In 2000, the Commission for the Conservation of Southern Bluefin Tuna (CCSBT) agreed to embark upon a process to develop a Management Procedure for the SBT (*Thunnus maccoyi*) fishery (Anon., 2000), with a target adoption date of 2004. A Management Procedure (MP), is defined as a simulation-tested decision rule (or Harvest Control Rule), and the requisite methods of data collection and analysis, which together are used to calculate a management recommendation (e.g. Total Allowable Catch (TAC)) for a fishery (e.g. de la Mare, 1986; Butterworth et al., 1997; Smith et al., 1999). At the outset, considerable optimism existed within the CCSBT that a jointly developed MP could break the dysfunctional cycle of contested stock assessments and failure to reach consensus on management decisions that had prevailed since the mid-1990s. The development process suffered from a number of setbacks, culminating in revelations of substantial data problems in 2005–2006 that undermined confidence in the agreed MP. As a result, MP implementation was suspended until the implications of the data problems can be formally admitted within the simulation testing process. MP implementation is now expected to begin in 2011 at the earliest (Anon., 2007a).

In this paper, we identify the positive outcomes from the process, and share a number of lessons that are likely to be relevant for other (particularly international) fishery management organizations. While many of the details are specific to the CCSBT, we note that similar general lessons have been described in other fisheries. As one set of contributors among many participants in the CCSBT MP process, we are not the sole originators of the methods and ideas described. We do, however, take responsibility for the opinions expressed herein, and recognize that other CCSBT participants may have different views.

1.1. Background to SBT fishery assessment and management

SBT are long-lived, migratory, high-value fish, found throughout most of the southern temperate oceans, except for the more easterly regions of the South Pacific. Surface and longline commercial fisheries for SBT began in the 1950s, with peak catches occurring around 1960 (Fig. 1). The stock has been very heavily fished

* Corresponding author. Tel.: +61 3 6232 5121; fax: +61 3 6232 5012.
E-mail address: dale.kolody@csiro.au (D. Kolody).

0165-7836/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.fishres.2008.08.016

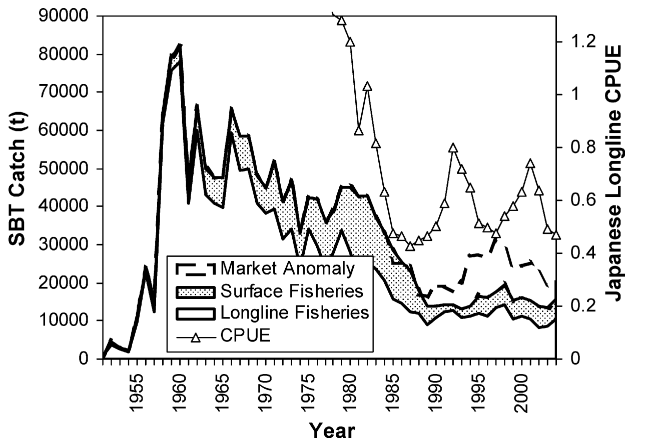

Fig. 1. Catch history of the SBT fishery by gear type. The longline catch under-reporting estimates derived from the Japanese market review have been included as a separate category (extracted from Fig. 2 of attachment 4 of Anon., 2006c). Surface fisheries include purse seine, pole and line, gillnet and handline. The Japanese longline CPUE is the nominal catch rate for SBT aged >4 in months April–September, normalized to mean of unity (observations from 1969 to 1979 are off the scale).

(Caton, 1991) and is currently perceived to be at or near historical low levels (Anon., 2006b). A major component of the surface fishery (off the southeast coast of Australia) collapsed in the late 1970s, and tagging studies demonstrated very high exploitation rates on juveniles in the early 1980s. Informal international management arrangements involving Australia, Japan and New Zealand were initiated in the early 1980s, and were subsequently formalised with the establishment of the CCSBT in 1993 (Caton, 1991; Anon., 1994). Australia introduced catch limits on its fishery in 1984 and international TACs with member allocations were introduced in 1985 under the informal tri-nation arrangement. The TACs were progressively lowered with a major reduction of approximately 50% for the 1989 fishing year. TACs limited catches from the Australian (primarily surface) fishery beginning in 1984 (Caton, 1991). However, it was not until the 1989 fishing year that the catch limits became restrictive for the Japanese longline fleet (i.e. the Japanese longline fishery reported that it was not able to catch its limit prior to this year; Caton, 1991). Formal catch limits essentially remained fixed from 1989 to 2006 (sometimes by agreement and sometimes by voluntary consent in the absence of an agreement).

During the 1990s, managers from Australia, Japan and New Zealand consistently had diverging opinions about the interpretation of the scientific advice, and disagreement about the setting of the TAC (e.g. Anon., 1995a,b, 1996, 1997a,b, 1998, 1999a,b). Initially, the stock assessments tended to predict a high probability of rapid stock rebuilding. However, as the years passed, the projected rebuilding was not evident from the data and different assessments yielded divergent predictions (e.g. Klaer et al., 1996). This led to a number of initiatives to attempt to improve the stock assessment, including the development of a joint experimental fishing program (Polacheck, 2002 and references therein). When agreement on this failed, Japan undertook unilateral experimental fishing, which was viewed by other CCSBT members primarily as a means of increasing catches. This resulted in a legal dispute in the International Tribunal for the Law of the Sea (ITLOS) (Firestone and Polacheck, 2003). The ITLOS hearing resulted in a temporary suspension of Japanese experimental fishing in 2000, however, the Arbitral Panel subsequently ruled that it did not have legal jurisdiction to resolve the case.

During the 1990s, catches of SBT by non-CCSBT parties (principally Taiwan, Korea and Indonesia) increased to substantial levels, adding impetus for the Commission to resolve its problems. In 2000, the CCSBT members negotiated a settlement to the experimental fishing dispute. It also agreed to (i) the appointment of an independent scientific advisory panel, (ii) development of a Scientific Research Program aimed at improving data for stock assessments, and (iii) the development of an MP. The concept of a simulation-tested, management decision rule had been formally introduced into the CCSBT process in 1993 (Sainsbury and Polacheck, 1993) and the CCSBT agreed to hold a Management Strategy Workshop in 1996 (Anon., 1996). However, the first CCSBT Management Strategy Workshop was not held until 1999 (Anon., 2000). Prior to this, an in depth illustration of the applicability of an MP approach to the SBT stock had been completed (e.g. Polacheck et al., 1999). The Commission agreed to develop an MP in 2000 (Anon., 2000).

1.2. Development of the CCSBT management procedure

The MP approach is often promoted in terms of potential advantages relative to an iterative stock assessment and *ad hoc* decision-making process (we use this latter terminology to distinguish the traditional alternative to MP management, but recognize that the term *ad hoc* can imply an avoidance of long-term policy that is not relevant for all fisheries that fall into this category). We consider the most important MP advantages to be:

- The MP is evaluated in relation to attainment of management objectives using long-term projections, in which management decisions are simulated in a feedback loop along with the fishery dynamics.
- An MP can be designed to be reasonably robust to the uncertainties inherent in the system (e.g. current population status, future production dynamics, sampling errors).
- Since the decision rules have to be agreed in advance, industry should have confidence in a stable decision-making process that allows them to make strategic investment decisions, while it reduces the scope for controversial decisions to be contested on the basis of "political" arguments about short-term consequences.

Butterworth (2007) provides a more comprehensive description of advantages and disadvantages expected for MPs.

The first CCSBT workshop of the MP development process was held in 2002, and there was a target for the Commission to adopt the MP in 2004. The process was intended to cover the following steps (e.g. Anon., 2002):

1. Identification of management objectives and quantifiable performance measures.
2. Development of decision rules (candidate MPs to be developed by member scientists).
3. Development of a range of operating models conditioned (fit) to real SBT data and representing the plausible uncertainty in the fishery.
4. Simulation testing of candidate MPs with the operating models.
5. Selection of an MP on the basis of the performance measures calculated from the simulation testing.
6. Implementation of the MP.

These steps were all iteratively revisited, with the exception of the final selection and implementation.

Beginning in 2002, the CCSBT Scientific Committee, and Stock Assessment Group meetings dedicated less time to the traditional annual stock assessment process, and focused on MP development. In addition, MP workshops were held in 2002, 2003, and 2004 and inter-sessional technical working groups were also convened on an *ad hoc* basis. It was hoped that the process would be concluded in

2004, with the CCSBT expected to begin setting the TAC using the MP in the following year. However, two notable problems arose in 2004 (Anon., 2004). First, the observed Catch Per Unit Effort (CPUE) relative abundance index for the most recent year (2003) was outside of the envelope of uncertainty that had been encompassed by the operating models (which were, at that time, conditioned with data only up to 2000). The second problem involved strong evidence for two or more years of recruitment at levels lower than had ever been observed before. An additional year of development was undertaken to try to account for these problems, and a new adoption date was set for 2005.

In 2005, the CCSBT Scientific Committee reached a consensus agreement, and recommended a single MP to the Commission (Anon., 2005a). In addition to the MP evaluation, the 2005 Stock Assessment Group and Scientific Committee meetings involved a review of stock status indicators, with a particular focus on recent recruitment. Given the strong evidence for two or more very weak cohorts, the Scientific Committee recommended that, in addition to adoption of the preferred MP, there should also be substantial reductions in catches in the next 1–2 years, to reduce the short-term risk of further stock declines (Anon., 2005a).

However, at the 2005 Scientific Committee and Commission meetings, concerns were raised about the validity of the fisheries data that were used to underpin the MP evaluation process (Anon., 2005a,b; Australia, 2005; JFA, 2005). The Commission agreed to adopt the recommended MP and the additional quota reductions, conditional on the outcome of a review of the data problems (Anon., 2005b). Australia and Japan initiated two independent reviews into possible under-reporting of catches, and the outcomes were presented to a special meeting of the Commission in June 2006. These review documents remain confidential within the CCSBT. The Commission reported the major findings (Anon., 2006a,d), and the implications of the outcomes for stock status and management advice were considered by the Scientific Committee (Anon., 2006b,c).

The first data review involved comparing official SBT catch statistics with the reported sales in the Japanese tuna markets. This review revealed substantial under-reporting of SBT longline catches since at least the early 1990s (see Fig. 1). There remains uncertainty about the identity of fleets contributing to the catch under-reporting. However, the CCSBT Stock Assessment Group and Scientific Committee adopted the operational assumption that a substantial proportion of the under-reported catches were taken by the Japanese longline fleet (Anon., 2006b,c). The specific characteristics of these catches (season, location, size composition) have important implications for the validity of the evaluation and implementation of the MP as: (i) the operating model used to evaluate the MP was conditioned on the historical catches (numbers and size composition) of all fleets, and standardized CPUE from the Japanese longline fleet, and (ii) all of the decision rules tested by the Commission used future catches (assumed to be perfectly reported) and CPUE from the Japanese fleet as the primary data inputs. As a result, uncertainty about the specific nature of the under-reported catches resulted in the Scientific Committee recommending to the Commission that the MP process be suspended until more comprehensive investigation of the implications of the catch under-reporting could be completed (Anon., 2007a).

The second data review involved an analysis of the sampling procedure used to enumerate the catch and estimate the size composition to provide an estimate of total catch in weight for the Australian purse seine fishery. The review concluded that the estimated number of fish transferred into the tow cages was likely reliable, but with respect to the size composition sampling "*...sampling biases could not be estimated satisfactorily from the available data*" (Anon., 2006b).

We consider it unlikely that the historical under-reported catch, and the potential impacts on Japanese CPUE, will ever be satisfactorily quantified, given the lack of progress toward this objective since the market review (Anon., 2007b). However, there remains the expectation within the CCSBT that effective compliance measures will be established, and new data will be collected, such that an MP could be implemented in the future. The current schedule is to have the first TAC set by an MP in 2011 or 2012 (Anon., 2007a).

2. Lessons from the CCSBT MP process

"The only real mistake is the one from which we learn nothing."
John Powell

Given that the CCSBT MP is not currently being used, and faces an uncertain future, some observers might describe the process as an expensive failure. However, it is important to recognize that the general understanding of the fishery is actually better now than at any time in the last 15–20 years, and this might not have been the case if the MP development process had not been undertaken.

There remains an unresolved question in the use of MPs: How uncertain can one be about the fishery data and system dynamics, and still implement an effective MP? It was initially assumed that the CCSBT would be able to develop an MP. However, given the uncertainties in the data quality that were revealed as a result of the MP development process, the consensus view was changed to admit that implementation was not possible on the original timeline. It seems likely that two decades of catch under-reporting would probably not have been recognized, if it had not been for the sequence of events initiated by the MP development process. The process of operating model conditioning (to simulation test the MP) led to a consensus view about the status and uncertainty of the SBT stock. All of the candidate MPs indicated that serious catch reductions would be required to halt the stock decline and meet even minimal stock rebuilding objectives. Given the history of the issue within the CCSBT, it seems plausible that the risk of imminent TAC reductions may have ultimately motivated Australian industry to formally document their concerns about under-reported longline catches. The resulting market review confirmed that the official catch statistics provided to the CCSBT and Scientific Committee had been grossly inaccurate.

The feasibility of MP implementation might be in doubt for the CCSBT (and other fisheries), and this raises a further question: What is the alternative to an MP that is going to provide more effective management? The CCSBT history indicates that the traditional assessment and management cycle had not worked effectively for many years. During the MP development process, the official CCSBT catch statistics were revealed to be seriously flawed, and the Commission produced the first consensus agreement to substantively change the TAC (~20% reduction) since 1989. But can we actually credit the MP development process with this tangible progress? A number of confounding factors occurred in parallel with the MP process: (i) a panel of independent consultant scientists was formally established within the CCSBT Scientific Committee, and was able to act in a mediating role (or offer a neutral opinion to the Commission if consensus was not achievable among member scientists), (ii) the use of jointly developed (and independently coded) software for operating model conditioning was instrumental in reaching a quantitative consensus view of the stock status, and (iii) multiple years of very low recruitment elevated concerns about the potential for a stock collapse. Some or all of these factors might have led to a similar outcome irrespective of the MP process. However, the MP did facilitate the explicit link between the consensus stock status estimates and the long-term management implications. We cannot be certain to what extent the MP process was responsible for this progress, however, we can conclude that progress was realized during the MP process, and we encourage further pursuit of the MP process within the CCSBT.

Irrespective of the future fate of the CCSBT MP implementation, a number of valuable lessons were learned, and we feel that it is important to share this experience with the broader fisheries community. We identified 10 lessons, which are loosely aggregated under the following four headings:

(I) When should one engage in an MP process?
(II) MPs: a practical interface between fisheries science and management.
(III) Representing uncertainty in the fishery system.
(IV) Selection of a decision rule.

The MP approach requires substantial technical expertise, and the practitioners engaged at the technical level will typically focus most of their efforts on the last two topics in this list. However, the first two topics are at least as important in the successful development and implementation of an MP, and these are relevant for both technical and non-technical participants. Consideration of these topics may help all stakeholders to acquire a realistic appreciation of the limits of what can be achieved with an MP.

3. (I) When should one engage in an MP process?

> "[When choosing] between two evils, I generally like to pick the one I never tried before." Mae West

There are a number of potential advantages that can be identified for engaging in an MP process (e.g. as outlined in the introduction). However, it is important to realize that MPs are not a panacea for resolving all fisheries management problems. The first three lessons relate to foreseeable problems that may be debilitating—irrespective of whether management is based on an MP, or the traditional iterative cycle of stock assessment and *ad hoc* decision making.

3.1. Lesson 1: the management procedure development period should not be seen as an opportunity for delaying inevitable management action

Undoubtedly, many of the participants engaged in the CCSBT MP process had genuine motives broadly consistent with the objectives of effective fisheries management, but others might have recognized that the lengthy MP development process represented a convenient mechanism with which to forestall an inevitable and disruptive management decision. The MP process was undertaken with the expectation of 2–3 years of development, during which all members of the CCSBT could be seen to be cooperatively working toward a common goal, and during which there would likely not be any pressure to change the TAC. Given the perceived status of the SBT fishery prior to the initiation of the MP process (i.e. historically low and reasonably stable stock size through the 1990s, with few scenarios predicting recovery to the Commission rebuilding objective; Anon., 2001), it is hard to imagine that there was any real doubt that substantial quota cuts would be required to achieve stock rebuilding in the timeframe defined by the Commission (i.e. return to 1980 levels by 2020, Anon., 1994). The MP might have provided a mechanism for fine tuning the rebuilding rate and achieving a long-term strategic target with a certain probability. However, on the basis of the consistent advice provided by the Scientific Committee, it can be argued that the CCSBT would have been acting consistently with its stated management objectives if it had taken the decision to cut quotas (at least moderately) in parallel with the MP development process.

Ideally, an MP should be designed and implemented in the early stages of fishery development, or at least long before overfishing and overcapitalization have created a political problem. Unfortunately, such fisheries are increasingly rare, and the MP process may be embraced as a tool for crisis management (see Constable et al., 2000 for an exception). In a crisis situation (in particular) it would be naïve to assume that all MP participants would enter into the MP process with similar expectations and universal goodwill. An MP development process should not be allowed to delay (and exacerbate the magnitude of) inevitable management actions.

3.2. Lesson 2: simulation testing of management procedures is not likely to be effective if the historical data are seriously flawed

The implementation of a fully tested MP is often regarded as a tool that will lead to robust management in situations with poor data quality and/or large uncertainty (i.e. MPs can be explicitly designed to incorporate principles of precautionary management). In this section we focus on the need for historical data in the context of the decision of whether or not to go down the path of an MP development process.

The consensus opinion within the CCSBT Scientific Committee is that the historical catch and effort data are not of sufficient quality to support an MP at this time (Anon., 2006c). A recent review of the CCSBT Scientific Research Program concluded that it had failed to meet its objectives for improving catch and CPUE data quality (Anon., 2007b; Davies et al., 2007). The SBT catch under-reporting controversy has been publicized in terms of the overall magnitude of the problem, and accusations of economic foul-play among CCSBT members. Unfortunately, the reconstructed fishery data series also contain poorly quantified temporal trends in biases (e.g. little information is available to quantify the proportion of fish bypassing the reasonably well-documented sales records from the Japanese tuna auctions). Temporal trends in biases also complicate the interpretation of the Japanese longline CPUE series (Fig. 1), which was used as the primary relative abundance index in the stock assessment modelling (and the MP decision rule). One can speculate about a broad range of alternative plausible CPUE trends depending on the extent to which the logbook CPUE records are believed to be affected by the catch under-reporting, and what is assumed about the trends in the catch under-reporting. Similar uncertainties exist in relation to the size composition of the under-reported catch, and these implications have not been examined to date.

Fisheries researchers are often forced to make recommendations that are based on analyses of poor and unverified data, often collected by people who may have obvious economic incentives for under-reporting catches (and discards). The fisheries scientists are usually acting in good faith with an implicit (or explicit) justification—"*a decision has to be made, so we will do the best that we can to provide advice with what we have got*." This capitulation is often coincident with an optimistic expectation that clever scenario modelling and statistical methods can compensate for poor quality data. However, it is rare to see any convincing demonstration that these methods are effective.

The process of modelling data uncertainty is further complicated by the political implications of legitimately questioning reporting standards. In an international organization, political pressures might restrict how alternative catch scenarios can be investigated. This has been the experience within the CCSBT. During the 1990s, some member scientists explored the potential stock assessment implications of relatively modest catch under-reporting scenarios (modest relative to the under-reporting revealed in the market review). These scenarios included under-reported catches in both the longline and surface fisheries, and were motivated, in part, on the basis of 50 longline vessels caught fishing out of season in 1996 (Anon., 1997a). Evaluating the effects of catch uncertainty was agreed, in principle, to be important for the stock assessment process by the Scientific Committee (Anon., 1997b). However, consensus among all parties could not be reached on the specific under-reporting scenarios that should be considered. The reasons for this are not documented, but (from our experience) they seem unlikely to have related purely to scientific or technical factors. This continued to be a politically sensitive issue in the subsequent MP development process, and again proved intractable, even though the Scientific Committee agreed that alternative catch time series were a source of uncertainty that needed to be explored (Anon., 2002).

As the CCSBT and other experiences demonstrate (e.g. International Whaling Commission—Ivashchenko et al., 2007), modelling input data scenarios is a poor substitute for acquiring reliable, high quality data in the first place. Accepting the validity of such an approach to modelling is likely to result in a serious misrepresentation of the real uncertainty. It weakens the argument for improved data provision/validation and thus reduces the likelihood of finding real solutions. There is a need for the explicit recognition within the fisheries community that a small number of modellers and statisticians simply cannot replace adequate field sampling and compliance programs. Without systematic monitoring programs and independent data verification (international observers, etc.), scientists should question whether they are able to provide meaningful management advice or whether they are actually lending credibility to a misguided process. In principle, it might be argued that the precautionary approach could be invoked to ensure that management actions are more (biologically) conservative when the data are more uncertain. However, if there is no way to meaningfully speculate about the magnitude of the uncertainty in the data, there is not much basis for deciding how conservative the management actions should be. There is an urgent need for the fisheries science community to be more pro-active in the development and implementation of independent ways to verify the basic data used in stock assessments and management procedures.

3.3. *Lesson 3: management procedure implementation requires a reliable data stream*

In the CCSBT, all candidate MPs involved a decision-making process that (directly or indirectly) related total catch removals to changes in Japanese longline CPUE (i.e. as with most tuna fisheries, there is assumed to be a quantifiable relationship between commercial fishery CPUE and abundance). If one or both of these data sources is a poor representation of reality, we could not expect the MP to operate effectively. Furthermore, there would be no reliable basis with which to retrospectively evaluate how well the MP had been performing. Unfortunately, this situation would probably not be recognized as long as the data problems remained unknown (unless perhaps the fishery collapsed).

The operation of the Japanese SBT fleet has changed substantially as a result of the adoption of an individual quota management system in 2006, and the substantive quota reductions implemented for 2007–2009. This is expected to have important implications for the interpretation of catch rates as indices of abundance (Anon., 2006c, 2007b), because operational changes in this fleet have caused ambiguities in the past. It was the reduction of the seasonal and spatial distribution of Japanese longline effort in the 1990s that led to speculation of an SBT range contraction. This uncertainty eventually led to the unilateral Japanese Experimental Fishing Program and the ITLOS hearing. While we do not know how the fleet will respond to the new management regime, there will presumably be some sort of discontinuity in the time series. This is under the optimistic assumption that the old CPUE series was not affected by the catch under-reporting problem. However, there is not likely to be much confidence in the old CPUE series unless additional data can be uncovered to further quantify the mechanism of the catch under-reporting and how it affected the CPUE. Furthermore, even if the data reporting problems can be eliminated in the future, it might still be optimistic to expect that the fleet behaviour will rapidly converge toward a stable and informative pattern. Further management actions could disrupt the fleet behaviour again in the near future (e.g. if and when an MP is adopted). In light of this, the CCSBT members are actively investigating alternative methods for obtaining abundance estimates that do not rely on CPUE as a primary input, including aerial surveys (Eveson et al., 2007), conventional tagging (Polacheck and Eveson, 2007) and close-kin genetics (Bravington and Grewe, 2007).

Without reliable catch data and abundance indices (or fishing mortality estimates), it seems unlikely that an MP could actually deliver robust management performance (at least in the context of the type of fishery management discussed in this paper). Some signal is required to quantify the relationship between fishing activity and the fish population. The need for accurate and verifiable catch reporting is a paramount priority. For many fisheries, we would expect that this is only possible with some combination of meaningful at sea observations (either via observers or automated electronic monitoring methods), Vessel Monitoring Systems, and integrated Catch Documentation Schemes. But even with reliable catch data, there is no guarantee that reliable abundance indices can be generated from fisheries data, and whenever possible, fisheries independent methods of monitoring fish populations should be sought.

4. (II) MPs: a practical interface between fisheries science and management

> "For if we were looking for gold, you can't suppose that we would willingly let mutual politeness hinder our search and prevent our finding it." Socrates

In our experience, fisheries scientists have tended to be the greatest proponents of MP development, but for the process to be effective, there needs to be an ongoing dialogue among scientific and non-scientific participants. This not only facilitates trust in the process, but also helps to keep the scientists grounded in the real world.

4.1. *Lesson 4: management procedures encourage useful dialogue among industry, managers and scientists*

When first initiated, the MP development process was unfamiliar to most of the SBT stakeholders, and many CCSBT scientists. An iterative consultative process was instituted for the duration of the MP development and was essential for gaining the participation and support from the various stakeholders. The initial emphasis was on stakeholder education, in which the inexperienced scientists, managers, industry and other stakeholder (e.g. conservation group) representatives acquired familiarity with the process. In later iterations, the emphasis shifted towards feedback, in which scientists attempted to outline different options of what was possible, and other stakeholders attempted to articulate their preferences among competing options.

In theory, the different participants have distinct roles:

- Scientists are primarily responsible for providing the best quantitative description of the fishery system (including the representation of uncertainties).
- Managers are responsible for articulating management objectives, including the levels of biological and economic risk that they are prepared to accept on behalf of the public, which includes balancing such diverse interests as those of conservation NGOs and the fishing industry.
- Industry and other stakeholder representatives have to work with the managers to ensure that the levels of risk and trade-offs are adequately defined, and that their practical knowledge of the fishery system is captured by the scientists.

In practice the distinctions among these roles are not likely to be easily delineated.

The CCSBT was very pro-active in engaging the wider stakeholder community with iterative consultations. In the MP process, managers are effectively being asked to make an immediate decision that results in a loss of future decision-making power. Without a full understanding and trust of the process, it would not be surprising if there was reluctance to embrace the final result. Similarly, industry may also be reluctant to agree to the process as it may foreclose their ability to contest scientific advice and recommendations at the political level (which might otherwise be recognized as a possible option in an *ad hoc* decision-making process). During the 4 year CCSBT MP development process, there were 7 separate meetings with full scientific and technical participation, plus smaller technical meetings with limited participation. There were also 3 formal meetings of the CCSBT Commission in which the MP process featured in the discussions. One of the scientists from the independent advisory panel held individual educational briefings for each of the CCSBT members and cooperating non-members. Australia (and presumably most other CCSBT members) also held regular consultations among the domestic stakeholders. Provided that the same individuals were involved with these proceedings, all of the members should have had a reasonable understanding of the process, and there was ample opportunity for the MP developers to solicit and receive relevant feedback from domestic managers and stakeholders.

In the CCSBT, the task of developing and testing the MPs fell primarily to the scientists, because they were the most familiar with the MP process in general, and the technical details of the simulation models in particular. In principle however, any of the stakeholders could have been intimately involved with developing and testing alternative MP algorithms. If a user-friendly interface to the MP software had been developed and distributed, this would have been a powerful educational tool and might have led to some new ideas. While the scope for creative decision rules in the CCSBT context was rather limited due to the (perceived) state of the spawning stock (see lesson 10), other fisheries might gain more benefit from a broader involvement in decision rule formulation.

We would argue that one of the greatest benefits of the CCSBT MP process was the facilitation of these stakeholder interactions. It helped to increase the general understanding of the fishery for all involved, and provided the mechanism and context for feedback between stakeholders and MP developers. It would be unfortunate if the recent revelations of the historical data problems have undermined the benefits of communication and cooperation that were gained through these consultations.

4.2. Lesson 5: poorly defined management objectives do not necessarily prevent the development of effective management procedures

Advocates of the MP approach often describe a straightforward sequence of steps, one of which involves a precise articulation of management objectives. This is typically identified as the responsibility of the managers, who must attempt to operationalize legislative obligations and institutional guidelines. However, we would not expect that objectives can be easily defined in quantitative terms. Multiple objectives will usually be relevant, and they will inevitably conflict to varying degrees. The failure of the CCSBT to articulate specific objectives was initially frustrating for the MP developers, but this did not prove to be a serious impediment to the formulation of a sensible range of decision rules.

The CCSBT has only ever clearly articulated one management objective (rebuilding the spawning biomass (SSB) to 1980 levels by 2020). Early in the MP process, it became evident that this objective was not likely to be achievable (even with zero catches). It was also generally accepted that the substantial TAC reductions that would have been required to minimize the stock rebuilding time would be severely disruptive to industry. Provided that the biological risk (i.e. the probability of decreasing the biomass to a sufficiently low level that the long-term recruitment potential, or stock structure, is impaired) could be managed through effective, but slower rebuilding, it was considered preferable not to induce major fishery disruptions. The rebuilding objective provided the primary objective with which developers could begin to explore MP behaviour, however, the Commission initially provided no guidance about the relative importance of other management trade-offs. While numerous performance measures were calculated, most of the discussion focused on three (relative) trade-offs:

- Reduced biological risk results in reduced long-term average catch.
- Large initial TAC cuts result in higher overall catch (over the 20–30 year time horizon evaluated) due to more rapid stock rebuilding.
- Reduced inter-annual catch variability results in reduced average catches (greater catch stability is achieved at the expense of a lessened capacity to respond rapidly to risky situations or take advantage of strong year classes).

As discussed in lessons 9 and 10, the range of MP behaviour that was judged to be sensible within this trade-off space was rather limited.

Some of the scientists indicated a desire to add another dimension to the problem, by examining the implications of alternative national quota allocations among CCSBT members. Given that the Australian (predominantly purse seine) catches are mostly aged 2–4 years, and the longline fisheries target fish of age >5, it is conceivable that large shifts in allocations among the fleets could have substantial impacts on management outcomes. It might be argued that the choice to illustrate the implications of alternative allocation options was no different than exploring alternative rebuilding options. However, given that catch allocations were the result of a protracted negotiation during the founding of the CCSBT, and subsequent negotiations with the addition of new members, other scientists felt that it would a waste of resources to provide advice on options that were extremely unlikely to be adopted. In this respect, the scientists prioritized and constrained the workload in relation to the political reality of the situation (but would have reprioritized the task list if the Commission requested advice in this respect).

Clearly articulated management objectives may not be obtainable at the outset of an MP development process. However, this does not preclude the development and selection of decision rules. Some management objectives will conflict in ways that are not likely to be quantitatively predictable unless they are explicitly examined in an MP evaluation process. The iterative exploration and illustration of trade-offs provides a contextual framework for managers to make risk-weighted decisions in relation to the specific circumstances of the individual fishery.

4.3. Lesson 6: management procedures may not reduce the scientific resources required to manage a fishery

One of the advantages occasionally listed for the MP process is the expectation that there will be a reduction in the resource-consuming processes of fisheries stock assessment, and the frequently acrimonious debates involved with deciding on management measures (e.g. TAC setting). However, observers of the CCSBT MP process might be left with the impression that it is actually a more resource intensive and controversial process than the dysfunctional process that it was intended to replace.

One of the problems in the CCSBT case involved the cyclic process of conditioning operating models (which is somewhat analogous to the stock assessment process), and retesting management procedures. If this process takes too long to complete, new data become available, and it might be argued that conditioning should be repeated with the new data. One of the problems arising in the CCSBT case arose in 2004, when there was an unexpectedly high CPUE observation corresponding to 2003 data (which was not included in the original operating model conditioning). Most, or all, of the candidate MPs examined at that time would have recommended an immediate TAC increase, despite the prevailing perceptions of sustained low stock biomass and the recent evidence for poor incoming recruitment (e.g. Kolody and Hartog, 2004). This problem of new data and information has been recognized in other systems, and in the case of the International Whaling Commission, has led to the recommendation that firm time limits should be applied to some phases of MP development (e.g. Punt and Donovan, 2007). While this is likely the only viable way to proceed in practice, problems will arise if the MP has to be revised as soon as it is implemented and negates one of the primary advantages of an MP over stock assessment advice. It could also undermine manager and stakeholder confidence in the MP approach. We would expect this sort of problem to arise frequently if the uncertainty in the underlying operating model has not been sufficiently quantified.

In addition to the increased scientific work involved with undertaking the initial MP evaluation, there are several situations in which we would expect the implemented MP to deviate from the automated ideal and require additional effort. These situations are discussed under the following three headings.

4.3.1. Developing a framework or "meta-rule" for handling exceptional circumstances

The CCSBT spent considerable time developing a framework (i.e. a "meta-rule") for handling "exceptional circumstances" (i.e. situations which were not covered in the MP simulation testing). The meta-rule allows the automated TAC setting from the MP to be over-ridden, and provides assurance to managers and industry that they would not be forced into accepting ridiculous decisions if unforeseen situations arise. Several general classes of exceptional circumstance were identified for the SBT fishery:

- New data might suggest that the stock status is outside of the envelope of plausible uncertainty that was admitted during the decision rule evaluation.
- The data underpinning the decision rule may no longer be available (or have fundamentally changed) so that it no longer has the information content assumed during simulation testing.
- Nations outside of the CCSBT could begin catching substantial numbers of SBT, and hence, the CCSBT might not be able to regulate a large portion of the catch.

The CCSBT MP was intended to have regular oversight by the Scientific Committee. Every year there was to be a qualitative review of the SBT fisheries indicators to determine if there was any evidence that the assumptions underpinning the MP development were being violated. While it would be difficult to identify every possible problem, these events would be expected to be substantial and rare, and not simply a means for any member to "tinker" with a specific MP decision.

If exceptional circumstances were identified, CCSBT MP metarules were to be invoked, and the normal MP cycle suspended. Additional analyses would be undertaken under the direction of the Scientific Committee, and an *ad hoc* decision-making framework would be adopted until such time as the MP was judged to be appropriate again, or a revised MP could be implemented if required.

4.3.2. Retrospective MP evaluation

In addition to the regular MP oversight, it was an explicit expectation that the behaviour of the MP would be thoroughly reviewed at regular intervals (~5 years). This would be expected to involve a comprehensive stock assessment, and an evaluation of whether the MP had been providing adequate performance with respect to management objectives. If the performance was not adequate, this would likely initiate MP revision.

4.3.3. MP revision

We would not expect that any MP would be sufficient to automate the management decision process in perpetuity. MP Revision would be expected in several circumstances: (i) exceptional circumstances might indicate that the MP is no longer appropriate, (ii) retrospective evaluation might indicate that management objectives are not being met, (iii) the Commission might decide that management objectives have changed, or (iv) new data and analytical methods might become available with which the performance of the MP would be expected to be improved. Revising the MP would, in many respects, involve a full repetition of the development process (though perhaps expedited as a result of experience gained previously). Revision should form part of an agreed schedule for the management of the fishery, and not be used as a pretext for interfering with specific decisions that are within the range of situations that the MP was designed to cope with.

An MP that automates the assessment and management decision process is an ideal to aim for, but the reality is likely to be more complicated. Thus, it is important to recognize that the MP process may not substantially reduce the need for ongoing fisheries science, and could actually increase the resources required (e.g. the IWC MP requires routine fishery independent abundance surveys, a requirement that did not exist previously). However, the MP process does provide a contextual framework to help prioritize research (e.g. Punt and Donovan, 2007). In principle, it should be possible to use the MP simulation process to evaluate the relative costs and benefits associated with reducing uncertainty in relation to the attainment of management objectives (although fisheries research frequently reveals that we are more uncertain about some things than had been assumed before the research was undertaken).

5. (III) Representing uncertainty in fisheries systems

"...we know there are some things we do not know. But there are also unknown unknowns, the ones we don't know we don't know." Donald Rumsfeld

The dynamics of most fisheries systems cannot be quantified very reliably. Fisheries data can be relatively uninformative, sometimes grossly inaccurate, and often the statistical properties are poorly understood. The mathematical models used in stock assessment invoke numerous untested (often untestable) assumptions, and population parameter estimates are frequently sensitive to these assumptions (e.g. Schnute and Richards, 2001). The MP approach is promoted as a method for identifying management systems that should be robust in the face of uncertainty. However, there is no simple prescription for quantifying the uncertainty. Participants should be careful to recognize the implications of inevitable arbitrary decisions, and the potential for the process to be subtly influenced (consciously or unconsciously) along the way. Unfortunately, there often seems to be a natural tendency for group dynamics to gravitate towards optimistic interpretations (e.g. Bax et al., 2005), and efforts must be made to ensure that a broad range of plausible sources of uncertainty are included in fisheries models.

5.1. Lesson 7: there is a large element of subjectivity in the quantification of stock status uncertainty

In the CCSBT MP process, there was a strong focus on attempting to quantify the uncertainty in stock status and production dynamics as revealed through a procedure largely analogous to a comprehensive stock assessment. While this explicit conditioning (fitting to data) is not the only way that operating models can be formulated, it seems like a sensible starting point to at least ensure that the representation of reality is reasonably consistent with whatever data are available. However, given the somewhat arbitrary decisions involved with model formulation, it is difficult to obtain a conviction that alternative representations of reality have not been overlooked. These missing models could be equally consistent with the data, but have substantially different management implications.

The development of CCSBT operating models involved a reasonably extensive coverage of model and parameter uncertainty. Several hundred models were defined in a multidimensional “grid” that consisted of different plausible combinations of assumptions, fixed parameter values and data sets (Anon., 2004). The best fit (Maximum Posterior Density) state/parameter estimates from each model were combined to describe the current stock status uncertainty and future stock/fishery dynamics. The statistical estimation uncertainty associated with each individual model was not considered, but important elements of the parameter uncertainty were encompassed by the combination of models. The different models were also weighted differently in the aggregated uncertainty set (e.g. models judged to be more plausible were given a higher weighting, and hence were more frequently represented in the simulation testing, than less plausible models). These relative plausibility weightings were determined by model likelihood values, if it was believed that the data were likely to be informative with respect to the parameter in question. Otherwise, weightings were determined by prior opinions expressed by participants (either by consensus, or the average of a secret ballot). The likelihood weightings were calculated using conventional statistical techniques. However, it was recognized that the modelling approach was a pragmatic and *ad hoc* process, in which statistical assumptions and inferences could not be rigorously defended. For example, catch-at-length sample sizes in the multinomial likelihood terms deviated from the true sample sizes in recognition of the fact that (i) fishery size composition sampling is rarely truly random, and (ii) “separable” fishery selectivity assumptions in the population dynamics models are never entirely correct. In some instances, input priors (e.g. variances for different input data) were varied to ensure an acceptable distribution of likelihood values for parameter values estimated by the model. The CCSBT had originally intended to include only a limited subset of conditioned operating models, and sought to encompass the stock status uncertainty described by the Bayesian posterior distributions. However, it proved to be infeasible to adequately quantify the posteriors of some key parameters, and the CCSBT participants felt that the MPD grid was a reasonable and feasible alternative (Anon., 2004).

Following the approach of the International Whaling Commission, there were actually two operating model grids in the CCSBT process, a high priority “reference set” and a lower priority “robustness set”. The reference set was originally intended to represent a sort of “best estimate” of the system uncertainty (i.e. including reasonably probable alternative models with differing management implications). The robustness set was a much more loosely defined set of models that included potentially troublesome, but (supposedly) much more unlikely scenarios (e.g. to test how the MPs behave if there is poor recruitment for the next 6 years). Originally, the robustness set was expected to represent a second tier criterion for selecting MPs, i.e. MPs that performed similarly with respect to the reference set of operating models could be compared on the basis of performance with respect to the robustness set. The robustness set was also intended to test whether “pathological” behaviour would result if some of the more unlikely scenarios did in fact occur (i.e. to ensure that the MPs were robust in these situations). The main focus at the end of this process was on the reference set, and the potential usefulness of the robustness set never seemed to reach fruition.

Reaching agreement on what to include in the MPD grid was a protracted process that required several iterations. For those that were not involved, this process may create the impression of objectivity, and a comforting conviction that rigorous mathematical methods were used to bound and characterize the uncertainty of the SBT fishery system. In reality, the exercise of bounding uncertainty was rather subjective, and the probability distributions could be substantially skewed one way or the other, depending on what area of the plausible model/parameter space was searched most thoroughly, or how strongly different proponents argued for inclusion of different models. There was actually surprisingly little disagreement about which models to include once the individual models were fit. But none of the model fits were entirely satisfactory (e.g. they exhibited some degree of systematic lack of fit between model predictions and observations), and there were divergent views about how much exploration should be undertaken. In the end, pragmatic time constraints limited the uncertainty exploration.

The operating model conditioning (and prior experience of participants) indicated that three of the most important factors to explore in quantifying uncertainty for the SBT stock included the stock recruitment relationship, natural mortality, and nonlinear relationships between abundance and catch rates. Not surprisingly, these were three of the main uncertainties that the Scientific Committee had identified at least as early as 1997 (e.g. Anon., 1997b). With hindsight, however, it is interesting to note that none of the scientists expressed any serious concern about not including catch under-reporting scenarios, even though under-reporting had been agreed by the Scientific Committee as a major source of uncertainty in the stock assessment process (Anon., 1997b).

We consider it essential to quantify uncertainty in operating models through the process of conditioning to historical data. However, given that there seems to be a disturbingly high probability of not adequately quantifying fishery system dynamics, we question whether the importance of historical conditioning has been over-emphasized relative to the importance of constructing plausible scenarios about the future (as discussed in the subsequent lesson).

5.2. Lesson 8: historical data are not sufficient to describe future uncertainty

With suitable effort, one can generally identify a broad range of operating models that are plausibly consistent with the historical data. However, there still seems to be a tendency for fisheries scientists to be overconfident in their predictions about the future (e.g. Hutchings and Reynolds, 2004 discuss difficulties in predicting recovery rates of exploited fish populations), and this has negative implications for the reliability of MP simulation testing. We speculate that there may be merit in decreasing the effort spent conditioning operating models to historical data (particularly if the data are of dubious quality), and increasing the effort spent exploring alternative plausible future dynamics scenarios.

In the CCSBT operating models, future dynamics were assumed to have similar characteristics to the historical estimates. Functional relationships and parameters were generally assumed to be stationary. Scenario-specific variability in stochastic model components was estimated from the conditioned model, and assumed to continue into the future (e.g. variances of recruitment deviations, CPUE observation errors, etc.). There was an attempt to admit time series structure to some stochastic processes in the models (e.g. auto-correlated errors in projected CPUE observations, and recruitment deviations). Auto-correlated processes are expected in ecological systems, unfortunately, these processes are usually difficult to quantify. In the SBT operating models, auto-correlated error distributions were adopted to compensate for (what would otherwise have been described as a) systematic lack of fit between model predictions and observations. The inclusion of auto-correlation had the desirable effect of lessening the sudden time series discontinuity that can result in the first year(s) of projections. However, there was not much effort spent considering plausible operating models with non-stationary characteristics (one exception involved brief periods of successive low recruitment).

The Scientific Committee implicitly adopted the position that it would be preferable to limit simulation testing to scenarios that could be quantified on the basis of the historical data, and rely on meta-rules to overcome unexpected circumstances in the future. This is the position advocated by Butterworth (2008), with the justification that it is preferable to have the MP process adopted with a relatively high probability of experiencing exceptional circumstances, rather than for the process to not be accepted at all. While there is pragmatic justification for this argument, there is a quantitative question of degree that needs to be addressed (i.e. ignoring plausible uncertainty for expediency would likely bring scientific integrity into question). Three lines of evidence suggested exceptional circumstances during the last 2 years of the CCSBT MP development process: recruitment at very low (unprecedented) levels, a new CPUE observation outside the bounds of any operating model predictions, and the revelation of catch under-reporting. SBT may be an extreme case (as the MP was suspended before the first TAC setting), but the point remains: if the probability of exceptional circumstances is too high, *ad hoc* decision-making becomes the norm, and the main advantage of the MP approach will not be realized. This is likely to undermine the credibility in the MP for the specific fishery, and the broader reputation of the MP process. Since the suspension of the MP, feedback to the authors from scientists, managers and policy makers (that are outside of the CCSBT MP process) have suggested that this has been the case.

We speculate that it would be worthwhile entertaining broader notions about future uncertainty than can be justified solely on the basis of the data from any specific fishery (particularly those with short time series, and/or questionable data quality). The substantial volume of fisheries management case histories might be used to quantify the probability of observing supposedly "unlikely" events. There may be a relatively small number of recurrent fisheries modelling problems that might be simulated in future projections (despite not having been observed in a specific fishery) and which might reduce the frequency of MP exceptional circumstances to an acceptable level (e.g. plausible model elements to focus on include: extreme and temporally correlated recruitment events, commercial CPUE hyper-stability). Butterworth et al. (1996) suggest a useful framework for considering how experience might be translated across systems in the context of the International Whaling Commission. Toward this end, we would encourage a comprehensive quantitative meta-analysis or systematic review of fisheries modelling problems relevant to the development of MPs.

Butterworth (2008) advocates against the position of expanding uncertainty, because (i) increased uncertainty will usually lead to more conservative TACs, and (ii) the more conservative TACs are likely to cause industry stakeholders to question the arbitrary nature of the uncertainty expansion. With respect to point (i): we would argue that if prior experience across a broad range of fisheries indicates that biological risk tends to be underestimated, then more precautionary TAC settings are justified. Undoubtedly some fisheries would suffer a short-term loss of economic opportunity as a result, but on average, these losses should be offset through future TAC increases, and reduced risk of fishery collapse (particularly in overfished situations). With respect to point (ii), we would argue that any attempt at quantifying fisheries uncertainty inevitably involves an element of arbitrariness. If relevant experience from other systems exists, it would not be an arbitrary decision to include it in the MP testing, it would be a negligent decision to exclude it.

The nature of the disagreement about where to focus the MP uncertainty quantification is currently qualitative in nature, and based on differing individual experiences. If the debate could be quantified (e.g. through a comprehensive fisheries meta-analysis or systematic review), perhaps much of the apparent disagreement would disappear. We advocate that scientists should be striving for the best representation of system uncertainty at each level of the MP process. Rather than institutionalizing the understatement of risk in the formulation of MP simulation models (in order to expedite adoption of an MP), fisheries managers should be made aware (to the extent possible) of the actual level of risk that they are accepting when making their decisions. MP developers should provide an explicit estimate of the probability (or target frequency) with which exceptional circumstances are expected to arise (at least with respect to the biological elements of the system). This is obviously difficult to quantify accurately. However, with an agreed target, the different participants at least have a common objective in mind, a quantity that they can relay to managers (so that managers do not have to second guess what scientists are referring to when they describe risk), and a benchmark against which scientists can retrospectively evaluate the effectiveness of the uncertainty quantification approach.

6. (IV) Choosing a decision rule

> "If one does not know to which port one is sailing, no wind is favourable." Lucius Annaeus Seneca

In 2005, the MP developers were tasked with recommending a single MP to the Commission, from a pool of four competing candidates developed by the CCSBT member scientists (Anon., 2005c, mathematical details and performance characteristics are summarized in Branch and Parma, 2005). While the development of the decision rules was undertaken independently, there was considerable discussion and transfer of ideas among developers, and all of the candidate MPs were thought to have benefitted as a result of the parallel and competing development processes. The decision rule that was ultimately selected involved the fitting of a simple age-aggregated (Fox) production model to the total catch and Japanese longline CPUE data (i.e. refitting the model each time a TAC recommendation is required). The TAC recommendation was a function of the estimated surplus production, the current level of depletion, size composition in the longline catch (as an index of recent recruitment), and the ratio of current CPUE to initial CPUE (plus additional control parameters that moderated the rate of TAC change). Other model-based and data-based decision rules were judged to have superior performance with respect to individual management objectives (e.g. catch stability among years, capacity to rebuild the stock and increase exploitation in the long-term), but the recommended MP was accepted as having the best performance trade-off overall. Selecting this MP from among the candidates was unexpectedly easy, as described in the last two lessons.

6.1. *Lesson 9: tuning management procedures to remove one dimension of the trade-off problem*

MP decision rules should have at least one parameter that can be adjusted to change their performance (e.g. if no internal 'gain' parameter in the rule is appropriate, a coefficient can be added to proportionally increase or decrease the TAC that would otherwise have been recommended). The selection of an MP involves two basic decisions: the choice of the functional form of the decision rule, and the specific set of parameter values to use with that decision rule. The evaluation of the performance of different decision rules is confounded with the choice of parameter values. "Tuning" (used extensively in the International Whaling Commission) usually refers to the process of adjusting the parameters of a decision rule so that one of the more important management performance measures (e.g. median biomass rebuilding) attains a precise value. With the different MPs tuned to the same criterion, their performance can be compared with respect to other (and potentially less important) criteria. Tuning removes one dimension from the evaluation problem, and allows one to make comparative statements such as: for the same level of biological risk, rule B can be expected to provide higher catches than rule A over the next 10 years.

Tuning in the case of SBT proved an extremely useful approach for comparing decision rules. Some decision rules could be eliminated simply on the basis of being unable to achieve a requisite tuning level (e.g. adequate stock rebuilding not possible). Other decision rules could be tuned and eliminated from the candidate pool, as it was readily evident that they provided generally worse performance across the other high priority performance indicators. The value of tuning in the CCSBT case was enhanced by the fact there were only three fundamental dimensions to the defined management objectives: biological risk, magnitude of catches and inter-annual catch stability. A large range of performance measures were actually calculated, but it was recognized that each measure tended to relate to one of these general objectives, and the different performance indices tended to be highly correlated with other indices that related to the same general objective.

It was relatively easy to agree on the tuning objective for the CCSBT MP, because the only clearly articulated management objective related to the rebuilding of the spawning stock. Decision rules were tuned to provide the same median spawning stock rebuilding by 2022 relative to the current level (though there was some disagreement as to whether the median was more appropriate than a lower, more precautionary, percentile). Different values of the tuning target were explored in successive consultative iterations, and periodically revised in relation to Commission requests (and the biological productivity limits that were resulting from the evolving operating models).

The choice of tuning criterion was not expected to influence the MP selection process. However, in some cases, the relative performance of the candidate MPs did vary appreciably, depending on (i) the actual value of the rebuilding target (e.g. biomass in 2022 either 10% or 30% larger than in 2004), or (ii) whether the rebuilding objective was based on the 50th or lower 10th percentile of the projected biomass distribution. Unless the management objectives are clearly defined in relation to a specific tuning objective, MP developers should compare decision rules across a range of tuning levels to prevent promising candidate MPs from being eliminated prematurely (i.e. relative performance might not be estimated very well by extrapolation).

6.2. *Lesson 10: in selecting a management procedure, the form of the decision rule might be less important than the management objective*

Despite the varying levels of complexity in the candidate CCSBT MPs, the differences in performance were not overwhelming. The pessimistic operating models in the final round of the MP process required substantive quota cuts from all candidate MPs to meet even minimal rebuilding objectives. The major differences in the behaviour of the decision rules were related to relatively minor differences in the magnitude of the initial TAC reductions. Rules with larger initial TAC reductions allowed more rapid rebuilding, lowered the biological risk, and allowed slightly higher cumulative catches over the 20 years projection period. Rules which allowed greater variability in TAC among years posed less biological risk, in that they could respond more quickly to stock declines (though most of the risk was associated with the actions in the initial TAC settings).

Coincident with the final MP selection, the Scientific Committee actually made a recommendation for immediate TAC reductions that exceeded the limits that the MPs were allowed to make. This might be described as a meta-rule application on top of the MP adoption. If this additional action was not taken, it is not clear that the MP that was selected would have remained preferable to other MPs. The catch stability offered by the selected rule resulted in a greatly elevated biological risk in the first few years relative to other rules with reduced catch stability.

In hindsight, these results are qualitatively consistent with what one would expect for any overfished stock that requires most of its surplus production to remain unharvested in order to meet a rebuilding objective. Other factors were identified which would be expected to constrain the variability in performance among decision rules:

- Amount of uncertainty embedded in the operating model—if there is little variability among the operating model scenarios, there is little scope for decision rules to achieve different behaviour (while still meeting the same tuning objective).
- Information content in the future data—if the decision rules cannot extract useful information about the system from the data, there is not much scope to differentially act on the data.
- Time horizon—a short time horizon offers little opportunity for MPs to "learn" about the system (e.g. distinguish whether a scenario involves a high or low productivity population) and hence little opportunity to exploit this information. Similarly, with a relatively long-lived species, the age structure of the population at the beginning of the MP simulations provides considerable inertia, such that the implications of future stochastic recruitment have relatively minor influence on the spawning stock for several years.

- Constraints imposed on the maximum amount of change in TAC allowed from year to year (i.e. if all rules recommend the same maximum change, the outcome will be the same).

We did not anticipate the degree of similarity in management performance arising from the different SBT MPs. However, this does offer some hope that if an MP process is engaged to resolve a crisis, it might not be prohibitively difficult to reach agreement on a decision rule (provided a management objective can be agreed). But it also suggests that if managers want to have real choice in the behaviour of the MP, the process should be investigated well before a crisis arises.

7. Conclusions

The development of an MP for the SBT fishery has taken a circuitous path. The idea was first promoted in the early 1990s, MP development was actively pursued during 2002–2006, and the future is uncertain, with the first TAC recommendation now scheduled to be reported to the Commission in 2011–2012. It is easy for critics to describe the suspended implementation of the CCSBT MP as a failure, however, we recognize that tangible progress was made during this process, including: (i) 15–20 years of substantial catch under-reporting problems were identified, and (ii) the CCSBT reached its first consensus agreement to substantively change the TAC since 1989. The extent to which the MP was directly responsible for this progress is debatable, but we were left with the impression that the MP development process facilitated a consensus view of the stock status and management implications that might not have been achieved within the conventional CCSBT assessment process. The main lessons from the CCSBT MP experience can be summarized in the following points:

1. An MP/MSE approach to fishery management will not alleviate the need for difficult management decisions. There may be pressure to view the period of MP development as a period where management action is not required, but succumbing to this temptation will likely result in more severe decisions when the MP is finally activated.
2. The effectiveness of an MP will be dependent on the quality of the data that are used to test it. Sophisticated scenario modelling and statistical methods should not be embraced as an adequate substitute for field sampling, monitoring and compliance programs that are required to collect and verify high quality informative data.
3. The effectiveness of an MP will be dependent on the quality of the data that are used to implement it. Even if the reliability is excellent, reliance on fisheries dependent data will not guarantee an effective MP because fishing operations are bound to change over time (perhaps in response to the MP). Fisheries independent (or semi-independent) data should be collected to ensure consistency with the MP requirements.
4. MP development and implementation should be undertaken with a spirit of cooperation among scientists, managers, industry and other stakeholders, each of whom have an important role to play in the process. Consultation is critical for gaining a trust in the process, and producing a more robust result.
5. It may not be possible for managers and stakeholders to articulate precise management objectives at the outset of the MP process, but this does not necessarily preclude a successful development process. Management trade-offs can be illustrated in the context of long-term management outcomes, such that decision makers can identify preferable trade-offs among objectives, even if they could not have explicitly defined them *a priori*.
6. The adoption of an MP process will not necessarily result in a substantial reduction in scientific investment in fishery management, because the automated system still needs to be monitored to check for exceptional circumstances (i.e. situations might arise for which the MP was not tested), the MP performance will need to be periodically evaluated to assess whether the management objectives are being met, and the MP may need to be revised in relation to changing management objectives and/or, important new information.
7. It is difficult to develop operating models that adequately reflect the uncertainty in fisheries systems. While the mathematics might provide the appearance of rigorous objectivity, the decisions made in the process of model selection (in particular) include a large element of subjectivity, and potentially could distort the uncertainty, either through chance or deliberate manipulation.
8. If fisheries uncertainty is understated in the MP development process, this tends to result in higher than intended biological risk, increased frequency of exceptional circumstances occurring once the MP is operational, and potentially lost economic opportunity in the longer term. Conversely, over-stating future uncertainty will probably lead to overly conservative management outcomes (likely lower short-term catches, and possibly lower long-term catches as well). Either extreme has the potential to undermine scientific credibility in the MP process. Scientists should strive to convey the best representation of uncertainty and risk at all levels of the MP development process, and we recommend that a meta-analysis or systematic review of fisheries modelling problems relevant to MP evaluation should be conducted to improve the quantification of uncertainty.
9. The process of "tuning" MPs facilitates the selection process by illustrating the management behaviour of different MPs that are adjusted so that they provide equal performance with respect to the highest priority management objective. In this way, performance with respect to secondary (and tertiary) management objectives can be used as selection criteria (e.g. given that two tuned rules provide identical behaviour with respect to acceptable biomass risk, one can choose the rule that provides better economic performance).
10. The management performance of the MPs might be largely determined by the management objectives (tuning criteria above), rather than the capacity of the decision rule to extract useful information from the data. The extent that this is true will depend on (i) whether the decision rule is constrained by rebuilding objectives (e.g. if rapid rebuilding is required, the only choice is to cut catches so that the stock rebuilds), (ii) the amount of uncertainty in the population (e.g. if all operating models are almost the same, none of the rules will have much capacity to distinguish among alternative population dynamics), and (iii) how informative the data are (e.g. if data are highly uncertain, none of the decision rules will be able to distinguish among operating models and differentially exploit the differences).

Acknowledgements

We are indebted to numerous colleagues in the CCSBT Scientific Committee for the colourful exchange of ideas over many years that led to the development of this manuscript. Thanks to Doug Butterworth, Martin Dorn, Sally Wayte, and two anonymous reviewers, for useful suggestions (particular thanks to MD for his diplomatic thesaurus and title suggestion). This paper is based on projects partially funded by the Australian Commonwealth Scientific and Industrial

Research Organization, Australian Fisheries Management Authority and the Department of Agriculture, Fisheries and Forestry over a number of years.

References

Anon., 1994. First Meeting of the Commission for the Conservation of Southern Bluefin Tuna, May 1994. Wellington, New Zealand.

Anon., 1995a. CCSBT Report of the Second Annual Meeting. Tokyo, Japan, 12–15 September.

Anon., 1995b. CCSBT Report of the First Special Meeting. Canberra, Australia, 3–6 October.

Anon., 1996. CCSBT Report of the Third Annual Meeting. Canberra, Australia, 24–28 September.

Anon., 1997a. CCSBT Report of the Resumed Third Annual Meeting (Revised). Canberra, Australia, 18–22 February.

Anon., 1997b. CCSBT Report of the Fourth Annual Meeting. First Part. Canberra, Australia, 8–13 September.

Anon., 1998. CCSBT Report of the Fourth Annual Meeting. Second Part. Canberra, Australia, 19–22 January.

Anon., 1999a. CCSBT Report of the Fifth Annual Meeting. First Part. Tokyo, Japan, 22–26 February.

Anon., 1999b. CCSBT Report of the Sixth Annual Meeting. First Part. Canberra, Australia, 29, 30 November.

Anon., 2000. CCSBT Report of the Special Meeting. Canberra, Australia, 16–18 November.

Anon., 2001. CCSBT Report of the Second Stock Assessment Meeting. Tokyo, Japan, 19–28 August.

Anon., 2002. CCSBT Report of the First Management Procedure Workshop. Tokyo, Japan, 3–4 and 6–8 March.

Anon., 2004. CCSBT Report of the Fifth Stock Assessment Group. Seogwipo City, Republic of Korea, 6–11 September.

Anon., 2005a. CCSBT Report of the Tenth Meeting of the Scientific Committee. Narita, Japan, 9 September.

Anon., 2005b. CCSBT Report of the Twelfth Annual Meeting of the Commission. Narita, Japan, 15 October.

Anon., 2005c. CCSBT Report of the Sixth Meeting of the Stock Assessment Group. Taipei, Taiwan, 29 August–3 September.

Anon., 2006a. CCSBT Report of the Special Meeting of the Commission. Canberra, Australia, 18–19 July.

Anon., 2006b. CCSBT Report of the Seventh Meeting of the Stock Assessment Group. Tokyo, Japan, 4–11 September.

Anon., 2006c. CCSBT Report of the Eleventh Meeting of the Scientific Committee, September.

Anon., 2006d. CCSBT Report of the Thirteenth Annual Meeting of the Commission. Miyazaki, Japan, 10–13 October.

Anon., 2007a. CCSBT Report of the Fourteenth Annual Meeting of the Commission. Canberra, Australia, 16–19 October.

Anon., 2007b. CCSBT Report of the Twelfth Meeting of the Scientific Committee. 10–14 September.

Australia, 2005. Comparison of CCSBT catch data with Japanese auction sales of frozen SBT. Commission for the conservation of southern bluefin tuna working paper CCSBT-ESC/0510/25.

Bax, N.J., Tilzey, R., Lyle, J. Wayte, S.E., Kloser, R.J., Smith A.D.M., 2005. Providing management advice for deep-sea fisheries: lessons learned from Australia's orange roughy fisheries. In: Shotton, R. (Ed.). Deep Sea 2003, pp. 259–272. Conference on the Governance and Management of Deep-sea fisheries. Part 1: Conference Reports, Queenstown December 1–5, 2003. FAO Fisheries Proceedings No 3/1, Rome FAO, 718 pp.

Branch, T., Parma, A., 2005. Performance of the final candidate management procedures selected at the 4th management procedure workshop. Commission for the conservation of southern bluefin tuna working paper CCSBT-ESC/0509/15.

Bravington, M., Grewe, P., 2007. A method for estimating the absolute spawning stock size of SBT, using close-kin genetics. Commission for the conservation of southern bluefin tuna working paper CCSBT-ESC/0709/18.

Butterworth, D.S., 2007. Why a management procedure approach? Some positives and negatives. ICES J. Mar. Sci. 64, 613–617.

Butterworth, D.S., 2008. Some lessons from implementing management procedures. In: Tsukamoto, K., Kawamura, T., Takeuchi, T., Beard Jr., T.D., Kaiser, M.J. (Eds.), Fisheries for Global Welfare and Environment, 5th World Fisheries Congress, 2008. TERRAPUB. pp. 381–397.

Butterworth, D.S., Cochrane, K.L., Oliveria, J.A.A., 1997. Mangement procedures: a better way to manage fisheries? The South African experience. In: Pikitch, E.K., Huppert, D.D., Sessenwine, M.P. (Eds.), Global Trends: Fisheries Management, 20. American Fisheries Society Symposium, Bethesda, Maryland, pp. 83–90.

Butterworth, D.S., Punt, A.E., Smith, A.D.M., 1996. On plausible hypotheses and their weighting, with implications for selection between variants of the Revised Management Procedure. Rep. Int. Whal. Commn. 46, 637–640.

Caton, A.E., 1991. Review of aspects of southern bluefin tuna biology, population and fisheries. In: Deriso, R.B., Bayliff, W.H., La Jolla (Eds.), World Meeting on Stock Assessment of Bluefin Tunas: Strengths and Weaknesses. Special report. Inter-American Tropical Tuna Commission, California, pp. 181–357.

Constable, A.J., de la Mare, W., Agnew, D.J., Everson, I., Miller, D., 2000. Managing fisheries to conserve the Antarctic marine ecosystem: practical implementation of the Convention on the Conservation of Antarctic Marine Living Resources (CCAMLR). ICES J. Mar. Sci. 57, 778–791.

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

de la Mare, W.K., 1986. Simulation studies on whale management procedures. Thirty sixth Report of the International Whaling Commission. Document SC/37/O14.

Eveson, P., Bravington, M., Farley, J., 2007. Aerial survey: updated index of abundance and preliminary results from calibration experiment. Commission for the conservation of southern bluefin tuna working paper CCSBT-ESC/0709/12.

Firestone, J., Polacheck, T., 2003. The effectiveness of the UN convention on the law of the sea in resolving international fisheries disputes: the Southern Bluefin Tuna case. In: Harrison, Neil E., Bryner, Gary C. (Eds.), Science and Politics in the International Environment. Rowman and Littlefield, Lanham, pp. 241–270.

Hutchings, J.A., Reynolds, J.D., 2004. Marine fish population collapses: consequences for recovery and extinction risk. Biol. Sci. 54, 297–309.

Ivashchenko, Y.V., Clapham, P.J., Brownell Jr., R.L. (Eds.), 2007. Scientific reports of Soviet whaling expeditions in the North Pacific, 1955–1978. U.S. Dep. Commer., NOAA Tech. Memo. NMFSAFSC-[175,] 36 pp. (Translation: Y.V. Ivashchenko) + Appendix.

JFA, 2005. Preliminary analysis on growth rates of farmed SBT through trade data and other related information. Commission for the conservation of southern bluefin tuna working paper CCSBT-ESC/0510/29.

Klaer, N., Sainsbury, K., Polacheck, T., 1996. A retrospective examination of southern bluefin tuna VPA and projection results, 1992–1995. Commission for the conservation of southern bluefin tuna working paper CCSBT/SC/96/31.

Kolody, D., Hartog, J., 2004. Examples of management procedure behaviour changes in response to operating model updating. Commission for the conservation of southern bluefin tuna working paper CCSBT-ESC/0409/25.

Polacheck, T., 2002. Experimental catches and the precautionary approach: The Southern Bluefin Tuna Dispute. Mar. Policy 26, 283–294.

Polacheck, T., Eveson, P., 2007. Analyses of tag return data from the CCSBT SRP tagging program – 2007. Commission for the conservation of southern bluefin tuna working paper CCSBT-ESC/0709/19.

Polacheck, T., Klaer, N.L., Millar, C., Preece, A.L., 1999. An initial evaluation of management strategies for the southern bluefin tuna fishery. ICES J. Mar. Sci. 56, 811–826.

Punt, A.E., Donovan, G., 2007. Developing management procedures that are robust to uncertainty: lessons from the Internatioanl Whaling Commission. ICES J. Mar. Sci. 64, 603–612.

Sainsbury, K., Polacheck, T., 1993. Development of a stock rebuilding strategy for Southern Bluefin Tuna. In: 12th SBT Trilateral Meeting Scientific Meeting, SBFWS/93/19, Hobart, Australia, 13–19 October.

Schnute, J.T., Richards, L., 2001. Use and abuse of fishery models. Can. J. Fish. Aquat. Sci. 58, 10–17.

Smith, A.D.M., Sainsbury, K.J., Stevens, R.A., 1999. Implementing effective fisheries-management systems—management stratey evaluation and the Australian partnership approach. ICES J. Mar. Sci. 56, 967–979.