Developing management procedures that are robust to uncertainty: lessons from the International Whaling Commission

André E. Punt and Greg P. Donovan

Punt, A. E. and Donovan, G. P. 2007. Developing management procedures that are robust to uncertainty: lessons from the International Whaling Commission. – ICES Journal of Marine Science, 64: 603–612.

Traditionally, fisheries management advice has been based on stock assessments that considered merely the "best" set of assumptions while uncertainty arising only from observation and process error was quantified, if considered at all. Unfortunately, uncertainty attributable to lack of understanding of the true underlying system and to ineffective implementation may dominate the sources of error that must be accounted for if management is to be successful. The management procedure approach is advocated as the appropriate way to develop management advice for renewable resources. This approach, pioneered by the International Whaling Commission (IWC) Scientific Committee, takes politically agreed management objectives and incorporates all scientific aspects of management including data collection and analysis, development of robust harvest control laws or effort regulations, and monitoring. A primary feature is that uncertainty (including that arising from sources conventionally ignored) is taken into account explicitly through population simulations for a variety of scenarios. The nature of the management procedures developed for commercial and aboriginal subsistence whaling and the processes by which they have been developed is highlighted. We also identify lessons that have been learned from two decades of IWC experience and suggest how these can be applied to other fishery situations.

Keywords: baleen whales, error, management procedure, monitoring, stock assessment, uncertainty.

Received 30 June 2006; accepted 2 February 2007; advance access publication 25 April 2007.

A. E. Punt: School of Aquatic and Fishery Sciences, Box 355020, University of Washington, Seattle, WA 98195-5020, USA and CSIRO Marine and Atmospheric Research, GPO Box 1538, Hobart, Tasmania 7001, Australia. G. P. Donovan: International Whaling Commission, The Red House, 135 Station Road, Impington, Cambridge CB4 9NP, UK. Correspondence to A. E. Punt: tel: +1 206 2216319; fax: +1 206 6857471; e-mail: aepunt@u.washington.edu

Introduction

Management of renewable marine resources to the satisfaction of all is difficult, even if managers are provided with exact information on the status and likely future trends of the resources being exploited, because it is necessary to balance various, often conflicting, "resource orientated" and "user-orientated" management objectives (Hall and Donovan, 2002). Additionally, different sources of uncertainty complicate attempts to manage the resources. The sources of uncertainty can be classified broadly (Francis and Shotton, 1997) as: (i) observation error (arising from sampling and monitoring of resources); (ii) model structure error (arising from lack of knowledge of population dynamics processes); (iii) process error (arising from seemingly unpredictable natural variability in population parameters affecting abundance, particularly recruitment); and (iv) implementation error (arising from problems in enforcement of measures taken).

Management decisions are often based on scientific assessments of stock status and the predicted consequences of alternative management actions. Although the need for scientific advice when managing renewable resources is now widely accepted, one of the earliest examples of this requirement enshrined in an international convention is the International Convention for the Regulation of Whaling (IWC, 2005a), signed in 1946, which states that amendments to its regulations shall be "based on scientific findings" (Donovan, 1992). National legislation may also include such exhortations (Anon., 1996).

Historically, stock assessments have been based on only the "best" set of assumptions, irrespective of how good they may be. Scientific uncertainty, if considered at all, quantified only the uncertainty arising from observation and process error. In addition, assessments usually attempted to relate then current stock status to what now are termed "biological reference points". These can be catch-, biomass-, or fishing-mortality-based. Although catch-based reference points are now generally regarded as insufficiently precautionary (Larkin, 1977), the use of biomass- and fishing-mortality-based reference points remains widespread. For example, the Pacific Fishery Management Council (PFMC) defines a stock as overfished (and in need of a formal rebuilding plan) when stock biomass drops below 25% of the average value in an unfished state, and overfishing to be occurring when the rate of fishing mortality (F) exceeds F_{MSY} (F at which the maximum sustainable yield is achieved).

Reference points alone are not sufficient to provide a scientific basis for making management decisions, so harvest control rules (HCRs; see Figure 1 for an example) that use reference points are commonly applied. Unfortunately, although the use of HCRs

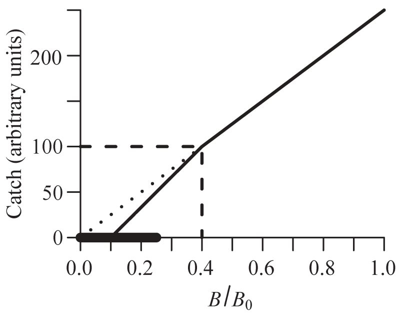

Figure 1. The harvest control rule (solid curve) used by the Pacific Fishery Management Council for stocks not designated as overfished (the solid bar indicates the range of stock sizes corresponding to being in an overfished state). Catch is reduced faster than linearly if a stock is assessed to be below the target biomass of 40% of the averaged unfished biomass ($0.4B_0$). Catch limits are not necessarily set to zero if a stock is depleted below $0.1B_0$; rather, a rebuilding plan is mandated to be developed for stocks depleted below $0.25B_0$.

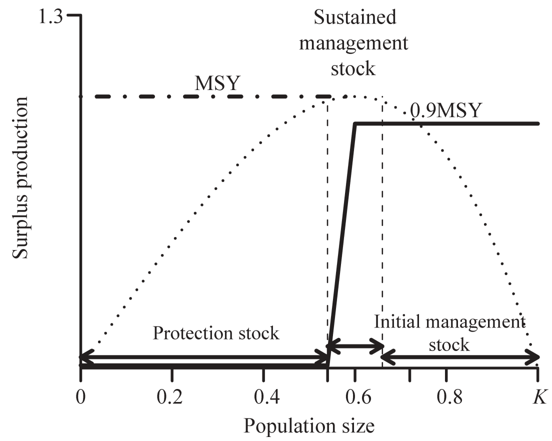

Figure 2. The catch control law for the NMP (solid lines) for the case in which MSYL is assumed to be 0.6*K*.

provides a means of specifying scientific management advice in a more objective manner, uncertainty traditionally has not been explicitly accounted for in their application. The HCR in Figure 1 is fully specified when augmented with specifications related to how to conduct assessments and rebuilding analyses (Anon., 2005). The parameters required include the stock biomass in an unfished state (B_0), the current stock biomass (B_{current}), and F_{MSY}. Estimates of B_0 and B_{current} are usually obtained by applying statistical catch-at-age analysis (e.g. Methot, 2006), and proxies for F_{MSY} are available, based on the relationship between spawning-biomass-per-recruit and F (Ralston, 2002). In 2005, this HCR could be applied to just 22 of the 80 species in the PFMC Groundfish Fishery Management Plan, i.e. insufficient information was available for all others. A further problem is that this HCR does not explicitly include a way to deal with uncertainty. Although scientists conducting assessments are encouraged to provide assessment scenarios that "bracket" uncertainty (Anon., 2006), there has been little consistency to date in how uncertainty has been bracketed for west coast groundfish, nor is there a formal way to use information from multiple alternative assessments when making management decisions.

Unfortunately, there is no guarantee that the management goals for a fishery will be satisfied even if (i) the HCR is fully-specified; (ii) data are available that, in principle, allow it to be applied; and (iii) some attempt is made to quantify uncertainty (Kirkwood, 1996). The only way to determine the effectiveness of a management process is to test it fully, using the simulation modelling approach pioneered by the Scientific Committee (SC) of the International Whaling Commission (IWC) when developing the Revised Management Procedure (RMP).

The IWC began using an HCR to provide formal management advice for commercial whaling in 1974, when it adopted what was termed the "New Management Procedure" (NMP; Figure 2). Problems with the NMP led to the development and adoption by the IWC of the "management procedure approach" (for a detailed account, see Hammond and Donovan, in press). Under this approach, management advice is based on a fully specified set of rules that have been tested in simulations of a wide variety of scenarios that specifically take uncertainty into account. The full procedure includes specifications for the data to be collected and how those data are to be used to provide management advice, in a manner that incorporates a feedback mechanism. Increasingly, there are examples of this approach being applied elsewhere (Kell *et al.*, 2006; Punt, 2006).

The problems with the NMP and the process of its replacement by the management procedure approach provide valuable lessons for other renewable resources. We review these problems (identifying the lessons learned), the process by which management procedures are evaluated by the SC, and how uncertainty is treated to ensure that the resulting management recommendations are sufficiently robust.

The New Management Procedure

The NMP was developed by Australian scientists as one response to calls for a moratorium on commercial whaling made at the United Nations conference on the environment and development, held in Stockholm in 1972 (Donovan, 1992, 1995). Although implicitly rather than explicitly expressed at the time, its two objectives were subsequently defined (IWC, 1981) as: (i) to ensure that the risks of extinction to individual stocks are not seriously increased by exploitation; and (ii) to maintain the status of whale stocks so as to make possible the highest continuing yield so far as the environment permits.

The conceptual basis for the NMP was the relationship between surplus production and population depletion (Allen, 1980). This relationship was assumed to be governed by a Pella–Tomlinson (1969) function, where the MSY level (MSYL) (the population size at which MSY is achieved) for whales was conventionally assumed to be 60% of the carrying capacity, *K* (Figure 2). Based on this, the NMP requires populations to be classified as either:

(i) Protection stocks–stocks depleted $<$0.9MSYL (the "protection level");

(ii) Sustained management stocks–stocks $>$0.9MSYL, but $<$1.2MSYL (a stock might also be classified in this category if it had been stable for "a considerable period" under a regime of constant catches);

(iii) Initial management stocks–stocks $>$1.2MSYL.

In addition, catch limits were constrained not to exceed the lower of 5% of the initial stock size and 0.9MSY.

The NMP was revolutionary for its time because it: (i) included a relatively high protection level (0.9MSYL, or 54% of the estimated unexploited level) at which catch limits were set to zero–although originally primarily justified for catch maximization rather than risk prevention (Butterworth and Best, 1994), many current management procedures still fail to include the concept of a protection level, or if they do, set the protection level considerably lower than 0.9MSYL; (ii) imposed a maximum catch limit that was less than the MSY estimate; (iii) aimed (even if implicitly) to leave stocks above (rather than at) MSYL; and (iv) appeared to take the decision-making process away from the politicians and to leave it with the scientists by "mechanizing" the provision of advice on catch limits.

Although seemingly well specified, the NMP definition is formally inconsistent because the constraint that catches cannot exceed 0.9MSY means that stocks will not be reduced to MSYL (de la Mare, 1986). In practice, however, this was one of the ways in which the NMP attempted to account for uncertainty by recognizing that determining whether a stock was at MSYL was difficult. A more serious problem was the lack of any formal (and agreed) basis to determine management units and to estimate the parameters needed to apply the NMP for each management unit, e.g. mortality, reproductive rates, and MSYL, with the required level of precision (Donovan, 1992). Although it is difficult to simulate the NMP decision-making process, de la Mare (1986) and IWC (1992a) evaluated the performance of one possible implementation using simulations and found it to behave poorly.

One consequence of these problems was that the SC was frequently unable to agree on catch limits when using the NMP. This was one of a complex set of reasons that the IWC introduced a moratorium on commercial whaling in 1982 to take effect in 1986 (e.g. Donovan, 1995). For example, in 1984, three species in four areas were considered in detail: sperm whales (*Physeter macrocephalus*) in the western North Pacific; Antarctic minke whales (*Balaenoptera bonaerensis*) in the southern hemisphere; common minke whales (*B. acutorostrata*) in the North Atlantic; and common minke whales in the North Pacific (IWC, 1985). However, the SC (i) was unable to provide estimates of initial and current population size for sperm whales in which it had confidence, and consequently did not provide advice on catch limits; (ii) was unable to agree on a classification for southern hemisphere minke whales (at that time, believed to have increased to above the initial carrying capacity in response to an expected surplus of krill resulting from the depletion of other large baleen whales, which made it impossible to be classified according to the NMP specifications), or on catch limits; (iii) could not classify either of the two stocks of minke whales in the northeastern Atlantic subject to commercial whaling (although ranges for catch limits were recommended despite dissenting views); and (iv) was able to classify only one of the two stocks of common minke whales in the North Pacific.

The Revised Management Procedure

Conceptual basis

The introduction of the moratorium led to the process of developing the RMP over a 6-y period (Hammond and Donovan, in press), which was finalized in 1994 with a written specification. The experience gained made it possible to develop a list of the steps needed to define and evaluate management procedures. Since then, these have been described extensively elsewhere (Kirkwood, 1996; Cooke, 1999; Sainsbury *et al.*, 2000; Donovan and Hammond, 2004; Kell *et al.*, 2006; Punt, 2006), but are summarized briefly here:

(i) qualitative specification and prioritization of the management objectives, as derived from legislation, legal decisions, and international standards and agreements;

(ii) quantification of the qualitative management objectives in the form of performance measures;

(iii) development and parameterization of a set of "operating models" that represent different plausible alternatives to the dynamics of the "true" resource and fishery being managed;

(iv) identification of candidate management procedures, including monitoring strategies;

(v) simulation of the future use of each candidate management procedure, involving for each time-step during the projection period: (a) generation of assessment data; (b) determination of the management action (i.e. assessment and application of some HCR); and (c) evaluation of the biological implications of the management action by removing the catch from the population as represented in the operating model;

(vi) summary of the performance of the candidate management procedures in terms of values for the performance measures; and

(vii) selection of the management procedure that best meets the specified objectives.

Most management procedures developed to date have focused on management using catch limits (e.g. Butterworth and Bergh, 1993; Geromont *et al.*, 1999; Punt and Smith, 1999), although some have been based on effort controls and forms of spatial management (e.g. Dichmont *et al.*, 2005). The major difference between these two types pertains to how management decisions are imposed (i.e. implementation error). For example, the performance of management systems based on catch limits can be affected by "highgrading" and "quota busting", while those based on effort regulations can be affected by "effort creep" and uncertainty in the relationship between fishing effort and fishing mortality. Key to the success of any evaluation of a management system based on a management procedure is the selection of the most important uncertainties to be reflected in the alternative operating models. These uncertainties should capture the major (but nevertheless plausible) factors that may affect the ability of each management procedure to satisfy the prescribed objectives.

Development

The development process of the RMP involved a series of workshops and discussion sessions during annual meetings of the SC. The initial focus was on developing a generic method for calculating safe catch limits that could be applied to any baleen whale population on its feeding grounds given perfect knowledge of stock structure (referred to as the "Catch Limit Algorithm", CLA). The CLA lies at the core of the RMP, which also includes

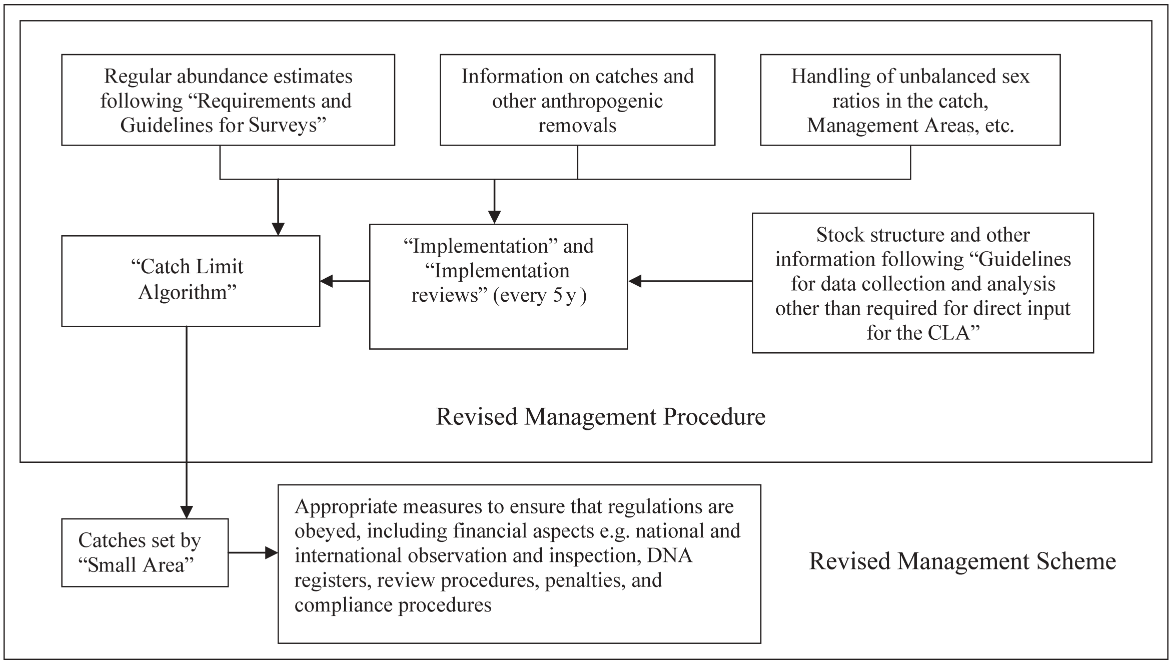

Figure 3. Schematic diagram of the management-procedure approach for commercial whaling (after Hammond and Donovan, in press) demonstrating the relationship between the CLA, the RMP, and the RMS.

rules on other scientific aspects of management, including multi-stock rules, and data and analysis requirements (Figure 3). Focus then shifted to developing ways of handling situations in which stock structure was uncertain. The approach chosen was to allow selection between a number of "variants". These variants are based on first setting catch limits for spatial/spatio-temporal strata that are small enough to ensure that whales of different stocks taken within each stratum will be taken in proportion to the abundances of those stocks (to avoid unintentionally taking whales out of proportion with their abundance in the area surveyed; IWC, 1999). Once these "Small Areas" have been determined, there are several ways (variants) in which catch limits may be combined to give a total limit for a wider area, depending, *inter alia*, on the available data for a particular species/region. Therefore, although the RMP is largely generic, it has case-specific aspects, because the choice between these multi-stock variants depends on further simulations (see the section "Implementation process" below).

The five CLAs developed by scientists from Australia, UK, South Africa, Iceland, and Japan initially differed quite markedly, e.g. in terms of the desired trade-off between risk and reward, whether population models were fitted to data on relative as well as absolute abundance indices, and whether relative abundance (catch per unit effort, cpue) data were used at all (Donovan, 1989; Hammond and Donovan, in press). However, the approaches converged over time. For example, it was rapidly agreed that there was little point in using cpue data because of disagreements regarding their reliability (IWC, 1989), and that future management procedures should be based on survey estimates of absolute abundance only.

This general approach of different teams developing candidate procedures has also been followed during the development phase of an Aboriginal Subsistence Whaling Management Procedure (AWMP), which eventually led to the adoption of "Strike Limit Algorithms" (SLAs) for the Bering–Chukchi–Beaufort Seas stock of bowhead whales (IWC, 2003a) and the eastern North Pacific stock of gray whales (IWC, 2005b). In the AWMP case, SLA development was pursued somewhat more cooperatively than had been the case for the RMP.

Discussion

Generic vs. case-specific approaches

An important difference between the AWMP and RMP is that the former followed a case-specific rather than a generic approach, the main reasons being that only a small number of aboriginal subsistence operations have been identified by the IWC and that it is unlikely that more will be accepted. In contrast, the number of potential commercial operations is substantially larger (in principle, all populations of all species of baleen whales could be subject to harvest), so that some standardization seemed desirable for reasons of efficiency. Furthermore, the management objectives for subsistence whaling differ from those for commercial whaling (e.g. rather than maximizing yield, the aim is to satisfy a pre-specified "need" in perpetuity, provided that certain conservation performance measures are met). Finally, aboriginal subsistence fisheries differ considerably in terms of the nature of the operations, the data available for assessment, and knowledge of stock structure. The adoption of a case-specific approach for data-rich operations (such as the bowhead and gray whale fisheries) accelerated the development of SLAs that met the management objectives. Had a generic management procedure robust to short time-series of abundance estimates and stock structure uncertainty been developed, the ability to satisfy subsistence needs for the data-rich aboriginal fisheries might well have been compromised.

The generic and the case-specific approaches have both advantages and disadvantages. For example, the amount of data available for individual species in finfish and invertebrate fisheries around the world ranges from extensive to almost nil, which would plead for case-specific approaches, because almost inevitably, generic procedures would have to be very conservative to achieve reasonable conservation performance in all cases. However, developing case-specific management procedures for all (or even the major) exploited stocks would be an immense undertaking that might be made lighter by first developing more generally applicable generic approaches. At the least, there may be value in developing generic management procedures that can be applied while case-specific procedures are being developed.

Implementation process

The RMP is considered final and has been adopted, but has not been used to set catch limits because the commercial whaling moratorium is still in place and no requests for advice on catch limits have been issued by the IWC. Before recommending that the RMP be applied to a species in a "region" (generally part of an ocean basin), simulation trials must be developed and run to capture the uncertainties deemed to be the most important for that stock complex/region. This process, referred to as an "Implementation" (in the IWC context, meaning that the SC notifies the Commission that it could produce information on catch limits if asked to do so), focuses primarily on uncertainties about stock structure, in particular temporal and spatial variation in the mixing of stocks in areas where whaling is to take place.

Many "Implementation Simulation Trials" (ISTs) may be required for specific cases if there are many alternative hypotheses related to stock structure, mixing, and other uncertainties such as the impact of bycatch of whales in fisheries. The process of designing, running, and interpreting ISTs can be onerous. For example, it took 12 y to complete the Implementation for western North Pacific minke whales (66 ISTs; IWC, 2003b). The reasons for this were that whaling occurs during migration (rather than on the feeding grounds, the situation for which the RMP had originally been designed) and that, because of the complex stock structure, new research conducted during the Implementation process led to a need to revise the hypotheses on which trials were based. The time required led to considerable frustration in the SC, and even to questions whether the RMP could be Implemented at all. Consequently, a rigorous set of requirements and guidelines was developed on how Implementations are to be conducted, so that the process could be completed within 2 y (Donovan and Hammond, 2004; IWC, 2005c). The guidelines also identify the information needed before an Implementation can commence (Figure 4). This information includes hypotheses about possible stock structure, specification of likely future removals (by both whaling and other anthropogenic causes), hypotheses about the size and spatial distribution of historical catches, and the abundance and migration data that will be used in the trials. The hypotheses identified during this "Pre-Implementation Assessment" should be sufficiently broad to prevent potential new information from leading to new hypotheses (but rather narrow or remove hypotheses). The guidelines also impose a temporal restriction on data that can be used. Data collected after a specified date can be used only when Implementations are reviewed.

The Implementation process focuses on developing ISTs to reflect plausible hypotheses and on assigning weights to each, based on perceived plausibility. The plausibility issue is one of the most difficult aspects to resolve and was a particular problem in the case of the western North Pacific common minke whale. Interpretation of the IST results (i.e. whether a particular variant performs adequately in terms of conservation objectives) is facilitated by a pre-specified set of rules. The first full application of the guidelines specified by IWC (2005c), from the Pre-Implementation Assessment through Implementation, is for western North Pacific Bryde's whales, and is scheduled for completion in 2007. North Atlantic fin whales have just completed the Pre-Implementation Assessment stage (IWC, 2007). The process used to develop and select management procedures is clearly far more formal at the IWC than that used to develop most fisheries management procedures. However, we believe that a process such as that adopted by the IWC would lead to a more rapid (and perhaps better documented) development and selection process.

Implementation Reviews

Although the simulation trials are based on 100-y projection periods (selected because of the slow dynamics of whales), the RMP and the AWMP include the requirement that "Implementation Reviews" be conducted every 5 y (IWC, 1999). North Atlantic common minke whales have already been through one review (IWC, 2004a). The aim of such reviews is to check that research conducted since the original Implementation does not reveal that hypotheses used in previous ISTs were not sufficiently broad to encompass reality or are no longer considered plausible. Basic changes to the RMP (as opposed to, for example, stock structure hypotheses in ISTs) are expected to arise rarely, and stringent conditions have been set on how proposed changes are reviewed (IWC, 1994, 2007).

Management procedures that have been implemented for finfish and invertebrate fisheries are generally revised irregularly (Punt, 2006), outside a formal structure and without formal requirements. An exception to this is the process established in South Africa (MCM, 2006).

Transparency

The taking of cetacean species for aboriginal and (particularly) commercial purposes remains a politically sensitive issue. Although the SC does not address issues regarding the politico-ethical acceptability of whaling (see Donovan, 1992), it is important that the process of developing and evaluating management procedures is wholly transparent. Transparency has been facilitated by having the chair of the group responsible for developing management procedures being considered both objective and independent (i.e. neither having expressed pro- or anti-whaling views nor being involved in developing one of the competing candidate management procedures). Transparency has also been achieved by having the computing manager at the IWC Secretariat responsible for coding (and testing) the operating models specified by the SC, validating the code for the CLA and SLAs, and conducting all calculations when recommendations on catch or strike limits are made. In contrast, such a clear separation of tasks is rarely the case for fisheries management. The development of a management procedure for southern bluefin tuna appears to be an exception (CCSBT, 2005a, 2005b).

Data collection and availability

The IWC has set standards for data collection. Specifically, a set of guidelines has been adopted for how surveys are to be conducted if the results are to be used in the RMP (IWC, 2005d). These guidelines also specify that catch limits will be reduced linearly to zero

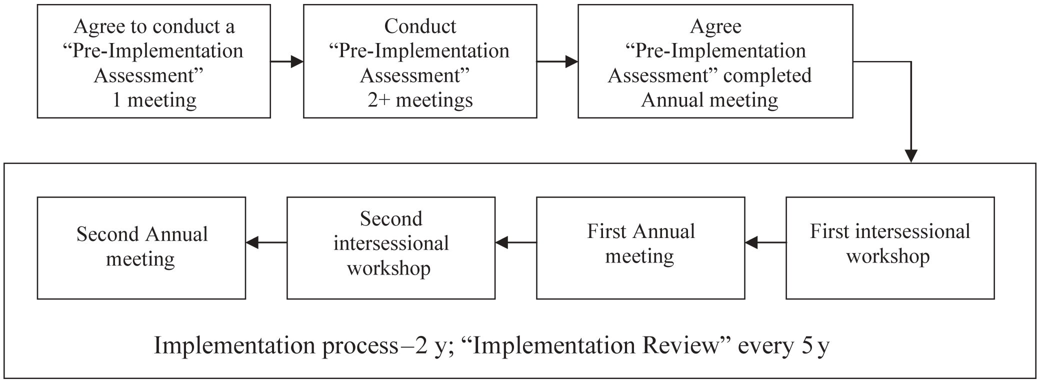

Figure 4. Conceptual overview of the procedure adopted by the IWC to facilitate implementation of the RMP (see IWC, 2005c).

over a period of 9–13 y after the last survey of an area (IWC, 1999). A data availability agreement (IWC, 2004a) specifies that all data used in the Implementation process must be freely available to members of the SC, although formal safeguards apply (e.g. with respect to publication rights). To date, no management procedure for a fish or invertebrate species has adopted such data collection standards or rules if the data required to apply the management procedure are not forthcoming.

Uncertainty factors considered

The uncertainties captured in simulation trials cover a number of factors. In general, the trials that capture uncertainty are divided into those considered most likely (the base-case or "evaluation" trials) and those considered less plausible, but for which performance should be adequate ("robustness" trials). The factors considered during the development of the CLA (Table 1) focused on those aspects considered most likely to affect performance. As expected, the factor having the greatest impact was the productivity of the resource. Therefore, the CLA was developed to perform adequately in the face of few data (only one estimate of absolute abundance when first applied) for stocks with perceived low productivity (an MSY that is only 1% of the number of mature animals at MSYL). Consequently, the resource is greatly underutilized when in fact productivity is higher. Other factors such as survey bias and temporal change in biological parameters also affected performance. Trials also included changes in carrying capacity and episodic events.

The uncertainties considered in developing SLAs for the Bering–Chukchi–Beaufort Seas stock of bowhead whales and the eastern North Pacific stock of gray whales (IWC, 2003b) were similar to those in Table 1, except that the impacts of demographic stochasticity, time-dependence in survey bias, and changes over time in the rate of natural mortality were also explored.

Stock structure

The multi-stock trials developed for specific Implementations have been much more complex than the single-stock trials used for developing the CLA and for the case-specific SLAs. For example, the trials developed for western North Pacific minke whales considered four major stock-structure assumptions based on an operating model that included 13 areas and a monthly time-step. The hypotheses underlying those assumptions were developed primarily from interpretations of genetic and length frequency data. The discussions of their relative plausibility were particularly fraught and led to the development of the guidelines referred to above (Donovan and Hammond, 2004; IWC, 2005e).

In fact, the major difference between the way management procedures have been evaluated for cetaceans and for fish and invertebrates is the focus on uncertainty in relation to stock structure. IWC (1992b, 1993, 1994) showed that a management procedure that performs adequately when stock structure is known can perform poorly when this is not the case. Specifically, conservation performance is poor when two stocks are assessed and managed as a unit, but catches are only taken (unintentionally) from one stock (Hall and Donovan, 2002). The lack of robustness to this type of uncertainty was the reason for the development of the multi-stock

Table 1. Factors and levels considered in the trials used to select the CLA (after IWC, 1992a). Underlined values denote those used to evaluate the candidate procedures, whereas the others were used to examine robustness.

Factor	Levels
Productivity	MSY rates of 1%; 4%; 7%; varying from 1% to 4%
Initial depletion, P_0	0.05, 0.3, 0.6, 0.99
Survey bias	0.5, 1, 1.5
Period of protection prior to management	Yes, No
Catches in error	No, half the true values
Age-at-maturity	7, 10
Episodic events	No, 50% of the population dies if an episodic event occurs at some specified frequency
MSYL	0.4*K*, 0.6*K*, 0.8*K*
Carrying capacity	Constant, increasing linearly over time, declining linearly over time
Survey intensity	Every fifth year, every tenth year

rules discussed above, and explains the need to conduct case-specific trials for any new Implementation of the RMP. An important new area of research being undertaken by the SC is the development of an individual-based simulation framework to investigate the performance of methods for analysing genetic data that may be used to inform stock-structure discussions in a management context (IWC, 2004b).

In reviewing how management procedures have been evaluated, Butterworth and Punt (1999) noted little progress in evaluating the impacts of spatial and stock structure on the performance of management procedures outside the IWC. More progress has been made on this front since then for shark populations off southern Australia (Punt *et al.*, 2005), groundfish species off southern Australia (Punt *et al.*, 2002), prawns off northern Australia (Dichmont *et al.*, 2005), and rock lobster off southern New Zealand (Bentley *et al.*, 2003). Plagányi *et al.* (2007) note that spatial structure will be considered when the management procedures for rock lobster, hake, sardine, and anchovy off South Africa are next revised.

Multispecies interactions

Interactions among species occur through bycatch, predation, or competition. Increasingly, such interactions are being explicitly included in operating models (Punt, 1993; Schweder *et al.*, 1998; Punt *et al.*, 2002; De Oliveira and Butterworth, 2004; Dichmont *et al.*, 2005; Plagányi and Butterworth, 2006). However, most focus has been on technical (i.e. bycatch) rather than biological (predation/competition) interactions. Adding biological interactions makes an operating model much more complicated, and requires more data to specify its parameters. The IWC has not attempted to build multispecies operating models, but rather has chosen to examine the impact of temporal changes in biological parameters that might be affected by biological interactions (such as carrying capacity, productivity, and natural mortality). Although it remains to be confirmed whether the implications of complicated biological interactions on the performance of management procedures can be captured adequately by varying parameter values of single-species operating models, the difficulties of predictive multispecies modelling approaches are well known, both in terms of assumptions and data requirements (e.g. IWC, 2004c).

Other features

The management procedures for commercial and aboriginal subsistence whaling share some common features, although their underlying objectives differ.

(i) The estimation methods that underlie the management procedures incorporate Bayesian aspects, by being based either on a conventional Bayesian assessment (Dereksdóttir and Magnússon, 2003, 2005), a Bayesian assessment that downweights the data to prevent large changes in stock status caused by noisy data (IWC, 1999), or on maximum-likelihood techniques with a penalty on deviations in parameter estimates from prior values (Johnston and Butterworth, 2004). The use of such methods is not related to philosophy, but rather to ensure that the parameters on which catch limits are based are set to conservative default values in the absence of informative data.

(ii) The RMP and Dereksdóttir–Magnússon SLAs account for parameter uncertainty by setting the catch limit based on a percentile of a posterior distribution <0.5. Consequently, increased uncertainty leads to lower catch limits.

(iii) The data used represent only a restricted subset of all data sources. For example, data on absolute abundance from surveys or (age-, sex-, and size-compositions of) catches, on relative abundance (e.g. from analyses of cpue data), and on fecundity rates exist for many whale stocks. However, only data on absolute abundance are used for setting catch limits, because other data sources can be subject to considerable uncertainty in interpretation. Hence, use of such data can lead to poorer performance than when they are ignored when the assumptions on which their use is predicated are wrong (e.g. that cpue is linearly proportional to abundance). Of course, ignoring additional data sources when they do provide useful information on status and trends may lead to some loss of yield (on average) for the same perceived risk to the resource. Although not used in the CLA, these data are used as part of the Implementation Review process to ensure that the parameter space of the uncertainty tested is still applicable.

(iv) All management procedures adopted involve fitting population dynamics models to data, because early work during RMP development suggested that model-based management procedures lead to less interannual variation in catch limits (IWC, 1992a). In contrast, many management procedures used in fisheries are based on changing catch limits in a direct relationship with the extent of changes in directly measurable quantities, such as cpue or survey estimates of abundance (De Oliveira and Butterworth, 2004; Breen *et al.*, 2006). One reason for this is that these more "empirical" approaches can be explained more easily to stakeholder groups. Although the need for simplicity is acknowledged by the IWC, stakeholder groups at the IWC generally have scientific advisors well versed in management procedures and their evaluation. In addition, particularly with respect to the AWMP development process, there has been a consistent effort to explain all stages of the process to the users and to consult with them on practical issues or design features (use of block quotas rather than annual quotas, carry-over provisions where catch limits are not reached in a particular year, etc.) (Donovan, 2006).

Given the recent adoption by the IWC of a formal structure to implement the RMP for a specific species and region and, in particular, the idea that there should be pre-specified standards of performance before the SC can recommend an Implementation, recognition of the value of research in reducing key uncertainties that lead to poor performance has increased. Specifically, IWC (2005c) allows for the use of a "research-conditional" option (Donovan and Hammond, 2004). Under strict conditions, this allows for temporary use of a variant that does not satisfy the pre-specified conservation performance standards, if this use is accompanied by an SC-approved research programme that should be able to determine whether or not the hypotheses on which performance is poor are indeed plausible. If the research fails to show within a specified time frame that the hypotheses are implausible, catch limits will be reduced to account for any catches above those that would have been set had a more conservative RMP variant been adopted that was robust to these hypotheses being true or not.

Available expertise and resources

The lack of people trained in the area of stock assessment and fisheries management is well recognized (Mace *et al.*, 2001). Unfortunately, the construction of alternative operating models that capture the key uncertainties requires considerable modelling experience. Consequently, efforts to develop appropriate management procedures are limited in many parts of the world more by an absence of suitable people than by financial resources. The lack of sufficient qualified personnel is one reason that the SC is reticent to conduct more than two Implementations at a time (IWC, 2005e), and why the Implementation process for North Atlantic fin whales will be delayed until after 2007 (both the western North Pacific Bryde's whale Implementation and a major bowhead whale Implementation Review are scheduled for completion in 2007).

Conclusions and lessons learned

The long development process of the management procedure approach within the SC of the IWC has been both painful and sometimes exhilarating. Nevertheless, the approach may well represent one of the most important recent advances in the management of renewable resources. The degree of rigour employed, particularly the explicit manner with which inevitable scientific uncertainty is dealt, is perhaps a by-product of the controversy surrounding whaling. In many respects, the framework is the first to have been specified for applying the precautionary principle to the management of renewable resources in a quantifiable manner (Garcia, 2000). Ironically though, the approach pioneered by the SC is not being used yet to manage commercial whaling, but is being applied to manage aboriginal subsistence fisheries as well as some commercial finfish and invertebrate fisheries (Hammond and Donovan, in press). In our opinion, the approach, despite the manifold difficulties that have been and will be encountered, represents the way forward to manage natural resources properly. The primary lessons learned may be summarized as follows.

(i) Management procedures should incorporate a degree of Bayesian philosophy; specifically the parameter values on which catch limits (or other regulations) are based should default to conservative values (e.g. a percentile <0.5 of their posterior distributions) until available data indicate otherwise. As a rule, if the point estimates of the parameters on which some management measure is based are identical for two stocks, the management measure taken should be less restrictive for the stock for which the information available is more precise. Including this feature in a management procedure encourages additional data collection.

(ii) Major gains may be made by having more than one team, each with a broad range of expertise, participating in the development process. Ensuring that the developers span a range of disciplines (biological, mathematical, and statistical) is likely to enhance the chances of the development of a management procedure that better satisfies the management objectives. Also, interaction among developers from different backgrounds, whether in a combative or collaborative environment, leads to innovative solutions. Close collaboration with representatives of all stakeholders and communication of the process to them will allow appropriate specification and quantification of the management objectives as well as the development of a procedure that is practical and carries a broad degree of support among users.

(iii) Management procedures should clearly specify their requirements for data and analysis, and include specific rules to handle situations in which the data needed to apply the management procedure are not available.

(iv) The time and effort required to develop sound management procedures should not be underestimated.

(v) Most non-IWC applications focus on the impact of observation and process error. However, structural error is likely to have a greater impact on performance, specifically with respect to spatial and stock structure.

(vi) A formal and well-specified process for evaluation is needed, particularly if the implementation is likely to be highly contentious. Of particular importance are a formal process for assigning weights to alternative simulation trials, a set of rules to determine when performance is considered adequate, and a "temporal science barrier"–a time limit after which new information will not be permitted to change how simulation trials are developed and performance is evaluated.

Acknowledgements

Funding for this work was provided to AEP by NMFS Grant NA04NMF4550330. We thank the many members of the RMP and AWMP groups within the IWC Scientific Committee who have contributed to what is, in our opinion, the major advance in management science in the past two decades. We dedicate this paper to two friends and colleagues who have recently died, Kjartan Magnússon and Geoff Kirkwood. Their contributions to the development of the RMP and AWMP processes were immense. Doug Butterworth (UCT) and the guest editor are thanked for useful comments on an earlier version of this manuscript.

References

Allen, K. R. 1980. A basis for calculating catch limits for previously unexploited stocks. Paper MP 20 presented to the Special Scientific Working Group on Management Procedures, Honolulu, HI, March 1980. 2 pp.

Anon. 1996. Magnuson–Stevens Fishery Conservation and Management Act. Section 301.

Anon. 2005. SSC Terms of Reference for Groundfish Rebuilding Analyses. Agenda Item F.3. March 2005. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 200, Portland, OR 97220.

Anon. 2006. Groundfish Stock Assessment and Review Process for 2007–2008. Agenda Item F.3.c. April 2006. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 200, Portland, OR 97220.

Bentley, N., Breen, P., and Starr, P. 2003. Design and Evaluation of a Revised Management Decision Rule for Red Rock Lobster (*Jasus edwardsii*) Fisheries in CRA 7 and CRA 8. New Zealand Fisheries Assessment Report 2003/30. 44 pp.

Breen, P., Kim, S., and Starr, P. 2006. Management Procedure Evaluations for Rock Lobsters in CRA 3 (Gisborne). New Zealand Fisheries Assessment Report 2005/61. 71 pp.

Butterworth, D. S., and Bergh, M. O. 1993. The development of a management procedure for the South African anchovy resource. *In* Risk Evaluation and Biological Reference Points for Fisheries Management. pp. 83–99. Ed. by S. J. Smith, J. J. Hunt, and D. Rivard. Canadian Special Publication of Fisheries and Aquatic Sciences, 120.

Butterworth, D. S., and Best, P. B. 1994. The origins of the choice of 54% of carrying capacity as the protection level for baleen whale stocks, and the implications thereof for management procedures. Reports of the International Whaling Commission, 44: 491–497.

Butterworth, D. S., and Punt, A. E. 1999. Experiences in the evaluation and implementation of management procedures. ICES Journal of Marine Science, 56: 985–998.

CCSBT. 2005a. Configuration of the Operating Model for the Reference Set. Attachment 5 of the Report of the Special Management Procedure Technical Meeting, 15–18 February 2005, Seattle, WA. Commission for the Conservation of Southern Bluefin Tuna.

CCSBT. 2005b. Draft CCSBT Management Procedure Specification. Attachment 6 of the Report of the Extended Scientific Committee for the Tenth Meeting of the Scientific Committee, 5–8 September 2005, Taipei, Taiwan. Commission for the Conservation of Southern Bluefin Tuna. 86 pp.

Cooke, J. G. 1999. Improvement of fishery-management advice through simulation testing of harvest algorithms. ICES Journal of Marine Science, 56: 797–810.

de la Mare, W. K. 1986. Simulation studies on management procedures. Reports of the International Whaling Commission, 36: 429–450.

De Oliveira, J. A. A., and Butterworth, D. S. 2004. Developing and refining a joint management procedure for the multispecies South African pelagic fishery. ICES Journal of Marine Science, 61: 1432–1442.

Dereksdóttir, E. H., and Magnússon, K. G. 2003. A strike limit algorithm based on adaptive Kalman filtering with an application to aboriginal whaling of bowhead whales. Journal of Cetacean Research and Management, 5: 29–37.

Dereksdóttir, E. H., and Magnússon, K. G. 2005. Application of a strike limit algorithm based on adaptive Kalman filtering to the eastern North Pacific stock of gray whales. Journal of Cetacean Research and Management, 7: 85–96.

Dichmont, C., Deng, A., Venables, W., Punt, A. E., Haddon, M., and Tattersall, K. 2005. A New Approach to Assessment in the NPF: Spatial Models in a Management Strategy Environment that includes uncertainty. Report of FRDC 2001/2002. CSIRO Marine Research, Cleveland, OH. 187 pp.

Donovan, G. P. 1989. The comprehensive assessment of whale stocks: the early years. Report of the International Whaling Commission (Special Issue 11). p[vi] + 210 pp.

Donovan, G. P. 1992. The International Whaling Commission: given its past, does it have a future? *In* Whales: Biology–Threats–Conservation. pp. 23–44. Ed. by J. J. Symoens. Royal Academy of Overseas Sciences, Brussels, Belgium.

Donovan, G. P. 1995. The International Whaling Commission and the revised management procedure. *In* Additional Essays on Whales and Man, pp. 4–10. Ed. by E. Hallenstvedt, and G. Blichfeldt, Alliance High North Alliance, Lofoten, Norway.

Donovan, G. P. 2006. Incorporation of user knowledge into the management of cetaceans: experience from the AWMP of the IWC. *In* Proceedings of the NAMMCO Conference on User Knowledge and Scientific Knowledge in Management Decision-Making, Reykjavík, Iceland, 4–7 January 2003, pp. 53–58. Ed. by G. K. Hovelsrud, and C. Winsnes. The North Atlantic Marine Mammal Commission.

Donovan, G. P., and Hammond, P. S. 2004. From pre-implementation to implementation: setting guidelines for practical application of the RMP. Paper SC/56/RMP6 presented to the IWC Scientific Committee, July 2004, Sorrento, Italy. 9 pp.

Francis, R., and Shotton, R. 1997. "Risk" in fisheries management. Canadian Journal of Fisheries and Aquatic Sciences, 54: 1699–1715.

Garcia, S. M. 2000. The precautionary approach to fisheries: progress review and main issues (1995–2000). *In* Current Fisheries Issues and the Food and Agriculture Organization of the United Nations, pp. 479–560. Ed. by. M. N. Nordquist and J. N. Moore. Center for Oceans Law and Policy, University of Virginia Law School.

Geromont, H. F., De Oliveira, J. A. A., Johnston, S. J., and Cunningham, C. L. 1999. Development and application of management procedures for fisheries in southern Africa. ICES Journal of Marine Science, 56: 952–966.

Hall, M. A., and Donovan, G. P. 2002. Environmentalists, fishermen, cetaceans and fish: is there a balance and can science find it? *In* Marine Mammals: Biology and Conservation, pp. 491–521. Ed. by P. G. Evans and J. A. Raga. Kluwer Academic/Plenum Publishers, New York.

Hammond, P. S., and Donovan, G. The RMP: managing whales in an uncertain world. Journal of Cetacean Research and Management, Special Issue 3 (in press).

IWC. 1981. Third Meeting of the Special Scientific Working Group on Management Procedures, Honolulu, HI, 20–26 March 1980. Reports of the International Whaling Commission, 31: 41–49.

IWC. 1985. Report of the Scientific Committee. Reports of the International Whaling Commission, 35: 31–58.

IWC. 1989. Report of the Scientific Committee. Reports of the International Whaling Commission, 39: 33–157.

IWC. 1992a. Report of the Third Comprehensive Assessment Workshop on Management Procedures, Oslo, 1–8 February 1990. Reports of the International Whaling Commission, 42: 271–304.

IWC. 1992b. Report of the Scientific Committee, Annex D. Report of the Sub-Committee on Management Procedures. Reports of the International Whaling Commission, 42: 87–136.

IWC. 1993. Report of the Scientific Committee, Annex I. Report of the Working Group on Implementation Trials. Reports of the International Whaling Commission, 43: 153–196.

IWC. 1994. Report of the Scientific Committee, Annex D. Report of the Sub-Committee on Management Procedures. Reports of the International Whaling Commission, 44: 74–92.

IWC. 1999. Report of the Scientific Committee, Annex N. The revised management procedure (RMP) for baleen whales. Journal of Cetacean Research and Management, 1 (Suppl.): 251–258.

IWC. 2003a. Chair's Report of the Fifty-Fourth Annual Meeting. Annual Report of the International Whaling Commission, 2002: 1–53.

IWC. 2003b. Report of the scientific committee, Annex E. Report of the standing working Group on the development of an aboriginal subsistence whaling management procedure (AWMP). Journal of Cetacean Research and Management, 5(Suppl.): 154–255.

IWC. 2004a. Report of the scientific committee. Journal of Cetacean Research and Management, 6(Suppl.): 1–60.

IWC. 2004b. Report of the modelling workshop on cetacean–fishery competition. Journal of Cetacean Research and Management, 6(Suppl.): 413–426.

IWC. 2004c. Report of the workshop to design simulation-based performance tests for evaluating methods used to infer population structure from genetic data, 21–24 January 2003, La Jolla, CA. Journal of Cetacean Research and Management, 6(Suppl.): 469–485.

IWC. 2005a. International Convention for the Regulation of Whaling. Annual Report of the International Whaling Commission, 2004: 157–168.

IWC. 2005b. Chair's Report of the Fifty-Sixth Annual Meeting. Annual Report of the International Whaling Commission, 2004: 1–58.

IWC. 2005c. Report of the scientific committee, Annex D. Report of the sub-committee on the revised management procedure. Appendix 2. Requirements and guidelines for implementation. Journal of Cetacean Research and Management, 7(Suppl.): 84–92.

IWC. 2005d. Report of the scientific committee, Annex D. Report of the sub-committee on the revised management procedure. Appendix 3. Requirements and guidelines for conducting surveys and analysing data within the revised management scheme. Journal of Cetacean Research and Management, 7(Suppl.): 92–101.

IWC. 2005e. Report of the scientific committee, Annex D. Report of the sub-committee on the revised management procedure. Journal of Cetacean Research and Management, 7(Suppl.): 77–113.

IWC. 2007. Report of the scientific committee. Journal of Cetacean Research and Management, 9 (Suppl.). 403 pp.

Johnston, S. J., and Butterworth, D. S. 2004. Final set of strike limit algorithms for the Eastern North Pacific stock of gray whales. Paper SC/56/AWMP6 presented to the IWC Scientific Committee, July 2004, Sorrento, Italy. 7 pp.

Kell, L. T., De Oliveira, J. A. A., Punt, A. E., McAllister, M. K., and Kuikka, S. 2006. Operational management procedures: an introduction to an evaluation framework for management systems. *In* The Knowledge Base for Fisheries Management, pp. 379–407. Ed. by L. Motos and D. Wilson. Elsevier, Amsterdam.

Kirkwood, G. P. 1996. Assessing the precautionary nature of fishery management strategies. FAO Fisheries Technical Memorandum, 350/2: 141–158.

Larkin, P. A. 1977. An epitaph for the concept of maximum sustainable yield. Transactions of the American Fisheries Society, 106: 1–11.

Mace, P. M., Bartoo, N. W., Hollowed, A. B., Kleiber, P., Methot, R. D., Murawski, S. A., Powers, J. E. *et al.* Marine Fisheries Stock Assessment Improvement Program. NOAA Technical Memorandum NMFS–F/SPO–56. 69 pp.

MCM. 2006. Procedures for deviating from OMP output for the recommendation of a TAC, and for initiating an OMP review. Unpublished MCM [Marine and Coastal Management, South Africa] Document. 9 pp.

Methot, R. 2006. User Manual for the Integrated Analysis Program Stock Synthesis 2 (SS2). Model Version 1.23a. Northwest Fisheries Science Center. 54 pp.

Pella, J. J., and Tomlinson, P. K. 1969. A generalised stock production model. Bulletin of the Inter-American Tropical Tuna Commission, 13: 421–496.

Plagányi, É. E., and Butterworth, D. S. 2006. A spatial multi-species operating model (SMOM) of krill-predator interactions in small-scale management units in the scotia sea. CCAMLR WG–EMM document WG/EMM/06/12. 27 pp.

Plagányi, É. E., Rademeyer, R. A., Butterworth, D. S., Cunningham, C. L., and Johnston, S. J. 2007. Making management procedures operational–innovations implemented in South Africa. ICES Journal of Marine Science, 64: 626–632.

Punt, A. E. 1993. The comparative performance of production-model and *ad hoc* tuned VPA based feedback-control management procedures for the stock of Cape hake off the west coast of South Africa. *In* Risk Evaluation and Biological Reference Points for Fisheries Management, pp. 283–299. Ed. by S. J. Smith, J. J. Hunt, and D. Rivard. Canadian Special Publication of Fisheries and Aquatic Sciences, 120.

Punt, A. E. 2006. The FAO precautionary approach after almost 10 years: have we progressed towards implementing simulation-tested feedback-control management systems for fisheries management. Natural Resource Modeling, 19: 441–464.

Punt, A. E., Pribac, F., Taylor, B. L., and Walker, T. I. 2005. Harvest strategy evaluation for school and gummy shark. Journal of Northwest Atlantic Fisheries Science, 35: 387–406.

Punt, A. E., and Smith, A. D. M. 1999. Harvest strategy evaluation for the eastern stock of gemfish (*Rexea solandri*). ICES Journal of Marine Science, 56: 860–875.

Punt, A. E., Smith, A. D. M., and Cui, G. 2002. Evaluation of management tools for Australia's South East Fishery. 1. Modelling the South East Fishery taking account of technical interactions. Marine and Freshwater Research, 53: 615–629.

Ralston, S. 2002. West coast groundfish harvest policy. North American Journal of Fisheries Management, 22: 249–250.

Sainsbury, K. J., Punt, A. E., and Smith, A. D. M. 2000. Design of operational management strategies for achieving fishery ecosystem objectives. ICES Journal of Marine Science, 57: 731–741.

Schweder, T., Hagen, G. S., and Hatlebaak, E. 1998. On the effect on cod and herring fisheries of retuning the revised management procedure for minke whaling in the greater Barents Sea. Fisheries Research, 37: 77–95.

doi:10.1093/icesjms/fsm035