ICES Journal of Marine Science, 56: 985–998. 1999
Article No. jmsc.1999.0532, available online at http://www.idealibrary.com on IDEAL®

Experiences in the evaluation and implementation of management procedures

D. S. Butterworth, and A. E. Punt

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.

A synthesis of the authors' experience with the evaluation and implementation of management procedures in Australasia, southern Africa, and the International Whaling Commission (IWC) is presented. The development of operating models for testing such procedures for the fisheries in question over their respective ranges of uncertainty, together with the statistics used to assess procedure performances, are considered first, and then suggestions are made that increasing experience is making it possible to develop a minimal set of key factors to include in such robustness trials. Some general lessons are drawn, primarily from the IWC's process of developing its Revised Management Procedure. Further implementation issues discussed are: candidate procedure selection in principle and practice, the extent of robustness testing desirable, the link to the evaluation of research priorities, and the reception accorded the management procedure approach by industry and decision-makers. Management procedures are seen to have potential benefits over the annual assessment basis for determination of Total Allowable Catch, but key problem areas that remain concern the definition of risk and the relative weights to be accorded to the various scenarios (of differing plausibilities) considered in robustness tests.

Key words: management procedure, operating model, simulation.

Received 31 October 1998; accepted 26 May 1999.

D. S. Butterworth: MARAM (Marine Resource Assessment Management Group), Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch 7701, South Africa. Tel: +27 21 6502343; fax: +27 21 6502334; e-mail: dll@maths.uct.ac.za. A. E. Punt: Division of Marine Research, CSIRO Marine Laboratories, GPO Box 1538, Hobart, Tasmania 7001, Australia. Tel: +61 3 62325492; fax: +61 3 62325000; e-mail: andre.punt@marine.csiro.au

Introduction

Butterworth *et al.* (1997) characterize a management procedure as a set of rules which utilize pre-specified data to provide recommendations for management actions, where the performance of the rules has been evaluated by simulation (see Cochrane *et al.*, 1998, for further elaboration). There are two main reasons for evaluating alternative candidate management procedures by simulation: (a) their relative performances can be assessed, and (b) their anticipated performance relative to specified management objectives can be determined.

The ability to make these comparisons (particularly the second) depends on the extent to which the set of operating models used to represent the true underlying situation in the fishery captures the full range of uncertainty pertinent to that fishery. The first part of this paper therefore explores the bases for the evaluation of management procedures in the development of which the authors have been closely involved. For the most part these relate to Southern Hemisphere fisheries, primarily in Australasia and southern Africa, but they include also examples from the International Whaling Commission (IWC). In some cases, references to other similar developments are incorporated in the interests of completeness. This exploration considers primarily the features included in the operating models, but it also remarks on the choice of performance statistics used.

The second part of the paper describes the authors' experiences of attempts to implement management procedure approaches in these fisheries. Some general lessons about the development and implementation process are drawn, based primarily upon the IWC's Revised Management Procedure (RMP) development process. Other topics covered are candidate procedure selection

Table 1. Species/stocks for which the authors are aware that an evaluation of management procedures has been conducted.

Stock/species	References
Cape hake off South Africa	Punt (1992, 1993); Punt *et al.* (1995); Punt and Butterworth (1995); Geromont and Butterworth (1998a, b)
Cape hake off Namibia	Butterworth and Geromont (1997)
South African anchovy	Bergh and Butterworth (1987); Butterworth and Bergh (1993); Butterworth *et al.* (1993); Cochrane *et al.* (1998)
South African sardine	Cochrane *et al.* (1998); De Oliveira *et al.* (1998a)
South African rock lobster (*Jasus lalandii*)	Johnston and Butterworth (1997); Johnston (1998); De Oliveira *et al.* (1998b)
Namibian Cape fur seals	Butterworth *et al.* (1998)
Namibian orange roughy (*Hoplostethus atlanticus*)	Branch (1998)
New Zealand rock lobster (*Jasus edwardsii*)	Starr *et al.* (1997)
New Zealand sea lion	Maunder *et al.* (in press)
Minke whales	Donovan (1989); Kirkwood (1997); Magnusson (1992); Smith *et al.* (in press); IWC (1992b, 1993, 1994b, 1997a)
Marine mammals	Wade (1998)
North Sea cod	Pelletier and Laurec (1992)
Australian gemfish	Punt and Smith (1999)
Southern bluefin tuna	N. L. Klaer (pers. comm.)

in principle and practice, the extent of robustness testing desirable, the link to evaluation of research priorities, and the reception accorded the approach by industry and decision-makers.

Bases for the evaluation of management procedures

Table 1 lists the stocks for which the authors are aware that management procedures have been evaluated (though not always, as yet, implemented). Management procedures must be fully operational and should therefore have as input data only information that is actually collected for the fishery in question. Therefore, studies in which it is assumed that a stock assessment has been conducted, and hence that an estimate of current biomass (and its age structure) is available with a given CV (e.g. Hollowed and Megrey, 1993; Lowe and Thompson, 1993; Baldursson *et al.*, 1996; Christensen, 1997) are excluded from Table 1. This is because previous studies (e.g. Punt, 1993) have shown that the results from stock assessments can be biased and that the extent of bias will change over time. In addition to the studies that are specific to particular species, several general examinations of the performances of management procedures have been conducted (e.g. Rosenberg and Restrepo, 1993; Horwood, 1994; Punt, 1995, 1997). Table 1 does not include studies aimed primarily at assessing the value of research for management (e.g. Cochrane and Starfield, 1992; McDonald and Smith, 1997; McDonald *et al.*, 1997) and those designed to evaluate adaptive management experiments (e.g. Collie and Walters, 1991; Mapstone *et al.*, 1996; Sainsbury *et al.*, 1997) even though conceptually these differ little from the studies evaluating alternative management procedures.

Operating model complexity

Francis and Shotton (1997) identify six sources of error and uncertainty that can be incorporated into a risk assessment. Two of these (error structure uncertainty and estimation uncertainty) relate to fitting models to data. The other four (process uncertainty, observation uncertainty, implementation uncertainty and model uncertainty) relate directly to the specification of operating models for the evaluation of management procedures. We review the first three of these briefly, and the remainder of this section then considers model uncertainty.

Process uncertainty (or process error) relates to "natural" (i.e. uncontrollable) variation in the underlying demographic rates and processes (Francis and Shotton, 1997). The most common example of process error is variation in recruitment about the value expected from the stock-recruitment relationship. However, some studies have also considered variation and trends in natural mortality, mass-at-age (Horwood, 1994; Baldursson *et al.*, 1996; Johnston and Butterworth, 1997; Johnston, 1998) and selectivity-at-age (Punt, 1993, 1995, 1997).

Observation error relates primarily to the error that arises when populations are sampled. Examples of observation error include variation about survey results, errors in measuring landed catches, and variation between the trends in standardized catch rates and those in exploitable biomass caused by changes in fishing efficiency or environmentally induced fluctuations in catchability. Hilborn (1997) includes slightly different forms of error under this heading, relating to uncertainty whether surveys provide absolute or relative indices of abundance and whether catch rate is related linearly to abundance. The former is the major source of uncertainty for Cape hake *Merluccius capensis* and *M. paradoxus* off Namibia (Butterworth and Geromont, 1997) and South African sardine *Sardinops sagax* populations (De Oliveira *et al.*, 1998a, b), while the latter will always be a major source of uncertainty unless estimates of relative abundance are available from fishery-independent surveys. In principle, estimates of observation error variance are available for some types of data (e.g. survey estimates, age-length keys). However, allowance needs to be made in the operating model that estimates of sampling variability (particularly for surveys) may be negatively biased because factors such as interannual fluctuations in the fraction of the population in the area surveyed are not incorporated in sampling variation. Cochrane *et al.* (1998) and De Oliveira *et al.* (1998a) include such "additional variance" in their evaluation of management procedures for sardine.

As part of its evaluation of management procedures for baleen whale stocks, the IWC has considered robustness to, *inter alia*, the magnitude of and trends in possible survey bias, the frequency of surveys, and the precision of future surveys (e.g. IWC, 1992a, 1993).

Implementation uncertainty relates to "the extent to which management policies will be successfully implemented" (Rosenberg and Brault, 1993). Factors likely to lead to implementation uncertainty include black-market landings, misreporting of the species composition of the catch, high-grading, and discarding. To date, implementation uncertainty has largely been ignored in the evaluation of management procedures (though see Powers and Restrepo, 1998).

The operating model designed for the North Sea (Horwood, 1994) includes discarding as well as uncertainty in the relationship between the total allowable catch (TAC) recommended by the management procedure and the actual catch. Johnston and Butterworth (1997) include the impact of catches by poachers and discard mortality in their evaluation of a management procedure for the rock lobster (*Jasus lalandii*) resource off South Africa's west coast. Implementation uncertainty is implicitly acknowledged in several studies by placing lower bounds on the catch. This reflects, *inter alia*, the reality that even if a management procedure indicates a zero TAC, this is unlikely to be enforced for socio-economic reasons (e.g. Cochrane *et al.*, 1998).

Model uncertainty relates to lack of knowledge concerning the correct form of the model to describe the resource (and fishery) dynamics. The studies considered in Table 1 vary considerably in terms of the extent to which different structural models are considered. In most cases, a single (usually age-structured) model is used to represent the population. It is assumed that the population is closed with respect to immigration and emigration, and unaffected by multispecies interactions and spatial substructure.

The most common form of "model uncertainty" relates to making allowance for uncertainty about the values for the model parameters within a single structural model. The IWC has, as part of its evaluation of candidate management procedures for baleen whale stocks, considered this form of uncertainty to the greatest extent. For example, Smith *et al.* (in press) examine the robustness of the IWC's RMP to uncertainties (in combination) about historical catches, different levels of productivity (and the population size at which this occurs), cycles in carrying capacity and in productivity (to mimic trends in the environment), and different forms for the stock-recruitment relationship. The next most common form of this uncertainty relates to the status of the resource when the management procedure is first applied. For example, Johnston and Butterworth (1997) and Starr *et al.* (1997) examine the implications if the actual resource abundance is half the current "best estimate", whereas Punt (1995, 1997) considers the robustness of different management procedures to a variety of initial stock sizes.

Many of the studies in Table 1 examine robustness to different forms for the stock-recruitment relationship, including relationships that exhibit depensation (Punt and Smith, 1999) and correlations among the deviations about the relationship. Such correlations are to be expected if the recruitment is 'driven' by some (autocorrelated) environmental factors. Butterworth and Geromont (1997), Geromont and Butterworth (1998a) and Punt and Smith (1999) all examine the consequences of non-stationarity in the stock-recruitment relationships for Cape hake and gemfish (*Rexea solandri*) respectively by considering the impact of a "regime" shift. Johnston and Butterworth (1997) and Butterworth *et al.* (1998) consider the implications of "episodic events" which lead to mass mortalities of rock lobster and Cape fur seals (*Arctocephalus pusillus pusillus*), respectively; such events were also considered in trials of the IWC's RMP (1992b).

Most evaluations of management procedures assume that age-specific selectivity is (on average) invariant over time. However, the performance of management procedures can be sensitive to the impact of changes in selectivity (Punt, 1992; Punt and Smith, 1999). Two scenarios can be envisaged in terms of the time-dependence of selectivity: one in which the reasons for any such changes are known (e.g. if the change is a consequence of an alteration in the gear-types used in the fishery), and the other in which those reasons are unknown. An example of the latter is the eastern stock of gemfish for which selectivity appears to have changed so that younger fish are increasingly being selected (possibly as a consequence of a reduction in abundance leading to a density-dependent response in the form of a reduction in the age-at-first-maturity) (Punt and Smith, 1999).

There are two main types of multispecies interaction that can be considered when constructing operating models. Technical interactions relate to the fact that changes to the management arrangements for one species will impact the catch of other species. Cochrane *et al.* (1998) and De Oliveira *et al.* (1998a) acknowledge this problem for the South African pelagic fishery and consequently evaluate 'joint' management procedures for the anchovy (*Engraulis capensis*) and sardine resources in which the expected by-catch of juvenile sardine is related to the size of the anchovy TAC, and this in turn affects the TAC for adult sardine. Maunder *et al.* (in press) examine the trade-off between minimizing the by-catch of New Zealand sea lions (*Phocarctos hookeri*), to permit this population to recover further, and the consequent loss in commercial catch that results from occasionally having to close the squid (*Nototodarus sloanii*) fishery that takes most of the by-catch of these sea lions.

Consideration of biological interactions (predation, competition, etc.) is far more complicated, given the poor understanding of how these interactions operate. Hence it is not surprising that, to date, little effort has been directed at including such interactions in operating models except by making allowance for age-specific natural mortality. A notable exception to this is the operating model developed by Horwood (1994) for North Sea cod (*Gadus morhua*), haddock (*Melanogrammus aeglefinus*), and whiting (*Merlangius merlangus*). Punt (1992, 1993) examined the performances of the management procedures for Cape hake given cannibalism among the two Cape hake species, while Punt and Butterworth (1995) evaluated the sensitivity of the performance of the then current hake management procedure to different levels of cull and hence population trajectories for the Cape fur seal population using an operating model that incorporates both hake species, the Cape fur seal, and "other predatory fish".

To the authors' knowledge, the only serious attempt to date to consider spatial/stock structure in operating models has been made by the IWC. The operating models developed for Southern Hemisphere, North Atlantic, and North Pacific minke whales (*Balaenoptera acutorostrata*) (IWC, 1993, 1994a, 1997a) consider the dynamics of multiple stocks/substocks, make allowance for errors in the (assumed) boundaries between putative stocks, and incorporate time-dependent mixing on the feeding grounds. The results of trials based on these more complex operating models suggest that the performance of management procedures that perform adequately if stock structure is certain can deteriorate markedly if this is not the case (IWC, 1992b, 1993, 1994b).

Work is currently in progress in Australia to develop operating models for several further stocks. All of these operating models will incorporate spatial structure to some extent. For example, although it is known that adult rock lobsters (*Jasus edwardsii*) off Tasmania do not move substantial distances (Pearn, 1994), recent assessments (Punt and Kennedy, 1997) suggest substantial spatial correlations in recruitment. The fishery for coral trout (*Plectropomas leopardus*) on the Great Barrier Reef provides an even more extreme example of multi-stock management in that, even though adults appear sedentary and permanent movement among reefs is rare (Davies, 1995), there is substantial larval drift among reefs. An appropriate operating model for this problem must account for its metapopulation structure. The operating model currently being developed for southern bluefin tuna (*Thunnus maccoyii*; N. L. Klaer, pers. comm.) includes spatial structure in order to develop models that relate commercial catch rates to abundance more appropriately.

A problem particular to the evaluation of management procedures occurs when the data for assessment purposes are contradictory (in terms of the models/assumptions used). For example, inferences based on catch rates for Cape hake imply a substantially more productive population than inferences that take account of the age structure of the catch (Punt, 1994). It is clearly inappropriate simply to base the evaluation on only one of these data sets (or perhaps an assessment based on some "average" of the two). Instead, simulations should be conducted assuming that one or the other of the data sets is correct (e.g. Punt, 1992, 1993), and also attempts should be made to explore a wider range of models which might be able to resolve the conflict. Although it is desirable that a management procedure be robust to such uncertainties, it is unclear how results based on conflicting data sets should be weighted.

Few of the operating models considered to date explicitly incorporate socio-economic factors. At present, economic implications are summarized by the average catch, the variation in catches, and (on occasion) the future-discounted catch. There are many economic models (e.g. O'Boyle *et al.*, 1991) that could be used in conjunction with biological models (whether similar social models are available is unclear). However, the closest attempts at linking biological and economic models (e.g. Baldursson *et al.*, 1996; Christensen, 1997) have not incorporated operational management procedures, but have assumed instead that estimates of biomass are available with known bias (taken to be zero) and precision.

Performance statistics

The statistics used to evaluate the performance of alternative candidate management procedures should be chosen so that they are easy for managers and stakeholders to interpret (Francis and Shotton, 1997). For example, although quantities such as the standard deviation or coefficient of variation of catch limits have been used in some studies (e.g. Hall *et al.*, 1988; Quinn *et al.*, 1990; Pelletier and Laurec, 1992; Overholtz, 1993), it is rare for non-specialists (e.g. managers and members of industry) to find these statistics particularly meaningful (Francis and Shotton, 1997). Consequently, for example, variation in catch is usually measured as the average of the absolute change in catch expressed as a proportion of average catch.

Performance statistics usually fall into three categories, with improved performance in one area generally leading to worse performance in at least one of the other two. These three relate to the general objectives of maximizing catch, minimizing risk to the resource, and maximizing industrial stability. Statistics related to catch maximization are the most straightforward to define. However, an exception is provided by the current evaluation of the Scientific Committee of the IWC of management procedures for aboriginal subsistence whaling, which do not aim to maximize catches (or profits), but rather attempt to satisfy ‘‘need’’. Therefore a subset of the performance statistics for evaluating candidate management procedures involves comparing the catch limits (strictly strike limits, as animals struck but lost are assumed to die) set with the catch level corresponding to aboriginal ‘‘need’’ (IWC, 1997b, c).

High interannual variability in catches is undesirable for the fishing industry because it leads to marketing problems and social dislocation. The interannual variation in catches can be described using a wide variety of performance measures (e.g. IWC, 1992b). However, many of these measures are difficult to interpret, so most studies have been based on the average absolute change in catch from one year to the next scaled by the total catch. This statistic captures short-term fluctuation in catch, but may not capture long-term trends in catch adequately. Nevertheless, the IWC (1991) reports that most of the different performance statistics it considered were highly correlated for the scenarios examined. Catch variability is also sometimes conveniently assessed by visual examination of time trajectories of catch from a subset of the simulations.

One of the performance statistics considered by Starr *et al.* (1997) was the probability of a reduction in the TAC. In New Zealand, ‘‘risk’’ to the fishery is also quantified by the probability that the quota will exceed a pre-specified fraction (67%/80%) of the exploitable biomass, so that the industry may be unable to take the quota (Francis, 1992; Hilborn, 1997).

The probability that the biomass will drop below some pre-specified threshold is the most common performance statistic used to quantify the ‘‘risk’’ to the resource. This statistic is designed to reflect concern about the possible consequences of low biomass. Such consequences include stock collapse due to recruitment failure, species replacement or depensatory processes (Hilborn, 1997), and impacts on the rest of the ecosystem (Corten, 1993). The threshold used most commonly when assessing ‘‘risk’’ is 20% of the average virgin level, B_0 (e.g. Beddington and Cooke, 1983; Francis, 1992; Punt, 1995, 1997). However, other levels have been considered (10% B_0 – Bergh and Butterworth, 1987; 25% B_0 – Hall *et al.*, 1988; Quinn *et al.*, 1990; 54% B_0 for baleen whales, although Butterworth and Best (1994) question whether this constitutes a misinterpretation of the reasons underlying the original choice of this figure). Variants on this theme include considering the current biomass (Punt and Walker, 1998) or the minimum historical biomass (Sigler and Fujioka, 1993).

Care needs to be taken when selecting this threshold to account for the biological characteristics of the species under consideration. For example, for short-lived species, the probability of dropping below a particular threshold may be high even in the absence of a fishery. This suggests, for example, that the larger the extent of natural variation in recruitment, the lower the level below which the stock should be considered to be in an undesirable state. Butterworth *et al.* (1993, 1997) argue that populations that fluctuate considerably naturally are more likely to be resilient to depletion to a specified level.

Hilborn (1997) criticizes the use of 20% of B_0 as a performance statistic because (a) 20% (or any other level) is arbitrary, (b) some stocks that have been depleted to very low levels have nevertheless recovered, and (c) stocks below 20% of the virgin biomass may be capable of producing high sustainable yields. He notes, furthermore, that maximum sustainable yield (MSY) may even occur below 20% of the virgin level. Myers *et al.* (1994) argue against the use of the 20% B_0 threshold because (a) it can perform very poorly empirically, (b) estimation of B_0 can be poor, and (c) 20% of B_0 is not a universal threshold. They instead advocate the use of the biomass at which recruitment drops to half its maximum level as the basis for a risk performance measure. To date, however, this suggestion has not been included in evaluations of management procedures in southern Africa and Australasia.

An alternative to the approach of specifying a threshold or a target biomass is to display the whole biomass distribution. For example, Punt (1995, 1997) quantifies the performance of a range of management procedures by displaying the 5%ile, the median, and the 95%ile of the distribution of the biomass in the final year of the projection, along with the 5%ile, 25%ile, and median of the distribution of the lowest biomass during the projection. This approach ''lays bare'' the whole distribution. Although more difficult to interpret, the approach does allow a more explicit consideration of the trade-off between population size and catch because it does not simply summarize the biomass distribution by a single number.

Hilborn (1997) argues that there is really no need for performance statistics to quantify ''risk'', and suggests instead that the mechanisms that may lead to ''undesirable states'' should be specifically included in operating models. Therefore, the implications of undesired depletion would be reflected in the sizes of future catches. While this approach may seem reasonable in principle, it is not clear whether the processes that act at low population size are sufficiently well known that it would be reasonable to include them in operating models. To date, very few studies have even allowed for the possibility of depensation at low population size. Furthermore, Butterworth *et al.* (1997) argue that similar average catch and short-term catch variability levels can result from very different long-term catch trends, and hence that dropping a ''risk'' statistic necessitates its replacement by a surrogate in the form of long-term catch variability.

Implementation in principle and practice

Although management-procedure-based management initiatives are underway for a number of fisheries, particularly in the Southern Hemisphere (see Table 1), their application in practice is perhaps most advanced in southern Africa. Inevitably, therefore, certain sections below show a rather heavy focus towards experiences in that region.

General lessons

The exercise carried out by the IWC's Scientific Committee to develop a RMP for commercial (baleen) whaling differed in some important aspects from the other initiatives listed in Table 1. It was probably the most thorough and detailed of all these processes. Furthermore, it sought a general procedure, rather than one separately tailored to each particular (potential) fishery, although case-specific options were introduced subsequently to deal with multi-stock considerations (IWC, 1994c). Finally, five separate groups ''competed'' (through working interactively and to some extent cooperatively) to produce the best performing procedure, and attempted rather different approaches to achieve this (IWC, 1992a; Kirkwood, 1997). This allows several more general insights about management procedures to be drawn.

First, three sets of developers based their approaches on fitting population models to the data, while the other two (Magnusson-Stefansson and Sakuramoto-Tanaka) used more empirical approaches based primarily on trends in abundance indices. The model-based approaches outperformed (though not to that large an extent) the empirical, particularly in regard to variability levels (IWC, 1992a). Essentially, it seems that the variance reduction provided by imposing the population model structure more than offset any bias introduced by possible mis-specification of this model.

A second general lesson provided by the IWC exercise was that basing the procedure on more complex population models offers few gains, if any, over simpler approaches. Thus, the Cooke procedure lost nothing in performance compared to that of the Punt-Butterworth approach, although it used a much simpler population model, despite the associated potential bias introduced (IWC, 1992a). The reason for this was that the observation errors in typical input data are too large for model estimates to be sensitive to such features. A further example of this lesson is provided by Punt (1993), who showed that, for the South African hake resource, production-model-based procedures which ignored catch-at-age data outperformed *ad hoc* tuned VPA-based procedures which used these data. The reason was that the latter procedures tended to follow noise rather than signal, so increasing interannual variation in TACs without any compensating gains for other performance statistics.

A corollary to this result relates to the frequent request by managers that the algorithm underlying a management procedure should contain ''design features''. The most common such feature is that the management procedure includes some biological reference point (e.g. fishing mortality not greater than F_{MSY}), or that they be ''precautionary''. Unfortunately, as noted by Kirkwood and Smith (1996), even though a management procedure contains such a ''design feature'', it does not follow that it will behave as intended thereby in practice. Generally, this is because of the estimation imprecision that results from the stochastic nature of input data. For example, de la Mare's RMP candidate was designed so that the probability of the resource dropping below a 54% B_0 ''protection level'' did not exceed a given value. However, Cooke's approach, which did not include this design feature, outperformed de la Mare's in terms of satisfying this criterion, and without sacrificing comparative performance on some other performance statistics (IWC, 1992a).

A final lesson from the IWC exercise seems to the authors to be that the development of ''case-specific'' rather than ''generic'' procedures is to be preferred. Though valuable insights were provided by the generic manner in which the RMP development process was pursued, case-specific simulation trials eventually had to be performed for individual regional species-stock complexes anyway (such as the various minke whale populations – IWC, 1993, 1994a, 1997a). With the wisdom of hindsight, prior testing of a generic algorithm has provided little saving of analysis time, while necessarily leading to non-optimal solutions for particular cases. Rather, to achieve the most efficient procedure (e.g. maximum expected catch for the same perceived risk), it seems preferable to condition the operating models used for testing immediately on the data already available for the resource in question – in other words, to use for operating models those estimated in the conventional stock assessment process. This avoids the inefficiency of use of a generic procedure designed to be robust to some uncertainty that cannot (given the data already available) pertain to the situation under consideration.

Basis for choosing between alternative candidate management procedures

The aim that a chosen management procedure should show performance that is robust across the range of alternative plausible scenarios consistent with available data begs two questions: how robust, and how plausible? Clearly, performance must deteriorate to some extent as the operating model is changed further from that for which the procedure provides optimal performance – but how much deterioration can be admitted before a candidate procedure would be deemed unacceptable?

If no limits are placed on plausibility, and a "worst-case scenario" approach adopted whereby adequate robustness is insisted upon for the most pessimistic situation imaginable, clearly no candidate management procedure (except an indefinite suspension of harvesting) could ever be deemed acceptable. To avoid this impasse, Butterworth *et al.* (1996) argue that performance statistics need to be considered as weighted summations over the different scenarios postulated, where the weights are proportional to the relative plausibilities (posterior probabilities should formal statistical comparisons be possible) accorded to each scenario.

While acknowledging that these aspects give rise to difficulties, one advantage of pursuing a management procedure basis for a fishery is an automatic linkage with the precautionary approach. FAO (1996) state that the precautionary approach to fisheries management requires "prior identification of undesirable outcomes and of measures that will avoid them or correct them promptly" (p. 4) and note (p. 8) that it "should involve the formulation of decision rules, which specify in advance what action should be taken when specified deviations from operational targets and constraints are observed". FAO (1996) notes further (p. 9) that: "A management plan should not be accepted until it has been shown to perform effectively in terms of its ability to avoid undesirable outcomes. The evaluation can be used to determine whether the data and assessment methods available for management are sufficient to meet management objectives. The evaluation should attempt to determine if the management plan is robust to statistical uncertainty and to incorporate knowledge on factors such as uncertain stock identity and abundance, stock dynamics, and the effects of environmental variability." In other words, in the view of the meeting upon which that report was based, the process to be followed to be consistent with the precautionary approach is exactly that pursued in the robustness-testing of a management procedure.

Selection in practice

Norms for the processes described above are still in the process of establishment so that, at the scientific level, there is currently a degree of ambivalence on how best to proceed. Thus far, the Scientific Committee of the IWC has seemingly adopted the approach of satisfying a risk criterion for the worst-case scenario (e.g. IWC, 1994b, for RMP implementations for minke whales) – an approach this conservative would not likely be acceptable in active fisheries.

Perhaps fortunately, in many cases considered so far, only one or two factors tend to dominate when the sensitivity of risk evaluations to alternative scenarios is considered (see Table 2). For example, the primary scientific focus in development of a management procedure for the South African west coast rock lobster resource was to secure a reasonable degree of robustness to a possible decrease in (or positively biased estimate of) current recruitment levels (Johnston and Butterworth, 1997; Johnston, 1998). This feature, in principle, eases the difficulty of according weights to alternative scenarios, as suggested by Butterworth *et al.* (1996), but as yet no-one appears to have been sufficiently brave to attempt this weighting.

Given the linkage to the precautionary approach discussed above, these difficulties translate directly into identical problems in providing objective criteria in terms of which to apply this approach consistently. One positive advance might be provided by developments in the management of Australia's southern shark (*Galeorhinus galeus*) fishery, where a recent selection was made on the basis of a marginal expected gain in risk avoidance equal to the marginal expected loss in average catch. The expectation in this instance was determined by weighting the results for 22 scenarios equally (Punt and Walker, 1998).

Moving from scientific fora to those involving industry or decision-makers, further difficulties arise as regards the interpretation of performance statistics by non-specialists (unsurprisingly, as even scientists active

Table 2. Uncertainties to which performance is most sensitive for a subset of studies in Table 1.

Cape hake off South Africa
Extent of future recruitment variability
Extent of future changes in fishing efficiency (if no future research surveys take place)
Natural mortality-at-age
Cape hake off Namibia
Bias in survey-based abundance estimates
Extent of recruitment variability
South African sardine
Form of the stock-recruitment relationship
Bias in the survey-based abundance estimates
Natural mortality-at-age
South African rock lobster
Future somatic growth rate
Current abundance and future recruitment levels
Namibian Cape fur seals
Probability of environmentally caused mass mortality
Natural mortality rate of adult females
Australian gemfish
Form of the stock-recruitment relationship
Whether selectivity is density-dependent
Historical catches
Namibian orange roughy
Absolute estimate of virgin biomass
Natural mortality rate
Minke whales
MSY rate
Bias in survey-based abundance estimates

in the field are feeling their way!). The most acute problem is a definition of risk that is meaningful to such audiences (Butterworth *et al.*, 1997).

For long-lived species with consequent low levels of abundance fluctuations, comparison of the median population projections corresponding to different candidate procedures can provide helpful insights, but problems arise in communicating the implications of variability in pelagic fisheries for short-lived species. For example, for the South African anchovy, such trajectories would be meaningless, but equally statistics of distributions prove difficult for decision-makers to assimilate. One approach to this problem which met with some success was the development of computer simulation games to familiarize industry with the range of possible future occurrences in the South African anchovy fishery under different management procedures (Butterworth *et al.*, 1997).

The difficulties that lay persons have in interpreting performance statistics have seemingly frustrated meaningful deliberation on and selection between alternative procedures at decision-maker level. For example, given alternative options, the IWC Commissioners consistently apparently simply chose the one reflecting the lowest risk, and hence the lowest catch. Effectively, therefore, the scientists who selected the range of alternatives to present were making the key trade-off choice. In contrast, the South African Sea Fisheries Advisory Council (SFAC) did the reverse in its recent final selection between candidates for the South African west coast lobster resource, and virtually duplicated this in its choice for Cape hake on the west coast (Geromont *et al.*, 1999). Again, therefore, rather than actually make a risk-reward trade-off choice, decision-makers effectively passed responsibility regarding risk back to scientists on the basis that all candidates presented had been deemed by those scientists to have shown satisfactory risk-related performances.

In South Africa, logistics (particularly the fact that the decision-making body, the SFAC, met only a few times each year) have made it virtually impossible to separate the long-term considerations appropriate to management procedure selection from the short-term interest in the size of the TAC for the next year. Inevitably then, procedure selection has in practice been unduly influenced by this last consideration. This fact of life, while "lamentable" in principle, should perhaps rather be viewed pragmatically in the positive light that adoption of a management procedure at least means that TACs for subsequent years will be determined by trends in resource status, rather than by similar short-term considerations with their associated higher levels of risk to the resource.

Robustness testing

There is a trade-off between the time required for a complete and thorough examination of operating models/candidate management procedures and the urgency that may exist for getting a management procedure in place. In Australia, the preference has been to delay consideration of the adoption of management procedures until: (a) a wide variety of operating models has been considered, (b) many candidate management procedures have been considered, and (c) all stakeholders are fully aware of the assumptions underlying the analyses. Thus, candidate management procedures for the eastern stock of gemfish are still under evaluation after 2 years, and a similar exercise for the two main species of Australia's southern shark fishery has yet to commence. At the IWC, development of the (generic) RMP took some 6 years. The design, implementation, and results review of subsequent case-specific trials for North Atlantic and Southern Hemisphere minke whales took some 3 years (IWC 1993, 1994b). At the other extreme, management procedures for Namibian hake and seals have been put in place in a matter of months after less extensive evaluation.

The IWC example is perhaps of limited relevance to commercial fisheries: given the existing moratorium on commercial whaling, there is no immediate pressure to (finalize a basis to) provide catch limit recommendations. Aboriginal whaling remains in progress, but given that multi-year block quotas are in place (and quite defensibly so for long-lived species on which the fishing mortality rate is low), there is similarly little real time-pressure to complete an Aboriginal Whaling Management Procedure. In these circumstances, the widening range of investigations pursued in the Scientific Committee of the IWC may owe more to scientific curiosity than to real-time management needs and concerns. The Australian fisheries mentioned do not enjoy this ‘‘luxury’’, and accordingly require interim bases for their management pending finalization of a management procedure. The general danger here is that the resources required to achieve such interim management are removed from the (limited) resources available for the exercise of management procedure development, so further delaying its finalization.

However, there are dangers in putting a management procedure in place too rapidly. With the benefit of hindsight, the adoption of a joint sardine–anchovy management procedure in South Africa in 1994 was probably premature. *Ad hoc* modifications rapidly became necessary, primarily because the sardine/anchovy abundance ratio rose markedly above its range during the preceding decade, to which robustness trials had been conditioned. The situation was exacerbated by a lack of specification of objectives in a situation where maximal simultaneous utilization of both species was impossible. Trade-off decisions are necessary because greater anchovy catches mean larger juvenile sardine by-catches in the fishery, necessitating smaller directed sardine catch allocations (Cochrane *et al.*, 1998). Such rapid adjustments can undercut the ‘‘respect’’ which adopted procedures and the recommendations they provide should be accorded, and lead to a lessening of scientists’ credibility.

An argument against the management procedure approach, and its supposed greater time-efficiency compared to conventional annual assessments, is that it simply ‘‘shifts the goalposts’’. Instead of arguments by parties with interests in either higher or lower TACs in the short term being focused on aspects of the assessment, they concentrate instead on the data input to the management procedure. Certainly there is evidence for this practice in the IWC, where Norway’s unilateral application of the RMP to set catch limits for its harvesting (legal in terms of its standing objection to the IWC’s commercial whaling moratorium) led to an enormously detailed investigation into the analysis of sighting survey data used to provide abundance estimates input to the RMP (IWC, 1997d). Similarly, in South Africa the process of adoption/revision of management procedures for west coast rock lobster and hake has seen debate concentrate on the standardization and reliability of c.p.u.e.-based indices input to these procedures. Nevertheless, these detailed reviews have proved valuable exercises, and hopefully constitute investments that will both improve management and save debating time in the future.

It is clear from the preceding section that some issues (see Table 2) have a large impact on the performance of management procedures while others do not. For example, Johnston and Butterworth (1997) found that the management procedures under consideration for the South African west coast rock lobster were robust to plausible levels of illegal catch and changes thereof, and Butterworth *et al.* (1998) report that the results of the evaluations for the Namibian seal resource were insensitive to many changes in the assumptions used to condition the simulations. The results from the increasing number of management procedure evaluations help to identify a ‘‘minimum set of questions that require addressing’’ to consider when evaluating a management procedure for a new species/region. However, some care should be taken in extrapolating the results for certain species/regions to others. For example, possible past changes in the stock-recruitment relationship appear to be of relative minor importance for Namibian seals, but not for eastern gemfish. Therefore, such a ‘‘minimal set’’ might be closer to the union than the intersection of the sets of factors listed for various resources considered in Table 2.

What of the resources required to undertake management procedure evaluations? A general lack of modelling skills has been identified as a major impediment to the provision of effective stock assessment advice in even a developed country such as Australia (Lyle, 1998). This problem is exacerbated for the evaluation of management procedures because the operating models used are often markedly more complex than the models underlying most stock assessments, and because no general software packages are available that implement ‘‘generalized’’ operating models. FAO (1996) note (p. 9) that ‘‘for small or artisanal fisheries, computationally intensive management analyses are often not possible or cost-effective’’. Therefore, the management procedure approach seems unlikely to replace conventional assessment-based management globally. A more realistic scenario would be many intermediate cases where the approach is implemented, but without as complex a robustness-testing process as some might argue necessary, simply because of limitations in the resources available to conduct analyses.

Evaluating research priorities

An important by-product of the management procedure evaluation process is that it provides a basis for prioritizing research. If meeting a risk criterion in robustness tests for a particular scenario necessitates the adoption of a relatively conservative management procedure in allowable catch terms, then if subsequent research can disprove the hypothesis underlying that scenario, a revised procedure could allow larger catches for the same perceived risk. Research priorities should therefore be accorded to addressing (resolvable) uncertainties for which robustness tests show the largest impacts of this nature on anticipated procedure performance. Possibly the greatest area of concern in this regard is uncertainty about stock structure, as evidenced in many IWC studies (1993, 1994a, b). This is because the methods commonly used for stock discrimination (e.g. genetics, morphology) generally have low power, so that while a statistically significant result implies some stock structure within the resource, usually little can be concluded from a non-significant result.

Evaluating the benefits of research (McDonald and Smith, 1997; McDonald *et al.*, 1997) is a key issue in Australian fisheries management, because the fact that many of those fisheries are relatively small by international standards does not change the amount of work required to conduct assessments. Efficient use of limited resources is therefore essential (a formal objective for federally managed fisheries in Australia is to "implement efficient and cost effective management" (Commonwealth of Australia, 1998)), and the Australian Fisheries Management Agency (AFMA) Research and Environment Committee recently agreed (for fisheries for which this was feasible) to base research prioritization on the outputs from management procedure development processes. As a corollary, those applying for funding of a research project are provided a framework within which to motivate: to show how resolution of the hypothesis they are to examine will impact procedure performance.

This basis for evaluation is already underway in Australia, with consideration of the merits of an independent shark survey in progress. Despite its longer history of management procedure implementation, moves in South Africa towards using the results of the evaluation process as a formal guide to research prioritization have, however, been disappointingly slow.

The management procedure approach is argued to be more cost-effective, reducing the time currently spent on arguing typically annual assessment exercises, and so freeing resources to address ultimately more important issues which require longer time-spans than 1 year to resolve (Butterworth *et al.*, 1997). It is perhaps premature to assess whether this goal is actually being achieved. Initiation of the approach requires considerable resources, and the "goalpost moving" aspects discussed above increase such requirements. Hopefully though, these will prove in time to be transient effects only, and the savings anticipated will be achieved.

Reception accorded the approach by industry and decision-makers

In South Africa, the pelagic industry has welcomed the opportunity to participate in the trade-off debates implicit in management procedure development and selection processes. Of particular import has been choices of the limits on the extent to which TACs can be reduced from one year to the next. Smaller limits mean a more stable industry, but necessitate lower catches on average for invariant risk levels.

However, South African industries have shown resistance to the "locking-in" aspect of agreement to a management procedure, possibly because this reduces options to argue for TAC adjustments for purported socio-economic reasons when resource signals indicate contrary changes. Formal objections tend to centre on alleged unreliability of procedure inputs, particularly c.p.u.e. The argument given is that confidence in the procedure requires confidence in the inputs. However, it is unclear that some of the associated uncertainties are soluble (at least in the short term), so the counter-argument has been that the management procedure approach is one which formally addresses the aspect of necessary robustness to such existing uncertainties.

In both South Africa and Namibia, decision-makers in the form of the Advisory Councils (who advise the Ministers directly) seem to have appreciated the lessening of uncertainty and debate that management procedure adoption has brought to the TAC recommendation process. Formally, it is the Ministers of the responsible Departments who are empowered to set annual TACs in both countries, and it is unclear to what extent these Ministers see themselves bound to follow the recommendations provided by the procedures whose adoption their Advisory Councils recommend to them. Despite the clear administrative advantages, particularly in fisheries such as that for the South African anchovy where quick decisions are necessary because of the limited time between recruitment surveys and the fish remaining available to the fishery, apparently implicit reservations still exist about the concept of pre-adopting a "formula" in contrast to formally reviewing and approving its TAC output once survey data become available. Scientists still have work to do to convince senior administrators and politicians that their role in fisheries management needs to be strategic (choosing between procedures on the basis of desired trade-offs), and not tactical (adjusting and adopting TACs), *inter alia* because the extra flexibility that the latter implies must lead to lower average catches for the same perceived level of risk.

The application of the management procedure approach has helped to encourage industry and decision-makers in South Africa to focus on better specifying management objectives – a process in which stakeholders other than population modellers, including other scientists, industry, and decision-makers also need to be involved to ensure that their concerns are adequately addressed. However, a difficulty in the South African context has been the uncertainties associated with a concurrent programme of redistribution of fishing rights to previously disadvantaged groups. In the absence of clarity as to the extent of the latter process, existing industry has bemoaned the realism of expecting them to specify their long-term objectives. This lack of certainty about future rights has doubtless also played some role in the choice of procedures that give the (near) highest immediate catch levels from the suite of options presented, as discussed above.

The specification of these objectives becomes more difficult when different agencies are involved in the management of the same resource, having responsibility for different parts of its habitat. Hence, for example, both a State (Tasmania, within 12 miles) and the Commonwealth of Australia (from 12 to 200 miles) have responsibility for blue warehou (*Seriolella brama*) management in Australia, so that an Offshore Constitutional Settlement has to be negotiated politically for the adoption of a management procedure to apply to the whole resource. The multi-sector nature of some fisheries, e.g. bottom trawl and longlining for hake in South Africa, and both these gears plus gillnets for school shark off southern Australia, can also lead to implementation problems.

Concluding remarks

What follows is a cautious attempt to summarize some of the experiences related above.

- Management procedures are preferably population-model-based rather than empirical, case-specific rather than generic, and simple rather than complex.
- Increasing numbers of evaluations are making it possible to develop a minimal set of key factors necessary to include in robustness trials.
- The process of robustness testing required for management procedures dovetails with the requirements of the precautionary approach.
- The range of robustness tests considered in different applications has varied widely. In practice, limited resources in terms of modelling expertise will likely determine the extent to which such tests are pursued in most applications.
- Performance statistics provided by management procedure evaluations are more helpful to industry and decision-makers, as they relate more closely to issues of direct concern to them. The definition and interpretation of risk-related statistics are, however, problematic, and scientists (let alone lay decision-makers) are still struggling to establish appropriate norms.
- Approaches to accord weights that reflect the relative plausibilities of alternative scenarios will need to be developed to facilitate the interpretation of the results of robustness trials, so as to allow movement away from the "worst case scenario" basis for procedure selection.
- At least at the start of the process of implementation of a management procedure, a greater degree of focus on and critical examination by interested parties of the data inputs proposed for the procedure should be expected.
- Management procedure development incidentally provides a basis for evaluating research priorities, whose potential awaits full utilization.
- It is premature to judge whether implementation of a management procedure approach will achieve the intended goal of diverting research resources from annual assessment exercises to longer-period projects with better prospects of resolving key uncertainties.
- More interaction with decision-makers is required if they are to fully realize and hence achieve the benefits, particularly in terms of cost-efficiency, of moving from an annual assessment to a management-procedure-based approach to setting (say) TACs.

Have management procedures yet been demonstrated to outperform the standard annual-best-assessment basis for management advice *in practice*? To conclude, reference is made to the case of the Namibian resource of Cape hake (Butterworth and Geromont, 1997), for which decision-makers were faced near the end of 1997 with two diametrically opposing assessments of resource status, as illustrated in Figure 2 of Geromont *et al.* (1999).

The industry, believing that research survey abundance estimates should be treated as relative indices, argued that the resource was underutilized and that the then current TAC of 120 000 t should be doubled. In contrast, government scientists believed that the absolute values for these estimates were reasonably reliable, that the resource was accordingly highly depleted, and hence that the TAC should be halved. There was no immediate prospect of resolving the debate on whose view of the survey estimates was the more reliable. The issue was of national importance, the hake fishery being one of the largest single contributors to the Namibian GDP (the landed value alone of the catch constituting more than 5% thereof).

Conventional best-assessment approaches to setting TACs provide no basis for a solution in these circumstances. However, a simple management procedure approach (see Table 2 of Geromont *et al.*, 1999) offered a way forward. Evaluations of the anticipated performance of the procedure (which was based on trends in indices of abundance) showed anticipated catch increases over time if the industry view was correct, but also security against resource depletion if the scientific view was the more appropriate, because catches would be reduced sufficiently rapidly over time in response to downward index trends. Industry and scientists were therefore able to find a compromise choice of a value for the control parameter of the procedure which reflected a mutually acceptable trade-off as regards adequately addressing their differing concerns. Hence, the responsible Minister's eventual decision was based not on some *ad hoc* compromise, but on an appreciation of the trade-off between risk and future catch that the decision involved.

Acknowledgements

We thank many colleagues, too numerous to mention individually, for discussions and the data provision that contributed to the analyses that underlie our views expressed above, and also Victor Restrepo and Gunnar Stefansson for their comments on an earlier version of this paper.

References

Baldursson, F. M., Danielsson, A., and Steffanson, G. 1996. On the rational utilization of the Icelandic cod stock. ICES Journal of Marine Science, 53: 643–658.

Beddington, J. R., and Cooke, J. G. 1983. The potential yield of fish stocks. FAO Fisheries Technical Paper, 242, Rome, FAO. 47 pp.

Bergh, M. O., and Butterworth, D. S. 1987. Towards rational harvesting of the South African anchovy considering survey imprecision and recruitment variability. South African Journal of Marine Science, 5: 937–951.

Branch, T. A. 1998. Assessment and adaptive management of orange roughy off southern Africa. MSc thesis, University of Cape Town. 202 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 in 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., Cochrane, K. L., and De Oliveria, J. A. A. 1997. Management procedures: a better way to manage fisheries? The South African experience. *In* Global trends: fisheries management, pp. 83–90. Ed. by E. K. Pikitch, D. D. Huppert, and M. P. Sissenwine. American Fisheries Society Symposium, 20, Bethesda, Maryland.

Butterworth, D. S., Cowan, C. L., and Johnston, S. J. 1998. The development of a management procedure for Namibian seals. Report to the Ministry of Fisheries and Marine Resources, Namibia. 47 pp.

Butterworth, D. S., De Oliveria, J. A. A., and Cochrane, K. L. 1993. Current initiatives in refining the management procedure for the South African anchovy resource. *In* Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations, pp. 439–473. Ed. by G. Kruse, D. M. Eggers, R. J. Marasco, C. Pautzke, and T. J. Quinn. Alaska Sea Grant College Program Report No. 93-02. University of Alaska Fairbanks.

Butterworth, D. S., and Geromont, H. F. 1997. Evaluation of a range of possible simple interim management procedures for the Namibian hake fishery. Report to the Ministry of Fisheries and Marine Resources, Namibia. 28 pp.

Butterworth, D. S., Punt, A. E., and Smith, A. D. M. 1996. On plausible hypotheses and their weighting, with implications for selection between variants of the Revised Management Procedure. Reports of the International Whaling Commission, 46: 637–640.

Christensen, S. 1997. Evaluation of management strategies – a bioeconomic approach applied to the Greenland shrimp fishery. ICES Journal of Marine Science, 54: 412–426.

Cochrane, K. L., Butterworth, D. S., De Oliveria, J. A. A., and Roel, B. A. 1998. Management procedures in a fishery based on highly variable stocks and with conflicting objectives: experiences in the South African pelagic fishery. Reviews in Fish Biology and Fisheries, 8: 177–214.

Cochrane, K. L., and Starfield, A. M. 1992. The potential use of predictions of recruitment success in the management of the South African anchovy resource. South African Journal of Marine Science, 12: 891–902.

Collie, J. S., and Walters, C. J. 1991. Adaptive management of spatially replicated groundfish populations. Canadian Journal of Fisheries and Aquatic Sciences, 48: 1273–1284.

Commonwealth of Australia 1998. Fishery Management Act 1991. Reprinted as in force on 31 January 1998 (including amendments up to and including Act no. 120 of 1997). Reprint No. 2. Australian Government Printing Service, Canberra. vii+178 pp.

Corten, A. 1993. The use of the MBAL concept in management advice. ICES CM 1993/H:19. 9 pp.

Davies, C. R. 1995. Patterns of movement of three species of coral reef fish on the GBR. PhD thesis, Department of Marine Biology, James Cook University, Townsville, Australia. 212 pp.

De Oliveira, J. A. A., Butterworth, D. S., Roel, B. A., Cochrane, K. L., and Brown, J. P. 1998a. The application of a management procedure to regulate the directed and bycatch fishery of South African sardine *Sardinaps sagax*. South African Journal of Marine Science, 19: 449–469.

De Oliveira, J. A. A., Butterworth, D. S., and Johnston, S. J. 1998b. Progress and problems in the application of management procedures to South Africa's major fisheries. *In* Fishery stock assessment models, pp. 513–530. Ed. by F. Funk, T. J. Quinn, J. Heifetz, J. N. Ianelli, J. E. Powers, J. F. Schweigert, P. J. Sullivan, and C. I. Zhang. University of Alaska Sea Grant, AK-SG-98-01, Fairbanks.

Donovan, G. P. (ed.) 1989. The comprehensive assessment of whale stocks: the early years. Reports of the International Whaling Commission, Special Issue 11. 210 pp.

FAO 1996. Precautionary approach to fisheries. Part 1: Guidelines on the precautionary approach to capture fisheries and species introductions. FAO Fisheries Technical Paper, 350, Part 1. Rome, FAO. 52 pp.

Francis, R. I. C. C. 1992. Use of risk analysis to assess fishery management strategies: a case study using orange roughy (*Hoplostethus atlanticus*) on the Chatham Rise, New

Zealand. Canadian Journal of Fisheries and Aquatic Sciences, 49: 922–930.

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

Geromont, H. F., and Butterworth, D. S. 1998a. Initial evaluation of a range of possible management procedures for west coast hake. SFRI Document WG/04/98/D:H:18. 26 pp.

Geromont, H. F., and Butterworth, D. S. 1998b. Management procedure robustness trials for west coast hake. SFRI Document WG/06/98/D:H:24. 17 pp.

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, D. L., Hilborn, R., Stocker, M., and Walters, C. J. 1988. Alternative harvest strategies for Pacific herring (*Clupea harengus pallasi*). Canadian Journal of Fisheries and Aquatic Sciences, 45: 888–897.

Hilborn, R. 1997. Uncertainty, risk and the precautionary approach. *In* Global trends: fisheries management, pp. 100–106. Ed. by E. K. Pikitch, D. D. Huppert, and M. P. Sissenwine. American Fisheries Society Symposium, 20, Bethesda, Maryland.

Hollowed, A. B., and Megrey, B. A. 1993. Evaluation of risk associated with application of alternative harvest strategies for Gulf of Alaska walleye pollock. *In* Proceedings of the international symposium on management strategies for exploited fish populations, pp. 291–320. Ed. by G. Kruse, D. M. Eggers, R. J. Marasco, C. Pautzke, and T. J. Quinn. Alaska Sea Grant College Program Report No. 93-02. University of Alaska Fairbanks.

Horwood, J. W. 1994. Modelling of fisheries management strategies. Privately published. 85 pp.

IWC 1991. Report of the scientific committee, Annex D. Report of the Sub-Committee on Management Procedures. Reports of the International Whaling Commission, 41: 90–112.

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

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 1994a. Report of the Scientific Committee, Annex G. Report of the Working Group on North Pacific Minke Whales Management Trials. Appendix 3. Specifications of the North Pacific minke whale whaling trials. Reports of the International Whaling Commission, 44: 133–143.

IWC 1994b. 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 1994c. Report of the Scientific Committee, Annex H. The Revised Management Procedure (RMP) for Baleen Whales. Reports of the International Whaling Commission, 44: 145–152.

IWC 1997a. Report of the Scientific Committee. Annex J. Report of the Working Group on North Pacific Minke Whale Trials. Reports of the International Whaling Commission, 47: 203–226.

IWC 1997b. Report of the Scientific Committee, Annex I. Report of the Workshop on the Development of an Aboriginal Subsistence Whaling Management Procedure (AWMP). Reports of the International Whaling Commission, 47: 192–202.

IWC 1997c. Report of the Scientific Committee, Annex P. Report of the Working Group on Aboriginal Subsistence Whaling Management Procedures. Reports of the International Whaling Commission, 47: 243–249.

IWC 1997d. Report of the Abundance Estimation Working Group. Reports of the International Whaling Commission, 47: 261–290.

Johnston, S. J. 1998. The development of an operational management procedure for the South African west coast rock lobster fishery. PhD thesis, University of Cape Town. 371 pp.

Johnston, S. J., and Butterworth, D. S. 1997. OMP results for a number of "strategies". SFRI Document WG/07/97/WCL23. 23 pp.

Kirkwood, G. P. 1997. The Revised Management Procedure of the International Whaling Commission. *In* Global trends: fisheries management, pp. 41–99. Ed. by E. K. Pikitch, D. D. Huppert, and M. P. Sissenwine. American Fisheries Society Symposium, 20, Bethesda, Maryland.

Kirkwood, G. P., and Smith, A. D. M. 1996. Assessing the precautionary nature of fishery management strategies. *In* Precautionary approach to fisheries. Part 2: scientific papers, pp. 141–158. FAO Fisheries Technical Paper, 350, Part 2, Rome, FAO.

Lowe, S. A., and Thompson, G. G. 1993. Accounting for uncertainty in the development of exploitation strategies for the Atka mackerel resources of the Aleutian Islands. *In* Proceedings of the international symposium on management strategies for exploited fish populations, pp. 203–231. Ed. by G. Kruse, D. M. Eggers, R. J. Marasco, C. Pautzke, and T. J. Quinn. Alaska Sea Grant College Program Report No. 93-02. University of Alaska Fairbanks.

Lyle, J. (comp.) 1998. Providing stock assessments for Australian fisheries. Australian Society for Fish Biology, Tasmanian Aquaculture and Fisheries Institute. 46 pp.

Magnusson, K. G. 1992. A feedback and probing strategy to regulate harvesting from a renewable resource. IMA Journal of Mathematics Applied in Medicine & Biology, 9: 43–65.

Mapstone, B. D., Campbell, R. A., and Smith, A. D. M. 1996. Design of experimental investigations of the effects of line and spear fishing on the Great Barrier Reef. CRC Reef Research Centre Technical Report, 7. Townsville, CRC Reef Research Centre. 86 pp.

Maunder, M. N., Starr, P. J., and Hilborn, R. in press. A Bayesian analysis to estimate loss in squid catch due to the implementation of a sea lion population management plan. Marine Mammal Science.

McDonald, A. D., and Smith, A. D. M. 1997. A tutorial on evaluating expected returns from research for fishery management. Natural Resource Modelling, 10: 185–216.

McDonald, A. D., Smith, A. D. M., Punt, A. E., Tuck, G. N., and Davidson, A. J. 1997. Empirical evaluation of expected returns from research on stock structure for determination of total allowable catch. Natural Resource Modelling, 10: 3–29.

Myers, R. A., Rosenberg, A. A., Mace, P. M., Barrowman, N., and Restrepo, V. R. 1994. In search of thresholds for recruitment overfishing. ICES Journal of Marine Science, 51: 191–205.

O'Boyle, R. N., Sinclair, A. F., and Hurley, P. C. F. 1991. A bioeconomic model of an age-structured groundfish resource exploited by a multi-gear fishing fleet. ICES Marine Science Symposia, 193: 264–274.

Overholtz, W. J. 1993. Harvesting strategies and fishing mortality reference point comparisons for the Northwest Atlantic stock of Atlantic mackerel (*Scomber scombrus*). Canadian Journal of Fisheries and Aquatic Sciences, 50: 1749–1756.

Pearn, R. 1994. Rock lobster tagging shows movement occurs. Fishing Today, 7: 27–29.

Pelletier, D., and Laurec, A. 1992. Management under uncertainty: defining strategies for reducing overexploitation. ICES Journal of Marine Science, 49: 389–401.

Powers, J. E., and Restrepo, V. R. 1998. An examination of operational procedures in management of marine fish stocks. *In* Abstracts of ICES Symposium "Confronting uncertainty in the evolution and implementation of fisheries-management systems, Cape Town, ICES. p. 18.

Punt, A. E. 1992. Management procedures for Cape hake and baleen whale resources. Benguela Ecology Programme Report, 23. 689 pp.

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 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 in Fisheries and Aquatic Sciences, 120.

Punt, A. E. 1994. Assessments of the stocks of Cape hake *Merluccius* spp. off South Africa. South African Journal of Marine Science, 14: 159–186.

Punt, A. E. 1995. The performance of a production-model management procedure. Fisheries Research, 21: 349–374.

Punt, A. E. 1997. The performance of VPA-based management. Fisheries Research, 29: 217–243.

Punt, A. E., and Butterworth, D. S. 1995. The effects of future consumption by the Cape fur seal on catches and catch rates of the Cape hakes. 4. Modelling the biological interaction between Cape fur seals *Arctocephalus pusillus pusillus* and Cape hakes *Merluccius capensis* and *M. paradoxus*. South African Journal of Marine Science, 16: 255–285.

Punt, A. E., Butterworth, D. S., and Martin, J. 1995. The effects of errors in the placement of the boundary between the West and South Coast hake *Merluccius* spp. stocks on the performance of the current hake management procedure. South African Journal of Marine Science, 15: 83–98.

Punt, A. E., and Kennedy, R. B. 1997. Population modelling of Tasmanian rock lobster, *Jasus edwardsii*, resources. Marine and Freshwater Research, 48: 967–980.

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., and Walker, T. I. 1998. Stock assessment and risk analysis for the school shark (*Galeorhinus galeus*) off southern Australia. Marine and Freshwater Research, 49: 719–731.

Quinn, T. J., Fagen, R., and Zheng, J. 1990. Threshold management polices for exploited populations. Canadian Journal of Fisheries and Aquatic Sciences, 47: 2016–2029.

Rosenberg, A. A., and Brault, S. 1993. Choosing a management strategy for stock rebuilding when control is uncertain. *In* Risk evaluation and biological reference points for fisheries management, pp. 243–249. Ed. by S. J. Smith, J. J. Hunt, and D. Rivard. Canadian Special Publication in Fisheries and Aquatic Sciences, 120.

Rosenberg, A. A., and Restrepo, V. 1993. The eloquent shrug: expressing uncertainty and risk in stock assessments. ICES CM 1993/D:12. 15 pp.

Sainsbury, K. R. A., Campbell, R., Lindholm, R., and Whitelaw, A. W. 1997. Experimental management of an Australian multispecies fishery: examining the possibility of trawl-induced habitat modification. *In* Global trends: fisheries management, pp. 107–112. Ed. by E. K. Pikitch, D. D. Huppert, and M. P. Sissenwine. American Fisheries Society Symposium, 20, Bethesda, Maryland.

Sigler, M. F., and Fujioka, J. T. 1993. A comparison of policies for harvesting stablefish, *Anoplopoma fimbria*, in the Gulf of Alaska. *In* Proceedings of the international symposium on management strategies for exploited fish populations, pp. 7–19. Ed. by G. Kruse, D. M. Eggers, R. J. Marasco, C. Pautzke, and T. J. Quinn. Alaska Sea Grant College Program Report No. 93-02. University of Alaska Fairbanks.

Smith, T., Schweder, T., Allison, C., Palka, D., Punt, A. E., and Butterworth, D. S. in press. Effects of combinations on the risk performance of the IWC's commercial whaling management procedure. Journal of Cetacean Research and Management (special issue).

Starr, P. J., Breen, P. A., Hilborn, R. H., and Kendrick, T. H. 1997. Evaluation of a management decision rule for a New Zealand rock lobster substock. Marine and Freshwater Research, 48: 1093–1101.

Wade, P. 1998. Calculating limits to the allowable human-caused mortality of cetaceans and pinnipeds. Marine Mammal Science, 14: 1–37.