

Tips and tricks in designing management procedures

Rebecca A. Rademeyer, Éva E. Plagányi, and Doug S. Butterworth

Rademeyer, R. A., Plagányi, É. E., and Butterworth, D. S. 2007. Tips and tricks in designing management procedures. – *ICES Journal of Marine Science*, 64: 618–625.

Management procedures (MPs) are becoming widely used in fisheries management, but guidelines to assist in their construction, evaluation, and implementation are few. We provide simple guidelines by drawing on experience from developing and applying MPs in southern Africa and internationally. Suggestions are provided on how to choose between candidate MPs and on key trade-offs in selecting between data-based (empirical) and model-based formulations. Assistance is also provided in dealing with different sources of uncertainty, such as deciding which operating models should be included in a reference set used for primary simulation testing and tuning (in contrast to robustness or sensitivity tests), and on how weights for the associated alternative hypotheses are most practically assigned. Finally, some guidelines are given for presenting the results effectively, which is one of the key challenges of a successful implementation process.

Keywords: management procedure, operating model, robustness, simulation testing, uncertainty.

Received 30 June 2006; accepted 22 January 2007; advance access publication 3 May 2007.

R. A. Rademeyer, É. E. Plagányi and D. S. Butterworth: MARAM (Marine Resource Assessment and Management Group), Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch 7701, South Africa. Correspondence to R. A. Rademeyer: tel: +27 21 650 3656; fax: +27 21 686 0477; e-mail: rebecca.rademeyer@uct.ac.za

Introduction

Management procedures (MPs) (Butterworth and Punt, 1999) and similar frameworks such as management strategy evaluation (MSE) (Smith *et al.*, 1999) are becoming more widely used in fishery management because they provide formalizations of long-term, robust strategies that are designed to satisfy multiple conflicting objectives. There are few guidelines available, however, to assist in their construction, evaluation, implementation, and presentation. We provide practical guidelines for new developers by drawing on experience in southern Africa (Plagányi *et al.*, 2007) and internationally.

MPs involve assessing the consequences of alternative options for management actions for both the target resource(s) and associated fisheries. Simulation trials ensure that the associated decision rules lead to performance that is robust to uncertainties about the dynamics of the resource being managed. The simulation framework essentially consists of an operating model (OM) to simulate the “true” system of resource dynamics and fishery and generate future resource-monitoring data typical of what would become available in practice, an estimator that provides information on resource status and productivity from these data, and a harvest control rule (HCR) that outputs a management action in the form of a total allowable catch (TAC) or allowable fishing effort (Kell *et al.*, 2006). Key steps in designing MPs are described, with suggestions for selecting the best option at each step.

Given that the evolution of the MP approach has been accompanied by the introduction of several technical terms with specific meanings in an MP context, a glossary is provided in the Appendix to assist readers.

Constructing OMs

The first step in assessing the consequences of different management options is to model several possible scenarios for the underlying true dynamics for the resource population(s) of interest and the impact of exploitation. These OMs are used as the basis to compute how the resource responds to different future levels of catch or effort. Typical population dynamics models include age structure, growth, natural mortality, and a stock–recruitment relationship with associated variability, but they may also include associated species or even the entire ecosystem (Smith *et al.*, 2007). The models are fit to data just as in a typical stock assessment process (Geromont *et al.*, 1999; Rademeyer, 2003; De Oliveira and Butterworth, 2004). Robustness to alternative models needs to extend only to those consistent with available data. This fitting process is also termed conditioning the OM to the available information (Butterworth, 1999; IWC, 2005).

The reason a range of OMs is required is that various uncertainties, which are always present in any assessment of the status and productivity of a resource, can affect the consequences of management measures. These uncertainties relate not only to the fit of the model to the data (i.e. uncertainty in the parameter values within a single model structure), but also to specification of the model structure (i.e. uncertainty about the processes operating in the real world; Butterworth and Punt, 1999).

In the initial phase of evaluating candidate MPs (CMPs), a single OM may typically be selected as a reference case. However, experience suggests that usually it is desirable to select a core set of OMs, termed the reference set (Rademeyer and Butterworth, 2006a), which includes the most important uncertainties, i.e. alternative scenarios that are both highly plausible

and have major impacts on results. For example, in the current development of an MP for the South African hake resource (comprising two species, *Merluccius capensis* and *M. paradoxus*, that are not distinguished in the commercial catches), three key aspects of the assessment account for most of the uncertainty regarding resource status and productivity (Rademeyer and Butterworth, 2006a). The reference set has been constructed by incorporating variations around these three aspects: (i) two (age-dependent) upper bounds for natural mortality; (ii) three assumptions about the species split in pre-1978 catches (surveys provide information on species composition thereafter); and (iii) four upper bounds for the steepness parameter of two stock-recruitment functions. Therefore, the reference set consists of 24 components.

Generally, one should always ensure that the final choice of OMs in the reference set covers a sufficiently representative range of potential estimates of current population status and productivity (Cooke, 1999). The most uncertain parameters in terms of population productivity tend to be the steepness parameter of the recruitment relationship and natural mortality. OMs, therefore, should cover the full range of plausible values for these parameters. In applications to whale populations, alternative hypotheses about population structure generally score high for inclusion in the reference set (IWC, 2000, Section 10; IWC, 2004a; Danielsdottir *et al.*, 2006). Recent debates (e.g. BCLME, 2006) indicate that such considerations may become more important in future revisions of MPs for South African hake as well as sardine (*Sardinops sagax*) and anchovy (*Engraulis encrasicolus*) fisheries.

Once a sufficiently promising reference set has been agreed, a wider range of robustness-test scenarios needs to be identified (Cooke, 1999). Such scenarios reflect the true dynamics that may vary more widely and be less plausible or have less impact than those included in the reference set. The types of robustness tests may differ among fisheries, but as a guideline, experience suggests that they could include different hypotheses on:

- *past data*: bias in survey estimates of absolute abundance (IWC, 2003), undetected trends in catch efficiency affecting catch per unit effort (cpue) (Rademeyer, 2003), or errors in catch statistics (Punt and Smith, 1999);
- *future availability of data*: the consequences of anticipated resource-monitoring data not becoming available (Geromont and Butterworth, 1998);
- *resource dynamics*: different forms of the stock–recruitment relationship (Punt and Smith, 1999) and incorporation of spatial structure or species interactions;
- *the environment*: changes in productivity/recruitment levels (Johnston and Butterworth, 2005) or carrying capacity (Rademeyer, 2003);
- *dynamics of the fishery*: changes in selectivity-at-age.

Currently, few ecosystem models can be applied reliably in a management context, so are able to serve as OMs to take account of the ecosystem effects of fishing. As a first step, proxies may be included in robustness tests to mimic ecosystem effects (such as time-dependent changes in natural mortality or carrying capacity to reflect an increased abundance of an important predator).

Because of time constraints, a practical suggestion is to perform initial evaluations of alternative CMPs only for those robustness tests yielding the most contrasting results from those of the reference case/set (Butterworth and Punt, 1999). For example, in the current revision of the MP for the South African hake fishery, the resource was projected forward under a constant catch for a fixed period for each of 28 different robustness tests initially identified (Rademeyer and Butterworth, 2006b). After comparing performance statistics, it was agreed to discontinue tests for a subset that provided results very similar to those of the reference set. However, before the ultimate MP is selected, a final check is desirable to confirm that this MP indeed performs robustly across the full set of robustness scenarios.

In principle, the performance of each CMP should be integrated over all possible scenarios considered, with relative weights assigned to the output statistics, to account formally for the relative likelihoods of the hypotheses postulated (Butterworth *et al.*, 1996; McAllister and Kirchner, 2002). These relative weights may prove important in evaluating the results of CMPs because they can affect the balance in choosing the appropriate trade-off between higher catches and lower risk of unintended depletion of the resource.

Such integration is helpful because a consistent evaluation of the results for individual tests is an onerous and difficult task. Overall performance statistics across the entire set of trials aid managers in making the final selection among candidates. By giving plausibility weights to all trials, their performance statistics can be ranked simply (of course, this requires that OMs are sufficiently similar; e.g. performance statistics for OMs ranging across different numbers of populations would be difficult to combine meaningfully). However, integrated statistics may obscure differences in expected performance between MPs by “hiding” low plausibility tests in tails of the distribution where they receive little attention. If performance deteriorates appreciably under specific scenarios for some MPs, this should be brought to the attention of decision-makers before they make their choice.

To assign weights to all tests, formal methods (such as likelihood-based or Bayesian methods) can rarely be used. One practical approach that has been used to reach agreement on numerical weights is a Delphi-like method in which committee members independently table their initial suggestions, and consensus is then sought through subsequent debate (CCSBT, 2003; Johnston and Butterworth, 2005). IWC (2004b) has opted for a high, medium, and low plausibility classification rather than a numerical weighting, discarding the low-plausibility trials and assigning a risk standard that is lower for medium- than for high-plausibility trials.

Defining a CMP

Empirical vs. model-based MPs

The estimator is the population-model-based framework within which the data obtained from the fishery are analysed and the current status and productivity of the resource is assessed. Related outputs are then fed into the HCR to provide a recommendation for management action. The combination of estimator and HCR provides the feedback mechanism within the MP. Let us assume that the management action refers to the setting of TAC. If the monitoring data derived from the OM turn out to indicate a stock status that is worse than that predicted the previous year, the new assessment coupled to the HCR will

recommend a lower TAC (and vice versa). Hence, the MP is able to self-correct over time, even if some assumptions made in developing a “best assessment” (typically corresponding to the OM given most weight) were wrong (Butterworth and Punt, 1999).

An MP of this estimator-plus-HCR type incorporates estimation of the status of the resource through the use of some population model and is referred to as model-based. The model may be an age-structured population model or an age-aggregated production model (e.g. Fox, 1970) fitted to relative or absolute abundance data (IWC, 1994; Geromont *et al.*, 1999). Examples of HCRs that convert outputs from the estimator into a recommended management action are constant fishing mortality/effort approaches (Kell *et al.*, 2006). More conservative approaches reduce catch to zero if abundance is estimated to drop below some threshold, as, for example, in the revised MP (RMP) for baleen whales (IWC, 1994).

In contrast, MPs can also be constructed, which are “model-free” (data-based, empirical) and which provide TAC recommendations directly (rather than through a two-stage model-based process), for example, through appropriate feedback in the form of recent upward or downward trends in abundance indices. HCRs for both model-based and empirical MPs typically include several free parameters that can be adjusted to tune their performances to achieve the desired balance among performance statistics over the range of scenarios simulated (discussed subsequently).

Which approach to select?

Model-free approaches are typically simple to develop and easily understood by stakeholders (such as the industry). Moreover, they require relatively little computer power for testing (because no iterative minimization routines are required for fitting models to data) and consequently allow for many simulations to be performed quickly (McAllister *et al.*, 1999). This approach has been used in the interim MP developed by Butterworth and Geromont (2001) for the Namibian hake resource, whose aim was to provide TAC recommendations that would perform well (in terms of catch and risk of resource depletion) across the wide range of possible levels of status of the resource argued for at that time. The inputs were measures of the recent trend (relative change over five years) in survey and cpue abundance indices. The first MP developed for the South African west coast rock lobster was also empirical (Johnston, 1998). However, with the longer time-series of cpue and survey data available in 2000, the RMP for that species was able to move to a model-based approach that allowed more data to be considered, and hence produced reduced variance (Johnston and Butterworth, 2005).

McAllister *et al.* (1999) suggest that the performance of model-free estimators may not prove entirely satisfactory in the long term, particularly if there are large uncertainties about bias and variance in the abundance index used. However, such estimates can provide good results if abundance estimates are in absolute numbers and if associated errors are small. An MP developed for Namibian seals was based directly on triennial aerial counts of pup production, providing estimates with relatively small coefficient of variations (CVs) (Butterworth *et al.*, 1998).

Although the empirical approach may move a resource in the desired direction (such as reversing a declining trend in an over-exploited population), it has the disadvantage that information on the level at which resource abundance will eventually equilibrate is lacking. Therefore, if the management objective is to drive resource abundance to a level at which it provides MSY

(MSYL), one can never be sure whether an empirical MP might stabilize abundance below MSYL (so forfeiting higher cpue) or above MSYL (so sacrificing potential catches).

Although empirical MPs have the advantage of simplicity, population-model-based approaches tend to perform better, for instance, in terms of less interannual variability in TACs (Butterworth and Punt, 1999; Punt and Smith, 1999). This outcome can be explained by the tendency of empirical approaches to estimate short-term trends, considering only data for the most recent years; in contrast, population-model-based MPs reflect the behaviour of the resource over much longer periods, and hence exhibit less variability in forecasts.

Importantly, the objective in choosing a particular model for use as an estimator in an MP is not to achieve a high degree of realism, but rather to ensure, in combination with a suitable HCR, good management performance (Cooke, 1999). Estimators based on simple population models have often been shown to perform as well or better than those based on more complex ones (Punt, 1993; Punt and Smith, 1999). Comparing MP performances based on an age-aggregated production model and on an *ad hoc* tuned Virtual Population Analysis (VPA) for the South African west coast hake resource, Punt (1993) showed that the former performed better, although it lacked the realism of the underlying dynamics of the OM. The latter MP tended to follow noise rather than signal, resulting in an increased interannual variation in TACs without being compensated by any gains in other performance statistics.

Sometimes, however, more complex estimators prove necessary. For example, for species where year-class strength is important, a model accounting for age structure may be required (Cooke, 1999). Rademeyer (2003) developed an MP for the Namibian hake resource in which the estimator was an age-structured production model of the same structural form as used for the OM. However, only two parameters (the pre-exploitation spawning biomass and the steepness parameter of the stock–recruitment relationship) were estimated annually in the simulation tests, whereas natural mortality, selectivity parameters, and variance, in addition to the sampling CV in survey estimates of abundance and historical stock–recruitment residuals, were all fixed, so would not always equate to those used in the OM in any particular trial. Fixing certain parameter values was done not only to avoid the high variance that may result from multiparameter estimation in the face of limited data, but also to keep computing time within reasonable bounds. In this instance, a problem arose with multimodality of the likelihood functions associated with an age-aggregated production-model approach.

In summary, experience suggests that it is useful to investigate the performance of both types of approach, at least in the early phase of the development process, but to use simplicity as a guiding principle in making the final choice(s). A helpful diagnostic is whether the estimators (or, typically, the trend lines estimated for empirical approaches) provide reasonable fits to the simulated data—if not, more complexity is usually required. A further guideline (if the CMPs allow) is to experiment first with generating noise-free future data: if a CMP does not perform reasonably for “perfect” data, it certainly should not be retained for further testing against more realistic noisy-data situations.

Other aspects

The form of the MP to be tested determines the required input data. Their nature, quantity, and statistical properties have to be

specified clearly (Cooke, 1999). Typical future data assumed to be available as input are indices of absolute and relative abundance (e.g. cpue from the fishery and biomass estimates from fishery-independent surveys, possibly by age or size group). In simulated projections, the variances used to generate future data are typically set to those estimated from the fits of the model to past data. Changes in estimates of resource status and productivity, and hence in future TACs, depend on these future data.

Further constraints may be applied to the initial output from the MP, such as limitations on the maximum permissible change in TAC from one year to the next (Geromont *et al.*, 1999; Kell *et al.*, 2006), in the interests of greater industry stability. Inputs from industry are desirable to guide the selection of such control-parameter values.

Evaluating CMPs

The performance of each CMP has to be evaluated through simulation over a range of scenarios. This is achieved by projecting the biomass forward under the prescribed HCR for a period defined by the longevity of the resource (typically 10–20 years for fish, but longer for long-lived resources with slow dynamics, such as whales). The performance is then assessed by inspecting the values of a set of performance statistics developed to measure the different management objectives predefined by decision-makers. The candidate providing the best trade-off between performances for what are often conflicting objectives is selected as the most appropriate.

For each MP/scenario combination, multiple replicate projections (typically 100 or more) have to be run to account for stochastic effects, particularly those resulting from observation errors and process errors (Butterworth and Punt, 1999; Cooke, 1999). To this end, random noise is generated when simulating future data and process effects, such as variations about stock–recruitment relationships. Typically, coefficients of variation will be set to values estimated from historical data.

Performance statistics

The objective of many fishery management policies is to balance three conflicting objectives: high annual catches, low risk of unintended depletion, and maximum industrial stability (Punt, 1993). Social objectives, such as fairness and equity in resource allocation among users, might also be considered. The performance statistics used in the evaluation have to be specified in relation to these three objectives:

- *catch-related*: an obvious choice is the average catch obtained over the projection period;
- *stability-related*: the average annual variation (AAV) in TAC from one year to the next (usually expressed as a proportion of the average annual catch) gives an indication of the extent of industrial stability to be anticipated;
- *risk-related*: commonly used statistics include the probability of depleting the (spawning-stock) biomass below some threshold or the median biomass expected at the end of the simulation period (compared with the biomass at the onset of this period).

In some cases, sufficient reliable information might be available for economic performance measures to be used, such as current net value and number of loss-making years (Holland *et al.*, 2005).

Communication of results

Performance statistics should be chosen to relate readily to the fishery, so that they are meaningful to managers and stakeholders (Francis and Shotton, 1997). Our experience is that statistics supplementary to those indicated above may be helpful, including projected cpue because this provides the industry with a simple proxy for economic performance.

When testing a particular CMP, once stochastic effects and model uncertainty across the reference set are accounted for, the resulting statistics represent distributions arising from alternative realizations. Results are frequently reported in the form of the medians and 95% quantiles of these distributions. However, tabulations of such statistics can be voluminous and difficult to interpret, even for scientists with experience in this area, so the form of plot shown in Figure 1 has become widely used to summarize such results conveniently. The example shown compares different CMPs for a single trial, but this format is also useful for comparing the performances of the same MP across several trials. It is always useful, as a reality check, to include the results for a zero catch and for a constant-catch or constant-effort strategy for comparison: CMPs warranting consideration should always perform better in terms of all objectives except variation in catch or effort. Plots of some individual trajectories of projected catch or biomass (“worm plots”) are usually easier for stakeholders to understand

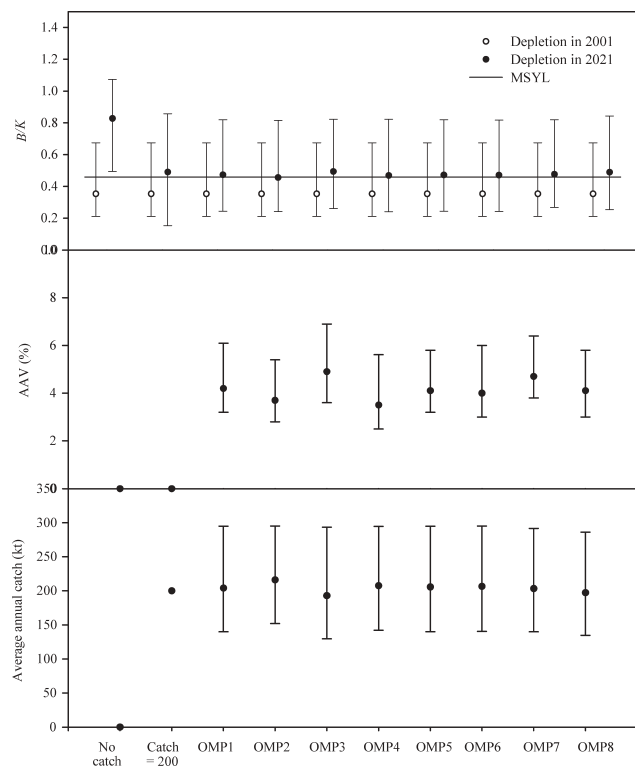


Figure 1. An example (in this case, for Namibian hake; from Rademeyer, 2003) of a widely used form of graphical summarization of the values of performance statistics for comparing across CMPs (in this case eight, OMP8 being the final choice; a no catch and a constant catch of 200 000 t option are included in the comparison): (a) initial (2001) and projected final (2021) depletion (B/K ; B , simulated biomass; K , virgin biomass) and level of biomass providing MSY (MSYL); (b) AAV in catch; (c) average annual catch. Bars show the 90% probability intervals.

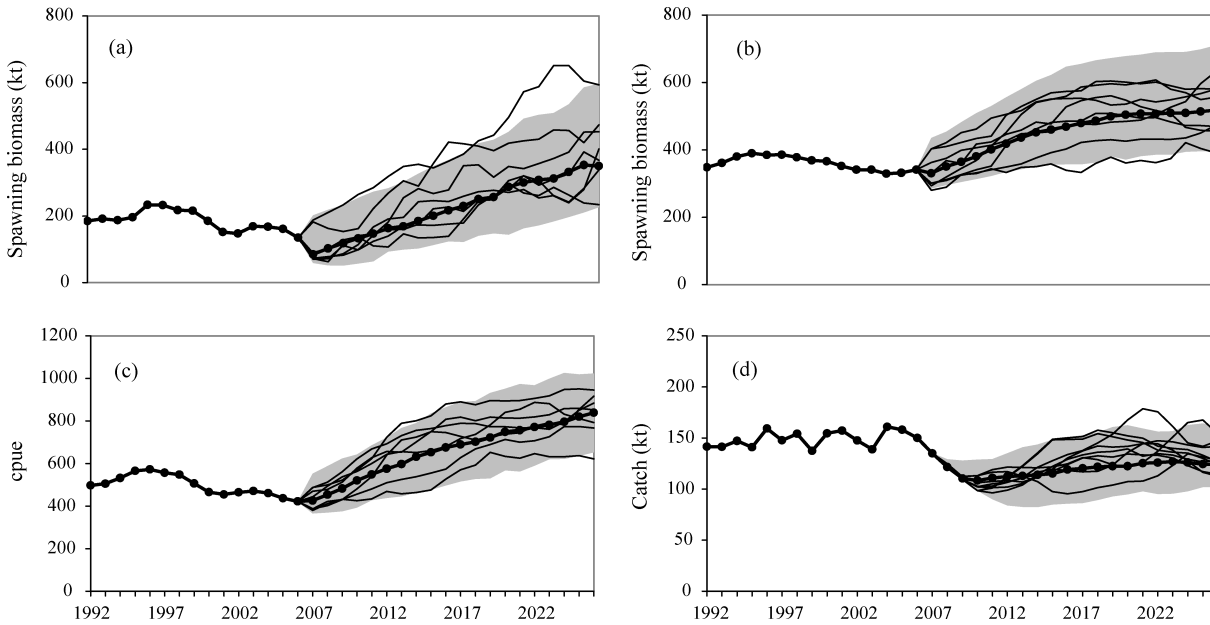


Figure 2. “Worm plots” showing ten possible trajectories of spawning biomass for the two hake species off South Africa: (a) *M. paradoxus* and (b) *M. capensis*; (c) offshore trawler cpue for species combined (proxied by exploitable biomass); and (d) total catch for one CMP and the reference set of OMs (from Rademeyer and Butterworth, 2006a). Annual medians (connected dots) and 90% probability envelopes (shaded) are also indicated.

than numerical statistics, and combining the two types of information in a single plot (Figure 2) may lead to better insight into the extent of the variability to be expected.

The primary purpose of computing performance measures is to permit stakeholders (and particularly the decision-makers, who have the responsibility for selecting the appropriate trade-off) to compare the different MP algorithms and their variants (Cooke, 1999). The challenge then is to summarize a large amount of information in a way as brief, meaningful, and stakeholder-friendly as possible. Even so, the choice may prove difficult, with many performance statistics to compare across many trials. The best way to undertake this process is case-dependent, but our experience is that a straightforward procedure is to select a few of the most important performance statistics (weighted across the scenarios comprising the reference set) and to use these in the comparison. The associated robustness trials then play the role of “tick tests”: to check (after an initial choice of MP has been made) that it does not result in an unacceptably large drop in performance for any of these trials (while accounting for their relative plausibilities). The approach of insisting that a minimum-risk standard be met for every trial (the so-called worst-case scenario management) is not recommended, because it fails to take due account of anticipated trade-offs with performance on meeting other objectives or of the relative plausibilities of different trials.

Summary

The steps necessary in developing an MP are shown in Figure 3. We suggest the following guidelines for the simulation-testing process:

- If uncertainties in the resource assessment are large, the construction of a reference set of OMs is preferable to the use of a single reference case OM. CMPs are then tuned to secure

the desired trade-offs. Work should focus first on developing CMPs that perform satisfactorily for the reference set.

- Initial evaluations of CMPs should focus on robustness tests against OMs, demonstrating the widest difference in resource behaviour from the reference set.
- The basis for selecting the final MP among CMPs has to be clear to all stakeholders and should be made as simple as can be justified. A useful approach is to focus on a few key performance statistics whose results are combined over all OMs included in a reference set, after appropriate weighting by their relative plausibilities.
- It is always useful to compare performances for both empirical and model-based MPs, but the latter, when based on an

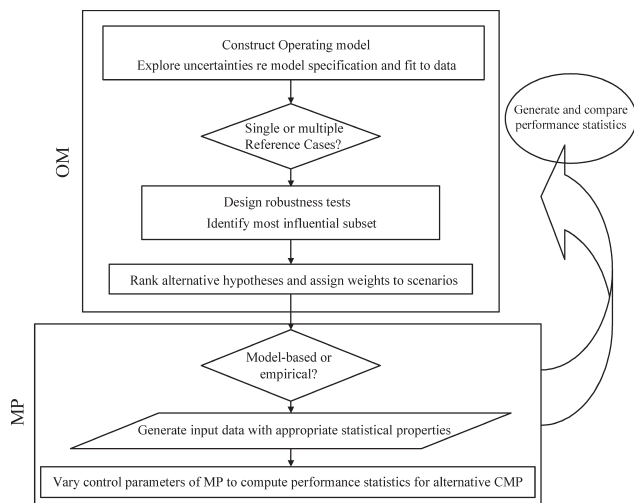


Figure 3. Flowchart to guide the MP development process.

age-aggregated population model, often prove a prudent choice.

- The performance statistics chosen to aid a selection among CMPs need to be meaningful to all stakeholders, and careful thought needs to be given on how best to present these to permit easy comparison.

Acknowledgements

The manuscript was prepared with funding support from the South African National Research Foundation and the Marine and Coastal Management Section of the Department of Environmental Affairs and Tourism. We also thank Trevor Branch, Anabela Brandão, Carryn Cunningham, and Susan Johnston of the MARAM group at the University of Cape Town for their input. Further, we thank Beth Fulton and an anonymous reviewer, as well as Greg Donovan and André Punt, for their constructive comments. In addition, Kevern Cochrane, Phil Hammond, Laurie Kell, Ana Parma, Tony Smith, and Niels Daan contributed in the development of the Glossary.

References

- BCLME. 2006. Report of the Joint Hake Research Planning Workshop, Cape Town, 9–12 May 2006. Benguela Current Large Marine Ecosystem Programme, Windhoek. www.benefit.org.na/whatsnews/Draft%20REPORT%20OF%20THE%20JOINT%20HAKE%20RESEARCH%20PLANNING%20WORKSHOP.doc.
- Butterworth, D. S. 1999. Questions that need to be addressed when specifying Implementation Simulation Trials. *Journal of Cetacean Research and Management*, 1(Suppl.): 61–116.
- 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., and Geromont, H. F. 2001. Evaluation of a class of possible Interim Management Procedures for the Namibian hake fishery. *South African Journal of Marine Science*, 23: 357–374.
- 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.
- 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. Report of the International Whaling Commission, 46: 637–640.
- CCSBT. 2003. Specification of the final reference set. Attachment D of the Report of the Fourth Meeting of the Stock Assessment Group, 25–29 August 2003, Christchurch, New Zealand. Commission for the Conservation of Southern Bluefin Tuna. 36 pp.
- Cooke, J. G. 1999. Improvement of fishery-management advice through simulation testing of harvest algorithms. *ICES Journal of Marine Science*, 56: 797–810.
- Danielsdottir, A. K., Gunnlaugsson, Th., Ólafsdottir, D., and Víkingsson, G. A. 2006. Stock structure hypotheses for North Atlantic fin whales. Appendix 4 of the Report of the Sub-Committee on the Revised Management Procedure (RMP). *Journal of Cetacean Research and Management*, 8(Suppl.): 78–90.
- 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.
- 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.
- Fox, W. W. 1970. An exponential surplus-yield model for optimizing exploited fish populations. *Transactions of the American Fisheries Society*, 99: 80–88.
- Geromont, H. F., and Butterworth, D. S. 1998. Management procedure robustness trials for west coast hake. *Marine and Coastal Management South Africa Document WG/06/98/D:H:36*. 13 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.
- Holland, D. S., Bentley, N., and Lallemand, P. 2005. A bioeconomic analysis of management strategies for rebuilding and maintenance of the NSS rock lobster (*Jasus edwardsii*) stock in southern New Zealand. *Canadian Journal of Fisheries and Aquatic Sciences*, 62: 1553–1569.
- IWC. 1994. The Revised Management Procedure (RMP) for baleen whales. Report of the International Whaling Commission, 44: 145–152.
- IWC. 2000. Report of the Sub-Committee on the Revised Management Procedure. *Journal of Cetacean Research and Management*, 6(Suppl.): 79–124.
- IWC. 2003. Fishery type 2: Implementation for Bering-Chukchi-Beaufort bowhead whales. Appendix 3(i) of the Report of the Standing Working Group (SWG) on the Development of an Aboriginal Subsistence Whaling Management Procedure (AWMP). *Journal of Cetacean Research and Management*, 6(Suppl.): 155–225.
- IWC. 2004a. North Pacific minke whale Implementation Simulation Trial specifications. Appendix 10 of the Report of the Sub-Committee on the Revised Management Procedure. *Journal of Cetacean Research and Management*, 6(Suppl.): 75–184.
- IWC. 2004b. Decision guidelines for selecting North Pacific minke RMP Implementation options. Appendix 4 of the Report of the Sub-Committee on the Revised Management Procedure. *Journal of Cetacean Research and Management*, 6(Suppl.): 75–184.
- IWC. 2005. Recommended schedule for an Implementation and subsequent Implementation Reviews. Adjunct 1 of Appendix 2: Requirements and guidelines for Implementations of the Report of the Sub-Committee on the Revised Management Procedure. *Journal of Cetacean Research and Management*, 7(Suppl.): 77–113.
- 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, South Africa. 875 pp.
- Johnston, S. J., and Butterworth, D. S. 2005. Evolution of operational management procedures for the South African West Coast rock lobster (*Jasus lalandii*) fishery. *New Zealand Journal of Marine and Freshwater Research*, 39: 687–702.
- Kell, L. T., Pilling, G. M., Kirkwood, G. P., Pastoors, M. A., Mesnil, B., Korsbrekke, K., Abaunza, P., et al. 2006. An evaluation of multi-annual management strategies for ICES roundfish stocks. *ICES Journal of Marine Science*, 63: 12–24.
- McAllister, M. K., and Kirchner, C. 2002. Accounting for structural uncertainty to facilitate precautionary fishery management: illustration with Namibian orange roughy. *Bulletin of Marine Science*, 70: 499–540.
- McAllister, M. K., Starr, P. J., Restrepo, V. R., and Kirkwood, G. P. 1999. Formulating quantitative methods to evaluate fishery management systems: what fishery processes should be modelled and what trade-offs should be made. *ICES Journal of Marine Science*, 56: 900–916.
- 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 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., 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.
- Rademeyer, R. A. 2003. Assessment of and management procedures for the hake stocks off southern Africa. MSc thesis, University of Cape Town, South Africa. 209+xxxvi pp.
- Rademeyer, R. A., and Butterworth, D. S. 2006a. Summary of the most recent South African hake assessments. BCLME Joint Hake Research Planning workshop, Cape Town, South Africa, Document BCLMEHW/MAY06/7.1.
- Rademeyer, R. A., and Butterworth, D. S. 2006b. Management procedure robustness tests for the South African *Merluccius paradoxus* and *M. capensis* resources under a constant catch scenario. Marine and Coastal Management South Africa Document WG/03/06/D:H:6. 26 pp.
- Smith, A. D. M., Fulton, E. J., Hobday, A. J., Smith, D. C., and Shoulder, P. 2007. Scientific tools to support the practical implementation of ecosystem-based fisheries management. ICES Journal of Marine Science, 64: 633–639.
- Smith, A. D. M., Sainsbury, K. J., and Stevens, R. A. 1999. Implementing effective fisheries-management systems—management strategy evaluation and the Australian partnership approach. ICES Journal of Marine Science, 56: 967–979.

Appendix

Glossary of common terminology used in conjunction with “MPs”

[NB: some of the terms (e.g. assessment, harvest strategy, management strategy) are accorded a wider range of meanings when used in a broader fisheries context; note further that when definitions use other terms defined elsewhere in this Glossary, these terms are shown in italics to ease cross-reference.]

AAV: Average annual variation in a *TAC* from one year to the next (expressed as a proportion of the average annual catch); this *performance statistic* is often used to measure the attainment of an *objective* related to minimizing catch variability.

Assessment: A mathematical population model coupled to a statistical estimation process that integrates data from a variety of sources to provide estimates of reference points and past and present abundance, fishing mortality, and productivity of a resource.

CMP: Candidate *MP*—one of a set of *MPs* under consideration for *implementation* to manage a resource.

Conditioning: An *OM* is “conditioned” on available information by adjusting the parameter values to ensure that it is consistent with this information, and hence reflects assumptions that are plausible—this process is similar (sometimes identical) to an *assessment*; the *conditioning* provides the *initial conditions* for projecting resource dynamics forward.

Error: Differences, reflecting uncertainties, between the actual dynamics of the resource (described by the *OM*) and perceptions arising from observations and assumptions. Four types of error may be distinguished, and *simulation trials* may take account of one or more of them:

- *Estimation error:* differences between the actual values of the parameters of the *OM* and those provided by the *estimator* when fitting a model to available data;
- *Implementation error:* differences between intended limits (as output by an *MP*) and those actually achieved;

- *Observation error (or measurement error):* differences between the measured value of some resource index and the corresponding actual value in the *OM*;
- *Process error:* natural variations in resource dynamics or systematic errors in outputs from an *estimator* arising from the use of a parameter value or model structure different from that of the *OM*.

Estimator: The statistical estimation process within an *assessment*; in a *MP* context, the component that provides information on resource status and productivity from past and generated future resource-monitoring data for input to the *HCR*.

Feedback control: Rules or algorithms based directly or indirectly on trends in observations of resource indices, which adjust the values of management measures such as *TACs* in directions intended to reverse inferred trends in abundance away from the target level reflecting decision-makers’ *objectives*.

FLR: Fisheries Library in R, a generic toolbox that can be used to construct *OMs* for *MSE*.

Generic MP: An *MP* that has been tested for potential use for a wide range of resources (e.g. the single-stock component of the IWC’s *RMP*), as distinct from a case-specific *MP* tested using *OMs conditioned* on data for a specific resource.

Harvest strategy: Intended meaning may be synonymous with *MP*, *MP (implicit)*, *HCR*, or *HCR+assessment*; in the last case, the *assessment* method may change at each application rather than remain fixed as for an *MP*.

HCR: Harvest control rule (also termed *harvest control law*)—a set of well-defined rules used for determining a management action in the form of a *TAC* or allowable fishing effort given input from an *estimator* or directly from data.

Implementation: The process of testing followed by practical application of an *MP* to provide resource management recommendations.

MP: Management procedure—the combination of pre-defined data, together with an algorithm to which such data are input to provide a value for a *TAC* or effort control measure; this combination has been demonstrated, through *simulation trials*, to show robust performance in the presence of uncertainties. Additional rules may be included, for example to spread a *TAC* spatially to cater for uncertainty about stock structure. Two types of *MP* may be distinguished:

- *Empirical MP:* An *MP* where resource-monitoring data (such as survey estimates of abundance) are input directly into a formula that generates a control measure such as a *TAC* without an intermediate (typically population-model based) *estimator*;
- *Model-based MP:* An *MP* where the process used to generate a control measure such as a *TAC* (this process is sometimes termed a catch limit algorithm or *CLA*) is a combination of an *estimator* and an *HCR*.

MP approach: Management of a resource using a fully specified set of rules incorporating *feedback control*; the approach is explicitly precautionary through its requirement for *simulation trials* to have demonstrated robust performance across a range of uncertainties about resource status and dynamics.

MSE: Management strategy evaluation—usually synonymous with *MP approach*; also often used to describe the process of testing *generic MPs* or *harvest strategies*.

MP (implicit): A set of rules for management of a resource that contains all the elements of an *MP*, but has not yet been evaluated through *simulation trials*.

Management strategy: Usually synonymous with *MP* but some authors use it to mean an *HCR*.

Observation model: The component of the *OM* that generates fishery-dependent and/or fishery-independent resource monitoring data for input to an *MP*.

Objectives: General goals for managing a resource as set by decision-makers—these often include the aims of maximizing catches, minimizing interannual changes in catch limits and the risk of unintended depletion of the resource and related species, and considerations of transparency and cost effectiveness.

OM: Operating model—a mathematical–statistical model used to describe the actual resource dynamics in *simulation trials* and to generate resource monitoring data when projecting forward.

OMP: Operational management procedure—analogue to an *MP*, except that this term is typically reserved to signify *MPs* that have actually been *implemented*, in contrast to the ones that are conceptual only.

Performance statistics: Statistics that summarize different aspects of the results of a *simulation trial* used to evaluate how well a specific *MP* achieves some or all of the general *objectives* for management for a particular *scenario*.

Plausibility: The likelihood of a *scenario* considered in *simulation trials* representing reality relative to other *scenarios* also under consideration; *scenarios* considered implausible (e.g. because of incompatibility with available data) are eliminated from the *simulation trials*.

Reference case: A single, typically central, *conditioned OM* for evaluating *CMPs* that provides a pragmatic basis for comparison with results of other *OMs*.

Reference set (also termed **base-case** or **evaluation scenarios**): a limited set of *scenarios*, with their associated *conditioned OMs*, which include the most important uncertainties in the model structure, parameters, and data, i.e. alternative *scenarios* which

have both high *plausibility* and major impacts on *performance statistics*.

Research-conditional option: Temporary application of an *MP* that does not satisfy conservative performance criteria, provided accompanied by both a research programme to check the *plausibility* of the *scenarios* that gave rise to this poor performance and an agreed subsequent reduction in catches should the research prove unable to demonstrate implausibility.

Robustness tests: *Tests* to examine the performance of an *MP* across a full range of *plausible scenarios*.

Scenario: A hypothesis concerning resource status and dynamics, represented mathematically as an *OM*.

Selection: The choice of an *MP* from a set of *CMPs* through comparing *performance statistics* from *tests* over a wide range of *scenarios*.

Simulation trial (or test): A computer simulation to project resource dynamics for a particular *scenario* forward for a specified period, under controls specified within an *MP*, to ascertain performance; such projections will typically be repeated a large number of times to capture stochasticity.

TAC: Total allowable catch (or catch limit) to be taken from a resource within a specified period.

Trade-offs: Comparisons of gains in some *performance statistics* against losses in others when selecting among *CMPs*; these *trade-offs* arise because some *objectives* for management conflict (e.g. maximizing catch vs. minimizing risk of unintended depletion).

Weights: Either qualitative (e.g. high, medium, low) or quantitative measures of relative *plausibility* accorded across a set of *scenarios*.

Worm plots: Plots showing a number of possible realizations of simulated projections of, for example, catch or spawning biomass under application of an *MP*.

doi:10.1093/icesjms/fsm050