K. Tsukamoto, T. Kawamura, T. Takeuchi, T. D. Beard, Jr. and M. J. Kaiser, eds.
Fisheries for Global Welfare and Environment, 5th World Fisheries Congress 2008, pp. 139–152.

Refocusing Stock Assessment in Support of Policy Evaluation

André E. Punt

School of Aquatic and Fishery Sciences
University of Washington
Box 355020, Seattle
WA 98195, USA

Email: aepunt@u.washington.edu

Fisheries stock assessment has traditionally been focused on providing the basis for tactical fisheries management advice. However, there is an increasing demand from decision makers for feedback control management strategies evaluated using Management Strategy Evaluation, MSE. MSE can be used to identify which uncertainties are likely to lead to poor performance of current management strategies and to which uncertainties such strategies are robust, and hence to the extent to which current management strategies need to be modified (if at all) to perform satisfactorily in the face of key uncertainties. Given this demand, there is a need for a refocus of the priorities of stock assessment science. In particular, more attention needs to be placed on exploring alternative model structures, including those that take spatial and multispecies considerations into account, while there is also a need to develop tools to assign weights or probabilities to alternative model structures using Bayesian and meta-analytic techniques. Finally, care needs to be taken not to overuse model selection methods when selecting model structures to use as the basis for the evaluation of management strategies.

KEYWORDS management strategies, meta-analysis, model selection, Monte Carlo simulation, simulation, stock assessment, uncertainty

1. Introduction

The objectives for conducting fisheries stock assessments during the 20th and early 21st centuries have been primarily tactical, such as providing management advice to decision makers regarding current stock size, in absolute terms, and relative to target, limit or threshold reference levels, as well as yields (current and long-term) according to harvest control rules (such as $F_{0.1}$). Other outputs from stock assessments have included information on trends in fishing mortality, and fishing mortality relative to reference points, such as F_{MSY}.

Fisheries stock assessment involves a wide variety of techniques ranging from catch curve analysis to estimate exploitation rates under the assumption that the population and fishery are in equilibrium, to the application of spatially-, age- and size-structured population dynamics models (such as CASAL (Bull *et al.* 2005); GADGET (Begley 2003; Stefansson 2003); MULTIFAN-CL (Hampton and Fournier 2001); and Stock Synthesis 2 (Methot 2005, 2007)) which estimate trends in a variety of measures of stock size and stock status (and their uncertainty). 'State-of-the-art' stock assessment continues to evolve to towards more complicated models that can use a broader range of data types in an increasingly statistically rigorous manner.

There is an increasing focus in stock assessment on quantifying and representing uncertainty, and including uncertainty in management advice, and the broad scope of assessment methods reflects this to some extent. Methods for representing uncertainty range include providing confidence intervals, calculating the probability of achieving various management goals, and showing the sensitivity of key model outputs to changing some of the assumptions of the assessment model, its data inputs and the emphasis placed on fitting different data inputs. One particularly common method for representing uncertainty in stock assessments and hence the consequences of future management actions is the decision table (see, for example, Table 1) which highlights the implications of choosing one set of assumptions (or state of nature) as the basis for management advice when another set of assumptions is true.

Although the information in a decision table has the potential to provide decision makers with the ability to evaluate the consequences of basing management decisions on one or other sets of assumptions, management decisions are nevertheless usually based on the "central" (or most likely) set of assumptions. This is often because assessment authors are unable (or unwilling) to assign probabilities to states of nature. Moreover, although management is based on a feedback control system, i.e. decisions are made, data are collected, and assessments are updated, there is no consideration of the implications of feedback in decision tables. Rather, decision tables are commonly based on the assumption that pre-specified

Table 1. A decision table for black rockfish (*Sebastes melanops*) off California and Oregon (Sampson 2007). The entries indicate the spawning biomass from 2009 under a variety of catch series. The three states of nature reflect assumptions regarding values for sex-specific natural mortality and the sizes of the historical catches. The high, medium and low catch series are based on applying the $F_{50\%}$ harvest strategy (Ralston 2002) to the high, medium, and low states of natures respectively.

Year	Catch	State of nature		
		Low	Medium	High
Low catches				
2009	909	2195	3284	5710
2010	831	2099	3168	5518
2011	782	1981	3015	5258
2012	765	1860	2855	4982
2013	772	1756	2714	4737
2014	789	1683	2614	4555
Medium catches				
2009	1454	2195	3284	5710
2010	1303	2007	3077	5428
2011	1203	1804	2844	5092
2012	1156	1612	2616	4753
2013	1146	1450	2422	4458
2014	1153	1329	2277	4237
High catches				
2009	2660	2195	3284	5710
2010	2333	1802	2876	5231
2011	2112	1416	2467	4726
2012	1994	1072	2096	4252
2013	1945	796	1791	3854
2014	1930	583	1557	3551

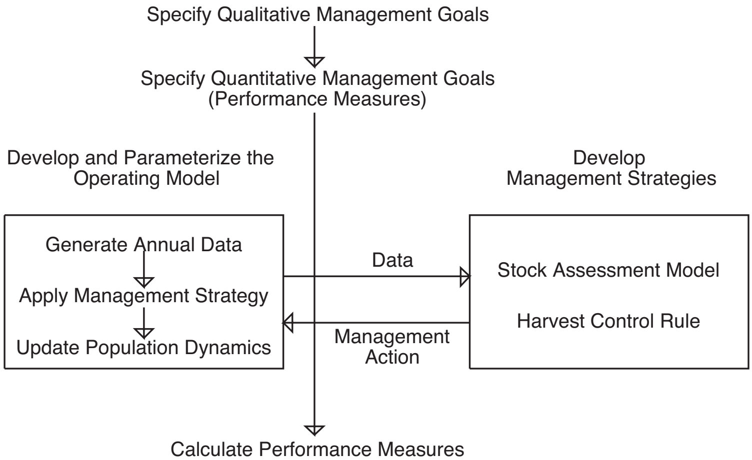

Fig. 1. Flowchart of the approach used to evaluate management strategies.

sequences of catches will be followed irrespective of the data that will be collected from the fishery in the future.

In cases in which management decisions are based on well-specified harvest control rules, it is possible to broaden the notion of stock assessment so that fuller account can be taken of uncertainty and the implications of feedback control. The Management Strategy Evaluation (MSE) approach is the use of simulation to evaluate the performance of combinations of stock assessment methods and harvest control rules (management strategies) in the face of uncertainty given (pre-agreed) objectives (Smith 1994; Smith *et al*. 1999; Kell *et al*. 2006). However, a focus on MSE rather than tactical management advice will require a change to the focus of stock assessment science. This paper first outlines MSE and then discusses how conventional stock assessment approaches will need to be modified so that they can be used to form the basis for MSE. The concepts outlined in this paper are illustrated using experiences from the International Whaling Commission, and the Pacific and North Pacific Fishery Management Councils as well as a simple illustrative situation (Appendix).

2. Management Strategy Evaluation

The MSE approach involves developing models that capture the entire management system, including the underlying fish population dynamics, the data collection scheme, the method of stock assessment used when providing management advice, and any harvest control rules. The MSE approach differs from earlier methods for evaluating management performance by explicitly including the impact of noise in the assessment data and by examining the use of incorrect assumptions when conducting stock assessments. MSE is considered to be state-of-the-art for evaluating and contrasting management strategies, and is fully consistent with the FAO precautionary approach to fisheries management (Punt 2006; Marasco *et al*. 2007). The steps followed when applying the MSE approach are (see also Fig. 1):

1) Identification of the management objectives and representation of these using a set of quantitative performance measures.
2) Development and parameterization of a set of alternative structural models (called operating models) of the system

under consideration; each of these operating models represents an alternative, yet nevertheless plausible, representation of reality.

3) Identification of the alternative management strategies which have the potential to satisfy the management objectives.
4) Simulation of the future use of each management strategy to manage the system (as represented by each operating model scenario) under feedback control. Each simulation trial usually involves about 100 replicates for a particular operating model specification. The following three steps occur for each year of the projection period.
   - Generation of the types of data available for assessment purposes.
   - Application of the method of stock assessment and the harvest control rule to determine a management measure for the next year (usually a Total Allowable Catch or Total Allowable Effort).
   - Determination of the implications of this management measure by setting the catch for the next year from the 'true' population represented in the operating model based on it.
5) Summary of the results of the simulations by means of the performance measures.

Conventionally, MSE, and hence the objectives and performance measures, has pertained to the direct effects of fisheries for target species (yield, stock status, and catch variability) (Butterworth *et al*. 1997; Punt and Smith 1999; Dichmont *et al*. 2006; Punt 2006), although, recently, it has also been proposed as the basis to evaluate whole-of-ecosystem management strategies and objectives (e.g. Sainsbury *et al*. 2000; Marasco *et al*. 2007; Fulton *et al*. 2007), including non-use objectives (Mapstone *et al*. in press). Equally importantly, MSE is considered a way to evaluate the implications of error and uncertainty when conducting stock assessments and applying harvest control rules. Specifically, attempts are usually made when applying MSE to account for all major sources of uncertainty (*cf*. Francis and Shotton 1997), *viz*.

- **Process error**: variability in population dynamic processes such growth, recruitment and natural mortality due to unpredictable environmental factors.
- **Observation error**: variability in the data on which stock assessments (or management strategies more generally) are based.
- **Model error**: structural differences between the operating model and the models on which stock assessments are based.
- **Implementation error**: error that arises from an inability to adequately implement and enforce regulations.

Specification of the values for the parameters of the operating model ("conditioning" the operating model) is generally based on fitting it to existing data. However, this may lead to an overly restrictive set of operating model scenarios and it is often the case that operating model scenarios are also developed based on expert beliefs and other *a priori* information to reflect scenarios, while not the most likely, that could nevertheless occur in the future (Kell *et al*. 2006).

3. Implications of MSE for Stock Assessment Science

3.1. Multispecies and spatial models

Almost all of the population dynamics models on which fisheries stock assessments are based assume that the population being assessed constitutes a single homogenous population while only very limited account (if any) is taken of predation mortality when providing management advice. Also, it is well known that that conservation performance can be poor when multiple stocks are assessed and managed together, but the harvest is not proportional to stock size or the intrinsic rate of growth (and hence productivity) differs among stocks. The poor performance of the IWC's Revised Management Procedure in the face of stock structure uncertainty (e.g. IWC 1992, 1993) was one reason for the development by the Scientific Committee of the International Whaling Commission (the IWC SC) of operating models that are specifically designed to examine the implications of stock-structure uncertainty, and the considerable focus in the IWC SC in recent years on developing processes for identifying the "full range" of "plausible" hypotheses, particularly those related to stock structure uncertainty (Punt and Donovan 2007).

The values for the parameters of the operating models developed by the IWC SC used when evaluating candidate management strategies are estimated, and their uncertainty quantified, by fitting them to data on absolute abundance by modeled region, proportions of different stocks in different areas based on analyses of (for example) genetic data, estimates of total mortality based on inferences from age-composition data, and movement rates based on tagging data (e.g. IWC 2004, 2007). Other studies of the performances of management strategies have considered the implications of spatial structure (e.g. Bentley *et al.* 2003; Punt *et al.* 2005; SC-CAMLR 2006; Hobday and Punt 2006), but the number of applications remains very low.

Given that the performance of management strategies can be sensitive to how spatial structure is treated, it is perhaps surprising that the number of spatially-structured stock assessment models which could form the basis for spatially-structured operating models is still relatively small. Examples, outside of the IWC of spatially-structured stock assessment models include those developed for hoki (*Macruronus novaezelandiae*) in New Zealand (e.g. Francis *et al.* 2003) and school shark (*Galeorhinus galeus*) in Australia (Punt *et al.* 2000). Assessment platforms such as CASAL, GADGET, and SS2 all allow spatial-structure to be represented, but this feature has not been used extensively to date in actual applications.

Several assessment models have been developed that consider multispecies interactions (e.g. Magnússon 1995; Begley 2003; Jurado-Molina *et al.* 2005). However, their use for management advice has been limited, although they clearly have a role to play to when evaluating the robustness of management strategies (which usually ignore multispecies interactions). MSE analyses based on operating models which explicitly include biological multispecies interactions are rare (Schweder *et al.* 1998; Fulton *et al.* 2007 being notable exceptions). In contrast, an increasing number of evaluations of management strategies (e.g. De Oliveira *et al.* 1998; De Oliveira and Butterworth 2004; Punt *et al.* 2005; Dichmont *et al.* 2006; Fulton *et al.* 2007) have accounted for technical interactions.

Environmental variables are being included in an increasing number of stock assessments, for example as proxies for recruitment success (e.g. Maunder and Watters 2003; Schirripa 2007). In principle, if these environmental variables can be forecasted, based, for example, on downscaled IPCC predictions, the robustness of current management strategies to what is effectively non-stationarity in the parameters of the operating model can be evaluated.

Assessments that incorporate spatial structure, multispecies interactions, and environmental forcing will likely lead to less precise estimates of model parameters and estimates of quantities of management interest and so may not be appropriate as the basis for tactical fisheries management advice. However, considering such models routinely as part of assessments will allow two key questions to be addressed: (a) Are perceptions of stock status radically altered by including such factors, and (b) are the likely consequences of management actions in the form of management strategies likely highly sensitive to these sources of model uncertainty?

3.2. The role of parameter estimation and weighting of models

The ideal management action or management strategy for a particular case will depend on the goals of management and the range of scenarios considered when evaluating alternative actions and strategies. However, if the aim of evaluating management strategies is to identify those that are (sufficiently) robust to uncertainty, each operating model scenario (and perhaps the objectives themselves) needed to be weighted. Table 2(a) contrasts the expected yield under five exploitation rates for four scenarios related to the steepness of the stock–recruitment relationship (see Appendix for technical details). The first four exploitation rates were determined by maximizing the expected yield for each of the four steepness values in turn, and the fifth exploitation rate by maximizing the expected yield over all the four values of steepness after giving each value for steepness equal weight. The optimal strategy differs among steepness values and there can be a large loss in yield if the exploitation rate used to determine management actions is based on the optimal exploitation rate for one value of steepness when another is true. This is quantified in Table 2(a) by the difference between expected value of perfect information about steepness (when each steepness value is equally weighted) and the expected yield for each of the five exploitation rates. The ability to avoid losses due to uncertainty can be reduced (and poor decisions avoided) if (appropriately selected) weights can be assigned to each set of assumptions about the dynamics of the system. This is, however, not always the case. For example, Table 2(b) shows that the expected yield (but not necessarily risk-related performance measures) is insensitive to the extent of variation in recruitment.

Butterworth *et al*. (1996) proposed the following four-level scheme to assign “plausibility ranks” to the hypotheses underlying operating model scenarios:

1) How strong is the basis for the hypothesis in the data for the species or region under consideration;
2) How strong is the basis for the hypothesis in the data for a similar species or another region;
3) How strong is the basis for the hypothesis for any species; and
4) How strong or appropriate is the theoretical basis for the hypothesis?

Although this scheme is semi-quantitative, weights (if assigned at all) tend to be assigned to scenarios qualitatively in MSE using a ‘Delphi type’ approach. For example, the weighting scheme developed for use by the International Whaling Commission involves assigning operating model scenarios weights of ‘high’, ‘medium’, ‘low’ and ‘no agreement’ (with ‘no agreement’ being treated as ‘medium’) (IWC 2005). Nevertheless, stock assessment science has an important role to play in terms of the first two levels of this scheme. Specifically, given that appropriate care is taken to ensure that the likelihood function is valid for the problem at hand (many of the data sets used for stock assessment are grossly overdispersed given the probability distributions assumed for them), and operating model scenarios are only assigned low weight because the data suggest that other scenarios fit the data better (rather than the data having no ability to discriminate among different scenarios—see “Complex versus simple models” below), it should be possible to identify “unlikely” scenarios.

The development of Bayesian methods allows a focus on key (often structural) uncertainties by “integrating out” other sources of uncertainty (such as process error). In particular, basing projections on samples

Table 2. Relationship between exploitation rate, *E*, and expected yield for: (**a**) four scenarios related to stock-recruitment steepness, *h*, when $\sigma_R = 0.6$, and (**b**) four scenarios related to the extent of variation in recruitment, σ_R when $h = 0.7$. The five exploitation rates were chosen to maximize yield under each of four scenarios and when the results from each scenario are weighted equally. The shaded cells are those that were maximized to obtain the five exploitation rates.

(a) Stock-recruitment steepness (expected value of perfect information 70.9).

E	*h*				Expected
	0.9	0.7	0.5	0.3	
0.049	37.9	36.7	34.0	20.4	32.3
0.162	82.4	74.1	55.3	0.0	53.0
0.321	108.0	85.7	35.6	0.0	57.3
0.681	122.1	60.5	0.0	0.0	45.6
0.272	102.5	84.8	44.9	0.0	58.0

(b) Extent of recruitment variability (expected value of perfect information 86.8).

E	σ_R				Expected
	0.2	0.4	0.6	0.8	
0.364	90.9	88.7	85.2	80.0	86.2
0.346	90.8	88.8	85.5	80.7	86.4
0.321	90.4	88.6	85.7	81.4	86.5
0.294	89.5	88.0	85.4	81.7	86.1
0.328	90.5	88.7	85.7	81.3	86.5

from posterior distributions which represent parameter uncertainty reduces the number of scenarios which need to be examined, making to communication of results to decision makers substantially more straightforward and also reducing the number of scenarios that need to be weighted. Care should, however, be taken to avoid treating parameters which could have a substantial impact on the performance of alternative management strategies as estimable parameters that are automatically "integrated out" when constructing Bayesian posterior distributions. For example, the IWC SC based the analyses used to evaluate management strategies (Strike Limit Algorithms) for the Bering–Chukchi–Beaufort Seas bowhead whales (*Balaena mysticetus*) and the eastern North Pacific gray whales (*Eschrichtius robustus*) on samples from Bayesian posteriors, but based operating model scenarios on specific choices for the parameter that determines productivity to allow the impact of this parameter on performance to be represented explicitly (even though the data appeared to be very informative about this parameter).

The use of Bayesian stock assessment methods requires that prior probability distributions be developed for all of the parameters of the model and these priors updated using data for the case in question. Moreover, given that prior probabilities can be assigned to models, Bayesian techniques can be used directly to compute the relative weight that should be assigned to alternative models. Priors can be developed subjectively, by analyzing auxiliary information (e.g. Givens *et al.* (1995) for the Bering–Chukchi–Beaufort Seas bowhead whales) and using meta-analysis. Meta-analysis provides a formal basis for developing priors as well implementing step 3 of the approach to weighting developed by Butterworth *et al.* (1996). To date, meta-analyses based on data for multiple stocks have been conducted which could be used to assign priors to the parameters of the stock–recruitment relationship (Myers *et al.* 1995; Liermann and Hilborn 1997), the form of the relationship between catch-rate and abundance (Harley *et al.* 2001), and survey catchability (Harley and Myers 2001; Millar and Methot 2002). The further development of meta-analytic techniques and their application to fisheries data sets should provide a more rigorous (and replicable) basis for assigning weights to alternative models.

3.3. Complex versus simple models

One of the central tenants of contemporary fisheries stock assessment, and statistical modelling in general, is the desire to identify models that provide a parsimonious representation of the available data, using approaches such as likelihood ratio tests and the Akaike Information Criterion (Akaike 1974; Burnham and Anderson 2002). While simpler models may provide better short-term forecasts and, in fact, estimate biomass more accurately (Ludwig and Walters 1985; Punt 1993), ignoring models which include biological processes which are poorly supported by the data, but to which the performance of management strategies may be very sensitive may lead to the selection of a management strategy that is not sufficiently robust to uncertainty.

There are several areas where models that would not be supported by standard model selection approaches should not be ignored. For example, inclusion of depensation in operating models. Most fisheries stock and recruitment data sets are incapable of selecting between alternative two-parameter stock–recruitment relationships such as Ricker and Beverton–Holt. Not surprisingly, therefore, only very few data sets provide statistically significant evidence for depensation (Myers *et al.* 1995) although, as noted by Liermann and Hilborn (1997), stock and recruitment data are generally uninformative about the presence of depensation rather than suggesting a lack of depensation. Ignoring depensation in the absence of information supporting such lack can, however, markedly under-estimate risk. For example, Fig. 2 contrasts two stock–recruitment relationships which differ in terms of whether there is depensation. The solid lines in Fig. 2 are based on the standard Beverton–Holt stock–recruitment relationship ($\gamma = 1$ in Eq. (A.1)), while the dashed lines are based on a stock–recruitment relationship that has the same equilibrium point and recruitment at 20% of the equilibrium unfished spawning stock biomass ($0.2B_0$) as the solid line, but for which the recruitment at $0.1B_0$ is 0.25 of that of the standard Beverton–Holt stock–recruitment relationship. The two stock–recruitment relationships are similar (arguably indistinguishable over a broad range of spawning stock biomass levels given typical stock and recruitment data sets), but have profound implications for both the level of catch that is sustainable and the fishing mortality associated with Maximum Sustainable Yield. Ignoring the possibility of depensation, given lack of (informative) data can consequently lead to unduly aggressive management strategies.

A perhaps more serious example of the impact of selecting perhaps overly parsimonious models when identifying operating models occurs in relation to stock structure uncertainty. The western North Pacific minke whales (*Balaenoptera acutorostrata*) provides an example of how stock-structure uncertainty can impact the selection of a

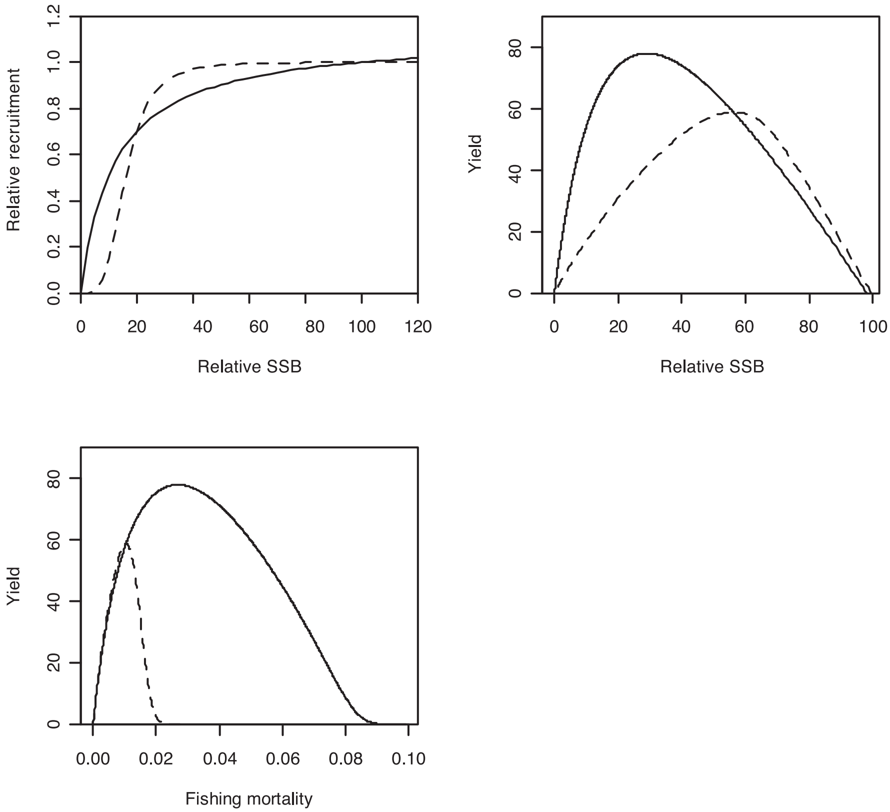

Fig. 2. Stock–recruitment relationships that include (dashed lines) and ignore (solid lines) depensation (upper left panel), and the implications of accounting for depensation for the (mean) surplus production function (upper right panel) and the relationship between yield and fishing mortality (lower left panel).

management strategy. Four broad stock-structure hypotheses were developed for this group of minke whales (IWC, 2004). These stock-structure hypotheses covered a range from two to four stocks in the region to be managed. The least complicated stock-structure hypotheses (two and three stocks) were justified primarily by the lack of evidence from statistical hypothesis tests applied to a range of genetic and non-genetic data. The most complicated hypothesis (four stocks) arose from the application of a clustering algorithm (Martien and Taylor 2001) to mtDNA data, noting that the power of most methods to detect stock structure can be very poor (Martien and Taylor 2003). Perhaps not unexpectedly, the operating models based on four stocks posed more of a challenge to the candidate management strategies, and uncertainty about stock structure led, in part, to the IWC SC being unable to reach agreement on relative weights for different stock structure hypotheses.

4. Concluding Remarks

The discussion above might seen to suggest that improved management advice will require more sophisticated and complicated stock assessments. In contrast, stock assessment science needs to consider two distinct types of assessments. Those which will form the basis for tactical management advice (e.g. setting of Total Allowable Catches and other management measures) by being part of management strategies, and those that will form the basis for the evaluation of candidate management strategies. In principle, the "stock assessment" component of a management strategy can be fairly simple (such as the largely empirical approaches on which management advice for sardine (*Sardinops sagax*) and anchovy (*Engraulis encrasicolus*) off south Africa are based; De Oliveira and Butterworth 2004) while great complexity should ideally only be found in the operating models used to evaluate management strategies.

The increased focus on an evaluation of feedback-control management strategies is likely to increase the demands on the stock assessment community owing to the move from trying to a identify (and then justify) a single "best" model to identifying an appropriate of range of plausible models, parameterizing the models, and assigning weights to them. However, the move to MSE has several advantages. Specifically, the evaluation of management strategies forces decision makers to be explicit regarding their management objectives and also helps to train them regarding the trade-offs among these management objectives (Walters 1994). In addition, attempting to explore a broad range of uncertainties has the advantage that it becomes possible to determine the relative importance of different factors in relation to achieving management goals. One almost general result of MSE analyses is that the presence of model uncertainty has a larger impact on the achieving management objectives than those sources of uncertainty considered routinely in stock assessments (observation and process error).

There are disadvantages associated with evaluating management strategies in addition to the increased technical demands on stock assessment scientists. Specifically, it is necessary to clearly define "risk". Unfortunately, while decision makers are often willing to express an interest in "minimizing risk", they are seldom willing to provide (let alone agree to) a definition for "risk" (and hence "acceptable risk"). Workshopping (e.g. Mapstone *et al.* in press) provides one way to explore different stakeholder objectives and attempt to find common ground among stakeholders who have perhaps markedly different objectives.

Data-poor situations pose major challenges for the conventional approach to fisheries stock assessment, and these problems are not alleviated when attention focuses on the performance of management strategies. For example, it is likely that the only management strategies that will perform adequately for all plausible scenarios are those which are very precautionary or which incorporate mechanisms (such as large MPAs) to protect the resource against the lack of information.

Schnute *et al.* (2007) identify the need for a global effort within the stock assessment community to develop software to implement general MSE frameworks. However, an equally important need at present is to identify whether there are "universal laws" which pertain to management strategies. Although they remain to be proved, such rules might include "management strategies based on empirical indicators are more likely than model-based management strategies to respond to major shifts in population abundance albeit at the cost of larger inter-annual variation in catches and stock sizes". A global meta-analytic analysis based on MSEs for a variety of regions could be used to "test" such proposed "laws". The availability of sets of such "laws" could be used for regions for which the resources needed to conduct MSEs are lacking.

Finally, Quinn (2003) speculates that the "Golden Age" of fisheries population dynamics models may be over, noting that attacks on modelling and assessment paradigms based only on single-species considerations will continue because of the increasing number of well-publicized fishery collapses and the fact that increasingly it is becoming obvious to the broader scientific community that the single-species models conventionally used as the basis of management advice make assumptions (such as that natural mortality is time-invariant) that are not valid. While the push towards assessment and management paradigms which explicitly include a greater variety of biological processes (in particular multi-species interactions and climate impacts on recruitment, growth and movement) will continue, it seems unlikely that management advice will be based on such paradigms for many years (Marasco *et al*. 2007). Rather, the fact that "complexity is not necessary better" means that changes to current management systems should continue to be made incrementally, using the MSE approach to confirm that proposed changes are both needed and beneficial. In fact, I predict that the MSE approach will show that in many cases well-designed management strategies based on single-species analyses are, in fact, fairly robust to the types of violations of the assumptions highlighted so frequently as long as management continues to be based on feedback control management systems and the decision makers follow the outputs from management strategies and enforce management regulations adequately.

Acknowledgements

This work was supported by NMFS grant NA04NMF4550330. Toshihide Kitakado (Tokyo University of Marine Science and Technology) is thanked for his comments on an earlier version of this paper.

References

Akaike H. A new look at the statistical model identification. *IEEE Trans. Automatic Cntl.* 1974; **19**: 716–723.

Begley J. Gadget User Manual. 2003 (available at: http://www.hafro.is/gadget).

Bentley N, Breen P, Starr P. Design and evaluation of a revised management decision rule for red rock lobster (*Jasus edwardsii*) fisheries in CRA 7 and CRA 8. 2003 New Zealand Fisheries Assessment Report 2003/30; 44 pp.

Bull B, Francis RICC, Dunn A, McKenzie A, Gilbert DJ, Smith MH. CASAL (C++ algorithmic stock assessment laboratory): CASAL user manual. v2.07–2005/08/21. 2005; NIWA Technical Report 127.

Burnham KP, Anderson DR. *Model Selection and Inference: A Practical Information-Theoretic Approach*. 2nd ed. Springer-Verlag, New York. 2002.

Butterworth DS, Punt AE, Smith ADM. On plausible hypotheses and their weighting, with implications for selection between variants of the Revised Management Procedure. *Rep. Int. Whal. Comm.* 1996; **46**: 637–640.

Butterworth DS, Cochrane KL, De Oliveira JAA. Management procedures: A better way to manage fisheries? The South African experience. *In*: Pikitch EK, Huppert DD, Sissenwine MP, Duke M (eds). *Global trends: fisheries management*, *Am. Fish. Soc. Symp.* 1997; **20**: 83–90.

De Oliveira JAA, Butterworth DS. Developing and refining a joint management procedure for the multispecies South African pelagic fishery. *ICES J. Mar. Sci.* 2004; **61**: 1432–1442.

De Oliveira JAA, Butterworth DS, Roel BA, Cochrane KL, Brown JP. The application of a management procedure to regulate the directed and bycatch fishery of South African sardine *Sardinops sagax*. *S. Afr. J. Mar. Sci.* 1998; **19**: 449–469.

Dichmont CM, Deng R, Punt AE, Venables W, Haddon M. Management Strategies for short lived species: the case of Australia's Northern Prawn Fishery. 1. Modelling the fishery taking account of the input control nature of management. *Fish. Res.* 2006; **82**: 204–220.

Francis RICC, Shotton R. "Risk" in fisheries management: a review. *Can. J. Fish. Aquat. Sci.* 1997; **54**: 1699–1715.

Francis RICC, Haist V, Bull B. Assessment of hoki (*Macruronus novaezelandiae*) in 2002 using a new model. 2003. New Zealand Fisheries Assessment Report 2003/6; 69 pp.

Fulton EA, Smith ADM, Smith DC. Alternative management strategies for Southeastern Australian Commonwealth Fisheries: Stage 2: Quantitative Management Strategy Evaluation. Report to the Australian Fisheries Management Authority and the Fisheries Research and Development Corporation. CSIRO Marine and Atmospheric Research. 2007; 400 pp.

Givens GH, Raftery AE, Zeh JE. Assessment of the Bering–Chukchi–Beaufort Seas stock of bowhead whales using the BALEEN II model in a Bayesian synthesis framework. *Rep. Int. Whal. Commn.* 1995; **45**: 345–364.

Hampton J, Fournier DA. A spatially disaggregated, length-based, age-structured population model of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. *Mar. Freshw. Res.* 2001; **52**: 937–963.

Harley SJ, Myers RA. Hierarchical Bayesian models of length-specific catchability of research trawl surveys. *Can. J. Fish. Aquat. Sci.* 2001; **58**: 1569–1584.

Harley SJ, Myers RA, Dunn A. Is catch-per-unit-effort proportional to abundance? *Can. J. Fish. Aquat. Sci.* 2001; **58**: 1760–1772.

Hobday DK, Punt AE. Spatial modelling and assessment of the Victorian Southern Rock Lobster (*Jasus edwardsii*) fishery. Final report to Fisheries Research and Development Corporation Project No. 2004/037. Primary Industries Research Victoria, Queenscliff, Australia. 2006; v + 199 pp.

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

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

IWC. Report of the Scientific Committee, Annex D. Report of the Sub-Committee on Management Procedures. *J. Cetacean Res. Manage.* 2004; **6**(suppl.): 75–184.

IWC. Report of the Scientific Committee. Annex D. Report of the sub-committee on the revised management procedure. Appendix 2. Requirements and guidelines for implementation. *J. Cetacean Res. Manage.* 2005; **7**(suppl.): 84–92.

IWC. Report of the Scientific Committee, Annex D. Report of the sub-committee on management procedures. *J. Cetacean Res. Manage.* 2007; **9**(suppl.): 88–128.

Jurado-Molina J, Livingston PA, Ianelli JN. Incorporating predation interactions in a statistical catch-at-age model for a predator-prey system in the eastern Bering Sea. *Can. J. Fish. Aquat. Sci.* 2005; **62**: 1865–1873.

Kell LT, De Oliveira JAA, Punt AE, McAllister MK, Kuikka S. Operational management procedures: An introduction to the use of management strategy evaluation frameworks. *In*: Motos L, Wilson DC (eds). *The Knowledge Base for Fisheries Management, Developments in Aquaculture and Fisheries Science*, 36. Elsevier, Amsterdam. 2006; 379–407.

Liermann M, Hilborn R. Depensation in fish stocks: a hierarchic Bayesian meta-analysis. *Can. J. Fish. Aquat. Sci.* 1997; **54**: 1976–1984.

Ludwig D, Walters CJ. Are age-structured models appropriate for catch-effort data? *Can. J. Fish. Aquat. Sci.* 1985; **42**: 1066–1072.

Magnússon KG. An overview of the multispecies VPA—theory and applications. *Rev. Fish. Biol. Fish.* 1995; **5**: 195–212.

Mapstone BD, Little LR, Punt AE, Davies CR, Smith ADM, Pantus F, McDonald AD, Williams AJ, Jones A. Balancing conservation and fishery objectives across multiple institutions and diverse stakeholders: Management strategy evaluations for line fishing in the Great Barrier Reef Marine Park and World Heritage Area. *Fish. Res.* (in press).

Marasco RJ, Goodman D, Grimes CB, Lawson PW, Punt AE, Quinn TJ II. Ecosystem Based Fisheries Management: Some practical suggestions. *Can. J. Fish. Aquat. Sci.* 2007; **64**: 928–939.

Martien KK, Taylor BL. A new method of generating hypothesized population structures for continuously distributed species using genetic data. Paper SC/53SD7 presented to the IWC Scientific Committee, July 2001; 16 pp.

Martien K, Taylor BL. Limitations of hypothesis-testing in defining management units for continuously distributed species. *J. Cetacean Res. Manage.* 2003; **5**: 213–219.

Maunder, MN, Watters, GM. A general framework for integrating environmental time series into stock assessment models: model description, simulation testing, and example. *Fish Bull. US.* 2003; **101**: 89–99.

Methot R. Technical description of the stock synthesis II assessment program. Version 1.17. March 2005.
Methot R. User manual for the Integrated analysis program stock synthesis 2 (SS2). Model version 2.00c. March 2007.
Millar RB, Methot, RD. Age-structured meta-analysis of U.S. West Coast rockfish (*Scorpaenidae*) populations and hierarchical modeling of trawl survey catchabilities. *Can. J. Fish. Aquat. Sci.* 2002; **59**: 383–392.
Myers RA, Barrowman NJ, Hutchings JA. Rosenberg AA. Population dynamics of exploited fish stocks at low population levels. *Science* (Washington DC), 1995; **269**: 1106–1108.
Punt AE. 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*: Smith SJ, Hunt JJ, Rivard D (eds). *Risk Evaluation and Biological Reference Points for Fisheries Management*. Can. J. Fish. Aquat. Sci. Spec. Publ. Vol. 120. 1993; 283–299.
Punt AE. The FAO Precautionary Approach after almost 10 years: Have we progressed towards implementing simulation-tested feedback-control management systems for fisheries management? *Nat. Res. Model.* 2006; **19**: 441–464.
Punt AE, Donovan G. Developing management procedures that are robust to uncertainty: Lessons from the International Whaling Commission. *ICES J. Mar. Sci.* 2007; **64**: 603–612.
Punt AE, Smith ADM. Harvest strategy evaluation for the eastern stock of gemfish (*Rexea solandri*). *ICES J. Mar. Sci.* 1999; **56**: 860–875.
Punt AE, Pribac F, Walker TI, Taylor BL, Prince JD. Stock assessment of school shark *Galeorhinus galeus* based on a spatially-explicit population dynamics model. *Mar. Freshw. Res.* 2000; **51**: 205–220.
Punt AE, Pribac F, Taylor BL, Walker TI. Harvest strategy evaluation for school and gummy shark. *J. Northw. Atl. Fish. Sci.* 2005; **32**: 387–406.
Quinn TJ II. Ruminations on the development and future of population dynamics models in fisheries. *Nat. Res. Model.* 2003; **16**: 341–392.
Ralston S. West coast groundfish harvest policy. *N. Am. J. Fish. Manage.* 2002; **22**: 249–250.
Sainsbury KJ, Punt AE, Smith ADM. Design of operational management strategies for achieving fishery ecosystem objectives. *ICES J. Mar. Sci.* 2000; **57**: 731–741.
Sampson DB. The status of black rockfish off Oregon and California in 2007. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 100, Portland, OR 97220, USA. 2007; 236 pp.
SC-CAMLR. Report of the Second Workshop on Management Procedures. *In*: *Report of the Twenty-fifth Meeting of the Scientific Committee (SC-CAMLR-XXV)*. Annex 4, Appendix D. CCAMLR, Hobart, Australia. 2006; 227–258.
Schirripa, MJ. Status of the sablefish resource off the continental U.S. Pacific coast in 2007. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 100, Portland, OR 97220, USA. 2007. 119 pp.
Schnute JT, Maunder MN, Ianelli JN. Designing tools to evaluate fishery management strategies: Can the scientific community deliver? *ICES J. Mar. Sci.* 2007; **64**: 1077–1084.
Schweder T, Hagen GS, Hatlebaak E. On the effect on cod and herring fisheries of retuning the Revised Management Procedure for minke whaling in the greater Barents Sea. *Fish. Res.* 1998; **37**: 77–95.
Smith ADM. Management strategy evaluation: the light on the hill. *In*: Hancock DA (ed). *Population Dynamics for Fisheries Management*. Australian Society for Fish Biology, Perth. 1994; 249–253.
Smith ADM, Sainsbury KJ, Stevens RA. Implementing effective fisheries-management systems—management strategy evaluation and the Australian partnership approach. *ICES J. Mar. Sci.* 1999; **56**: 967–979.
Stefansson, G. Issues in multispecies models. *Nat. Res. Model.* 2003; **16**: 415–437.
Walters CJ. Use of gaming procedures in evaluation of management experiments. *Can. J. Fish. Aquat. Sci.* 1994; **51**: 2705–2714.

Appendix: A simple MSE analysis

The population dynamics are assumed to be governed by:

$$B_{t+1} = B_t e^{-M} + \frac{(5^\gamma - 1) h R_0 \left(B_{t-L}/B_0\right)^\gamma}{(1-h) + (5^\gamma h - 1)\left(B_{t-L}/B_0\right)^\gamma} e^{\varepsilon_t - \sigma_R/2} - C_t; \quad \varepsilon_t \sim N\left(0; \sigma_R^2\right) \tag{A.1}$$

where B_t is the biomass at the start of year t, M is the instantaneous rate of nature mortality (set to 0.2 yr^{-1} for the analyses of this paper), R_0 is the recruitment at unfished equilibrium ($=B_0(1 - e^{-M})$), B_0 is the average unfished biomass, C_t is the catch during year t:

$$C_t = EB_t \tag{A.2}$$

E is the exploitation rate, L is the lag between spawning and recruitment to the exploitable (equals mature) biomass (set to 3 yr for the analyses of this paper), h is the "steepness" of the stock–recruitment relationship (the recruitment at $0.2B_0$ expressed as fraction of R_0), γ is the parameter that determines the shape of the stock–recruitment relationship (set to 1 for the Beverton–Holt form of the stock–recruitment relationship), and σ_R is the standard deviation of the random fluctuations in recruitment.