Comment on "Purported flaws in management strategy evaluation: basic problems or misinterpretation?" by Butterworth *et al.*

Marie-Joëlle Rochet[1, *] and Jake C. Rice[2]

[1] IFREMER, Département Ecologie et Modèles pour l'Halieutique, BP 21105, Nantes cedex 03, France
[2] Department of Fisheries and Oceans, 200 Kent Street, Ottawa, Ontario, Canada K1A 0E

*: Corresponding author : M-J. Rochet: tel: +33 240 374121; fax: +33 240 374075; email address : mjrochet@ifremer.fr.

Abstract:

Simulation-based management strategy evaluation is a valuable tool, when appropriately implemented. Implementation, however, may not always have been appropriate, and some reasons are provided why perhaps there is incomplete faith in certain of its technical aspects, such as knowing the distribution of the parameters of population processes from the information in limited datasets. A management strategy that has been evaluated by simulation should not be used as an "autopilot", because even the most competent of experts can develop autopilots with imperfect and incomplete knowledge of reality, and all information should be incorporated when decisions have to be made.

Keywords: decision-making, distribution models, management strategy evaluation

The manuscript by Butterworth *et al.* (2010) is a response to a paper published in the *ICES Journal of Marine Science* by Rochet and Rice (2009), referred to hereafter as R&R. The authors make five substantive criticisms of R&R. They argue that most concerns regarding management strategy evaluation (MSE) raised in the R&R paper are addressed when the method is properly implemented. They argue that the paradox identified by R&R, namely that modelling a distribution implies knowing more than only estimating its expected value, is a misunderstanding; instead they assume that modelling uncertainty is always preferable to ignoring it. They challenge R&R's interpretation that their simulations showed re-sampling methods may underestimate the probabilities of low values. They further argue that R&R's criticism in terms of implementation uncertainty can in fact be addressed by simulations as part of an MSE. They suggest that the alternative approaches proposed by R&R for the evaluation of management options are not appropriate for incorporating uncertainty nor the effects of feedback. Several points are discussed without direct criticism of R&R, particularly the robustness of MSE conclusions to various operating models, and on whether simulation results can be used to identify the best MSE.

R&R did not assert that there are fundamental flaws in MSE (a point conceded by Butterworth *et al.*, 2010, notwithstanding their title), and we agree that if simulation-based methods were always appropriately implemented, there would have been no need to write the R&R paper. Doug Butterworth and his colleagues are among the leaders in MSE development, and implementations they have led are generally not implicated in the issues raised by R&R. However, although analytical stock assessments carried out carefully by assessment scientists meet high standards, there is no guarantee that all stock assessments are analytically without flaws, and the fact that good MSEs can be produced by expert scientists does not ensure that all MSEs will avoid the types of flaw discussed in R&R. Although R&R did not go to the extreme criticism of "garbage in – garbage out", as used by Butterworth *et al.* (2010), we have seen the types of flaws discussed in R&R in work presented in various fora. As we all agree they should be avoided, if this exchange leads to increased vigilance by practitioners and reviewers, the discipline and the resources will be much better off.

With agreement on many core points between the Rochet and Rice (2009) and Butterworth *et al.* (2010) papers, it is difficult to identify the source of the "misinterpretation" referred to in the title of the latter paper, and the most important misunderstanding of definition could be that the definition in Butterworth *et al.* (2010) is itself consistent with R&R's usage. However, Butterworth *et al.* (2010) might misunderstand or misrepresent some key parts of R&R's thesis. Some are details, such as their criticism of R&R's interpretation of the use of re-sampling methods in MSEs, where we believe that Figure 3 of R&R does demonstrate the possibility (but not the certainty) of underestimating the likelihood of rare events. Some are more conceptual, but more serious. Butterworth *et al.* (2010) represent R&R's cautions regarding MSE as a desire to restrict practice to only "traditional stock assessment". Rather, unlike Butterworth *et al.* (2010), who suggest that MSE is the only alternative to traditional stock assessment, the main point by R&R was that there are different, complementary approaches. Each, including MSE, will have strengths and drawbacks that need to be considered on a case-by-case basis, and no one should be considered the universal solution to fisheries assessment and advice. For example, *post hoc* analyses suffer problems of attribution of cause, as outlined by Butterworth *et al.* (2010). However, such analyses are required to form the knowledge base underlying operating models and MSEs: if causes are not understood in real cases, there is little hope that they can be simulated adequately in models. Likewise, when R&R stated that, paradoxically, modelling a distribution implies knowing more than just taking its expected value, R&R did not argue or imply that it was preferable to ignore uncertainty. Rather they cautioned about the risks of assuming that uncertainty could be captured fully and accurately by using distributions for key parameters, especially if the distributions were not securely supported by adequate data. R&R were not specific about what alternative was a better way to deal with uncertainty, again because they felt that it was unlikely that there was some single best way. For example, in their

commentary on R&R's reference to implementation uncertainty, Butterworth *et al.* (2010) argue for considering implementation within the MSE, and provide examples where it has been done. We do not share the belief that we should be confident of modelling accurately the complex human processes underlying implementation uncertainty. Rather, we encourage efforts aimed at using MSE to establish likely implementation errors that would cause specific management strategies to fail, a point made in R&R. We advise caution in believing that human behaviour and economic externalities can be modelled well enough to develop confidence that a management strategy will be found that is reliable enough to use as an "autopilot".

It is perhaps this search for an autopilot that best differentiates Butterworth *et al.* (2010) from R&R. The latter argue for maximizing intelligence in science support for fisheries decision-making. Contrary to the discussion of this point by the former, this is not arguing for the annual debate of assessment results that opens the door to partisanship – be it power-brokering by industry lobbies or advocacy science by special interest groups. Intelligence is getting as much information as possible into public dialogue and decision-making, ensuring that all of it is understood by all parties, and designing decision-making processes that are inclusive and equitable. We think that this should not be done once, during an MSE – even one that, as Butterworth *et al.* (2010) say, can engage stakeholders – but each time a decision has to be made. We are well aware that Sparholt *et al.* (2007) found that implementation of the precautionary approach at ICES did not prevent demersal stocks from declining, but Gerjan Piet and some Dutch colleagues show, in a manuscript currently under preparation ("The necessity of response indicators as part of an ecosystem approach to fisheries management"), that for 121 stocks for which ICES provided advice over the period 1987–2006, the official TAC was on average 30% higher than the scientific advice, suggesting that quantitative advice is used qualitatively anyway. We agree that such a system needs to be improved. However, we believe that improvement can be made in many ways, and certainly not by adopting a system built on an autopilot. Even with stakeholders involved in the process, the details and parameterization of the operating model that determine the final answer are proposed in most cases by scientific and technical experts whom stakeholders trust and feel they are not competent to contest, or in some instances do not fully understand. As a consequence, dictatorship of technical expertise is difficult to avoid, even in the best of faith.

1. References

Butterworth, D. S., Bentley, N., De Oliveira, J. A. A., Donovan, G. P., Kell, L. T., Parma, A., Punt, A. E., *et al.* 2010. Purported flaws in management strategy evaluation: fundamental problems or misinterpretation? ICES Journal of Marine Science, 67: 000–000.

Rochet, M-J., and Rice, J. C. 2009. Simulation-based management strategy evaluation: ignorance disguised as mathematics? ICES Journal of Marine Science, 66: 754–762.

Sparholt, H., Bertelsen, M., and Lassen, H. 2007. A meta-analysis of the status of ICES fish stocks during the past half century. ICES Journal of Marine Science, 64: 707–713.