Conclusions and Recommendations from the 2018 Joint tuna RFMO Management Strategy Evaluation Working Group Meeting in Seattle, USA – 13-15 June 2018

MSE process and stakeholder dialogue

1) The Group **stresses** that a successful and efficient MSE process should not be assigned to a single individual – it is an iterative process that should involve a consistent, core group of experts that regularly reports on progress to other scientists, managers and other stakeholders and implements their feedback. In addition, experience with previous MSE initiatives has highlighted the value of a 'guillotine' mechanism if the whole process is to avoid back-tracking and to meet deadlines for completion within a reasonable period of 2-3 years:
    a) the first guillotine should apply to data selection, after which no new data may be taken into account in the process;
    b) the second guillotine applies to agreement on a set of satisfactorily conditioned operating models, after which MPs testing is based on those accepted models alone.

   Further data or scenarios with their associated OMs that are forthcoming after these guillotine dates can be taken into account when the accepted MP enters a subsequent review and revision process (under an agreed schedule).
2) The Group **recommends** that each RFMO identifies all stakeholders, ideally at the outset, and clarifies their role and input within their MSE process. Not all stakeholders need to be involved in all aspects of the process; however, transparency and trust is critical and must be established. Mechanisms to achieve this, such as the use of "intermediary groups" (e.g. Miller et al; submitted) should be established.
3) In addition to scientists, the Group **advises** that consideration should be given to the use of other experts (e.g. managers, industry and/or conservation representatives) with experience of the MSE implementation process, to provide capacity building workshops for managers. This may facilitate better targeted information sharing as scientists may have a tendency to concentrate more on technical issues. In addition to the present capacity building efforts, consideration should also be given to more targeted approaches to individuals closer to decision process; this could include one-on-one meetings (with either a single individual or a group from a single country).
4) Small technical task groups to discuss and advance key aspects of the MSE process that are of common interest to the Tuna RFMOs are beneficial (and see 5 below). Care should be taken to ensure communication is maintained and that the work of these task groups is presented back to the larger tRFMO MSE WG and appropriate RFMO working groups.
5) Reviews of an MSE can be considered at 3 levels:
    i. *Broad:* the overall MSE process (i.e. the rationale, framework and workplan);
    ii. *More detailed:* specific MSE components e.g. review of operating models (OMs) and their conditioning (see 7 below); and
    iii. *Specific:* validation of the technical code developed for MSEs at the various RFMOs, i.e. confirm that the code is correct and consistent with the equations documented in the full 'trial specifications document' (see 12 below).

   The Group **recommends** that RFMOs should decide at an early stage how this review process will occur (including internal review through Scientific and other RFMO Committees and groups and/or the appointment of independent external experts on technical and process aspects of

MSE), noting that review must be iterative, not occur only at the end of the process. Should one or more RFMOs request that this Group is involved in the review process (this would provide a level of consistency amongst RFMOs), then long-term funding, support and expertise will be required. One option is that this Group could be directly involved in the first two levels as an advisory body, providing advice and facilitating contact with key experts to conduct one, or more, stages of the review and recommend appropriate sources of expertise for the validation exercise. The group **noted**, that transparency for the wider community was an important aspect of the review process and that the use of experts independent of the RFMO has been valuable in a number of case studies.

6) The Group recognizes that obtaining MSC (or similar) certification is a key motivator for fishing industries. However, concern has arisen about the applicability of the current MSC guidelines/criteria to fisheries managed under approaches developed using MSE. This is because the MSC's approach seems to be based on the "best assessment plus HCR" paradigm with its associated reference points, and these concepts often do not translate readily to the rather different management framework based on the precautionary MSE process. The Group therefore **recommends** that dialogue takes place with the MSC (perhaps leading to a joint workshop) to discuss their criteria for certification in an MSE context.

Conditioning operating models

7) With respect to OMs, the Group **advises** that it is valuable to limit their number to that needed to adequately address the key uncertainties, with a focus on those that may have management implications in the future (see 9 below). However, it **stresses** that this limitation should not be taken too far – the OMs should consider a range of plausible scenarios which is sufficiently broad that tested MPs or HCRs[1] do not require amendment or retesting too often.
8) The Group also **stresses** that it **essential** that all OMs are adequately conditioned i.e. ensure that they are sufficiently consistent with the historical data to be considered plausible. Whilst conditioning is a case-specific process, there are some general guidelines that should be followed including: the use of standard model fit diagnostics for indications of model mis-specification (automated where possible); focusing on the conditioning of 'limit' cases and which may be sufficient to justify the assumption that conditioning in between these is adequate.
9) Stock structure has been found elsewhere (e.g. with whales) to be a major source of uncertainty with strong conservation and management implications. It is also difficult to model. Thus far, this issue has not been given much emphasis in fishery MSE development. The Group **recommends** that much more attention is dedicated to this issue, including a focus on the research needed to provide the necessary data to develop and parametrize the OMs needed.
10) Shortage of time precluded discussion of the topic of how to weight the scenarios for which OMs are developed in relation to their relatively plausibility. The Group **agrees** that this is an important and difficult issue that should be taken up with high priority in future meetings.

[1] Harvest Control Rules with discontinuities (leading to potentially large changes in recommended catches) should generally be avoided.

11) With respect to multispecies MSE, the Group **recommends** that initial OM developments focus on technical interactions (i.e., fleet and fishing operation levels with fleets focusing primarily on one species being unable to avoid catching others).

Computational aspects

12) The Group **stresses** that it is **essential** that the mathematical specifications for all code developed for MSE purposes is fully documented as part of a trials specification document; and that the code is validated and made publicly available, since it will ultimately be used to provide management recommendations.

Dissemination of results

13) The Group **recommends** that visualization approaches for presenting MSE results (e.g. for case study stocks) should be tested with various focus groups to check their suitability for each forum/stakeholder group. Feedback will help to develop more effective and targeted formats.
14) The Group **recommends** the creation of a common 'GitHub' or similar site to submit code for individual components of the various sets of MSE initiatives, including graphical presentations of results. This will also facilitate sharing of code on a modular basis that could be helpful to MSEs across RFMOs.

Further Work

15) In terms of its further work, the Group **recommends** that**:**
    i. refinements to the draft glossary be made such that it can be finalized in three months' time (this should eventually include a lay glossary as well as a technical glossary);
    ii. it continues to discuss the topic of 'Exceptional Circumstances'; this will be coordinated by Ann Preece and David Die;
    iii. further consideration is given to the relative merits of model-based vs empirical MPs as it has been suggested (e.g. see Punt 2018) that there is little difference between model-based and empirical MPs in terms of performance, but that the latter have advantages in terms of easier understanding by stakeholders and simpler testing;
    iv. a comprehensive joint TRFMO MSE WG website is developed that provides information and updates regarding the activities of the Group in a clear manner, as well as links to each RFMO's MSE webpages (a website manager should be identified and supported); and
    v. the Chair develops a workplan, possibly in conjunction with a steering committee, to develop an agenda for the next meeting as well as a workplan and priorities for further activities.

References

Miller, S. K., Anganuzzi, A., Butterworth, D. S., Davies, C. R., Donovan, G. P., Nickson, A., Rademeyer, R. A. and Restrepo, V. Submitted. Improving communication: the key to increasing the effectiveness of MSE processes.

Punt. A. E. 2018. What Makes an MP an MP and an MSE an MSE?. Unpublished Presentation. University of Washington, 14 January 2018.

Glossary of terms for harvest strategies, management procedures and management strategy evaluation

- This glossary was developed to encourage a consistent use of terms associated with harvest strategies, management procedures and management strategy evaluation processes underway across the five tuna RFMOs.
- It was developed from a range of sources, including ISSF, Rademeyer *et al*. 2007, IOTC, PEW Charitable Trust and a range of MSE practitioners with broad experience across tuna and other fisheries.
- A draft of the glossary was reviewed by participants in the 2018 Joint tuna RFMO Management Strategy Evaluation Working Group Meeting in Seattle and adopted for the purposes of improving consistency and clarity of communication in tRFMO MSE processes.
- The glossary is available for use by others with appropriate acknowledgement. (Anon. 2018. Glossary of terms for harvest strategies, management procedures and management strategy evaluation, http://www.tuna-org.org/Documents/MSEGlossary_tRFMO_MSEWG2018.pdf.)

Terms commonly used in Management Strategy Evaluation or Management Procedure literature

Term	Definition	Abbreviation/Symbol
Average Annual Variation (in catch/TAC)	The absolute value of the proportional TAC change each year, averaged over the projection period.	**AAV**
Biomass	Stock biomass, which may refer to various components of the stock. Often spawning stock biomass (SSB) of females is used, as the greatest conservation concern is to maintain the reproductive component of the resource.	**B**
Candidate Management Procedure	An MP (defined below) that has been proposed, but not yet adopted.	**CMP**
Conditioning	The process of fitting an Operating Model (OM) of the resource dynamics to the available data on the basis of some statistical criterion, such as a Maximum Likelihood. The aim of conditioning is to select those OMs consistent with the data and reject OMs that do not fit these data satisfactorily and, as such, are considered implausible.	
Error	Differences, primarily reflecting uncertainties in the relationship between the actual dynamics of the resource (described by the OMs) and observations. Four types of error may be distinguished, and simulation trials may take account of one or more of these: • Estimation error: differences between the actual values of the parameters of the OM and those provided by the estimator when fitting a model to the available data; • Implementation error: differences between intended management actions (as output by an MP) and those actually achieved (e.g. reflecting over-catch); • Observation error (or measurement error): differences between the measured value of some resource index and the corresponding value calculated by the OM; • Process error: natural variations in resource dynamics (e.g., fluctuations about a stock-recruitment curve or variation in fishery or survey selectivity /catchability).	

Term	Definition	Abbreviation/Symbol
Estimator	The statistical estimation process within a population model (assessment or OM); in a Management Strategy Evaluation (MSE) context, the component that provides information on resource status and productivity from past and generated future resource-monitoring data for input to the Harvest Control Rule (HCR) component of an MP in projections.	
Exceptional circumstances	Specifications of circumstances (primarily related to future monitoring data falling outside the range covered by simulation testing) where overriding of the output from a Management Procedure should be considered, together with broad principles to govern the action to take in such an event.	
Feedback Control	Rules or algorithms based, directly or indirectly, on trends in observations of resource indices, which adjust the management actions (such as a TAC change) in directions that will change resource abundance towards a level consistent with decision makers' objectives.	
Harvest Control Rule (also Decision Rule)	A pre-agreed and well-defined rule or action(s) that describes how management should adjust management measures in response to the state of specified indicator(s) of stock status. This is described by a mathematical formula.	**HCR**
Harvest Strategy	Some combination of monitoring, assessment, harvest control rule and management action designed to meet the stated objectives of a fishery. Sometimes referred to as a Management Strategy (see below). A fully specified harvest strategy that has been simulation tested for performance and adequate robustness to uncertainties is often referred to as a Management Procedure.	**HS**
Implementation	The practical application of a Harvest Strategy to provide a resource management recommendation.	
Kobe Plot	A plot that shows the current stock status, or a trajectory over time for a fished population, with abundance on the horizontal axis and fishing mortality on the vertical axis. These are often shown relative to B_{MSY} and to F_{MSY}, respectively. A Kobe plot is often divided into four quadrants by a vertical line at $B=B_{MSY}$ and a horizontal line at $F=F_{MSY}$.	
Limit Reference Point	A level of biomass below, or fishing mortality above, which an actual value would be considered undesirable, and which management action should seek to avoid.	**LRP**
Management Objectives	The social, economic, biological, ecosystem, and political (or other) goals for a given management unit (i.e. stock). These typically conflict, and include concepts such as maximising catches over time, minimising the chance of unintended stock depletion, and enhancing industry stability through low inter-annual variability in catches. For the purposes of Management Strategy Evaluation (MSE) these objective need to be quantified in the form of Performance statistics (see below).	**Objectives, MOs**

Term	Definition	Abbreviation/Symbol
Management Plan	In a broad fisheries governance context, a Management Plan is the combination of policies, regulations and management approaches adopted by the management authority to reach established societal objectives. The management plan generally includes the combination of policy principles and forms of management measures, monitoring and compliance that will be used to regulate the fishery, such as the nature of access rights, allocation of resources to stakeholders, controls on inputs (e.g. fishing capacity, gear regulations), outputs (e.g. quotas, minimum size at landing), and fishing operations restrictions (e.g. closed areas and seasons). Ideally, the Management Plan will also include the Harvest Strategy for the fishery or a set of principles and guidelines for the specification, implementation and review of a formal Management Procedure for target and non-target species.	
Management Procedure	A management procedure has the same components as a harvest strategy. The distinction is that each component of a Management Procedure is ***formally specified,*** *and* ***the combination*** of monitoring data, analysis method, harvest control rule and management measure has been simulation tested to demonstrate adequately robust performance in the face of plausible uncertainties about stock and fishery dynamics.	**MP**
Management Strategy	Synonymous with harvest strategy. (But note that this is also used with a broader meaning in a range of other contexts.)	
Management Strategy Evaluation	A process whereby the performances of alternative harvest strategies are tested and compared using stochastic simulations of stock and fishery dynamics against a set of performance statistics developed to quantify the attainment of management objectives.	**MSE**
Maximum Economic Yield	The (typically annual) yield that can be taken continuously from a stock sustainably (i.e. without reducing its size) that maximizes the economic yield of a fishery in equilibrium. This yield occurs at the effort level that creates the largest positive difference between total revenues and total costs of fishing (including the cost of labor, capital, management and research etc.), thus maximizing profits.	**MEY**
Maximum Sustainable Yield	The largest (typically annual) yield that can be taken continuously from a stock sustainably (i.e. without reducing its size). In real, and consequently stochastic situations, this is usually estimated as the largest average long-term yield that can be obtained by applying a constant fishing mortality F, where that F is denoted as F_{MSY}.	**MSY**
Observation Model	The component of the OM that generates fishery-dependent and/or fishery-independent resource monitoring data from the underling true status of the resource provided by the OM, for input to an MP.	

Term	Definition	Abbreviation/Symbol
Operating Model(s)	A mathematical–statistical model (usually models) used to describe the fishery dynamics in simulation trials, including the specifications for generating simulated resource monitoring data when projecting forward in time. Multiple models will usually be considered to reflect the uncertainties about the dynamics of the resource and fishery.	**OM(s)**
Performance statistics/measures	A set of statistics used to evaluate the performance of Candidate MPs (CMPs) against specified management objectives, and the robustness of these MPs to important uncertainties in resource and fishery dynamics.	
Plausibility (weights)	The likelihood of a scenario considered in simulation trials representing reality, relative to other scenarios also under consideration. Plausibility may be estimated formally based on some statistical approach, or specified based on expert judgement, and can be used to weight performance statistics when integrating over results for different scenarios (OMs).	
Precautionary Approach	An approach to resource management in which, where there are threats of serious irreversible environmental damage, lack of full scientific certainty is not used as a reason for postponing cost-effective measures to prevent environmental degradation.	**PA**
Reference case (also termed reference scenario or base case)	A single, typically central, conditioned OM for evaluating Candidate MPs (CMPs) that provides a pragmatic basis for comparison of performance statistics of the CMPs.	**RC (or BC)**
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 of Candidate MPs).	**RS**
Research-conditional option	Temporary application of an MP that does not satisfy conservation performance criteria, 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 (i.e. beyond the range of the Reference Set of models alone) of plausible scenarios. While plausible, robustness test OMs are typically considered to be less likely than the reference set OMs, and often focus on particularly challenging circumstances with potentially negative consequences to be avoided.	
Scenario	A hypothesis concerning resource status and dynamics or fishery operations, represented mathematically as an OM.	

Term	Definition	Abbreviation/Symbol
Simulation trial/test	A computer simulation to project stock and fishery dynamics for a particular scenario forward for a specified period, under controls specified by a HS or MP, to ascertain the performance of that HS or MP. Such projections will typically be repeated a large number of times to capture stochasticity.	
Spawning Biomass, initial	Initial spawning biomass prior to fishing as estimated from a stock assessment.	**SSB_0**
Spawning Biomass, current	Spawning biomass (SSB) in the last year(s) of the stock assessment.	**$SSB_{current}$**
Spawning Biomass at MSY	The equilibrium spawning biomass that results from fishing at F_{MSY}. In the presence of recruitment variability, fishing a stock at F_{MSY} will result in a biomass that fluctuates above and below SSB_{MSY}.	**SSB_{MSY}**
Stationarity	The assumption that population parameter values are fixed (at least in expectation), and not varying systematically, over time. This is a standard assumption for many aspects of stock assessments, OMs and management plans.	
Stock assessment	The process of **estimating** stock abundance and the impact of fishing on the stock, similar in many respects to the process of conditioning OMs.	
Target Reference Point	The point which corresponds to a state of a fishery and/or resource which is considered desirable and which management aims to achieve.	**TRP**
Trade-offs	A balance, or compromise, achieved between desirable but conflicting objectives when evaluating alternative MPs. Trade-offs arise because of the multiple objectives in fisheries management and the fact that some objectives conflict (e.g. maximizing catch vs minimizing risk of unintended depletion).	
Tuning	The process of adjusting values of control parameters of the Harvest Control Rule in a Management Procedure to achieve a single, precisely-defined performance statistic in a specified simulation test. This reduces confounding effects to allow the performance of different candidate MPs to be compared more readily with respect to other management objectives. For example, in the case of evaluating rebuilding plans, all candidate MPs might be tuned to meet the rebuilding objective for a specified simulation trial; then the focus of comparisons among MPs is performance and behaviour with respect to catch and CPUE dimensions.	
Weight(s)	Either qualitative (e.g. high, medium, low) or quantitative measures of relative plausibility accorded across a set of scenarios.	
Worm plot	Time series plots showing a number of possible realizations of simulated projections of, for example, catch or spawning biomass under the application of an MP for a specific OM or weighted set of OMs.	