REPORT OF THE 2022 INTERSESSIONAL MEETING OF THE WORKING GROUP ON STOCK ASSESSMENT METHODS

(Online, 31 May - 3 June 2022)

The results, conclusions and recommendations contained in this Report only reflect the view of the Working Group on Stock Assessment Methods. Therefore, these should be considered preliminary until the SCRS adopts them at its annual Plenary meeting and the Commission revise them at its Annual meeting. Accordingly, ICCAT reserves the right to comment, object and endorse this Report, until it is finally adopted by the Commission.

1. Opening, adoption of agenda and meeting arrangements and assignment of rapporteurs

The 2022 Intersessional Meeting of the Working Group on Stock Assessment Methods (WGSAM, "the Group") was held online from 31 May to 3 June 2022. Dr Michael Schirripa (USA), the Rapporteur for the WGSAM, opened the meeting and served as Chair.

The ICCAT Executive Secretary and the SCRS Chair welcomed and thanked the participants. The Group Chair proceeded to review the Agenda which was adopted after some changes (**Appendix 1**).

The List of Participants is included in **Appendix 2**. The List of Documents and Presentations provided at the meeting is attached as **Appendix 3**. The abstracts of all SCRS documents and presentations provided are included in **Appendix 4**. The following served as rapporteurs:

Sections	*Rapporteur*
Items 1, 8	A. Kimoto
Item 2	C. Peterson
Item 3	D. Courtney, N. Fisch and E. Babcock
Item 4	E. Babcock
Item 5.1	B.T. Moffat and E. Babcock
Item 5.2	K. Gillespie
Items 6, 7	M. Schirripa

2. Harvest Control Rules, Limit Reference Points and Management Strategy Evaluation

MSE updates

Each of Species Group Rapporteurs provided an update of the current status and future directions of their respective Management Strategy Evaluation (MSE):

The Rapporteur of the Atlantic albacore stocks (ALB) provided an overview of the North Atlantic albacore MSE (SCRS/P/2022/029). The presentation included management objectives; Management Procedure including data and stock assessment specifications, harvest control rule, Limit Reference Points, and allowable % TAC change; exceptional circumstances protocol; OM reference grid; performance statistics; and plots and tables used to present results for the North Atlantic albacore (N-ALB) MSE since the MSE was initiated in 2011. The N-ALB MSE timeline was presented, along with future directions which include detection of Exceptional Circumstances and preparing for an updated assessment and "MSE Round 2."

The Rapporteur of the western bluefin tuna (BFT) stock presented materials of BFT MSE (SCRS/P/2022/038): updates on the MSE framework including updated performance metrics and Limit Reference Points; updated Candidate Management Procedure (CMP) results and the process for CMP refinement and selection; the process for CMP finalization; and presentation of MSE results. It was highlighted the processes of development and performance CMP tuning the value of the BFT MSE web-dedicated (e.g., Splash) page (which followed North Atlantic swordfish MSE) and online Shiny apps; the refinement of selected performance metrics used for decision making; and the development of empirical (non-model based) MPs. The Rapporteur discussed the experience of the BFT Technical Sub-group on MSE interacting iteratively with managers through Panel 2 and with stakeholders through the communications group and Ambassador meetings. The significance of iterative engagement with Panel 2 was demonstrated through an example of Panel 2 changing their preference on allowable downward %TAC change restrictions after seeing the potential gains in average catch that could be obtained by allowing for a larger downward % TAC change. Link to web-dedicated MSE page is available at https://iccat.github.io/abft-mse/.

The Rapporteur of the North Atlantic swordfish stock (N-SWO) provided the updates of N-SWO MSE (SCRS/P/2022/036). The presentation covered the history of the N-SWO MSE to date, including key 2021/2022 developments and MSE structure, including OM axes of uncertainty and performance metrics; interactions with Panel 4; and the future workplan, which includes OM reconditioning, revisiting the axes of uncertainty, development of CMPs, and N-SWO MSE communications (which is intended to follow the example set by BFT). The Rapporteur also noted that the protocol for Exceptional Circumstances would undergo further development after development of candidate management procedures. The SWO Species Group anticipates providing management advice based on a management procedure by the end of 2023. Link to web-dedicated MSE page is available at https://iccat.github.io/nswo-mse/.

The Coordinator of the tropical tunas stocks presented preliminary results from the 2022 Intersessional tropical tunas (TT) MSE technical sub-group meeting held on 19-20 May 2022. The presentation included the details of the TT MSE roadmap; progress-to-date of the West skipjack (SKJ) MSE and the multi-species MSE (bigeye, yellowfin, East SKJ), which has included initial conditioning of the OMs, identification of major sources of uncertainty, identification of initial performance metrics and MPs, and MSE training; and future activities, including reconditioning of the OMs and communication with Panel 4. Future activities will include reconditioning of OMs. The Coordinator of the TT Species Group also highlighted capacity building by implementing an Ambassador program following the application by BFT and support through MSE training courses, which had been conducted in Brazil and should be applied to other CPCs.

The Group discussed the added complexity of a multi-stock MSE. Namely, challenges associated with equitable identification and weighting of management objectives across species, identifying multi-stock Reference Points (like MSY), analyzing and/or grouping performance metrics across species, and technical concerns of shifting stock boundaries between the West and East SKJ. Ultimately, many of these reflect decision points for Panel 1, and the MSE structure will need to reflect the main goal of the multi-stock MSE, which is to capture the real-world dynamics wherein any management applied to one stock will affect the other stocks in the MSE.

Communication

The Group noted the importance of communications, both with stakeholders and associated Panels. Notably, the MSE process is conceptually challenging, particularly for managers, stakeholders, and scientists more accustomed to the best assessment paradigm, wherein the best estimate of current status and future forecasts are presented. Contrarily, within the MSE paradigm, management control rules are developed and tested within a wide range of plausible current and future stock and fishery conditions. The aim of an MSE is to test the MP, not to provide advice on the current state of the stock.

Communication within an MSE process is multi-fold, and includes:

- Communication and interaction between stakeholders and managers
- Communication, training, and interaction between scientists and managers
- Communication, training, and interaction between stakeholders and scientists
- Continued communication after an MP is initially adopted

Iterative communication with the Panel, who are the key decision makers of the MSE process, is critical, particularly since management decisions are outside the scope of the SCRS. The Group discussed ways of most effectively interacting with their Panels and requesting additional meetings as necessary. To help overcome the formal Panel meeting structure, which does not always allow for easy decision-making, the Group highlighted the value of briefing with the Panel Chair ahead of the meeting to discuss what (if any) key decisions will need to be made and what should be communicated to update and help alleviate concerns of the Panel during the meeting.

Stakeholder outreach and communication are key to the MSE process. The BFT MSE Technical Sub-group developed a dedicated communications group that serves to develop communication materials appropriate to each audience and appointed Ambassadors (English, French, and Spanish-speaking) to present standardized overviews and training on the MSE process designed for stakeholders. The TT MSE Technical Sub-group also highlighted the value of MSE training courses given in Brazil and aims to replicate the training in other CPCs. Stakeholder communication is not standardized across CPCs, and the Group questioned how these communications could be standardized. The Group agreed that the Ambassador approach has been productive for BFT and should be continued for all MSE efforts. These communication efforts may be associated with additional costs.

The Group highlighted the value in developing unified and consistent presentation materials across all species to disseminate MSE results. Standardizing MSE reporting could allow managers and stakeholders to acclimate and more easily interpret MSE results. Consistent presentation materials can include standardized plots, MSE contractor-hosted Shiny app tools and web-dedicated MSE pages. Web-dedicated MSE pages have proven particularly useful for external communication and internally for MSE analysts and participants. It was understood, however, that alternative plots may be added as necessary for each stock.

The Group also questioned how communication would continue after an MP is adopted. As observed in ALB, the MSE process continues after MP adoption. The Group noted that web-dedicated MSE pages hosted by MSE Consultants could play this role, and a more thorough understanding of the next steps would be helpful, including who continues to update future MSEs and whether code reviews will become outdated.

Sharing of information and Standardization across MSEs

The Group highlighted the value of cross-referencing the experiences of each species' MSE process and identifying strengths and weaknesses that can be applied across MSEs. Across all MSEs, performance metrics tended to reflect safety, status, stability, and yield. Exceptional Circumstances should identify values not observed within the MSE projections, which would trigger a requisite of revisiting the MP.

The Group suggested that some form of nonstationarity should be considered in each MSE development (including but not limited to spatial distribution, recruitment, natural mortality, etc.), though it should not be required for all species. Inclusion or exclusion of nonstationary considerations should reflect the priorities of the MSE and balance the model complexity. For example, consider that multispecies Reference Points may be of greater interest within the TT mixed stock MSE rather than nonstationarity.

Recommendations for MSE best practice

The Group supported the following three recommendations:

1. Exploration of supplementary funding options
2. Review of overall MSE efforts in ICCAT
3. Formulation of a dedicated Reference Point study group

Recommendation 1: Exploration of supplementary funding options

Motivation:

ICCAT has been involved in the development of MSE frameworks for six key species and eight stocks (N-ALB, N-SWO, E and W-BFT and TT (Yellowfin, Bigeye and E and W-SKJ) for about a decade. During this period the level of effort attributed to each MSE has varied depending upon the funding available. In this regard, by far the greatest contribution to the MSE process has been the commitment of CPC's through their scientific and management staff involved in this long-term process (1000's of hours). Without this commitment, the MSEs could not have advanced to potentially providing TAC advice in the relatively near future for all Species Groups. The first was N-ALB (interim TAC) in 2019 to be followed by BFT in 2022.

In addition to the in-kind contribution of CPC's there have been a number of specific funding programs to support MSE development and implementation. By far the greatest financial contribution has been made by the EU through the GBYP program with voluntary contributions from several other CPC's throughout the BFT MSE development. For the other Species Groups, the majority of funding has been from the annual science budget and voluntary contributions, primarily from the EU and US. For 2022 the approved budget for MSE, excluding BFT, was €160,000. Other programs such as the JCAP (ICCAT-Japan Capacity-building Assistance Project) have supported very important initiatives on Capacity Building. **Table 1** summarizes the funding by year and species.

Recommendation:

The Group acknowledged the important contribution, both in-kind and financial, by all parties to the overall MSE approach and explored ideas on what may be needed in the future to ensure the continued successful implementation of MSE. The Group also noted that different levels of funding will be required for each of the ongoing MSE process. The Group discussed several aspects of MSE in terms of consolidation of MSE programs across the six species, recognizing the inter-species differences, to optimize the available resources.

The Group recommended the following:

- Commission/SCRS explore opportunities for additional funding to continue the MSE development and to support ongoing and future activities, as well as the anticipated requirements for completion of all MSE frameworks currently under development. Comparatively speaking, the complications of a three species eastern TT MSE will require more support and funding than a single species initiative.

- Explore opportunities for the efficient use of the limited resources, such as collaborating with international institutes, programs and researchers.

All the CPCs should continue to support and further expand their contributions to the MSE processes through a greater involvement of their national experts.

Recommendation 2: Review of overall MSE efforts in ICCAT

Motivation:

Though there is some cross-referencing of MSE activities between species, thus far, each MSE process has been largely conducted independently.

Recommendation:

The Group continued to recommend an overall review of ICCAT MSE activities by an external expert. This overall review would serve to identify potential improvements, highlight any missing components or shortcomings of the current process, achieve efficiencies across species and promote standardization of the MSE process across species, refine and standardize MSE communication and stakeholder engagement, and provide guidance on what the future of MSE looks like within ICCAT. This could include the way the MSE processes are supported and how resources are divided, and how the MSE processes should be structured and supported after MP adoption.

Recommendation 3: Formulation of a dedicated Reference Point study group

Motivation:

It was noted that selection of Limit Reference Points within MSEs across ICCAT species are largely based on previous precedent (e.g., $0.4B_{MSY}$ for ALB, SWO, BFT). For example, current B_{LIM} values for ALB, SWO, and BFT are set equal to $0.4B_{MSY}$, and the origination of this value is not fully clear (Preece *et al.,* 2011; Kell *et al.,* 2012; Anon. 2015). The Group indicated that it would be helpful to better understand the origination and logistics of these values.

The Group also noted the distinction between passive (purely observational) versus active (wherein some management action is triggered) Limit Reference Points. Limit Reference Points should be less strict within the MSE paradigm compared to the best assessment paradigm given the need to account for behavior across multiple OMs, and the “tail” or extreme probabilities (e.g., 5th and 95^{th} percentiles) of observed B_{LIM} should not be ignored.

Recommendation:

The Group recommended that a dedicated Reference Point study group be formed to investigate how Reference Points should be identified for each species. The Group should consider how Limit Reference Points should be calculated across life history strategies, and if they feel so inclined, could expand their investigation beyond Limit Reference Points to more broadly consider other types of Biological Reference Points (e.g. MSY-based Reference Points). Also, the Group may investigate dynamic, time-varying Reference Points.

3. Stock assessment

3.1 Validation of integrated stock assessment model ensembles

Presentation SCRS/P/2022/034 provided two examples of the use of structural uncertainty grids in regional fisheries management organisations (RFMO) shark assessments. The examples were provided based on recent assessments of blue shark (*Prionace glauca*) conducted for the Indian Ocean Tuna Commission (IOTC; Rice and Sharma 2015, Rice 2017, Rice 2021) and of oceanic whitetip (*Carcharhinus longimanus*) conducted for the Western and Central Pacific Fisheries Commission (WCPFC; Rice and Harley 2012, Trembly-Boyer *et al.,* 2019, Rice *et al.,* 2021). Details of the presentation covered the use of multiple axes of uncertainty, the use of the results to identify key uncertainties, and how inter-assessment research led to reduced uncertainty. The presentation discussed the use of weights on individual model runs and how to interpret the results of a weighted group of runs. The presentation demonstrated the use of a structural uncertainty grid to evaluate potential management actions and reduce, over time, assessment uncertainty.

The Group noted that in addition to evaluating alternative catch trends within a structural uncertainty grid, other methods exist for including error in the catch such as a catch multiplier in Stock Synthesis. The Group noted that for a recent ICCAT white marlin stock assessment (Anon., 2020), the use of the catch multiplier feature in Stock Synthesis (in the years following regulation and allowing the catch multiplier to be year specific) provided a plausible explanation for a flat CPUE trend while the observed catch was decreasing. Consequently, the Group suggested that in cases where species identification or landings uncertainty is an issue, the use of a catch multiplier (introducing error) could be used in testing.

The Group discussed the merits of alternative grid weighting determined using an expert opinion (WCPFC oceanic whitetip example) versus other methods such as equal weighting typically adopted by ICCAT. The author noted that weights assigned to grids in the WCPFC oceanic whitetip example were developed by the analytical team to down weight marginal life history characteristics within the grid (such as those associated with high M and high steepness which were determined to have a low plausibility for long-lived and low fecundity sharks), and were subsequently adopted by the scientific committee associated with the WCPFC assessment process. The Group discussed whether it would be possible to use the assessment results to get plausibility values and use those as weights within a structural uncertainty grid. The Group noted that the choice of which parameters, or model formulations, to include within a structural uncertainty evaluation is an important consideration that is not reflected in the alternative grid weightings and that the uniqueness of alternative model formulations may not be captured in alternative grid weightings.

The Group noted and then discussed the merits of using an iterative approach to reduce assessment uncertainty over time. For example, the IOTC blue shark structural uncertainty grid identified catch as a major uncertainty, which was subsequently reduced by over time with focused intersessional research to reduce uncertainty in catch and life history, which reduced the range of uncertainty in subsequent assessments.

The Group noted that for data-poor stock assessment models many uncertain parameters may need to be fixed, including the stock recruitment steepness parameter (h) and natural mortality parameter (M). In such cases, the profile likelihood plots of estimated parameters such as the equilibrium unfished recruitment, R0 against fixed parameters, such as h and M, are not directly comparable to stock assessment models that do freely estimate such parameters.

Presentation SCRS/P/2022/023 provided examples of model validation diagnostics for integrated stock assessments. The methods presented are applicable to multiple modelling frameworks e.g. Bayesian biomass dynamic as well as integrated assessments models. Multiple ICCAT Species Groups, and RFMOs have identified the need for objective criteria to assess model plausibility and validation during stock assessment. Akaike information criterion (AIC), retrospective bias and model residuals are commonly used as model diagnostics. However, AIC-based selection is not suitable for comparing across models with different datasets and weighting, residual patterns can be removed by misspecification, and retrospective patterns can be removed by the data. Therefore, neither alone can be used for validation, which requires assessing whether it is plausible that a system equivalent to the model generated the data (Kell *et al*. 2021). Validation, therefore, requires that the system is observable and measurable, and observations should be used, rather than model-based quantities (e.g. SSB, recruitment, or F) unless these are well known. Validation is, therefore, best performed by hindcasting where predictions of observations not used in fitting are compared to their known values.

It was noted that this presentation was based on a presentation previously provided to the Center for the Advancement of Population Assessment Methodology (CAPAM) workshop on Model Diagnostics in Integrated Stock Assessments (online, 31 January – 3 February 2022) following which the authors (Kell *et al.*, 2022) were asked to develop guidelines for the use of the hindcast as part of the selection, rejection, weighting and extension of models in ensembles.

The Group discussed that where there is large structural uncertainty about system processes, elicitation (e.g., such from a basic data review to drive hypothesis development and testing of alternative model scenarios) can be used to identify, describe, and catalogue the major sources of uncertainty that are likely to impact management objectives. The Group discussed that if structural model uncertainty among plausible model scenarios obtained from hypothesis development and testing of alternative scenarios substantially exceeds the estimation uncertainty from a single model, then ensembles should be developed, and simple prediction skill weighting could be used to integrate across models when providing advice. However, the Group discussed that an ensemble may be biased if selected models are a subset of all plausible models or are not unique. Within this context, it was noted that those weighting schemes based on prediction skill and uniqueness (Sanderson et al., 2017) have been shown to perform well for modelling climate.

It was noted that this presentation was also previously presented during the 2022 Intersessional Meeting of the Shark Species Group (Online, 16-18 May 2022) in preparation for an upcoming blue shark assessment (Anon., 2022). It was noted that this presentation was developed in response to the Recommendation in 2021. It was noted that feedback was requested from the Group on developing ensemble model guidelines based on prediction skills in general for the upcoming blue shark assessment as a test case.

The Group identified topics for further discussion following the presentation including: 1) Developing ensemble model guidelines based on prediction skills and developing model validation procedures based on prediction skills are both areas of ongoing research within fish stock assessment; 2) Model validation procedures based on prediction skill may be difficult to operationalize within the time constraints of a linear Species Group stock assessment timeline (typically a data preparatory meeting followed by an assessment meeting); 3) Model validation procedures based on prediction skill would be beneficial if they are robust and transparent and can be shared across Species Groups, and 4) Development of model validation procedures based on prediction skill could also benefit from being iteratively updated based on lessons learned from their application across Species Groups and other RFMOs and CAPAM.

The Group discussed that it may be difficult to interpret mean absolute scaled error (MASE) statistic results when there is data conflict among CPUE indices, when there is excessive smoothing of annual variability in a CPUE series, when CPUE series are short and noisy, and when the data are correct but the model may be wrong, among other situations. The Group discussed how the MASE statistic results should be interpreted within the context of other common model diagnostics as well as basic data review. It was agreed that model validation is a research area that would benefit from collaboration with others, as well as simulation studies (e.g. Carvalho, *et al*., 2017).

SCRS/P/2022/022 presented analyses of trends in the estimated recruitment deviations on the tropical tuna stock assessments with a view to developing a new diagnostic for integrated models. Results of the case studies indicated that in general, when there is a statistically significant increasing trend in recruitment deviations, productivity parameters (unfished recruitment, MSY) are estimated to be low compared to recent catch history. This indicates that deviates may be compensating for the loss of biomass in periods of high catch even though they are estimating lower productivity. Results from the simulation study corroborated the case studies and further indicated that trends in recruitment deviates can be caused by bias in the biological parameters used as fixed values in integrated assessment models. Comparisons between the recruitment deviations diagnostic and other diagnostics showed large consistency agreement with the age-structured production model (ASPM) diagnostic, however not with MASE, Mohn's rho, or the runs tests. The authors suggest that the recruitment deviation diagnostic can provide statistical support for hypotheses and assumptions when selecting ensembles of models to develop fisheries management advice. The Group noted that if the alpha significance level were to be dropped to its more common 0.05 value, fewer trends would be identified as significant. This would lead to fewer failures of the diagnostic.

The Group also discussed the consideration of recruitment deviations in stock assessment projections, emphasizing that they need be considered if there is a trend, or productivity may be under or over estimated and catch advice overly conservative or risky. Although it was also noted that if used as a diagnostic, a model run with a trend in residuals would fail the diagnostic and might not be included for use in projections.

The Group discussed whether the trend in recruitment deviates could simply be caused by environmental effects that results in a trend in productivity, however others argued that this should be modeled directly, rather than captured through recruitment deviations which are meant to be unbiased process errors.

The Group discussed that the recruitment deviations showed largely consistent agreement with the ASPM diagnostic (Carvalho et al., 2017) and was faster to run. However, a point was made that recruitment deviates should not be used in isolation for validation for model weighting because they are not observed data (Kell *et al.*, 2021 and Carvalho *et al.,* 2021). Recruitment is likely to vary independently of spawning fish abundance, due to environmental conditions and species interactions (Cury *et al*., 2014). For example, for tropical tunas, Wu *et al.,* (2020) showed that the Atlantic Multidecadal Oscillation (AMO) is likely to affect the recruitment, growth rates, and abundances of yellowfin. Lehodey *et al*., (2003) in the Central and Western Pacific showed that there was high recruitment during La Niña, and Dell'Apa *et al.*, (2018) showed that increasing Sea Surface Temperature in the Gulf of Mexico can potentially increase habitat suitability for larvae of tropical skipjack tuna. It was noted that in integrated assessments, process error is mainly modelled by the recruitment deviates, but random variation in other processes such as M (natural mortality) can occur.

The Group discussed that caution was therefore suggested as the characteristics of time-series of stock dynamics may be determined by the model used to generate them, rather than by underlying ecological phenomena. This can be especially true when information about cohort abundance is noisy or lacking (Dickey-Collas, *et al.,* 2015), as is often the case in tuna assessments. Therefore, it was argued that while looking at trends in recruitment deviates from integrated assessments is of value in flagging concerns it should be used as a symptom of problems rather than for weighting ensembles. For example, by prioritizing the use of an age-structured production model (ASPM) as a diagnostic of processes that control the expected dynamics through a production function.
The Group supported highlighting the following three discussion topics:

- Coupling the Jackknife procedure (removal of one data source at a time) with standard diagnostics to evaluate data conflicts and model misspecification.

- Assessment teams should write a concise summary of the structural uncertainties that were identified during the assessment and that were not properly considered in the uncertainty grid used for management advice. If time allows, this should be done during the assessment meeting and incorporated in a special section of the meeting report. If this is not done during the assessment meeting it should be done in a short SCRS paper to be presented during the Species Group meeting in September 2022.

- In reference to trends of the recruitment deviation presentation, the Group highlighted the consistency in agreement between the recruitment deviation diagnostics and the age-structured production model (ASPM), and that both may be useful as a diagnostic of processes to be included in stock assessment runs to help diagnose model misspecification, in combination with other diagnostics such as those in Carvalho *et al.,* (2021). The Group also recommended that this diagnostic could benefit from further simulation testing.

4. Bycatch estimation

SCRS/2022/105 presented a newly developed R library for semi-automated total bycatch estimation using model-based and design-based estimators. The method was tested using simulated observer and logbook data from three simulated fleets in the Atlantic generated by LLSIM (longline simulator, Goodyear *et al.,* 2021) with a decreasing trend in blue marlin abundance over time. Simulations included a range of observer coverage levels, with random allocation or biased allocation proportional to swordfish or blue marlin catch. The bycatch estimation tool performed well and produced reasonably unbiased estimates of the total blue marlin bycatch for all methods. Precision improved with coverage as expected, particularly for the ratio estimator and for the model-based estimators when they were only used to predict the bycatch in unsampled trips. The estimated variances were approximately correct, particularly for the models preferred by the model selection algorithm using cross-validation.

The Group discussed the details of the bycatch estimation tool, and what the next steps should be for the project.

In general, the amount of bias in the estimates was driven more by the observer allocation scheme than the specific allocation method. If an observer program has non-random sampling, it may be possible to correct for bias by including a model of the probability of each trip being sampled, but the tool does not currently include this option.

The tool does not currently impute bycatch ratios in unsampled strata for the stratified ratio estimator. Some kind of imputation method (e.g. borrowing bycatch rates from nearby strata) would be needed to apply this method to real data from most fleets, since it is common for both fishing effort and observer coverage to be distributed such that some strata (e.g. seasons, spatial areas) are unlikely to be sampled in all years.

When simulated observer coverage levels increased over time to mimic the historical expansion of observer coverage in Atlantic longline fisheries, the increasing coverage improved the precision of the estimates, despite the fact that marlin was becoming rarer in the catch. It would be interesting to run the simulations with a flat trend in blue marlin abundance instead of decreasing to separate the effects of changing observer coverage levels and the effects of changing abundance. It is possible that different estimation methods would perform better with different trends in abundance.

More variables could be added to the model, such as the LLSIM habitat variable, which would be highly informative about the probability of catching blue marlin. This would require set-by-set rather than trip-by-trip analysis and would only be realistic if such data were available for a particular fishery.

The method uses prediction intervals rather than the more restrictive confidence intervals of the mean to generate the predictions of total bycatch because the model is used to predict new catches in the unsampled trips. This is different from CPUE indices, in which the standard error of the mean is used as a measure of precision. The tool could be useful for comparing the estimated standard errors in CPUE indices across methods, as this standard error is a key input to stock assessment.

The tool could use target species catches rather than the effort as the denominator when expanding bycatch. This might be useful for non-target species like some sharks or some small tunas. Using the LLSIM data, swordfish is not a good candidate for the expansion variable for blue marlin bycatch, because historical fishing mortality in the longline fishery of swordfish is not particularly correlated with the longline fishing mortality of blue marlin. LLSIM includes a third species, a generic tuna, which might be a good expansion variable.

An important next step is to apply the method to real CPC data. There is also a need for beta testing, and training. These could all be combined, in that CPC scientists or Species Group members with data could work with the developers to learn the tool, apply it to their data, and identify any issues that need to be addressed in the software development. Having CPC scientists apply the tool to their own data would be one solution to the problem that observer data is often confidential. They could also supply data for the developers to apply the tool with a confidentiality agreement. It is important to have participation by the data holders so the tool can be made robust to the specific issues that come up in each data set.

There may be some overlap with the proposed study group on CPUE standardization, such as the development of diagnostics. Semi-automated diagnostics for CPUE standardization could be added to the tool, including new diagnostics proposed by the study group. The tool currently has an option to exclude methods that fail a test of whether the data are consistent with the assumed error model from the DHARMa R library (residual diagnostics for hierarchical (multi-level/mixed) regression models). However, this tends to be too restrictive with large datasets. More diagnostic plots could be added, and the tool could be used to identify red flags that indicate mis-specified models for both CPUE standardization and bycatch estimation. Many diagnostics used in CPUE standardization are available.

The tool does not currently include all the methods used in CPUE standardization, such as random effects for interactions, GAMs (generalized additive model), or geostatistical models. These could be added but would take some time, particularly for the geostatistical models. Geostatistical models are useful for both CPUE standardization and bycatch estimation, particularly for species with range shifts. R packages that do this, such as VAST, INLA, or sdmTMB (Anderson *et al.* 2022, Rue *et al.* 2009, Thorson 2019) could potentially be integrated with the bycatch estimation tool.

The simulations could be made more accurate to what CPCs actually do, including how the observer coverage is allocated across fleet components and the number of sets in a trip. This would benefit from the participation of CPC scientists well familiar with the National observer programs.
There is a need to make the tool more widely useful, by adding an online training module or online class and applying it to more types of fisheries.

5. CPUE standardization/incorporation of oceanographic and environmental changes into the assessment process

5.1 Characterization of multi-decadal oceanographic indicators

Presentation SCRS/P/2022/039 described an investigation of the projected Atlantic Multidecadal Oscillation (AMO), an index traditionally defined using sea surface temperature, at depths inhabited by highly migratory species. When recreating the sea surface AMO index from a new climate dataset, it was discovered that the signal did not align with others previously published. It became evident that the AMO is expressed differently dependent on how the simulated or observed dataset was created. The AMO signal was then characterized at several depths. Signals were analyzed for regime shifts of statistically significant warm and cool phases. Notable results include a compression of the signal with depth as well as a delayed regime shift of cool to warm phase with depth. Signals at depth from a peak warm and cool phase year were mapped for a spatial temperature comparison illustrating a greater temperature discrepancy at lower latitudes in the North Atlantic as well as towards the sea surface.

The Group acknowledged the concerns of the large variations in AMO signals for fisheries inference and the lack of consistency in the AMO index dependent on the dataset, time scale and detrending methods. A conversation ensued on the greater concern of poorly understood uncertainty in environmental data used to inform fisheries management. The AMO has been used in previous fisheries studies and assessments such as for swordfish in the North Atlantic. There was a debate on wether the trends would be similar if another dataset was used in the analysis and it was thought that the major trends may be similar while the details tend to have the most variability.

The Group raised the issue of the AMO being incorporated into species distribution models, CPUE standardization and abundance estimations. There was a debate as to whether it is still appropriate to incorporate the AMO due to the newly understood uncertainty. It was argued to be appropriate for use only if a full understanding and a clear hypothesis exist before incorporation. The question about what dataset is best to use going forward is ongoing. Subsequently, the Group emphasized the importance of clearly stating the data used and all assumptions and accompanying uncertainty of the model when using environmental data for fisheries.

SCRS/2022/106 presented a Management Strategy Evaluation framework called "EcoTest" to inform decision makers of Ecosystem Based Fisheries Management (EBFM). A multi-species framework that supports tactical decision making can make significant progress toward the essential goals of EBFM. The authors demonstrated the use of EcoTest using the Atlantic longline fishery as a case study. This is an extension to openMSE software, used for single-species modeling, that simulates multi-species fisheries dynamics. A range of features are possible in EcoTest, such as the ability to evaluate current indicators as well as design new indicators and identify the conditions under which indicators operate reliably.

The Group was informed that this document was also presented at the concurrent Subcommittee on Ecosystems meeting because it is important to establish challenging and plausible scenarios for ecosystem dynamics and then test whether current and potential indicators can reflect stock status.

This proof-of-concept application was based on the dynamics of multiple species caught in Atlantic longline fisheries, for which Stock Synthesis assessments were available. In the projection scenarios, there was a strong correlation between the fishing mortality rates of the target species and indicators such as mean length and CPUE for the secondary species, implying that multispecies indicators may be informative for EBFM (Ecosystem-Based Fisheries Management).

The Group discussed the use of this tool and possible future developments. There is a need to develop new tools to operationalize EBFM, and this tool provides a useful step forward.

Future work may include full management strategy evaluation with a closed loop, and cross-validation to test proposed indicators. A spatial operating model would be useful to incorporate targeting or avoidance of species and range changes. Spatial correlation is complex, in that some target and secondary species are correlated with each other in some places but not others. There is a need to consolidate the indicators used in the current Eco-card into specific information on management-relevant measures of trends, perhaps using analytic methods such as neural nets.

In response to a question about why the tool was based on stock assessment models rather than ecosystem models (e.g. Atlantis, EcoSpace), the authors explained that the model is intended to be able to pass the stringent peer reviews of the stock assessment process and to be consistent with the assessments. Ecosystem models may be too slow and complex to be useful tools for this context, and it is difficult to control the statistical properties of data generated by ecosystem models. The EcoTest framework can be set up with an appropriate level of complexity, including elements like species interactions and habitat, to test hypotheses related to ecosystem processes. Lack of progress in EBFM using models is partly due to the difficulty of using ecosystem models and getting them to pass review through the stock assessment process with appropriate diagnostics. Models must be responsive to the information needed for management, rather than necessarily the best models from a scientific perspective.

The framework can include uncertainty by incorporating multiple assessment models for each species. If the grid of uncertainty is large, due to having multiple assessment models for each species, it is possible to sample the grid to get stable conclusions. The method would benefit from collaborations with the species groups to identify the key uncertainties including the range of assessment models that should be included for each species.

5.2 CPUE standardization diagnostics

The Group was presented with a proposal for the creation of a study group tasked with developing standards and best practices for CPUE model diagnostics. The presenter noted that there are many strengths to the current ICCAT approach whereby CPC scientists prepare indices as assessment inputs, however, there is currently a lack of guidance on appropriate model diagnostics for these CPUEs. This can sometimes lead to CPUE papers requiring revisions and delays in the adoption of indices. The presenter briefly reviewed some of the diagnostics currently in place and emphasized the need for minimum standards for these diagnostics. Noting existing species distribution models and a longline simulator created by members of the Group, the presenter and Chair suggested that these data sets be used for a simulation study. In this scenario, the study group would work alongside WGSAM Group members currently engaged in bycatch estimation analyses using the longline simulator datasets with known properties. Differing error models would be evaluated using various diagnostic tools to examine which diagnostics are most useful for model validation. The study group deliverables could include an SCRS paper describing minimum reporting requirements for CPUE analyses and a peer reviewed paper with results from the simulation study.

The Group welcomed the proposal for the creation of the study group with some noting the importance of broad membership - particularly from those who are not currently represented in WGSAM study groups. There was agreement that the WGSAM Chair and study group lead would seek out participants to be involved in this effort. The study group is open to all interested participants, indicating their interest by email to the Chair, and communications will be through the WGSAM. It was also suggested that this effort could lead to CPUE capacity building workshop(s) in the future but that these would benefit from participation from developing regions from the outset of the work.

6. Recommendations

Recommendation with financial implications

1. The Group recognized the usefulness of the bycatch estimation tool presented to the Group and recommended it continue to be funded for further development as a means to address the SCRS general needs to estimate bycatch of species such as, but not limited to, billfish and sharks. The Group further recommends that this work be carried out using the 2022 WGSAM funds.

2. The Group continued to recommend an overall review of ICCAT MSE activities by an external expert. This overall review would serve to identify potential improvements, highlight any missing components or shortcomings of the current process, achieve efficiencies across species and promote standardization of the MSE process across species, refine and standardize MSE communication and stakeholder engagement, and provide guidance on what the future of MSE looks like within ICCAT. This could include the way MSEs are supported and how resources are divided, and how the MSE process should be structured and supported after MP adoption. The amount of funding required for this recommendation should be estimated by the ICCAT Secretariat in coordination with the SCRS Chair*.

Working Group on Stock Assessment Methods 2022	**2023**
Other fisheries related studies (including data recovery, experts, etc.)	
Tool for estimate bycatch of species	**€35,000**
Overall review of ICCAT MSE activities*	
TOTAL	**€35,000**

General recommendations

1. The Group recognized the need for further attention to the manner in which CPUE diagnostics are formulated and presented for use in ICCAT stock assessments. In response, the WGSAM recommended the formation of a study group on CPUE diagnostics. The study group would work closely with the bycatch estimation tool contractor to develop a path forward to improve CPUE diagnostic interpretation, creation of guidelines and best practices. The study group should make a concerted effort to ensure involvement of scientists from several CPCs via direct invitation and outreach.

2. The Group recommended that a dedicated Limit Reference Point study group be formed to investigate how Reference Points should be identified in general as well as for each species. The Group should consider how Limit Reference Points should be calculated across life history strategies, and if they feel so inclined, could expand their investigation beyond Limit Reference Points to more broadly consider also Target Reference Points (e.g. MSY), including dynamic, time-varying Reference Points.

3. Assessment teams should write a concise summary of the structural uncertainties that were identified during the assessment and that were not considered in the uncertainty grid used for management advice. If time allows it, this should be done during the assessment meeting and incorporated in a special section of the meeting report. If this is not done during the assessment meeting it should be done in a short SCRS paper to be presented during the Species Group meeting in September.

4. The Group recommended that the scheduling of the WGSAM meeting concurrent with the Subcommittee on Ecosystems meeting be discontinued. The concurrent timing was an impediment to the progress of both groups. CPC scientists were forced to attend either one or the other of the meetings despite the relevance of their contributions to each meeting. While the Group did see some utility of having perhaps one overlapping day, the advantages of this option would be dependent on the overlap of the two groups' specified agenda items.

7. Other matters

The Chair presented the tentative terms of reference (TORs) for the continued investigation into bycatch estimation methodologies to be developed in the second half of 2022 and throughout 2023, that the Group adopted. The Group also discussed the workplan and its budget estimate for 2023, and they were adopted during the meeting (**Appendix 5**).

A technical matter was brought to the attention of the Group, related to the estimation of stock status and Kobe matrices using different assessment model platforms. The Group recommends that modelers be aware of what biomass and fishing mortality are reported in the results and ensure that biomass and fishing mortality corresponds to the correct reference year in both Kobe plots, matrices, and summary tables. In some platforms, in order to account for landings in the terminal year of the model on biomass estimates, a minimum 1-yr simple forecast is needed to generate the SSB at the end of the terminal year (beginning of the first forecast year). In these cases, the stock status estimate for the terminal year, t, should be based on the F ratio estimate for year t and the biomass ratio estimate for year t+1. The Group recommended analysts define the biomass estimate timing and show diagnostics to account for time lags and that the timing remain consistent throughout the stock assessment document (i.e. figures, tables and text) and between modeling platforms used for that stock assessment.

8. Adoption of the report and closure

The Report of the 2022 Intersessional Meeting of the Working Group on Stock Assessment Methods was adopted. Dr Michael Schirripa, the SCRS Chair, and the ICCAT Executive Secretary thanked the participants and the Secretariat for their hard work and collaboration to finalize the report on time. The meeting was adjourned.

References

Anderson, S.C., E.J. Ward, P.A. English, L.A.K. Barnett. 2022. sdmTMB: an R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. bioRxiv 2022.03.24.485545; doi: https://doi.org/10.1101/2022.03.24.485545.

Anonymous. 2015. IOTC Resolution 15/10 on target and limit reference points and a decision framework.

Anonymous. 2020. Report of the 2019 ICCAT white marlin stock assessment meeting (Miami, USA, 10-14 June 2019). Collect. Vol. Sci. Pap. ICCAT, 76(4): 97-181.

Anonymous. 2022. Intersessional Meeting of the Shark Species Group (Online, 16 - 18 May 2022).

Carvalho, F., Punt, A.E., Chang, Y.J., Maunder, M.N. and Piner, K.R., 2017. Can diagnostic tests help identify model misspecification in integrated stock assessments? Fisheries Research, 192, pp.28-40.

Carvalho, F., Winker, H., Courtney, D., Kapur, M., Kell, L., Cardinale, M., Schirripa, M., Kitakado, T., Yemane, D., Piner, K. R., Maunder, M. N., Taylor, I., Wetzel, C. R., Doering, K., Johnson, K. F., and R. D. Methot. 2021. A Cookbook for Using Model Diagnostics in Integrated Stock Assessments. Fisheries Research, 240, 105959. https://doi.org/10.1016/j.fishres.2021.105959.

Cury, P.M., Fromentin, J.M., Figuet, S. and Bonhommeau, S. 2014. Resolving Hjort's dilemma: how is recruitment related to spawning stock biomass in marine fish? *Oceanography*, *27*(4), pp.42-47.

Dell'Apa, A., Carney, K., Davenport, T.M. and Carle, M.V. 2018. Potential medium-term impacts of climate change on tuna and billfish in the Gulf of Mexico: A qualitative framework for management and conservation. *Marine Environmental Research*, *141*, pp.1-11.

Dickey-Collas, M., Hintzen, N.T., Nash, R.D.M., Schön, P.J. and Payne, M.R. 2015. Quirky patterns in time-series of estimates of recruitment could be artefacts. *ICES Journal of marine Science*, *72*(1), pp.111-116.

Goodyear, C.P. 2021. Development of new model fisheries for simulating longline catch data with LLSIM. Collect. Vol. Sci. Pap. ICCAT, 78(5): 53-62.

Kell, L. T. Mosqueira I., De Bruyn P., and Magnusson A. 2012. An evaluation of limit and target reference points as part of a Harvest Control Rule: an Atlantic swordfish example. Collect. Vol. Sci. Pap. ICCAT, 68(4): 1630-1644.

Kell, L. T., Sharma, R., Kitakado, T., Winker, H., Mosqueira, I., Cardinale, M., and D. Fu. 2021. Validation of stock assessment methods: is it me or my model talking? ICES Journal of Marine Science. ICES Journal of Marine Science, 78:2244–2255.

Kell, L. T., Mosqueira, I., Winker, H., Sharma, R., Kitakado, T., and M. Cardinale. [accepted 2022]. Validation of integrated stock assessment model ensembles. Fish and Fisheries.

Lehodey, P., Chai, F. and Hampton, J. 2003. Modelling climate-related variability of tuna populations from a coupled ocean–biogeochemical-populations dynamics model. *Fisheries Oceanography*, *12*(4-5), pp.483-494.

Preece A., Hillary R., and Campbell D. 2011. Identification of candidate limit reference points for the key target species in the WCPFC. Scientific Committee Seventh Regular Session 9-17 August 2011At: Pohnpei, Federated States of Micronesia.

Rice, J., and R. Sharma. 2015. Stock assessment of blue shark (*Prionace glauca*) in the Indian Ocean using Stock Synthesis. IOTC-2015-WPEB11-28 Rev_1. Working paper submitted to the 11th Session of the Indian Ocean Tuna Commission (IOTC) Working Party on Ecosystems and Bycatch (WPEB), 7–11 September 2015, Olhão, Portugal. Available: http://www.iotc.org/documents/stock-assessment-blue-shark-pironce-glauca-indian-ocean-using-stock-synthesis (October 2018).

Rice, J. 2017. Stock assessment blue shark (Prionace glauca) in the Indian Ocean using Stock Synthesis. IOTC–2017–WPEB13–33. Working paper submitted to the Indian Ocean Tuna Commission 13th Working Party on Ecosystems and Bycatch, 9-11 March 2017, San Sebastián, Spain. Available: http://iotc.org/documents/stock-assessment-blue-shark-prionace-glauca-indian-ocean-using-stock-synthesis (October 2018).

Rice, J. 2021. Stock assessment of blue shark (*Prionace glauca*) in the Indian Ocean using Stock Synthesis. IOTC-2021-WPEB17(AS)-15_Rev1. Working paper submitted to the Indian Ocean Tuna Commission 17th Working Party on Ecosystems and Bycatch6 – 10 September 2021.Virtual Meeting. Available: https://iotc.org/sites/default/files/documents/2021/10/IOTC-2021-WPEB17-15_Rev2.pdf

Rice, J., and Harley S. 2012. Stock assessment of oceanic whitetip sharks in the western and central Pacific Ocean WCPFC-SC8-2012/SA-WP-06 Rev 1. Working paper submitted to the Western and Central Pacific Fisheries Commission (WCPFC) Science Committee Eighth Regular Session (SC8), 7-15 August 2012 Busan, Republic of Korea. Available: https://www.wcpfc.int/doc/sc8-sa-wp-06/oceanic-whitetip-shark-stock-assessment (October 2018).

Rice J., Carvalho F., Fitchett M., Harley S., and Ishizaki A. 2021. Future stock projections of oceanic whitetip sharks in the western and central Pacific Ocean. WCPFC-SC17-2021/SA-IP-21. Seventeenth regular session of the Scientific Committee. Electronic meeting 11-19 August 2021.

Rue,H., S. Martino, and N. Chopin. 2009. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). Journal of the Royal Statistical Society, Series B,71(2):319-392.

Sanderson, B. M., Wehner, M., and Knutti, R. 2017. Skill and independence weighting for multi-model assessments, Geosci. Model Dev. 10:2379–2395. https://doi.org/10.5194/gmd-10-2379-2017

Thorson, J. T. 2019. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fisheries Research 210,143–16.1.

Tremblay-Boyer, Laura; Felipe Carvalho; Philipp Neubauer; Graham Pilling (2019). Stock assessment for oceanic whitetip shark in the Western and Central Pacific Ocean, 98 pages. WCPFC-SC15-2019/SA-WP-06. Report to the WCPFC Scientific Committee. Fifteenth Regular Session, 12–20 August 2018, Pohnpei, Federated States of Micronesia.

Wu, Y.-L., K.-W. Lan, *et al.,* 2020. "Determining the effect of multiscale climate indices on the global yellowfin tuna (*Thunnus albacares)* population using a time series analysis." Deep Sea Research Part II: Topical Studies in Oceanography 175:104808.

Table 1. ICCAT MSE funding over the past 9 years by process and source.

<table>
<tr><th>Year/ MSE</th><th>Bluefin tuna</th><th>North Atlantic Albacore</th><th>North Atlantic Swordfish</th><th>Tropical tunas (Multi-species+ Western Skipjack)</th></tr>
<tr><td>2014</td><td>GBYP: €119,400</td><td></td><td></td><td></td></tr>
<tr><td>2015</td><td>GBYP: €183,343</td><td></td><td></td><td></td></tr>
<tr><td>2016</td><td>GBYP: €132,965</td><td></td><td></td><td></td></tr>
<tr><td>2017</td><td>GBYP: €89,042</td><td></td><td></td><td></td></tr>
<tr><td rowspan="2">2018</td><td>GBYP: €163,845</td><td>EU voluntary contribution: €9,521
ICCAT regular Budget: €1,639</td><td>EU voluntary contribution: €38,405
ICCAT regular Budget: €6,613</td><td>EU voluntary contribution: €66,477
ICCAT regular Budget: €11,447</td></tr>
<tr><td colspan="4">Regular budget (three MSE training courses to build capacity on Management Strategy Evaluation” – €12,174)</td></tr>
<tr><td>2019</td><td>GBYP: €103,965</td><td>EU voluntary contribution: €9,500
ICCAT regular Budget: €500</td><td>EU voluntary contribution: €51,150
ICCAT regular Budget: €3,850</td><td></td></tr>
<tr><td>2020</td><td>GBYP: €109,850</td><td>EU voluntary contribution: €18,624
ICCAT regular Budget: €5,76</td><td>EU voluntary contribution: €106,700
ICCAT regular Budget: €3,300</td><td>EU voluntary contribution: €14,545
ICCAT regular Budget: €455</td></tr>
<tr><td rowspan="2">2021</td><td>GBYP: €176,123</td><td>EU voluntary contribution: €27,851
US voluntary contribution: €10,149
ICCAT regular Budget: €10,149</td><td>EU voluntary contribution: €130,534
US voluntary contribution: €47,566
ICCAT regular Budget: €47,566</td><td>EU voluntary contribution: €51,305
US voluntary contribution: €18,695
ICCAT regular Budget: €18,695</td></tr>
<tr><td colspan="4">JCAP (three MSE training courses to build capacity on Management Strategy Evaluation methods for fisheries scientists, managers and stakeholders”: €21,176</td></tr>
<tr><td>2022</td><td>GBYP: €160,000</td><td>€20,000</td><td>€90,000</td><td>€70,000</td></tr>
<tr><td>Sub-Total</td><td>€1,259,533</td><td>€108,509</td><td>€525,684</td><td>€251,619</td></tr>
<tr><td>TOTAL</td><td colspan="4">€2,145,345</td></tr>
</table>

Appendix 1

Agenda

1. Opening, adoption of agenda and meeting arrangements and assignment of rapporteurs
2. Harvest Control Rules, Limit Reference Points and Management Strategy Evaluation
3. Stock assessment
    3.1 Validation of integrated stock assessment model ensembles
4. Bycatch estimation
5. CPUE standardization/incorporation of oceanographic and environmental changes into the assessment process
    5.1 Characterization of multi-decadal oceanographic indicators
    5.2 CPUE standardization diagnostics
6. Recommendations
7. Other matters
8. Adoption of the report and closure

Appendix 2

List of Participants

CONTRACTING PARTIES

CANADA

Gillespie, Kyle
Aquatic Science Biologist, Fisheries and Oceans Canada, 125 Marine Science Drive, St. Andrews, NB, E5B 0E4
Tel: +1 506 529 5725, E-Mail: kyle.gillespie@dfo-mpo.gc.ca

Carruthers, Thomas
2150 Bridgman Ave, Vancouver Columbia V7P 2T9, Canada
Tel: +1 604 805 6627, E-Mail: tom@bluematterscience.com

CHINA, (P.R.)

Feng, Ji
Shanghai Ocean University, 999 Hucheng Huan Rd, 201306 Shanghai
Tel: +86 159 215 36810, E-Mail: fengji_shou@163.com; 276828719@qq.com; f52e@qq.com

Huang, Yucheng
Shanghai Ocean University, 999 Hucheng Huan Road, Shanghai, 201306
Tel: +86 177 989 21637, E-Mail: yuchenhuang0111@163.com

Yang, Shiyu
Shanghai Ocean University, 999 Hucheng Huan Road, Shanghai, 201306
Tel: +86 185 021 91519, E-Mail: yangshiyu_shou@163.com

Zhang, Fan
Shanghai Ocean University, 999 Hucheng Huan Rd, 201306 Shanghai
Tel: +86 131 220 70231, E-Mail: f-zhang@shou.edu.cn

EL SALVADOR

Chavarría Valverde, Bernal Alberto
Asesor en Gestión y Política Pesquera Internacional, Centro para el Desarrollo de la Pesca y Acuicultura (CENDEPESCA), Final 1ª Avenida Norte, 13 Calle Oriente y Av. Manuel Gallardo, 1000 Santa Tecla, La Libertad
Tel: +506 882 24709, Fax: +506 2232 4651, E-Mail: bchavarria@lsg-cr.com

Galdámez de Arévalo, Ana Marlene
Jefa de División de Investigación Pesquera y Acuícola, Ministerio de Agricultura y Ganadería, Final 1a. Avenida Norte, 13 Calle Oriente y Av. Manuel Gallardo. Santa Tecla, La Libertad
Tel: +503 2210 1913; +503 619 84257, E-Mail: ana.galdamez@mag.gob.sv; ana.galdamez@yahoo.com

EUROPEAN UNION

Biagi, Franco
Senior Expert Marine & Fishery Sciences, Directorate General for Maritime Affairs and Fisheries (DG-Mare) - European Commission, Unit C3: Scientific Advice and data collection, Rue Joseph II, 99, 1049 Brussels, Belgium
Tel: +322 299 4104, E-Mail: franco.biagi@ec.europa.eu

Barciela Segura, Carlos
ORPAGU, C/ Manuel Álvarez, 16. Bajo, 36780 Pontevedra, Spain
Tel: +34 627 308 726, E-Mail: cbarciela@orpagu.com; septimocielo777@hotmail.com

Celona, Antonio
NECTON Marine Research, V. le A. de Gaperi 187, 98165 Catania, Italy
Tel: +39 339 124 9021, E-Mail: nectonricerche@gmail.com; info@necton.it

Merino, Gorka
AZTI - Tecnalia /Itsas Ikerketa Saila, Herrera Kaia Portualdea z/g, 20100 Pasaia - Gipuzkoa, Spain
Tel: +34 94 657 4000; +34 664 793 401, Fax: +34 94 300 4801, E-Mail: gmerino@azti.es

Ortiz de Urbina, Jose María
Ministerio de Ciencia, Innovación y Universidades, Instituto Español de Oceanografía, C.O de Málaga, Puerto Pesquero s/n, 29640 Fuengirola, Málaga, Spain
Tel: +34 952 197 124, Fax: +34 952 463 808, E-Mail: urbina@ieo.csic.es

Ortiz de Zárate Vidal, Victoria
Investigadora, Ministerio de Ciencia, Innovación y Universidades, Instituto Español de Oceanografía, C.O. de Santander, Promontorio de San Martín s/n, 39004 Santander, Cantabria, Spain
Tel: +34 942 291 716, Fax: +34 942 27 50 72, E-Mail: victoria.zarate@ieo.csic.es

Rueda Ramírez, Lucía
Ministerio de Ciencia, Innovación y Universidades, Instituto Español de Oceanografía Málaga, Puerto pesquero s/n, 29640 Fuengirola Málaga, Spain
Tel: +34 952 197 124, E-Mail: lucia.rueda@ieo.csic.es

Winker, Henning
Joint Research Centre - European Commission, Ispra, Italy, TP 051, Via Enrico Fermi 2749, 21027 Ispra, VA, Italy
Tel: +39 351 525 2370, E-Mail: henning.winker@ec.europa.eu

MOROCCO

Abid, Noureddine
Chercheur et ingénieur halieute au Centre Régional de recherche Halieutique de Tanger, Responsable du programme de suivi et d'étude des ressources des grands pélagiques, Centre régional de l'INRH à Tanger/M'dig, B.P. 5268, 90000 Drabed, Tanger
Tel: +212 53932 5134; +212 663 708 819, Fax: +212 53932 5139, E-Mail: nabid@inrh.ma; noureddine.abid65@gmail.com

Bensbai, Jilali
Chercheur, Institut National de Recherche Halieutique à Casablanca - INRH/Laboratoires Centraux, Ain Diab près du Club équestre OULAD JMEL, Rue Sidi Abderrhman / Ain Diab, 20100 Casablanca
Tel: +212 661 59 8386, Fax: +212 522 397 388, E-Mail: bensbaijilali@gmail.com

Serghini, Mansour
Institut national de recherche halieutique, Route Sidi Abderrahmane Club équestre Ould Jmel, 20000 Casablanca
Tel: 0660 455 363, E-Mail: serghini@inrh.ma; serghini2002@yahoo.com; serghinimansour@gmail.com

Yassir, Anass
Ingénieur Statisticien, Institut National de Recherche Halieutique, Route Sidi Abderrahmane Club équestre Ould Jmel, 20000 Casablanca
Tel: +212 618 392 738, E-Mail: yassir@inrh.ma; yassiranas.insea@gmail.com

SENEGAL

Ba, Kamarel
Docteur en Sciences halieutiques et modélisation, Ministère de l'Agriculture et de l'Equipment Rural, Institut Senegalais de Recherches Agricoles (ISRA), Centre de Recherches Oceanographiques de Dakar Thiaroye (CRODT), Pôle de Recherches de Hann, Route du Front de Terre, 2241 Dakar
Tel: +221 76 164 8128, Fax: +221 338 328 262, E-Mail: kamarel2@hotmail.com

Ndiaye, El Hadji
Direction des Pêches maritimes, 20000 Dakar
Tel: +221 77 543 6301, E-Mail: elhandiaye@yahoo.fr

UNITED KINGDOM OF GREAT BRITAIN AND NORTHERN IRELAND

Walker, Nicola
Centre for Environment, Fisheries and Aquaculture Science (CEFAS), Lowesfolk Suffolk NR33 0HT
Tel: +44 1502 524450, E-Mail: nicola.walker@cefas.co.uk

UNITED STATES

Brown, Craig A.
Chief, Highly Migratory Species Branch, Sustainable Fisheries Division, Southeast Fisheries Science Center, NOAA, National Marine Fisheries Service, 75 Virginia Beach Drive, Miami, Florida 33149
Tel: +1 305 586 6589, E-Mail: craig.brown@noaa.gov

Babcock, Elizabeth
Professor, Rosenstiel School of Marine and Atmospheric Science, University of Miami, Department of Marine Biology and Ecology, 4600 Rickenbacker Causeway, Miami Florida 33149
Tel: +1 305 421 4852, Fax: +1 305 421 4600, E-Mail: ebabcock@rsmas.miami.edu

Courtney, Dean
Research Fishery Biologist, NOAA/NMFS/SEFSC Panama City Laboratory, 3500 Delwood Beach Road, Panama City Beach Florida 32408
Tel: +1 850 234 6541, E-Mail: dean.courtney@noaa.gov

Die, David
Research Associate Professor, Cooperative Institute of Marine and Atmospheric Studies, University of Miami, 4600 Rickenbacker Causeway, Miami, Florida 33149
Tel: +1 305 421 4607, E-Mail: ddie@rsmas.miami.edu

Fisch, Nicholas
Southeast Fisheries Science Center, 101 Pivers Island Road, Beaufort, North Carolina 28516
Tel: +1 727 798 8424, E-Mail: nickcfisch@gmail.com; nicholas.fisch@noaa.gov

Lauretta, Matthew
Fisheries Biologist, NOAA Fisheries Southeast Fisheries Center, 75 Virginia Beach Drive, Miami, Florida 33149
Tel: +1 305 361 4481, E-Mail: matthew.lauretta@noaa.gov

Moffat, Ben
4600 Rickenbacker Causeway, Miami, Florida 33149
Tel: +1 240 688 7472, E-Mail: btm76@rsmas.miami.edu

Peterson, Cassidy
NOAA Fisheries, 101 Pivers Island Rd, Miami, FL 28516
Tel: +1 910 708 2686, E-Mail: cassidy.peterson@noaa.gov

Schirripa, Michael
Research Fisheries Biologist, NOAA Fisheries, Southeast Fisheries Science Center, 75 Virginia Beach Drive, Miami, Florida 33149
Tel: +1 305 445 3130; +1 786 400 0649, Fax: +1 305 361 4562, E-Mail: michael.schirripa@noaa.gov

Walter, John
Research Fishery Biologist, NOAA Fisheries, Southeast Fisheries Center, Sustainable Fisheries Division, 75 Virginia Beach Drive, Miami, Florida 33149
Tel: +305 365 4114; +1 804 815 0881, Fax: +1 305 361 4562, E-Mail: john.f.walter@noaa.gov

OBSERVERS FROM COOPERATING NON-CONTRACTING PARTIES, ENTITIES, FISHING ENTITIES

CHINESE TAIPEI

Lee, Yi-Jou
Phd student, Department of Environmental Biology and Fisheries Science, No. 2 Pei-Ning Rd. Keelung, 202301 Keelung
Tel: +886 2 2462 2192 ext. 5046, E-Mail: li00131022@gmail.com

Su, Nan-Jay
Assistant Professor, Department of Environmental Biology and Fisheries Science, National Taiwan Ocean University, No. 2 Pei-Ning Rd. Keelung, Zhongzheng Dist., 202301
Tel: +886 2 2462 2192 #5046, Fax: +886-2-24622192, E-Mail: nanjay@ntou.edu.tw

COSTA RICA

Lara Quesada, Nixon
Biólogo Marino, INCOPESCA, 125 metros este y 75 metros norte de planta de atún Sardimar, 60101 Puntarenas
Tel: +506 831 12658, E-Mail: nlara@incopesca.go.cr; nixon.lara.21@gmail.com; nlara@incopesca.go.cr

Pacheco Chaves, Bernald
Instituto Costarricense de Pesca y Acuicultura, INCOPESCA, Departamento de Investigación, Cantón de Montes de Oro, Puntarenas, 60401
Tel: +506 899 22693, E-Mail: bpacheco@incopesca.go.cr

Umaña Vargas, Erik
Jefe, Oficina Regional de Limón
E-Mail: eumana@incopesca.go.cr

OBSERVERS FROM INTERGOVERNMENTAL ORGANIZATIONS

INDIAN OCEAN TUNA COMMISSION - IOTC

Fu, Dan
Stock Assessment Expert, IOTC, Victoria, Mahe, Seychelles
Tel: +248 252 5471, E-Mail: dan.fu@fao.org

OBSERVERS FROM NON-GOVERNMENTAL ORGANIZATIONS

EUROPÊCHE

Kell, Laurence
Visiting Professor in Fisheries Management, Centre for Environmental Policy, Imperial College London, Henstead, Suffolk SW7 1NE, United Kingdom
Tel: +44 751 707 1190, E-Mail: laurie@seaplusplus.co.uk; l.kell@imperial.ac.uk; laurie@kell.es

INTERNATIONAL SEAFOOD SUSTAINABILITY FOUNDATION – ISSF

Scott, Gerald P.
FAO Common Oceans /ABNJ Tuna project, 11699 SW 50th Ct, Cooper City, Florida 33330, United States
Tel: +1 954 465 5589, E-Mail: gpscott_fish@hotmail.com

PEW CHARITABLE TRUSTS - PEW

Galland, Grantly
Officer, Pew Charitable Trusts, 901 E Street, NW, Washington, DC 20004, United States
Tel: +1 202 540 6953; +1 202 494 7741, Fax: +1 202 552 2299, E-Mail: ggalland@pewtrusts.org

EXTERNAL EXPERT

Gedamke, Todd
Stuart FL 5521 SE NASSAU TERRACE 34997, United States
Tel: +1 804 684 9522, E-Mail: todd@merconsultants.org

Harford, Bill
President, Nature Analytics, 551 Lakeshore Rd E, Suite 105, Mississauga Ontario L5G 0A8, Canada
Tel: +1 905 452 2113, E-Mail: bill@natureanalytics.ca

SCRS CHAIRMAN

Melvin, Gary
SCRS Chairman, St. Andrews Biological Station - Fisheries and Oceans Canada, Department of Fisheries and Oceans, 285 Water Street, St. Andrews, New Brunswick E5B 1B8, Canada
Tel: +1 506 652 95783; +1 506 651 6020, E-Mail: gary.d.melvin@gmail.com; gary.melvin@dfo-mpo.gc.ca

SCRS VICE CHAIRMAN

Arrizabalaga, Haritz
Principal Investigator, SCRS Vice-Chairman, AZTI Marine Research Basque Research and Technology Alliance (BRTA), Herrera Kaia Portualde z/g, 20110 Pasaia, Gipuzkoa, Spain
Tel: +34 94 657 40 00; +34 667 174 477, Fax: +34 94 300 48 01, E-Mail: harri@azti.es

*****

ICCAT Secretariat

C/ Corazón de María 8 – 6th floor, 28002 Madrid – Spain
Tel: +34 91 416 56 00; Fax: +34 91 415 26 12; E-mail: info@iccat.int

Manel, Camille Jean Pierre
Neves dos Santos, Miguel
Ortiz, Mauricio
Kimoto, Ai

Appendix 3

List of Papers and Presentations

Number	Title	Authors
SCRS/2022/105	Efficacy of a bycatch estimation tool	Babcock E.A., Harford W.J., Gedamke T., Soto D., and Goodyear C. P.
SCRS/2022/106	ECOTEST, a proof of concept for evaluating ecological indicators in multispecies fisheries, with the Atlantic longline fishery case study	Huynh Q.C., Carruthers T., and Taylor N.G.
SCRS/P/2022/022	Analysis of recruitment deviates of tropical tuna stock assessments	Merino G., Urtizberea A., Fu D., Winker H., Cardinale M., Lauretta M.V., Murua H., Kitakado T., Arrizabalaga H., Scott R., Pilling G., Minte-Vera C., Xui H., Laborda A., Erauskin-Extraminiana M., and Santiago J.
SCRS/P/2022/023	Model Diagnostics in Integrated Stock Assessments	Kell L.T.
SCRS/P/2022/029	Updated summary on North Atlantic ALB MSE	Arrizabalaga H., Merino G.
SCRS/P/2022/034	Examples of the use of Structural Uncertainty in RFMO Shark Assessments	Rice J.
SCRS/P/2022/036	Updated summary on North Atlantic SWO MSE	Gillespie K.
SCRS/P/2022/038	BFT Management Strategy Evaluation (MSE)	Walter J.
SCRS/P/2022/039	Atlantic Multidecadal Oscillation: A Clarification and Projection at Depths	Moffat B.T.

Appendix 4

SCRS Document and Presentations Abstracts as provided by the authors

SCRS/2022/105 - The bycatch estimation tool developed by Babcock (2022) was subjected to simulation testing using the species distribution model and longline simulator (LLSIM) developed by Goodyear (2021). To evaluate the efficacy of the bycatch estimation tool, generalized representations of ICCAT CPC longline fisheries were created using LLSIM and were coupled with alternative representations of observer programs to produce simulated logbook and observer databases for a range of observer coverage levels and allocation methods. Using a semi-automated model selection process, linear predictors based on negative binomial and delta lognormal models were used to predict total annual bycatch of blue marlin from the simulated datasets. A stratified ratio estimator was also used for comparison. Across representations of observer programs, bycatch estimates were reasonably unbiased, with diminishing variation in bias estimates as observer coverage increased. The use of simulated data sets provides a demonstration of the utility of the bycatch estimation tool as well as evaluation of its reliability.

SCRS/2022/106 - There is a need for rigorous science to inform decision makers for Ecosystem Based Fisheries Management (EBFM). It is important to establish challenging and plausible scenarios for ecosystem dynamics and then test whether current and potential indicators can reflect stock status. Without the validation of indicators and the testing of relevant policy guidance to mitigate ecosystem impacts, there is a credibility gap between scientific practitioners of ecosystem science and decision makers that need to defend their actions in large multi-party negotiations. A multi-species framework that supports tactical decision making can make significant progress towards the essential goals of EBFM. We present a management strategy evaluation framework called "EcoTest". This is an extension to openMSE software, used for single-species modeling, that simulates multi-species fisheries dynamics. A range of features are possible in EcoTest, such as the ability to evaluate current indicators as well as design new indicators and identify the conditions under which indicators operate reliably. Here we demonstrate the use of EcoTest using the Atlantic longline fishery as a case study.

SCRS/P/2022/022 - Presentation provided a study examining a novel model diagnostic proposed for fisheries stock assessments. The diagnostic involves analyzing whether there is a trend in the estimated recruitment deviations over time from an assessment using linear models. As a case study, this diagnostic was applied to tropical tuna assessments fit using different growth, steepness, and natural mortality assumptions from the Atlantic, Indian, Western and Central Pacific, and East Pacific Oceans, particularly focusing on the assessment of the Indian Ocean yellowfin. The diagnostic was also applied within a simulation study of the Indian Ocean yellowfin and compared to other diagnostics such as the ASPM, Mohn's rho, MASE, runs tests, and autocorrelation and variance (Carvalho et al., 2021). Results for the case studies indicated that in general, when there was a statistically significant increasing trend in recruitment deviations, productivity parameters (unfished recruitment, MSY) were estimated to be lower. The opposite was true when statistically significant decreasing trends were found in recruitment deviates over time; they generally resulted in larger estimates of productivity parameters. This indicates that when recruitment deviates show an increasing trend, they may be compensating for the loss of biomass in periods of high catch even though they are estimating lower productivity. In these cases, recruitment deviates are not random process error but a factor that drives the responses of fish stocks to levels of high catch. Results from the simulation study corroborated the case studies and further indicated that trends in recruitment deviates can be caused by bias in the biological parameters used as fixed values in integrated assessment models. Comparisons between the recruitment deviations diagnostic and other diagnostics showed large consistency agreement with the ASPM diagnostic, however not with MASE, Mohns rho, or the runs tests. The authors suggest that the recruitment deviation diagnostic can provide statistical support for hypotheses and assumptions when selecting ensembles of models to develop fisheries management advice.

SCRS/P/2022/023 - Presentation provided examples of model validation diagnostics for integrated stock assessments. The methods presented are applicable to multiple modelling frameworks e.g. Bayesian biomass dynamic as well as integrated assessments models. Multiple ICCAT working groups, and RFMOs have identified the need for objective criteria to assess model plausibility and validation during stock assessment. Generic tools are being developed for model diagnostics that can be used to identify uncertainties, biases and misspecifications when developing a base case (e.g. the SS3diags package in R). Recent work has shown that model selection is an iterative process that is difficult to automate (Carvalho *et al*. 2021). Akaike information criterion (AIC), retrospective bias and model residuals are commonly used as model diagnostics. However, AIC-based selection is not suitable for comparing across models with different datasets and weighting, residual patterns can be removed by adding more parameters than justified by the data, and retrospective patterns removed by ignoring the data. Therefore, neither alone can be used for validation, which requires assessing whether it is plausible that a system identical to the model generated the data (Kell et al. 2021b). Validation, therefore, requires that the system is observable and measurable, and so observations should be used, rather than model-based quantities (e.g. SSB, recruitment, or F) unless these are well known. Model validation is best performed by hindcasting and cross validation where predictions of observations not used in fitting are compared to their known values. Hindcasting can be used to help select, weight and extend models i.e. to explore alternative formulations. Examples were provided from recent work on developing model ensembles (Kell et al. In Prep) and weighting schemes based on Retrospective Analysis (Mohn's ρ) and Prediction Skill (MASE) were compared.

SCRS/P/2022/029 - Presentation provided an update of the ALB MSE process that lead to the adoption of the first "full" management procedure (MP) for northern albacore (Rec 21-04), including a Harvest control rule, the way to determine stock status and a protocol for exceptional circumstances. The MSE process lasted at least 10 years, since the Commission requested to the SCRS to develop a limit reference point for this stock (Rec 11-04). The presentation showed a summarized chronology of key actions by Panel 2 (e.g. definition of management objectives in 2015, the adoption of performance statistics in 2016), the interactions between scientists and managers (e.g. communication of results about MP performance and advice to develop the exceptional circumstances protocol), and some technical characteristics of the MSE framework (e.g. Reference set of Operating models and characteristics of MPs tested). The primary document for the ALB MSE is the ALB MSE consolidated report (Merino et al., 2020).

SCRS/P/2022/034 - Presentation provided examples of the use of structural uncertainty grids in regional RFMO shark assessments. Two examples were provided based on recent assessments of blue shark (Prionace glauca) conducted for the Indian Ocean Tuna Commission (IOTC; Rice and Sharma 2015, Rice 2017, Rice 2021) and of oceanic whitetip (*Carcharhinus longimanus*) conducted for the Western and Central Pacific Fisheries Commission (WCPO; Rice and Harley 2012, Trembly-Boyer et al 2019, Rice et al 2021). Details of the presentation covered the use of multiple axes of uncertainty, the use of the results to identify key uncertainties, how inter-assessment research led to reduced uncertainty. The presentation discussed the use of weights on individual model runs and how to interpret results of a weighted group of runs. The goal of the presentation was to provide examples of the use of a structural uncertainty grid where over time the range of uncertainty is reduced, though analysis of factors (axes) that contribute to large variation. Limitations of the grid were discussed in relation to assessment uncertainty and compared that you would get from MCMC or similar methods. The use of a structural uncertainty grid in the evaluation of possible management actions was demonstrated. General recommendations were to identify major axes of uncertainty prior to analysis and to use an iterative approach to reduce uncertainty over time via structural uncertainty evaluation which can identify areas where research is most needed.

SCRS/P/2022/036 - No text provided by the author.

SCRS/P/2022/038 - No text provided by the author.

SCRS/P/2022/039 - Presentation overviewed a study that investigated the projected Atlantic Multidecadal Oscillation (AMO), an index traditionally defined using sea surface temperature, at depths where highly migratory species (HMS) tend to inhabit. This was possible using volumetric temperature data from the NOAA Geophysical Fluid Dynamics Laboratory Earth System Model version 4.1 (GFDL-ESM4.1), which includes simulated temperatures from 1900 to 2050 at 46 depth layers ranging from 0 to 1968 meters deep. When recreating the sea surface AMO index with the ESM4.1 dataset, it was discovered that the signal did not align with others previously published. It became evident that the AMO is expressed differently, in some cases producing conflicting warm-cool phases, depending on how the simulated or observed data was created. This is likely caused by model simulations only prescribing external forcing (CO_2, volcanoes, methane), leading to flaws of uncertainty because of internal variability that has limited predictability and errors in the forced response. Other data are based on in situ observations that are then gridded and flawed or uncertain mainly due to gaps in data coverage, although less so than model simulations. It emphasized the paramount importance of clearly stating the data used and all assumptions of the model when using environmental data for fisheries inference and understanding all accompanying uncertainty. Then using ESM4.1, it characterized the AMO signal at depths of interest to HMS distribution. Signals were analyzed for regime shifts of statistically significant warm and cool phases using a sequential t-test algorithm. Notable results include a compression of the signal with depth as well as a delayed regime shift of cool to warm phase with depth. Lastly, signals at depth from a peak warm and cool phase year were mapped for a spatial temperature comparison illustrating a greater temperature discrepancy at lower latitudes in the North Atlantic towards the sea surface.

Appendix 5

Workplan 2023

1. Evaluation of the products provided by the bycatch estimation methodology contract
2. Development of advice and/or guidelines on bycatch estimation
3. Initiation of the Study Group on CPUE standardization diagnostics
4. Initiation of the Study Group on Reference Points
5. Addressing within year issues referred to WGSAM by other Species Groups