REPORT OF THE 2020 SECOND INTERSESSIONAL MEETING OF THE ICCAT BLUEFIN TUNA MSE TECHNICAL GROUP

(Online, 28-30 September 2020)

"The results, conclusions and recommendations contained in this Report only reflect the view of the Bluefin Management Strategy Evaluation, MSE, Technical Group. Therefore, these should be considered preliminary until the SCRS adopts them at its annual Plenary meeting and the Commission revises 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 second intersessional Bluefin MSE Technical Group ("the Group") meeting was held online from 28 to 30 September 2020. Dr John Walter (USA), the Rapporteur for the western Atlantic stock, opened the meeting and served as Chair. On behalf of the Executive Secretary, the Assistant Executive Secretary welcomed the participants to the meeting. The Chair proceeded to review the Agenda which was adopted after minor changes (**Appendix 1**). Due to the time constraints, the Group focused on the main outputs from the meeting in this report and any technical aspects were expanded in Appendices. It was noted that this meeting does not have any authority to make final decisions, but rather its purpose is to prepare the material required for the bluefin tuna (BFT) intersessional meeting to be held in December 2020.

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

Sections	Rapporteur
Items 1, 10	A. Kimoto
Items 2 - 6	C. Fernandez, D. Butterworth, J. Walter
Items 7, 8	N.G. Taylor
Item 9	T. Carruthers, J. Walter

2. Consultant's update on work since July meeting

2.1 TSD finalization

Progress has continued on updating the Trial Specification Document (TSD). The current version (September 24, 2020) is TSD 20-3 (**Appendix 5**) and contains the following updates relative to earlier versions: in July, the distribution of catches among fleets, seasons, areas and projections was incorporated; in September, a performance metrics table was added.

2.2 Coding of robustness tests

2.2.1 General

The three main robustness tests, senescence, growth curve for eastern stock, "Brazilian" catches, had already been coded in July 2020 (with 4 Operating Models (OMs) each), but the "Brazilian" catches test needed to be redone. Robustness tests on time-varying mixing and persistent change in mixing have now also been coded (with 2 OMs each) and will be included in the Atlantic Bluefin Tuna Management Strategy Evaluation, ABTMSE, package when a new version (ver. 6.6.18) is released. Robustness tests on catchability increases or decreases and indices related non-linearly to abundance have yet to be coded but are forthcoming. Additional robustness tests are potentially still going to be considered (such as tests assuming a single stock or developing a single Candidate Management Procedure or CMP, and hence TAC) for the entire Atlantic), but due to the complexities of developing these scenarios, they can be considered only in a second round of the MSE, after an initial CMP is developed and presented to the Commission. This could be in a time frame of approximately 5 years from the present. For now, the immediate main focus for this first round will be on the robustness tests that have been identified as top priority (TSD table 9.3).

The consultant will provide the OM fitting reports for the robustness tests on senescence and eastern growth, whereas the other OM fitting reports will become available once they have been produced.

- The Group agreed to examine the OM fits for robustness trials prior to the December meeting and to send any comments to the Consultant.

2.2.2 Specification of "Brazilian catches" trial

The decision of the Group is to adopt the approach in SCRS/2020/152 that identified the catch in the West Atlantic between 20 N and 20 S that might be eastern origin, modifying the total tonnages allocated to the two stocks (see **Appendix 6**). In relation to the robustness test specifications, the Group agreed that the catch transfer (from the West to the East Atlantic) would be in terms of tonnage only, without modifying the age composition of the catches in either area.

2.3 Shiny App development

The Consultant demonstrated the most recent version of the Shiny App, illustrating how it can be used to compare among CMPs. In the future, automatic reports could be built directly from the Shiny App.

The Group found the Shiny App extremely useful. Clarifications about certain figures and other features of the results displayed were requested. Particularly, there were questions about how the probability quantiles shown were calculated, and about the meaning of the axes' values in the radar plots.

- It was agreed that captions for all figures and tables should be included, which should explain clearly what each represents and how it should be interpreted.
- The Group agreed that comments/suggestions for additions be sent to the Consultant during and soon after the meeting.

3. Progress of developers on CMP refinement

3.1 Report by each developer

The Group focused on what the CMP developers have updated in their studies.

SCRS/P/2020/063 presented some progress since the BFT intersessional meeting in May 2020 (Anon. 2020b). Different indices are weighted based on CV values reported in the BFT MSE Technical Group (BFT MSE TG) February meeting report (Anon. 2020a), and different target values for the East and West indices are now considered. This leads to improved performance, but considerable development is still needed.

SCRS/2020/147 showed that the fixed proportion (FXP) CMP resulted in very low Br30 (biomass relative to dynamic B_{MSY} after projection year 30) values for some OMs. Ways of improving this behaviour were explored, essentially by modifying the FXP CMP such that the TAC could be reduced by larger amounts and the fishing proportion be reduced when the stock biomass indicator falls below some threshold. The main problems found were for Recruitment level 3 scenarios (regime shift). The western stock proved to be more problematic than the eastern one, particularly for OMs with smaller scale SSB for the western stock.

SCRS/2020/150 has not made many changes since the May meeting (Anon. 2020b), but the work is ongoing.

SCRS/2020/145 used five OMs selected from medoids of clusters, as explained in earlier meetings. Both empirical and model-based (delay-difference models) CMPs were considered. The aim now is to refine these initial CMPs, so as to reduce the variance of Br30 across OMs. A possibility to achieve this may be to weight the five TACs in an adaptive manner, rather than use fixed weights, and focusing more on the higher or the lower TACs depending on observed stock size signals.

SCRS/2020/144 found that the most difficult situation was found to be the changing recruitment regime. To improve CMP performance in these cases, the possibility of excluding indicators with more autocorrelation in favour of others with less autocorrelation will be investigated.

SCRS/2020/127 considers a constant harvest rate in East and West areas, as well as a version with constant harvest rate in the East and a rebuilding approach in the West.

SCRS/2020/151 showed some changes since the previous presentation, with a threshold for recruitment in the West being added. No results are yet available from the latest package, but some results are provided in the document.

Group discussion focused on:

- Frequency of TAC updates: for interim CMP development, the Group agreed to use TAC updates of 2 years as a default to allow for comparison between alternative CMP; CMP developers will check to determine how the TAC interval affects the performance of their CMP. The frequency of TAC updates in practice, i.e. for management application, will have to be reviewed by the Commission for final approval.
- The Group agreed that some flexibility for maximum TAC change is needed, because some CMPs may use these constraints in different ways (e.g. depending on stock size or trend), aiming to improve CMP performance.
- Some interaction with Panel 2 is also needed on these items (see Section 8).

3.2 Comparison of results using Shiny App

The Consultant showed the results from all CMPs in the Shiny App (except SCRS/2020/151, whose results had not yet been included). These "Zeh plots" of AvC30 (average catch over years 1-30 by area) and Br30 (by stock) showed the expected trade-off between these two metrics, both in the East and West, which motivated the usefulness of development tuning (i.e. a choice of a point, or a small set of different points (maybe 3 different ones), in that trade-off space) to help compare the different CMPs.

3.3 Discussion

The Group noted that there also seemed to be a trade-off between performance for the East and West.

Therefore, three main issues to address are: development tuning, East-West trade-off, and multiple CMPs and how to cull them. At this point, it is still expected that the Catch versus Biomass target will be the main trade-off to be resolved.

In relation to the East-West trade-off, the Group proposed constructing radar plots showing Catch and Biomass in the eastern and western stocks, i.e. four measures in the same radar plot. One would then examine both the area and symmetry of radar envelopes. It was noted that catch variability is another important consideration, in addition to average catch and stock biomass, that might need to be included.

- The Consultant will include a way to compare East and West in the Shiny App, and to show the trade-off between them.
- To facilitate displays, it was agreed that CMPs should be named using four characters: the first two to identify the team, and last 2 to identify the CMP version (**Table 1**). CMP developers should use those in all their results displays.
- A table of the various CMPs presented, indicating some basic features (e.g. empirical or model-based) will be prepared and included in this report (**Table 1**).

4. Towards grid finalization

4.1 Consideration of further "does it matter" contributions

The question posed is whether it is possible to remove some uncertainty axes from the interim OM grid because they matter very little for CMP performance.

Butterworth and Rademeyer, 2020 was reviewed; this work had already been presented in the May meeting (Anon. 2020b). For five different CMPs, it examines how much the Br30 value changes across each of the uncertainty axes while integrating over the other ones. The results are displayed in **Figure 1** of the document and indicate that the "Recruitment" axis is the most influential, the "Scale" and "LHw" (likelihood weight for the length composition data) axes matter less, and the "Mixing" and "Biology" axes appear to matter very little. The Group was reminded that the "LHw" axis was to some extent linked with the level of mixing of the eastern stock in the West area. The Group concluded that the "Mixing" and "Biology" axes could be potential candidates for removal from the OM grid.

SCRS/2020/126 was presented to the Group because it was discussed only partially during the second BFT intersessional meeting in July 2020 (Anon. 2020c).

SCRS/2020/117 applied a GLM approach to examine if the OM grid could be reduced. Nine different CMPs were considered and, for each of them, the Br30 and Av30 values analyzed with GLMs. The GLM structure was: Response ~ mu + Stock + Recruitment + Scale + Biology + Mixing + LHw + 2-way interactions + 3-way interactions + error. Stepwise forward selection was used to select the factors for inclusion. Overall, it was found that what is influential varies depending on the performance metric examined, and also differs from CMP to CMP. The analysis so far suggests that mixing may be the least influential axis. Interactions exist for many CMPs and are influential, which suggests limiting changes to the OM grid. The method could also be used to rank CMPs and to see how fixing an axis to a particular level affects CMP ranking.

The main concern raised by the Group was the large number of free (estimable) parameters in the GLM relative to the number of "observations", which could lead to overparameterization.

The authors concluded with the following summary of what needs to occur in order for this GLM approach to inform on what is influential:

- Decide on interim performance metrics, solely for the purposes of compilation of results, for safety and stability (AAVC (average variation in catch between TAC changes over 30 year time period) and LD (lowest depletion over 30 years) by Oct 3rd).
- The Group agreed that the results would be compiled and made available for the CMP results received so far.
- Apply GLM approach to data (December 15):
    - Address potential over parameterization. Use AICc or increase AIC threshold (2x)
    - Plot effect sizes with s.d. (degree of significance)
    - Identify the maximum difference among factors. Provide ranked % deviance explained by factor (or change in AIC)
    - Work with the residuals, identify the "important factor components" thus avoiding the issue of degrees of freedom.

4.2 Interim grid first revision

The Group discussed if some axes could be removed from the interim grid and has identified further approaches for exploring that issue. At this time, the Group did not make any changes to axes for revision. The Group also did not make any changes to levels within each axis.

5. Development tuning

SCRS/2020/146 considers a particular CMP (FXP) and tunes it (Br30 as close to 1 as possible for eastern and western stocks) in three different ways: using only OM1, using the median across all 96 OMs, and using only the 5 OMs selected in SCRS/2020/145. The resulting percentiles of Br30 across the 96 OMs, for the tuned CMPs, were very similar when the tuning was based on OM1 or on the median across all 96 OMs, but differed for the western stock (in terms of median, but hardly of variance) when the tuning was based on the 5 OMs. The authors of the analysis concluded that it was appropriate and more practical to conduct development tuning based on a single OM (although not necessarily OM1 - this should be discussed, with a particular focus on the western stock, which is seen to be the most problematic in all the analyses conducted thus far).

The Group discussed whether using a single (and common) OM for development tuning for all the CMPs may not be appropriate. Whereas for a particular CMP, tuning based on OM1 may lead to similar Br30 results to tuning based on the median across the 96 OMs in the grid, this may not be the case for other CMPs. Hence, CMPs tuned based on a single OM may not be as comparable as intended under this process of development tuning.

The Group investigated this issue further during the meeting, by considering a range of CMPs tuned to achieve Br30=1 under various choices of single OM and examining the corresponding (Br30, AvC30) values under other single OMs or as the median over the entire grid of 96 OMs. The results showed that conclusions were not straightforward, and that selecting a single OM, or a small set of OMs, as the basis of the development tuning would be very difficult.

The following points were raised in the Group discussion:

1) Finalization tuning will (likely) be for a range from conservative to aggressive across a final plausibility-weighted grid
2) Development tuning is a means to an end (not an end in itself), and intends:
   a) to provide insight into what features of CMPs give better performance
   b) to aid in the CMP culling process
   c) that CMP developers, while undertaking the development tuning agreed by the Group, are also free to try others also and report back why they consider them better
3) It was unlikely that the Group would be able to agree on a single OM for development tuning, and hence there was a need to find another way
4) The best option for development tuning would be to use the median over the interim grid (as nearest to what will be used for finalization tuning), but technically that would be difficult and time consuming for tuning in two dimensions (i.e. for both eastern and western stocks)
5) From the analyses that had been conducted to date, the western stock proves much more "difficult" for achieving conservation/rebuilding compared to the eastern stock.

The following approach for development tuning was agreed by the Group:

a) Tune only to the western stock, using 3 alternative target Br30 levels for it (1, 1.25, and 1.5 to a guideline tolerance of 0.01 – at *this* time). Leave the eastern stock to each developer group. This tuning will be completed by the date specified in the workplan (see section 9) using the determinist OMs, with perfect implementation (according to specification listed in section 1.1 of the CMP developers guide package version 6.6.18).
b) The tuning for the western stock will be based on the median Br30 across the entire OM interim grid (96 OMs) – that is more practical as it is one-dimensional (i.e. only for the western stock).

SCRS/2020/148 and SCRS/2020/149 were discussed at an informal meeting with CMP developers on September 17, 2020.

6. Plausibility

The BFT MSE TG February meeting (Anon. 2020a) considered options for plausibility weighting of OMs (see Section 7 of that report). Essentially, IWC-like and CCSBT-like approaches were discussed in that meeting, noting that the IWC-like approach would classify OMs into high/medium/low plausibility categories, whereas the CCSBT-like approach would assign numerical weights (based either on statistical or Delphi-type approaches linked to individual expert judgement).

The BFT MSE TG February meeting (Anon. 2020a) suggested that a hybrid approach involving selected elements of IWC-like and CCSBT-like approaches could be used to provide a flexible OM plausibility weighting for bluefin tuna. For example, qualitative high/medium/low plausibility categories (i.e. IWC-like) could be assigned to uncertainty axes that are more "discrete" and/or where experts hold strongly different views. Quantitative OM weights (i.e. CCSBT-like) could be applied to uncertainty axes that are "continuous" and/or lack strongly opposing views among experts.

In the current meeting, the Group agreed to go over each of the uncertainty axes (the 5 axes in the interim grid + the 3 main robustness tests) and to consider, for each of them, whether the CCSBT-like approach seemed possible or whether the IWC-like approach should be applied.

After considering the eight uncertainty axes above, the Group agreed that use of an IWC-like approach was not needed, and that the CCSBT-like approach could be applied with some customization for the Group´s application.

A basis for plausibility weighting over a grid of OMs was proposed to the Group. Example of plausibility weights is shown in **Table 2**.

1) The grid of OMs is specified by:
   uncertainty axes a (1, …. N) and the levels l (1, ….. $n(l)$) within each axis

2) Plausibility weights $w_{a,l(a)}$ are assigned to the levels l along each axis a by some method (e.g. likelihood based; expert judgement finalized using a Delphi approach), where these weights are normalized to sum to 1, i.e.:
   Sum over l of $w_{a,l(a)} = 1$ for each axis a

3) Each operating model (OM) is characterized by a vector of l values:
   $\{l(a=1), l(a=2), \ldots l(a=N)\}$ which can be written more simply as $\{l_1, l_2, \ldots l_N\}$

   so that each OM can be designated as OM($l_1, l_2, \ldots l_N$).

4) The plausibility weight W given to that OM when integrating over the grid is given by the product of the corresponding levels weight over all the axes, i.e.:
   W[OM($l_1, l_2, \ldots l_N$)] = Product over a of [la] = Product over a of [$w_{a,l(a)}$]

   For example, given an example W matrix (**Table 2**),
   OM[(a=1, l=2); (a=2, l=3); (a=3, l=2) ; (a=4, l=2) ; (a=5, l=1)] would have a weight W of 0.01125 (0.3* 0.25 * 0.5 * 0.5 * 0.6).

5) Note that it follows from this method of constructing these weights that the sum of the weights W over all the OMs in the grid will be 1, i.e.:
   Sum over a from 1 to N and of $l(a)$ over 1 to $n(l)$ of [$w_{a,l(a)}$] = 1

   i.e. the assumption is made that the contribution to the overall weighting from each uncertainty axis can be taken to be independent of that from the others.

The Group agreed to a conceptual plan to move forward with the weighting exercise that would score each level within the axes of uncertainty. After creating the poll, Step 1 in the proposal for weighting is to conduct a silent poll where experts would score each factor level for plausibility with percentages summing to 100 as outlined in **Table 2**. Guidance (or rules) for scoring is presented in **Table 3**. For any axes with wide distributions in scoring, the process could be iterated following a Delphi-type approach with the protocols for this second iteration to be discussed at the December 2020 online meeting of the BFT Species Group (BFTSG).

A version of the candidate poll that will be reviewed by the BFTSG will have to be prepared by mid-November so that it can be considered by the BFTSG. Between December and the March intersessional, the poll would be conducted for review in early 2021 by the BFTSG; the reconciliation of divergent scores would occur at the March 2021 intersessional (exact timing subject to Commission approval).

A final step, to be conducted after the initial weighting defined in the first round, will be to determine if there are OMs that (a) exhibit particularly implausible behavior and (b) are not sufficiently weighted in the first round. It is possible that, if the weighting in Step 1 does not sufficiently address issues, then further reweighting could be considered. However, to do so will require sufficient justification that the particular OMs do not reflect a biologically plausible situation, e.g. too great a proportion of fish in a particular stratum.

The Group further agreed that although such strategic decisions could be taken at this meeting, further processes needed would be implemented afterwards and should be reflected in the workplan.

7. Culling of CMPs

Due to time constraints this discussion did not occur.

8. Interactions with the Commission and other stakeholders

The Group discussed that a dialogue with Panel 2 March meeting could be useful in 2021, focused primarily on an overview of the MSE progress and challenges to date. This could be undertaken by the SCRS Chair.

The BFT MSE TG noted that a separate day prior to the November 2021 Commission meeting would be an important opportunity for dialogue with Panel 2 / November Commission. Dialogue with Panel 2 in 2021 is absolutely critical to beginning the process that will occur in 2022 for possible MP adoption at the 2022 Commission meeting.

9. Workplan and Roadmap

Workplan

a) The ICCAT Secretariat will provide catch data files for an updated conditioning of the 'Brazilian catch' robustness OMs. The Consultant will share all robustness OM conditioning reports with BFTSG members who will take this opportunity to comment on these results.
b) The Consultant will add computer code to implement robustness OM features that specify changing index catchability and non-linearity in indices.
c) The BFTSG will make suggestions regarding improvements and additions to the Shiny App. The Consultant will record suggestions and make updates to the Shiny App where possible. The Consultant will develop tools for appending Shiny App MSE results, and include these in an updated ABTMSE package including documentation of these features.
d) All CMP developers will implement development tuning to achieve multiple biomass targets for the western stock (see section 5). At this time developers should implement a 2-year TAC change (by November 15).
e) CMP developers are encouraged to continue electronic dialogue or to hold informal webinars, as desired, to facilitate development, share coding tips and clarify tasks, as needed.
f) An investigation of the importance of OM reference grid factors by Hanke *et al*. (SCRS/2020/117) will be broadened to include results from all CMP developers.
g) An initial weight poll will be developed (November 15) for the initial consideration of the BFTSG in December.
h) A BFTSG intersessional meeting will be held on December 1, 2, and 3.

Suggested TOR for the December BFTSG meeting

i. Further discussion on grid finalization and possible reconditioning (if so, when and how)
ii. Update from CMP developers on progress and summarization of development tuning
iii. BFT MSE TG presents initial proposal for plausibility weighting to BFTSG
iv. Agree to protocols and guidelines for plausibility weighting
v. Agree on which axes are to be scored and how scoring would be conducted, particularly details of the poll
vi. Discussion of GBYP proposal to the EU
vii. Time permitting, discussion of other GBYP matters.

MSE Roadmap

The BFT MSE TG considered the roadmap and identified a number of priority items from the MSE roadmap. The BFT MSE TG did not fully revise the roadmap as this should be a task of the BFTSG. The priority items noted were as follows:

- Code review and peer review of MSE (terms of reference to be determined in December 2020)
- Aim to provide CMP to Commission for adoption in 2022
- Timeline for finalizing exceptional circumstances should occur after adoption of an MP (2023)
- Interactions with Panel 2 (TBD in December, but necessary in 2021)

11. Adoption of the report and closure

The report was adopted during the meeting. The Chair of the Group and SCRS Chair thanked all the participants for their efforts. The meeting was adjourned.

References

Anon 2020a. Report of the 2020 Intersessional Meeting of the ICCAT Bluefin Tuna MSE Technical Group *(Madrid, Spain, 24-28 February 2020*). ICCAT Collect. Vol. Sci. Pap. 77(2): 1-74.

Anon. 2020b. Report of the 2020 Intersessional Meeting of the ICCAT Bluefin Tuna Species Group (*Online, 14-22 May 2020*). ICCAT Collect. Vol. Sci. Pap. 77(2): 96-214.

Anon. 2020c. Report of the Second 2020 ICCAT Intersessional Meeting of the Bluefin Tuna Species Group (Online, 20-28 July 2020).

Butterworth and Rademeyer. 2020. What do current results using the package indicate regarding which uncertainty axes "matter" regarding CMP performance, and what are the next steps needed in the ABFT MSE process. ICCAT Collect. Vol. Sci. Pap. 77(2): 434-440.

Table 1. A list of CMPs currently under development (or in use for comparative purposes) with some of their basic features.

CMP Names	*SCRS documents*	*Authors*	*Type of MP*
Zero	-	-	zero catch
H_1	SCRS/2020/144	Hanke	empirical
H_2	SCRS/2020/144	Hanke	empirical
EA_1	SCRS/P/2020/063	Andonegi, Fernandez, Arrizabalaga, Rouyer, and Gordoa	empirical
EA_2	SCRS/P/2020/063	Andonegi, Fernandez, Arrizabalaga, Rouyer, and Gordoa	empirical
BR_1	SCRS/2020/147	Butterworth and Rademeyer	empirical
BR_2	SCRS/2020/147	Butterworth and Rademeyer	empirical
BR_3	SCRS/2020/147	Butterworth and Rademeyer	empirical
BR_4	SCRS/2020/147	Butterworth and Rademeyer	empirical
LW_1	SCRS/2020/127	Lauretta and Walter	empirical
LW_2	SCRS/2020/127	Lauretta and Walter	empirical
C50	-	-	50% of current TAC
C100	-	-	current TAC
C_75	SCRS/2020/150	Carruthers	empirical
C_100	SCRS/2020/150	Carruthers	empirical
C_125	SCRS/2020/150	Carruthers	empirical
CJR_1	SCRS/2020/145	Cox, Johnson, and Rossi	model based
TN_1	SCRS/2020/151	Tsukahara and Nakatsuka	empirical

Table 2. Plausibility weighting/polling proposal. Example score for each level (within each axis) according to the plausibility criteria, then multiply scores to achieve the overall weighting for the OM. Note that the values presented here are simply placeholders.

			levels			
Factor	***Axis***	***Number of levels***	***1***	***2***	***3***	***4***
interim grid	(a=1) Recruitment	3	0.3	0.3	0.4	
interim grid	(a=2) Scale	4 (--, -+, +-, ++)	0.25	0.25	0.25	0.25
interim grid	(a=3) Spawn. Frac. / M	2 (A, B)	0.5	0.5		
interim grid	(a=4) Mixing	2 (I, II)	0.5	0.5		
interim grid	(a=5) Length Comp Wt	2 (L, H)	0.6	0.4		
robustness	Brazil catches	2	0.5	0.5		
robustness	Senescence	2	0.5	0.5		
robustness	East growth curve = West	2	0.5	0.5		

Table 3. Initial proposal from BFT MSE TG for plausibility weighting considerations and polling.

Rules for plausibility weighting	***Decisions***
1. Participants	Strawman- Participants at the December BFTSG online meeting. The BFTSG would ratify the decisions.
2. Blind/anonymous scoring	Individual respondents cannot be influenced by poll results of other poll respondents.
3. Axes to be scored	a. Score all reference and robustness axes. b. Equal weighting for recruitment and spawning fraction/*M* axes. i.e. omit these from polling.
4. Guidance on plausibility weighting	• Based on whether factor achieves its biological objective; this can consider performance in conditioning (e.g. if the factor level moves fish in areas not considered to be plausible) and participants' *a priori* biological expertise. • The range of the axis levels should also be considered, e.g. does it span range of plausibility, which might apply to the biology factors rather than to each individual level of this axis. • Any asymmetry in the probability range of axis levels should be considered. • Weights across levels sum to 1.
5. Plausibility should be based on CMP performance	No, not on CMP performance.
6. 'weight' the levels within an axis, not the axis	Yes, levels within the axis.
7. Should different axes be weighted?	No need to weight each individual axis – already subsumed in factor weighting.
8. Process for conducting poll and iteration of plausibility weighting	*To be determined at next BFTSG meeting.*

Appendix 1

Agenda

1. Opening, adoption of agenda and meeting arrangements
2. Consultant's update on work since July meeting
   - 2.1 TSD finalization
   - 2.2 Coding of robustness tests
     - 2.2.1 General
     - 2.2.2 Specification of "Brazilian catches" trial
   - 2.3 Shiny app development

3. Progress of developers on CMP refinement
   - 3.1 Report by each developer
   - 3.2 Comparison of results using Shiny app
   - 3.3 Discussion

4. Towards grid finalization
   - 4.1 Consideration of further "does it matter" contributions
   - 4.2 Interim grid first revision

5. Development tuning
   - 5.1 Specification of a set of targets
   - 5.2 Discussion related to next steps by CMP developers

6. Plausibility
   - 6.1 Discussion of process suggested at February meeting
   - 6.2 First steps in implementation of that process

7. Culling of CMPs
   - 7.1 Initial discussion of basis to undertake this

8. Interactions with the Commission and other stakeholders

9. Workplan

10. Adoption of the report and closure

Appendix 2

List of Participants

CONTRACTING PARTIES

ALGERIA

Ferhani, Khadra
Centre National de Recherche et de Développement de la Pêche et de l'Aquaculture (CNRDPA), 11 Boulevard Colonel Amirouche, BP 67, 42415 Tipaza Bou Ismail
Tel: +213 550 735 537, Fax: +213 24 32 64 10, E-Mail: ferhani_khadra@yahoo.fr; dpmo@mpeche.gov.dz

Kouadri-Krim, Assia
Chef de Bureau, Ministère de la Pêche et des Productions Halieutiques, Direction du développement de la pêche, Rue des 04 Canons, 16000
Tel: +213 558 642 692, Fax: +213 21 43 31 97, E-Mail: dpmo@mpeche.gov.dz; assiakrim63@gmail.com

CANADA

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

Duprey, Nicholas
Senior Science Advisor, Fisheries and Oceans Canada - Fish Population Science, Government of Canada, 200-401 Burrard Street, Vancouver, BC V5V 4V1
Tel: +1 604 499 0469; +1 250 816 9709, E-Mail: nicholas.duprey@dfo-mpo.gc.ca

Hanke, Alexander
Scientist, St. Andrews Biological Station/ Biological Station, Fisheries and Oceans Canada, 125 Marine Science Drive, St. Andrews, New Brunswick E5B 2T0
Tel: +1 506 529 5912, Fax: +1 506 529 5862, E-Mail: alex.hanke@dfo-mpo.gc.ca

Maguire, Jean-Jacques
1450 Godefroy, Québec G1T 2E4
Tel: +1 418 527 7293, E-Mail: jeanjacquesmaguire@gmail.com

EUROPEAN UNION

Álvarez Berastegui, Diego
Instituto Español de Oceanografía, Centro Oceanográfico de Baleares, Muelle de Poniente s/n, 07121 Palma de Mallorca, España
Tel: +34 971 133 720; +34 626 752 436, E-Mail: diego.alvarez@ieo.es

Andonegi Odriozola, Eider
AZTI, Txatxarramendi ugartea z/g, 48395 Sukarrieta, Bizkaia, España
Tel: +34 661 630 221, E-Mail: eandonegi@azti.es

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

Di Natale, Antonio
Aquastudio Research Institute, Via Trapani 6, 98121 Messina, Italy
Tel: +39 336333366, E-Mail: adinatale@acquariodigenova.it

Fernández Llana, Carmen
Instituto Español de Oceanografía, Avda. Príncipe de Asturias, 70 bis, 33212 Gijón, España
Tel: +34 985 309 780, Fax: +34 985 326 277, E-Mail: carmen.fernandez@ieo.es

Gordoa, Ana
Centro de Estudios Avanzados de Blanes (CEAB - CSIC), Acc. Cala St. Francesc, 14, 17300 Blanes, Girona, España
Tel: +34 972 336101, E-Mail: gordoa@ceab.csic.es

Guerin, Benoît
1407 Chemin des Maures, 83400 Hyeres, France
Tel: +33 632 02 68 15, E-Mail: bgseaconsulting@gmail.com

Molina Schmid, Teresa
Subdirectora General Adjunta, Subdirección General de Acuerdos y Organizaciones Regionales de Pesca, Dirección General de Recursos Pesqueros, Ministerio de Agricultura, Pesca y Alimentación, Secretaría General de Pesca, C/ Velázquez, 144 2ª Planta, 28006 Madrid, España
Tel: +34 91 347 60 47; +34 656 333 130, Fax: +34 91 347 60 42, E-Mail: tmolina@mapa.es

Rouyer, Tristan
Ifremer - Dept Recherche Halieutique, B.P. 171 - Bd. Jean Monnet, 34200 Sète, Languedoc Rousillon, France
Tel: +33 782 995 237, E-Mail: tristan.rouyer@ifremer.fr

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

Tugores Ferra, Maria Pilar
ICTS SOCIB - Sistema d'observació y predicció costaner de les Illes Balears, Parc Bit, Naorte, Bloc A 2ºp. pta. 3, E-07121 Palma de Mallorca, España
Tel: +34 971 439 998, E-Mail: pilar.tugores@gmail.com

JAPAN

Butterworth, Douglas S.
Emeritus Professor, Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch, 7701 Cape Town, South Africa
Tel: +27 21 650 2343, E-Mail: doug.butterworth@uct.ac.za

Nakatsuka, Shuya
Director, Pacific Bluefin Tuna Resources Group, National Research Institute of Far Seas Fisheries, Japan Fisheries Research and Education Agency, 5-7-1 Orido, Shizuoka Shimizu 424-8633
Tel: +81 543 36 6035, Fax: +81 543 36 6035, E-Mail: snakatsuka@affrc.go.jp

Rademeyer, Rebecca
Marine Resource Assessment and Management Group, Department of Mathematics and Applied Mathematic - University of Cape Town, Private Bag, 7700 Rondebosch, South Africa
Tel: +651 300 442, E-Mail: rebecca.rademeyer@gmail.com

Tsukahara, Yohei
ScientistHighly Migratory Resources Division, Fisheries Stock Assessment Center, Fisheries Resources Institute, Highly Migratory Resources Division, Fisheries Stock Assessment Center, Fisheries Resources Institute, Japan Fisheries Research and Education Agency, 5-7-1 Orido, Shizuoka Shimizu-ku 424-8633
Tel: +81 54 336 6000, Fax: +81 54 335 9642, E-Mail: tsukahara_y@affrc.go.jp

KOREA REP.

Lee, Mi Kyung
National Institute of Fisheries Science, Distant Water Fisheries Resources Research Division, 216 Gijanghaean-ro, Gijang-eup, Gijang-gun, 46083 Busan
Tel: +82 51 720 2332, Fax: +82 51 720 2337, E-Mail: ccmklee@korea.kr; cc.mklee@gmail.com

NORWAY

Mjorlund, Rune
Directorate of Fisheries, Strandgaten 229, 5804 Bergen
Tel: +47 952 59 448, E-Mail: rune.mjorlund@fiskeridir.no

Nottestad, Leif
Principal Scientist, Institute of Marine Research Research Group on Pelagic Fish, P.O. Box 1870 Nordnesgaten, 33, 5817 Bergen, Hordaland county
Tel: +47 5 99 22 70 25, Fax: +47 55 23 86 87, E-Mail: leif.nottestad@hi.no

TUNISIA
Zarrad, Rafik
Institut National des Sciences et Technologies de la Mer (INSTM), BP 138 Ezzahra, Mahdia 5199
Tel: +216 73 688 604; +216 972 92111, Fax: +216 73 688 602, E-Mail: rafik.zarrad@instm.rnrt.tn; rafik.zarrad@gmail.com

UNITED STATES
Brown, Craig A.
Chief, Highly Migratory Species Branch, Sustainable Fisheries Division, NOAA Fisheries Southeast Fisheries Science Center, 75 Virginia Beach Drive, Miami, Florida 33149
Tel: +1 305 586 6589, Fax: +1 305 361 4562, E-Mail: craig.brown@noaa.gov

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

Peterson, Cassidy
NOAA Fisheries, 75 Virginia Beach Drive, Miami, FL 33149
E-Mail: cassidy.peterson@noaa.gov

Schalit, David
176 Mulberry Street - 4th floor, New York 10013
Tel: +1 917 573 7922, E-Mail: dschalit@gmail.com

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

OBSERVERS FROM NON-GOVERNMENTAL ORGANIZATIONS

PEW CHARITABLE TRUSTS - PEW
Johnson, Samuel D.N.
School of Resource and Environmental Management, 8888 University Drive, Burnaby, BC V5A1S6, Canada
Tel: +1 604 365 7133, E-Mail: samuelj@sfu.ca

THE OCEAN FOUNDATION
Miller, Shana
The Ocean Foundation, 1320 19th St., NW, 5th Floor, Washington, DC 20036, United States
Tel: +1 631 671 1530, E-Mail: smiller@oceanfdn.org

Pipernos, Sara
The Ocean Foundation, 1320 19th St. NW, Washington DC 20036, United States
Tel: +1 860 992 6194, E-Mail: spipernos@oceanfdn.org

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, E-Mail: gary.d.melvin@gmail.com; gary.melvin@dfo-mpo.gc.ca

INVITED EXPERT
Parma, Ana
Principal Researcher, Centro para el Estudio de Sistemas Marinos, CONICET (National Scientific and Technical Research Council), Blvd. Brown 2915, U 9120 ACF Puerto Madryn, Chubut, Argentina
Tel: +54 (280) 488 3184 (int. 1229), Fax: +54 (280) 488 3543, E-Mail: parma@cenpat-conicet.gob.ar; anaparma@gmail.com

*****

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

Neves dos Santos, Miguel
Ortiz, Mauricio
Kimoto, Ai
Taylor, Nathan
Alemany, Francisco
Pagá, Alfonso

Appendix 3

List of Papers and Presentations

Number	Title	Authors
SCRS/2020/117	A GLM approach for determining the influence of operating model features on management procedure performance	Hanke A.R., Arrizabalaga H., Andonegi E., and Duprey N.
SCRS/2020/126	PART 1: Investigation of the impact of spatial distribution of mean available biomass on Operating Model projection outcomes	Carruthers T., Butterworth D., and Rademeyer R.
SCRS/2020/127	Atlantic bluefin tuna constant harvest rate and index-based Candidate Management Procedures	Lauretta M., and Walter J.
SCRS/2020/144	A description of 4 candidate management procedures for bluefin tuna	Hanke A.R.
SCRS/2020/145	Two classes of multi-model candidate management procedures for Atlantic bluefin tuna	Cox S.P. Johnson S.D.N., and Rossi S.P.
SCRS/2020/146	Implications of alternative choices of OMs for development tuning targets	Butterworth D.S., and Rademeyer R.A.
SCRS/2020/147	Refining the FXP (fixed proportion) CMP	Butterworth D.S., and Rademeyer R.A.
SCRS/2020/148	Selecting an OM for ABFT MP development tuning	Butterworth D.S., and Rademeyer R.A.
SCRS/2020/149	Demonstration of CMP development tuning for Atlantic bluefin tuna	Carruthers T., Butterworth D.S., and Rademeyer R.A.
SCRS/2020/150	Designing and testing a multi-stock spatial management procedure for Atlantic bluefin tuna	Carruthers T.
SCRS/2020/151	Tentative concepts for CPUE-based simple candidate management procedure for MSE of Atlantic Bluefin Tuna	Tsukahara Y., and Nakatsuka S.
SCRS/2020/152	Bluefin tuna catch review of the “Brazilian episode” for the MSE OM robustness test	Kimoto A., and Ortiz M.

SCRS/P/2020/063	Latest progress on refining and tuning the index-based EU-cMP	Andonegi E., Fernandez C., Arrizabalaga H., Rouyer T., and Gordoa A.

Appendix 4

SCRS Document and Presentations Abstracts as provided by the authors

SCRS/2020/117 - GLMs were fit to the performance metrics generated by management procedures applied to models of the Atlantic Bluefin tuna fishery in a closed loop simulation. The models identified the features of the population model that accounted for the most variability in the average catch and biomass ratio over 30 years of simulated management. The variability in the performance metrics of the alternative management procedures tested was attributed to a differing set of population model features, i.e. the most influential axes of uncertainty in the population model were management procedure dependent.

SCRS/2020/126 - No text provided by the author.

SCRS/2020/127 - We evaluated two candidate management procedures for Atlantic bluefin tuna using the ABT_MSE package in R. The first procedure applied a constant harvest rate strategy for both the east and west stocks. The second procedure evaluated the ability to achieve SSB of the West stock at or above current estimates (measured by stock-of-origin indices). Observations from indices of abundance were assumed proportional to spawning biomass and juvenile abundance for each stock and area, respectively, with no observation error (observation model = Perfect_Obs). Each procedure was compared against zero-catch scenarios for comparison of trade-offs among strategies. The constant procedures were tuned to the median Br30=1 across five selected OMs that characterized the general clusters in the larger OM grid.

SCRS/2020/144 - Four candidate management procedures for Western and Eastern Bluefin tuna stocks were are described.

SCRS/2020/145 - Two classes of multi-model candidate management procedures for Atlantic bluefin tuna were developed and tested. Procedures were based on spawning biomass estimation methods scaled to five operating models selected via cluster analysis from the reference OM grid. For the empirical class, OM catchability and a constant stock mixing distribution were used to estimate area biomass from the larval indices. For model-based MPs, five delay difference assessment models were scaled to each of the five operating models, matching stock recruit steepness and biomass for the recent historical period from 1965 - 2016. At each time step, estimates of current (empirical) or projected (model-based) biomass were generated from approved management indices and used in harvest control rules to generate area-specific TACs, and the five TACs were averaged to produce harvest advice for the East and West area. Multi-model CMPs scaled to the five OMs performed well across the full range of 96 operating models with minimal tuning; however, some CMPs were overly conservative and would benefit from refinement to reduce overfishing when stock biomass is overestimated.

SCRS/2020/146 - Questions have been raised about how the choice of OM (or group of OMs) for development tuning impacts the distribution of results across the full set of 96 OMs of the interim grid. The results for three different choices are examined for the FXP CMP for 100 tuning (Br30 = 1 for both Eastern and Western stocks for deterministic projections). Although the Br30 distributions for the whole interim grid shift up or down to different extents, relative to each of their medians, these distributions are unchanged for all practical purposes. This argues for using a single OM as the basis for development tuning, in the interests of simplicity of implementation.

SCRS/2020/147 - Various refinements to the FXP (fixed proportion of abundance as indicated by an aggregate of abundance indices for the area concerned) CMP are considered in an initial attempt to improve its performance, particularly with regard to resource conservation. For the "100 tuning" variant (Br30 = 1 for both Eastern and Western stocks for deterministic projections of OM1), there are OMs in the interim grid for which either or both of the two stocks can be rendered (near) extinct. The OMs that lead to the greatest difficulties for these CMPs in that conservation context are those with a low abundance scale for the Western stock, and further those incorporating the R3 scenario where a regime shift occurs in the future. Introducing two modifications to the CMP – a linear decrease in the fishing proportion when the aggregate abundance index drops below a threshold for the area concerned, and allowance for greater than 20% decreases in TACs in similar circumstances – considerably improves such conservation performance for the Eastern stock. However, some problematic R3-scenarios OMs remain when considering conservation performance for the Western stock. Placing a cap on the TAC for the East area can assist marginally in that respect, but also introduces some associated disadvantages. Suggestions are made for additional areas of investigation to refine the performance of this CMP further.

SCRS/2020/148 - An initial suggestion is put forward for an Operating Model (OM) to be used for development tuning. This is based on ordering median Br30 performance statistics for two CMPs applied to the 96 OMs of the interim grid of OMs. An OM is sought for which Br30 is near to the median value for all the OMs – this both for the Eastern and Western stocks, and for the two CMPs considered. The OM put forward on this basis is OM1.

SCRS/2020/149 - When evaluating Candidate Management Procedures (CMPs), a fundamental trade-off exists between catch performance (what is taken from a fish stock) and biomass performance (what remains after catches). CMPs typically include control parameters that alter how management advice is calculated from data, for example providing higher catches at the cost of long-term biomass. The control parameters of two functionally different CMPs were tuned so that the CMPs obtained comparable biological performance outcomes. In doing so the performance of the CMPs could be more clearly evaluated on a 'level playing field' at the same location in the catch biomass performance trade-off.

SCRS/2020/150 - A candidate management procedure to set total allowable catch advice from indices of abundance was designed that has three novel aspects. Firstly, it combines catch rate indices by area and spawning biomass indices by stock to infer regional abundance. This configuration has the advantage that TACs are set according to multiple sources of information and mixing is accounted for, for example allowing TACs in the western area to respond to fluctuations in productivity in the Eastern stock. Secondly, the P implements a harvest control rule that can account for both stock status (B/BMSY) and exploitation rate (F/FMSY). The advantage of this approach is that for example, a stock that is overfished and recovering (underfishing) does not necessarily incur a TAC reduction. Thirdly, the MP includes protocols for detecting and adjusting for chronic overfishing due to miscalibration of indices or large reductions in stock productivity. A preliminary test of six variants of the MP was carried out for the 96 operating models of the interim grid and the 12 primary robustness operating models.

SCRS/2020/151 - This document describes details of calculation for a candidate management procedure CMP) for management strategy evaluation of Atlantic bluefin tuna. The basic concept of this CMP is easy to understand, simple to use and cheap to maintain. TAC from this CMP could be determined by three indices for eastern and western area, respectively. This document consists of tentative flow chart and detailed equations for this CMP.

SCRS/2020/152 - The current assumptions for the Atlantic Bluefin tuna catch include within the West stock unit the catches off Brazil in the late 1950's to the early 1960's, commonly known as the "Brazilian episode". Under the development of the BFT MSE, a robustness test has been requested that proposes that all these catches are assumed to be part of the eastern stock. This document reviews the ICCAT database used in the current OMs, and provides the catch at age for the robustness test.

SCRS/P/2020/063 - Latest progress on refining and tunning the index based, which has been previously proposed for Atlantic bluefin tuna was provided with the most available version of the software (version 6.6.17). This CMP was index based, using 4 indices for the Eastern stock and 4indices for the Western stock. Following the suggestions at the previous BFTWG meetings, the method of weighting indices has been modified. The presentation shows estimated catch and SSB by stock using this CMP for some OMs.

Appendix 5

Specifications for MSE Trials for Bluefin Tuna in the North Atlantic
Version 20-3: September 24, 2020

Specifications for the MSE trials are contained in a living document that is under constant modification. The most recent version of the document (Version 20-3: September 24, 2020) can be found here.

Appendix 6

Specification of "Brazilian catches" robustness trial

For the "Brazilian catches" robustness test, the Group agreed that the following catches will be reallocated from the west stock (OM area 2) to the east stock units (OM area 4):

a) Catches between 1950 and 1970.
b) Catches between 20 N latitude and 20 S latitude in the West stock area.
c) All flags and gears, remove minor catches (< 10 t in total) if the catches occurred only for a several years.

Total catch amount in **Table 1** is transferred from the West to the East Atlantic for this robustness test. **Figure 1** shows the total catch in the West or in the East stocks before and after this transfer.

Table 1. Total catch amount transferred from the West to the East Atlantic.

Year	Catch (ton)
1957	30.00
1958	32.00
1959	199.85
1960	339.00
1961	373.00
1962	1212.42
1963	5997.09
1964	8660.32
1965	2328.08
1966	138.39
1967	49.54
1968	8.60
1969	14.16
1970	11.13

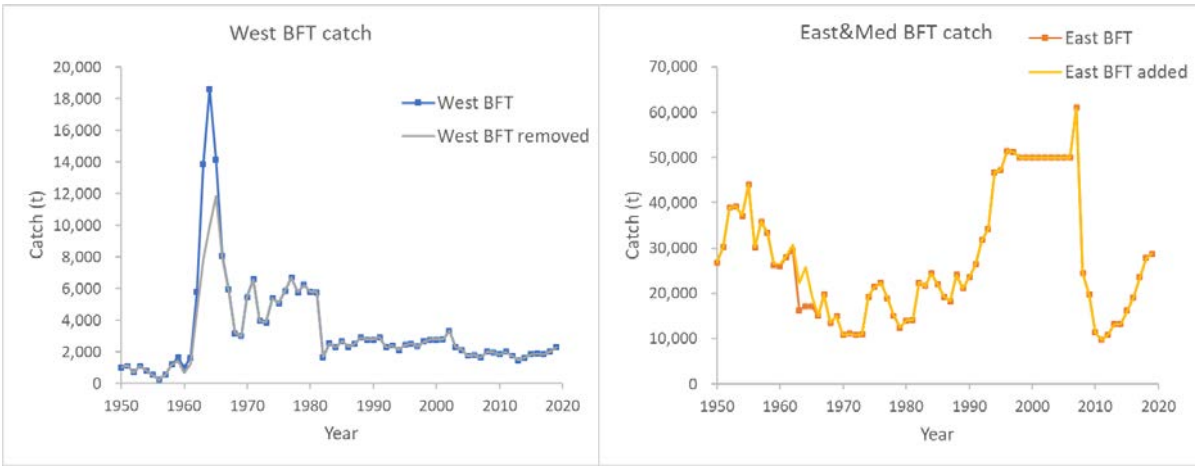

Figure 1. Total bluefin catch by stock in the period between 1950 and 2019, after moving the West BFT catches in the area between 20N and 20S latitude between 1950 and 1970 to the East Atlantic. The grey line for the West catch (left) and the yellow line for the East catch (right) are used in the robustness test for the "Brazilian catches".