

SPECIFYING AND WEIGHTING SCENARIOS FOR MSE ROBUSTNESS TRIALS

Levontin, P.¹, Leach, A.W.¹, Holt, J.¹ and Mumford, J.D.¹, Kell, L.²

SUMMARY

The GBYP Bluefin tuna rebuilding plan uses stochastic projections that do not capture all the uncertainty associated with stock assessment/ management variables. This could mean that the outcomes predicted by the projections are more optimistic or pessimistic than those that will be achieved in practice. A methodology was sought to capture stakeholder perceptions of particular uncertainties that should be included in stock assessments of Bluefin tuna and then to provide preliminary quantification of their relative importance impact on achieving management objectives. Ultimately, this will allow risk-based scenarios to be specified for the Operating Models (OM) used as part of a Management Strategy Evaluation and enable the SCRS and the GBYP Steering Committee to prioritise research. Given that the combinations of scenarios for inclusion in an MSE can grow exponentially with each extra variable, it may not be possible to evaluate the quantitative impact of all sources of uncertainties identified, or even prioritised. Therefore discussions with assessment scientists were conducted to reduce the initial list to those variables most amenable for further evaluation using simpler quantitative modelling approaches such as elasticity or scenario-based sensitivity analysis. In elasticity analysis the proportional change of derived values relative to changes in the input parameters allows the relative impact of the different inputs to be evaluated. Having determined which of the uncertainties have greater impact on derived values, measured using a utility function, discussions can be initiated with the stakeholders to elicit which of the shortlisted uncertainties should have priority for further quantitative investigations. Finally, a representative 'reference' set of Operating Models can be selected based on analysis of interactions among uncertainties. The plausibility weights for this reference set of OM provide another opportunity to engage stakeholders, and to elicit their views as to how robustness trials with the MSE should be 'tuned'. Having thus established an MSE framework, other sources of uncertainty from the qualitative analysis stage can be quantitatively addressed but it is still unlikely that every single one can be given a quantitative treatment. Therefore, elicitation process will also serve to document what is missing from the quantitative risk assessment, giving a more transparent and comprehensive view of uncertainties in the scientific advice to managers and other stakeholders.

RÉSUMÉ

Le plan de rétablissement du thon rouge fait appel à des projections stochastiques qui ne saisissent pas toute l'incertitude liée à l'évaluation des stocks / aux variables de gestion. Cela pourrait signifier que les résultats prévus par les projections sont plus optimistes ou pessimistes que ceux qui seront atteints dans la pratique. Une méthodologie a été recherchée afin de permettre aux intervenants de saisir les perceptions d'incertitudes particulières qui devraient être incluses dans les évaluations de stocks du thon rouge, puis de fournir une quantification préliminaire de leur incidence relative sur la réalisation des objectifs de gestion. En fin de compte, cela permettra de préciser les scénarios fondés sur les risques pour les modèles opérationnels (OM) utilisés dans le cadre d'une évaluation de la stratégie de gestion et de permettre au SCRS et au Comité de direction du GBYP d'établir l'ordre de priorité de la recherche. Étant donné que les combinaisons des scénarios à des fins d'inclusion dans une MSE peuvent croître de façon exponentielle avec chaque variable supplémentaire, il n'est peut-être pas possible d'évaluer l'impact quantitatif de toutes les sources d'incertitudes identifiées, ou même hiérarchisées. C'est pourquoi des discussions avec des scientifiques de l'évaluation ont

¹ Centre for Environmental Policy, Imperial College London, Silwood Park, Buckhurst Road, Ascot, Berkshire, SL5 7PY, United Kingdom.
j.mumford@imperial.ac.uk

² ICCAT Secretariat; Laurie.kell@iccat.int

été menées afin de réduire la liste initiale à ces variables plus propices à une évaluation plus approfondie en utilisant des approches de modélisation quantitatives plus simples, telles que les analyses de sensibilité fondées sur l'élasticité ou les scénarios. Dans l'analyse de l'élasticité, le changement proportionnel des valeurs calculées par rapport aux changements dans les paramètres d'entrée permet d'évaluer l'impact relatif des différents intrants. Ayant déterminé quelles incertitudes ont plus d'impact sur les valeurs calculées, mesurées à l'aide d'une fonction d'utilité, les discussions peuvent être lancées avec les parties prenantes pour déterminer quelles des incertitudes énumérées devraient avoir priorité pour de nouvelles recherches quantitatives. Enfin, un jeu de modèles opérationnels représentatif « de référence » peut être choisi sur la base de l'analyse des interactions entre les incertitudes. La pondération de plausibilité pour ce jeu de référence des modèles opérationnels fournit une nouvelle occasion de faire participer les intervenants et de solliciter leur avis quant à comment les essais de robustesse avec la MSE doivent être « calibrés ». Après avoir ainsi établi un cadre MSE, d'autres sources d'incertitude de l'étape de l'analyse qualitative peuvent être adressées quantitativement mais il est encore improbable que chacun puisse bénéficier d'un traitement quantitatif. Par conséquent, ce processus servira également à documenter ce qui est absent de l'évaluation quantitative des risques, donnant une vue plus transparente et plus détaillée des incertitudes présentes dans l'avis scientifique donné aux gestionnaires et aux autres parties prenantes.

RESUMEN

El plan de recuperación de atún rojo del GBYP utiliza proyecciones estocásticas que no capturan toda la incertidumbre asociada a la evaluación de stock/variables de ordenación. Esto podría significar que los resultados predichos por las proyecciones sean más optimistas o pesimistas que los que se lograrían en la práctica. Se buscaba una metodología para capturar las percepciones de las partes interesadas de las incertidumbres particulares que deberían ser incluidas en las evaluaciones de stock de atún rojo y posteriormente facilitar una cuantificación preliminar de su importancia relativa a la hora de lograr los objetivos de ordenación. Finalmente, esto permitirá especificar escenarios basados en el riesgo para los modelos operativos (OM) utilizados como parte de una evaluación de la estrategia de ordenación y permitirá al SCRS y al Comité directivo del GBYP establecer prioridades en la investigación. Teniendo en cuenta que las combinaciones de escenarios para su inclusión en una MSE pueden crecer exponencialmente con cada variable extra, podría no ser posible evaluar el impacto cuantitativo de todas las fuentes de incertidumbre identificadas, o incluso priorizadas. Por tanto, se mantuvieron discusiones con los científicos de la evaluación para reducir la lista inicial a aquellas variables que más se prestan a una evaluación más profunda utilizando enfoques de modelación cuantitativa más simples como análisis de sensibilidad de elasticidad o basados en el escenario. En el análisis de elasticidad el cambio proporcional de los valores derivados en relación con los cambios en los parámetros de entrada permite evaluar el impacto relativo de los diferentes datos de entrada. Una vez determinado cuáles de las incertidumbres tienen más impacto en los valores derivados, medido utilizando una función de utilidad, pueden iniciarse las discusiones con las Partes interesadas para dilucidar cuáles de las incertidumbres incluidas en la lista debería tener prioridad para realizar más investigaciones cuantitativas. Por último, un conjunto representativo de modelos operativos de "referencia" puede seleccionarse basándose en un análisis de las interacciones entre las incertidumbres. Las ponderaciones de plausibilidad para este conjunto de modelos operativos de referencia proporciona otra oportunidad de involucrar a las partes interesadas, y para obtener sus opiniones respecto a cómo deberían "ajustarse" los ensayos de robustez con la MSE. Habiendo establecido así un marco de MSE, pueden abordarse cuantitativamente otras fuentes de incertidumbre del análisis cuantitativo, pero sigue siendo poco probable que cada una de ellas pueda recibir un tratamiento cuantitativo. Por tanto, el proceso de dilucidación servirá también para documentar lo que falta en la evaluación del riesgo cuantitativa, ofreciendo una perspectiva más amplia y transparente de las incertidumbres existentes en el asesoramiento científico a los gestores y otras partes interesadas.

KEYWORDS

*Bluefin tuna, Thunnus thynnus, GBYP,
Stock assessment, Risk analysis, MSE, Uncertainty*

1. Introduction

Although several sources of uncertainty were considered when formulating the East Atlantic and Mediterranean Bluefin Tuna Recovery Plan, not all sources were explicitly considered. Therefore, a contract for a Risk Assessment was awarded under Phase III to identify the main sources of uncertainty and the legitimate concerns of a wide range of stakeholders. Subsequently the meeting on bluefin stock assessment methods (Anon., 2014) endorsed this work and recommended that the major sensitivities for both separate and mixed stock assessments (e.g., M, fecundity schedule, SRR and alternative mechanism of population regulation) should be identified in a paper on a Risk Assessment. It was also recommended that this paper be used to inform the choice of Operating Model (OM) scenarios to be used in the bluefin Management Strategy Evaluation (MSE). Therefore in this study we develop tools to help turn a qualitative study into a quantitative one that can be used to help identify simulation trials (i.e., scenarios) to be used in a MSE. We do this using the Bluefin population model described in Kell (2014).

As discussed in Kell (2014) when building an OM it is necessary to develop hypotheses about system dynamics that can be run as part of stochastic Monte Carlo simulations. However, Monte Carlo simulations are costly in terms of time and resource to conduct. Therefore there is benefit in first running deterministic (or a limited number of stochastic) simulations to identify main effects or important interactions. Following this, fully stochastic simulations can be run for the trials (i.e. scenarios) that are considered to be important.

To do this deterministic runs are conducted initially to explore the dynamics and the effect of model and value uncertainty. Stochastic simulations that include observation and process error and assessment procedures with feedback will be done later as part of the MSE. This approach means that rather than running all possible combinations of treatments, an experimental design can be used to run only main effects and selected interactions.

2. Methods

2.1 *Qualitative analysis: identified and prioritized uncertainties based on stakeholder perceptions*

In the previous phases of the Risk Assessment (Leach *et al.*, submitted) a list of uncertainties relevant to the management of the Eastern bluefin tuna stock were identified. This list was completed through interviews with stakeholders: GBYP scientists, managers and NGO observers. Semi-quantitative feedback was elicited from 28 stakeholders which allowed a prioritisation of 33 uncertainties in terms of their urgency for testing in a quantitative manner. It was discovered that individual stakeholder opinions differed on all three dimensions in which uncertainties were analysed: 1). importance in the management context - what impact a source of uncertainty is anticipated to have on the probability of achieving management objectives; 2). how likely it is that a particular uncertainty could be reduced by investing in research; 3). to what extent it was already represented in the current GBYP assessment. In order to understand the reasons for disagreements and explore the possibility of achieving consensus in a larger group, a focus workshop group of five people (four scientists and an NGO representative) was conducted (Tenerife, Spain). Through a group discussion, we succeeded in achieving consensus opinion on the most relevant dimension of these uncertainties to risk management, Importance. Using both these consensus scores and the overall stakeholder responses, we compiled a 'high value' target list of 20 uncertainties that are most important candidates for the quantitative treatment (**Table 1**).

In this paper we discuss a methodology for identify the uncertainties that have the biggest impact on management advice and which should be included in the MSE trials and provide an illustrative example.

2.2 *Quantitative Analysis: approaches to quantifying uncertainty*

The list of priorities (**Table 1**) identified in the previous step is subject to computational constraints as some are more difficult than others to model. Thus following discussions with modellers, an initial set of uncertainties was selected for the quantitative assessment. Each of these uncertainties was broken into scenarios or levels to characterise it prior to expressing these differences mathematically within a simpler model used to re-assess the importance of selected uncertainties by quantifying the impact it has on key model outputs. It was proposed that, prior to inclusion in operating model scenarios, these are first evaluated with a simpler approach to gauge the quantitative impact each source of uncertainty is likely to have on the operating model. Further, each was measured against a base case scenario before looking at the interactive effects these uncertainties might produce.

2.2.1 Model description

Biological parameters are taken from the ICCAT assessment (Kell *et al.*, 2011), i.e. average-at-age vectors for mass, proportion mature, natural mortality and selectivity derived from the Adapt-VPA stock assessment. These values can be changed to evaluate the sensitivity of the utility functions to the assumptions.

2.2.1.1 Model

Life-history traits of bluefin as assumed in the assessment are:

- annual spawning (1 cohort per year),
- 50% maturity at age 4, 100% maturity at ages 5+,
- fecundity is linearly proportional to weight,
- growth following the von-Bertalanffy equation used in the ICCAT working group (with the following parameters: $L_{\infty} = 318.85$, $k=0.093$, $t_0=-0.97$),
- length-weight relationship used in the ICCAT working group ($W=2.95 \cdot 10^{-5} \cdot L^{2.899}$),
- lifespan of 40 years,
- age-specific, but time-invariant, natural mortality based on tagging experiments on the southern bluefin tuna and used in the ICCAT working group (i.e. $M=0.49$ for age 1, $M=0.24$ for ages 2 to 5, $M=0.2$ for age 6, $M=0.175$ for age 7, $M=0.15$ for age 8, $M=0.125$ for age 9 and $M=0.1$ for ages 10 to 20).

Given the selection pattern (s) of a fishery, and the catchability (q) of a population for a given effort (E), the fishing mortality rate ($F_{a,y,j}$) for age a , year y , and population j is given by:

$$F_{a,y,j} = E_y * q_j * s_{a,j}$$

The abundance (N_j) at age $a+1$, at the start of year $y+1$, in sub-population j , is:

$$N_{a+1,y+1,j} = N_{a,y,j} * \exp(-F_{a,y,j} - M_{a,y,j})$$

Recruitment was modelled as a Beverton and Holt relationship with a fixed value of steepness.

See Kell (2014) for full details.

2.2.1.2 Alternative scenarios

Kell (2014) summarised the different sources of uncertainty, i.e. process, observation, estimation, implementation, model, value, translational or institutional. In this study (as explained above) only model and value uncertainty are considered and possible modelling approaches are summarised in the final column. These are summarised as factors with levels in **Table 2**. The 3rd column summarises the number of levels for each factor. The 4th column shows the cumulative number of scenarios if only the main effects are modelled, i.e. a single level is varied in the base case at a time.

The first two factors **Historic Catch** and **Future Recruitment** were the sources of uncertainty included in the assessment and projections used to calculate the K2SM. The values of steepness chosen were 1 (as assumed in the assessment) and 0.7 an arbitrary value to provide some contrast. Natural mortality was either that assumed by the working group (SCRS) or derived from weight-at-age (Lorenzen, 1996). To evaluate the effect of artisanal fisheries juvenile mortality was increased by a factor of 1.5. Plus group dynamics were evaluated for an increase in mortality and by setting the F_{ratio} to 1. The working group had estimated the F_{ratio} and that value (SCRS) was used for the base case.

2.3 Utility function

A decision is needed on how to evaluate model sensitivity because we are interested in a wide range of outputs, each of which could have a different sensitivity/elasticity to a given source of uncertainty. Examples of such model outcomes include time taken to rebuild the stock, risk of stock collapse and short versus long term yield; each of which may be assigned different importance (and weights) by different stakeholders. In economics Utility represents a relative measure of the satisfaction experienced by the consumer of a commodity. Different stakeholders may have different utilities, for example long term yield and time taken to rebuild the stock may be more important to some stakeholders than short term yield, although reducing the risk of stock collapse may be equally important to both groups. To have a measure of model sensitivity for each group we need to define a utility function (ICES, 2007) and then weight the different components. These utility functions can then be used to evaluate the impact of the different sources of uncertainty.

It has been highlighted in several studies that eliciting management objectives which are often ambiguous, conflicting or simply never made completely explicit is often an impediment to modeling or risk analysis in general (ICES, 2007; Leach *et al.*, submitted).

To construct a utility function we first need to identify management objectives.

2.3.1 Objectives

First we summarise the “explicit” management objectives for bluefin i.e. those in the ICCAT Basic Texts and in the Commission Recommendations made when the bluefin recovery plan was implemented. We then discuss “implicit” objectives based on The Principles Of Decision Making For ICCAT Conservation And Management Measures consistent with the Precautionary approach and the Straddling stocks agreement and other conventions such as CITES and IUCN.

2.3.1.1 Explicit

The main management objective of ICCAT is to maintain the populations of tuna and tuna-like fishes at levels which will permit the maximum sustainable catch. Originally interpreted as using F_{MSY} as a target, in 2007 a 15 year Recovery Plan was implemented with the goal of achieving B_{MSY} with at least a 60% probability by 2022. B_{MSY} was based on $F_{0.1}$, a proxy for F_{MSY} . $F_{0.1}$ is the point on the yield per F curve where the slope equals 10% of that at the origin, B_{MSY} is estimated by multiplying the spawner-per-recruit at $F_{0.1}$ by the assumed level of recruitment. The corresponding objectives are to achieve the maximum long-term yield and ensure that by 2022 the stock is greater than B_{MSY} with a 60% probability.

2.3.1.2 Implicit

We may also consider objectives based on a variety of agreements, that although not explicitly included in the recovery plan or in the Basic Text are included in recent conventions such as the United Nations Conference on Straddling Fish Stocks and Highly Migratory Fish Stock Agreement (UNFSA). The objective of UNFSA is to ensure the long-term conservation and sustainable use of straddling fish stocks and highly migratory fish stock consistent with the precautionary approach (see <http://daccess-ods.un.org/TMP/8829557.8956604.html>).

Both the Straddling Stocks agreement and the PA were signed after the Basic Text of ICCAT. However, the principles of decision making [Rec 11-13] note that management decisions should be based upon scientific advice consistent with the precautionary approach. Therefore although not explicitly stated in the bluefin recovery plan in this study we consider management objectives based on the PA and Straddling Fish Stocks agreements.

There are other Conventions which could potentially impact on the management of bluefin, e.g. CITES and the IUCN redlist. A proposal for listing Atlantic bluefin on CITES appendix I and II was made in 2009, the criteria for a CITES listing for a commercial species are given in footnote 2 in CITES Conf 9.24. Atlantic bluefin is also classed as ‘Endangered’ on the IUCN Red list (Collette *et al.*, 2011; IUCN, 2013) based on a combination of factors including limited range, inferred low densities and presumed unsustainable interactions with fisheries.

2.3.2 Constructing a utility function

For a stochastic model, we can construct annual utility function considering four components (risk of overfishing, risk of being below biological reference point, a measure of yield relative to some desired level, and variability of recruitment), as follows:

$$U_y = \text{prob}(F < F_{MSY}) + \text{prob}(SSB > B_{MSY}) + \frac{\bar{Y}}{Y_{max}} + \frac{1}{CV_{rec}} = U_{F,y} + U_{B,y} + U_{Y,y} + U_{R,y}$$

However, this might reflect how satisfied we are with the stock in a particular year but more distant outcomes are less valuable to us than the current ones. Thus over a period of time we propose a utility function with built-in discount rates for each of the four component measures relating to objectives. So over the simulated time period the total utility will be:

$$U = \sum_y [U_{F,y}(1 - d_{S,F})^{y-1} + U_{B,y}(1 - d_{S,B})^{y-1} + U_{Y,y}(1 - d_{S,Y})^{y-1} + U_{R,y}(1 - d_{S,R})^{y-1}]$$

Where $U_{F,y}$ means Utility with respect to F and is for each year equals to $\text{prob}(F < F_{MSY})$.

This is just one possible representation of a utility function and will use several variations of this general form to assess the sensitivity of the base case model to alternative scenarios described in **Table 2**.

2.4 Progress score with respect to incorporating the uncertainty

We developed a measure of where we are with respect to including uncertainty relative to an ideal situation where we have quantified the impact of every source of uncertainty and using research reduced uncertainty itself as much as is possible. In our expert elicitations we considered the possibility of reducing uncertainty which is represented in **Table 1** by the size of the hoops. Small hoops corresponded to views that uncertainty could not be reduced further, large hoops reflected the view that a lot can be done still to reduce a particular source of uncertainty (through data collection, research, *etc.*). The position of the hoops with respect to the Y-axis depicted the importance or the weight each source of uncertainty was believed to have in the assessment/management of the stock. Utility analysis supersedes expert opinion on the weight each source of uncertainty is believed to carry for the assessment because it could be demonstrated that the impact of such uncertainty on the utility function is of particular magnitude. Assuming that having a quantitative analysis is preferable to merely a qualitative one, we can now calculate how far we are from achieving a relative score of 100% of incorporating all important sources of uncertainty. A tentative score function is as follows:

$$P = \sum_u W_u * V_u * A_u$$

Where:

P Progress

W_u Weight (Importance) of uncertainty

V_u Value of reducing uncertainty

A_u Assessment level (binary function (qualitatively assessed = 2, quantitatively assessed = 1))

Given the weights, we calculate the minimum score possible = P_{min} , this is achieved when all ratings for the value of reducing uncertainty are 1 (implying that it could not be reduced further through research) and all sources have been quantitatively assessed. P_{max} is the maximum score where all ratings for the value of reducing uncertainty are set a maximum and no variable had been quantitatively assessed.

The overall current score is then:

$$Score = 1 - \frac{P - P_{min}}{P_{max} - P_{min}}$$

That is, when $P = P_{min}$ the score is 100%.

Using this method of measuring the progress towards reducing, quantifying and including uncertainty we are currently at 35% (footnote of **Table 1**). This score can be improved in two ways: conducting quantitative assessments and investing in research to reduce uncertainty if there is scope for those sources with high importance weightings.

3. Results of Illustrative Trial

Tables 3, 4 and 5 below compare three different methods to measure the ‘importance’ of a particular source of uncertainty. Eight scenarios are compared to the base case using three different utility functions each a combination of ‘Biological’, ‘Economic’ and ‘Social’ components. In the ‘Combined’ column, the three components are simply added implying they have equal weights in the overall utility. The first ‘Absolute’ utility function uses absolute values for the biomass of fish, for yield and for fishing mortality which is a proxy for employment, hence the ‘social’ component, as maintaining employment is one of the management objectives, **Table 3**. We assume that catchability is time invariant and so F is a proxy for effort and hence an index of employment; although this is unlikely to be true and is used here for illustrative purposes only. The second utility function (**Table 4**) uses relative values, where historical averages presented in part a) of the table, are used to normalise biomass, yield and fishing mortality. In the third version of a utility function, the values are normalised using MSY reference points, (**Table 5**). In general, the utility function we used in illustrative trials has the form in the equation below:

$$U = \sum_{y=1}^{11} [U_{Biol,y}(1-d)^{y-1} + U_{Econ,y}(1-d)^{y-1} + U_{Soc,y}(1-d)^{y-1}]$$

Where the sum covers 11 years for the period 2013-2023 ($y=1$ corresponds to 2013, *etc.*) and $U_{Biol,y}$ is related to the sum of SSB and Plus Group Biomass for the respective year, $U_{Econ,y}$ to yield, and $U_{Soc,y}$ equals fishing mortality (F) in ‘absolute’ case, F/F_{MSY} in the MSY case and F divided by the average fishing mortality for 2010-2012 in the historical reference points formulation of the utility function.

Using absolute values is inconvenient if we want to consider trade-offs by combining different components, in that case adding biological and social components makes little sense even if quantitatively they were on the same scale (which is clearly not the case in **Table 3**) since they have incomparable units of measure. But normalising (dividing) values by either historically meaningful quantities (**Table 4**) or MSY based calculations (**Table 5**) makes all components of the utility function unit-less thus enabling us to make sense of the overall, or ‘Combined’ utility, the last column of **Table 4 and 5**.

The percentage values show how each component of the overall utility and the total were affected by each scenario relative to the base case: the difference between utility of the scenario and the base case divided by the value of the utility function (component) for the base case.

The last column shows how the overall utility function has changed with each scenario relative to the base case, and if we had an agreed with the stakeholders form of the utility function these would be the primary numbers of interest. We can see that uncertainty over reported catches has a lot of influence, no matter which utility function is used. Recruitment and Steepness scenarios appear important when the utility function is constructed out of values normalised by MSY reference points which are allowed to change with the scenario. Clearly, the scenario influences the reference points and hence our perception of how satisfied we are with achieving management goals. It is suggested by the results that scenarios related to natural mortality are not important. This is in clear contradiction with the expert opinion which rated the importance of uncertainty related to natural mortality at age as ‘major’, **Table 1**.

The revenue component does not change in the first two variants of the utility function because the model uses constant catch projections, however in the utility function that evaluates scenarios relative to perceived MSY reference points (which themselves change depending on the scenario) the revenue component does vary, suggesting for instance that the fishery is underperforming from the economic standpoint in the ‘high recruitment’ scenario. This is because the high recruitment causes the estimated Yield at MSY to be higher and so the same level of catch looks worse in comparison with a higher standard.

All three utility functions use the same discount rate ($d = 5\%$) and equal weights for the components. However, we don't know stakeholder preferences. Different stakeholders might give different weights to different components and also prefer different discounting factors for each component. Higher discount rates may be preferred by stakeholders interested in economic returns whereas stakeholders with longer term environmental and sustainability interests may prefer the use of smaller, zero or even negative discount rates (Teh *et al.*, 2011; HM Treasury, 2011).

The tables below show how sensitive our conclusions about the importance of different sources of uncertainties are to the choice of the utility function: what would be vital to understand also is how robust the quantification of uncertainty is to the choice of the modelling approach. How do these results depend on the model itself and the modelling approach more generally?

Figures 1-4 show the time series of SSB, Yield, Fishing mortality and Plus Group Biomass under all nine scenarios. The three panels in each Figure correspond to utility function's specifications: the top presents absolute values (**Table 3**), the middle – values relative to a historical average (**Table 4**) and the bottom panel – values relative to calculated MSY reference points (**Table 5**).

4. Discussion

The scenario analysis is insufficiently informative about uncertainties since there is a potential, given the nature of non-linear models, that the combined impact of some uncertainties is far more significant than singular deviations from the base-case can reveal. But in order to include all these combinations, we would need to consider an impractical number of model runs. So how do we choose which combinations to investigate further?

Deterministic analysis with a simple model, as was used for illustration purposes in this paper, can form a basis for the selection. For instance, those uncertainties in which the model was found not particularly sensitive could be fixed in the future analysis. However, as we have demonstrated, quantification of uncertainties is dependent upon the choice of a utility function which should be specified through a transparent stakeholder elicitation process. As well as needing to select a specific utility function, a model needs to be chosen to form a basis for quantifying uncertainties. It is possible to employ different modelling approaches for different types of uncertainties but it is desirable to understand how robust quantification of uncertainty is to the modelling approach. Do we get the same ranking of uncertainties if we use quick and simple deterministic evaluations as we would have gained from a more detailed stochastic modeling approach? The next step would be to repeat the analysis done in this paper with a different modelling approach and compare the results. We have already pointed to the disagreements between the model based and the expert based rankings of uncertainties. Which is more reliable? Would the experts concede that the model provides a greater insight or would they be suspicious of model assumptions that might mislead the analysis of uncertainties?

Discussions with modellers and/or stakeholders should inform the process of finalising the list of scenario combinations to test following a further evaluation of a more complete set of uncertainties. If we are interested in pursuing other uncertainties which were highlighted in the qualitative analysis, such as risk attitudes of managers or impacts of the regulations, we would need to reduce the list of possible OM models still further. This final reference set for the MSE framework could consist of five or fewer models chosen out of a larger list ranked according to utility function values so that the reference set covers the entire range of utility scores. If resources permit, expert views would be solicited and accounted for at this final stage of the analysis instead. **Appendix 1** lists the forms of uncertainty and possible ways to include them in the MSE.










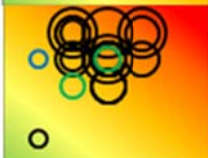
Acknowledgements

This work was funded by short-term contract for the Stock Assessment Modelling (GBYP 04/2011) of the Atlantic-Wide Research Programme on Bluefin Tuna (ICCAT GBYP – Phase 2). The authors would like to give special thanks to Laurie Kell from ICCAT for making invaluable contributions to the concepts, explanation and discussion in the paper and for providing the modelling results as well as graphics used in the paper.

References

- Anon. 2014. Report of the 2013 Meeting on Bluefin Stock Assessment Methods. Gloucester, Massachusetts, United States – July 20 to 22, 2013. ICCAT Col. Vol. Sci. Pap. 70(1): 160-189.
- Beverton, R.J.H. and Holt, S.J. (1957) On the Dynamics of Exploited Fish Populations. London: Chapman and Hall, 533 pp.
- CITES (2010). Proposal to include Atlantic Bluefin Tuna (*Thunnus thynnus* (Linnaeus, 1758)) on Appendix I of CITES in accordance with Article II 1 of the Convention. Consideration of Proposals for amendment of Appendices I and II. Convention on International Trade in Endangered Species of Wild Fauna And Flora, fifteenth meeting of the Conference of the Parties Doha (Qatar), 13-25 March 2010. CoP15 Prop. 19. pp. 34. <http://www.cites.org/eng/cop/15/prop/E-15-Prop-19.pdf> (last accessed 11 November 2013).
- Collette, B., Amorim, A.F., Boustany, A., Carpenter, K.E., de Oliveira Leite Jr., N., Di Natale, A., Die, D., Fox, W., Fredou, F.L., Graves, J., Viera Hazin, F.H., Hinton, M., Juan Jorda, M., Kada, O., Minte Vera, C., Miyabe, N., Nelson, R., Oxenford, H., Pollard, D., Restrepo, V., Schratwieser, J., Teixeira Lessa, R.P., Pires Ferreira Travassos, P.E. & Uozumi, Y. 2011. *Thunnus thynnus*. In: IUCN 2013. IUCN Red List of Threatened Species. Version 2013.2. www.iucnredlist.org. Downloaded on 26 November 2013.
- Kell, L.T., Fromentin, J.-M., Bonhommeau, S. 2011. An Evaluation of the Implications of Population Structure on the Current Bluefin Tuna Advice Framework. ICCAT Col. Vol. Sci. Pap. 68(1): 370-378.
- Kell, L.T., Merino, G., De Bruyn, P., Ortiz de Urbinaz, J, Arrizabalaga, H., Santiago, J., Murua H. 2013. An Example Management Strategy Evaluation of a Harvest Control Rule. ICCAT Col. Vol. Sci. Pap. 70(5): 2096-2110.
- Kell, L.T. Identification of the major sensitivities in the East Atlantic and Mediterranean bluefin assessment. ICCAT Collective Volume of Scientific Papers, 71. *In this Volume*.
- Leach, A.W., Levontin, P., Holt, J., Kell, L. T. & Mumford, J.D. Identification and prioritisation of uncertainties for management of Eastern Atlantic Bluefin Tuna (*Thunnus thynnus*). Submitted to Ecological Applications, 20/11/2013.
- Lorenzen, K. 1996. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. *Journal of fish biology* 49:4: 627-642.
- HM Treasury. 2011. The Green Book: Appraisal and Evaluation in Central Government. TSO, London. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220541/green_book_complete.pdf
- ICES. 2007. Report of the Study Group on Risk Assessment and Management Advice (SGRAMA), 5–9 February 2007, Cape Town, South Africa. RMC:02. 68 pp.
- IUCN 2013. The IUCN Red List of Threatened Species. Version 2013.2. <http://www.iucnredlist.org>. Downloaded on 21 November 2013.
- Marshall, T., (2009). Implementing information on stock reproductive potential in fisheries management: The motivation, challenges and opportunities. In Jakobsen, T., Fogarty, M.J., Megrey, B. A., Monksness, E., (Eds), *Fish reproductive biology. Implications for assessment and management*. Wiley-BlackWell. JohnWiley & Sons, Ltd. Oxford, UK. pp. 395-420.
- Teh, L.S.L., Teh L.C.L. and Sumaila U. R. Low Discounting (2011) Behavior among Small-Scale Fishers in Fiji and Sabah, Malaysia. *Sustainability* 3: 897-913 doi:10.3390/su3060897.

Table 1. Prioritised list of top 20 uncertainties that were agreed by the focus group to be either massively or majorly important. The Quantified columns are all “2” at this baseline stage and as these are included in the assessments many of these will become “1” (full description in Section 2.4).

| Variable | Individual elicitations | Uncertainty medians | Weights (Importance) | Quantified: (1 = Yes, 2 = No) | Value |
|--|---|---------------------|----------------------|--------------------------------|-------|
| Catch under-reporting - in particular of juvenile catch in artisanal fisheries |  | 4 | 5 | 2 | 40 |
| Variability in migration patterns |  | 4 | 5 | 2 | 40 |
| Risk attitudes of managers |  | 3 | 5 | 2 | 30 |
| Social impacts of regulations and its affect on small local communities |  | 3.5 | 5 | 2 | 35 |
| Environmentally driven recruitment variability and density dependence |  | 4 | 4 | 2 | 32 |
| Natural mortality at age, variability, age related senescence |  | 4 | 4 | 2 | 32 |
| Steepness (meta-analysis) |  | 4 | 4 | 2 | 32 |
| Standardisation across gear, countries, areas and time |  | 3.5 | 4 | 2 | 28 |
| Generation of age data, age-length keys, slicing |  | 4 | 4 | 2 | 32 |
| Growth |  | 3 | 4 | 2 | 24 |











| | | | | | |
|--|---|-----|---|---|----|
| Migration between ICCAT agreed stock units |  | 3 | 4 | 2 | 24 |
| Maturation and fecundity |  | 3 | 4 | 2 | 24 |
| Changes in regulations translating into changes of fishing practices |  | 3 | 4 | 2 | 24 |
| Environmental factors that influence migration patterns |  | 4 | 4 | 2 | 32 |
| Complexity of tuna habitat |  | 4 | 4 | 2 | 32 |
| Ecological/environmental (other than climate change) potential to change population dynamics |  | 4 | 4 | 2 | 32 |
| Climate change and/or increased variability's potential to change population dynamics |  | 4 | 4 | 2 | 32 |
| Spawning, periodicity, aggregation and location of spawning areas |  | 3 | 4 | 2 | 24 |
| Group dynamics, skipped spawning, density dependence |  | 4 | 4 | 2 | 32 |
| Impacts of regulations and its effect on the species' apparent global distribution. |  | 3 | 4 | 2 | 24 |
| <hr/> | | | | | |
| Minimum possible score (all sources of uncertainty are very well understood and their impact quantified): | | 84 | Score achieved: 575 | | |
| Maximum possible score (qualitative assessment shows that all sources of uncertainty are poorly understood): | | 840 | Progress (percentage) towards quantifying all important sources of uncertainty: 35% | | |

Table 2. Scenarios. Bold text indicates parameter values used for the base case.

| <i>Factor</i> | <i>Levels</i> | <i>N</i> | <i>Σ Main Effects</i> |
|------------------------|----------------------------|----------|-----------------------|
| Historic Catch | Reported , Inflated | 2 | 2 |
| Future Recruitment | Medium , Low, High | 3 | 4 |
| Steepness | 1 , 0.7 | 2 | 5 |
| Natural Mortality | SCRS , Life History | 2 | 6 |
| Juvenile Mortality | $M_1 \times (1, 1.5)$ | 2 | 7 |
| Plus Group F_{ratio} | SCRS , 1.0 | 2 | 8 |
| Plus Group Mortality | $M_{PG} \times (1, 2)$ | 2 | 9 |

Table 3. Utility function using absolute values.

| a) | Scenario | Utility | | | Combined |
|----|---------------------|---------------|---------------|-------------|-----------------|
| | | Biological | Economic | Social | |
| | Base case | 4.E+09 | 1.E+08 | 0.36 | 3.96E+09 |
| | + Catch inflated | 8.E+09 | 1.E+08 | 0.14 | 7.76E+09 |
| | + Recruitment low | 4.E+09 | 1.E+08 | 0.41 | 3.70E+09 |
| | + Recruitment high | 4.E+09 | 1.E+08 | 0.32 | 4.28E+09 |
| | + Steepness 0.7 | 4.E+09 | 1.E+08 | 0.36 | 4.00E+09 |
| | + M Lorenzen | 4.E+09 | 1.E+08 | 0.37 | 3.93E+09 |
| | + M_1 1.5 | 4.E+09 | 1.E+08 | 0.36 | 3.96E+09 |
| | + F_{ratio} false | 4.E+09 | 1.E+08 | 0.36 | 3.96E+09 |
| | + PlusGrpM 2 | 4.E+09 | 1.E+08 | 0.37 | 3.89E+09 |

| b) | Scenario | % Change in Utility | | | Combined |
|----|---------------------|---------------------|----------|--------|----------|
| | | Biological | Economic | Social | |
| | Base case | | | | |
| | + Catch inflated | 99% | 0% | -62% | 96% |
| | + Recruitment low | -7% | 0% | 13% | -6% |
| | + Recruitment high | 8% | 0% | -12% | 8% |
| | + Steepness 0.7 | 1% | 0% | -2% | 1% |
| | + M Lorenzen | -1% | 0% | 1% | -1% |
| | + M_1 1.5 | 0% | 0% | 0% | 0% |
| | + F_{ratio} false | 0% | 0% | 0% | 0% |
| | + PlusGrpM 2 | -2% | 0% | 1% | -2% |

Table 4. Utility function using values relative to historical reference points.

| a) | Ref points | Mean | Meaned years | b) | Scenario | Utility | | | Combined |
|----|------------|-------------|--------------|----|----------------------------|-------------|------------|------------|-------------|
| | | | | | | Biological | Economic | Social | |
| | SSB | 263,000,000 | 1950-1980 | | Base case | 18.5 | 5.3 | 9.2 | 33.0 |
| | PlusGroup | 158,000,000 | 1950-1980 | | + Catch inflated | 36.4 | 5.3 | 3.5 | 45.2 |
| | Yield | 23,000,000 | 1950-1980 | | + Recruitment low | 17.4 | 5.3 | 10.4 | 33.1 |
| | F | 0.0396 | 2010-2012 | | + Recruitment high | 19.9 | 5.3 | 8.1 | 33.3 |
| | | | | | + Steepness 0.7 | 18.7 | 5.3 | 9.0 | 33.0 |
| | | | | | + M Lorenzen | 18.4 | 5.3 | 9.3 | 32.9 |
| | | | | | + M ₁ 1.5 | 18.5 | 5.3 | 9.2 | 33.0 |
| | | | | | + F _{ratio} false | 18.5 | 5.3 | 9.2 | 33.0 |
| | | | | | + PlusGrpM 2 | 18.2 | 5.3 | 9.3 | 32.8 |

| c) | Scenario | % Change in Utility | | | Combined |
|----|----------------------------|---------------------|----------|--------|----------|
| | | Biological | Economic | Social | |
| | Base case | | | | |
| | + Catch inflated | 97% | 0% | -62% | 37% |
| | + Recruitment low | -6% | 0% | 13% | 0% |
| | + Recruitment high | 7% | 0% | -12% | 1% |
| | + Steepness 0.7 | 1% | 0% | -2% | 0% |
| | + M Lorenzen | -1% | 0% | 1% | 0% |
| | + M ₁ 1.5 | 0% | 0% | 0% | 0% |
| | + F _{ratio} false | 0% | 0% | 0% | 0% |
| | + PlusGrpM 2 | -2% | 0% | 1% | -1% |

Table 5. Utility function using values relative to MSY reference points (updated for each scenario).

a)

| Scenario | Utility | | | Combined |
|----------------------------|-------------|------------|------------|-------------|
| | Biological | Economic | Social | |
| Base case | 43.9 | 4.8 | 2.7 | 51.3 |
| + Catch inflated | 82.7 | 3.2 | 1.0 | 86.9 |
| + Recruitment low | 57.1 | 6.6 | 3.0 | 66.7 |
| + Recruitment high | 35.0 | 3.6 | 2.4 | 40.9 |
| + Steepness 0.7 | 19.4 | 5.2 | 5.0 | 29.6 |
| + M Lorenzen | 43.5 | 4.8 | 2.7 | 51.0 |
| + M ₁ 1.5 | 43.9 | 4.8 | 2.7 | 51.3 |
| + F _{ratio} false | 43.9 | 4.8 | 2.7 | 51.3 |
| + PlusGrpM 2 | 43.1 | 4.8 | 2.7 | 50.5 |

b)

| Scenario | % Change in Utility | | | Combined |
|----------------------------|---------------------|----------|--------|----------|
| | Biological | Economic | Social | |
| Base case | | | | |
| + Catch inflated | 88% | -32% | -64% | 69% |
| + Recruitment low | 30% | 38% | 13% | 30% |
| + Recruitment high | -20% | -25% | -12% | -20% |
| + Steepness 0.7 | -56% | 9% | 89% | -42% |
| + M Lorenzen | -1% | 0% | 1% | -1% |
| + M ₁ 1.5 | 0% | 0% | 0% | 0% |
| + F _{ratio} false | 0% | 0% | 0% | 0% |
| + PlusGrpM 2 | -2% | 0% | 1% | -2% |

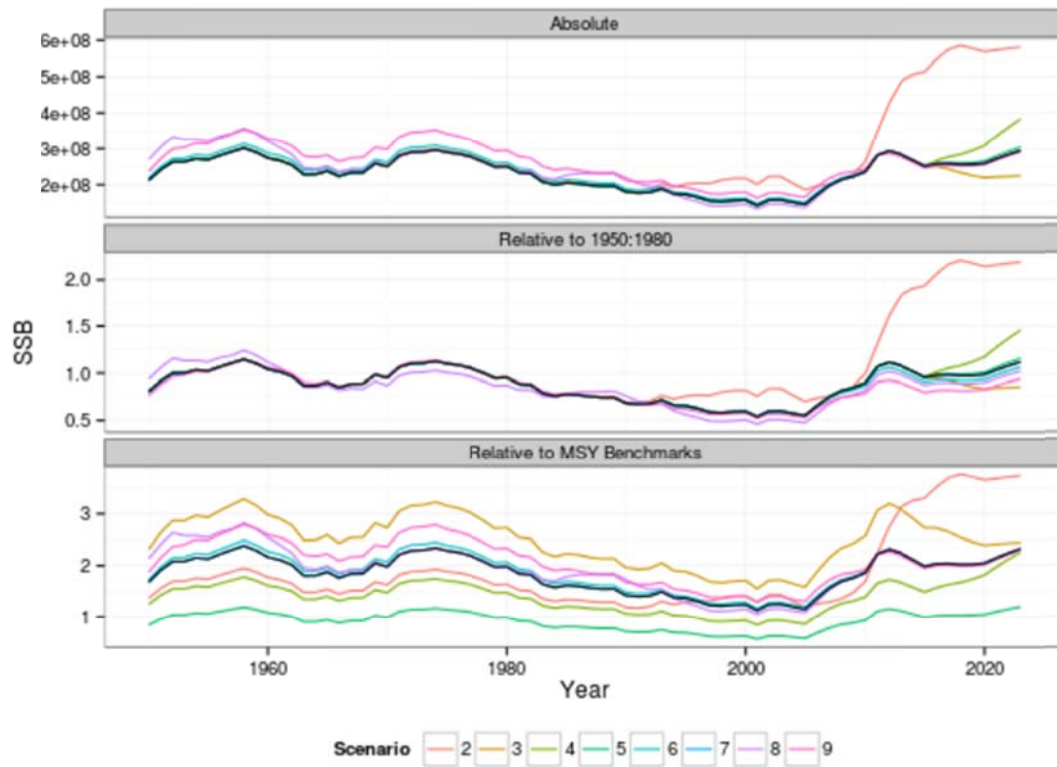


Figure 1. Spawning stock biomass in 9 scenarios from 1950-2023. Scenario 1 (the base case) is shown as the heavy black line.

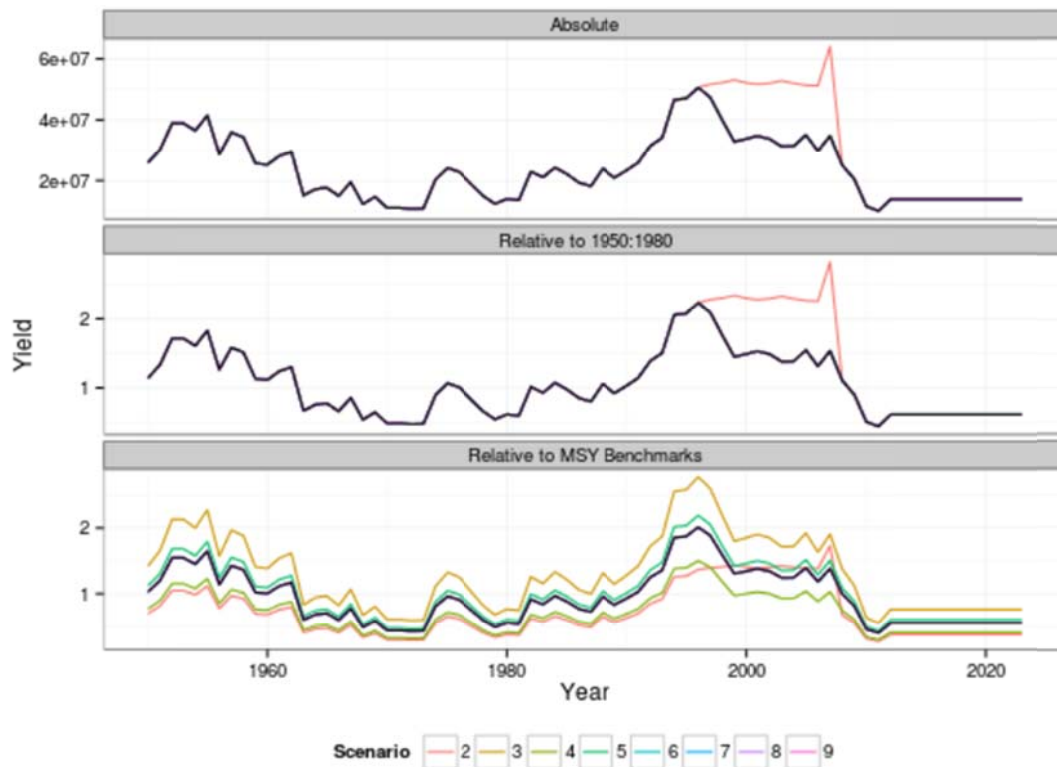


Figure 2. Yield in 9 scenarios from 1950-2023. Scenario 1 (the base case) is shown as the heavy black line.

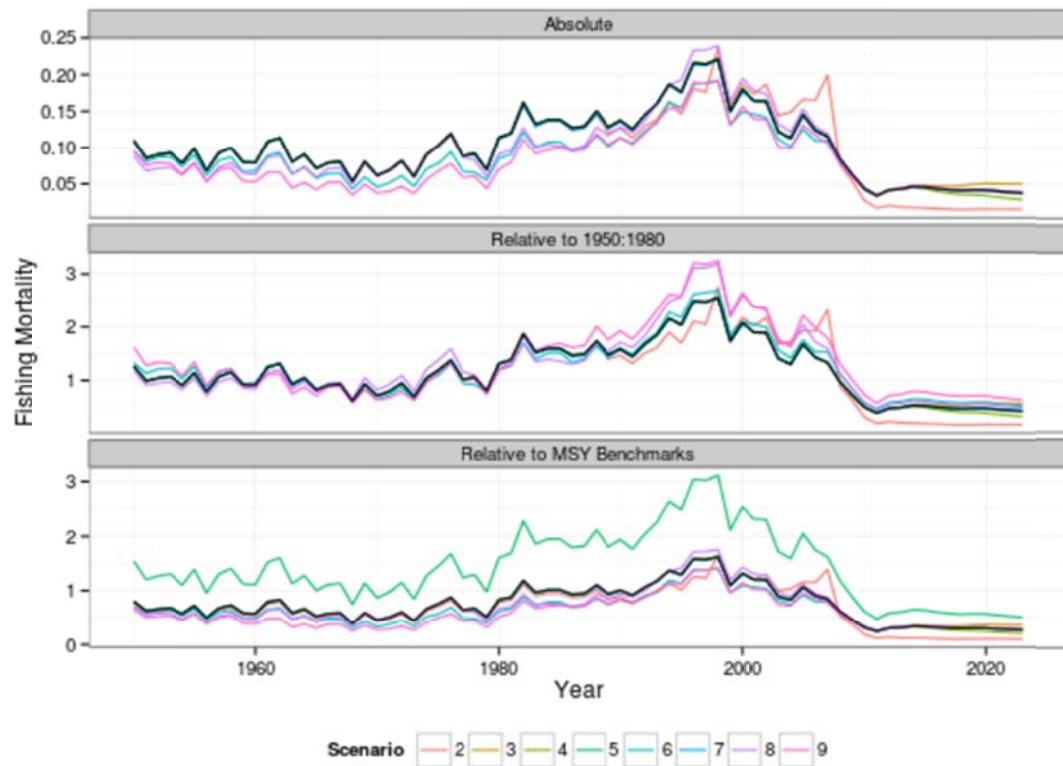


Figure 3. Fishing mortality in 9 scenarios from 1950-2023. Scenario 1 (the base case) is shown as the heavy black line.

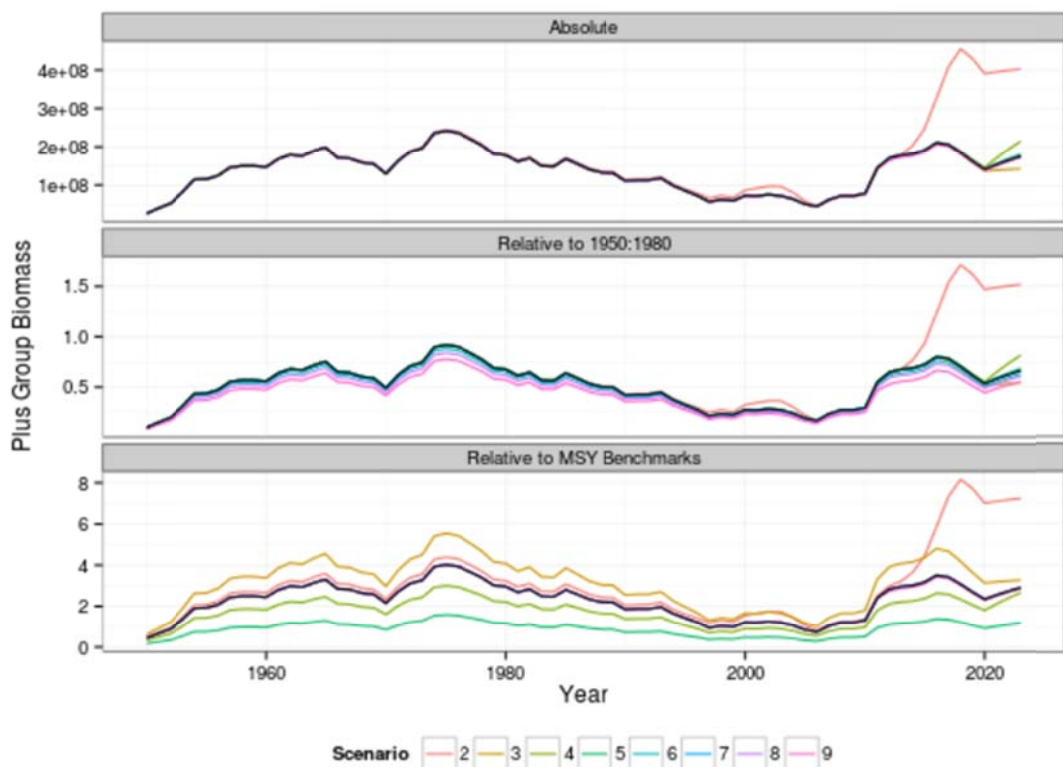


Figure 4. Plus Group Biomass in 9 scenarios from 1950-2023. Scenario 1 (the base case) is shown as the heavy black line.

Possible modelling approaches to variable short-listed as important by group elicitation

| No Variable | | MSE | | | | | | Possible Modelling Approaches | |
|-------------|--|---------------------|------------|----------------|-------|-------|---------------|-------------------------------|--|
| | | Observation Process | Estimation | Implementation | Model | Value | Translational | Institutional | |
| 1 | Catch under-reporting in particular of juvenile catch in artisanal fisheries | ✓ | | ✓ | ✓ | ✓ | | | Include an extra term mortality of juveniles I.e. $R(a,y)$. Since there is no catch quota for juveniles then mortality is a function of effort of |
| 2 | Variability in migration patterns | ✓ | | | ✓ | | | | In the 2 population stock model, allow trends in catchability, i.e. $q(i,y)$ |
| 3 | Risk attitudes of managers | | | | | | ✓ | ✓ | Issue for Utility Function |
| 4 | Social impacts of regulations and its affect on small local communities | | | | | | | ✓ | Issue for Utility Function |
| 5 | Environmentally driven recruitment variability and density dependence | ✓ | | | ✓ | | | | Regime shift and steepness |
| 6 | Natural mortality at age, variability, age related senescence | | | | ✓ | ✓ | | | Shape of $M(a,y)$ vector |
| 7 | Steepness (meta-analysis) | | | | ✓ | ✓ | | | Change the value, also assumes that recruitment is a function of SSB |
| 8 | Standardisation across gear, countries, areas and time | ✓ | ✓ | ✓ | | | | | How well is CPUE a proxy for stock abundance |
| 9 | Generation of age data, age-length keys, slicing | ✓ | ✓ | | ✓ | ✓ | | | How well is size mapped to age? |
| 10 | Growth | ✓ | | | ✓ | ✓ | | | Alternative growth models |

| | | | | | | | | | |
|----|--|---|--|---|---|---|--|--|--|
| 11 | Migration between ICCAT agreed stock units | | | ✓ | ✓ | | | | This is mainly an issue for the Western stock and so ignore here |
| 12 | Maturation and fecundity | | | | ✓ | ✓ | | | Alternative models for Spawner Recruit Potential, i.e. egg production v SSB |
| 13 | Changes in regulations translating into changes of fishing practices | | | ✓ | | | | | Ignore |
| 14 | Environmental factors that influence migration patterns | ✓ | | | | | | | As 2, In the 2 population stock model, allow trends in catchability, i.e. $q(i,y)$ |
| 15 | Complexity of tuna habitat | | | | ✓ | | | | Cannot include |
| 16 | Ecological/environmental (other than climate change) potential to change population dynamics | ✓ | | | | | | | As 2 but assume stationarity |
| 17 | Climate change and/or increased variability's potential to change population dynamics | ✓ | | | ✓ | | | | Non-stationarity, Regime shift |
| 18 | Spawning, periodicity, aggregation and location of spawning areas | ✓ | | | | | | | CV on $N(0,y)$ and closed areas |
| 19 | Plus Group dynamics, skipped- spawning, density dependence | | | | ✓ | | | | M, fecundity, growth, ... |
| 20 | Impacts of regulations and its effect on the species' apparent global distribution. | | | ✓ | | | | | Cannot include |