REPORT OF THE 2014 MEETING OF THE ICCAT WORKING GROUP ON STOCK ASSESSMENT METHODS (WGSAM)

(Dublin, Ireland – April 7 to 11, 20143)

1. Opening, adoption of Agenda and meeting arrangements

The Meeting was held at the Irish Sea Fisheries Board (BIM) office in Dublin, Ireland from April 7- 11, 2014. Local arrangements were made by Dr. Michael Keatinge. Dr. Pilar Pallarés, on behalf of the ICCAT Executive Secretary, thanked BIM for hosting the meeting and providing all logistical arrangements.

Dr. Michael Schirripa, the Stock Assessment Methods Working Group Rapporteur, chaired the meeting. Dr. Schirripa welcomed meeting participants ("the Group") and proceeded to review the Agenda which was adopted without changes (**Appendix 1**).

The List of Participants is attached as **Appendix 2**.

The List of Documents presented at the meeting is attached as **Appendix 3**.

The following participants served as Rapporteurs for various sections of the report:

Section	*Rapporteurs*
1,10	P. Pallarés
2	L. Kell
3	C. Porch
4	H. Arrizabalaga and G. Melvin
5	C. Brown
6	G. Díaz
7	M. Schirripa and V. Ortiz de Zárate
8-9	N. Abid

2. Review of current ICCAT method for estimating Effort Distribution (EFFDIS)

SCRS/2014/026 presented a brief review of previous methods of estimation of the total longline effort as well as suggestions on improving this estimation as discussed by the 2013 WGSAM. The most recent calculation of EffDIS was conducted in 2009 and used the nine major ICCAT tuna and tuna like species to obtain Task-1 global nominal catches (in weight) and CPUE's from partial catch and effort (Task-2) statistics. The current model basic assumption considers that catch rates are equivalent at partial and global level. Comparing the results with previous estimates (obtained during inter-sessional meetings on Ecosystems in 2007 and 2008) the global results do not show major differences. However, at more disaggregated levels the differences are larger for certain flags. In the majority of the cases, large relative variations are usually associated with various improvements and corrections made to some specific datasets. Global geographical distribution shows a small increase near the Venezuela waters due to various corrections made to spatial distributions of catch and effort statistics from Venezuela in several years. Recommended areas for improvement of EffDIS estimation are presented in order to stimulate discussion and finalise the methods that should be used in the future.

It the importance was also recognized of accounting for changes in fishing operations and characteristics of the main fleets from each CPC operating within the ICCAT Convention area as these affect the efficiencies of the fleets for catching target and by-catch species. Documentation of these technological and behavioral changes is particularly important to understand the national reports of catch and effort annually submitted (Task II-CE). SCECO had also made a number of recommendations related e.g. to the i) potential limitations and or restrictions of the data and information ii) characterized of uncertainty e.g. substitutions, raising ratios and proportion of unclassified fleets ("other") and that iii) additional methods of raising data should be explored and iv) methods such as cross-validation should be explored.

WGSAM discussed the use of EffDIS and issues related to its estimation. In particular is the current spatial and temporal level of aggregation (i.e. 5 degree squares and month) adequate for the intended use? this could be evaluated by conducting studies at different levels of aggregation and comparing the conclusions obtained with the current data set. Also the classification of fleets was important and it was also thought that external support would be required to help produce the produce EffDIS in the future.

EffDIS is an important resource for the SCECO (Subcommitte on Ecosystems) and in 2013 SCECO (ref) recommended that efforts should be made to develop similar EffDIS estimates for the BB and PS gears. This would be particularly useful for evaluations of management measures of time-area closures. It will be important to identify what measure of overall effort to use for these two gears.

Despite the importance of EffDIS for the work of the SCRS the last update was conducted in 2007. This is mainly due to several methodological issues that still need to be resolved (see SCRS\2013\36 for a full summary) and the lack of resources to conduct the analysis. In particular there is a need to validate the assumption used to construct the data base which is aggregated to a course level (i.e. 5 degree squares) and evaluate it fitness for purpose. Therefore the group recommended that

1) Methods for classifying fleets be explored; and that
2) External support is required to produce EffDIS in the future.

3. Quantification of uncertainty in ICCAT assessments

The Commission expects risk-based advice on management measures as prescribed in the Kobe II Strategy Matrix and as embedded in its Decision Framework (Rec. 11-13). An important aspect of providing such scientific advice is adequate quantification of uncertainty in stock condition and future prospects under future management option scenarios. With the advent of more commonly applied, highly parameterized stock assessment models, the computational investment in quantifying uncertainty in stock status and future prospects is quite heavy. This is also the experience at other tRFMOs and a number of approximations for quantifying both process and observational uncertainty are being applied to develop risk-based management advice. The WGSAM was therefore asked to provide guidance on the evolution of and possibility of harmonizing methods to apply for uncertainty characterization across species groups. In this regard the WGSAM noted that all assessment models represent simplifications of a more complex fishery system and therefore are intrinsically limited in their ability to reflect the inherent variability of that system. In other words, managers should not expect that the uncertainty in fisheries scientific advice will ever be completely quantified. Nevertheless, the group agreed that improvements can be made in the way uncertainty is represented in the various species groups.

The group identified three basic approaches to characterizing uncertainty in the outputs of stock assessment models; (1) a “model-based” approach that explicitly accounts for the perceived major sources of uncertainty using a single, versatile modelling platform; (2) model averaging, i.e., combining alternative runs from multiple model platforms, and (3) an “empirical” approach that uses existing historical information on the consistency of stock assessment outputs through time. The first approach has been the most common in practice. Typically a single base model is developed that includes estimable parameters representing what are perceived to be the key sources of uncertainty (either with or without informative priors). Measures of the uncertainty in key management parameters are then calculated through standard methods (inverse-Hessian, bootstrapping, Bayesian integration) and incorporated into the Kobe matrices. Sensitivity runs or alternative model platforms may be run to gain further insights into potential uncertainties, but are often used primarily to qualify the advice from the base model and not incorporated directly in the Kobe matrices. The efficacy of this approach hinges on the ability to modify the assessment model in a way that reasonably approximates the source of uncertainty.

A second strategy for incorporating the information provided by alternative models that has sometimes been used is model-averaging, where frequency weights are assigned to each candidate model, perhaps based on some measure of the fit to the data (e.g., AIC, inverse variance weighting) or expert opinion (see also section 5). A disadvantage is that this approach will likely result in a multimodal distribution of the measures of stock status being considered (e.g., limit reference point) or else require assumptions about its distributional form (e.g., normal or lognormal). The group considered that the success of either model-based approach strongly depends on the ability of the assessment group to develop informative priors or postulate reasonable alternative states of nature before discovering the management implications of those states. Otherwise, there is a danger of producing biased estimates of both central tendency and uncertainty by culling models considered to be uninformative or intentionally introducing bias by adding models or altering priors. It was mentioned that the U.S. National Hurricane Center uses model averaging routinely and derives the weights for each candidate model from their performance in past forecasts (i.e., by comparing the model predictions with actual hurricane tracks). Unfortunately, unlike the paths of hurricanes, the true state of a fish stock is rarely known.

Empirical approaches examine past model performance to infer the envelope of total uncertainty. Ideally, one would compare the performance of the model with the true disposition of the stock, but again this is not possible in practice. Alternatively one may compare the performance of two or more historical benchmark stock assessments over each year they have in common. The resulting envelope of total uncertainty would implicitly include "within-model" estimation errors as well as systemic errors as might occur with changes in the models or philosophies of the assessment team. This approach has already been adopted by the Pacific Fishery Management Council in the United States (see Ralston et al. 2011) and is being considered by other councils. Some potential problems with empirical method were raised during the working group discussions. Concern was expressed that such an approach would not reflect the tendency of stock assessments to improve in accuracy and precision with time through the addition of new data, better models and improved data collection. However, it was pointed out by others in the working group that data improvements may only account for some of the uncertainty and that the composition of scientists on the stock assessment team, the composition of the assessment review panel, and changes in stock assessment methods could be equally as influential on the consistency of stock assessment outputs. It was suggested that empirical analyses could at least be used to help ground truth the variance estimates derived from current model-based procedures.

A presentation was made entitled "Quantifying uncertainty due to data processing in age-structured stock assessments". Preliminary work was presented detailing a method for imputing missing size data from observations based on their spatio-temporal proximity (assuming observations were approximately multivariate normal with estimated covariances). Estimated distributions were then used to generate many total size datasets (sampling imputation model parameters and then sampling from the weighted size observations), which were converted to age distributions using several aging methods (e.g. cohort slicing and age-length keys). Each set of age data was then used in a VPA with different natural mortality rates and relative abundance indices. Preliminary results suggest that size imputation with effective sample size of 100 is too low to generate any variability in the derived catch-at-size datasets. Natural mortality rate (80% 100% and 120% of the levels of the most recent assessment) strongly determines estimated FMSY from XSA analysis.

The group agreed that variations on this approach would be a useful way to attempt to quantify how uncertainty in the data will propagate into uncertainty in the assessment. Such an approach would also help identify the relative importance of a suspected source of uncertainty and the corresponding need to incorporate it explicitly in developing the Kobe matrices used for scientific advice to managers. It was also pointed out during the discussion of management strategy evaluations that operating models should be constructed with the major perceived sources of uncertainty in mind and that the same could be used to help quantify the extent to which these uncertainties propagate into the uncertainty in scientific advice (which is in fact a variation of the approach discussed above). It was also pointed out that risk analyses such as that conducted for bluefin tuna (Leach et al. 2014) are useful for identifying the major sources of uncertainty perceived by scientists and other stakeholders, which may have the added benefit of affording broader acceptance when the results are presented.

4. Characterization of quality of the fisheries data and biological information

Document SCRS/2014/035 was presented with the aim of revising background information and promoting discussion under this agenda item. Resolution [13-15] for standardization of scientific information in the SCRS Annual Report requires "… to score the quality of the fisheries data and related knowledge of the species (e.g. biological parameters, fishery distribution patterns historical data, selectivity) used as inputs to stock assessments. Qualitative scores on input data and assumptions may be detailed and should summarize the state of knowledge of the different inputs…". The 2013 Report of the Subcommittee on Statistics proposed a method to score different data elements. The method consisted of a table with a list of categories and items (within a category) that were evaluated and weighted (e.g. according to the relative importance of the items within each category, in the stock assessment). The score for each item was the product of its value and its weight, and the score for each category was the sum of the different items scores within a category. The global quality of the information for a given stock assessment was suggested to be calculated as the average score across categories. The proposal lacked, however, predefined criteria to base the scoring, such as those used in IOTC Species Data Catalogues, where the bounds of the scores are determined by qualitative statements about availability/quality of the data (e.g. "available according to standards", "not available according to standards", or "not available at all"). Thus, SCRS/2014/035 proposed a generic framework, somewhat similar to the proposal by the SC Statistics, but with predefined, qualitative scoring guidelines. The proposed quality score (Q) of the different information sets was:

$$Q=D*I$$

Where:

D=availability and quality of Data.

With

- 1= no data or bad/unknown quality
- 2= if data exists, but quality is not very good
- 3= if good quality data exists

I=Impact of the data on assessment outcomes (how important is to have that information set).

With:

- 1= low impact on assessment results
- 2= moderate impact on assessment results
- 3= high/unknown impact on assessment results

The document included several examples where the scoring was tentatively applied to different albacore and skipjack stocks. Different items related to fisheries and biology were scored and the total score for each stock reflecting the sum of Q values in the different items.

The group found this generic framework useful as a starting point and discussed several issues around it. There was agreement that scores should be based upon objective, quantifiable, and scientifically defensible methodology. While this can be easier to achieve when it comes to scoring data availability, it was recognized that objectively measuring data quality might be slightly more difficult. However, the group suggested mechanisms to make progress on scoring to the degree possible.

At this stage, one important consideration was the purpose or usefulness of this exercise. On one hand, there is a need to respond to Resolution [13-15] so that the Commission has, besides the Kobe matrix, information about the quality of the data used to derive the Kobe matrix. For this purpose, some simple representation (similar to the examples provided in SCRS/2014/035) were considered. However, the group noted that the quality scores could also be of use for other purposes. For instance, they could be useful to the scientific working groups to characterize the main sources of uncertainty and to improve, to the degree possible, the provision of scientific advice (e.g. by selecting appropriate models and/or weighting different scenarios).

The group noted that the importance of different datasets ("I" scores) depend on the model used as well as on the life history of the stock. Thus, total "Q" scores might not be comparable across stocks. To overcome this, the group recommended to quantify the ratio between the total score and the maximum potential score given the specific I vector for each stock, which would allow for some colour based comparison across stocks.

Finally, the group recommended developing a metadatabase that contains information on the quality and quantity of biological and fishery data available for the assessments. This, together with the existing mechanism to evaluate CPUEs and fishery independent abundance indices, would help to more objectively score the quality and quantity of fishery related information used in the assessment of ICCAT species to inform stock assessment scientists and to provide a basis for an overall view of data quality to the Commission i.e. the value of the "D" scores. The group discussed the potential structure of a metadatabase, that can then be populated and updated as new information becomes available.

The metadatabase would have three main components, one dealing with fishery data, another one dealing with biological data and a third one dealing with mark recapture studies (**Appendices 4** and **5**, respectively). The suggested structure is flexible and could incorporate new fields as the need is envisaged. In the case of the biological metadata, it was considered that the fields under "General information" plus some information about what parameters and how well they are estimated would already be useful. The fields under "Parameter estimates" would allow metaanalyses to be conducted that could help in several ways, e.g. characterizing uncertainty on given parameters. However, feeding these fields would of course be labor intensive. The group also felt that, in particular cases, having access to the raw data of some studies might be of use. The metadatabase would allow to easily identify the sources of those raw data.

5. Reconcile the results when dealing with Multiple Modeling Methods

SCRS assessments have often, and increasingly, included the use of multiple modelling methods (*i.e.* different model types, alternate hypotheses) to estimate the status of the stock relative to ICCAT conservation benchmarks. The Group discussed the various reasons why this is done. For instance, model types may differ in underlying assumptions, and the assessment participants may be unable to determine which is more appropriate. In such cases, incorporating the results of multiple models into the management advice is intended to better reflect the uncertainty in the results. In other cases, alternative configurations of the same model platform or type may be regarded as sufficiently plausible to include in the results. The Group agreed that there is merit to this approach, but recognized that there may be difficulties in determining how to most appropriately combine such results, including assigning suitable weighting, especially in those instances when the various model results may be conflict.

No new SCRS papers were available on this topic. However, the Group reviewed the recent work by Deroba et al. (2014) employing simulation to test the robustness of stock assessment models to error. The conclusions of this paper that are relevant to this topic include:

- The biggest differences occurred when comparing results from different types of models (surplus production model, age-based, etc.); models tended to perform similarly to other models of the same type
- Self-testing is useful and recommended, and self- and cross-testing frequently highlighted divergence in the most recent years of the time series
- Among model variability can be considered as a type of uncertainty in the assessment; this has implications when considering whether to base management advice on only one model configured to be the “best fit” for the assessment, or to whether to incorporate the results of different models (for example, the application of model averaging).

The conclusions of Deroba et al. (2014) were consistent with experiences of SCRS stock assessments, and provide an analytical confirmation of the importance of model selection (especially with respect to the type of model), examination of diagnostics and validation of results, and the implications of being more or less inclusive of models for the final advice.

The Group considered that an essential initial component in reconciling different model results is the selection process of the most appropriate models from which to develop management advice. This process should begin during the data preparatory meeting, where time should be set aside to determine which models will be used in the assessment. As described previously, the various model data requirements, assumptions and capability to take into account important changes in the fishery known to have occurred should be considered in light of the quality and quantity of available data. The involvement of CPC scientists is key to this process, as they are most familiar with changes in the various fisheries that might warrant consideration in the assessment process, and are well suited to identify any concerns with the appropriateness of particular data sets for use in the various models. The performance of the various models employed in the previous assessment should also be considered.

During the actual assessment meeting, sufficient time should allotted for the careful review of model diagnostics and performance. Performance measures could include an evaluation of whether results appear consistent with what is known of the species biology as well as fishery trends (an example of one such approach, the “retrospective prediction performance measure”, is detailed below). The likelihood and consequences of violation of assumptions for each model should be considered in light of available information on the population and the fishery. One potential consequence of this process is that model(s) identified during the data preparatory meeting for use in the assessment may be (appropriately) rejected for inclusion in the final management advice.

The Group discussed whether or not it would be appropriate to consider results from simpler models together with those from more complex models (*e.g.* using results from a surplus production model as well as those from a more fully integrated model, such as stock synthesis). In other words, if the data were sufficient to run the more complex models (which, in theory, could take into account more aspects and changes in the stock and fishery parameters), should the advice be based instead solely on the results from those complex models? The Group considered that there may be circumstances where it is appropriate to consider results across such divergent model types.

The Group also discussed options for additional criteria or procedures for selecting models. One such procedure, based on “retrospective prediction performance measure”, was suggested. The idea of this method is similar with the retrospective analysis traditionally used in the VPAs but is extended to the comparison of multiple models. For example, assume that two models are used in the assessment and provide two different results. The steps of the retrospective prediction performance measure process would be:

- truncate the available data, dropping the most recent XX years (as appropriate for the stock),

- rerun each particular model to get an estimate of the population dynamics XX years in the past
- project forward from that model using actual catches to the current year
- calculate a CPUE trend (or perhaps just the estimated biomass trend, as appropriate to the complexity of the model) that is predicted for those projected years to the current
- compare that CPUE trend to the actual observed (standardized) CPUE trend(s).

With this approach, models could be eliminated from consideration if the results of this analysis demonstrates inconsistency with the observed (standardize) CPUE trend(s). The method can be further extended to develop weights for the various model results (to be used when combining results across models) by defining a measure of discrepancy between the predicted and observed CPUE series (e.g. sum of squared residuals divided by SEs of observed CPUE). It is noted that this approach has no requirement that models utilize the same datasets. The applicability of this approach is dependent upon the actual observed (standardized) CPUE tend(s) adequately reflecting the population and the absence of conflicting CPUE indices. The Group found this approach to be promising, and recommended that further research be conducted regarding its suitability for use in model selection and/or weighting.

The group noted that the selection of which models to use for projections is also important. In some circumstances, a different model than the one used to estimate current stock status may be used to conduct projections. This can occur, for example, if the model used from estimating stock status cannot be used for projections, or requires additional processing time (beyond what can be accomplished during the assessment meeting) to produce certain output elements that would be required for projection. The Group noted that there has been some concern regarding the appropriateness of this approach. The Group identified the need to develop approaches to investigate this. One potential approach suggested would involve rerunning the model used to estimate current stock status, dropping some number of recent years of data (as in a retrospective analysis). The model being considered for the projections could then be run, projecting from the last year of the retrospective model run, using the actual catch history up to the current year. The performance of the projection model could be evaluated through comparison of the projected stock status/condition in the current year to that being estimated by the assessment model.

Once the assessment scientists have agreed on the models to be included for developing the management advice (e.g. estimates of stock status relative to biological reference points and associated uncertainty, Kobe strategy matrix), there remains the issue of how to present these results. A generally preferred option would be to somehow combine these results to provide benchmarks and appropriate representations of uncertainty. This is particularly difficult when the various results are conflicting among the models. The Group expressed a strong caution that it is generally inappropriate to combine results which are entirely inconsistent between each model; for example, when the different models are structured based on two alternative hypotheses reflecting very different assumed states, with little or no overlap in probability distributions. In such cases, combining such results equally would produce estimates of stock status and/or trends averaged between the trends, which is not consistent with either alternative hypothesis. In such instances, scientists are encouraged to rigorously examine the underlying assumptions of each hypothesis and consider carefully if each is supported by existing data and current science. If this process is insufficient to reject one hypothesis, scientists should consider if relative probabilities of each hypothesis can be assigned. Further research to collect data to assist in accepting/rejecting hypotheses or to assign likelihoods may be advisable.

The usual practice of SCRS when combining results from multiple models is to give equal weighting to the results from each model. As mentioned in Section 3 of this report, weights could be assigned to results from each model based on some measure of fit to the data (*e.g.* AIC, inverse variance weighting). In addition to the concerns described in Section 3, this approach is limited in that it is appropriate only when the candidate models are using the same datasets.

The Group also discussed the potential to use MSE to inform model selection/weighting. The Group identified a potential problem with this approach in that the operational model for the MSE may influence perception of model performance. It was suggested that the use of surplus production models may be problematic if that the productivity functions are strongly skewed to the left and productivity soars as stock levels are depleted.

The Group acknowledged that an element of subjectivity is inevitable in any weighting scheme. Even the default assignment of equal weighting is implicitly a decision that may give more weight to some results than might be warranted if the totality of factors such as model assumptions, capability, data quality, fishery trends and biology could be accounted for in an objective manner. Therefore, the Group considers that expert opinion may be appropriate for the assignment of weights, if supported by knowledge of these factors and the postulation of reasonable mechanisms that would support a conclusion that certain model(s) are more likely to reflect the true situation than others.

Another common practice in SCRS assessments is to ultimately reject from inclusion in the management advice various models that produce plausible results, but may be considered somewhat less likely. The Group noted that there have been concerns that the existing practice of selecting only the subset of models deemed most likely into the final results may result in an underestimation of the range of uncertainty. Further research is required.

6. Limit Reference Points, Harvest Control Rules, and Management Strategy Evaluations

The evaluation of Limit Reference Points (LRP) and Harvest Control Rules (HCR) through the use of Management Strategy Evaluation (MSE) is increasingly being recognized by global tuna RFMOs as an effective means to advance their fishery management process. The 2013 assessments of albacore and swordfish were used as examples of how an MSE process could possibly be formally included in the management of those stocks. The WGSAM plans to continue this effort by (1) continuing to refine the methods within the MSE process, (2) introduce MSE more assessments when and where appropriate, and (3) foster lines of communication that keep managers informed of their benefits and weaknesses. Regarding dialog and communication, ICCAT has recently adopted the [Rec. 13-18] Recommendation by ICCAT for Enhancing the Dialogue between Fisheries Scientists and Managers that aims to enhance communication and foster mutual understanding between fisheries managers and scientists in order to facilitate more streamlined, science-based decision-making. As well, the Recommendation outlines specific tasks to be achieved during the first meeting of the Standing Working Group for Enhancing the Dialogue between Fishery Scientists and Managers (SWGSM) in 2014.

An empirical Harvest Control Rule has been adopted for southern bluefin tuna (SBT) to set Total Allowable Catches (TACs). The HCR is based on year-to-year changes in the indices from a fisheries dependent CPUE index of adult abundance and a fisheries independent aerial survey of juveniles. Before the HCR can be implemented appropriate reference levels, for catch and the indices must be selected and HCR parameters tuned to meet management objectives using management strategy evaluation (MSE).

In SCRS/2014/036 a preliminary MSE was conducted for Mediterranean bluefin using the SBT HCR as part of a Management Procedure (MP, i.e. the combination of pre-defined data, together with an algorithm to which the data are input to provide a value for a TAC or effort control measure). Next steps will be to:

1) Identify management objectives and map these to performance measures in order to quantify how well they have been achieved.
2) Select hypotheses about system dynamics.
3) Condition Operating Models (OMs) on data and knowledge and possible rejecting and weighting of the different hypotheses.
4) Identify candidate management strategies and code these up as MPs.
5) Project the OMs forward using the MPs as feedback control procedures; and
6) Agree the MPs that best meet management objectives.

The Group emphasized the need to carefully select the assumptions of the operating models as wrong assumptions can affect both stock assessment and Management Strategy Evaluations (MSE) results. It was discussed that the selection of the range of scenarios used in the operating model is important (Fromentin et al., 214; Leach et al., 2014). The Group agreed that in a MSE framework, the goal is not to choose the management procedure or harvest control rules (HCR) that gives the best performance, but the one that is most robust across all chosen scenarios. For example, if the operating model is run with 4 level of steepness, which are all assumed equally plausible the chosen HCR should be that which reaches the management goals under those 4 levels of steepness. The Group also discussed that many operating models assume that population regulation only occurs at recruitment; however, other possibilities of population regulation should also be capture by the operational model (e.g., natural mortality).

It was agreed by the Group that the operating model should be more complex than the estimation model. For example, a surplus production model can be used as the estimating model while a more complex model, like SS3, can be used as the operating model. The document presented an example where 2 surveys were used as indices for recruits and adults. The possibility of not having to carry out these types of surveys every year to obtain indices of biomass that inform the HCR was discussed. It was indicated to the Group that using HCR might reduce the data requirements compared with what currently is used in some stock assessments. The Group also discussed if the HCR should include more precautionary limit reference points for those cases when the biomass index estimated from the surveys have higher uncertainty.

The presentation that the SCRS Chair made to the Commission at its 2013 meeting was reviewed. The Group emphasized the importance of the upcoming first meeting of the ICCAT 'Standing Working Group to Enhance Dialogue between Fishery Scientists and Managers' where issues related to Limit Reference Points, probabilities associated to the Kobe matrix, rebuilding periods for different stock, etc., will be discussed. The Group emphasized that without this dialogue between scientists and manager, HCR cannot be developed.

Document SCRS/2014/025 describes how the authors search for evidence of a stock recruitment relationships for bluenfin, yellowfin and albacore tuna. The authors concluded that evidence of the existence of a SRR for any of the stock was weak and the data could also support other hypotheses (Vert-pre, 2014) e.g. that of Gilbert (1998) that recruitment fluctuates around a mean level for a period of time and then a regime shift occurs. This has obvious important implications for stock assessment and management advice.

The Group discussed that the 'regime shift' hypothesis presented in the paper could be an artifact of the estimating models. In other words, although the 'regime shift' is not inconsistent with the lack of a SRR observed in the mentioned stocks, that *per se* is not proof that a 'regime shift' actually occurred.

The paper suggested that SSB is a function of recruitment (instead of recruitment being a function of SSB). The Group inquired why there seemed to be autocorrelation in the recruitment patterns. It was discussed that the observed autocorrelation might be due to using cohort slicing to age the studied stocks or due to other model assumptions.

The document suggested that for the purpose of stock status projections, more weight should be given to recent observed recruitments instead of weighting all the recruitment patterns in the time series the same (in other words, tomorrow's recruitment is most probably to be more similar to yesterday's than recruitment 25 years ago).

In document SCRS/2014/037 the SSB and R series resulting from the case base scenarios for bluefin tuna Eastern Atlantic and Medterranean stock using "Inflated" and "Reported" catches, were used to fit three different S-R models (Beverton & Holt, Ricker and smooth Hockey-stick). The results show that the stock has maintained its full reproductive capacity throughout the time series from the 1950s and that a Ricker S-R relationship cannot be rejected for this stock. The document argues that this last result adds more uncertainty to the estimate of B_{01}, which makes it useless as a biomass reference point for this stock. As an alternative, the document proposes the use of B_{loss} as B_{lim} and then estimating B_{pa} from this value. This allowed the authors to simulate the behavior of some HCR and select the most suitable for this stock

The Group discussed that ICES is moving away from the limit reference point framework, which is opposite to what is proposed in the paper. Regardless, the utility of the approach proposed in the document can be tested through simulation. also discussed the validity of making assumptions in the operation models to be more precautionary (like using a hockey-stick recruitment relationship). In general, the Group disagreed with this approach. The Group also discussed that factors that might contribute to a Ricker type of S-R relationships in tunas without reaching a final conclusion.

7. Incorporation of Ecosystem, Climate, and Habitat (ECH) information into stock assessments

Discussion on the incorporation of ecosystem, climate and habitat (ECH) information in to stock assessments centered around two approaches that can be used to enhance the stock assessment process of ICCAT species. Two approaches were identified; one qualitative and the other the quantitative with the latter aimed to reduce the uncertainty in the stock assessment process.

A single presentation entitled "A hypothesis of a recent poleward shift in the distribution of North Atlantic swordfish" was made on this agenda item based on the SRCS\2013\161 report. The report describes changes in swordfish abundance indices relative to indicators of broad scale environmental process (eg., AMO, NAO, AWP). Observations of opposing trends in abundance for northern Swordfish suggested the possibility of a shift in abundance from warm, southern latitudes to cooler, more northern latitudes. Several of the observed indices of abundance changed sharply in direction in from negative to positive, while others showed an opposite change.

The observed changes in the direction of the abundance indices correspond with changes in trends in the size of the Atlantic Warm Pool (AWP), the change in sign of the Atlantic Multidecadal Oscillation (AMO), and the North Atlantic Oscillation (NAO). To quantify a possible relation between the changes in abundance and the various candidate environmental indices, we ran the assessment model without the influence of the environmental data and regressed the residuals of the fit to the CPUEs to the various environmental indices. Given the suspected temperature tolerance limits of Swordfish, it is possible that either their preferred habitat has moved north, a preferred prey species, or both.

With these results as background a discussion was held on the advantages of aggregating the different CPUE´s by area instead of the current aggregation by fleet available to the SWO WG. This would require access to the fleet set-by-set data and allow the incorporation of area and fleet effect as well as spatial structure into the assessment model (i.e. CPUE´s east versus CPUE´s west of Atlantic). It was further highlighted that there are other approaches to incorporate the area as an interaction factor in the modeling of the CPUE´s by using the mixed models. Weighting approaches could be calculated once the CPUE´s are standardized by area.

The WG also introduced the use of habitat modeling to incorporate environmental driven variability by determining spatial and temporal distribution, in the modeling of CPUE´s.

There was an extended discussion on how to deal with the aggregation of set by set data for longline fisheries targeting swordfish as a continuation of the study presented. A major area of discussion was on how to establish an appropriate collaboration framework among national scientists to obtain set by set data and how the data could be jointly analyzed. This exchange could be made under the ICCAT Confidentiality Agreement already in place. It was recognized that this would be an ongoing process, likely to involve future support of the Secretariat, working in the "cloud" with restriction access to the scientist involved in the collaborative analysis of data. These data are necessary to address a number of concerns/interactions associated with the fishery and the indices for improvement of the assessment and incorporation of environmental factors.

8. Other matters

8.1 Collaboration with ICES SISAM

A presentation was made to the group by one of the co-chairs of the ICES SISAM (Strategic Initiative on Stock assessment Methods). This outlined the work conducted under this initiative to date, and a discussion was held on the future direction of SISAM and its relevance to the ICCAT scientific community. ICCAT scientists participated in SISAM work in 2013 through the World Congress on Stock assessment Methods (WCSAM). SISAM developments in 2014 include completing papers for a special edition of the ICES Journal of Marine Science (based on the WCSAM), and workshops by partner organisations such as CAPAM, as well as an open session which is being held by ICES at its ASC in September 2014. The discussion at the meeting outlined that at least ICES and ICCAT face similar issues in terms of the development of stock assessment methods and the provision of fisheries advice, and that these issues may be common to many RFMO's globally. These issues related to the advancement of methodologies and the challenge to maintain enough technically experienced scientists in the organisations network to run these methods in a considered way. Along with this shared endeavours in developing assessment and advice frameworks dealing with situations where data is limited, was discussed. It was agreed that it would be of positive benefit for ICCAT to maintain involvement with the SISAM initiative and through this to work collaboratively with scientists from other RFMO's to help resolve these issues.

9. Recommendations

- The Group recommended to encourage CPCs to report their Task 2 catch and effort data at a finer geographical stratification (e.g. 1° by 1°) instead of reporting these data at 5° by 5° as in some places this scale might be too coarse. WGSAM further requests that subcommittee STAT consider if reporting should be required at this finer scale resolution.

- The Group also agreed that the implementation of the management strategy evaluation approach (MSE) and promotion of the dialogue between scientists and fisheries managers on the Harvest Control Rules and MSE should be encouraged to improve the scientific advice given to the Commission. These efforts should include a review of MSE efforts so far in light of successes, lack of successes and the resources limiting future MSE progress.

- The Group felt that simple criteria could be used by the different working groups to start scoring the quality of the information used in different stock assessments. Meanwhile, the group recommends to continue developing further detailed, scientifically based, objective ways to provide such scores. This includes developing criteria to evaluate the importance of the different data elements depending on the life history and/or assessment model used. Along that line, the group recommended to continue developing a metadatabase with information on the quantity and quality of available fisheries and biological information.

- The Group encouraged that the retrospective prediction performance measure, as described in Section 5 of this report, be evaluated through simulation studies, potentially with models conditioned on past ICCAT assessments. If possible, a case study should be conducted for a stock not having conflicting observed (standardized) CPUE series.
- The Group again encourages CPCs to provide limited access to CPUE set-by-set data according to the needs and priorities identified by the different species groups and the subcommittees. This would enable SCRS to produce a wide variety of indices on a more informative spatial scale. Initially information on a single species and/or by specific fleet could be used to illustrate the benefits. Use of the existing “cloud” opportunities maintained by the Secretariat for storage and access was suggested for ease of multi-lateral collaborations. This exchange could be made under the ICCAT Confidentiality Agreement already in place.

- The Group also considered that the SCRS should continue to participate in the ICES SISAM initiative in order to further promote collaborative work in developing assessment methodologies, to share and develop knowledge on how to communicate uncertainty to managers, to foster closer collaboration on joint assessments (e.g. porbeagle) as well as practical initiatives such as sharing the agenda for ICCAT and ICES methods working group. This will provide standardised and advertised open access to assessment input and output data, as well as sharing more detailed data. These elements could be progressed through involvement in the Global Assessments Methods working Group (GAME).

10. Adoption of the report and closure

The report was adopted during the meeting. The Convener of WGSAM thanked the local organizers for the excellent meeting arrangements and the participants for their efficiency and hard work. The Secretariat reiterated his thanks to the Irish Sea Fisheries Board (BIM) for the exceptional organization of the meeting and for the warm support provided to participants. The meeting was adjourned.

References

Leach, Adrian W., et al. "Identification and prioritization of uncertainties for management of Eastern Atlantic bluefin tuna (Thunnus thynnus)."Marine Policy 48 (2014): 84-92.

Ralston, S., Punt, A. E., Hamel, O. S., DeVore, J. D., and R. J. Conser. A meta-analytic approach to quantifying scientific uncertainty in stock assessments. Fish. Bull. 109:217–231 (2011)

Fromentin J.-M., S. Bonhommeau, H. Arrizabalaga, and L. L. Kell. (2014). The spectre of uncertainty in management of exploited fish stocks: the illustrative case of atlantic bluefin tuna. Marine Policy, 47: 8–14, 2014.

Gilbert D.. Towards a new recruitment paradigm for fish stocks. Canadian Journal of Fisheries and Aquatic Sciences, 54(4):969–977, 1997.

Leach A., P. Levontin, J. Holt, L. Kell, and J. Mumford. (2014) Identification and prioritization of uncertainties for management of eastern atlantic bluefin tuna (Thunnus thynnus). Marine Policy, 48: 84–92.

Ralston, S., Punt, A. E., Hamel, O. S., DeVore, J. D., and R. J. Conser. A meta-analytic approach to quantifying scientific uncertainty in stock assessments. Fish. Bull. 109:217–231 (2011)

Vert-pre K. A., R. O. Amoroso, O. P. Jensen, and R. Hilborn. Frequency and intensity of productivity regime shifts in marine fish stocks. Proceedings of the National Academy of Sciences, 110(5):1779–1784,2013.

Appendix 1

AGENDA

1. Opening, adoption of agenda and meeting arrangements
2. Review of current ICCAT method for estimating Effort Distribution (EFFDIS)
3. Quantification of uncertainty in ICCAT assessments
4. Characterization of quality of the fisheries data and biological information
5. Reconcile the results when dealing with Multiple Modeling Methods
6. Limit Reference Points, Harvest Control Rules, and Management Strategy Evaluations
7. Incorporation of Ecosystem, Climate, and Habitat (ECH) information into stock assessments
8. Other matters
9. Recommendations
10. Adoption of the report and closure

Appendix 2

LIST OF PARTICIPANTS
SCRS CHAIRMAN

Santiago Burrutxaga, Josu
SCRS Chairman, Head of Tuna Research Area, AZTI-Tecnalia, Txatxarramendi z/g, 48395 Sukarrieta Bizkaia, Spain
Tel: +34 94 6574000 (Ext. 497); 664303631, Fax: +34 94 6572555, E-Mail: jsantiago@azti.es

CONTRACTING PARTIES

CANADA

Melvin, Gary
Biological Station - Fisheries and Oceans Canada, Department of Fisheries and Ocenas531 Brandy Cove Road, St. Andrews, New Brunswick, E5B 2L9, CANADA
Tel: +1 506 529 5874, Fax: +1 506 529 5862, E-Mail: gary.melvin@dfo-mpo.gc.ca

EUROPEAN UNION

Arrizabalaga, Haritz
AZTI - Tecnalia /Itsas Ikerketa Saila, Herrera Kaia Portualde z/g, 20110 Pasaia Gipuzkoa, Spain
Tel: +34 94 657 40 00, Fax: +34 94 300 48 01, E-Mail: harri@azti.es

Ortiz de Zárate Vidal, Victoria
Ministerio de Economía y Competitividad, Instituto Español de Oceanografía, C.O. de Santander, Promontorio de San Martín s/n, 39012 Santander Cantabria, ESPAÑA
Tel: +34 942 291 716, Fax: +34 942 27 50 72, E-Mail: victoria.zarate@st.ieo.es

Keatinge, Michael
BIM (The Irish Seafisheries Board), Crofton Road, Dun Laoghaire, Dublin, Ireland
Tel: +353 1 214 4230, Fax: +353 1 230 0564, E-Mail: keatinge@bim.ie

JAPAN

Kimoto, Ai
Researcher, Bluefin Tuna Group, National Research Institute of Far Seas Fisheries, 5-7-1 Orido Shimizu, Shizuoka
424-8633, JAPAN
Tel: +81 543 36 6036, Fax: +81 543 35 9642, E-Mail: aikimoto@affrc.go.jp

Kitakado, Toshihide
Associate Professor, Faculty of Marine Science, Tokyo University of Marine Science and Technology, Department of Marine Biosciences5-7, Konan 4, Minato-ku, Tokyo, Japan 108-8477, JAPAN
Tel: +81 3 5463 0568, Fax: +81 3 5463 0568, E-Mail: kitakado@kaiyodai.ac.jp

MOROCCO

Abid, Noureddine
Chercheur au Centre Régional de Recherche Halieutique de Tanger, Center Regional de L'INRH á Tanger/M'dig, B.P. 5268, 90000 Drabed Tanger, Morocco
Tel: +212 53932 5134, Fax: +212 53932 5139, E-Mail: abid.n@menara.ma; noureddine.abid65@gmail.com

Ben Mhamed, Abdelouahed
Institut national de recherche halieutique, 2, Rue Tiznit, 20000 Casablanca, MARRUECOS
Tel: Fax: E-Mail:

Serghini, Mansour
Institut national de recherche halieutique, 2, Rue Tiznit, 20000 Casablanca , MARRUECOS
Tel: Fax: E-Mail:

UNITED STATES
Brown, Craig A.
NOAA Fisheries Southeast Fisheries Center, Sustainable Fisheries Division, 75 Virginia Beach Drive, Miami, Florida 33149
Tel: +1 305 361 4590, Fax: +1 305 361 4562, E-Mail: Craig.brown@noaa.gov

Díaz, Guillermo
NOAA-Fisheries, Southeast Fisheries Science Center, 75 Virginia Beach Drive, Miami Florida 33149, ESTADOS
UNIDOS
Tel: +1 305 361 4277, Fax: E-Mail: guillermo.diaz@noaa.gov

Porch, Clarence E.
Chief, Sustainable Fisheries Division, Southeast Fisheries Science Center, National Marine Fisheries Service, 75 Virginia Beach Drive, Miami, Florida 33149
Tel: +1 305 361 4232, Fax: +1 305 361 4219, E-Mail: clay.porch@noaa.gov

Schirripa, Michael
NOAA Fisheries, Southeast Fisheries Science Center75 Virginia Beach Drive, Miami Florida 33149, ESTADOS
Tel: +1 305 361 4568, Fax: +1 305 361 4562, E-Mail: michael.schirripa@noaa.gov

OBSERVERS FROM NON-GOVERNMENTAL ORGANIZATIONS

WWF
Quílez Badia, Gemma
WWF Mediterranean Programme Office, c/ Carrer Canuda, 37 3er,, 08002 Barcelona , ESPAÑA
Tel: +34 93 305 6252, Fax: +34 93 278 8030, E-Mail: gquilez@atw-wwf.org

****

ICCAT SECRETARIAT
c/ Corazón de María, 8 - 6th -7th fl. 28002 Madrid, Spain
Tel: + 34 91 416 5600, Fax: +34 91 415 2612, E-Mail: info@iccat.int

Pallarés, Pilar
Kell, Laurence

Appendix 3

LIST OF DOCUMENTS

SCRS/2014/025 Which Came First? The Chicken, The Egg or The Tortilla? Kell L.T., Fromentin J.M. and Szuwalski C.S.

SCRS/2014/026 Proposals for the improvement of the estimation of the overall longline effort distribution (EffDIS) in the ICCAT area. de Bruyn P., Palma C. and Gallego J.L.

SCRS/2014/035 Characterizing quality of data used in ICCAT assessments. Arrizabalaga H.,Santiago J. , Scott G. and Murua H.

SCRS/2014/036 An Example Management Strategy Evaluation Of A Model Free Harvest Control Rule . Kell L.T., Hillary R., Fromentin J.M. and Bonhommeau S.

SCRS/2014/037 Comment on the Eastern Atlantic and Mediterranean Bluefin tuna. de Cardenas E.

Appendix 4

Suggested fields for the fisheries and mark-recapture metadatabase

Fishery data:

Structure based on Gear type followed by:

CPC (Filter % Catch or number of CPC's)

Data type:

- Size Composition
- Age composition
- Catch
- Catch effort
- Mark-recapture studies

Size Composition, Age composition

- Years of coverage
- Number of samples (Average sample size)
- Sampling Coverage (i.e. percent coverage from samples)
- Representativeness
- Caveats
- Spatial distribution

Catch

- Years of coverage
- Percent of catch
- Landings: Precision – Census, High medium and low, or unknown
- Percent of discard
- Discards: Precision – Census, High medium and low, or unknown
- Caveats
- Spatial distribution

Effort

- Years of coverage
- Unit of measure
- Spatial scale;
- Precision:
- Representativeness
- Spatial distribution

Mark-recapture studies

Fields to be discussed

Appendix 4

Biological Metadabase for individual studies

General Information

Authors
Pulbication year
Citation
Species
Stock
Type of study
Number of samples
Area Covered
Fisheries independent?
Flag
Gear
Size Range
Period of study
Methodology
Comments
Availability of raw data

Parameter estimates table for a field rather than different fields

Growth	Parameter	s.d.	units	Sex			
Reproduction	Parameter	s.d.	units	Sex			

Natural mortality	M	s.d	units	Sex		Age

Stock structure & movements	Nº of subgroups					units	
	Movement rates	From	To	Migration rate	s.d	units	Age
		Where	ProporitonE	ProportionW	s.d	units	Age

List of parameters

Growth	Reproduction
K	A50
linf	L50
t0	A100
variation of size at age	L100
longevity	Spawning season
a	Batch Fecundity
b	Spawning interval
convertion factors	F/M sex ratio
…	…