

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

(Dublin, Ireland – April 7 to 11, 2014)

SUMMARY

The meeting was held in Dublin, Ireland, 7-11 April 2014 with the objective, among others, of providing guidance on the evolution and possibility of harmonizing methods to apply for uncertainty characterization across species groups in order to provide risk-based advice to the Commission, to continue the evaluation of advice using Management Strategy Evaluation (MSE) and to establish guidance on best practices to reconcile or combine assessment results when multiple modelling methods to estimate the status of the stock relative to ICCAT conservation benchmarks are used.

RÉSUMÉ

La réunion a eu lieu à Dublin (Irlande) du 7 au 11 avril 2014 et visait entre autres à fournir une orientation sur l'évolution des méthodes et sur leur éventuelle harmonisation pour caractériser l'incertitude existant au sein des groupes d'espèces dans le but de formuler un avis basé sur les risques à l'intention de la Commission, à poursuivre l'évaluation de l'avis au moyen de l'évaluation des stratégies de gestion (MSE) et à établir des orientations sur les meilleures pratiques permettant de réconcilier ou de combiner les résultats d'évaluation lors de l'utilisation de plusieurs méthodes de modélisation visant à estimer l'état des stocks par rapport aux paramètres de conservation de l'ICCAT.

RESUMEN

La reunión se celebró en Dublín, Irlanda, del 7 al 11 de abril de 2014 entre sus objetivos figuraban proporcionar orientación sobre la evolución y la posibilidad de armonizar métodos para caracterizar la incertidumbre en los diferentes grupos de especies con el fin de facilitar a la Comisión un asesoramiento basado en el riesgo, continuar la evaluación del asesoramiento utilizando una evaluación de estrategias de ordenación (MSE) y establecer orientaciones sobre mejores prácticas para conciliar o combinar los resultados de la evaluación cuando se utilizan múltiples métodos de modelación para estimar el estado del stock en relación con los elementos de referencia de conservación de ICCAT.

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:

<i>Section</i>	<i>Rapporteurs</i>
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 (Sub-committee 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 be 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.*, 2014; 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 bluefin, 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 Mediterranean 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. It was 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 SCRS/2013/161 report. The report describes changes in swordfish abundance indices relative to indicators of broad scale environmental process (e.g., 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.

RAPPORT DE LA RÉUNION DE 2014 DU GROUPE DE TRAVAIL ICCAT SUR LES MÉTHODES D'ÉVALUATION DES STOCKS (WGSAM)

(Dublin, Irlande, 7 - 11 avril 2014)

1. Ouverture, adoption de l'ordre du jour et organisation des sessions

La réunion s'est tenue dans les bureaux de l'Irish Sea Fisheries Board (BIM) à Dublin, en Irlande, du 7 au 11 avril 2014. Le Dr Michael Keatinge s'est occupé de la logistique au niveau local. La Dre Pilar Pallarés, au nom du Secrétaire exécutif de l'ICCAT, a remercié le BIM d'accueillir la réunion et de fournir toute la logistique.

Le Dr Michael Schirripa, rapporteur du groupe de travail sur les méthodes d'évaluation des stocks, a présidé la réunion. Le Dr Schirripa a souhaité la bienvenue aux participants de la réunion (le « groupe ») et a passé en revue l'ordre du jour qui a été adopté sans modification (**Appendice 1**).

La liste des participants est jointe en tant qu'**Appendice 2**.

La liste des documents présentés à la réunion est jointe à l'**Appendice 3**.

Les participants suivants ont assumé la tâche de rapporteur des divers points du rapport :

<i>Point</i>	<i>Rapporteur</i>
1,10	P. Pallarés
2	L. Kell
3	C. Porch
4	H. Arrizabalaga et G. Melvin
5	C. Brown
6	G. Díaz
7	M. Schirripa et V. Ortiz de Zárate
8-9	N. Abid

2. Examen de la méthode actuelle de l'ICCAT visant à estimer la distribution de l'effort (EFFDIS)

Le document SCRS/2014/026 présente brièvement les méthodes utilisées antérieurement d'estimation de l'effort palangrier total ainsi que des suggestions d'amélioration de cette estimation, comme le WGSAM en a débattu en 2013. Le calcul le plus récent de EffDIS a été réalisé en 2009 et utilisait les neuf principales espèces de thonidés et d'espèces apparentées de l'ICCAT pour obtenir les prises nominales globales de la Tâche I (en poids) et les CPUE à partir des statistiques partielles de prise et d'effort (Tâche II). Le postulat fondamental du modèle actuel considère que les taux de capture sont équivalents au niveau partiel et global. Si l'on compare les résultats avec des estimations antérieures (obtenus pendant les réunions intersessions du Sous-comité des écosystèmes en 2007 et 2008), les résultats globaux ne font pas apparaître de grandes différences. Or, à des niveaux plus désagrégés, les différences sont importantes pour certains pavillons. Dans la majorité des cas, de grandes variations relatives sont habituellement associées à diverses améliorations et corrections effectuées à certains jeux de données spécifiques. La distribution géographique globale fait apparaître une faible augmentation près des eaux vénézuéliennes en raison de diverses corrections apportées aux distributions spatiales des statistiques de prise et d'effort du Venezuela au cours de plusieurs années. Les domaines recommandés d'amélioration de l'estimation de EffDIS sont présentés afin de stimuler les discussions et d'achever les méthodes qui devraient être utilisées à l'avenir.

On a reconnu l'importance de tenir compte des changements dans les opérations de pêche et les caractéristiques des principales flottilles de chaque CPC opérant dans la zone de la Convention de l'ICCAT, étant donné que ceux-ci affectent l'efficacité des flottilles pour capturer les espèces cibles et les espèces accessoires. Il est important de documenter ces changements technologiques et stratégiques afin de comprendre les rapports nationaux de prise et d'effort présentés tous les ans (Tâche II-CE). Le Sous-comité des écosystèmes a également formulé plusieurs recommandations concernant par exemple i) les limites potentielles et/ou les restrictions des données et des informations ii) la caractérisation de l'incertitude, par exemple les substitutions, les ratios d'extrapolation et la proportion des flottilles non classifiées (« autres »), iii) le fait que d'autres méthodes d'extrapolation des données devraient être explorées et iv) le fait que les méthodes telles que la validation par recoupement devraient être explorées.

Le WGSAM a examiné la question de l'utilisation de EffDIS et les questions liées à son estimation. Il convient en particulier de se demander si le niveau d'agrégation spatio-temporel utilisé actuellement (à savoir carrés de 5°x5° et mois) est adapté à l'objectif poursuivi. Ce point pourrait être évalué en réalisant des études à différents niveaux d'agrégation et en comparant les conclusions tirées avec le jeu de données actuel. De plus, la classification des flottilles était importante et on a également pensé qu'un support externe pourrait être requis pour contribuer à produire EffDIS à l'avenir.

EffDIS est une ressource importante pour le Sous-comité des écosystèmes. En 2013, ce Sous-comité a recommandé que des efforts soient déployés pour élaborer des estimations similaires de EffDIS pour la canne et la senne. Ceci serait particulièrement utile pour les évaluations des mesures de gestion des fermetures spatio-temporelles. Il sera important d'identifier la mesure de l'effort de pêche global à utiliser pour ces deux engins.

En dépit de l'importance que revêt EffDIS pour le travail du SCRS, la dernière mise à jour a été réalisée en 2007. Ceci est dû en grande mesure à plusieurs problèmes méthodologiques qui doivent encore être résolus (voir SCRS/2013/036 qui contient un résumé complet) et aux ressources insuffisantes pour réaliser l'analyse. En particulier, il est nécessaire de valider le postulat utilisé pour construire la base de données agrégée à un niveau de carrés de 5°x5° et d'évaluer s'il convient à l'emploi visé. Le groupe recommande dès lors :

- 1) que des méthodes de classification des flottilles soient explorées et
- 2) qu'un soutien externe soit apporté pour élaborer EffDIS à l'avenir.

3. Quantification de l'incertitude dans les évaluations de l'ICCAT

La Commission espère un avis sur les mesures de gestion, lequel serait basé sur les risques, tel qu'énoncé dans la matrice de stratégie de Kobe II et inscrit dans son cadre de décision (Rec. 11-13). Un aspect important lié à la formulation de cet avis scientifique est la quantification adéquate de l'incertitude entourant l'état des stocks et les perspectives futures selon les scénarios de futures options de gestion. Avec l'arrivée de modèles d'évaluation des stocks appliqués plus communément et hautement paramétrés, l'investissement informatique que représente la quantification de l'incertitude entourant l'état des stocks et les perspectives futures est assez lourd. D'autres ORGP thonières font le même constat et un certain nombre d'approximations pour quantifier les deux processus et l'incertitude par observation sont appliquées afin de formuler un avis de gestion basé sur les risques. Il avait dès lors été demandé au WGSAM de fournir une orientation sur l'évolution des méthodes et sur leur éventuelle harmonisation afin de demander aux groupes d'espèces de décrire l'incertitude. À cet égard, le WGSAM a fait remarquer que tous les modèles d'évaluation représentent des simplifications d'un système de pêche plus complexe et qu'ils sont donc intrinsèquement limités en ce qui concerne leur capacité de refléter la variabilité inhérente de ce système. En d'autres termes, les gestionnaires ne devraient pas s'attendre à ce que l'incertitude entourant l'avis scientifique sur les pêcheries soit un jour totalement quantifiée. Néanmoins, le groupe a convenu que des améliorations peuvent être apportées de manière à ce que les différents groupes d'espèces représentent l'incertitude.

Le groupe a identifié trois approches de base visant à caractériser l'incertitude dans les sorties des modèles d'évaluation de stock : (1) une approche « fondée sur le modèle » qui tient explicitement compte des principales sources d'incertitude perçues utilisant une seule plateforme de modélisation polyvalente ; (2) une moyenne des modèles, c'est-à-dire, une combinaison des scénarios alternatifs des diverses plates-formes de modélisation, et (3) une approche « empirique » qui utilise des informations historiques sur la cohérence des sorties des modèles d'évaluation de stocks au cours du temps. La première approche est l'approche la plus communément utilisée. De manière générale, un seul modèle de base est développé comprenant des paramètres estimables représentant ce qui est perçu comme étant les principales sources d'incertitude (avec ou sans priors informatifs). Les mesures de l'incertitude dans les paramètres clés de gestion sont ensuite calculées avec des méthodes standard (inverse du Hessien, bootstrap, intégration bayésienne) et intégrées dans les matrices de Kobe. Des scénarios de sensibilité ou d'autres plateformes de modélisation peuvent être exécutés afin de se faire une meilleure idée des incertitudes potentielles, mais sont souvent utilisés principalement pour qualifier l'avis du modèle de base et ne sont pas incorporés directement dans les matrices de Kobe. L'efficacité de cette approche dépend de la capacité de modifier le modèle d'évaluation de façon à se rapprocher raisonnablement de la source d'incertitude.

La moyenne des modèles est la deuxième stratégie, qui a parfois été utilisée pour incorporer les informations fournies par d'autres modèles. Dans ce cas, des pondérations par fréquence sont attribuées à chaque modèle concourant, reposant sans doute sur quelques mesures de l'ajustement aux données (p.ex. AIC, pondération par l'inverse de la variance) ou l'opinion d'experts (voir également le point 5). Cette approche présente un

inconvenient, en ce sens qu'elle donnera probablement lieu à une distribution multimodale des mesures de l'état du stock à l'étude (p. ex., point de référence limite) ou nécessitera des postulats sur son mode de distribution (p. ex., normale ou lognormale). Le groupe a estimé que le succès de l'une ou l'autre approche « fondée sur le modèle » dépend fortement de la capacité du groupe d'évaluation d'élaborer des priors informatifs ou des postulats alternatifs raisonnables des états de la nature avant de découvrir les implications de gestion de ces états. D'autre part, il existe un risque de produire des estimations biaisées de la tendance centrale et de l'incertitude en dépurant des modèles jugés non informatifs ou en introduisant intentionnellement des biais en ajoutant des modèles ou en altérant les priors. Il a été mentionné que le Centre national des ouragans des États-Unis utilise systématiquement une moyenne des modèles et obtient les pondérations de chaque modèle concourant à partir des résultats obtenus dans des prévisions antérieures (c.-à-d. en comparant les prévisions du modèle avec les trajectoires réelles des ouragans). Malheureusement, contrairement aux trajectoires des ouragans, l'état réel d'un stock de poissons est rarement connu.

Les approches empiriques examinent les résultats obtenus antérieurement par un modèle afin de déduire les marges de l'incertitude totale. Idéalement, il faudrait comparer le rendement du modèle avec l'état réel du stock, ce qui n'est pas possible dans la pratique. On peut également comparer le rendement de deux ou plusieurs évaluations de stocks historiques de référence en ce qui concerne les années communes. L'enveloppe résultante de l'incertitude totale inclurait implicitement les erreurs d'estimation « inhérente au modèle » ainsi que les erreurs systémiques que les changements des modèles ou les changements philosophiques de l'équipe d'évaluation peuvent entraîner. Cette approche a déjà été adoptée par le Conseil de gestion des pêcheries du Pacifique des États-Unis (cf. Ralston et al. 2011) et d'autres conseils étudient actuellement son adoption. Quelques problèmes potentiels découlant de la méthode empirique ont été soulevés pendant les discussions du groupe. On a exprimé la crainte que cette approche ne reflète pas la tendance des évaluations de stocks d'améliorer la précision temporelle en ajoutant de nouvelles données, en améliorant les modèles et en augmentant la collecte de données. Toutefois, d'autres membres du groupe de travail ont souligné que les améliorations des données ne dissiperont qu'une partie de l'incertitude et que la composition des scientifiques formant l'équipe d'évaluation des stocks, la composition du groupe d'examen de l'évaluation et les changements des méthodes d'évaluation des stocks pourraient avoir la même influence sur la cohérence des résultats de l'évaluation de stocks. Il a été suggéré que les analyses empiriques pourraient au moins être utilisées pour aider à fournir une vérité de terrain des estimations de la variance dérivées des procédures actuelles fondées sur le modèle.

Une présentation intitulée « Quantification de l'incertitude causée par le traitement des données dans les évaluations de stocks structurées par âge » a été réalisée. Des travaux préliminaires ont été présentés décrivant une méthode d'imputation des données de taille manquantes provenant d'observations reposant sur la proximité spatio-temporelle (en postulant que les observations étaient approximativement multivariées et normales avec des covariances estimées). Des distributions estimées ont ensuite été utilisées pour créer plusieurs jeux de données de taille complets (en échantillonnant les paramètres attribués au modèle et en échantillonnant ensuite à partir des observations des tailles pondérées) qui ont été convertis en distributions des âges au moyen de plusieurs méthodes de détermination de l'âge (p.ex. découpage des cohortes et clés âge-taille). Chaque jeu de données d'âge a ensuite été utilisé dans une VPA avec différents taux de mortalité naturelle et indices d'abondance relative. Les résultats préliminaires donnent à penser que l'imputation des tailles avec un échantillon effectif de taille de 100 est trop faible pour générer une variabilité dans les jeux de données de prise par taille dérivés. Le taux de mortalité naturelle (80, 100 et 120% des niveaux de l'évaluation la plus récente) détermine en grande mesure la F_{PME} estimée sur la base de l'analyse XSA.

Le groupe a convenu que les variations de cette approche seraient utiles pour tenter de quantifier la façon dont l'incertitude entourant les données se propagera dans l'incertitude entourant l'évaluation. Une telle approche aiderait également à déterminer l'importance relative d'une source présumée d'incertitude et le besoin correspondant de l'incorporer explicitement dans le développement des matrices de Kobe utilisées aux fins de la formulation de l'avis scientifique destiné aux gestionnaires. On a également souligné pendant le débat sur les évaluations des stratégies de gestion que les modèles opérationnels devraient être construits en tenant compte des principales sources d'incertitude perçues et que celles-ci pourraient être utilisées pour aider à quantifier la mesure dans laquelle ces incertitudes se propagent dans l'incertitude entourant l'avis scientifique (qui correspond de fait à une variation de l'approche abordée ci-dessus). On a également fait remarquer que les analyses de risque telles que celle réalisée sur le thon rouge (Leach et al. 2014) permettent d'identifier les principales sources d'incertitude perçues par les scientifiques et les autres parties prenantes, ce qui peut présenter l'avantage d'offrir une plus grande acceptation lors de la présentation des résultats.

4. Description de la qualité des données des pêcheries et des informations biologiques

Le document SCRS/2014/035 a été présenté dans le but de réviser les informations de référence et de promouvoir la discussion au titre de ce point de l'ordre du jour. Aux termes de la Résolution en vue de standardiser la présentation des informations scientifiques dans le rapport annuel du SCRS (Rés. 13-15), le SCRS devrait évaluer plus en profondeur la qualité des données halieutiques et des données se rapportant à la connaissance de l'espèce (par exemple, les paramètres biologiques, l'historique des modes de pêche, la sélectivité) utilisées comme données d'entrée des évaluations des stocks. Les notes qualitatives des données d'entrée et des postulats pourraient être détaillées et devraient résumer l'état des connaissances sur les différentes données d'entrées. Le rapport du Secrétariat sur les statistiques de 2013 proposait une méthode d'attribution de notes aux différents éléments de données. La méthode consistait à utiliser un tableau contenant une liste de catégories et d'éléments (au sein d'une catégorie) qui étaient évalués et pondérés (par exemple, en fonction de l'importance relative des éléments au sein de chaque catégorie, dans l'évaluation du stock). La note attribuée à chaque élément était le résultat de sa valeur et de sa pondération, et la note de chaque catégorie était la somme des différentes notes des éléments au sein d'une catégorie. On a suggéré que la qualité globale de l'information d'une évaluation de stock donnée soit calculée comme la note moyenne des catégories. La proposition ne prévoyait toutefois pas de critères prédéfinis servant de base à la notation, tels que ceux utilisés dans les catalogues de données des espèces de la CTOI, où les limites des notes sont déterminées par des énoncés qualitatifs de la disponibilité/qualité des données (p ex. « disponible suivant les normes », « non disponible suivant les normes » ou « non disponible »).

Par conséquent, le document SCRS/2014/035 proposait d'utiliser un cadre générique, semblable à la proposition avancée par le Sous-comité des statistiques, mais incluant des directives prédéfinies de notation qualitative. La note qualitative (Q) proposée des différents jeux d'informations était :

$$Q = D * I$$

Où :

D= disponibilité et qualité des données

Avec :

- 1= pas de données ou données de mauvaise qualité/qualité non connue
- 2= les données existent, mais leur qualité n'est pas très bonne
- 3= des données de bonne qualité existent

I= impact des données sur les résultats de l'évaluation (l'importance de ce jeu de données).

Avec :

- 1= faible impact sur les résultats de l'évaluation
- 2= impact modéré sur les résultats de l'évaluation
- 3= impact élevé/non connu sur les résultats de l'évaluation

Le document incluait plusieurs exemples où la notation était provisoirement appliquée à différents stocks de germon et de listao. Différents éléments concernant les pêcheries et la biologie ont été soumis au système de notation et la note totale pour chaque stock est le résultat de la somme des valeurs Q des différents éléments.

Le groupe a estimé que ce cadre générique pouvait servir de point de départ utile et a discuté de plusieurs questions s'y rapportant. Il a été convenu que les notes doivent être fondées sur une méthodologie objective, quantifiable et scientifiquement défendable. Alors que cet exercice pourrait être plus facile à réaliser lorsqu'il s'agit d'attribuer une note à la disponibilité des données, il a été reconnu qu'il pourrait être légèrement plus difficile d'évaluer objectivement la qualité des données. Toutefois, le groupe a suggéré des mécanismes permettant de faire progresser la notation dans la mesure du possible.

À ce stade, il convient de tenir compte de l'objectif ou l'utilité de cet exercice. D'une part, il est nécessaire de répondre à la Résolution 13-15 afin que la Commission dispose, outre la matrice de Kobe, d'informations sur la qualité des données utilisées pour dériver la matrice de Kobe. À cette fin, une représentation simple (similaire aux exemples présentés dans le document SCRS/2014/035) a été considérée. Néanmoins, le groupe a estimé que les notes qualitatives pourraient également être appliquées à d'autres fins. À titre d'exemple, elles pourraient être utiles pour que les groupes de travail scientifiques décrivent les principales sources d'incertitude et améliorent, dans la mesure du possible, la formulation de l'avis scientifique (p.ex. en sélectionnant les modèles appropriés et/ou en pondérant différents scénarios).

Le groupe a noté que l'importance des différents jeux de données (notes « I ») dépend du modèle utilisé ainsi que du cycle vital du stock. Ainsi, les notes « Q » totales peuvent ne pas être comparables d'un stock à l'autre. Afin de surmonter cet écueil, le groupe a recommandé de quantifier le ratio entre la note totale et la note maximale potentielle compte tenu du vecteur I spécifique à chaque stock, ce qui permettrait de réaliser une comparaison basée sur les couleurs entre les stocks.

Finalement, le groupe a recommandé de développer une base de métadonnées contenant des informations sur la qualité et la quantité des données biologiques et halieutiques disponibles pour les évaluations. Cette idée, conjuguée au mécanisme actuel d'évaluation des CPUE et des indices d'abondance indépendants des pêcheries, contribuerait à évaluer plus objectivement la qualité et la qualité des informations halieutiques utilisées dans l'évaluation des espèces de l'ICCAT pour éclairer les scientifiques qui effectuent les évaluations de stock et pour fournir une base d'une vue d'ensemble de la qualité des données à la Commission, c.-à-d. la valeur des notes « D ». Le groupe a discuté de la structure que pourrait présenter une base de métadonnées, qui peut ensuite être alimentée et mise à jour dès que de nouvelles informations sont disponibles.

La base de métadonnées aurait trois composantes principales, la première concernerait les données des pêcheries, la deuxième les données biologiques et la troisième les études sur les récupérations de marques (**Appendices 4** et **5**, respectivement). La structure suggérée est souple et pourrait intégrer de nouveaux champs en tant que de besoin. Dans le cas des métadonnées biologiques, il a été estimé que les champs « Informations générales » ainsi que certaines informations indiquant quels sont les paramètres utilisés et la mesure dans laquelle ils sont estimés seraient déjà très utiles. Les champs « estimations des paramètres » permettraient de réaliser des méta-analyses qui pourraient être utiles de plusieurs façons, par exemple pour caractériser l'incertitude de paramètres donnés. Toutefois, alimenter ces champs impliquerait bien entendu un important travail. Le groupe a également estimé qu'il pourrait être utile, dans certains cas, d'avoir accès aux données brutes de quelques études. La base de métadonnées permettrait d'identifier facilement les sources de ces données brutes.

5. Rapprochement des résultats lorsque plusieurs méthodes de modélisation sont utilisées

Les évaluations du SCRS ont souvent et, de plus en plus, inclus l'utilisation de plusieurs méthodes de modélisation (c.-à-d. différents types de modèle, autres hypothèses) pour estimer l'état du stock par rapport aux points de référence de conservation de l'ICCAT. Le groupe s'est penché sur les différentes raisons de procéder de la sorte. Par exemple, les types de modèle peuvent différer des postulats sous-jacents et les participants à l'évaluation peuvent être incapables de déterminer la solution la plus appropriée. Dans ces cas-là, l'intégration des résultats de plusieurs modèles dans l'avis de gestion vise à mieux refléter l'incertitude entourant les résultats. Dans d'autres cas, d'autres configurations de la même plateforme ou type de modélisation peuvent être considérées comme étant suffisamment plausibles pour les inclure dans les résultats. Le groupe a convenu que cette approche est utile, mais a reconnu qu'il peut s'avérer difficile de déterminer la façon la plus appropriée de combiner ces résultats, y compris l'attribution d'une pondération adéquate, notamment lorsque les divers résultats du modèle peuvent être contradictoires.

Aucun nouveau document SCRS n'était disponible à ce sujet. Cependant, le groupe a passé en revue les travaux récents de Deroba et al. (2014) utilisant des simulations pour tester la solidité des modèles d'évaluation des stocks à l'erreur. Les conclusions tirées dans ce document se rapportant à ce sujet sont les suivantes :

- Les différences les plus importantes sont apparues lorsque l'on compare les résultats de différents types de modèles (modèle de production excédentaire, modèle fondé sur l'âge, etc.). Les modèles du même type avaient tendance à procéder de la même manière.
- L'auto-évaluation est utile et recommandée et l'auto-évaluation par recoupement a souvent mis en évidence des divergences dans les séries temporelles des années les plus récentes.
- La variabilité entre les modèles peut être considérée comme un type d'incertitude entourant l'évaluation, ce qui a des incidences lorsque l'on envisage de faire reposer l'avis de gestion uniquement sur un seul modèle configuré de manière à être le « meilleur ajustement » aux fins de l'évaluation ou d'incorporer les résultats des différents modèles (par exemple, l'application de la moyenne des modèles).

Les conclusions de Deroba et al. (2014) coïncidaient avec les expériences des évaluations de stocks du SCRS et fournissent une confirmation analytique de l'importance du choix du modèle (en particulier en ce qui concerne le type de modèle), l'examen des diagnostics et la validation des résultats, et les implications de l'intégration ou non des modèles dans l'avis final.

Le groupe a estimé qu'une composante initiale essentielle de rapprochement des différents résultats des modèles est le processus de sélection des modèles les plus appropriés sur lesquels faire reposer l'avis de gestion. Ce processus devrait commencer pendant la réunion de préparation des données, au cours de laquelle il conviendrait de prendre le temps de déterminer les modèles qui seront utilisés dans l'évaluation. Comme décrit ci-dessus, les différentes exigences des modèles en matière de données, leurs postulats et leur capacité de tenir compte des importants changements de la pêche s'étant produits, à notre connaissance, devraient être considérés à l'aune de la qualité et de la quantité des données disponibles. La participation des scientifiques des CPC est fondamentale dans ce processus, car ils sont les plus familiarisés aux changements des différentes pêcheries qui pourraient justifier un examen dans le processus d'évaluation, et sont les plus à même d'identifier les préoccupations au sujet du bien-fondé de certains jeux de données à utiliser dans les différents modèles. Les résultats des différents modèles utilisés dans l'évaluation précédente devraient également être examinés.

Pendant la réunion d'évaluation, suffisamment de temps devrait être consacré à l'examen minutieux des diagnostics et des résultats du modèle. Les mesures des résultats pourraient inclure l'évaluation de la question de savoir si les résultats sont compatibles avec les connaissances sur la biologie de l'espèce ainsi que les tendances des pêcheries (un exemple de cette approche, « la mesure des résultats de la prévision rétrospective », est détaillé ci-dessous). La probabilité et les conséquences de la violation des postulats de chaque modèle doivent être considérées en tenant compte des informations disponibles sur la population et la pêche. Le fait que le modèle identifié pendant la réunion de préparation des données aux fins de son utilisation dans l'évaluation puisse être (correctement) rejeté dans l'avis de gestion final est l'une des conséquences possibles de ce processus.

Le groupe s'est demandé s'il serait approprié d'examiner les résultats de modèles plus simples ainsi que ceux de modèles plus complexes (par exemple, en utilisant les résultats d'un modèle de production excédentaire ainsi que ceux d'un modèle plus pleinement intégré, tel que le Stock Synthèse). En d'autres termes, si les données sont suffisantes pour exécuter des modèles plus complexes (qui, en théorie, pourraient prendre en compte davantage d'aspects et les changements du stock et des paramètres halieutiques), l'avis ne devrait-il pas reposer uniquement sur les résultats de ces modèles complexes ? Le groupe a estimé que dans certains cas il convient de tenir compte des résultats de ces types de modèles divergents.

Le groupe a également discuté des options pour des critères ou des procédures additionnelles pour sélectionner les modèles. Une telle procédure, basée sur une « mesure des performances de la prédiction rétrospective », a été suggérée. L'idée de cette méthode est similaire à l'analyse rétrospective traditionnellement utilisée dans les VPA mais elle s'étend à la comparaison des modèles multiples. À titre d'exemple, on postule que deux modèles sont utilisés dans l'évaluation et fournissent deux résultats différents. Les étapes du processus de mesure des performances de la prédiction rétrospective consisteraient à :

- tronquer les données disponibles, abandonnant les xx dernières années (tel qu'approprié pour le stock),
- ré-exécuter chaque modèle approprié afin d'obtenir une estimation de la dynamique des populations pendant xx années dans le passé,
- projeter en avant ce modèle en utilisant les prises réelles correspondant à l'année en cours,
- calculer une tendance de la CPUE (ou peut-être tendance de la biomasse estimée, en fonction de la complexité du modèle) qui est prédite pour ces années projetées par rapport à l'année en cours et
- comparer cette tendance de CPUE à la/aux tendances(s) observée(s) réelle(s) de la CPUE (standardisée).

Avec cette approche, les modèles pourraient ne pas être retenus si les résultats de cette analyse font apparaître des incohérences avec la/les tendance(s) observées de la CPUE (standardisée). La méthode peut être amplifiée pour développer des pondérations pour les divers résultats des modèles (à utiliser lors de la combinaison des résultats entre les modèles) en définissant une mesure de divergence entre les séries de CPUE prédites et observées (p.ex. somme des carrés des valeurs résiduelles divisée par SE de la CPUE observée). Il convient de noter que cette approche n'exige pas que les modèles utilisent les mêmes jeux de données. L'applicabilité de cette approche dépend de la question de savoir si les tendances de CPUE réelles observées (standardisées) reflètent adéquatement la population et l'absence d'indices de CPUE contradictoires. Le groupe a trouvé que cette approche était prometteuse et il a recommandé de poursuivre la recherche afin de déterminer si elle pouvait s'appliquer à la sélection et/ou à la pondération des modèles.

Le groupe a noté que la sélection des modèles à utiliser pour les projections était également importante. Dans certaines circonstances, on pourrait utiliser pour réaliser les projections un modèle différent à celui employé pour estimer l'état actuel du stock. Cela peut arriver par exemple si le modèle utilisé pour estimer l'état des stocks ne peut pas servir pour les projections ou a besoin d'un temps de traitement supplémentaire (au-delà de ce qui peut

être accompli pendant la réunion d'évaluation) afin de produire certains résultats qui seraient requis pour la projection. Le groupe a fait remarquer que quelques préoccupations ont été exprimées concernant le caractère approprié de cette approche. Le groupe a identifié la nécessité de développer des approches afin d'enquêter sur cette question. Une approche potentielle suggérée consisterait à ré-exécuter le modèle utilisé pour estimer l'état actuel du stock, en supprimant un certain nombre d'années récentes de données (comme dans une analyse rétrospective). Le modèle envisagé pour les projections pourrait ensuite être exécuté, se projetant à partir de la dernière année du scénario du modèle rétrospectif, utilisant l'historique des captures réelles jusqu'à l'année en cours. La performance du modèle de projection pourrait être évaluée en comparant l'état/situation du stock projeté dans l'année en cours avec ce qui est estimé par le modèle d'évaluation.

Une fois que les scientifiques réalisant l'évaluation se seront mis d'accord sur les modèles à inclure pour formuler l'avis de gestion (p.ex. estimations de l'état des stocks par rapport aux points de référence biologiques et à l'incertitude associée, matrice de stratégie de Kobe), reste à savoir comment présenter ces résultats. Une option généralement préférée consisterait à combiner d'une certaine manière ces résultats afin de fournir des paramètres et des représentations appropriées de l'incertitude. Ceci est particulièrement difficile lorsque les divers résultats sont contradictoires entre les modèles. Le groupe a signalé qu'il était généralement inapproprié de combiner des résultats qui sont complètement incompatibles entre chaque modèle ; par exemple, quand les différents modèles sont structurés sur la base de deux hypothèses alternatives qui reflètent des états présumés très différents, avec peu ou aucun chevauchement dans les distributions de probabilités. En pareils cas, le fait de combiner ces résultats de façon égale produirait des estimations de l'état des stocks et/ou des tendances égalisées entre les tendances, ce qui n'est compatible avec aucune des hypothèses alternatives. Dans ces circonstances, les scientifiques sont encouragés à examiner rigoureusement les postulats sous-jacents de chaque hypothèse et envisager prudemment si chacun est étayé par les données existantes et les connaissances scientifiques actuelles. Si ce processus ne suffit pas à rejeter une hypothèse, les scientifiques devraient se demander si des probabilités relatives de chaque hypothèse peuvent être assignées. Il pourrait être conseillé de réaliser des recherches approfondies afin de recueillir des données en vue d'aider à accepter/rejeter les hypothèses ou d'attribuer des vraisemblances.

Lorsque les résultats sont combinés à partir de plusieurs modèles, le SCRS a pour pratique habituelle d'attribuer une pondération égale aux résultats de chaque modèle. Comme il est mentionné à la rubrique 3 du présent rapport, des pondérations pourraient être affectées aux résultats de chaque modèle sur la base d'une certaine mesure d'ajustement aux données (p.ex. AIC, pondération de variance inverse). En plus des préoccupations décrites à la rubrique 3, cette approche est limitée en ce sens qu'elle est appropriée seulement quand les modèles alternatifs utilisent les mêmes jeux de données.

Le groupe a également discuté de l'emploi éventuel de la MSE pour informer la sélection/pondération du modèle. Le groupe a identifié un problème potentiel avec cette approche en ce sens que le modèle opérationnel de la MSE risque d'influencer la perception des performances du modèle. Il a été suggéré que l'utilisation des modèles de production excédentaire pourrait s'avérer problématique si les fonctions de productivité sont fortement inclinées vers la gauche et que la productivité monte en flèche au fur et à mesure que les niveaux des stocks chutent.

Le groupe a reconnu qu'un élément de subjectivité est inévitable dans tout schéma de pondération. Même l'assignation par défaut de pondération égale constitue implicitement une décision susceptible de donner plus de poids que justifié à certains résultats si la totalité des facteurs, tels que postulats de modèles, capacité, qualité des données, tendances des pêcheries et biologie, pouvaient être objectivement pris en compte. C'est pourquoi le groupe estime que l'opinion des experts pourrait être appropriée pour l'assignation des pondérations, si elle était étayée par la connaissance de ces facteurs et la postulation de mécanismes raisonnables qui appuieraient une conclusion selon laquelle certains modèles sont plus susceptibles que d'autres de refléter la véritable situation.

Une autre pratique commune dans les évaluations du SCRS consiste à exclure finalement de l'avis de gestion divers modèles qui produisent des résultats plausibles, mais qui pourraient être considérés d'une certaine façon comme moins vraisemblables. Le groupe a noté que des inquiétudes existent quant au fait que la pratique actuelle de sélection d'un seul sous-ensemble de modèles jugés les plus vraisemblables dans les résultats finaux pourrait entraîner une sous-estimation de la gamme de l'incertitude. Une recherche plus poussée est requise.

6. Points limites de référence, normes de contrôle de la ponction et évaluations de la stratégie de gestion

L'évaluation des points limites de référence (« LRP ») et de normes de contrôle de la ponction (« HCR ») au moyen de l'utilisation d'une évaluation de la stratégie de gestion (« MSE ») est de plus en plus reconnue par les ORGP thonières mondiales comme étant un moyen efficace de faire avancer leur processus de gestion des pêcheries. Les évaluations de 2013 de germon et d'espadon ont servi d'exemple pour illustrer la façon dont un processus de MSE pourrait formellement être inclus dans la gestion de ces stocks. Le groupe de travail sur les méthodes d'évaluation envisage de poursuivre cet effort en (1) continuant à affiner les méthodes dans le cadre du processus de MSE, (2) introduisant la MSE dans davantage d'évaluations lorsque cela s'avère pertinent et (3) favorisant la communication en vue de tenir les gestionnaires informés de ses points forts et faibles. En ce qui concerne le dialogue et la communication, l'ICCAT a récemment adopté la *Recommandation de l'ICCAT visant à renforcer le dialogue entre les halieutes et les gestionnaires des pêcheries* (Rec. 13-18), qui vise de façon générale à renforcer la communication et à améliorer la compréhension mutuelle entre les gestionnaires des pêcheries et les halieutes afin de faciliter la prise de décisions reposant sur la science et à la simplifier. De plus, la Recommandation énumère les tâches spécifiques à réaliser pendant la première réunion du groupe de travail permanent dédié au dialogue entre halieutes et gestionnaires des pêcheries (SWGSM) en 2014.

Une norme de contrôle de la ponction empirique a été adoptée pour le thon rouge du sud afin d'établir le total des prises admissibles (TAC). La HCR est basée sur des changements d'année en année dans les indices à partir d'un indice de CPUE d'abondance d'adultes dépendant des pêcheries et d'une prospection aérienne de juvéniles indépendante des pêcheries. Avant que la HCR ne puisse être mise en œuvre, des niveaux de référence appropriés pour la capture et les indices doivent être sélectionnés et les paramètres de la HCR doivent être calibrés afin de répondre aux objectifs de gestion à l'aide de l'évaluation de la stratégie de gestion (MSE).

Dans le SCRS/2014/036, une MSE préliminaire a été réalisée pour le thon rouge de la Méditerranée en utilisant la norme de contrôle de la ponction du thon rouge du Sud dans le cadre d'une procédure de gestion (MP, c.-à-d. la combinaison de données prédéfinies conjointement avec un algorithme auquel les données sont assignées pour fournir une valeur pour un TAC ou une mesure du contrôle de l'effort). Les prochaines démarches seront les suivantes :

- 1) Identification des objectifs de gestion et mise en correspondance de ceux-ci avec les mesures du rendement afin de quantifier la mesure dans laquelle ils peuvent être atteints.
- 2) Sélection des hypothèses relatives aux dynamiques du système.
- 3) Conditionnement des modèles opérationnels (OM) en fonction des données et des connaissances et suppression potentielle et pondération des différentes hypothèses.
- 4) Identification des stratégies de gestion potentielles et codification de celles-ci en tant que procédures de gestion.
- 5) Projection des modèles opérationnels en avant en utilisant les procédures de gestion comme procédures de contrôle des réponses.
- 6) Détermination des modèles opérationnels qui remplissent le mieux les objectifs de gestion.

Le groupe a souligné la nécessité de sélectionner minutieusement les postulats des modèles opérationnels étant donné que des postulats erronés peuvent affecter à la fois les résultats des évaluations de la stratégie de gestion (MSE) et les résultats des évaluations de stocks. On a signalé qu'il était important de sélectionner la gamme des scénarios utilisés dans le modèle opérationnel (Fromentin et al., 214 ; Leach et al., 2014). Le groupe a convenu que, dans un cadre MSE, le but n'est pas de choisir la procédure de gestion ou la norme de contrôle de la ponction (HCR) qui donne les meilleures performances, mais celle qui est la plus solide parmi tous les scénarios sélectionnés. À titre d'exemple, si le modèle opérationnel est exécuté avec quatre niveaux de la pente à l'origine de la relation stock-recrutement (steepness), dont on postule qu'ils sont tous plausibles à parts égales, la HCR choisie devrait être celle qui atteint les objectifs de gestion en vertu de ces quatre niveaux de steepness. Le groupe a également discuté que de nombreux modèles opérationnels postulent que la régulation des populations n'a lieu qu'au recrutement ; toutefois, le modèle opérationnel devrait également tenir compte d'autres possibilités de régulation des populations (p.ex. mortalité naturelle).

Le groupe a reconnu que le modèle opérationnel devrait être plus complexe que le modèle d'estimation. À titre d'exemple, un modèle de production excédentaire peut être utilisé comme modèle d'estimation, tandis qu'un modèle plus complexe, comme SS3, peut servir de modèle opérationnel. Le document a présenté un exemple dans lequel deux prospections ont été utilisées comme indices pour les recrues et les adultes. On a évoqué la possibilité de ne pas devoir réaliser ces types de prospection tous les ans dans le but d'obtenir des indices de biomasse qui informent la HCR. On a signalé au groupe que le fait d'utiliser la HCR pourrait réduire les

exigences en matière de données par rapport à celles qui sont actuellement utilisées dans certaines évaluations des stocks. Le groupe s'est également demandé si la HCR devrait inclure plus de points limites de référence de précaution pour les cas où l'indice de biomasse estimé à partir des prospections contient plus d'incertitude.

On a passé en revue la présentation réalisée par le Président du SCRS à la réunion de 2013 de la Commission. Le groupe a souligné l'importance de la première réunion du groupe de travail ICCAT permanent dédié au dialogue entre halieutes et gestionnaires des pêcheries qui aurait prochainement lieu et porterait sur les thèmes suivants : points de référence limites, probabilités associées à la matrice de Kobe, périodes de rétablissement pour différents stocks, etc. Le groupe a souligné que sans ce dialogue entre les scientifiques et les gestionnaires, la HCR ne peut pas être mise au point.

Le SCRS/2014/025 décrit la façon dont les auteurs recherchent des éléments de preuve d'une relation stock-recrutement (SRR) pour le thon rouge, l'albacore et le germon. Les auteurs ont conclu que les preuves de l'existence d'une SRR pour l'un quelconque des stocks étaient faibles et que les données pouvaient également étayer d'autres hypothèses (Vert-pre, 2014) p.ex. celle de Gilbert (1998) selon laquelle le recrutement fluctue autour d'un niveau moyen pendant une période de temps, puis un changement de régime survient. Ceci a d'importantes implications évidentes pour l'évaluation des stocks et l'avis de gestion.

Le groupe a signalé que l'hypothèse du « changement de régime » présentée dans le document pourrait être le produit des modèles d'estimation. En d'autres termes, même si le « changement de régime » n'est pas incompatible avec l'absence d'une SRR observée dans les stocks mentionnés, ceci ne constitue pas en soi la preuve qu'un « changement de régime » a véritablement eu lieu.

Le document suggérait que la SSB est une fonction du recrutement (au lieu que le recrutement soit une fonction de la SSB). Le groupe s'est demandé pourquoi il semblait y avoir une auto-corrélation dans les schémas de recrutement. On a signalé que l'auto-corrélation observée pourrait être due à l'utilisation du découpage des cohortes afin de déterminer l'âge des stocks ou bien due à d'autres postulats du modèle.

Le document suggérait que, pour les besoins des projections de l'état des stocks, davantage de poids devrait être accordé aux recrutements récemment observés plutôt que de pondérer tous les schémas de recrutement dans les séries temporelles de la même façon (en d'autres termes, le recrutement de demain sera très vraisemblablement plus similaire à celui d'hier qu'à celui d'il y a 25 ans).

Dans le SCRS/2014/037, les séries de R et SSB résultant des scénarios du cas de base pour le stock de thon rouge de l'Atlantique Est et de la Méditerranée utilisant des prises « réajustées » et « déclarées » ont servi à ajuster trois différents modèles S-R (Beverton & Holt, Ricker et *smooth Hockey-stick* (bâton de hockey lisse)). Les résultats montrent que le stock a maintenu son entière capacité reproductive pendant toute la série temporelle à partir des années 50 et qu'une relation S-R de Ricker ne peut pas être rejetée pour ce stock. Le document explique que ce dernier résultat ajoute davantage d'incertitude à l'estimation de B_{01} , ce qui la rend inutile comme point de référence de la biomasse pour ce stock. Comme alternative, le document propose d'employer B_{loss} comme B_{lim} et ensuite d'estimer B_{pa} à partir de cette valeur. Ceci a permis aux auteurs de simuler le comportement de certaines HCR et de sélectionner la plus adéquate pour ce stock.

Le groupe a indiqué que la CIEM s'éloignait du cadre du point de référence limite, ce qui est à l'opposé de ce qui est proposé dans le document. Nonobstant, l'utilité de l'approche proposée dans le document peut être testée au moyen de la simulation. Le groupe s'est également penché sur la validité de formuler des postulats dans les modèles opérationnels afin de faire preuve de plus de précaution (comme en utilisant une relation de recrutement en forme de bâton de hockey). En règle générale, le groupe n'a pas été favorable à cette approche. Le groupe a également évoqué les facteurs susceptibles de contribuer à une relation S-R de type Ricker chez les thonidés, sans parvenir à une conclusion finale.

7. Incorporation des informations sur l'écosystème, le climat et l'habitat (ECH) dans les évaluations des stocks

Les discussions relatives à l'incorporation des informations sur l'écosystème, le climat et l'habitat (ECH) dans les évaluations des stocks se sont centrées sur deux approches qui peuvent être utilisées pour renforcer le processus d'évaluation des stocks des espèces relevant de l'ICCAT. Deux approches ont été identifiées : une approche qualitative et une approche quantitative, cette dernière visant à réduire l'incertitude entourant le processus d'évaluation des stocks.

Une seule présentation intitulée « Hypothèse d'un récent déplacement vers le pôle de la distribution de l'espadon de l'Atlantique nord » a été réalisée au titre de ce point de l'ordre du jour, inspirée du document SCRS/2013/161. Le document décrit les changements dans les indices d'abondance de l'espadon par rapport aux indicateurs des processus environnementaux à grande échelle (p.ex. AMO, NAO, AWP). Les observations des tendances opposées de l'abondance pour l'espadon du nord suggéraient la possibilité d'un déplacement de l'abondance depuis les latitudes chaudes du Sud vers les latitudes plus froides du Nord. Plusieurs indices d'abondance observés changeaient brutalement de direction, de valeurs négatives à positives, alors que d'autres présentaient un changement inverse. Les changements de direction observés des indices d'abondance correspondent aux changements des tendances de la taille de la piscine d'eau chaude de l'Atlantique (Atlantic Warm Pool - AWP), du changement de signe de l'oscillation atlantique multiséculaire (Atlantic Multidecadal Oscillation - AMO) et de l'oscillation nord-atlantique (North Atlantic Oscillation - NAO). Dans l'objectif de quantifier un éventuel rapport entre les changements de l'abondance et les divers indices environnementaux potentiels, nous avons exécuté le modèle d'évaluation sans l'influence des données environnementales et procédé à la régression des valeurs résiduelles de l'ajustement de la CPUE aux divers indices environnementaux. Compte tenu des limites de tolérance aux températures prévues pour l'espadon, il est possible que sa préférence en matière d'habitat ait évolué vers le nord ou qu'il ait changé de proie préférée, ou que ces deux facteurs soient réunis.

Avec ces résultats comme référence, une discussion a eu lieu sur les avantages de l'agrégation des différentes CPUE par zone plutôt que l'agrégation actuelle par flottille dont dispose le groupe d'espèces sur l'espadon. Ceci nécessiterait l'accès aux données des flottilles opération par opération et permettrait d'incorporer l'effet de la zone et de flottille ainsi que la structure spatiale dans le modèle d'évaluation (c.-à-d. CPUE de l'Atlantique Est par opposition à CPUE de l'Atlantique Ouest). On a de surcroît souligné qu'il existe d'autres approches visant à incorporer la zone comme facteur d'interaction dans la modélisation des CPUE en utilisant les modèles mixtes. Les approches de pondération pourraient être calculées une fois que les CPUE auront été standardisées par zone.

Le groupe de travail a également introduit l'emploi de la modélisation de l'habitat afin d'incorporer la variabilité environnementale en déterminant la distribution spatio-temporelle, dans la modélisation des CPUE.

Une longue discussion s'est ensuivie sur la façon d'aborder l'agrégation des données consignées opération par opération par les pêcheries palangrières ciblant l'espadon en tant que prolongement de l'étude présentée. Un grand domaine de discussion a porté sur la façon d'établir un cadre de collaboration approprié entre les scientifiques nationaux afin d'obtenir des données consignées opération par opération et sur la manière dont les données pourraient être conjointement analysées. Cet échange pourrait avoir lieu en vertu de l'Accord de confidentialité de l'ICCAT, déjà mis en place. Il a été reconnu que ceci serait un processus continu, qui ferait vraisemblablement appel au Secrétariat et qu'il s'agirait de travailler dans le « nuage » dont l'accès serait restreint aux scientifiques impliqués dans l'analyse collaborative des données. Ces données sont requises pour aborder un certain nombre de préoccupations/interactions associées à la pêche et aux indices afin d'améliorer l'évaluation et d'incorporer des facteurs environnementaux.

8. Autres questions

8.1 Collaboration avec SISAM de la CIEM

Une présentation a été faite devant le groupe par l'un des co-présidents de l'initiative stratégique sur les méthodes d'évaluation des stocks (SISAM) de la CIEM. Celle-ci a décrit les travaux menés à ce jour dans le cadre de cette initiative et une discussion a été tenue sur l'orientation future de la SISAM et son importance pour la communauté scientifique de l'ICCAT. Les scientifiques de l'ICCAT ont pris part aux travaux de la SISAM en 2013 en assistant à la Conférence mondiale sur les méthodes d'évaluation des stocks (WCSAM). Au nombre des réalisations de la SISAM en 2014, on peut citer la finalisation de documents pour une édition spéciale du Journal des sciences marines de la CIEM (reposant sur le WCSAM), des ateliers organisés par des organisations partenaires, telles que CAPAM, ainsi qu'une session ouverte que tiendra la CIEM à son ASC en septembre 2014. Les participants à la réunion ont souligné que la CIEM et l'ICCAT sont au moins confrontées à des questions similaires en termes de développement de méthodes d'évaluation des stocks et de formulation de l'avis sur les pêcheries, et que ces questions pourraient être communes à de nombreuses ORGP à l'échelle mondiale. Ces questions portaient sur l'avancée des méthodologies et le défi de maintenir suffisamment de scientifiques techniquement expérimentés dans le réseau d'organisations pour exécuter ces méthodes d'une façon appropriée. On a également discuté des initiatives communes visant à élaborer des évaluations et des cadres d'avis pour affronter les situations où les données sont limitées. Il a été convenu qu'il serait avantageux pour l'ICCAT de continuer à collaborer à l'initiative de la SISAM et, à travers cela, de travailler en coopération avec des scientifiques d'autres ORGP afin de contribuer à la résolution de ces questions.

9. Recommandations

- Le groupe a recommandé d'encourager les CPC à déclarer leurs données de prise et d'effort de Tâche II dans une stratification géographique plus fine (p.ex. 1° x 1°) au lieu de déclarer ces données par carrés de 5° x 5°, étant donné que cette échelle pourrait dans certains lieux être trop imprécise. Le WGSAM demande en outre au Sous-comité des statistiques d'envisager de solliciter la déclaration des données dans cette résolution à plus fine échelle.
- Le groupe a également convenu que la mise en œuvre de l'approche de l'évaluation de la stratégie de gestion (MSE) et la promotion du dialogue entre les scientifiques et les gestionnaires des pêcheries sur les normes de contrôle de la ponction et la MSE devraient être encouragées afin d'améliorer l'avis scientifique formulé à la Commission. Ces efforts devraient inclure l'examen des efforts de MSE déployés jusqu'à présent en tenant compte des succès, des échecs et des ressources limitant les progrès futurs de la MSE.
- Le groupe a estimé que des critères simples pouvaient être utilisés par les différents groupes de travail afin de commencer à attribuer une note à la qualité de l'information utilisée dans les différentes évaluations de stock. Entre-temps, le groupe recommande de poursuivre le développement de moyens plus détaillés, objectifs et fondés sur la science pour fournir de telles notes. Ceci prévoit l'élaboration de critères pour évaluer l'importance des différents éléments de données en fonction du cycle vital et/ou du modèle d'évaluation utilisé. Dans cette ligne d'idée, le groupe a recommandé de poursuivre la mise en place d'une base de métadonnées avec des informations sur la quantité et la qualité des informations disponibles sur les pêcheries et la biologie.
- Le groupe a encouragé que la mesure des performances de la prédiction rétrospective, telle que décrite à la Section 5 du présent rapport, soit évaluée par des études de simulation, éventuellement avec des modèles conditionnés par d'antérieures évaluations de l'ICCAT. Si possible, une étude de cas devrait être réalisée pour un stock pour lequel il n'a pas été observé de séries de CPUE (standardisées) contradictoires.
- Le groupe encourage, une fois de plus, les CPC à fournir un accès limité aux données de CPUE opération par opération selon les nécessités et les priorités identifiées par les différents groupes d'espèces et les sous-comités. Ceci permettrait au SCRS de fournir une grande variété d'indices sur une échelle spatiale plus informative. Initialement, l'information sur une seule espèce et/ou par flottille spécifique pourrait être utilisée pour illustrer les avantages. On a suggéré d'avoir recours aux opportunités existantes du « nuage » maintenu par le Secrétariat pour le stockage et l'accès afin de faciliter les collaborations multilatérales. Cet échange pourrait avoir lieu en vertu de l'Accord de confidentialité de l'ICCAT, déjà mis en place.
- Le groupe a également considéré que le SCRS devrait continuer à participer à l'initiative SISAM de la CIEM afin de promouvoir davantage le travail collaboratif en développant des méthodologies d'évaluation, mettre en commun et développer les connaissances sur la façon de communiquer l'incertitude aux gestionnaires, d'encourager une collaboration plus étroite aux évaluations conjointes (p.ex. au sujet du requin-taube commun), de prendre aussi des initiatives pratiques, telles que la mise en commun de l'ordre du jour du groupe de travail sur les méthodes de l'ICCAT et de la CIEM. Cela fournira un accès libre normalisé et public aux données d'entrée et de sortie des évaluations, ainsi qu'une mise en commun de données plus détaillées. Ces éléments pourraient être renforcés à travers la participation au groupe de travail mondial sur les méthodes d'évaluation (GAME).

10. Adoption du rapport et clôture

Le rapport a été adopté pendant la réunion. Le coordinateur du WGSAM a remercié les organisateurs locaux pour l'excellente logistique de la réunion ainsi que les participants pour l'excellent travail réalisé. Le Secrétariat a réitéré ses remerciements à l'*Irish Sea Fisheries Board* (BIM) pour l'organisation exceptionnelle de la réunion et le chaleureux appui fourni aux participants. La réunion a été levée.

Références

- 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.

INFORME DE LA REUNIÓN DE 2014 DEL GRUPO DE TRABAJO SOBRE MÉTODOS DE EVALUACIÓN DE STOCKS (WGSAM)

(Dublín, Irlanda, 7 - 11 de abril de 2014)

1 Apertura de la reunión, adopción del orden del día y disposiciones para la reunión

La reunión se celebró en las oficinas del Irish Sea Fisheries Board (BIM) en Dublín, Irlanda, del 7 al 11 de abril de 2014. Las disposiciones locales fueron llevadas a cabo por el Dr. Michael Keatinge. La Dra. Pilar Pallarés, en nombre del Secretario Ejecutivo de ICCAT, expresó su agradecimiento al BIM por acoger la reunión y facilitar las disposiciones logísticas.

El Dr. Michael Schirripa, Relator del Grupo de trabajo sobre métodos de evaluación de stocks, presidió la reunión. El Dr. Schirripa dio la bienvenida a los participantes (el Grupo) y procedió a revisar el orden del día que fue adoptado sin cambios (**Apéndice 1**).

La lista de participantes se adjunta como **Apéndice 2**.

La lista de documentos presentados a la reunión se adjunta como **Apéndice 3**.

Los siguientes participantes actuaron como relatores de las diversas secciones del informe:

<i>Sección</i>	<i>Relatores</i>
1, 10	P. Pallarés
2	L. Kell
3	C. Porch
4	H. Arrizabalaga y G. Melvin
5	C. Brown
6	G. Díaz
7	M. Schirripa y V. Ortiz de Zárate
8-9	N. Abid

2 Examen de los métodos actuales de ICCAT para estimar la distribución de esfuerzo (EFFDIS)

El SCRS/2014/026 presentaba una breve revisión de los métodos anteriores de estimación del esfuerzo total de palangre, así como sugerencias para mejorar esta estimación, tal y como discutió el WGSAM en 2013. El cálculo más reciente de EffDIS se realizó en 2009 y utilizó las nueve especies principales de túnidos y especies afines de ICCAT para obtener capturas nominales globales de Tarea I (en peso) y las CPUE a partir de estadísticas parciales de captura y esfuerzo (Tarea II). El supuesto base del modelo actual considera que las tasas de captura son equivalentes a nivel parcial y global. Al comparar los resultados con estimaciones anteriores (obtenidas durante reuniones intersesiones del Subcomité de ecosistemas en 2007 y 2008), los resultados globales no mostraban diferencias importantes. Sin embargo, a niveles más disgregados las diferencias eran mayores para algunos pabellones. En la mayoría de los casos, las importantes variaciones relativas se suelen asociar con diferentes mejoras y correcciones de algunos conjuntos de datos específicos. La distribución geográfica global muestra un pequeño incremento cerca de las aguas de Venezuela debido a las diferentes correcciones realizadas en las distribuciones espaciales de estadísticas de captura y esfuerzo de Venezuela durante varios años. Se presentan campos recomendados para mejorar la estimación de EffDIS con el fin de impulsar los debates y finalizar los métodos que deberían usarse en el futuro.

Se reconoce la importancia de tener en cuenta los cambios en las operaciones pesqueras y en las características de las flotas principales de cada CPC que opera en la zona del Convenio de ICCAT ya que estas afectan a la eficacia de las flotas a la hora de capturar especies objetivo y especies de captura fortuita. La documentación de estos cambios tecnológicos y de estrategia es especialmente importante para entender los informes nacionales de captura y esfuerzo que se presentan anualmente (Tarea II-CE). El Subcomité de Ecosistemas cuenta también con varias recomendaciones relacionadas, por ejemplo con i) restricciones y/o limitaciones potenciales de los datos y la información, ii) descripción de la incertidumbre, por ejemplo, sustituciones, ratios de extrapolación y proporción de flotas sin clasificar ("otras"), iii) deberían explorarse métodos adicionales de extrapolar los datos y iv) deberían explorarse métodos como la validación cruzada.

El WGSAM discutió el uso de EffDIS y los temas relacionados con su estimación. En particular, ¿es el actual nivel temporal y espacial de agregación (es decir, cuadrículas de 5 grados y mes) adecuado para su uso previsto? esto podría evaluarse llevando a cabo estudios con diferentes niveles de agregación y comparando las conclusiones obtenidas con el conjunto de datos actual. Asimismo, era importante la clasificación de las flotas y también se pensó que sería necesario apoyo externo para ayudar a realizar el EffDIS en el futuro.

EffDIS es un recurso importante para el Subcomité de Ecosistemas y en 2013, el Subcomité de Ecosistemas recomendó que se hicieran esfuerzos para desarrollar estimaciones EffDIS similares para los artes de cebo vivo y cerco. Esto resultaría especialmente útil para las evaluaciones de medidas de ordenación de vedas espacio-temporales. Será importante identificar qué medida de esfuerzo global utilizar para estos dos artes.

A pesar de la importancia de EffDIS para el trabajo del SCRS, la última evaluación se llevó a cabo en 2007. Esto se debe principalmente a varias cuestiones metodológicas que aún deben ser resueltas (véase SCRS/2013/036 para un resumen completo) y a la falta de recursos para llevar a cabo el análisis. En concreto, es necesario validar el supuesto utilizado para construir la base de datos que está agregada a nivel de cuadrículas de 5 grados) y evaluar si conviene. Por lo tanto, el Grupo recomendó que:

- 1) Se exploren métodos para clasificar las flotas y
- 2) Se necesita apoyo externo para producir EffDIS en el futuro.

3 Cuantificación de la incertidumbre en las evaluaciones de ICCAT

La Comisión espera un asesoramiento sobre medidas de ordenación basado en el riesgo, tal y como se establece en la matriz de estrategia de Kobe II y en su marco de toma de decisiones [Rec. 11-13]. Un aspecto importante a la hora de proporcionar dicho asesoramiento científico es cuantificar adecuadamente la incertidumbre sobre el estado del stock y las perspectivas futuras en escenarios de futuras opciones de ordenación. Con la llegada de modelos de evaluación de stock altamente parametrizados y más comúnmente aplicados, la inversión en informática para cuantificar la incertidumbre sobre el estado del stock y las perspectivas futuras es bastante elevada. Otras OROP de túnidos han tenido experiencias similares y se están aplicando una serie de aproximaciones para cuantificar ambos procesos y la incertidumbre por observación con el fin de desarrollar asesoramiento en materia de ordenación basado en el riesgo. Se pidió por tanto al WGSAM que proporcionara orientación sobre la evolución y la posibilidad de armonizar métodos para caracterizar la incertidumbre en los diferentes grupos de especies. En este sentido, el WGSAM indicó que todos los modelos de evaluación representan simplificaciones de un sistema pesquero más complejo y, por tanto, están intrínsecamente limitados en su capacidad de reflejar la variabilidad inherente de dicho sistema. En otras palabras, los gestores no deben esperar que la incertidumbre en el asesoramiento científico pesquero sea nunca completamente cuantificada. No obstante, el Grupo convino en que pueden hacerse mejoras en la forma en que se representa la incertidumbre en los diversos grupos de especies.

El Grupo identificó tres enfoques básicos para representar la incertidumbre en los resultados de los modelos de evaluación de stock: (1) un enfoque "basado en el modelo" que tenga en cuenta explícitamente las principales fuentes de incertidumbre percibidas utilizando una plataforma de modelación única y versátil, (2) promediando modelos, es decir, combinando ensayos alternativos de diferentes plataformas de modelación y (3) un enfoque "empírico" que utilice información histórica existente sobre la coherencia de los resultados de los modelos de evaluación de stock a lo largo del tiempo. El primer enfoque ha sido el más común en la práctica. Habitualmente, se desarrolla un solo modelo de base que incluye los parámetros estimables que representan lo que se percibe como las fuentes clave de la incertidumbre (con o sin distribuciones previas informativas). Posteriormente, se calculan las medidas de la incertidumbre en los parámetros clave de ordenación mediante métodos estándar (hessiano inverso, bootstrap, integración bayesiana) y se incorporan en las matrices de Kobe. Pueden ejecutarse los ensayos de sensibilidad o las plataformas alternativas del modelo para obtener más información sobre posibles incertidumbres, pero a menudo se usan principalmente para cualificar el asesoramiento del caso base del modelo y no se incorporan directamente en las matrices de Kobe. La eficacia de este enfoque depende de la capacidad para modificar el modelo de evaluación de una forma que aproxime razonablemente la fuente de incertidumbre.

Una segunda estrategia para incorporar la información facilitada por modelos alternativos, que se ha usado a veces, es la promediación de modelos, en la que se asignan ponderaciones por frecuencia a cada posible modelo, basándose tal vez en alguna medida del ajuste de los datos (por ejemplo, AIC, ponderación por la inversa de la varianza) o en la opinión de los expertos (véase también la sección 5). Una desventaja es que este enfoque tendrá

probablemente como resultado una distribución multimodal de las medidas del estado del stock que se están considerando (por ejemplo, punto de referencia límite) o requerirán supuestos acerca de su forma de distribución (por ejemplo, normal o lognormal). El Grupo consideró que el éxito de cualquier enfoque basado en un modelo depende enormemente de la capacidad del grupo de evaluación para desarrollar distribuciones previas informativas o postular hipótesis alternativas y razonables antes de descubrir las implicaciones en cuanto a ordenación de estas hipótesis. Si no, existe el peligro de producir estimaciones sesgadas tanto de la tendencia central como de la incertidumbre descartando modelos considerados no informativos o introduciendo intencionadamente sesgos añadiendo modelos o alterando las distribuciones previas. Se mencionó que el Centro nacional de huracanes de Estados Unidos utiliza la promediación de modelos de forma rutinaria y deriva las ponderaciones para cada modelo candidato de su rendimiento en pasadas predicciones (es decir, comparando las predicciones del modelo con las trayectorias reales del huracán). Lamentablemente, a diferencia de las trayectorias de los huracanes, rara vez se conoce el verdadero estado de un stock de peces.

Los enfoques empíricos examinan el rendimiento pasado del modelo para deducir los rangos de la incertidumbre total. De forma ideal, se compararía el rendimiento del modelo con la verdadera disposición del stock, pero una vez más, en la práctica esto no es posible. Por otro lado, se podría comparar el rendimiento de dos evaluaciones de stocks históricas de referencia durante cada año que tienen en común. La incertidumbre total resultante incluiría implícitamente errores de estimación "dentro del modelo", así como errores sistémicos que podrían producirse con cambios en los modelos o en las filosofías del equipo de evaluación. Este enfoque ha sido ya adoptado por el Consejo de ordenación pesquera del Pacífico en Estados Unidos (véase Ralston et al. 2011) y se está considerando en otros consejos. Durante las discusiones del Grupo de trabajo se plantearon algunos posibles problemas con el método empírico. Se expresó la inquietud de que dicho enfoque no reflejaría la tendencia de las evaluaciones de stock de mejorar en precisión y exactitud con el tiempo mediante el añadido de nuevos datos, mejores modelos y una mejor recopilación de datos. Sin embargo, otros en el grupo de trabajo señalaron que las mejoras en los datos solo pueden disminuir parte de la incertidumbre y que la composición de los científicos del equipo de evaluación de stock, la composición del panel de revisión de la evaluación y los cambios en los métodos de evaluación de stock podrían tener la misma influencia en la coherencia de los resultados de la evaluación de stock. Se sugirió que, al menos, los análisis empíricos podrían utilizarse para ayudar a proporcionar una base real para las estimaciones de varianza derivadas de los actuales procedimientos basados en modelos.

Se realizó una presentación titulada "Cuantificando la incertidumbre debida al procesamiento de los datos en las evaluaciones de stocks estructuradas por edad". Se presentó el trabajo preliminar detallando un método para asignar los datos de talla faltantes a partir de observaciones basadas en su proximidad espacio-temporal (asumiendo que las observaciones eran aproximadamente multivariantes normales con covarianzas estimadas). Posteriormente, se utilizaron las distribuciones estimadas para generar muchos conjuntos de datos de talla completos (muestreando los parámetros asignados al modelo y posteriormente muestreando a partir de las observaciones de talla ponderadas), que fueron convertidos a distribuciones de edad utilizando diversos métodos de determinación de la edad (por ejemplo, separación de cohortes y claves de edad-talla). Cada conjunto de datos de edad se usó posteriormente en un VPA con diferentes tasas de mortalidad natural e índices de abundancia relativa. Los resultados preliminares sugieren que la atribución de tallas con un tamaño de muestra efectivo de 100 es demasiado baja para generar cualquier variabilidad en los conjuntos de datos de captura por talla derivados. La tasa de mortalidad natural (80%, 100% y 120% de los niveles de la evaluación más reciente) determina enormemente la F_{RMS} estimada a partir del análisis XSA.

El Grupo se mostró de acuerdo en que las variaciones en este enfoque serían una forma útil de intentar cuantificar cómo la incertidumbre en los datos se convertiría en incertidumbre en la evaluación. Dicho enfoque ayudaría también a identificar la importancia relativa de una fuente de incertidumbre sospechada y la correspondiente necesidad de incorporarla explícitamente al desarrollar las matrices de Kobe utilizadas para formular el asesoramiento científico a los gestores. Durante la discusión de evaluaciones de estrategias de ordenación, se señaló también que los modelos operativos deberían construirse teniendo en mente las mayores fuentes de incertidumbre percibidas y que las mismas podrían usarse para ayudar a cuantificar el alcance de propagación de estas incertidumbres en la incertidumbre del asesoramiento científico (que, de hecho, es una variación del enfoque discutido más arriba). Se indicó también que los análisis de riesgo como los llevados a cabo para el atún rojo (Leach et al. 2014) son útiles para identificar las principales fuentes de incertidumbre percibidas por los científicos y otras partes interesadas y que podrían tener, además, el beneficio añadido de conseguir una mayor aceptación cuando se presentan los resultados.

4 Descripción de la calidad de los datos pesqueros y la información biológica

Se presentó el documento SCRS/2014/035 con el objetivo de revisar la información de contexto y fomentar las discusiones sobre este punto del orden del día. La Resolución [13-15] para la estandarización de la información científica en el informe anual del SCRS requiere "...calificar la calidad de los datos de las pesquerías y de los datos relacionados con los conocimientos sobre las especies (por ejemplo, parámetros biológicos, datos históricos de patrones de distribución de la pesquería, selectividad) utilizados como entradas para las evaluaciones de stock. Las calificaciones cualitativas de los supuestos y datos de entrada podrían ser detalladas y deberían resumir el estado de los conocimientos de las diferentes entradas...". El Informe de 2013 del Subcomité de estadísticas proponía un método para calificar los diferentes elementos de datos. El método consistía en una tabla con una lista de categorías y puntos (dentro de una categoría) que eran evaluados y ponderados (por ejemplo, de acuerdo con la importancia relativa de los puntos dentro de cada categoría, en la evaluación de stock). La calificación de cada punto era el producto de su valor y su ponderación, y la calificación de cada categoría era la suma de las diferentes puntuaciones de cada punto dentro de una categoría. Se sugirió calcular la calidad de la información para una determinada evaluación de stock como la clasificación media de las categorías. A la propuesta le faltaban, no obstante, criterios predefinidos para basar la clasificación, como los utilizados en los Catálogos de datos de especies de la IOTC, en los que los límites de las puntuaciones están determinados por afirmaciones cualitativas acerca de la disponibilidad/calidad de los datos (por ejemplo, "disponible de acuerdo con los estándares", "no disponible de acuerdo con los estándares" o "no disponible en absoluto").

Por tanto, el SCRS/2014/035 proponía un marco genérico, algo similar a la propuesta del SC de estadísticas, pero con directrices de puntuación predefinidas y cualitativas. La puntuación de calidad propuesta (Q) de los diferentes conjuntos de información era:

$$Q=D*I$$

donde:

D=disponibilidad y calidad de los datos
Con

- 1= no hay datos o son de calidad mala/desconocida
- 2=si existen datos, pero la calidad no es muy buena
- 3=si existen datos de buena calidad

I=impacto de los datos en los resultados de la evaluación (cuán importante es disponer de este conjunto de información).

Con:

- 1=escaso impacto en los resultados de la evaluación
- 2=impacto moderado en los resultados de la evaluación
- 3=impacto elevado/desconocido en los resultados de la evaluación

El documento incluía varios ejemplos en los que se aplicaba provisionalmente la puntuación a diferentes stocks de atún blanco y listado. Se puntuaron los diferentes puntos relacionados con las pesquerías y la biología y la puntuación total de cada stock reflejaba la suma de los valores de Q en los diversos puntos.

El Grupo consideró que este marco genérico era útil como punto de partida y debatió varios puntos sobre ello. Se llegó al acuerdo de que las puntuaciones deberían basarse en una metodología objetiva, cuantificable y científicamente defendible. Aunque esto puede ser más fácil de lograr cuando se trata de puntuar la disponibilidad de datos, se reconoció que medir objetivamente la calidad de los datos podría ser algo más difícil. Sin embargo, el grupo sugirió mecanismos para avanzar en la puntuación en la medida de lo posible.

En esta etapa, una consideración importante era el propósito o utilidad de este ejercicio. Por una parte, es necesario responder a la Resolución 13-15 con el fin de que la Comisión cuente, además de la matriz de Kobe, con información acerca de la calidad de los datos utilizados para derivar la matriz de Kobe. Con este fin, se consideró alguna representación simple (similar a los ejemplos aportados en el SCRS/2014/035). No obstante, el grupo indicó que las puntuaciones de calidad podrían usarse también para otros fines. Por ejemplo, podrían ser útiles para que los grupos de trabajo científicos describan las principales fuentes de incertidumbre y para mejorar, en la medida de lo posible, la formulación del asesoramiento científico (por ejemplo, seleccionando modelos adecuados y/o ponderando diferentes escenarios).

El Grupo constató que la importancia de los diferentes conjuntos de datos (puntuaciones "I") depende del modelo usado, así como del ciclo vital del stock. Por tanto, las puntuaciones totales "Q" podrían no ser comparables entre stocks. Para solucionar esto, el grupo recomendó cuantificar la ratio entre la puntuación total y la máxima puntuación posible teniendo en cuenta el vector I específico para cada stock, lo que permitiría alguna comparación basada en colores entre los stocks.

Por último, el Grupo recomendó desarrollar una base de metadatos que contenga información sobre la calidad y cantidad de los datos pesqueros y biológicos disponibles para las evaluaciones. Esto, junto con el mecanismo existente para evaluar los índices de abundancia independientes de la pesquería y de las CPUE, ayudaría a puntuar de forma más objetiva la calidad y cantidad de la información relacionada con la pesquería que se utiliza en la evaluación de las especies de ICCAT con el fin de aportar información a los científicos que hacen la evaluación de stock y proporcionar a la Comisión una base para que tenga una perspectiva global de la calidad de los datos, es decir, el valor de las puntuaciones "D". El Grupo discutió la posible estructura de una base de metadatos que puede posteriormente ser alimentada y actualizada a medida que se disponga de nueva información.

La base de metadatos constaría de tres componentes principales, uno sobre datos pesqueros, otro sobre datos biológicos y un tercero con estudios de marcado recaptura (**Apéndices 4 y 5**, respectivamente). La estructura sugerida es flexible y podría incorporar nuevos campos a medida que se prevea su necesidad. En el caso de los metadatos biológicos, se consideró que serían ya útiles los campos incluidos en "Información general", además de alguna información acerca de qué parámetros y cuán bien estimados están. Los campos incluidos en "estimaciones de parámetros" permitirían llevar a cabo metaanálisis que podrían ayudar de diversas formas, por ejemplo, describiendo la incertidumbre sobre determinados parámetros. No obstante, llenar estos campos sería, por supuesto, una intensa labor. El grupo consideró también que, en casos particulares, podría ser útil tener acceso a los datos en bruto de algunos estudios. La base de metadatos permitiría identificar fácilmente las fuentes de estos datos en bruto.

5 Conciliar los resultados cuando se utilizan múltiples métodos de modelación

Las evaluaciones del SCRS a menudo, y cada vez más, han incluido el uso de múltiples métodos de modelación (es decir, tipos diferentes de modelos, hipótesis alternativas) para estimar el estado del stock en relación con los elementos de referencia de conservación de ICCAT. El Grupo discutió las diversas razones de por qué se hace esto. Por ejemplo, los tipos de modelos pueden diferir en los supuestos subyacentes, y los participantes en la evaluación pueden no ser capaces de determinar cuál es más adecuado. En dichos casos, se prevé que incorporar los resultados de múltiples modelos en el asesoramiento sobre ordenación refleja mejor la incertidumbre en los resultados. En otros casos, las configuraciones alternativas de la misma plataforma o tipo de modelación pueden ser consideradas lo suficientemente plausibles para incluirlas en los resultados. El Grupo se mostró de acuerdo en que este enfoque tiene sus puntos buenos, pero reconoció que podría haber dificultades para determinar cuál es la forma más adecuada de combinar dichos resultados, lo que incluye asignar una ponderación adecuada, especialmente en aquellos casos en los que los diversos resultados del modelo pueden entrar en conflicto.

No hay ningún documento SCRS disponible sobre este tema. Sin embargo, el Grupo examinó el trabajo reciente de Deroba et al. (2014) empleando la simulación para probar la robustez de los modelos de evaluación de stock ante el error. Las conclusiones de este documento que están relacionadas con este tema incluyen:

- Las mayores diferencias se produjeron al comparar los resultados de diferentes tipos de modelos (modelo de producción excedente, basado en la edad, etc.); los modelos tendían a funcionar de forma similar a otros modelos del mismo tipo.
- La autoprueba es útil y recomendable, y la autoprueba y la prueba cruzada frecuentemente destacan la divergencia en los años más recientes de la serie temporal.
- Entre los modelos, la variabilidad puede ser considerada como un tipo de incertidumbre en la evaluación; esto tiene implicaciones al considerar si basar el asesoramiento sobre ordenación solo en un modelo configurado para ser el "mejor ajuste" para la evaluación o si incorporar los resultados de diferentes modelos (por ejemplo, la aplicación de la promediación de modelos).

Las conclusiones de Deroba et al. (2014) eran coherentes con las experiencias de las evaluaciones de stock del SCRS y proporcionan una confirmación analítica de la importancia de la selección del modelo (especialmente respecto al tipo de modelo), el examen de diagnósticos y la validación de los resultados, y las implicaciones de integrar o no los modelos en el asesoramiento final.

El Grupo consideró que un componente inicial esencial a la hora de reconciliar diferentes resultados del modelo es el proceso de selección de los modelos más adecuados a partir de los que desarrollar el asesoramiento en materia de ordenación. Este proceso debería iniciarse durante la reunión de preparación de datos, donde debe reservarse un tiempo para determinar qué modelos se usarán en la evaluación. Tal y como se ha descrito anteriormente, deberían considerarse para los diversos modelos los requisitos de datos, los supuestos y la capacidad de reflejar cambios importantes en la pesquería que se sabe que han ocurrido, teniendo en cuenta la calidad y cantidad de los datos disponibles. La implicación de los científicos de las CPC es clave en este proceso, ya que son los más familiarizados con los cambios en las diversas pesquerías que podrían merecer consideración en el proceso de evaluación, y están preparados para identificar cualquier inquietud con la adecuación de conjuntos de datos particulares para su uso en los diversos modelos. También debería considerarse el rendimiento de los diversos modelos utilizados en la evaluación anterior.

Durante la reunión de evaluación, debería asignarse tiempo suficiente para realizar un cuidadoso examen de los diagnósticos y el rendimiento del modelo. Las mediciones del rendimiento podrían incluir una evaluación de si los resultados parecen coherentes con lo que se conoce de la biología de la especie, así como de las tendencias pesqueras (un ejemplo de un enfoque así, "medición del rendimiento de la predicción retrospectiva", se detalla a continuación). Deberían considerarse la probabilidad y las consecuencias de la infracción de supuestos para cada modelo teniendo en cuenta la información disponible sobre la población y la pesquería. Una posible consecuencia de este proceso es que el modelo identificado durante la reunión de preparación de datos para su uso en la evaluación podría ser rechazado (adecuadamente) para su inclusión en el asesoramiento final en materia de ordenación.

El Grupo discutió si sería adecuado o no considerar los resultados de modelos más simples junto con los resultados de modelos más complejos (por ejemplo, usando los resultados de un modelo de producción excedente y los de un modelo más plenamente integrado, como stock synthesis). En otras palabras, si los datos fueran suficientes para ejecutar los modelos más complejos (que, en teoría, podrían tener en cuenta más aspectos y cambios en los parámetros del stock y la pesquería), ¿debería, en su lugar, basarse el asesoramiento solo en los resultados de los modelos más complejos? El Grupo consideró que podría haber circunstancias en las que es adecuado considerar los resultados de dichos tipos de modelos divergentes.

El Grupo debatió las opciones para procedimientos o criterios adicionales para la selección de los modelos. Se sugirió uno de estos procedimientos, basado en la medición del "desempeño de la predicción retrospectiva". La idea de este método es similar a la del análisis retrospectivo tradicionalmente utilizado en los VPA, pero ampliándola a la comparación de modelos múltiples. Por ejemplo, partiendo del supuesto de que se utilizan dos modelos en la evaluación que arrojan dos resultados diferentes. Las fases del proceso de medición de la predicción retrospectiva serían:

- truncar los datos disponibles, eliminando los xx años más recientes (en función de lo que sea apropiado para cada stock);
- volver a ejecutar cada modelo para obtener una estimación de la dinámica de la población durante xx años del pasado;
- proyectar el modelo hacia adelante utilizando las capturas reales hasta el año actual;
- calcular una tendencia de la CPUE (o quizá solo la tendencia de la biomasa estimada, lo que proceda en función de la complejidad del modelo) que se predice para los años proyectados hasta el año actual y
- comparar dicha tendencia de la CPUE con la(s) tendencia(s) real(es) de la CPUE (estandarizada) observadas.

Con este enfoque, pueden dejar de considerarse algunos modelos si los resultados de los análisis demuestran que existen incoherencias con la(s) tendencia(s) de la CPUE observada(s). El método puede ampliarse más para desarrollar ponderaciones para los diferentes resultados del modelo (que puede utilizarse cuando se combinan resultados entre modelos) mediante la definición de una medida de discrepancia entre las series de CPUE predichas y observadas (por ejemplo, suma de los residuos cuadrados dividida por SE de la CPUE observada). Se constató que este enfoque no requiere que los modelos utilicen los mismos conjuntos de datos. La aplicabilidad de este enfoque depende de que la(s) tendencia(s) de la CPUE real observada (estandarizada) refleje(n) adecuadamente la población y la ausencia de índices de CPUE en conflicto. El Grupo consideró que este enfoque era prometedor, y recomendó que continúen los trabajos de investigación relacionados con la idoneidad de su utilización en la selección del modelo y/o en la ponderación.

El Grupo constató que también es importante la selección de los modelos que se van a utilizar para las proyecciones. En algunas circunstancias, para realizar las proyecciones puede utilizarse un modelo diferente al utilizado para estimar el estado actual del stock. Esto puede suceder, por ejemplo, en los casos en los que el

modelo utilizado para estimar el estado del stock no puede utilizarse para las proyecciones o requiere un tiempo de procesamiento adicional (más del disponible durante la reunión de evaluación) para producir algunos elementos de salida que se requerirían para la proyección. El Grupo indicó que había surgido cierta inquietud en relación con la idoneidad de este enfoque. El Grupo identificó la necesidad de desarrollar enfoques para investigar esta cuestión. Se sugirió un enfoque potencial que implicaría volver a ejecutar el modelo utilizado para estimar el estado actual del stock, eliminando algunos de los años recientes de datos (como en el análisis retrospectivo). Entonces podría ejecutarse el modelo considerado para las proyecciones, proyectándose a partir del último año del ensayo del modelo retrospectivo, utilizando el historial de captura real hasta el año en curso. Los resultados del modelo de proyección podrían evaluarse mediante una comparación de la condición/estado proyectado en el año actual con la condición/estado estimado mediante el modelo de evaluación.

Después de que los científicos que realizan la evaluación se pongan de acuerdo sobre los modelos que se tienen que incluir para desarrollar el asesoramiento en materia de ordenación (por ejemplo, estimaciones del estado del stock con respecto a los puntos de referencia biológicos e incertidumbre asociada, matriz de estrategia de Kobe), queda la cuestión de la forma en que se tienen que presentar los resultados. Una opción que generalmente es la preferida sería combinar en cierto modo estos resultados para proporcionar niveles de referencia y representaciones adecuadas de la incertidumbre. Esto resulta particularmente difícil cuando los diversos resultados de los modelos entran en conflicto. El Grupo manifestó que resulta generalmente inapropiado combinar resultados que son totalmente incoherentes entre cada modelo, por ejemplo, cuando los diferentes modelos están estructurados basándose en dos hipótesis alternativas que reflejan estados asumidos muy diferentes, con poco o ningún solapamiento en las distribuciones de probabilidades. En dichos casos, combinar dichos resultados del mismo modo produciría estimaciones del estado del stock y/o tendencias promediadas entre las tendencias, lo que no es coherente con ninguna de las dos hipótesis alternativas. En dichos casos, se insta a los científicos a examinar con rigor los supuestos subyacentes de cada hipótesis y a considerar detenidamente si cada uno de ellos está respaldado por los datos existentes y la ciencia actual. Si este proceso resulta insuficiente para descartar una hipótesis, los científicos deberían considerar si se pueden asignar las probabilidades relativas de cada hipótesis. Sería aconsejable realizar más trabajos de investigación para recopilar datos que contribuyan a la aceptación/descarte de las hipótesis o a asignar probabilidades.

La práctica general del SCRS cuando se combinan resultados de múltiples modelos es ponderar igualmente los resultados de cada modelo. Tal y como se mencionó en la sección 3 de este informe, las ponderaciones podrían asignarse a los resultados de cada modelo basándose en alguna medición del ajuste a los datos (por ejemplo, AIC; ponderación de variación inversa). Además de las preocupaciones descritas en la sección 3, este enfoque es limitado en el sentido de que sólo es apropiado cuando los modelos potenciales utilizan los mismos conjuntos de datos.

El Grupo también debatió el potencial de utilizar la evaluación de estrategias de ordenación para aportar información a la ponderación/selección del modelo. El Grupo identificó un problema potencial en este enfoque en el sentido de que el modelo operativo de la MSE podría influir en la percepción del buen funcionamiento del modelo. Se sugirió que la utilización de modelos de producción excedente podría resultar problemática si las funciones de productividad presentan una fuerte asimetría hacia la izquierda y la productividad asciende a medida que los niveles del stock merman.

El Grupo reconoció que en un esquema de ponderación es inevitable que exista un elemento de subjetividad. Incluso la asignación por defecto de ponderación igual es implícitamente una decisión que podría dar más peso que el que se merece a algunos resultados, si la totalidad de los factores, como los supuestos del modelo, capacidad, calidad de los datos, tendencias de la pesquería y biología, pueden tenerse en cuenta de un modo objetivo. Por tanto, el Grupo considera que la opinión experta podría ser apropiada para la asignación de ponderaciones, si está respaldada por el conocimiento de estos factores y la postulación de mecanismos razonables que podrían respaldar una conclusión de que es más probable que ciertos modelos reflejen mejor la situación real que otros.

Otra práctica común en las evaluaciones del SCRS es descartar en última instancia para su inclusión en el asesoramiento en materia de ordenación varios modelos que producen resultados plausibles pero que podrían ser considerados en cierto modo menos probables. El Grupo constató que se había manifestado la preocupación con respecto a que la práctica existente de seleccionar únicamente un subconjunto de modelos considerados más probables en los resultados finales podría dar lugar a una subestimación de la gama de incertidumbre. Se requieren más trabajos de investigación.

6 Puntos de referencia límite, normas de control de la captura y evaluaciones de estrategias de ordenación

La evaluación de puntos de referencia límite (LRP) y normas de control de la captura (HCR) mediante el uso de evaluación de estrategias de ordenación (MSE) cuenta cada vez más con el reconocimiento de las OROP mundiales de túnidos como un medio eficaz de avanzar en el proceso de ordenación de pesquerías. Las evaluaciones de 2013 de atún blanco y pez espada se utilizaron como ejemplos de cómo un proceso MSE podría incluirse formalmente en la ordenación de estos stocks. El WGSAM prevé continuar con este esfuerzo mediante: (1) la continuación de la mejora de los métodos en el marco del proceso MSE; (2) la introducción de MSE en más evaluaciones cuando y donde proceda; y 3) el fomento de líneas de comunicación que mantengan informados a los gestores de sus beneficios y de sus puntos débiles. En lo que concierne al diálogo y la comunicación, hace poco ICCAT ha adoptado la *Recomendación de ICCAT para mejorar el diálogo entre los gestores y científicos pesqueros* [Rec. 13-18] que tiene como objetivo mejorar la comunicación y favorecer el entendimiento mutuo entre los científicos y gestores pesqueros con el fin de facilitar una toma de decisiones basada en la ciencia y simplificada. Asimismo, la Recomendación describe tareas específicas que se tienen que realizar durante la primera reunión del Grupo de trabajo permanente para mejorar el diálogo entre los gestores y científicos pesqueros (SWGSM) en 2014.

Se ha adoptado una norma de control de la captura (HCR) empírica para el atún rojo del sur (SBT) para establecer el total admisible de capturas (TAC). La HCR se basa en cambios de un año a otro en los índices a partir de un índice de CPUE de abundancia de adultos dependiente de la pesquería y de una prospección aérea de juveniles independiente de las pesquerías. Antes de que pueda implementarse la HCR, se deben seleccionar niveles de referencia apropiados para la captura y los índices; y los parámetros HCR deben adaptarse para cumplir los objetivos de ordenación mediante la evaluación de estrategias de ordenación (MSE).

En el documento SCRS/2014/036 se llevó a cabo una MSE preliminar para el atún rojo del Mediterráneo utilizando la HCR del SBT como parte de un procedimiento de ordenación (MP, a saber, la combinación de datos predefinidos, con un algoritmo al que se asignan los datos para proporcionar un valor para un TAC o una medida de control del esfuerzo). Los próximos pasos serán:

- 1) Identificar objetivos de ordenación y confrontarlos con las mediciones del desempeño con el fin de cuantificar en qué medida se han alcanzado.
- 2) Seleccionar hipótesis sobre la dinámica del sistema.
- 3) Modelos operativos (OM) en función de los datos y conocimientos, así como para el posible descarte y ponderación de las diferentes hipótesis.
- 4) Identificar estrategias de ordenación potenciales y codificarlas como procedimientos de ordenación.
- 5) Proyectar los modelos operativos hacia el futuro utilizando los MP como procedimientos de control de respuesta, y
- 6) Llegar a un acuerdo sobre los MP que cumplen mejor los objetivos de ordenación.

El Grupo resaltó la necesidad de seleccionar detenidamente los supuestos de los modelos operativos dado que los supuestos erróneos pueden afectar a los resultados de las evaluaciones de stock y de las evaluaciones de estrategias de ordenación (MS). Se debatió la importancia de la selección de la gama de escenarios utilizados en el modelo operativo (Fromentin et al., 2014; Leach et al., 2014). El Grupo acordó que en un marco MSE, el objetivo no es escoger el procedimiento de ordenación o las normas de control de la captura (HCR) que proporcionen los mejores resultados, sino los que sean más robustos en los escenarios escogidos. Por ejemplo, si el modelo operativo se ejecuta con cuatro niveles de inclinación, que se asume que son igualmente plausibles, la HCR elegida debería ser la que alcance los objetivos de ordenación con estos cuatro niveles de inclinación. El Grupo también debatió el hecho de que muchos modelos operativos asumen que la regulación de la población solo se produce en el reclutamiento, sin embargo, el modelo operativo debería captar también otras posibilidades de regulación (por ejemplo, mortalidad natural).

El Grupo acordó que el modelo operativo debería ser más complejo que el modelo de estimación. Por ejemplo, un modelo de producción excedente puede utilizarse como el modelo de estimación mientras que un modelo más complejo, como SS3, puede utilizarse como modelo operativo. El documento presentaba un ejemplo en el que se utilizaron dos prospecciones como índices para reclutas y adultos. Se debatió la posibilidad de no tener que realizar estos tipos de prospecciones cada año para obtener índices de biomasa que aporten información a la HCR. Se indicó al Grupo que utilizar una HCR podría reducir los requisitos de datos en comparación con los datos requeridos actualmente en algunas evaluaciones de stock. El Grupo también debatió si la HCR debería incluir puntos de referencia límite más precautorios para aquellos casos en los que el índice de biomasa estimado a partir de las prospecciones presenta una mayor incertidumbre.

Se revisó la presentación que el Presidente del SCRS realizó a la Comisión en 2013. El Grupo resaltó la importancia de la próxima primera reunión del Grupo de trabajo permanente dedicado al diálogo entre los gestores y científicos pesqueros (SWGSM), durante la cual se debatirán cuestiones relacionadas con los puntos de referencia límite, probabilidades asociadas con la matriz de Kobe, periodos de recuperación para los diferentes stocks, etc. El Grupo resaltó que la HCR no podría desarrollarse sin este diálogo entre científicos y gestores.

En el documento SCRS/2014/025 se describe cómo los autores buscan evidencias de una relación stock-reclutamiento para el atún rojo, rabil y atún blanco. Los autores llegaron a la conclusión de que la evidencia de la existencia de una SRR para cualquiera de los stocks era débil y que los datos podrían respaldar también otras hipótesis (Vert-pre, 2014), por ejemplo, la de Gilbert (1998) de que el reclutamiento fluctúa en torno a un nivel medio durante un periodo de tiempo y luego se produce un cambio de régimen. Esto tiene obvias implicaciones para la evaluación de stock y el asesoramiento en materia de ordenación.

El Grupo debatió que la hipótesis de “cambio de régimen” presentada en el documento podría ser un producto de los modelos de estimación. Dicho de otro modo, aunque el “cambio de régimen” no es incoherente con la ausencia de una SRR observada en los stocks mencionados, esto, en sí mismo, no es una prueba de que se haya producido realmente un “cambio de régimen”.

En el documento se sugería que la SSB es una función de reclutamiento (en vez de que el reclutamiento sea una función de la SSB). El Grupo preguntó por qué parecía haber autocorrelación en los patrones de reclutamiento. Se debatió el hecho de que la autocorrelación observada podría deberse a la utilización de separación de cohortes para asignar edades a los stocks estudiados o a otros supuestos del modelo.

El documento sugería que a efectos de las proyecciones del estado del stock, debería darse más peso a los reclutamientos recientes observados en vez de ponderar todos los patrones de reclutamiento de la serie temporal del mismo modo (dicho de otro modo, es más probable que el reclutamiento de mañana sea más similar al de ayer que al reclutamiento de hace 25 años).

En el documento SCRS/2014/037 se utilizaron las series de SSB y de R procedentes de los escenarios del caso base para stock de atún rojo del Atlántico este y Mediterráneo, utilizando capturas “infladas” y “comunicadas” para ajustar tres modelos diferentes S-R (Beverton & Holt, Ricker and smooth Hockey-stick). Los resultados muestran que el stock mantuvo toda su capacidad reproductiva a lo largo de la serie temporal, desde 1950 y que la relación S-R de Ricker no puede descartarse para este stock. En el documento se afirmaba que este último resultado añade más incertidumbre a la estimación de B_{01} , lo que hace que no pueda utilizarse como punto de referencia de biomasa para este stock. Como alternativa el documento propone la utilización de B_{loss} como B_{lim} y después estimar B_{pa} a partir de dicho valor. Esto permitió a los autores simular el comportamiento de algunas HCR y seleccionar la más adecuada para este stock.

El Grupo debatió el hecho de que ICES está dejando de lado el marco de punto de referencia límite, que es lo contrario a lo que se propone en el documento. Al margen de esto, la utilidad del enfoque propuesto en el documento puede probarse mediante simulación. También se debatió la validez de establecer supuestos en los modelos operativos para que sean más precautorios (por ejemplo utilizando una relación de reclutamiento en forma de palo de hockey). En líneas generales, el Grupo manifestó que no estaba de acuerdo con este enfoque. El Grupo también debatió los factores que podrían contribuir a relaciones S-R de tipo Ricker en los atunes, pero no llegó a una conclusión final.

7 Incorporación de información sobre ecosistemas, clima y hábitat (ECH) en las evaluaciones de stock

Los debates sobre la incorporación de información sobre ecosistema, clima y hábitat (ECH) en las evaluaciones de stock se centraron en los dos enfoques que pueden utilizarse para mejorar el proceso de evaluación de stock de las especies de ICCAT. Se identificaron dos enfoques, uno cualitativo y otro cuantitativo, y este último tenía como objetivo reducir la incertidumbre en el proceso de evaluación de stock.

Bajo este punto del orden del día se expuso una única presentación bajo el título “Hipótesis de un reciente desplazamiento hacia los polos en la distribución del pez espada del Atlántico norte”, basándose en el documento SCRS/2013/161. Se describían los cambios en los índices de abundancia de pez espada con respecto a los indicadores de procesos medioambientales a gran escala (por ejemplo, AMO, NAO, AWP). Basándose en observaciones de tendencias opuestas en la abundancia de pez espada septentrional, el documento sugería la

posibilidad de un desplazamiento de la abundancia de latitudes más cálidas y meridionales hacia latitudes más frías y septentrionales. Varios de los índices de abundancia observados cambiaban abruptamente de dirección de negativo a positivo, mientras que otros mostraban un cambio opuesto. Los cambios observados en la dirección de los índices de abundancia se corresponden con cambios en las tendencias en el tamaño de la Piscina Cálida del Atlántico (AWP), el cambio de signo de la Oscilación Multidecadal del Atlántico (AMO) y la Oscilación del Atlántico Norte (NAO). Para cuantificar una posible relación entre los cambios en la abundancia y los diversos índices medioambientales posibles, se ejecutó el modelo de evaluación sin la influencia de los datos medioambientales y se realizó una regresión de los residuos del ajuste de las CPUE a los diversos índices medioambientales. Teniendo en cuenta los presuntos límites de tolerancia del pez espada a la temperatura, es posible que o bien su hábitat preferido se haya desplazado hacia el norte o bien se haya desplazado su presa preferida o ambos.

Basándose en estos resultados, se inició un debate sobre las ventajas de la agregación de las diferentes CPUE por área, en vez de la agregación actual por flota disponible para el Grupo de especies sobre pez espada. Esto requerirá un acceso a los datos de la flota para cada operación de pesca, lo que permitirá la incorporación del efecto zona y flota, así como de la estructura espacial en los modelos de evaluación (a saber las CPUE del Atlántico este frente a las CPUE del Atlántico oeste). También se resaltó que hay otros enfoques para incorporar la zona como un factor de interacción en la modelación de las CPUE mediante la utilización de modelos mixtos. Los enfoques de ponderación podrían calcularse tras estandarizar las CPUE por zona.

El Grupo también presentó la utilización de la modelación del hábitat para incorporar la variabilidad regida por el medio ambiente, determinando la distribución espacial y temporal en la modelación de las CPUE.

Hubo una amplia discusión sobre cómo abordar la agregación de los datos disgregados por operaciones de pesca para las pesquerías de palangre dirigidas al pez espada, como continuación del estudio presentado. Un campo principal de discusión fue el modo de establecer un marco de colaboración apropiado entre los científicos nacionales para obtener datos de cada operación de pesca y analizar estos datos conjuntamente. Este intercambio podría realizarse en el marco del acuerdo de confidencialidad de ICCAT ya existente. Se reconoció que sería un proceso gradual, que posiblemente requerirá el apoyo de la Secretaría, así como trabajar en la “nube” con un acceso restringido a los científicos que participan en el análisis en colaboración de estos datos. Estos datos son necesarios para abordar una serie de preocupaciones/interacciones asociadas con la pesquería y los índices con el fin de mejorar la evaluación e incorporar en ella factores medioambientales.

8 Otros asuntos

8.1 Colaboración con SISAM-ICES

Se expuso al Grupo una presentación realizada por uno de los copresidentes de ICES SISAM (iniciativa estratégica sobre métodos de evaluación de stock). De este modo se resaltaba el trabajo realizado en el marco de esta iniciativa hasta la fecha, y se mantuvo un debate sobre la orientación futura de SISAM y su relevancia para la comunidad científica de ICCAT. Los científicos de ICCAT participaron en los trabajos de SISAM en 2013, a través del Congreso mundial sobre métodos de evaluación (WCSAM). Las actividades de SISAM en 2014 incluyen completar documentos para una edición especial del *ICES Journal of Marine Science* (basada en el WCSAM), y jornadas de trabajo de organizaciones asociadas como CAPAM, así como una sesión abierta que celebrará ICES en su ASC en septiembre de 2014. En los debates de la reunión se resaltó que ICES e ICCAT se enfrentan a cuestiones similares al menos en términos de desarrollo de métodos de evaluación de stocks y de provisión de asesoramiento sobre pesquerías, y que estas cuestiones podrían ser comunes a muchas OROP a nivel mundial. Estas cuestiones estaban relacionadas con los avances en las metodologías y el reto de mantener en la red de la organización a científicos con experiencia técnica suficiente para utilizar estos métodos de un modo adecuado. Junto con esto, se debatieron los esfuerzos compartidos para desarrollar marcos de evaluación y asesoramiento que abordan situaciones en las que los datos son limitados. Se acordó que sería beneficioso para ICCAT mantener la colaboración con la iniciativa SISAM y, a través de esta plataforma, trabajar colaborando con científicos de otras OROP para contribuir a la resolución de estas cuestiones.

9 Recomendaciones

- El Grupo recomendó que se inste a las CPC a comunicar sus datos de captura y esfuerzo de Tarea II en una estratificación geográfica más fina (por ejemplo 1°x1°) en vez de comunicar estos datos en estratos de 5°x5°, ya que en algunos lugares dicha escala podría resultar demasiado tosca. El WGSAM solicita además que el Subcomité de estadísticas considere si debería requerirse también la comunicación en esta resolución de escala más fina.
- El Grupo acordó también que debería fomentarse la implementación del enfoque de evaluación de estrategias de ordenación (MSE) y el diálogo entre científicos y gestores pesqueros sobre normas de control de la captura y MSE para mejorar el asesoramiento científico proporcionado a la Comisión. Estos esfuerzos deberían incluir una revisión de los esfuerzos MSE realizados hasta la fecha considerando los logros, fracasos y los recursos que limitan los progresos de MSE futuras.
- El Grupo manifestó que pensaba que los diferentes grupos de trabajo podrían utilizar criterios simples para empezar a asignar una puntuación a la calidad de la información utilizada en diferentes evaluaciones de stock. Entre tanto, el Grupo recomienda que se sigan desarrollando modos objetivos más detallados y basados en la ciencia de proporcionar dichos sistemas de puntuación. Esto incluye desarrollar criterios para evaluar la importancia de los diferentes elementos de datos dependiendo del ciclo vital y/o del modelo de evaluación utilizado. En esta misma línea, el Grupo recomendó que prosiga el desarrollo de la base de metadatos con información sobre la cantidad y calidad de la información biológica y pesquera disponible.
- El Grupo instó a que la medición de la realización de predicción retrospectiva, descrita en la sección 5 de este informe, se evalúe mediante estudios de simulación, potencialmente con modelos condicionados por evaluaciones pasadas de ICCAT. A ser posible, debería desarrollarse un caso de estudio para un stock para el que no se hayan observado series de CPUE contradictorias.
- El Grupo instó a las CPC a proporcionar un acceso limitado a los datos de CPUE para cada operación de acuerdo con las necesidades y prioridades identificadas por los diferentes grupos de especies y subcomités. Esto permitiría al SCRS producir una amplia variedad de índices en una escala espacial más informativa. En un primer momento podría utilizarse la información de una sola especie y/o flota específica para ilustrar los beneficios de este enfoque. Se sugirió la utilización de las oportunidades que brinda la “nube” mantenida por la Secretaría para fines de archivo y acceso, con el fin de facilitar la colaboración multilateral. Este intercambio podría realizarse en el marco del acuerdo de confidencialidad de ICCAT ya existente.
- El Grupo consideró también que el SCRS debería continuar participando en la iniciativa ICES_SISAM para seguir fomentando la colaboración para el desarrollo de metodologías de evaluación, para compartir y desarrollar los conocimientos sobre el modo de comunicar la incertidumbre a los gestores, para propiciar una colaboración más estrecha en las evaluaciones conjuntas (por ejemplo, marrajo sardinero), así como iniciativas prácticas, como compartir el orden del día de los grupos de trabajo sobre métodos de ICES y de ICCAT. Esto proporcionará un acceso abierto publicitado y estandarizado a los datos de entrada y salida de las evaluaciones, así como la posibilidad de compartir datos más detallados. Estos elementos podrían desarrollarse mediante la participación en el Grupo de trabajo global sobre métodos de evaluación (GAME).

10 Adopción del informe y clausura

El informe fue adoptado durante la reunión. El presidente del WGSAM expresó su agradecimiento a los organizadores locales por el excelente trabajo de organización de la reunión y a los participantes por su eficacia y el gran trabajo realizado. La Secretaría reiteró su agradecimiento a la *Irish Sea Fisheries Board* (BIM) por la excepcional organización de la reunión y por el cálido apoyo prestado a los participantes. La reunión fue clausurada.

Referencias

- 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.

APPENDICES

- Appendice 1.** Ordre du jour
- Appendice 2.** Liste des participants
- Appendice 3.** Liste de documents
- Appendice 4.** Suggestions de champs à inclure dans la base de métadonnées halieutiques et de marquage-recapture.
- Appendice 5.** Base de métadonnées biologiques pour des études individuelles

APÉNDICES

- Apéndice 1.** Orden del día
- Apéndice 2.** Lista de participantes.
- Apéndice 3.** Lista de documentos.
- Apéndice 4.** Campos sugeridos para la base de metadatos pesquerías y recuperación de marcas.
- Apéndice 5.** Bases de metadatos biológicos para estudios individuales.

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 Oceans 531 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, Spain
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, Morocco

Serghini, Mansour

Institut national de recherche halieutique, 2, Rue Tiznit, 20000 Casablanca, Morocco

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, USA
Tel: +1 305 361 4277, 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, USA
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 , Spain
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

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.

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

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	N° 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
conversion factors	F/M sex ratio
...	...