

**REPORT OF THE ICCAT WORKING GROUP ON
STOCK ASSESSMENT METHODS**
(Madrid, Spain - May 8 to 11, 2000)

1. Opening adoption of the agenda and arrangements for the meeting

In accordance with decisions made by the SCRS in its 1998 and 1999 meetings, the Working Group on Stock Assessment Methods met at the ICCAT Secretariat in Madrid, 8-11 May 2000. The meeting was chaired by V. Restrepo who opened the meeting and welcomed all the participants.

The primary objective of this new working group is *“To implement Quality Management for stock assessment methods, leading to the review, testing and documentation of assessment methods used by the SCRS.”* For this first meeting of the Working Group the following objectives were defined and circulated in advance:

- (a) Develop a protocol for the review of methods, and prioritize future work of the Working Group.*
- (b) Evaluate methods for CPUE standardization that take targeting into account.*
- (c) Evaluate methods for CPUE standardization that take spatial heterogeneity into account.*
- (d) To formulate advice on assessment mechanisms to monitor stock rebuilding.*

The Tentative Agenda was adopted by the Working Group and is attached as **Appendix 1**.

The following served as rapporteurs:

Sections 1, 7 and 8	V. Restrepo
Section 2	E. Babcock and J. Mejuto
Section 3	D. Die and N. Miyabe
Sections 4 and 5	C. Porch and C. Legault
Section 6	R. Pianet

It was decided that a short Executive Summary also be prepared for the SCRS and Commission use. The SCRS Chairman, J. Powers, agreed to prepare the initial draft.

2. Quality management systems

2.1. Presentation of Documents.

There were two documents presented under this agenda item. These papers were the basis for most of the discussion.

Document SCRS/00/38 described the ICES quality assurance policy for fish stock assessment and management advice. The focus of the policy is not on choosing assessment methods, but rather on the mechanics of determining whether a method performs as it should. The policy was adopted in October 1999, based on the objectives of scientific consensus, integrated (ecosystem) approaches, credibility to the clients, transparency, responsiveness, and independence of political influence.

The components of quality control are: (1) a quality procedure for the assessment groups; (2) timeliness; (3) transparency, including the ability to follow data flow through the databases, and the flow of assumptions and decisions through the assessments, through the working group reports, advisory reports, and manuals for general assessment procedure and handbooks for assessments currently in development; (4) internal peer review. The peer review takes place between the assessment and the presentation of management advice, to determine whether the assessment is reasonable, and that the projections are based on a reasonable assessment. This review has been made into a more formal, better documented procedure, and in small subsets of the Advisory Committee on Fisheries Management (ACFM) before the ACFM meeting. External reviews have not been implemented, but will be periodic.

Document SCRS/00/40 presented a proposed stock assessment quality control procedure for ICCAT. The levels at which quality control should be implemented were:

- (a) Collection of raw data.
- (b) Aggregation of (a) for submission to the Secretariat.
- (c) Summarization of analyses of (a) or (b) by the Secretariat for use by species working groups.
- (d) Analysis of (a) by national scientists for use by species working groups.
- (e) Development of assessment methods.
- (f) Implementation of assessment methods.
- (g) Preparation of inputs to the assessment programs
- (h) Selection of options for running the assessment programs.
- (i) Validation of the results in the light of program outputs, diagnostics and expert knowledge.
- (j) Summarizing and synthesizing the results
- (k) Formulating advice that is consistent with the results.

Items (a)-(c) are considered to be the responsibility of the Sub-committee on Statistics, (j)-(k) are the responsibility of the *Ad hoc* Working Group on SCRS Organization, and (d)-(i) are the responsibility of this Working Group.

2.2 Recommendations for Quality Control of Stock Assessments

The goal of the quality control protocol is to make sure that assessments are scientifically sound, follow established practices, and are transparent and well documented, without sacrificing methodological innovation. Some of the quality control would have to be done by national scientists and some by the ICCAT Secretariat. The protocol would include standards for inputs, diagnostics and programming. It is not the intent of the protocol to restrict the flexibility of assessment working groups in choosing assessment methodology.

It was recommended that a catalogue of ICCAT approved applications (i.e. software) be developed, which would include the elements listed in **Appendix 3** (see Section 2.2.3). The relative merits of the methods would not be evaluated, but rather whether the software implementing the method worked as intended and was adequately documented.

The protocol could be enforced either by specifying that an application could not be used unless it had been cataloged, or by requiring working group reports to specify whether or not each result had been produced with an approved application. The protocol could not be enforced for at least 2 years.

2.2.1 Analyses by national scientists.

Quality control of preparatory analyses would be mainly the responsibility of national scientists, who should document their methods and perform standard diagnostics. The issues to be considered are whether the method used was appropriate and whether any errors were made in the application. New analysis methods should be presented with simulation testing and documentation of their differences from previous methods. National laboratories should document their quality assurance processes (standards will need to be developed by the national laboratories; the SCRS may wish to consider developing guidelines for minimum standards).

The Working Group recommended that national scientists should: (1) attempt to independently replicate their results, (2) document their results and the reasons why alternatives were not chosen and (3) should run standard diagnostics for the methods used. It was recognized that many national scientists are already doing these things in the papers presented to assessment working groups, as well as bringing their data so that alternative methods can be applied during the assessment.

National scientists from different countries are encouraged to develop joint projects in order to solve common data analysis problems, such as CPUE standardization. While quality assurance would be improved if national scientists could share data so that others could replicate their analyses, there are confidentiality issues involved.

2.2.2 Development of assessment methods

ICCAT encourages the development of new methods. There was a discussion about whether this Working Group should provide criteria for evaluating the relative merits of a method, e.g. that the method must be published in a journal, or externally reviewed or internally reviewed. The Working Group also discussed whether there should be guidelines for evaluating new models, by looking at how sensitive estimates of reference points are to the structural changes in a model, whether there is “value added” by the model, etc. The conclusion was that: (1) the software to implement suggested new methods should be submitted for cataloguing 5-10 months before the assessment working group meeting so that the application can be approved, and (2) the authors should document how the method differs from the previous method and why it is an improvement.

There was discussion about whether applications used in sensitivity analyses should have to go through the bureaucratic process. It was pointed out that changes in model inputs that arise at an assessment may suggest changes in the model which should be used, at least in sensitivity analyses, although they could not be approved by the catalog process. However, presenting a method that uses untested software, even in a sensitivity analysis, may give the untested results undue weight. It is up to the Working Group to weigh the risk of introducing a bug against that of ignoring a plausible model. _ The Working Group agreed that methods should be reviewed both internally and externally.

The Working Group recommended that a list of standard diagnostics for each method be developed by the individuals who are proposing the new methods

2.2.3 Implementation of assessment methods

It was recommended that a catalogue of applications be developed, which would include the elements listed in **Appendix 3**. The Working Group recommends that a Coordinating Committee be set up to supervise this catalog. The Committee would be composed of three people: the Secretariat’s Population Dynamicist plus two national scientists to be nominated by the chair of the SCRS. When an application was submitted to the Coordinating Committee, the Committee would determine whether the authors had

adequately provided all of the elements in **Appendix 3**, and would evaluate whether the application was well documented and performed as claimed. The Committee's comments would be added to the information in **Appendix 3** as Item 16. Applications that were approved by the Committee would then be added to the catalog of applications available to working groups for use in assessments.

Methods accepted into the ICCAT catalog need to present all the information in **Appendix 3**, and justify the need for a new program. The Coordinating Committee should have the flexibility to decide what level of detail is needed for each item in **Appendix 3**. It was suggested that there should be some standard ICCAT data sets to test whether a new code gets the same results as the established code when applied to the same question. The test data should not be used to evaluate the merits of a method, since useful methods may perform badly with certain simulated data.

There was discussion about whether new applications could be sent to the Coordinating Committee by the authors, or whether applications would have to be nominated by either the chairs of working groups, or the heads of national delegations, or the SCRS chair, or some other nominating authority. Approval of an application would take some time (~8 months). Methods should be prioritized.

New versions of cataloged programs should be provided to the Coordinating Committee, along with text describing changes from previous versions.

It was suggested that object-oriented programming would ameliorate the problem of software testing, by allowing an application to be altered with a minimum of code changes. Spreadsheets and other high-level program language applications can be given to the Secretariat without providing all of the information in **Appendix 3**.

It is not recommended that users of the applications must be certified or that there be any formal training. However, it would be beneficial if more working group participants could use the applications, and the catalog will help people to learn.

Some Working Group members expressed concern that this protocol would slow down the process of innovation. However, if the catalog includes a "mother" program, approving a variation in the method should be fairly quick (a few months). The Committee will work by e-mail inter-sessionally.

2.2.4 Checklist for inputs

The Working Group recommended that the input files for each run should be circulated in assessment working groups, so that the group can check for errors as far as possible within the limited time available. Models should either have a standard input format, or spit out the inputs in a standardized format so that they can be checked. The run log file should specify the input files, output files and options chosen for each run, and be automatically generated. Either the output files should be marked, indicating whether they are from the final model, or some number identifying the final model should be included in the working group report.

2.2.5 Checking model options.

At assessments, the person nominated by the working group chair to check model inputs should also check the model options. The run log should record the model options.

2.2.6 Validation of the results in the light of program outputs, diagnostics and expert knowledge

For the assessments themselves, quality control would include

- (1) Verifying the inputs (e.g. does catch at age sum to total catch?)
- (2) Verification and logging of run options.
- (3) Cataloging, documenting and storing all runs.

The assessments should also be subject to occasional peer review, possibly every 5 years.

It was stressed that the species working groups must maintain the right to evaluate methods with their own expert knowledge, and determine which inputs and assumptions are appropriate.

If it is necessary for small changes to be made to an application during an assessment, assessment working groups must decide on the tradeoff between risk of programming errors and flexibility (although this Working Group could provide guidelines). Working group chairs should be wary of coding changes, as errors have an impact on the credibility of the assessment process. Method variations could be accepted during a working group if the change was small, transparent, offered substantial advantages and had been evaluated with standard diagnostics. It is desirable for two people at a meeting to make the change independently to avoid coding errors. If a program is updated during a meeting, the new version should be given to the Secretariat.

2.3. Top down quality control

The quality control procedures discussed above evaluate the assessments from the bottom up, but there was also some discussion on examining the assessment process from the top down. New methods would be evaluated in terms of the management objectives, then the goals of the method, then the suitability of the components of the model (model of fish, fishery, parameter estimation, inputs), the conceptual aspects of the model (statistical rigor, biological realism, handling of uncertainty) and finally the practical aspects (data requirements, ease of use, transparency). The Working Group concluded that while quality assurance is needed at all levels of the assessment process, the choice of method is at the discretion of the individual species working groups. The protocol developed here covers quality assurance of application software only.

Suggestions for choice of method (by species working groups) included the ICES criteria (documented, peer reviewed, simulation tested). Methods should also be clear in their assumptions, and should provide the information necessary for management.

2.4. Comments on other matters.

The Working Group recommended that the SCRS Subcommittee on Statistics incorporate quality assurance in a broad sense, considering the needs of the assessments. There was discussion of whether ICCAT could leverage minimum standards of data reporting and quality assurance by nations. It was recommended that estimates of uncertainty be provided with the data, so that uncertainty can be incorporated in the assessments. Improved quality assurance of the Secretariat catch at age data was also recommended. Future changes in the ICCAT database are expected to facilitate data checking by Secretariat staff.

It was also suggested that, to avoid the problem of new data leading to changes in methods, which in turn leads to ill-considered decisions at assessments, the data-compiling and assessment processes

should be separated. Data compilation and assessment sessions should either be held during separate meetings within a year, or assessments should use data compiled the previous year. Separate data and assessment meetings would be expensive. However, the issue is brought to the attention of the *ad hoc* Working Group on SCRS Organization.

The SCRS has discussed and is continuing to discuss how to implement reviews of assessments.

3. Standardization of cpue indices

The SCRS and two species working groups requested the ICCAT Working Group on Methods to evaluate and make specific recommendations on methods of CPUE standardization that take the issues of targeting and spatial heterogeneity into account. One specific document was prepared in order to help discussions at this meeting (SCRS/00/39), but several other papers dealing with this same issue, but not prepared for this meeting were presented as background papers and were considered by the Working Group.

3.1 *Linking stock assessment and CPUE standardization*

In SCRS/00/39 Medley notes that, in general, CPUE standardization is done independently of the stock assessment, as a first step in the analysis. Standardization is used to remove some of the variability in CPUE that is thought not to be related to fluctuations in abundance. Unfortunately some of the variability removed may have been informative for abundance estimation, and thus would be lost if standardization is done separately. Medley proposes that CPUE standardization should be simultaneously conducted with the stock assessment in a single analysis. This can be simply done by developing the CPUE model independently, but then integrating the CPUE model with the population dynamics model into a single estimation process. A strength of this approach is that the population model is consistent with the treatment of CPUE data. It must be noted, however, that this approach is not the only possible one to ensure consistency between the population model and the CPUE standardization.

These integrated models tend to have linear components for the effort standardization and non-linear components for the population model. There are packages that can be used to obtain numerical solutions for these hybrid models.

These sorts of integrated models can be used to address any modelling issue regarding effort standardization and thus can be used to deal with issues of targeting and spatial heterogeneity. More specific reference to such applications is made in the appropriate section below. It was suggested that a practical example be used to evaluate the approach.

3.2 *Dealing with targeting in mixed species fisheries*

Some ICCAT fisheries are characterized by the simultaneous catch of more than one species. Historically, many of these fisheries originated as single species or had a clear target species. As the fisheries evolved targeting practices often changed and may even change seasonally. Although in some rare cases there is direct information on the target species for any given catch record, most often the target species has to be inferred from information on gear characteristics, or may be totally absent and has to be inferred from the ratios of different species in the catch. This section presents first an account of the main issues regarding targeting in ICCAT fisheries, then an overview of methods that attempt to deal with inferring targeting practices from by-catch information and finally a discussion on how to incorporate this information into the standardization of fishing effort.

3.2.1 Overview of targeting information available for ICCAT fisheries

a) Japanese longline fishery

This fishery really started after World War II targeting initially yellowfin and albacore for canning. As the sashimi market developed and high quality tuna meat began to fetch high prices, the fishery started to shift its targeting towards bigeye and bluefin tuna. In addition to these tunas, other by-catch species (eg. billfishes, sharks) are caught. Currently longline sets contain 2000-3000 hooks per set and, on average, catches are commonly of 20-30 fish per day, although maximum catches are around 200 fish per day. Under such current conditions it does not seem that gear saturation is occurring during recent years, even though most fish could be captured within a small section of the longline.

Information on targeting can be obtained from information on hooks per basket and depth of longline. Baskets directed to bigeye contain more than 10 hooks, those used to catch bluefin have 7-9 hooks. It must be noted that the hook set-up can change within a trip because fishermen switch from targeting one species to another. There are strong seasonal changes in targeting, bluefin being targeted mostly during the winter and spring. It is well known that other species such as yellowfin and billfishes are rarely targeted although these species may constitute important components of the catch.

b) Swordfish longlines

There have been considerable modifications of gears over time, including the type of mainline. These modifications were partially driven by economic returns that drove fishermen to target sharks or other species by changing the time, areas and gear used in their longline sets. Most of these changes tend to be small and the available data do not help to clearly estimate targeting.

c) Purse seine tropical tuna fishery

The purse seine tropical tuna fishery has historically targeted yellowfin and skipjack. Bigeye and small tuna (AUX and LTA) have been the most important by-catch species, other by-catch species are not significant (around a 1% of total catch). Data available from logbooks include catches of the three main tuna species (yellowfin, skipjack and bigeye) and some of small tuna. Other by-catch species data are partially obtained by observers.

One characteristic of these fleets has been the continuous improvement in the technical equipment of boats that leads to the perception of continuous increase in their fishing power. Changes in the equipment occur in a very quick way for all fleets and as a consequence the different levels of the technology factors (bird radar, sonar, mesh size, depth, etc.) rarely overlap in time, making difficult the detection of their effects on CPUE. Also the skipper's skill is an important factor to take into account in the estimation of the fishing power of boats. Most of those data are not currently available.

The extension in the use of FADs for fishing since 1991 has created an important change in this fishery. Prior to the introduction of FADs, fishing on natural log was limited to a localised area and time and catches on logs represented around the 15% of the total catch. With the introduction of FADs those catches are now close to 50%. The species and size composition differs substantially depending on fishing modes: in the free school catch yellowfin is the dominant species followed by skipjack, on the contrary in the FAD's catch skipjack represent more than a 70% and small yellowfin and bigeye are close to 15% each. Since 1990, information on the fishing mode of catches is available in the logbooks. This fishing mode has introduced a lot of new elements and problems (added to the different species composition) to consider in the standardization of effort. Firstly, it is not possible to split effort by fishing mode from the information available in the logbooks, secondly there are a lot of interactions between factors such as

season and area, fishing mode, number of FADs used and skipper (because fishing free schools demands more skill). Oceanographic conditions (e.g. the depth of the thermocline, the wind's speed, the surface temperature, etc.) are also factors that may have to be taken into account in effort standardization.

In more recent years (1997-1999) the temporal (3 months) closure of a wide area around the Equator to fishing on FADs has introduced a new element to consider in the standardization procedure.

3.2.2 Inferring targeting from catch or catch rates of secondary target species and by-catch species

A direct way to estimate targeting is to use information on fishing practices that can define whether a given species is targeted or not. It is therefore recommended that when possible such information be collected. Often, however, this information is not provided with the catch records. In such cases you can attempt to use information from catch or catch rates of secondary target species and by-catch species to develop indices with which fishing effort can be assigned to a particular target species.

For CPUE data sets without enough direct measures of gear characteristics and other factors, proxies for targeting based on catches of the original target (hereafter denoted as target species) and companion species can and have been used in recent ICCAT stock assessments. However, concerns have been raised about the accuracy of such proxies and whether they correctly indicate changes in targeting and about the potential misrepresentation of trends in abundance when such proxies are applied to produce time series of relative abundance (see SCRS/99/125). For example, a biased proxy of targeting could cause a real decrease in abundance to be incorrectly dismissed as a change in targeting when no such change had actually occurred. In 1999 the Swordfish Working Group (SWO 1999 report) identified several methods to create such indices and evaluated them through simulation. Simulations made during that meeting revealed one of those indices as more robust than the others to different scenarios of historical targeting changes. These simulations were, however, largely focused on the swordfish case and therefore did not consider the full range of possible target and by-catch abundance trajectories that occur in other ICCAT fisheries. These simulations did not consider either all the possible scenarios of bias due to historical changes in targeting and under-reporting of by-catch. As a result during the present meeting it was decided that further simulations should be run to make the results applicable to a broader range of tuna stocks.

In order to determine whether various potential targeting indicators would be biased by systematic changes in target species biomass, non-target species biomass, total effort, the reporting of non-target species, and various combinations thereof, simulations were performed. These simulations projected deterministic population trajectories of hypothetical target and non-target species biomass, with different trends in targeting by a fishery, and calculated five different indicators of targeting throughout the time series (described below). The cases identified for evaluation were based on some ICCAT fisheries in which systematic changes in targeting has been suspected:

- (1) Swordfish (original target species) and blue shark (secondary target species) (See SCRS/99/125)
- (2) Swordfish (original target species), blue shark (secondary target species) and yellowfin tuna (non-target species)
- (3) Skipjack (original target species) and yellowfin tuna (secondary target species)
- (4) Skipjack (original target species) and yellowfin tuna (secondary target species), but where the secondary species is very low in abundance relative to the original target species.
- (5) Skipjack (original target species), yellowfin (secondary target species) and bigeye tuna (non-target species)

Original target species denotes the species that is viewed as the only target species for the first several years of a fishery. Secondary target species denotes a species that is caught as by-catch for the first several years of a fishery and later becomes a target species by at least some portion of the fishing fleet. Non-target species refers to a species that is not targeted but caught as by-catch.

In order to develop realistic population trajectories of the fish stocks, a Schaefer production model was used, with r and K values approximating those for the simulated species. The Schaefer model parameter values that were assumed for these different species are listed in **Table 1**.

The series of total annual effort was specified as either constant or a linearly increasing. The true targeting index is given by:

$$targ = \frac{Effort_o}{Effort_o + Effort_s}$$

where $Effort_o$ = the effort directed on the original target species and $Effort_s$ = effort directed on the secondary target species. In the simulations, the values for $Effort_o$ and $Effort_s$ were obtained by multiplying the total effort specified in each year by $targ$ and $(1-targ)$, respectively.

For each species, two different types of values for q were specified: One was for the situation in which the original target species was being targeted ($q_{species, o}$). The other was for when the secondary target species was being targeted ($q_{species, s}$). The latter value was set to a value much lower than the former value. Thus, the total annual catch of each species was:

$$C_{species} = (q_{species, o} * Effort_o + q_{species, s} * Effort_s) * B_{species}$$

The indices of targeting for the original target species considered were:

Index I: Catch of original target species/total catch of the other species where other species refers to the non-original target species

Index II: catch of original target species/total catch

Index III: 1/uncorrected catch rate of other = total effort /catch of other species

Index IV: 1/Catch of other

Index V: whether Catch of original target species is greater than Catch of other.

Each of these indices of targeting was computed over a 15-year period in each of the five species scenarios defined above and plotted along with the true index of targeting on the original target species. In each of the five species scenarios, three different targeting situations were simulated. In the first there was no change in targeting and reporting. In the second, the targeting decreased linearly from 100% on the original target species to 20% within five years. In the third, this decrease also occurred together with a simultaneous increase in reporting of the secondary target species from 20% to 100%.

For each species scenario, a number of different biomass trajectory scenarios were simulated by manipulating the values for q , the biomass relative to K for each species, and whether effort was constant or increasing. In each case, the indices of targeting were graphed, after normalising by dividing each value by the average of the series.

The CPUE series of the original target species that would be estimated by applying each targeting index were also calculated as catch of target species divided by (total effort times the targeting index normalized to be between zero and one).

In addition to plotting the targeting series and resulting CPUE series, the results of the various scenarios were tabulated in terms of the sum of squared residuals between each derived and the true relative abundance values for each series.

Figure 1 depicts the simulation results using targeting indices I to IV for the “skipjack” and “yellowfin tuna” case (Index V did very poorly in most cases, and was eliminated from further consideration). Each set of panels in **Figure 1** shows the estimated targeting index as well as the corresponding CPUE series. Figures for the other cases are not shown because they are similar to the two-“tuna” species case. **Table 2** shows the sum of squared residuals in the derived CPUE series from the various scenarios obtained for the four remaining indices that are computed from the catch values. **Table 3** shows the ranking of the four targeting indices based on the sum of squared residuals.

When targeting proxies based on the relative size of the catch of the targeted species compared to the others are used for annual totals only, they can be expected to produce highly biased results and should be avoided. However, if the method is applied on a set by set basis, this method may be more appropriate. The index in which the catch of the original target species was divided by the total catch performed best in over half of the case scenarios. However it also performed very poorly relative to the others in about a third of the scenarios. Moreover, in many of the scenarios, all of the indices, even the best performing index, performed very poorly and produced highly spurious trends in abundance.

The results of this evaluation suggest that there is no clearly "best" method that could be generally applied and that any one proxy, even if it performs best relative to the others, can still produce very biased results. Therefore, whenever it is of interest to try to correct for targeting when only total catch of each species and total effort is available, the results of this study should be consulted and a similar simulation exercise could be carried out. It would be advisable to identify the various particular conditions in the fishery of interest. This would include the potential trends in total effort, the relative abundances of the different species caught, their possible trends in abundance, and whether changes in reporting have occurred. A decision to apply one of the alternatives should be made only if one has been found to perform acceptably when applied to indicate potential trends in abundance.

This deterministic evaluation has evaluated only the potential bias in some alternative proxies for targeting. It has not evaluated the potential precision in the alternative proxies which is also an important consideration, once some alternatives have been identified that have acceptably low bias. This latter issue requires further attention in Monte Carlo simulations in which random variability in catch and effort data are simulated, in particular to identify how large the changes in targeting need to be before they can be reliably detected.

Also, the Schaefer model was used to produce biologically realistic biomass trends in these simulations. However, the performance of the targeting proxies depends on the biomass trends, not the population dynamics model, so it could be useful to perform simulations like this with simple linear trends in the biomass of each species.

The inconsistency in performance of the alternative targeting proxies across different scenarios and the large bias in results frequently obtained from them suggest that it is also desirable to work to identify other more reliable targeting indices. Proxies that could also be constructed and evaluated include those based on GLM analysis of catch rate data (Glazier 2000). It is also advisable for additional fishery data to be collected on a regular basis and analyzed to obtain a stronger empirical basis to identify changes in targeting. These additional data could include, for example, direct measures of gear characteristics and the way in which the gear is deployed (e.g., longline depth and hook spacing).

Of the proxy methods evaluated, the use of the ratio of catch of the target species to total catch (Index IV, **Table 3**), performed best on average and remains the preferred proxy, although this method may not provide the best performance in all cases. These evaluations have not tested the bias in using CPUE as relative abundance indices when the effect of targeting is neglected. This source of bias could be substantial and should be evaluated within the simulations outlined above to provide a basis for comparing the costs and benefits of applying proxies for targeting relative to ignoring targeting effects.

3.2.3 Incorporating targeting information during CPUE standardization

Glazer and Butterworth (*In prep.*) and Glazer (2000) considered a model that incorporated targeting when the catch of the by-catch species is small in comparison to that of the target species and the abundance of the by-catch species has no trend. This work showed that when there is positive correlation between catch rates of the target species and by-catch it is probably not appropriate to use by-catch information for the by-catch species as a covariate. On the other hand when there is negative correlation then it may be appropriate to incorporate the by-catch as a factor in the CPUE standardization. Unfortunately many of the targeting issues with tuna refer to species mixes where the abundance of the secondary species is not much smaller than that of the primary species so that the above models are not appropriate. In such cases an option is to carry out the CPUE standardization as a multivariate analysis where the catch rate of both target and by-catch species are dependent variables. Such analyses can be carried out by using special SPLUS libraries. Principal Component Analysis can also help accomplish this but care must be exerted in choosing the independent variables. It would be best if these variables are closely related to the process that controls changes in CPUE, e.g. the location of a front, rather than arbitrary grids based on latitude and longitude. Although conceptually this is desirable, we do not often have better proxies than location, and when we do have better proxies they often do not extend far in time for the analyses to be of much use for reconstructing historical CPUE data series.

Set by set data where CPUE of several species are obtained simultaneously make it possible to separate the effect of targeting of species (or substitution between species) from short term variability in overall catch rate, e.g. due to unobserved (or unobservable) environmental factors. **Appendix 4** presents in a multiplicative model framework how this might be done based on explicit modeling of the targeting process and short term environmental parameters accounting for overall good or bad hauls. More work is needed before the usefulness of the approach can be evaluated.

When the main reason to incorporate targeting into the CPUE standardization stems from the need to deal with the problem of the presence of records with zero catch of one of the species a delta log-normal model can be used to account for such records explicitly and there is no need to incorporate any other targeting variables.

Targeting can also be treated as a missing data problem. In general, target models are considered simple where the target species is explicitly available as a data variable. For example, within a trip, a model of CPUE could be:

$$m_i = (w e^{\hat{\alpha}_1} + (1-w) e^{\hat{\alpha}_2})$$

Where m is the expected CPUE in the GLM, β_1 the linear parameters (year, location, vessel and other effects) targeting species 1, β_2 for targeting species 2 and w the proportion of effort targeting the first species. The problem is to estimate the parameters where targeting data (w) is not recorded. Where a significant number of vessels have observer data where the species target is recorded, a good method to fill in the missing values may be the EM algorithm. The EM algorithm yields MLE parameters, and should account for targeting in the final model. The method imputes expected values for the target variable,

which is the proportion of effort units applied to target the species. However, the general method would also need to address:

- More complex models, including different effects of catchability and environmental effects
- Where there is no targeting data, only catch species composition

The first issue is an extension of the approach, and while it may present a more complex numerical analysis, the EM approach should in theory cope with it. The second issue is much more serious as through aliasing it may not be possible to solve the problem without making further significant assumptions within the model. The recommendation is to explore whether the EM, stochastic EM (Gilks, Richardson and Spiegelhalter 1996) or other missing data method may be a useful way to approach the problem of targeting.

Regardless of what targeting model is used it is important to remember that not all catch per unit of effort data series may constitute a good index of abundance. Landings of by-catch species per unit of effort are often more a reflection of marketing preferences than of abundance. In multispecies fisheries, catch limits imposed on specific species can produce changes in the fishery operations, change targeting practices and ultimately affect the consistency of CPUE data series.

3.3 Models dealing with spatial heterogeneity in CPUE data

There has been considerable work done on standardization of CPUE indices in a spatial context in tuna species, especially for southern bluefin tuna. The Japanese longline data from this fishery shows a significant shrinkage of the area fished over time. This may be explained by a contraction of the area where the stock is abundant or by the effects of fishery regulations (smaller TAC) that have led the fishermen to only fish in the most productive areas of the stock. In this fishery CPUE has been standardized by multiplying GLM estimates for the different regions and months by the area of the region, then adding for the different regions and averaging over months.

Unfortunately this analyses requires dealing with regions with no fishing effort, but within the historic fishing grounds. Two extreme hypotheses have been tried, to assume the abundance in such regions is zero or to assume that it equals the average of all fished grids. More recently a third hypothesis has been tried by using a geostatistical interpolation technique to estimate density in the areas where there is no fishing in a particular year. There has been no resolution in the debate of how appropriate any of these interpolation techniques are. The use of habitat information to help interpolate such densities has also been proposed. GAM or mixed models could be used for this purpose.

The abundance trends for southern bluefin tuna are, unfortunately, very sensitive to the above assumptions especially during the most recent 10 years. The above comments for southern bluefin tuna, however, may not always be applicable to Atlantic tuna stocks.

Once these spatial abundance factors (=density) are estimated from such models, it is better to calculate absolute abundance by simply multiplying the density by the size of the region and summing over all regions. When this is done, however, one has to be mindful of the spatial scale of the region. Regions that are too small may suffer from problems of local depletion and thus may not generate reliable estimates of density. Cooperation between fishermen (code groups) may significantly reduce the variance between vessels by the fact that they share information on the location of high catches. Research conducted in the IATTC and ICCAT suggests that cooperation increases efficiency so effectively that CPUE may even increase in the presence of increases in fishing effort. This could also be interpreted as more efficient searching by vessels that belong to a code group in comparison to those that fish individually.

Three issues are commonly faced, when GLM models are used to standardize CPUE for spatial heterogeneity:

- The opportunity to use random effects models to reduce parameter complexity (see O'Brien and Kell SCRS/96/173)
- Which interactions should be kept in the model
- Which diagnostics should be investigated to ensure that the model is appropriate for the application

In the case of interactions it is recommended to use likelihood ratio tests to see whether adding extra terms statistically helps the fit.

3.3.1 Incorporating spatial population dynamics in the CPUE standardization

If management is conducted in a spatial manner (spatial closures) it often justifies incorporating spatial structure into the stock assessment model. The same can be said when life history characteristics have a spatial component, for example when migrations are related to life history stages.

Medley (SCRS/00/39) shows that models can be developed to explicitly consider processes like spatial structure and migration together with a model that incorporates CPUE standardization. It can be shown that incorporating these components into the integrated model can improve the fit to simulated data. When using these models, however, attention must be paid to the spatial scale at which CPUE standardization is built in the integrated model. In principle there is no need that the spatial scale of the CPUE standardization matches that of the population dynamic model. Some level of aggregation of catch and effort data (preprocessing) will always be required, as it is general practice in effort standardization (e.g. it is likely that even if data is collected haul by haul, it will be aggregated at larger scales before it can be standardized.). It must be remembered that any standardized CPUE series obtained from an integrated model will be dependent on the specific population model used. These series, however, can still be used in other analyses.

4. Stock rebuilding and assessment methods

Continuing application of the assessment method that was in use at the time a specific rebuilding objective was adopted forms the basis for monitoring such rebuilding. This process should include retrospective analysis as a diagnostic to check reliability.

Further discussion on this topic centered on three major themes: 1) when is a change in models warranted, 2) how should the change be presented to managers, and 3) how can we improve the management paradigm.

4.1. When to change

If a rebuilding plan is in place already, is it better to stick with the current assessment method for monitoring or switch to new method that is perceived to be an improvement? Concern was expressed that the credibility of the assessment process can suffer with a change in methodology without sufficient justification, especially if the change is later shown to be wrong. The ICES Working Group prefers to be as consistent as possible from year to year, however it recognizes that some changes are inevitable if due to nothing more than new types of data. It was also argued that science can and must improve over time. However, due to changes in membership of the working groups over time, the perception of which model is "best" sometimes changes with each assessment. Furthermore, some members suggested that any major changes presented at a working group meeting not be used in that assessment but rather delayed

until the next assessment to give participants an opportunity to examine fully this major change. Another concern was that the new model may change the perception of absolute abundance. Since ICCAT uses relative benchmarks, this is perhaps less of a problem, although the problem of setting TACs remains. In the ICES experience, presentation of results from both the old model and the new model has been found to be necessary but not necessarily sufficient to convince management of the need to change to the new model. In summary, it was felt that the null hypothesis ought to be the *status quo* and that the evidence should be substantial for a change to take place, i.e. the working group should be convinced the change is warranted.

4.2. How to present change

When managers are accustomed to a method that has been in use for some time they may be more likely to listen to the scientific advice. Often when a change is made, the assessment scientists have not done an adequate job of convincing the managers that the change is appropriate. This becomes particularly evident when the new method is especially favorable or unfavorable to a constituency. It was felt that a consistent and streamlined format of the Executive Summary that includes the essential benchmarks for management but does not include details of the model or jargon might help in this regard. However, given that many constituencies are involved in the assessment, changes in the methodology should be documented and justified in the detailed report. The reasons for making a change should be fully transparent to avoid questions such as “What was wrong with last year’s model?”

4.3. Improvements

The Working Group felt that the management procedure approach has desirable qualities but reservations were expressed that it may be difficult to implement in practice. The management procedure approach is based upon simulation studies to evaluate management controls that manifest robust performance under a wide range of the key uncertainties pertaining to the fishery. Work is currently underway in this area with progress to be reported for the next meeting of this Working Group.

4.4. Recommendations

The current model should be the null hypothesis and the working group should be sufficiently convinced that any new method is an improvement in order for a change to occur, and should prepare sufficient documentation to that effect. The Executive Summary should be streamlined to avoid jargon and the reason for change should be transparent in the detailed report. Management procedure approach simulations can be used to help prioritize research in support of management.

5. Other matters

5.1 Review of papers

The Working Group reviewed several papers that did not directly address the original terms of reference but were nevertheless pertinent to the theme of a Working Group on Methods. Synopses of the comments from the group are presented below.

Gaertner and Fonteneau (SCRS/00/36).

Concern was expressed that, currently, more attention is often paid to the precision of a given estimator than to its accuracy. In some cases a simple estimator such as is presented here may be more robust in terms of accuracy than its more sophisticated counterparts. It is important, of course, to evaluate

carefully the underlying assumptions of any method. For example, the method advocated in this paper is equilibrium-based and as such may tend to over-estimate MSY in certain circumstances. This bias may be small for stocks with a high intrinsic rate of increase (such as skipjack) or perhaps mitigated by the smoothing procedure (equilibrium approximation) taken in the paper. The potential magnitude of the bias could be evaluated through simulation studies. Overall the approach appears promising, particularly when effort data are absent. It may also be useful as a qualitative decision rule for managers and simulation studies with this in mind should be encouraged.

Geromont and Butterworth (SCRS/00/33).

This paper presented a method for allowing the model to determine the relative weights assigned to indices of abundance used to tune virtual population analyses. A similar approach has been employed by the IWC to estimate the additional variance associated with a single time series and in South Africa for multiple series. The Working Group concurred that the proposed approach may in some cases be an improvement over the iterative reweighting and input variance (alone) methods and that it is worth examining further. However reservations were expressed that the approach is likely to be imprecise, especially where short time series are involved (the model will be over-parameterized), and perhaps inaccurate when the various time series indicate conflicting trends. It was suggested that the use of input variance information was desirable, but that consideration then had to be given to accounting for possible biases in such variance estimates, possibly by use of the additional variance method suggested.

The suggestion was also made that the criteria for weighting (or excluding) each index should largely be determined external to the model, including the knowledge of the investigators who developed them, and their likely representativeness of trends in abundance. Analyses which include indices that are statistically contradictory should be considered with care.

Legault and Porch (SCRS/00/35)

The Working Group lamented the all too frequent changes in the way the indices of abundance are weighted. For some species a new weighting scheme has been adopted at almost every assessment meeting and, as a consequence, the credibility of the assessment process has suffered. The Working Group agreed that simulation studies like those presented would provide useful advice concerning the relative merits of the various weighting schemes. It was suggested that the simulations be extended to examine the effect of having to estimate the selectivities (as opposed to assigning the correct values as done in the paper) and the degree of correlation between the ages in a given index that would be necessary to mitigate the advantage of age-specific indices over age-aggregated indices. It was also felt that it would be useful to broaden the scope of the simulations to include other stock and fishery characteristics. Simulations could also be used to examine the suitability of likelihood ratio tests as criteria for determining whether or not model estimated variances are statistically justifiable when the data are sparse or the model is mis-specified.

Maury (SCRS/00/37)

The Working Group found this model to be a useful exploratory tool for examining possible trends in the catchability coefficients of various fleets over time and the effect of changes in fishing area on the perception of MSY. The application to skipjack reported in the paper was intended in this light, however it was pointed out that it could also be used in the context of an assessment. Towards this end, the relatively small difference between the posterior Bayes factors (PBF) associated with the different model structures was taken as evidence that the model was over-parameterized. It was suggested that the number of parameters be limited to those which cause a significant change in the PBF. Concerns were

expressed that the priors used, particularly as regards the random walk catchability coefficients, could end up driving the results, and that care should be taken in their selection.

McAllister et al. (SCRS/00/34)

The Working Group agreed that the general approach of presenting analyses in the form of decision analysis tables was a promising aid for both scientists and managers. It might be particularly useful for highlighting the priorities for future research. The technical point was made that the prior for the MSY level was developed from data which of late has been heavily criticized. It may be appropriate to recast the model to use a prior for steepness based on Myers' (e.g., Myers *et al.* 1990) data. At this point the concern was reiterated that the choice of priors can heavily influence the results, particularly when the data are not especially informative. This led to a discussion of how to quality control expert knowledge. It is not just a problem for Bayesian methods; fixed parameters in frequentist methods are also based on expert knowledge. The criteria for choosing priors should be transparent. It was pointed out that in business applications there are ways of addressing expert knowledge through multicriteria decision making. Such protocols might avoid merely transferring the debate from the choice of the base case to the choice of priors.

5.2 Suggestions from the participants for the future of the Methods Working Group

A major focus of this first meeting was on quality control issues and the transferability of methodology between species working groups. While these same issues may not warrant such attention in the future, subsequent meetings of the Working Group on Methods must focus on specific issues in order for the Working Group to have an impact on SCRS policy. The terms of reference for each specific meeting should address priority issues identified by the SCRS, although perhaps 20 percent of the meeting could be devoted to other methodological matters of interest to the participants. Concrete advice must be generated, not simply an account of the debate.

5.3 Possible collaboration with other bodies

When methodological questions are held in common, it could be beneficial to collaborate with other international organizations. The IOTC, for example, has informally contacted several other international fisheries bodies, including the CCSBT, IATTC, SPC and ICCAT, to investigate the possibility of collaboration in this regard. It may not be necessary or even desirable to meet with other working groups on occasions where the topics are mostly specific to ICCAT. Moreover, logistical considerations such as the timing of the meeting may sometimes cause difficulties. Nevertheless, such collaboration could be especially helpful where some working groups include members with a particular expertise not shared by the others or when time constraints prevent a working group from exploring a matter in sufficient detail by itself. The Working Group recommended this form of collaboration be pursued by ICCAT. For such meetings the SCRS should consult with the other bodies to develop an agenda.

6. Recommendations

Agenda Item 2. – Quality management

1. All methods used for assessment should be documented, including Quality Control at all levels of contribution to the results.

2. A Coordinating Committee should be created in order to control the development of applications; it should remain small and work mainly by correspondence inter-sessionally.
3. To be accepted during a Working Group meeting, any new proposed methodology should clearly demonstrate that it is a real improvement compared to the traditionally used ones and should be carefully controlled. Proposals of new methods should ensure that applications fulfil QC criteria as quickly as possible.
4. Quality Control should be considered in a broad way, including both data collection in all of its components (catch and effort, but also size, sex and species composition) and data processing (e.g., catch at size and age), this being within the responsibilities of the Statistics Sub Committee. Consequent perceptions of data quality should be included in the assessment models.

Agenda Item 3. – CPUE Standardizations

1. Of the options examined, using the ratio of catch of the species of interest to the total catch as a variable to define targeting tended to perform better than the other options, but in some circumstances even that method could be misleading. The long term solution appears to be data collection of detailed effort characteristics. Taking in account the spatial distribution of the resources and its dynamics should be encouraged, including all the available environmental information and developing new types of non linear models such as GAM.

Agenda Item 4 – Stock rebuilding and assessment methods

1. It is recommended that assessment methods used to follow stock rebuilding maintain consistency during this process; the current model should be the null hypothesis and the working group should be sufficiently convinced that any new method is an improvement in order for a change to occur. Should a change occur, the results of the new model should also be considered in light of the old model.

Agenda Item 5 – Other matters

1. Subsequent meetings of the ICCAT Assessment Methods Working Group should focus on specific issues in order for the Working Group to have an impact on SCRS policy. The terms of reference for each specific meeting should address priority issues identified by the SCRS.
2. In order to further a broader international development in issues of stock assessment and related matters, the Working Group recommends that SCRS make contact with other scientific fisheries advisory organisations (e.g. IOTC, IATTC, ICES) in order to formulate terms of reference of common interest for a methodological Working Group with a wider membership.

7. Adoption of the report and closure

The report was adopted during the meeting. Due to the interest in this Working Group by the SCRS in general, it is important to have the report translated into all official languages well in advance of the SCRS.

At the closure of the meeting the Chairman thanked all the participants for their hard work and valuable contributions. The meeting was adjourned.

8. Literature cited and background documents

8.1 Documents prepared for the meeting

GAERTNER, D., and A. Fonteneau. Approximate estimate of the MSY from catch data without effort information. Application to tuna fisheries. **SCRS/00/36**. This volume.

GEROMONT, H.F., and D.S. Butterworth. Possible extensions to the ADAPT VPA model applied to western North Atlantic bluefin tuna, addressing in particular the need to account for “additional variance”. **SCRS/00/33** This volume.

LASSEN, H., and H. Sparholt. ICES quality assurance policy for fish stock assessment and management advice. **SCRS/00/38**. This volume.

LEGAULT, C.M., and C.E. Porch. Comparisons of weighting schemes for tuned virtual population analyses. **SCRS/00/35**. This volume.

MAURY, O. Multi-fleet non-equilibrium production models including stock surface to estimate catchability trends and fishery dynamics in a Bayesian context. Application to the skipjack tuna fishery (*Katsuwonus pelamis*) in the Atlantic Ocean. **SCRS/00/37**. This volume.

MCALLISTER, M., E.A. Babcock, and E.K. Pikitch. Using Bayesian methods and decision analysis as a rational basis for dealing with conflicting stock assessment results while providing management advice on stock rebuilding. **SCRS/00/34**. This volume.

MEDLEY, P.A.H. Integrating CPUE standardization within stock assessment. **SCRS/00/39**. This volume.

RESTREPO, V.R. Proposed stock assessment quality control procedures for ICCAT. **SCRS/00/40**. This volume.

8.2 Background documents distributed at the meeting

ANON. Report of the bluefin tuna methodology session (Excerpts). *ICCAT Coll. Vol. Sci. Pap.* Vol. 46-1, pp 187-268.

ANON. Detailed swordfish 1999 assessment report (Excerpts). In prep.

BROWN, C.A., and C.E. Porch. A numerical evaluation of lognormal, delta-lognormal and Poisson models for standardizing indices of abundance from West Atlantic bluefin catch per unit effort data (preliminary results). *ICCAT Coll. Vol. Sci. Pap.* Vol 46-4 . pp 476-482.

BUTTERWORTH, D.S., J.N. Ianelli, and R. Hilborn. In prep. A statistical time-series model for stock assessment of southern bluefin tuna. MARAM, University of Cape Town, Rondebosch 7701, South Africa.

BUTTERWORTH, D.S., and H.F. Geromont. In prep. Simulation testing as an approach to evaluate the reliability of assessment methods: An example involving initial consideration of the one/two

stock hypothesis for North Atlantic bluefin tuna. MARAM, University of Cape Town, Rondebosch 7701, South Africa.

COOKE, J.G. A procedure for using catch-effort indices in bluefin tuna assessments (revised). *ICCAT Coll. Vol. Sci. Pap.* Vol. 46-2. pp 228-232.

COOKE, J.G., and K. Lankester. Consideration of statistical models for catch-effort indices for use in tuning VPAs. *ICCAT Coll. Vol. Sci. Pap.* Vol. 45-2. pp. 125-131

GLAZER, J. 2000. Some thoughts on the by-catch issue for West Coast hake. Marine and Coastal Management, Private Bag X2, Rogge Bay, 8012, South Africa.

GLAZER, J., and D.S. Butterworth. In prep. GLM-based standardization of the South African West Coast hake catch per unit effort series, focusing on adjustments for targeting at other species. Marine and Coastal Management, Private Bag X2, Rogge Bay, 8012, South Africa.

MEJUTO, J., and J.M. de la Serna. Standardized catch rates by age and in biomass for the North Atlantic swordfish (*Xiphias gladius*) from the Spanish longline fleet for the period 1983-1998 and bias produced by changes in the fishing strategy. *ICCAT Coll. Vol. Sci. Pap* Vol. 51 pp. 1387-1411.

O'BRIEN, C.M., and L.T. Kell. The use of generalized linear models for the modelling of catch-effort series. I. Theory. *ICCAT Coll. Vol. Sci. Pap.* Vol. 46-4. pp. 476-482.

PALMA, C., J. Pereira, J. Mejuto, and M. Santos. Effect of simulating targeting levels on swordfish (*Xiphias gladius*) standardised CPUE estimates, caught by the Portuguese surface longline fleet in the North Atlantic. *ICCAT Coll. Vol. Sci. Pap* Vol 51 pp. 1718-1727.

PORCH, C.E. The implications of using the frequency of zero catches and other measures as indices of abundance. *ICCAT Coll. Vol. Sci. Pap.* Vol. 46-2. pp. 237-241

PUNT, A.E., R.W. Leslie, and A.J. Penney. A preliminary examination of the Taiwanese longline catch and effort data (1967 to 1992) for South Atlantic albacore (*Thunnus alalunga*). *ICCAT Coll. Vol. Sci. Pap.* Vol. 43 pp. 225-245.

TAKEUCHI, Y., and K. Yokawa. A note on methods to account targeting in CPUE standardization. *ICCAT Coll. Vol. Sci. Pap* Vol 51 pp. 2280-2286.

8.3 Other citations

GILKS, W.R., Richardson, S. and Spiegelhalter, D.J. 1996. Markov Chain Monte Carlo in Practice. Chapman and Hall, London. 486p.

MYERS, R.A., W.A. Blanchard, and K.R. Thompson. 1990. Summary of North Atlantic fish recruitment 1942-1987. Can. Tech. Rep. Fish. Aquat. Sci. 1743. 108pp.

RAPPORT DU GROUPE DE TRAVAIL ICCAT SUR LES MÉTHODES D'ÉVALUATION DES STOCKS (Madrid, 8-11 mai 2000)

1 Adoption de l'ordre du jour et organisation de la réunion

Conformément aux décisions prises par le Comité scientifique lors de ses réunions de 1998 et 1999, le Groupe de travail sur les Méthodes d'évaluation des stocks s'est réuni au Secrétariat de l'ICCAT les 8-11 mai 2000. La réunion a été présidée par le Dr Victor Restrepo, qui a ouvert les débats et souhaité la bienvenue à tous les participants.

L'objectif principal de ce nouveau groupe de travail est de *“mettre en place une gestion de qualité des méthodes d'évaluation des stocks, dans l'optique d'effectuer l'examen, la vérification et la documentation des méthodes d'évaluation qui sont utilisées à l'heure actuelle par le SCRS”*. Pour la première réunion du Groupe de travail, les objectifs suivants ont été définis et diffusés en avance:

- (a) *Élaboration d'un protocole de révision des méthodes et l'ordre de priorité des tâches futures du groupe,*
- (b) *Évaluation des méthodes de standardisation de la CPUE qui tiennent compte du ciblage,*
- (c) *Évaluation des méthodes de standardisation de la CPUE qui tiennent compte de l'hétérogénéité spatiale,*
- (d) *Avis à fournir sur les mécanismes d'évaluation permettant de suivre le rétablissement des stocks.*

L'ordre du jour provisoire a été adopté par le Groupe de travail et figure ci-joint en **Appendice 1**.

Les scientifiques suivants ont assumé la tâche de rapporteur:

Sections 1, 7 et 8	Victor Restrepo
Section 2	Elisabeth Babcock et Jaime Mejuto
Section 3	David Die et Naozumi Miyabe
Sections 4 et 5	Clay Porch et Christopher Legault
Section 6	Renaud Pianet

Il a été décidé de rédiger également un bref Résumé exécutif à l'intention du SCRS et de la Commission. Le Président du SCRS, J.E. Powers, a convenu d'en préparer la première ébauche.

2 Systèmes de gestion de la qualité

2.1 Présentation des documents

Deux documents ont été présentés au titre de ce point de l'ordre du jour. Ceux-ci ont servi de base à la plupart des débats.

Le document SCRS/00/38 décrit la politique de la garantie de qualité que le CIEM utilise dans l'évaluation des stocks de poisson, et les conseils de gestion. La politique ne met pas l'accent sur le choix

des méthodes d'évaluation, mais plutôt sur les mécanismes permettant de déterminer si une méthode fonctionne correctement. La politique a été adoptée en octobre 1999, sur la base des objectifs de consensus scientifique, d'approches (écosystème) intégrées, de crédibilité envers les clients, transparence, réaction, et indépendance vis-à-vis de toute influence politique.

Les composantes du contrôle de la qualité sont comme suit: (1) procédure de qualité pour les groupes d'évaluation; (2) intemporalité; (3) transparence, notamment la capacité de suivre le flux de données dans les bases de données, et le flux des postulats et décisions à travers les évaluations, les rapports de groupes de travail, les rapports consultatifs, les manuels aux fins de la procédure d'évaluation générale et les manuels pour les évaluations en cours de développement; et (4) peer review interne. La peer review intervient entre l'évaluation et la présentation des avis de gestion, afin de déterminer si l'évaluation est raisonnable et si les projections sont fondées sur une évaluation raisonnable. Cette peer review s'est convertie en une procédure plus officielle et mieux détaillée, et en sous-ensembles du Comité consultatif sur la Gestion des Pêcheries (ACFM), avant la réunion de l'ACFM. Des reviews externes n'ont pas eu lieu, mais celles-ci seront périodiques.

Le document SCRS/00/40 a présenté une proposition de procédure de contrôle de la qualité de l'évaluation des stocks par l'ICCAT. Les niveaux auxquels devrait être appliqué le contrôle de la qualité sont les suivants:

- (a) Collecte des données brutes;
- (b) Ajout de (a) pour soumission au Secrétariat;
- (c) Récapitulation des analyses de (a) ou (b) par le Secrétariat à l'intention des groupes de travail sur les espèces;
- (d) Analyse de (a) par des scientifiques nationaux à l'intention des groupes de travail sur les espèces;
- (e) Mise au point de méthodes d'évaluation;
- (f) Mise en oeuvre de méthodes d'évaluation;
- (g) Préparation des valeurs d'entrée pour les programmes d'évaluation;
- (h) Sélection des options disponibles pour exécuter les programmes d'évaluation;
- (i) Validation des résultats compte tenu des données de sortie du programme, des diagnostics et de l'expertise;
- (j) Récapitulation et synthèse des résultats;
- (k) Formulation d'avis cohérents avec les résultats.

Les points (a)-(c) tombent sous la responsabilité du Sous-Comité des Statistiques; les points (j)-(k) incombent au Groupe de travail *ad hoc* sur l'Organisation du SCRS; et les points (d)-(i) reviennent au présent Groupe de travail.

2.2 Recommandations en matière de contrôle de la qualité des évaluations des stocks

Le protocole du contrôle de la qualité a pour but de s'assurer que les évaluations sont scientifiquement fiables, qu'elles respectent les pratiques établies, qu'elles sont transparentes et bien documentées, sans pour autant sacrifier l'innovation méthodologique. Une partie du contrôle de la qualité devrait être réalisée par des scientifiques nationaux, et l'autre partie par le Secrétariat ICCAT. Le protocole établirait des normes pour les valeurs d'entrée, les diagnostics et la programmation. Le protocole n'a pas l'intention de restreindre la souplesse dont jouissent les groupes de travail dans le choix de la méthodologie d'évaluation.

Il a été recommandé de cataloguer les applications ICCAT qui ont été approuvées (c.-à-d. le logiciel). Le catalogue comprendrait les éléments énumérés en **Appendice 3** (voir Section 2.2.3). On n'évaluerait pas le bien-fondé relatif des méthodes, mais on vérifierait plutôt si le logiciel exécutant la méthode fonctionne correctement et est adéquatement documenté.

Le protocole pourrait être mis en vigueur, soit en spécifiant qu'une application ne peut être utilisée à moins qu'elle ne soit cataloguée, soit en exigeant que les rapports des groupes de travail précisent si chaque résultat a été produit avec une application approuvée. Le protocole ne pourrait entrer en vigueur avant au moins deux ans.

2.2.1 Analyses des scientifiques nationaux

Le contrôle de la qualité des analyses préparatoires incomberait principalement aux scientifiques nationaux, qui devraient documenter leurs méthodes et effectuer des diagnostics standard. Les questions à examiner seraient de savoir si la méthode utilisée était appropriée et si des erreurs avaient été commises dans l'application. Il faudrait présenter de nouvelles méthodes d'analyse avec essais de simulation, et documenter leurs différences par rapport aux méthodes antérieures. Les laboratoires nationaux devraient documenter leurs processus de contrôle de la qualité (ces derniers devront établir des normes; le SCRS souhaitera peut-être envisager d'établir des directives pour les standards minimum).

Le Groupe de travail a recommandé que les scientifiques nationaux: (1) tentent de reproduire indépendamment leurs résultats; (2) documentent leurs résultats et les raisons motivant leur refus d'alternatives; et (3) effectuent des diagnostics standard des méthodes employées. Il a été reconnu que de nombreux scientifiques nationaux le faisaient déjà dans les travaux présentés aux groupes de travail sur l'évaluation, et qu'ils apportaient également leurs données afin que d'autres méthodes puissent être appliquées lors de l'évaluation.

Les scientifiques nationaux de différents pays sont encouragés à élaborer conjointement des projets visant à résoudre des problèmes d'analyse de données communs, comme la standardisation de la CPUE. Si les scientifiques nationaux mettaient en commun leurs données afin que d'autres puissent reproduire leurs analyses, la garantie de qualité en serait améliorée. Toutefois, des questions de confidentialité entrent en jeu.

2.2.2 Élaboration de méthodes d'évaluation

L'ICCAT encourage l'élaboration de nouvelles méthodes. Des débats ont eu lieu sur la question de savoir si ce Groupe de travail devait fournir des critères permettant d'évaluer le bien-fondé relatif d'une méthode, par ex. si la méthode devait être publiée dans un journal, ou faire l'objet d'un examen externe ou interne. Le Groupe de travail s'est aussi demandé s'il fallait créer des directives pour évaluer les méthodes nouvelles, en étudiant dans quelle mesure les estimations de points de référence sont sensibles aux changements structuraux intervenus dans le modèle, s'il y avait une "valeur ajoutée" par le modèle, etc. La conclusion est que: (1) le logiciel à mettre au point a suggéré que les nouvelles méthodes devaient être soumises pour être cataloguées 5 à 10 mois avant la réunion du groupe de travail sur l'évaluation, de façon à ce que l'application puisse être approuvée; et (2) les auteurs devraient expliquer de quelle manière la méthode diffère de la précédente et pourquoi il s'agit d'une amélioration.

On s'est demandé si les applications utilisées dans les analyses de sensibilité devaient être soumises à un processus bureaucratique. On a souligné que des changements survenus lors d'une évaluation dans les valeurs d'entrée du modèle peuvent suggérer des changements dans le modèle qui devrait être utilisé, du moins dans les analyses de sensibilité, bien qu'ils ne puissent pas être approuvés par le processus de catalogage. Toutefois, le fait de présenter une méthode qui a recours à un logiciel non testé, même dans une analyse de sensibilité, risque de donner trop de poids aux résultats non testés. Il appartient au Groupe de travail de peser le pour et le contre, et de trancher entre le risque d'introduire un virus et celui d'ignorer un modèle plausible.

Le Groupe de travail est convenu que les méthodes devraient faire l'objet d'un examen interne et externe.

Le Groupe de travail a recommandé que les personnes qui proposent de nouvelles méthodes dressent une liste de diagnostics standard pour chaque méthode.

2.2.3 Mise en oeuvre des méthodes d'évaluation

Il a été recommandé d'élaborer un catalogue d'applications qui contiendrait les éléments répertoriés en **Appendice 3**. Le Groupe de travail recommande qu'un Comité de coordination soit mis en place pour contrôler ce catalogue. Le Comité se composerait de trois personnes: le Dynamiste de Population du Secrétariat, plus deux scientifiques nationaux désignés par le Président du SCRS. Lorsqu'une application serait soumise à l'examen du Comité de coordination, celui-ci déciderait si les auteurs avaient adéquatement fourni tous les éléments répertoriés en **Appendice 3**, et évaluerait si l'application était bien documentée et fonctionnait comme prévu. Les commentaires du Comité seraient ajoutés à l'information contenue en **Appendice 3** sous le point 16 de l'ordre du jour. Les applications approuvées par le Comité seraient ensuite ajoutées au catalogue d'applications qui serait mis à la disposition des groupes de travail pour mener à bien leurs évaluations.

Les méthodes acceptées dans le catalogue ICCAT doivent faire figurer toutes les informations en **Appendice 3**, et prouver le bien-fondé d'un nouveau programme. Le Comité de coordination devrait être en mesure de décider du niveau de détail requis pour chaque élément en **Appendice 3**. Il a été suggéré qu'il conviendrait d'établir quelques jeux de données ICCAT standard destinés à vérifier si un nouveau code obtenait les mêmes résultats qu'un code établi, lorsqu'on l'appliquait à la même question. Les données de test ne devraient pas être utilisées pour évaluer le bien-fondé d'une méthode, étant donné que des méthodes utiles risquaient de donner de mauvais résultats avec certaines données simulées.

On s'est demandé si les auteurs pouvaient envoyer les nouvelles applications au Comité de coordination, ou bien si les applications devraient être désignées par les présidents des groupes de travail, ou par les chefs de délégations nationales, ou par le président du SCRS, ou par une autre autorité. L'approbation d'une application prendrait un certain temps (environ 8 mois). La priorité devrait être donnée aux méthodes.

Le Comité de coordination devrait recevoir les nouvelles versions des programmes catalogués, ainsi qu'une description des changements intervenus depuis les versions précédentes.

Il a été suggéré qu'une programmation orientée sur l'objet améliorerait le problème du test par logiciel en permettant de modifier une application avec un minimum de changement de codes. Les tableurs et autres applications de langage de programmation sophistiquées peuvent être communiqués au Secrétariat sans que toute l'information soit fournie en **Appendice 3**.

Il n'est pas recommandé que les utilisateurs des applications soient certifiés ou qu'ils aient suivi une formation officielle. Toutefois, il serait utile que davantage de participants du groupe de travail puissent utiliser les applications; à cet égard, le catalogue est un bon outil d'enseignement.

Certains membres du Groupe de travail se sont déclarés préoccupés par le fait que ce protocole risquait de ralentir le processus d'innovation. Cependant, si le catalogue inclut un programme "mère", l'approbation d'une variation dans la méthode devrait être assez rapide (quelques mois). Le Comité travaillera par e-mail entre les sessions.

2.2.4 Liste de contrôle pour les valeurs d'entrée

Le Groupe de travail a recommandé que des fichiers de valeurs d'entrée pour chaque passage soient diffusés dans les groupes de travail sur l'évaluation, de façon à ce que le groupe puisse contrôler les erreurs autant que possible dans le temps disponible. Les modèles devraient avoir un format de valeurs d'entrée standard, ou dégager les valeurs d'entrée dans un format standardisé aux fins de leur

vérification. Le fichier logarithmique de passage devrait préciser les fichiers de valeurs d'entrée, les fichiers de sortie et les options choisies pour chaque passage, et devrait être créé automatiquement. Il faudrait soit marquer les fichiers de sortie de façon à ce qu'ils indiquent s'ils proviennent du modèle définitif, soit inclure dans le rapport du groupe de travail un numéro identifiant le modèle définitif.

2.2.5 Vérification des options du modèle

Lors des évaluations, la personne désignée par le président du groupe de travail pour vérifier les valeurs d'entrée du modèle devrait aussi vérifier les options du modèle. Le carnet de passage devrait enregistrer les options du modèle.

2.2.6 Validation des résultats compte tenu des sorties du programme, des diagnostics et de l'expertise

Pour les évaluations mêmes, le contrôle de la qualité consisterait à:

- (1) Vérifier les valeurs d'entrée (par ex. est-ce que la prise par âge totalise la prise totale?),
- (2) Vérifier et relever les options de passage,
- (3) Cataloguer, documenter et stocker tous les passages.

Les évaluations devraient également faire l'objet d'une peer review environ tous les 5 ans.

Il a été souligné que les groupes de travail sur les espèces devaient conserver le droit d'évaluer les méthodes d'après leur propre expertise, et de décider quelles données d'entrée et quels postulats étaient appropriés.

S'il s'avère nécessaire de modifier légèrement une application lors d'une évaluation, les groupes de travail concernés devront se prononcer entre le risque d'erreurs de programmation et la flexibilité (le présent Groupe de travail pourrait néanmoins fournir des directives). Les présidents des groupes de travail devraient se méfier des changements de codification, étant donné que les erreurs nuisent à la crédibilité du processus d'évaluation. On pourrait accepter des variations de méthode pendant un groupe de travail s'il s'agissait de changements modiques, transparents, présentant de réels avantages et qui auraient été évalués au moyen de diagnostics standard. Il est souhaitable qu'à la réunion deux personnes effectuent indépendamment les changements, ceci afin d'éviter des erreurs de codification. Si un programme est actualisé pendant une réunion, la nouvelle version devrait être remise au Secrétariat.

2.3 Contrôle de la qualité en aval

Les procédures de contrôle de la qualité examinées ci-dessus étudient les évaluations en amont. Or, on s'est également demandé s'il convenait d'examiner en aval le processus d'évaluation. On évaluerait les nouvelles méthodes en termes d'objectifs de gestion, ensuite les buts de la méthode, le caractère approprié des composantes du modèle (modèle de Poisson, pêche, estimation des paramètres, valeurs d'entrée), les aspects conceptuels du modèle (rigueur statistique, réalisme biologique, traitement de l'incertitude), et finalement les aspects pratiques (données requises, facilité d'utilisation, transparence). Le Groupe de travail a conclu que, si la garantie de qualité est indispensable à tous les niveaux du processus d'évaluation, il revient à chaque groupe d'espèce d'en choisir la méthode. Le protocole présentement élaboré ne couvre que la garantie de qualité du logiciel d'application.

Les choix suggérés de méthodes (par les groupes de travail sur les espèces) comprenaient les critères du Conseil international pour l'Exploration de la Mer (CIEM), documentés et révisés (peer review), et

dont la simulation a été testée. Les méthodes devraient également présenter des postulats clairs, et fournir l'information nécessaire pour la gestion.

2.4 Observations sur d'autres questions

Le Groupe de travail a recommandé que le Sous-Comité des Statistiques du SCRS incorpore la garantie de qualité dans un sens large, en tenant compte des besoins des évaluations. On s'est demandé si l'ICCAT pouvait influencer sur les standards minimum de la déclaration des données et de la garantie de qualité par nation. Il a été recommandé de fournir des estimations de l'incertitude avec les données, de façon à ce que l'incertitude puisse être incorporée aux évaluations. Il a également été recommandé d'améliorer la garantie de qualité des données de prise par âge du Secrétariat. Les modifications qui vont avoir lieu dans la base de données ICCAT devraient faciliter la vérification des données par le Secrétariat.

Il a aussi été suggéré que, pour éviter le problème qui se poserait si de nouvelles données modifiaient les méthodes, ce qui entraînerait des décisions peu réfléchies lors des évaluations, il fallait séparer la collecte des données des processus d'évaluation. La compilation des données et les séances d'évaluation devraient avoir lieu séparément lors de réunions tenues tout au long de l'année. Alternativement, les évaluations devraient avoir recours aux données compilées l'année précédente. Il serait coûteux de tenir séparément des réunions sur les données et sur les évaluations. Toutefois, la question est portée à l'attention du Groupe de travail *ad hoc* sur l'organisation du SCRS.

Le SCRS a examiné et continue d'examiner la question de savoir comment faire le point sur les évaluations.

3 Standardisation des indices de CPUE

Le SCRS et deux groupes de travail sur les espèces ont demandé au Groupe de travail ICCAT sur les Méthodes d'évaluer et de formuler des recommandations spécifiques sur les méthode de standardisation de la CPUE qui tiennent compte des questions de ciblage et d'hétérogénéité spatiale. Un document rédigé spécifiquement dans le but de faciliter les débats a été présenté à cette réunion (SCRS/00/39), mais plusieurs autres documents traitant de la même question, mais n'ayant pas été préparés pour cette réunion, ont été présentés comme documents de fond, et ont été examinés par le Groupe de travail.

3.1 Établir le lien entre l'évaluation des stocks et la standardisation de la CPUE

Dans le document SCRS/00/39, Medley fait observer qu'en général la standardisation de la CPUE se fait indépendamment de l'évaluation des stocks, comme première étape de l'analyse. On a recours à la standardisation pour atténuer la variabilité de la CPUE jugée ne pas être liée à des fluctuations de l'abondance. Malheureusement, il se peut qu'une partie de la variabilité disparue ait été informative pour les estimations d'abondance, et serait par conséquent perdue si la standardisation avait lieu séparément. Medley propose que la standardisation de la CPUE soit menée simultanément avec l'évaluation des stocks dans une seule analyse. Cette opération peut s'effectuer simplement en élaborant le modèle de la CPUE indépendamment, mais en intégrant ensuite le modèle de la CPUE avec le modèle de la dynamique de population dans un processus d'estimation unique. Le point fort de cette démarche est que le modèle de population concorde avec le traitement des données de CPUE. Il convient de noter, toutefois, que cette approche n'est pas la seule à garantir la cohérence entre le modèle de population et la standardisation de la CPUE.

Ces modèles intégrés tendent à avoir des composantes linéaires pour la standardisation de l'effort et des composantes non-linéaires pour le modèle de population. On peut utiliser des programmes pour obtenir des solutions numériques pour ces modèles hybrides.

Ces types de modèles intégrés peuvent servir à aborder toute question de modélisation en ce qui concerne la standardisation de l'effort, et peuvent donc être employés pour traiter des questions de ciblage et d'hétérogénéité spatiale. Ces applications font l'objet d'une plus ample explication dans la section ci-dessous. Il a été suggéré d'avoir recours à un exemple pratique pour évaluer la démarche.

3.2 Traitement du ciblage dans les pêcheries d'espèces mixtes

Les pêcheries ICCAT se caractérisent par la prise simultanée de plus d'une espèce. Historiquement, nombre de ces pêcheries constituaient à l'origine des espèces uniques ou possédaient des espèces cibles clairement établies. Au fur et à mesure que les pêcheries ont évolué, les pratiques de ciblage ont souvent changé et risquent même de changer de façon saisonnière. Bien que dans de rares cas un registre des prises fournisse une information directe sur l'espèce-cible, la plupart du temps cette dernière doit être déduite des informations sur les caractéristiques des engins. Elle peut aussi être totalement absente, auquel cas il faut l'inférer des ratios des différentes espèces capturées. Cette section fait d'abord une présentation des principales questions relatives au ciblage des pêcheries ICCAT; elle donne ensuite un aperçu général des méthodes qui tentent de déduire des pratiques de ciblage à partir des informations sur les prises accessoires; elle se penche enfin sur la question de savoir comment incorporer cette information dans la standardisation de l'effort de pêche.

3.2.1 Aperçu général de l'information sur le ciblage disponible pour les pêcheries ICCAT

a) Pêche palangrière japonaise

Cette pêcherie a véritablement démarré après la Deuxième Guerre Mondiale et a initialement ciblé l'albacore et le germon pour la mise en conserve. Au fur et à mesure que le marché du sashimi s'est développé et que les prix du thon de qualité ont grimpé, la pêcherie a amorcé un changement de ciblage au profit du thon obèse et du thon rouge. Outre ces thonidés, d'autres espèces accessoires (par ex. istiophoridés, requins) sont capturées. Actuellement, les jeux de palangre comprennent 2000-3000 hameçons par jeu et, en moyenne, les prises sont de 20-30 poissons par jour, bien que les prises maximum atteignent environ 200 poissons par jour. Dans les conditions actuelles, il ne paraît pas y avoir saturation des engins ces dernières années, même si la majeure partie du poisson pourrait être capturée dans une section plus petite de la palangre.

Les renseignements sur le ciblage peuvent être obtenus à partir de l'information sur le nombre de hameçons par panier et la profondeur de la palangre. Les paniers dirigés au thon obèse contiennent plus de 10 hameçons, ceux utilisés pour capturer le thon rouge comptent 7 à 9 hameçons. Il convient de noter que l'agencement des hameçons peut varier en fonction de l'espèce visée par les pêcheurs lors d'une sortie. On observe de fortes modifications saisonnières au niveau du ciblage, le thon rouge étant par exemple ciblé principalement en hiver et au printemps. Chacun sait que d'autres espèces, comme l'albacore et les istiophoridés, sont rarement ciblées, même si elles peuvent constituer un élément important de la prise.

b) Pêche palangrière d'espadon.

D'importantes modifications au niveau des engins ont vu le jour au fil des ans, notamment en ce qui concerne le type de ligne. Ces modifications ont été en partie motivées par des gains économiques qui ont conduit les pêcheurs à cibler des requins ou d'autres espèces en modifiant l'heure, les zones et l'engin utilisé dans leurs jeux de palangre. La plupart de ces modifications sont minimales, et les données disponibles ne permettent pas d'estimer clairement le ciblage.

c) Pêche à la senne de thonidés tropicaux.

Historiquement, la pêche à la senne des thonidés tropicaux a ciblé l'albacore et le listao. Le thon obèse et les petits thonidés (AUX et LTA) ont constitué les espèces accessoires les plus importantes, les

autres espèces accessoires sont négligeables (environ 1% de la prise totale). Les données disponibles dans les livres de bord comprennent les prises des trois principales espèces de thonidés (albacore, listao et thon obèse), et de certains petits thonidés. Les données concernant d'autres espèces accessoires proviennent en partie d'observateurs.

Une caractéristique de ces flottes est l'amélioration constante de l'équipement technique dont ont bénéficié les bateaux, perçue comme une augmentation constante de leur puissance de pêche. Les changements d'équipement intervenant très rapidement pour toutes les flottes, les différents niveaux des facteurs technologiques (radar à oiseaux, sonar, taille de la maille, profondeur, etc.) se chevauchent donc rarement dans le temps, ce qui entrave la détection de leurs effets sur la CPUE. En outre, l'habileté du capitaine constitue un facteur important qu'il convient de prendre en compte dans l'estimation de la puissance de pêche d'un bateau. La plupart de ces données ne sont pas actuellement disponibles.

Le recours accru à la pêche sous DCP (dispositif de concentration de poissons) depuis 1991 a entraîné un changement important dans cette pêcherie. Avant l'introduction des DCP, la pêche sous objet flottant naturel se limitait à une zone et à une période spécifiques, et les prises sous objet flottant naturel représentait environ 15% de la prise totale. Avec l'introduction des DCP, les prises atteignent maintenant presque 50%. La composition spécifique et des tailles varie sensiblement en fonction des modes de pêche: dans les prises réalisées en bancs libres, l'albacore constitue l'espèce dominante, suivie du listao. En revanche, dans les prises réalisées sous DCP, le listao représente plus de 70%, et les jeunes albacores et le thon obèse environ 15% chacun. Depuis 1990, l'information sur les modes de capture est disponible dans les livres de bord. Ce mode de pêche a donné lieu à un grand nombre d'éléments et de problèmes nouveaux (à ajouter aux différentes compositions spécifiques) qu'il convient d'examiner dans la standardisation de l'effort. Premièrement, il est impossible de séparer l'effort par le mode de pêche de l'information disponible dans les livres de bord; deuxièmement, il y a une grande interaction entre les facteurs, tels que saison et zone, mode de pêche, nombre de DCP utilisés et capitaine (car la pêche en bancs libre exige plus d'habileté). Les conditions océanographiques (par ex. la profondeur de la thermocline, la vitesse du vent, la température à la surface, etc.) sont également des facteurs susceptibles d'être pris en compte dans la standardisation de l'effort.

Plus récemment (1997-1999), la fermeture temporelle (3 mois) d'une vaste zone de pêche sous DCP située autour de l'Équateur a introduit un nouvel élément qu'il convient d'examiner dans la procédure de standardisation.

3.2.2 Déduction du ciblage d'après la capture ou le taux de capture des espèces-cibles secondaires et des espèces accessoires

On peut directement évaluer le ciblage grâce à l'information disponible sur les pratiques de pêche, qui indique si une espèce donnée est ciblée ou pas. Il est donc recommandé de recueillir dans la mesure du possible cette information. Or, bien souvent, celle-ci ne figure pas dans les registres de capture, auquel cas on peut tenter d'avoir recours aux informations sur les captures ou les taux de capture d'espèces-cibles secondaires ou d'espèces accessoires, afin d'élaborer des indices au moyen desquels l'effort de pêche peut être affecté à une espèce-cible particulière.

Pour les jeux de données de CPUE dépourvus de suffisamment de mesures directes des caractéristiques des engins et d'autres facteurs, les récentes évaluations des stocks ICCAT ont pu avoir recours à des indices approchants de ciblage fondés sur les prises de la cible d'origine (ci-après désignée comme l'"espèce-cible") et des espèces associées. Toutefois, des inquiétudes ont été exprimées quant à l'exactitude de ces indices approchants - on s'est demandé à cet égard si ces derniers indiquaient correctement les changements de ciblage - et quant à l'éventuelle présentation erronée de la tendance de l'abondance si ces indices approchants étaient appliqués pour produire des séries temporelles d'abondance relative (cf. SCRS/99/125). À titre d'exemple, si un indice approchant de ciblage est biaisé, la baisse réelle de l'abondance pourrait être incorrectement interprétée comme un changement de ciblage alors qu'en réalité ce changement n'aurait pas eu lieu. En 1999, le Groupe de travail sur l'Espadon

(Rapport 1999 sur l'Espadon) a identifié plusieurs méthodes visant à créer ces indices, et les a évaluées par simulation. Les simulations effectuées pendant cette réunion ont révélé qu'un de ces indices était plus robuste que les autres face à différents scénarios de modifications historiques du ciblage. Ces simulations se sont toutefois concentrées sur le cas de l'espadon, et n'ont pas par conséquent examiné toute la gamme des trajectoires éventuelles de l'abondance des espèces cibles et accessoires qui existe dans d'autres pêcheries ICCAT. Ces simulations n'ont pas non plus examiné tous les scénarios possibles de biais dû aux modifications historiques du ciblage et aux déclarations d'espèces accessoires inférieures à la réalité. Il a par conséquent été décidé à la présente réunion de réaliser de nouvelles simulations destinées à rendre les résultats applicables à une plus large gamme de stocks de thonidés.

Afin de déterminer si divers indicateurs de ciblage potentiels seraient biaisés par des changements systématiques intervenus dans la biomasse de l'espèce-cible, la biomasse de l'espèce non cible, l'effort total, la déclaration d'espèces non cibles, et diverses combinaisons de ce qui précède, des simulations ont été réalisées. Celles-ci ont projeté des trajectoires de population déterministes de la biomasse hypothétique d'espèces-cibles et non cibles, avec différentes tendances de ciblage par pêcherie, et elles ont calculé cinq indicateurs différents de ciblage à travers la série temporelle (décrite ci-dessous). Les cas identifiés pour l'évaluation se sont fondés sur certaines pêcheries ICCAT dans lesquelles on a soupçonné des changements systématiques de ciblage:

- (1) Espadon (espèce-cible primaire) et requin bleu (secondaire) (*cf.* SCRS/99/125)
- (2) Espadon (espèce-cible primaire), requin bleu (secondaire) et albacore (espèce non ciblée)
- (3) Listao (espèce-cible primaire) et albacore (secondaire)
- (4) Listao (espèce-cible primaire) et albacore (secondaire), mais où l'espèce secondaire est très faible en abondance par rapport à l'espèce-cible primaire
- (5) Listao (espèce-cible primaire), albacore (secondaire) et thon obèse (espèce non ciblée)

Une espèce-cible primaire est celle qui a été la seule espèce visée par une pêcherie pendant ses premières années. Une espèce-cible secondaire est celle qui est capturée en tant que prise accessoire pendant les premières années de la pêcherie, puis devient une espèce-cible pour une partie au moins de la flottille de pêche. Les espèces non ciblées sont celles qui ne sont pas visées, mais sont capturées en tant que prises accessoires.

Pour dresser des trajectoires réalistes de la population des stocks de poisson, un modèle de production de Schaefer a été utilisé, avec des valeurs de r et de K se rapprochant de celles des espèces simulées. Le **Tableau 1** indique les valeurs qui ont été postulées pour les paramètres du modèle de Schaefer pour ces différentes espèces.

La série d'effort total annuel a été spécifiée comme étant, soit constante, soit à croissance linéaire. Le véritable indice de ciblage est fourni par:

$$targ = \frac{Effort}{Effort_o + Effort_s}$$

où $Effort_o$ = l'effort portant sur l'espèce-cible primaire, et $Effort_s$ = l'effort portant sur l'espèce-cible secondaire. Dans les simulations, les valeurs d' $Effort_o$ et d' $Effort_s$ ont été obtenues en multipliant respectivement l'effort total spécifié pour chaque année par $targ$ et $(1-targ)$.

Deux types différents de valeurs de q ont été fixées pour chaque espèce: l'une lorsque l'espèce-cible primaire était visée ($q_{species, o}$), l'autre lorsqu'il s'agissait de l'espèce secondaire ($q_{species, s}$). Cette dernière valeur a été fixée à un niveau bien plus faible que la première. La prise annuelle totale de chaque espèce était donc:

$$C_{species} = (q_{species, o} * Effort_o + q_{species, s} * Effort_s) * B_{species}$$

Les indices de ciblage considérés pour l'espèce-cible primaire étaient les suivants:

Indice I Prise de l'espèce-cible primaire/prise totale des autres espèces lorsque le terme autres espèces se réfère à des espèces-cibles non-primaires

Indice II Prise de l'espèce-cible primaire/prise totale

Indice III $1/\text{taux de capture non-corrigé d'autres espèces} = \text{effort total}/\text{prise d'autres espèces}$

Indice IV $1/\text{prise d'autres espèces}$

Indice V Lorsque la prise de l'espèce-cible primaire dépasse la prise d'autres espèces

Chacun de ces indices de ciblage a été calculé sur une période de 15 ans selon chacun des cinq scénarios définis ci-dessus, puis mis en corrélation avec l'indice réel de ciblage de l'espèce-cible primaire. Trois situations différentes de ciblage ont été simulées pour chacun des cinq scénarios spécifiques. La première ne comportait pas de modification du ciblage ou de la déclaration. Dans la deuxième, le ciblage de l'espèce-cible primaire diminuait de façon linéaire de 100% à 20% en cinq ans. Dans la troisième, cette baisse se produisait également, mais accompagnée d'une hausse simultanée de 20% à 100% des déclarations concernant l'espèce-cible secondaire.

Un certain nombre de scénarios différents de trajectoires de la biomasse ont été simulés pour chaque scénario spécifique en manipulant les valeurs de q , la biomasse relative à K pour chaque espèce, et selon un effort constant ou croissant. Dans chaque cas, les indices de ciblage ont été portés sur graphique après avoir normalisé en divisant chaque valeur par la moyenne de la série.

Les séries de CPUE de l'espèce-cible primaire qui seraient estimées en appliquant chaque indice de ciblage ont également été calculées en tant que prise de l'espèce-cible divisée par (l'effort total x l'indice de ciblage normalisé entre 0 et 1).

Outre la mise en corrélation des séries de ciblage et des séries de CPUE qui en découlent, les résultats des divers scénarios ont été tabulés en termes de la somme des carrés des valeurs résiduelles entre chaque dérivée et les valeurs de l'abondance relative réelle dans chaque série.

La **Figure 1** illustre les résultats de la simulation qui utilise les indices I à IV pour les cas "listao" et "albacore" (l'indice V s'est avéré médiocre dans la plupart des cas, et a été éliminé des études ultérieures). Chaque jeu de panneaux de la **Figure 1** montre l'indice de ciblage estimé, ainsi que les séries correspondantes de CPUE. Les figures des autres cas ne sont pas représentées, car elles sont semblables au cas "deux espèces thonières". Le **Tableau 2** donne la somme des carrés des valeurs résiduelles des séries de CPUE découlant des divers scénarios obtenus par les quatre indices retenus calculés d'après les valeurs de capture. Le **Tableau 3** indique l'ordre d'importance des quatre indices de ciblage d'après la somme des carrés des valeurs résiduelles.

Lorsque l'on utilise, pour le seul total annuel, des indices approchants fondés sur le volume relatif de la prise de l'espèce-cible comparé à celle d'autres espèces, on peut s'attendre à ce qu'ils donnent des résultats fortement biaisés, et il faut donc les éviter. Toutefois, si la méthode est appliquée jeu par jeu, elle peut s'avérer plus appropriée. L'indice qui divise la prise de l'espèce-cible primaire par la prise totale est celui qui a le mieux fonctionné dans plus de la moitié des cas de scénarios. En revanche, il s'est avéré très médiocre par rapport aux autres dans environ un tiers des scénarios. En outre, dans nombre de scénarios, les indices, même celui qui fonctionnait le mieux, se sont avérés très médiocres et ont donné des résultats très spéculatifs quant à la tendance de l'abondance.

Les résultats de cette évaluation suggèrent qu'il n'existe pas clairement de "meilleure" méthode qui

puisse être appliquée de façon générale, et que tout indice approchant, même s'il fonctionne mieux que d'autres, peut encore donner des résultats très faussés. Il convient donc de consulter les résultats de la présente étude et d'effectuer un exercice similaire de simulation, lorsqu'il est intéressant de chercher à corriger le ciblage alors que l'on ne dispose que de la prise totale de chaque espèce et de l'effort total. Il serait à conseiller d'identifier les diverses conditions particulières de la pêcherie sous étude. Ceci comprendrait la tendance potentielle de l'effort total, l'abondance relative des différentes espèces capturées, la tendance éventuelle de leur abondance, et la question de savoir si des modifications sont intervenues dans les déclarations. Il faut se prononcer quant à l'application de l'une des alternatives si l'on a vu que l'une d'entre elles fonctionne de façon acceptable lorsque l'on utilise pour détecter la tendance potentielle de l'abondance.

Cette évaluation déterministe n'a évalué que le biais potentiel de quelques alternatives d'indice approchant du ciblage. Elle n'a pas évalué le degré potentiel de précision des alternatives d'indice approchant, ce qui est aussi un facteur important à considérer, une fois que l'on a identifié quelques alternatives avec un degré suffisamment faible de biais pour pouvoir être acceptées. Ce dernier point demande à être mieux suivi dans les simulations de Monte Carlo qui simulent la variabilité aléatoire des données de capture et d'effort, notamment pour définir quelle doit être la magnitude des changements du ciblage pour que ceux-ci puissent être détectés de façon fiable.

Par ailleurs, le modèle de Schaefer a servi à donner la tendance de la biomasse qui est réaliste du point de vue biologique dans ces simulations. Toutefois, la performance des indices approchant du ciblage dépend de la tendance de la biomasse, et non du modèle de dynamique des populations, si bien qu'il pourrait s'avérer utile de mener des simulations comme celle qui est effectuée ici, avec de simples tendances linéaires de la biomasse de chaque espèce.

La performance irrégulière des alternatives d'indice approchant de ciblage dans différents scénarios, et les biais importants que présentent fréquemment les résultats qui en découlent, suggèrent qu'il serait également souhaitable d'identifier d'autres indices, plus fiables, du ciblage. Les indices approchant qui pourraient aussi être élaborés et évalués comprennent ceux qui se fondent sur l'analyse des données sur le taux de capture par le modèle linéaire généralisé (GLM) (Glazier 2000). Il est également à conseiller de recueillir de façon régulière des données additionnelles sur la pêche, et de les analyser afin de disposer d'une base empirique solide pour identifier les changements du ciblage. Ces données additionnelles pourraient comprendre, par exemple, des mesures directes des caractéristiques des engins, et le mode de déploiement de ceux-ci (par exemple, profondeur des palangres et distance entre hameçons).

Parmi les méthodes d'indice approchant évaluées, l'utilisation du ratio de la prise de l'espèce-cible par rapport à la prise totale (indice IV, **Tableau 3**), était celle qui fonctionnait le mieux dans l'ensemble, et qui a réuni les suffrages, bien qu'il ne s'agisse pas forcément de la méthode qui assure la meilleure performance dans tous les cas. Ces évaluations n'ont pas testé les biais de l'utilisation de la CPUE en tant qu'indices de l'abondance relative lorsque l'on néglige l'incidence du ciblage. Cette source de biais pourrait s'avérer substantielle, et devrait être évaluée, dans le cadre des simulations décrites ci-dessus, afin de fournir une base pour la comparaison des coûts et bénéfices de l'application d'indices approchant du ciblage et du fait de ne pas tenir compte de l'incidence de ce dernier.

3.2.3 Introduction d'information sur le ciblage dans la standardisation de la CPUE

Glazer et Butterworth (en préparation) et Glazier (2000) ont envisagé un modèle qui tienne compte du ciblage lorsque la capture d'espèces accessoires est réduite par rapport à celle de l'espèce-cible, et que l'abondance des espèces accessoires ne montre pas de tendance. Ce travail montre que lorsqu'il y a une corrélation positive entre le taux de capture de l'espèce-cible et celui de la prise accessoire, il ne convient probablement pas d'utiliser l'information sur la prise accessoire pour l'espèce accessoire en tant que covariance. En revanche, lorsqu'il existe une corrélation négative, il peut alors s'avérer approprié d'inclure la prise accessoire en tant que facteur de standardisation de la CPUE. Malheureusement, nombre des questions concernant le ciblage des thonidés se réfèrent à des espèces mélangées où

l'abondance de l'espèce secondaire n'est que peu inférieure à celle de l'espèce primaire, si bien que les modèles ci-dessus ne conviennent pas. Dans ce cas, une option consiste à effectuer la standardisation de la CPUE en tant qu'analyse à variables multiples, où le taux de capture de l'espèce-cible comme des espèces accessoires sont des variables dépendantes. Ces analyses peuvent être menées à travers les bibliothèques spécialisées SPLUS. L'analyse de la composante principale (Principal Component Analysis) peut également aider à atteindre cet objectif, mais il faut choisir avec soin les variables indépendantes. Il serait préférable que les variables soient étroitement liées au processus de contrôle des modifications de la CPUE, par exemple la localisation d'un front, plutôt qu'à des grilles arbitraires basés sur la latitude et la longitude. Bien que ceci soit souhaitable du point de vue conceptuel, nous ne disposons pas souvent de meilleurs indices approchants que la location, et lorsqu'ils sont disponibles, ils ne remontent pas suffisamment dans le temps pour que les analyses soient vraiment utiles pour reconstituer les séries historiques de CPUE.

Les données jeu par jeu où l'on obtient de façon simultanée la CPUE de plusieurs espèces permettent d'isoler l'effet du ciblage des espèces (ou des indices approchants entre espèces) d'après la variabilité à court terme du taux global de capture, due par exemple à des facteurs environnementaux non-observés (ou qui ne peuvent pas l'être). L'**Appendice 4** présente, dans une structure modélifique multiplicative, la façon de réaliser ce travail selon la modélisation explicite du processus de ciblage et des paramètres de l'environnement à court terme tenant compte de l'ensemble des hâles, bons ou médiocres. Le travail doit être poursuivi avant de pouvoir évaluer le degré d'utilité de l'approche.

Lorsque la principale raison d'incorporer le ciblage dans la standardisation de la CPUE découle de la nécessité d'aborder le problème de la présence d'enregistrements de captures nulles de l'une des espèces, on peut utiliser un modèle delta lognormal pour prendre en compte ces enregistrements de façon explicite, et il n'est pas nécessaire d'incorporer d'autres variables de ciblage.

Le ciblage peut également être traité comme un problème de données manquantes. En général, les modèles de ciblage sont considérés comme simples lorsque l'espèce-cible est explicitement disponible en tant que variable statistique. Par exemple, dans le courant d'une sortie, un modèle de CPUE pourrait être:

$$m_i = (w e^{\hat{\alpha}_1} + (1-w) e^{\hat{\alpha}_2})$$

Où m est la CPUE attendue dans le GLM, β_1 les paramètres linéaires (effets année, position, bateau et autres) du ciblage de l'espèce 1, β_2 pour le ciblage de l'espèce 2 et w la proportion de l'effort ciblant la première espèce. Le problème est d'estimer les paramètres lorsque les données de ciblage (w) ne sont pas enregistrées. Si un nombre significatif de bateaux ont des données d'observateur lorsque l'espèce-cible est enregistrée, l'algorithme EM est peut-être une bonne solution pour combler les lacunes des données. L'algorithme EM donne les paramètres MLE et devrait tenir compte du ciblage dans le modèle définitif. La méthode impute des valeurs escomptées pour la variable cible, qui est la proportion des unités d'effort appliquées pour cibler les espèces. Cependant, la méthode générale devrait également aborder:

- des modèles plus complexes, y compris différents effets de capturabilité et des effets sur l'environnement,
- en l'absence de données de ciblage, seulement la composition spécifique des captures.

La première question est un élargissement de l'approche, et même si elle peut présenter une analyse numérique plus complexe, l'approche EM devrait en théorie être compatible. La deuxième question est beaucoup plus sérieuse, étant donné que par l'échantillonnage on risque de ne pas pouvoir solutionner le problème sans formuler de nouveaux postulats significatifs dans le modèle. Il est recommandé d'explorer la question de savoir si la méthode EM, EM stochastique (Gilks, Richardson et Spiegelhalter 1996) ou une autre méthode de données manquantes serait utile pour aborder le problème du ciblage.

Quel que soit le modèle de ciblage utilisé, il est important de se souvenir que toutes les séries de données CPUE ne peuvent donner un bon indice d'abondance. Les débarquements d'espèces accessoires par unité d'effort reflètent plus souvent des préférences commerciales que d'abondance. Dans les pêcheries plurispécifiques, les limites de capture imposées à des espèces spécifiques peuvent entraîner des changements dans les opérations de pêche et dans les pratiques de ciblage et, en dernier ressort, compromettre la cohérence des séries de données de CPUE.

3.3 Modèles traitant de l'hétérogénéité spatiale dans les données de CPUE

Un travail considérable a été réalisé dans la standardisation des indices de CPUE dans un contexte spatial pour les espèces de thonidés, particulièrement le thon rouge du Sud. Les données japonaises de pêche à la palangre indiquent un rétrécissement considérable de la zone de pêche au fil du temps. Ce phénomène peut s'expliquer par un resserrement de la zone où le stock abonde, ou par les effets de la réglementation en matière de pêche (TPA plus faible) qui ont conduit les pêcheurs à ne pêcher que dans les zones les plus productives du stock. Dans cette pêcherie, la CPUE a été standardisée en multipliant les estimations de la modélisation linéaire généralisée pour les différents mois et régions par la zone de la région, puis en faisant l'addition pour les différentes régions, et en calculant la moyenne des mois.

Malheureusement, ces analyses doivent traiter des régions dépourvues d'effort de pêche, mais à l'intérieur de zones de pêche historique. Deux hypothèses extrêmes ont été mises à l'essai: postuler que l'abondance dans ces régions est nulle, ou postuler qu'elle équivaut à la moyenne de toutes les grilles pêchées. Plus récemment, une troisième hypothèse a été tentée, à savoir faire appel à une technique d'interpolation géostatistique destinée à évaluer la densité dans les zones où l'on ne pratique pas la pêche une année donnée. On n'a pas trouvé de solution au débat portant sur le caractère approprié de ces techniques d'interpolation. Il a également été proposé d'utiliser l'information sur l'habitat pour aider à interpoler ces densités. Des modèles GAM ou mixtes pourraient être utilisés à cet effet.

Les tendances de l'abondance du thon rouge du Sud sont, malheureusement, très sensibles aux hypothèses susmentionnées, tout particulièrement au cours des dix dernières années. Les commentaires ci-dessus au sujet du thon rouge du Sud ne sont toutefois pas toujours applicables aux stocks de thonidés de l'Atlantique.

Une fois que ces facteurs d'abondance spatiale (=densité) ont été estimés d'après ces modèles, il vaut mieux calculer l'abondance absolue en multipliant simplement la densité par la taille de la région et en totalisant toutes les régions. Une fois fait, il ne faut pas toutefois oublier l'échelle spatiale de la région. Les régions trop petites pourraient connaître des problèmes de raréfaction locale, et risqueraient ainsi de générer des estimations de densité non fiables. La coopération entre pêcheurs (groupes codifiés) pourrait réduire considérablement la variance entre les bateaux du fait qu'ils partageraient l'information sur l'emplacement des grosses captures. Les recherches menées à l'IATTC et à l'ICCAT suggèrent que la coopération favorise tant l'efficacité que la CPUE peut même augmenter alors que l'effort de pêche est accru. On peut aussi interpréter ce phénomène en disant que les bateaux qui font partie d'un groupe codifié pêchent plus efficacement que ceux qui s'adonnent à la pêche individuellement.

Trois questions se posent habituellement lorsque des modèles GLM sont utilisés pour standardiser la CPUE à des fins d'hétérogénéité spatiale:

- l'opportunité d'avoir recours à des modèles d'effets aléatoires visant à réduire la complexité du paramètre (voir O'Brien et Kell SCRS/96/173)
- quelles interactions devraient être conservées dans le modèle
- quels diagnostics devraient être étudiés pour garantir que le modèle est approprié pour l'application

Dans le cas des interactions, il est recommandé de faire appel à des tests de ratio de probabilité destinés à vérifier si l'ajout de termes supplémentaires améliore statistiquement l'ajustement.

3.3.1 Incorporation de la dynamique spatiale des populations dans la standardisation de la CPUE

Si la gestion est menée de façon spatiale (fermetures spatiales), il est souvent justifié d'incorporer une structure spatiale dans le modèle d'évaluation des stocks. Il en va de même quand les caractéristiques du cycle vital comprennent une composante spatiale, par exemple lorsque les migrations sont reliées aux périodes du cycle vital.

Medley (SCRS/00/39) indique que l'on peut mettre au point des modèles destinés à examiner explicitement des processus tels que la structure spatiale et la migration, avec un modèle qui incorpore la standardisation de la CPUE. On peut démontrer que l'incorporation de ces composantes dans un modèle intégré peut améliorer l'ajustement des données simulées. Lorsqu'on utilise ces modèles, il faut toutefois tenir compte de l'échelle spatiale à laquelle la standardisation de la CPUE est établie dans le modèle intégré. En principe, il n'est pas primordial que l'échelle spatiale de la standardisation de la CPUE corresponde à celle du modèle de la dynamique des populations. Il sera toujours nécessaire d'obtenir des données de capture et d'effort regroupées (pré-traitement), comme cela se fait couramment dans la standardisation de l'effort (par exemple il est probable que même si les données sont recueillies à chaque remontée d'engins, celles-ci seront regroupées sur de plus grandes échelles avant de pouvoir être standardisées). Il ne faut pas oublier que toute série de CPUE standardisée obtenue à partir d'un modèle intégré dépendra du modèle de population spécifiquement utilisé. Or, ces séries peuvent toujours être employées dans d'autres analyses.

4 Rétablissement des stocks et méthodes d'évaluation

La poursuite de l'application de la méthode d'évaluation utilisée au moment de l'adoption d'un objectif spécifique de rétablissement est à la base du suivi de ce rétablissement. Ce processus devrait comprendre une analyse rétrospective en tant que diagnostic destiné à vérifier la fiabilité.

Les débats sur cette question se sont centrés sur trois thèmes principaux: 1) quand un changement de modèles est-il justifié, 2) comment le changement devrait-il être présenté aux gestionnaires, et 3) comment peut-on améliorer le paradigme de gestion.

4.1 Quand faut-il apporter un changement

Si un plan de rétablissement est déjà en place, est-il préférable de s'en tenir à la méthode d'évaluation en cours, ou bien d'adopter une nouvelle méthode perçue comme une amélioration? On s'est demandé avec inquiétude si un changement de méthodologie insuffisamment justifié ne risquait pas de porter atteinte à la crédibilité du processus d'évaluation, surtout si le changement donnait par la suite de mauvais résultats. Le Groupe de travail du CIEM préfère maintenir le plus de cohérence possible d'une année sur l'autre, mais il reconnaît néanmoins que certains changements sont inévitables, ne serait-ce qu'en raison de l'apparition de nouveaux types de données. On a également fait valoir que la science pouvait et devait améliorer avec le temps. Or, étant donné que les membres des groupes de travail ne sont pas toujours les mêmes au fil des ans, la perception de ce qui paraît être le "meilleur" modèle change parfois d'une évaluation à l'autre. En outre, certains membres ont suggéré que tout changement important présenté à la réunion d'un groupe de travail ne devrait pas être utilisé lors de cette évaluation, mais être plutôt renvoyé à la suivante, de façon à permettre aux participants de l'examiner en profondeur. Certains se sont demandé si le nouveau modèle ne risquait pas de modifier la perception de l'abondance absolue. Étant donné que l'ICCAT a recours à des repères relatifs, cela représente peut-être moins un problème, bien que le problème d'établir des TPA subsiste. Dans l'expérience du CIEM, la présentation des résultats à partir de l'ancien et du nouveau modèle s'est avérée essentielle, mais pas nécessairement suffisante pour convaincre les gestionnaires de la nécessité de recourir au nouveau modèle. En résumé, on a estimé que

l'hypothèse nulle devrait être le *statu quo*, et que la preuve devrait être considérable pour qu'un changement se produise, c.-à-d. que le groupe de travail devrait être convaincu de la justification du changement.

4.2 Comment présenter le changement

Si les gestionnaires sont habitués à une méthode qui est en vigueur depuis un certain temps, ils peuvent être plus à même d'écouter les conseils scientifiques. Souvent, quand un changement est réalisé, les scientifiques d'évaluation n'ont pas réussi à convaincre les gestionnaires du caractère opportun de ce changement. Ceci devient particulièrement évident lorsque la nouvelle méthode est spécialement favorable ou défavorable à un ensemble de communautés. On a estimé qu'un Résumé exécutif présenté dans un format cohérent et simplifié, et indiquant les repères essentiels pour la gestion, sans toutefois inclure les détails du modèle ou le jargon, serait utile à cet égard. Toutefois, étant donné que de nombreuses communautés participent à l'évaluation, tout changement de méthodologie devrait être documenté et justifié dans le rapport détaillé. Les raisons motivant un changement devraient être entièrement transparentes afin d'éviter des questions telles que "qu'est-ce qui n'allait pas avec le modèle de l'année dernière?".

4.3 Améliorations

Le Groupe de travail a estimé que la formule de gestion envisagée présentait des qualités avantageuses, mais il a ajouté qu'elle serait peut-être difficile à mettre en oeuvre dans la pratique. Celle-ci est fondée sur des études de simulation visant à évaluer les contrôles de gestion qui réalisent de solides performances dans un grand éventail d'incertitudes clés afférentes à la pêche. Des travaux sont actuellement en cours dans ce domaine, et les progrès seront présentés à la prochaine réunion du présent Groupe de travail.

4.4 Recommandations

Le modèle actuel devrait être l'hypothèse nulle; pour qu'un changement se produise, le groupe de travail devrait être suffisamment convaincu que la nouvelle méthode va apporter une amélioration, et il devrait élaborer suffisamment de documentation à cet égard. Le Résumé exécutif devrait être simplifié pour éviter tout jargon, et le motif du changement devrait apparaître de manière transparente dans le rapport détaillé. On peut avoir recours à des simulations de formule de gestion pour aider à établir l'ordre de priorité de la recherche en vue de fournir des avis de gestion.

5 Autres questions

5.1 Examen des documents

Le Groupe de travail a examiné plusieurs documents qui n'abordaient pas directement le mandat original, mais qui étaient néanmoins pertinents au thème traité par le Groupe de travail sur les Méthodes d'évaluation. Un résumé des observations du groupe figure ci-dessous.

Gaertner et Fonteneau (SCRS/00/36).

On a constaté avec inquiétude que, souvent, on tenait davantage compte de la précision d'un estimateur donné que de son exactitude. Dans certains cas, un simple estimateur comme celui qui est présenté ici peut être plus solide en termes d'exactitude qu'un autre plus sophistiqué. Il est bien entendu important d'évaluer attentivement les postulats sous-jacents de toute méthode. À titre d'exemple, la méthode préconisée dans le présent document se fonde sur l'équilibre et en tant que telle aurait tendance à surestimer la PME dans certaines circonstances. Ce biais peut être faible pour les stocks dotés d'un

taux de croissance intrinsèque élevé (comme le listao) ou peut-être atténué par la procédure de lissage (approximation aux conditions d'équilibre) adoptée dans le document. On pourrait évaluer la magnitude potentielle du biais par des études de simulation. De façon générale, l'approche paraît prometteuse, notamment en l'absence de données d'effort. Elle pourrait également s'avérer utile en tant que règle de décision qualitative à l'intention des gestionnaires, et les études de simulation avec cet aspect à l'esprit devraient être encouragées.

Geromont et Butterworth (SCRS/00/33).

Ce document a présenté une méthode destinée à permettre au modèle de déterminer les poids relatifs affectés aux indices d'abondance utilisés pour calibrer les analyses de population virtuelles. Une approche similaire a été utilisée par la Commission internationale de la Baleine (IWC) pour estimer la variance additionnelle associée à une série temporelle unique, et en Afrique du Sud, à plusieurs séries. Le Groupe de travail est convenu que l'approche proposée pourrait dans certains cas constituer une amélioration par rapport aux méthodes de repondération itérative et de variance d'entrée (seule), et qu'il convenait de l'examiner plus avant. Toutefois, on a craint que l'approche soit imprécise, surtout lorsqu'il s'agit de séries temporelles de courte durée (le modèle sera sur-paramétrisé), et peut-être même inexacte lorsque les diverses séries temporelles indiqueront des tendances contradictoires. Il a été suggéré qu'il était souhaitable d'utiliser l'information sur la variance d'entrée, mais qu'il fallait tenir compte des biais éventuels que pouvaient contenir ces estimations de variance, éventuellement en ayant recours à la méthode de variance additionnelle qui a été suggérée.

On a également suggéré que les critères de pondération (ou d'exclusion) de chaque indice devraient en grande partie être déterminés à l'extérieur du modèle, y compris l'expertise des enquêteurs qui les ont élaborés, ainsi que leur éventuelle représentativité des tendances de l'abondance. Les analyses qui incluent des indices statistiquement contradictoires devraient faire l'objet d'un examen attentif.

Legault et Porch (SCRS/00/35).

Le Groupe de travail a déploré les changements trop fréquents intervenus dans la façon de pondérer les indices d'abondance. Pour certaines espèces, un nouveau programme de pondération a été adopté à pratiquement chaque session d'évaluation, ce qui a eu pour conséquence de mettre en cause la crédibilité du processus d'évaluation. Le Groupe de travail est convenu que des études de simulation telles que celles qui ont été présentées fourniraient des avis utiles sur le bien-fondé relatif des divers programmes de pondération. Il a été suggéré d'élargir les simulations de façon à examiner l'effet de devoir évaluer les sélectivités (par opposition à l'affectation de données correctes, comme il est fait dans le document), et le degré de corrélation entre les âges dans un indice donné qui serait nécessaire pour atténuer l'avantage des indices spécifiques à l'âge sur les indices regroupés par âge. On a également estimé qu'il serait utile d'élargir le champ des simulations afin d'inclure d'autres caractéristiques de stock et de pêche. Les simulations pourraient aussi être utilisées pour examiner le caractère approprié des tests de ratio de probabilité comme critères visant à déterminer si les variances estimées d'un modèle sont statistiquement justifiables lorsque les données sont rares ou que le modèle est mal spécifié.

Maury (SCRS/00/37).

Le Groupe de travail a estimé que ce modèle constituait un outil exploratoire utile pour examiner la tendance éventuelle des coefficients de capturabilité des diverses flottes sur une période, et l'effet des changements survenus dans une zone de pêche sur la perception de la PME. C'est dans cette lumière qu'il a été appliqué au listao, dont mention est faite dans le document; on a toutefois souligné qu'il pouvait aussi être employé dans le contexte d'une évaluation. À cette fin, la différence relativement faible entre les facteurs Bayes de distribution *a posteriori* (PBF) associés aux différentes structures de modèles a servi à démontrer que le modèle était sur-paramétrisé. On a été suggéré de limiter le nombre de paramètres à ceux qui modifient sensiblement les PBF. On s'est montré préoccupé par le fait que la distribution *a priori*, notamment en ce qui concerne les coefficients de capturabilité à déroulement

aléatoire, risquait d'influencer les résultats, et qu'il fallait donc la choisir avec prudence.

McAllister et al. (SCRS/00/34).

Le Groupe de travail est convenu que l'approche générale consistant à présenter des analyses sous forme de tableaux d'analyse de décision représentait une aide prometteuse à la fois pour les scientifiques et les gestionnaires. Celle-ci s'avérerait particulièrement utile pour mettre en lumière les priorités des recherches futures. On a souligné, d'un point de vue technique, que la distribution *a priori* du niveau de PME s'est faite à partir de données qui ont récemment soulevé de fortes critiques. Il conviendrait peut-être de reformuler le modèle afin d'utiliser une distribution *a priori* pour l'inclinaison basée sur les données de Myers (par ex. Myers *et al.* 1999). Là encore, on a constaté avec inquiétude que le choix de la distribution *a priori* pouvait fortement influencer les résultats, surtout si les données ne sont pas particulièrement informatives. On en est venu à débattre de la façon de contrôler qualitativement l'expertise. Il ne s'agit pas uniquement d'un problème propre aux méthodes bayésiennes; les paramètres fixes dans les méthodes traditionnelles se fondent aussi sur l'expertise. Les critères de sélection de la distribution *a priori* devraient être transparents. On a souligné que dans des applications commerciales, on aborde l'expertise par un processus de prise de décision à critères multiples. De tels protocoles pourraient éviter que l'on passe simplement du choix du cas de base à celui de la distribution *a priori*.

5.2 Suggestions des participants pour l'avenir du Groupe de travail sur les Méthodes d'évaluation

Cette première réunion s'est concentrée principalement sur les questions de contrôle de la qualité et sur la transférabilité de la méthodologie entre les groupes de travail sur les espèces. Même si ces questions ne vont pas à l'avenir susciter une telle attention, il n'en reste pas moins que les prochaines réunions du Groupe de travail sur les Méthodes devront se pencher sur des questions spécifiques si le Groupe de travail veut avoir un impact sur la politique du SCRS. Le mandat de chaque réunion spécifique devrait aborder des questions prioritaires identifiées par le SCRS, bien que 20% de la réunion pourrait être consacré à d'autres questions méthodologiques intéressant les participants. Il faut apporter des avis concrets, et ne pas se contenter de présenter un compte rendu des débats.

5.3 Collaboration éventuelle avec d'autres organismes

Lorsque des questions méthodologiques sont d'un intérêt commun, il pourrait être avantageux de collaborer avec d'autres organisations internationales. Par exemple, l'IOTC a contacté à titre informel plusieurs autres organismes de pêche internationaux, notamment le CCSBT, l'IATTC, le SPC et l'ICCAT pour s'enquérir des possibilités de collaboration à cet égard. Il n'est pas nécessaire ni souhaitable de se réunir avec d'autres groupes de travail si les thèmes abordés sont principalement spécifiques à l'ICCAT. En outre, des considérations d'ordre logistique, comme la date de la réunion, peuvent parfois poser des difficultés. Néanmoins, cette collaboration pourrait s'avérer particulièrement utile lorsque certains groupes de travail ont des membres dotés d'une expertise distincte qu'ils ne partagent avec personne, ou bien lorsque les délais impartis ne permettent pas à un groupe de travail d'explorer suffisamment une question à lui tout seul. Le Groupe de travail a recommandé que l'ICCAT poursuive ce type de collaboration. Pour ces réunions, le SCRS devrait consulter les autres organismes afin d'élaborer un ordre du jour.

6 Recommandations

Point 2 de l'ordre du jour - Gestion de la qualité

- 1 Toutes les méthodes utilisées aux fins de l'évaluation devraient être documentées, y compris le contrôle de la qualité à tous les niveaux de contribution aux résultats.
- 2 Un Comité de coordination devrait être créé afin de contrôler la mise au point des

applications; celui-ci devrait être de petite taille et fonctionner principalement par correspondance entre les sessions.

- 3 Pour être acceptée pendant la réunion d'un Groupe de travail, toute nouvelle proposition de méthodologie doit clairement démontrer qu'il s'agit d'une véritable amélioration par rapport aux méthodologies traditionnellement utilisées. Elle doit, en outre, faire l'objet d'un contrôle attentif. Les propositions de nouvelles méthodes doivent veiller à ce que les applications remplissent les critères de contrôle de la qualité le plus rapidement possible.
- 4 Le contrôle de la qualité devrait être examiné de façon générale et inclure la collecte des données dans toutes ses composantes (prise et effort, mais aussi taille, sexe et composition spécifique), ainsi que le traitement des données (par ex. prise par taille et par âge). Cette tâche devrait incomber au Sous-Comité des Statistiques. Les perceptions de la qualité des données qui en résulteraient devraient être incluses dans les modèles d'évaluation.

Point 3 de l'ordre du jour - Standardisations de la CPUE

- 1 Des options à l'étude, celle qui consistait à utiliser le ratio de la prise des espèces d'intérêt à la prise totale comme une variable pour définir le ciblage a donné de meilleurs résultats que les autres options. Dans certains cas, toutefois, même cette méthode pourrait être trompeuse. La solution à long terme semble être la collecte de données des caractéristiques d'effort détaillées. Il conviendrait d'encourager la prise en compte de la distribution spatiale des ressources et ses dynamiques, sans oublier toute l'information environnementale disponible. Il faudrait aussi élaborer de nouveaux types de modèles non linéaires, comme le GAM.

Point 4 de l'ordre du jour – Rétablissement des stocks et méthodes d'évaluation

- 1 Il est recommandé que les méthodes d'évaluation utilisées pour suivre le rétablissement des stocks restent cohérentes pendant ce processus; le modèle actuel devrait être l'hypothèse nulle et le groupe de travail devrait être suffisamment convaincu que toute nouvelle méthode apportera une amélioration pour qu'un changement survienne. Si un changement devait intervenir, les résultats du nouveau modèle devraient également être examinés par rapport à l'ancien modèle.

Point 5 de l'ordre du jour - Autres questions

- 1 Les prochaines réunions du Groupe de travail ICCAT sur les Méthodes d'évaluation devraient se concentrer sur des questions spécifiques si le Groupe de travail veut avoir un impact sur la politique du SCRS. Le mandat de chaque réunion spécifique devrait aborder des questions prioritaires identifiées par le SCRS.
- 2 Afin d'élargir le forum international sur les questions relatives à l'évaluation des stocks et thèmes connexes, le Groupe de travail recommande que le SCRS se mette en contact avec d'autres organisations qui fournissent des avis scientifiques aux pêcheries (par exemple IOTC, IATTC, ICES) dans l'optique de formuler un mandat d'intérêt commun pour un Groupe de travail méthodologique de composition plus large.

7 Adoption du rapport et clôture

Le rapport a été adopté pendant la réunion. Étant donné que les travaux de ce Groupe de travail intéressent le Comité scientifique de façon générale, il est important que le rapport soit traduit dans toutes les langues officielles, et ce bien avant la tenue du SCRS.

À la clôture de la réunion, le Président a remercié tous les participants de leur travail considérable et de leurs précieuses contributions.

Les débats de la réunion ont été levés.

8 Bibliographie et documents de référence

8.1 Documents élaborés pour la réunion

GAERTNER, D., and A. Fonteneau. Approximate estimate of the MSY from catch data without effort information. Application to tuna fisheries. **SCRS/00/36**. Ce volume.

GEROMONT, H.F., and D.S. Butterworth. Possible extensions to the ADAPT VPA model applied to western North Atlantic bluefin tuna, addressing in particular the need to account for “additional variance”. **SCRS/00/33**. Ce volume.

LASSEN, H., and H. Sparholt. ICES quality assurance policy for fish stock assessment and management advice. **SCRS/00/38**. Ce volume.

LEGAULT, C.M., and C.E. Porch. Comparisons of weighting schemes for tuned virtual population analyses. **SCRS/00/35**. Ce volume.

MAURY, O. Multi-fleet non-equilibrium production models including stock surface to estimate catchability trends and fishery dynamics in a Bayesian context. Application to the skipjack tuna fishery (*Katsuwonus pelamis*) in the Atlantic Ocean. **SCRS/00/37**. Ce volume.

MCALLISTER, M., E.A. Babcock, and E.K. Pikitch. Using Bayesian methods and decision analysis as a rational basis for dealing with conflicting stock assessment results while providing management advice on stock rebuilding. **SCRS/00/34**. Ce volume.

MEDLEY, P.A.H. Integrating CPUE standardization within stock assessment. **SCRS/00/39**. Ce volume.

RESTREPO, V.R. Proposed stock assessment quality control procedures for ICCAT. **SCRS/00/40**. Ce volume.

8.2 Documentation de référence distribuée à la réunion

ANON. Report of the bluefin tuna methodology session (Excerpts). *ICCAT Coll. Vol. Sci. Pap.* Vol. 46-1, pp 187-268.

ANON. Detailed swordfish 1999 assessment report (Excerpts). En cours d'élaboration.

BROWN, C.A., and C.E. Porch. A numerical evaluation of lognormal, delta-lognormal and Poisson models for standardizing indices of abundance from West Atlantic bluefin catch per unit effort data (preliminary results). *ICCAT Coll. Vol. Sci. Pap.* Vol 46-4 . pp 476-482.

BUTTERWORTH, D.S., J.N. Ianelli, and R. Hilborn. In prep. A statistical time-series model for stock assessment of southern bluefin tuna. MARAM, University of Cape Town, Rondebosch 7701, South Africa.

BUTTERWORTH, D.S., and H.F. Geromont. In prep. Simulation testing as an approach to evaluate the reliability of assessment methods: An example involving initial consideration of the one/two stock

hypothesis for North Atlantic bluefin tuna. MARAM, University of Cape Town, Rondebosch 7701, South Africa.

COOKE, J.G. A procedure for using catch-effort indices in bluefin tuna assessments (revised). *ICCAT Coll. Vol. Sci. Pap.* Vol. 46-2. pp 228-232.

COOKE, J.G., and K. Lankester. Consideration of statistical models for catch-effort indices for use in tuning VPAs. *ICCAT Coll. Vol. Sci. Pap.* Vol. 45-2. pp. 125-131

GLAZER, J. 2000. Some thoughts on the by-catch issue for West Coast hake. Marine and Coastal Management, Private Bag X2, Rogge Bay, 8012, South Africa.

GLAZER, J., and D.S. Butterworth. In prep. GLM-based standardization of the South African West Coast hake catch per unit effort series, focusing on adjustments for targeting at other species. Marine and Coastal Management, Private Bag X2, Rogge Bay, 8012, South Africa.

MEJUTO, J., and J.M. de la Serna. Standardized catch rates by age and in biomass for the North Atlantic swordfish (*Xiphias gladius*) from the Spanish longline fleet for the period 1983-1998 and bias produced by changes in the fishing strategy. *ICCAT Coll. Vol. Sci. Pap* Vol. 51 pp. 1387-1411.

O'BRIEN, C.M., and L.T. Kell. The use of generalized linear models for the modelling of catch-effort series. I. Theory. *ICCAT Coll. Vol. Sci. Pap.* Vol. 46-4. pp. 476-482.

PALMA, C., J. Pereira, J. Mejuto, and M. Santos. Effect of simulating targeting levels on swordfish (*Xiphias gladius*) standardised CPUE estimates, caught by the Portuguese surface longline fleet in the North Atlantic. *ICCAT Coll. Vol. Sci. Pap* Vol 51 pp. 1718-1727.

PORCH, C.E. The implications of using the frequency of zero catches and other measures as indices of abundance. *ICCAT Coll. Vol. Sci. Pap.* Vol. 46-2. pp. 237-241

PUNT, A.E., R.W. Leslie, and A.J. Penney. A preliminary examination of the Taiwanese longline catch and effort data (1967 to 1992) for South Atlantic albacore (*Thunnus alalunga*). *ICCAT Coll. Vol. Sci. Pap.* Vol. 43 pp. 225-245.

TAKEUCHI, Y., and K. Yokawa. A note on methods to account targeting in CPUE standardization. *ICCAT Coll. Vol. Sci. Pap* Vol 51 pp. 2280-2286.

8.3 Autres textes cités

GILKS, W.R., Richardson, S. and Spiegelhalter, D.J. 1996. Markov Chain Monte Carlo in Practice. Chapman and Hall, London. 486p.

MYERS, R.A., W.A. Blanchard, and K.R. Thompson. 1990. Summary of North Atlantic fish recruitment 1942-1987. Can. Tech. Rep. Fish. Aquat. Sci. 1743. 108pp.

INFORME DEL GRUPO DE TRABAJO ICCAT SOBRE MÉTODOS DE EVALUACIÓN

Madrid, 8 a 11 de mayo de 2000

1 Apertura, adopción del Orden del día y disposiciones para la reunión

De acuerdo con las decisiones del SCRS en sus reuniones de 1998 y 1999, el Grupo de Trabajo sobre Métodos de Evaluación se reunió en la Secretaría de ICCAT los días 8 a 11 de mayo de 2000. La reunión estuvo presidida por Victor Restrepo, quien inauguró la sesión y dio la bienvenida a todos los participantes.

El objetivo primordial de este nuevo grupo de trabajo es *“Implementar una ordenación de calidad para los métodos de evaluación, con vistas a la revisión, comprobación y documentación de los métodos de evaluación empleados por el SCRS”*. Para esta primera reunión del Grupo de Trabajo, y con antelación a la misma, se definieron los siguientes objetivos que se hicieron circular de antemano:

- (a) *Establecer un protocolo para revisión de los métodos, y establecer prioridades respecto a las futuras tareas del Grupo de Trabajo.*
- (b) *Evaluar métodos para la estandarización de la CPUE que tengan en cuenta el objetivo del esfuerzo.*
- (c) *Evaluar métodos para la estandarización de la CPUE que tengan en cuenta la heterogeneidad espacial.*
- (d) *Formular asesoramiento sobre mecanismos de evaluación destinados a hacer un seguimiento de la recuperación del stock.*

Se adoptó el Orden del día provisional que se adjunta como **Apéndice 1**.

Se designaron los siguientes relatores:

Apartados 1, 7 y 8	Victor Restrepo
Apartado 2	Elisabeth Babcock y Jaime Mejuto
Apartado 3	David Die y Naozumi Miyabe
Apartados 4 y 5	Clay Porch y Chistopher Legault
Apartado 6	Renaud Pianet

Se decidió preparar también un breve Resumen Ejecutivo destinado al SCRS y a la Comisión. El presidente del SCRS, Joseph Powers, se encargaría de preparar el borrador inicial.

2 Sistemas de control de calidad

2.1 Presentación de documentos

En este punto del Orden del día se presentaron dos documentos, que fueron la base de la mayor parte de las discusiones.

El Documento SCRS/00/38 describía las normas de ICES para asegurar la calidad de la evaluación de stocks de peces y del asesoramiento respecto a ordenación. Las normas no se centran en la selección de métodos de evaluación, sino más bien en los mecanismos para determinar si un método opera en la

debida forma. Las normas fueron adoptadas en octubre de 1999, basándose en los objetivos de consenso científico, enfoques integrados (ecosistema), credibilidad de los clientes, transparencia, facultad de respuesta e independencia de toda influencia política.

Los componentes del control de calidad son: (1) un procedimiento de calidad para los grupos de evaluación; (2) puntualidad; (3) transparencia, incluyendo la habilidad para seguir el flujo de los datos por medio de las bases de datos y el flujo de supuestos y decisiones por medio de las evaluaciones, de los informes del grupo de trabajo, informes de asesoramiento y manuales de procedimiento generales de evaluación, así como guías prácticas para la evaluación que están actualmente en preparación; (4) examen interno por pares. Este examen por pares tiene lugar entre la evaluación y la presentación del asesoramiento científico, con el fin de determinar si éste es razonable y si las previsiones se basan en una evaluación también razonable. Este examen se ha integrado en un procedimiento más formal y documentado, y en pequeños subgrupos del Comité Asesor sobre Ordenación de Pesquerías (ACFM) antes de la reunión de este Comité. Las revisiones externas no están todavía implementadas, pero serán periódicas.

El documento SCRS/00/40 presentaba una propuesta de control de calidad de la evaluación de stock para ICCAT. Los niveles a los cuales se debería implementar este control de calidad son:

- (a) Recogida de datos brutos.
- (b) Agregación de (a) para su presentación a la Secretaría.
- (c) Resumen de los análisis de (a) o (b) por parte de la Secretaría para su uso en los grupos de estudio de especies.
- (d) Análisis de (a) por los científicos nacionales para su uso por los grupos de especies.
- (e) Desarrollo de métodos de evaluación.
- (f) Implementación de métodos de evaluación.
- (g) Preparación de incorporaciones a los programas de evaluación.
- (h) Selección de opciones para la ejecución de los programas de evaluación.
- (i) Validación de los resultados a la luz de los obtenidos por el programa, diagnósticos y opinión de los expertos.
- (j) Resumen y síntesis de los resultados.
- (k) Formular asesoramiento en consonancia con los resultados obtenidos.

Los puntos (a) a (c) son responsabilidad del Subcomité de Estadísticas, los puntos (j) a (k), del Grupo de Trabajo Ad hoc sobre Organización del SCRS y los puntos (d) a (i) responsabilidad del Grupo de Trabajo sobre Métodos de Evaluación.

2.2 Recomendaciones respecto a control de calidad de las evaluaciones de stock

El objetivo del protocolo de control de calidad es asegurar que las evaluaciones son valederas desde el punto de vista científico, que siguen prácticas establecidas y son transparentes y están bien documentadas, todo ello sin sacrificar la innovación de la metodología. Parte del control de calidad debería ser realizado por científicos nacionales y parte por la Secretaría de ICCAT. El protocolo incluiría normas para las entradas, diagnósticos y programación. La intención del protocolo no es limitar la flexibilidad de los grupos de trabajo de evaluación en la elección de la metodología de evaluación.

Se recomendó que se crease un catálogo de aplicaciones aprobadas por ICCAT (es decir, programas), que incluyese los elementos reseñados en el **Apéndice 3** (véase el Apartado 2.2.3). Las ventajas de los métodos no serían evaluadas, pero se estudiaría si el programa de implementación del método funcionaba en la forma deseada y si estaba debidamente documentado.

El protocolo podría ponerse en vigor, bien especificando que no podría hacerse uso de una aplicación a menos que estuviese catalogada, o bien estableciendo que los informes del grupo de trabajo concretasen si cada uno de los resultados se había o no obtenido por medio de una aplicación aprobada. El protocolo no podría entrar en vigor hasta al cabo de al menos 2 años.

2.2.1 Análisis por parte de científicos nacionales

El control de calidad de los análisis preparatorios sería sobre todo tarea de los científicos nacionales, quienes deberían documentar sus métodos y hacer diagnósticos estándar. Los puntos a considerar son, si el método aplicado era el adecuado y los posibles errores cometidos en la aplicación. Se deberían presentar nuevos métodos de análisis, con simulación de comprobaciones y documentación respecto a las diferencias existentes en relación con métodos anteriores. Los laboratorios nacionales deberían documentar sus procesos de control de calidad (no es necesario que los laboratorios nacionales establezcan estándares; el SCRS podría establecer pautas respecto a estándares mínimos).

El Grupo de Trabajo recomendó que los científicos nacionales: (1) intentasen repetir sus resultados de forma independiente; (2) que documentasen sus resultados, dando las razones por las cuales no se habían elegido alternativas, y (3) estableciesen diagnósticos estándar para los métodos aplicados. Se constató que muchos de los científicos nacionales ya siguen este procedimiento en los documentos que presentan a los grupos de trabajo de evaluación, y que también aportan sus datos con el fin de que en el curso de las evaluaciones puedan aplicarse métodos alternativos.

Se insta a científicos de diferentes países a que desarrollen proyectos conjuntos a fin de solventar los problemas comunes en el terreno del análisis de datos, tales como la estandarización de la CPUE. Si bien la seguridad respecto a calidad mejoraría si los científicos nacionales pudiesen compartir datos, con el fin de que otros pudiesen repetir sus análisis, esto implica asuntos relacionados con la confidencialidad.

2.2.2 Desarrollo de métodos de evaluación

ICCAT desea fomentar el desarrollo de nuevos métodos. Se entabló una discusión sobre si el Grupo de Trabajo debería facilitar criterios para evaluar las relativas ventajas de un método, por ejemplo, que el método apareciese en una publicación o que fuese revisado a nivel externo o interno. También se discutió la oportunidad de que existan pautas para evaluar nuevos modelos, examinando la sensibilidad de las estimaciones de puntos de referencia a los cambios estructurales en un modelo, estudiando si el modelo aporta algún “valor añadido”, etc. La conclusión fue: (1) que el programa destinado a implementar los nuevos métodos propuestos debería ser presentado para su catalogación entre 5 y 10 meses antes de la reunión del grupo de asesoramiento, con el fin de que fuese aprobado, y (2) que los autores deberían documentar las diferencias entre dicho método y el método aplicado anteriormente y demostrar en qué consiste la mejora.

Se entabló una discusión sobre si las aplicaciones en los análisis de sensibilidad deberían estar sometidas a un proceso burocrático. Se señaló que los cambios en las incorporaciones al modelo que surgen en el curso de una evaluación podrían sugerir cambios en dicho modelo que deberían usarse, al menos en los análisis de sensibilidad, si bien no podían ser aprobados por el proceso de catalogación. Sin embargo, el presentar un método que aplica un programa no comprobado, incluso en un análisis de sensibilidad, podría prestar un peso indebido a resultados no comprobados.

El Grupo acordó que los métodos deberían ser sujetos a una revisión tanto interna como externa.

El Grupo de Trabajo recomendó que las personas que propusiesen nuevos métodos estableciesen una lista de diagnósticos estándar para cada uno de los métodos.

2.2.3 Implementación de los métodos de evaluación

Se recomendó que se estableciese un catálogo de aplicaciones, que incluyese los elementos relacionados en el **Apéndice 3**. El Grupo recomendó que se crease un Comité de Coordinación encargado de supervisar este catálogo, formado por tres personas: el especialista en Dinámica de Poblaciones de la Secretaría y dos científicos nacionales nombrados por el presidente del SCRS. Al presentar una aplicación concreta al Comité de Coordinación, éste determinaría si sus autores habían presentado de forma adecuada todos los elementos del **Apéndice 3**, y estudiaría si la aplicación estaba bien documentada y funcionaba tal como se pretendía. Los comentarios del Comité se añadirían a la información en el **Apéndice 3** como punto 16. Las aplicaciones que fuesen aprobadas por el Comité se incorporarían entonces al catálogo de aplicaciones disponible para los grupos de trabajo en sus evaluaciones.

Los métodos incorporados al catálogo de ICCAT deberán presentar toda la información estipulada en el **Apéndice 3** y justificar la necesidad de un nuevo programa. El Comité Coordinador debería tener la suficiente flexibilidad para decidir acerca del nivel de detalle necesario para cada apartado del **Apéndice 3**. Se sugirió que deberían existir algunos conjuntos de datos ICCAT estándar para comprobar si un nuevo código obtiene los mismos resultados que el código establecido, cuando se aplica a la misma cuestión. Los datos de prueba no deberían usarse para evaluar las ventajas de un método, ya que métodos útiles en sí podrían funcionar mal con ciertos datos simulados.

Se discutió sobre si los autores podrían enviar las nuevas aplicaciones al Comité de Coordinación, o bien si las aplicaciones tendrían que ser designadas por los presidentes de los grupos de trabajo, por los jefes de las delegaciones de los distintos países o bien por el presidente del SCRS o alguna otra autoridad. La aprobación de las aplicaciones tomaría algún tiempo (- 8 meses). Los métodos deberían estar en escala de prioridades.

Al Comité de Coordinación se le deberían facilitar nuevas versiones de los programas catalogados, junto con el texto describiendo cuales eran los cambios en relación con versiones anteriores.

Se sugirió que una programación orientada hacia un objetivo determinado mejoraría el problema de la comprobación de programas, ya que permitiría alterar una aplicación con un mínimo de cambio de códigos. Se podrían facilitar a la Secretaría hojas de cálculo y otros programas de lenguaje de alto nivel, sin toda la información contenida en el **Apéndice 3**.

No se recomendó que los usuarios de las aplicaciones fuesen titulados, ni tampoco que recibiesen un entrenamiento especial. Sin embargo, sería conveniente que un buen número de participantes en los grupos pudiesen servirse de las aplicaciones y el catálogo les ayudaría en el aprendizaje.

Algunos de los miembros del Grupo manifestaron su preocupación por el hecho de que este protocolo fuese a dilatar el proceso de innovación. Sin embargo, si el catálogo incluye un programa “madre”, la aprobación de una variación en el método debería ser bastante rápida (unos meses). El Comité llevará a cabo sus tareas por correo electrónico durante el período que mediase entre sus reuniones.

2.2.4 Lista de comprobación de las entradas

El Grupo de Trabajo recomendó que los ficheros de entradas para cada ensayo fuesen distribuidos a los grupos de evaluación, con el fin de permitirles comprobar los errores, en la medida de lo posible y de acuerdo con el tiempo del que disponían. Los modelos deberían tener un formato de entrada estándar o bien devolver dichas entradas en un formato estandarizado, para que puedan ser comprobadas. El fichero “log” del ensayo debería concretar acerca de los ficheros de entrada, ficheros de salida y opciones seleccionadas para cada ensayo y generarse de forma automática. Los ficheros de salida deberían estar

marcados, indicando si proceden del modelo definitivo, o bien, en el informe del Grupo de Trabajo el modelo definitivo debería estar identificado con un número.

2.2.5 Comprobación de las opciones del modelo

En el curso de las evaluaciones, la persona designada por el presidente del grupo de trabajo para comprobar las entradas del modelo, debería asimismo comprobar las opciones del modelo. El fichero “log” del ensayo debería registrar dichas opciones.

2.2.6 Validación de los resultados a la luz de los resultados del programa, diagnósticos y experiencias

Respecto a las evaluaciones en sí mismas, el control de calidad incluiría:

- (1) Comprobación de las entradas (por ejemplo, ¿se suma la captura por edad a la captura total?).
- (2) Comprobación y registro de las opciones del ensayo.
- (3) Catalogación, documentación y archivo de todos los ensayos.

Las evaluaciones deberían también ser objeto de una revisión por pares, posiblemente cada 5 años.

Se insistió en que los grupos de trabajo sobre especies deben conservar el derecho a evaluar los métodos de acuerdo con su propia experiencia y determinar cuales son las entradas y supuestos adecuados.

Si en el curso de una evaluación, se considera necesario introducir pequeños cambios en una aplicación, los grupos que realizan la evaluación deberán decidir acerca de una solución intermedia entre el riesgo de errores de programación y la flexibilidad (si bien el Grupo de Trabajo podría facilitar pautas). Los presidentes de los grupos de trabajo deberán ser prudentes a la hora de cambiar códigos, ya que los errores repercuten sobre la credibilidad del proceso de evaluación. Se podrían aceptar cambios en el método en el curso de una sesión del grupo de trabajo, si éstos son de poca importancia, transparentes, ofrecen ventajas patentes y han sido evaluados con diagnósticos estándar. Es conveniente que en una reunión, dos personas hagan el cambio independientemente, para evitar errores en la codificación. Si un programa es actualizado durante una reunión, la nueva versión se facilitará a la Secretaría.

2.3 Control de calidad, desde arriba hacia abajo

Los procedimientos de control de calidad antes discutidos comprueban las evaluaciones desde abajo hacia arriba, pero se discutió también acerca de estudiar el proceso de evaluación desde arriba hacia abajo. Los nuevos métodos serían evaluados en relación con los objetivos de ordenación, después, los objetivos del método, después la adecuación de los componentes del modelo (modelo de peces, pesquería, estimación de parámetro, entradas), aspectos conceptuales del modelo (rigor estadístico, realismo desde el punto de vista biológico, forma de tratar las incógnitas) y finalmente, los aspectos prácticos (requisitos en cuestión de datos, facilidad de uso, transparencia). El Grupo de Trabajo llegó a la conclusión que si bien el asegurar la calidad es algo necesario a todos los niveles del proceso de evaluación, la elección de método queda a la discreción de cada uno de los grupos de trabajo sobre especies. El protocolo establecido aquí tan solo cubre el aspecto de asegurar la calidad del programa de aplicación.

Las sugerencias respecto a elección de método (por parte de los grupos de trabajo sobre especies) incluían los criterios de ICES (documentados, revisados por pares, comprobados por simulación). Los métodos deberían también ser claros en sus supuestos y facilitar la información necesaria para la ordenación.

2.4. Comentarios sobre otros asuntos

El Grupo de Trabajo recomendó que el Subcomité de Estadísticas del SCRS incorpore la seguridad respecto a calidad en un sentido amplio, teniendo en cuenta los requisitos de las evaluaciones. Se discutió sobre si ICCAT podría conseguir de los países unos niveles mínimos de comunicación de datos y de calidad. Se recomendó que, junto con los datos se facilitasen las estimaciones respecto a las incertidumbres, con el fin de incorporar esta incertidumbre en las evaluaciones. Asimismo se recomendó mejorar la calidad de los datos de captura por edad de la Secretaría. Se confía en que los cambios que se van a introducir en la base de datos ICCAT faciliten la comprobación de los mismos por parte del personal de Secretaría.

Se sugirió también que, con el fin de evitar el problema de que los nuevos datos conduzcan a un cambio de métodos, lo cual a su vez conduce a decisiones mal calibradas en las evaluaciones, los procesos de acopio y de evaluación de datos tuviesen lugar en reuniones aparte durante el año, o bien, que en las evaluaciones se hiciera uso de datos recopilados el año anterior. El organizar reuniones de evaluación separadas, con otros datos, resultaría caro. Sin embargo, esta cuestión se plantea al Grupo de Trabajo ad hoc sobre Organización del SCRS.

El SCRS ha discutido y sigue discutiendo acerca de la forma de implementar el examen de las evaluaciones.

3 Estandarización de los índices de CPUE

El SCRS y dos de los grupos de estudio de especies pidieron que el Grupo de Trabajo ICCAT sobre Metodología evaluase y formulase recomendaciones concretas sobre métodos de estandarización de la CPUE que tuviesen en cuenta la cuestión de objetivo del esfuerzo y la heterogeneidad espacial. Un documento concreto estaba destinado a las discusiones de la reunión en curso (SCRS/00/39) y además, otros documentos que trataban la misma cuestión aunque no habían sido preparados para esta reunión, presentaban antecedentes y fueron tenidos en cuenta por el Grupo de Trabajo.

3.1 Relación entre la evaluación de stock y la estandarización de la CPUE

En el documento SCRS/00/39, Medley observa que en general, la estandarización de la CPUE se hace independientemente de la evaluación de stock, como primer paso del análisis. La estandarización sirve para eliminar parte de la variabilidad en la CPUE que se piensa no está relacionada con las fluctuaciones en la abundancia. Desafortunadamente, parte de la variabilidad eliminada podría aportar información para la estimación de la abundancia, y se perdería en el caso de que la estandarización se hiciese por separado. Medley propone que la estandarización de la CPUE se hiciese simultáneamente con la evaluación de stock en un análisis único. Esto se puede llevar a cabo sencillamente desarrollando el modelo de CPUE de forma independiente, pero conjuntando el modelo de CPUE con el modelo de dinámica de poblaciones en un único proceso de estimación. El interés de este enfoque es que el modelo de población concuerda con el tratamiento de los datos de CPUE. Conviene observar, sin embargo, que este no es el único enfoque posible para asegurar la concordancia entre el modelo de población y la estandarización de la CPUE.

Estos modelos integrados tienden a tener componentes lineales para la estandarización del esfuerzo y componentes no lineales para el modelo de población. Hay conjuntos con los que se podrían obtener soluciones numéricas para estos modelos híbridos.

Estos modelos integrados podrían servir para tratar cualquier tema de modelación en relación con la estandarización del esfuerzo, y así, podrían aplicarse para tratar temas relacionados con el objetivo del esfuerzo y la heterogeneidad espacial. En el siguiente apartado se hace referencia concreta a tales aplicaciones. Se sugirió que se use un ejemplo práctico para evaluar el enfoque.

3.2 Sobre el objetivo del esfuerzo en las pesquerías de especies mixtas

Algunas de las pesquerías ICCAT se caracterizan por su captura simultánea de varias especies. En su origen, muchas de estas pesquerías estaban dirigidas a una sola especie, o bien tenían un objetivo concreto. En el curso de su desarrollo, las pesquerías fueron cambiando sus estrategias, cambios que incluso se modificaban cada temporada. Si bien en algunos casos raros se tiene información directa acerca de la especie-objetivo en un determinado registro de captura, con frecuencia, esta especie-objetivo tiene que deducirse de la información sobre características del arte de pesca, o bien se carece de toda información y ha de partirse de las proporciones de las diferentes especies en la captura. Este apartado presenta en primer lugar las cuestiones principales relacionadas con la especie-objetivo en las pesquerías ICCAT, después, una visión global de los métodos que intentan deducir cual es el objetivo basándose en la información sobre captura secundaria y finalmente, una discusión sobre cómo incorporar esta información en la estandarización del esfuerzo pesquero.

3.2.1 Visión global de la información sobre el objetivo del esfuerzo en las pesquerías ICCAT

a) Pesquería japonesa de palangre

Esta pesquería se inició realmente después de la Segunda Guerra Mundial y estaba dirigida en sus comienzos al rabil y al atún blanco, con destino a las conservas. Al desarrollarse el mercado de sashimi y el atún de calidad alcanzó precios altos, la pesquería empezó a dirigir su esfuerzo hacia el patudo y el atún rojo. Además de estas especies de túnidos se capturan otras especies de forma secundaria (por ejemplo, marlines, tiburones). Actualmente los lances de palangre emplean 2000-3000 anzuelos por lance y, como media, se pescan entre 20 y 30 peces por día, si bien las capturas máximas llegan a los 200 peces diarios. En estas circunstancias, no parece que haya habido una saturación del arte en los últimos años, incluso si la mayor parte de los peces pudieron ser capturados por una pequeña sección del palangre.

La información sobre el objetivo del esfuerzo puede obtenerse basándose en los anzuelos por cesta y en la profundidad del palangre. Las cestas destinadas al patudo tienen mas de 10 anzuelos, las destinadas al atún rojo tienen entre 7 y 9 anzuelos. Conviene observar que la colocación de los anzuelos puede cambiar en el curso de un viaje, porque los pescadores trasladan el esfuerzo de una especie a otra. Se producen importantes cambios en la dirección del esfuerzo según las temporadas, siendo el atún rojo la especie-objetivo sobre todo en invierno y primavera. Es bien sabido que otras especies, como el rabil y los marlines, rara vez constituyen el objetivo, si bien podrían ser un importante componente de la captura.

b) Palangre para el pez espada

Los artes han experimentado grandes cambios a lo largo del tiempo, incluso en el tipo de línea principal. Estos cambios se deben en parte a cuestiones económicas que impulsaban a los pescadores a dedicarse a la pesca de tiburones u otras especies, cambiando la época, zona y arte de sus lances de palangre. Casi todos estos cambios son pequeños y los datos disponibles no ayudan a obtener una clara estimación del objetivo del esfuerzo.

c) Pesquería de cerco para túnidos tropicales

La pesquería de cerco para túnidos tropicales se ha dirigido tradicionalmente al rabil y al listado. El patudo y los pequeños túnidos (AUX y LTA) han sido las principales especies en la captura secundaria; otras especies en esta misma categoría tienen escasa importancia (son aproximadamente el 1% de la captura total). Los datos disponibles en los cuadernos de pesca incluyen capturas de las tres principales especies de túnidos (rabil, listado y patudo) y de algunos túnidos menores. Los observadores obtienen datos sobre otras especies que son captura secundaria.

Una de las características de estas flotas ha sido la continua mejora del equipo técnico de los barcos, y se observa un aumento igualmente continuo de su potencia pesquera. Los cambios de equipo de todas las flotas se producen con gran rapidez y, como consecuencia, los diferentes niveles en la tecnología (radar de pájaros, sonar, tamaño de la malla, profundidad, etc.) rara vez se solapan en el tiempo, dificultando así la observación de su repercusión sobre la CPUE. Otro factor de importancia a tener en cuenta en la estimación de la potencia pesquera de los barcos es la propia habilidad del capitán. La mayor parte de estos datos no está disponible en la actualidad.

El incremento del uso de DCPs en la pesca, a partir de 1991, ha provocado un importante cambio en esta pesquería. Antes de su introducción, la pesca bajo troncos naturales se limitaba a una zona y época concretas, y las capturas obtenidas bajo estos troncos constituían un 15% del total de captura. Con el uso de DCPs, estas capturas se aproximan al 50%. La especie y la composición por talla difiere mucho, dependiendo del tipo de pesca: en la captura sobre cardúmenes libres, el rabil es la especie dominante, seguido del listado; por el contrario, en la pesca con DCPs, el listado representa más del 70% y el patudo y rabil pequeño, cerca del 15% respectivamente. A partir de 1990, la información sobre el tipo de pesca figura en los cuadernos de pesca. Este tipo de pesca ha introducido muchos elementos nuevos, así como problemas (añadidos a las diferencias en la composición por especies), que han de ser tenidos en cuenta al estandarizar el esfuerzo. En primer lugar, no es posible separar el esfuerzo por tipo de pesca en la información disponible en los cuadernos de pesca, en segundo lugar se producen muchas interacciones entre factores tales como temporada y zona, tipo de pesca, número de DCPs usados y el capitán (porque la pesca sobre cardúmenes libres exige más habilidad). Las condiciones oceanográficas (por ejemplo, profundidad de la termoclina, velocidad del viento, temperatura en superficie, etc.) son también factores que podría ser necesario tener en cuenta al estandarizar el esfuerzo.

En los últimos años (1997-1999), la veda temporal (tres meses) en una amplia zona en torno al ecuador, para los DCPs, ha introducido un nuevo elemento a tener en cuenta en el proceso de estandarización.

3.2.2 Deducción del objetivo del esfuerzo partiendo de la captura o tasas de captura de especies que son objetivo secundario y especies de captura secundaria

Una forma directa de estimar el objetivo del esfuerzo es usar información sobre prácticas pesqueras que pueden definir si una determinada especie constituye o no un objetivo. Por tanto se recomienda que siempre que sea posible se recoja esta información. Con frecuencia, sin embargo, esta información no se facilita junto con los registros de captura. En estos casos se puede intentar usar información sobre captura o tasas de captura de especies que son objetivo secundario y de especies de captura secundaria, para obtener índices con los cuales se puede asignar el esfuerzo de pesca a una especie-objetivo determinada.

Respecto a los conjuntos de datos de CPUE sin suficientes medidas directas de características del arte y otros factores, en las recientes evaluaciones de stock de ICCAT se han usado aproximaciones al objetivo del esfuerzo basadas en capturas del objetivo original (que denominaremos especie-objetivo) y especies afines. Pero se ha expresado preocupación respecto a la precisión de estas aproximaciones, sobre si señalan correctamente los cambios de objetivo del esfuerzo y sobre la posible representación errónea de las tendencias en la abundancia cuando las aproximaciones se aplican para obtener series temporales de la abundancia relativa (véase SCRS/99/125). Por ejemplo, una aproximación al objetivo del esfuerzo que esté sesgada, podría provocar que un auténtico descenso en la abundancia fuese considerado por error como un cambio en la dirección del esfuerzo, cuando en realidad tal cambio no ha tenido lugar. En 1999, el Grupo de Trabajo sobre Pez Espada (Informe SWO 1999), señaló varios métodos para crear tales índices y los evaluó por simulación. Las simulaciones realizadas durante dicha reunión revelaron que uno de los índices era más sólido que los otros con diferentes escenarios de cambios en la dirección del esfuerzo. Sin embargo, estas simulaciones se centraban sobre todo en el caso del pez espada y en consecuencia, no tenían en cuenta toda la gama de posibles trayectorias del esfuerzo y de la abundancia de la captura secundaria que tienen lugar en las pesquerías ICCAT. Tampoco

tomaban estas simulaciones en consideración todos los posibles escenarios de sesgo debidos a cambios históricos en la dirección del esfuerzo, ni la comunicación incompleta de la captura secundaria. Por ello, en la reunión en curso se decidió llevar a cabo nuevas simulaciones con el fin de poder aplicar los resultados a una gama más amplia de stocks de túnidos.

Se llevaron a cabo simulaciones, con el fin de determinar si varios indicadores potenciales del objetivo del esfuerzo se verían sesgados por cambios sistemáticos en la biomasa de especies-objetivo, biomasa de especies que no son objetivo, el esfuerzo total, la información sobre especies que no son objetivo, y varias combinaciones de todo ello. Estas simulaciones dieron trayectorias deterministas de población de la biomasa hipotética de especies-objetivo y no objetivo, con diferentes tendencias en la dirección del esfuerzo de una pesquería, y se calcularon cinco indicadores diferentes del objetivo del esfuerzo en toda la serie temporal (descrita más adelante). Estos casos identificados para la evaluación se basaban en algunas pesquerías ICCAT, en las cuales se sospecha que han tenido lugar cambios sistemáticos en la dirección del esfuerzo:

- 1) Pez espada (especie-objetivo original) y tintorera (especie-objetivo secundaria) (véase el SCRS/99/125).
- 2) Pez espada (especie-objetivo original), tintorera (especie-objetivo secundaria) y rabil (especie no objetivo).
- 3) Listado (especie-objetivo original) y rabil (especie-objetivo secundaria).
- 4) Listado (especie-objetivo original) y rabil (especie-objetivo secundaria), pero donde la abundancia de la especie secundaria es muy baja en relación con la especie-objetivo original).
- 5) Listado (especie-objetivo original), rabil (especie-objetivo secundaria) y patudo (especie no objetivo).

La especie-objetivo original indica la especie que se considera la única objeto de pesca selectiva durante los primeros años de una pesquería. La especie-objetivo secundaria se refiere, como su nombre indica, a una especie que se obtiene como captura secundaria durante los primeros años de una pesquería, y que posteriormente se convierte en especie objetivo, al menos por parte de una sección de la flota pesquera. La especie no objeto de pesca selectiva hace referencia a una especie a la cual no se dirige la captura, pero que se obtiene como captura secundaria.

Con el fin de desarrollar trayectorias de población realistas para los stocks de peces, se utilizó un modelo de producción de Schaefer, con valores de r y K que se aproximan a los de las especies simuladas. Los valores de los parámetros del modelo de Schaefer que se asumen para estas diferentes especies se relacionan en la **Tabla 1**.

Las series de esfuerzo anual total se especificaron como un incremento constante o lineal. El índice verdadero de captura selectiva viene dado por:

$$targ = \frac{Effort_o}{Effort_o + Effort_s}$$

donde $Effort_o$ = al esfuerzo dirigido sobre la especie-objetivo original, y $Effort_s$ = esfuerzo dirigido sobre la especie-objetivo secundaria. En las simulaciones, los valores de $Effort_o$ y de $Effort_s$ se obtuvieron multiplicando el esfuerzo total especificado en cada año por $targ$ y $(1-targ)$, respectivamente.

Para cada especie se especificaron dos tipos diferentes de valores de q : uno para aquella situación en la cual se perseguía la especie-objetivo original ($q_{\text{especie, o}}$). El segundo para cuando se perseguía la especie-objetivo secundaria ($q_{\text{especie, s}}$). Este último se estableció a un valor mucho más bajo que el primero. Por lo tanto, la captura total anual de cada especie era:

$$C_{\text{especie}} = (q_{\text{especie, o}} * \text{Effort}_o + q_{\text{especie, s}} * \text{Effort}_s) * B_{\text{especie}}$$

Los índices de selectividad de la especie-objetivo original que se tuvieron en cuenta eran:

Índice I Captura de la especie-objetivo original/captura total de las otras especies cuando otras especies se refiere a las especies-objetivo no originales.

Índice II Captura de especies-objetivo originales/captura total.

Índice III 1/tasa de captura no corregida de otra = esfuerzo total/captura de otra especie.

Índice IV 1/captura de otra especie.

Índice V Si la captura de la especie-objetivo original es superior o no a la captura de otra especie.

Cada uno de estos índices de captura dirigida fue calculado para un período de 15 años para cada uno de los escenarios de las cinco especies que se definen más arriba, y se representaron en gráficos con el índice verdadero de captura dirigida a la especie-objetivo original. En cada uno de los cinco escenarios de especies se simulaban tres situaciones diferentes de captura dirigida. En el primero no se produjeron cambios en la captura dirigida ni en la información. En el segundo, la captura dirigida descendía linealmente desde el 100% en la especie-objetivo original al 20% en un período de cinco años. En el tercero, este descenso también se producía conjuntamente con un incremento simultáneo en la información de la especie-objetivo secundaria, del 20% al 100%.

Para cada especie, se simuló un determinado número de diferentes escenarios de trayectorias de la biomasa mediante la manipulación de los valores de q , la biomasa relativa a K para cada especie, y si el esfuerzo era o no constante o creciente. En cada caso, se hizo un gráfico de los índices de la captura selectiva, una vez realizada su normalización al dividir cada valor por la media de las series.

También se calcularon las series de CPUE de la especie-objetivo original, que se estimarían aplicando cada índice de pesca selectiva como captura de la especie-objetivo dividida por las veces que el índice total normalizado se encuentre entre cero y uno.

Además de realizar un gráfico de las series de captura dirigida y de las series resultantes de CPUE, se tabularon los resultados de los diversos escenarios en términos de la suma cuadrática de los residuales entre cada derivada y los verdaderos valores relativos de abundancia de cada serie.

La **Figura 1** representa los resultados de las simulaciones que emplean índices de captura dirigida de I a IV para el caso del “listado” y del “rabil” (el índice V tuvo una actuación poco satisfactoria en la mayor parte de los casos, y fue eliminado para su posterior consideración). Cada conjunto de paneles de la **Figura 1** muestra el índice estimado de captura dirigida, así como las series correspondientes de CPUE. No se muestran las cifras de los otros casos porque son similares al caso de las dos especies de “túidos”. La **Tabla 2** muestra la suma cuadrática de los residuales de las series deducidas de CPUE a partir de los diversos escenarios obtenidos para los cuatro índices restantes que se calculan a partir de los valores de captura. La **Tabla 3** muestra el rango de los cuatro índices de captura dirigida basados en la suma cuadrática de los residuales.

Si se utilizan sólo para los totales anuales, los indicadores indirectos basados en el tamaño relativo de la captura de la especie-objetivo en comparación con el de las otras, cabe esperar que produzcan resultados fuertemente sesgados, y deberían evitarse. Sin embargo, si el método se aplica conjunto por conjunto, este método podría ser más adecuado. En más de la mitad de los escenarios, el mejor comportamiento correspondió al índice por el cual se dividía la captura de la especie-objetivo original por la captura total. No obstante, también actuó muy insatisfactoriamente en relación a los otros en un tercio de los escenarios. Además, en muchos de los escenarios, todos los índices, incluso el que presentaba una mejor actuación, presentaba una actuación muy pobre y produjeron tendencias espurias en la abundancia.

Los resultados de esta evaluación sugieren que no hay un método claramente “superior” que pueda aplicarse en general, y que cualquier indicador indirecto, incluso si se comporta mejor en relación a los otros, aún puede producir resultados muy sesgados. Por lo tanto, cuando sea de interés tratar de corregir la captura dirigida cuando se dispone sólo de la captura total de cada especie y del esfuerzo total, se deben consultar los resultados de este estudio y llevar a cabo un ejercicio de simulación similar. Sería aconsejable identificar las diversas condiciones particulares de las pesquerías concernidas. Esto incluiría las tendencias potenciales en el esfuerzo total, las abundancias relativas de las diferentes especies capturadas, sus posibles tendencias en la abundancia y si han tenido lugar o no cambios en la información. La decisión de aplicar una de las alternativas deberá tomarse sólo en el caso de que una de ellas actúe aceptablemente al aplicarla para indicar tendencias potenciales en la abundancia.

Esta evaluación determinista evalúa sólo el sesgo potencial en algunos indicadores indirectos alternativos en relación a la pesca selectiva. No ha evaluado la potencial precisión en cuanto a los indicadores indirectos alternativos, que es también una consideración importante, toda vez que se han identificado algunas alternativas que poseen un sesgo aceptablemente bajo. Este último tema requiere mayor atención en las simulaciones de Monte Carlo, en las cuales se simula la variabilidad aleatoria en los datos de captura y esfuerzo, en particular para identificar la magnitud que deben tener los cambios en la captura dirigida antes de que puedan ser detectados con fiabilidad.

Asimismo, se empleó el modelo de Schaefer para producir simulaciones biológicamente realistas en las tendencias de la biomasa. No obstante, la actuación de los indicadores indirectos de captura dirigida depende de las tendencias en la biomasa, no en el modelo de dinámica de población, de forma que podría ser útil para llevar a cabo simulaciones como ésta, con tendencias lineales simples en la biomasa de cada especie.

La falta de coherencia en la actuación de los indicadores indirectos alternativos a través de diferentes escenarios, y el amplio sesgo que frecuentemente aparece en los resultados, sugieren que también es deseable trabajar en la identificación de otros índices más fiables de captura selectiva. Los indicadores indirectos que también podrían ser creados y evaluados incluyen a aquellos que se basan en análisis GLM de datos de tasas de captura (Glazier 2000). Es también aconsejable la recolección regular de datos adicionales de pesquería, que deberían analizarse para obtener una base empírica más sólida, con el fin de identificar cambios en la captura dirigida. Estos datos adicionales podrían incluir, por ejemplo, la medición directa de las características del arte y su despliegue (por ejemplo, profundidad del palangre y espacio entre anzuelos).

De los métodos de indicadores indirectos evaluados, el empleo del ratio de la captura de las especies-objetivo a la captura total (Índice IV, **Tabla 3**), presentó la mejor media en su comportamiento, y sigue siendo el indicador directo preferido, si bien este método podría no facilitar la mejor actuación en todos los casos. Estas evaluaciones no han ensayado el posible sesgo al emplear la CPUE como índice de abundancia relativa cuando se descuida el efecto de la captura dirigida. Esta fuente de sesgo podría ser sustancial, y se le debería evaluar en el conjunto de las simulaciones que más arriba se indican, para facilitar una base que permita comparar los costes y los beneficios de aplicar indicadores indirectos a la captura dirigida en relación a despreciar los efectos de la misma.

3.2.3 Incorporación de información sobre captura dirigida durante la estandarización de la CPUE

Glazer y Butterworth (*In prep.*) y Glazer (2000) tomaron en cuenta un modelo que incorporaba la captura dirigida cuando la captura de la especie secundaria es pequeña en comparación con la de la especie-objetivo y la abundancia de la especie secundaria carece de tendencia. Este trabajo mostró que cuando hay una correlación positiva entre las tasas de captura de la especie-objetivo y las de la captura secundaria probablemente no resulta adecuado emplear información de la captura secundaria como covarianza. Por otra parte, cuando existe una correlación negativa, podría ser apropiado incorporar la captura secundaria como un factor en la estandarización de la CPUE. Desafortunadamente, muchos de los temas relativos a la captura dirigida a los túnidos se refieren a mezcla de especies donde la abundancia de la especie secundaria no es muy inferior a la de la especie principal, de forma que los modelos anteriormente citados no resultan idóneos. En tales casos, una opción podría ser llevar a cabo la estandarización de la CPUE como un análisis de variación múltiple, donde la tasa de captura de la especie-objetivo y de la especie secundaria sean variables dependientes. Estos análisis se pueden llevar a cabo empleando las librerías especiales SPLUS. También puede ayudar a cumplir esta tarea el Análisis de Componente Principal (Principal Component Analysis), pero se debe tener cautela a la hora de elegir las variables independientes. Lo mejor sería si estas variables estuvieran estrechamente relacionadas con el proceso que controla los cambios en la CPUE, por ejemplo, la localización de un frente, en vez de con cuadrículas arbitrarias basadas en la latitud y la longitud. Si bien en términos conceptuales esto es lo deseable, con frecuencia no disponemos de indicadores indirectos mejores que la localización, y cuando tenemos mejores indicadores indirectos frecuentemente no se extienden en el tiempo para que los análisis sean de mucha mayor aplicación para la reconstrucción de los datos de las series históricas de CPUE.

Los datos de lance por lance donde las CPUE de diversas especies se obtienen de forma simultánea, hacen posible separar el efecto de la captura selectiva de las especies (o sustituciones entre especies) de la variabilidad a corto plazo en la tasa de captura global, por ejemplo, debido a factores medioambientales no observados (o no observables). El **Apéndice 4** presenta en el marco de un modelo multiplicativo cómo se podría llevar esto a cabo a partir de modelizaciones explícitas del proceso de captura selectiva y en parámetros medioambientales a corto plazo que incluyan globalmente lances buenos o negativos. Es necesario trabajar más antes de que sea posible evaluar la utilidad del enfoque.

Cuando la razón principal para incorporar la captura selectiva en la estandarización de la CPUE emana de la necesidad de tratar el problema de la presencia de registros con captura cero de una de las especies, se puede utilizar un modelo delta lognormal para incluir explícitamente tales registros, y no es necesario incorporar otras variables de captura selectiva.

La captura dirigida también puede tratarse como un problema de falta de datos. En general, los modelos de captura dirigida se consideran sencillos cuando la especie-objetivo se encuentra explícitamente disponible como una variable de los datos. Por ejemplo, en el contexto de una marea, un modelo de CPUE podría ser:

$$m_i = (w e^{\hat{\alpha}_1} + (1-w) e^{\hat{\alpha}_2})$$

donde m es la CPUE prevista en el GLM de las especies-objetivo nº1, β_1 los parámetros lineales (año, lugar, barco y otros efectos), β_2 para las especies-objetivo nº2, y w la proporción de esfuerzo dirigido a la primera especie. El problema consiste en estimar los parámetros donde los datos de captura dirigida (w) no están registrados. Cuando un número significativo de barcos tienen datos de observadores donde se registran las especies-objetivo, un buen método para cumplimentar los datos que faltan podría ser el algoritmo de EM. El algoritmo de EM facilita parámetros de MLE, y debería incluir la captura selectiva en el modelo final. El método atribuye valores previstos para la variable de captura dirigida, que es la proporción de las unidades de esfuerzo aplicadas a la pesca selectiva de las especies. Sin embargo, el método general también necesitaría tratar sobre:

- Modelos más complejos, incluyendo diferentes efectos de capturabilidad y de medio ambiente.
- Donde no existan datos de captura dirigida, sólo la composición por especies en la captura.

La primera cuestión es una ampliación del enfoque, y si bien podría presentar un análisis numérico más complejo, el enfoque EM debería en teoría tener la capacidad suficiente para realizarlo. El segundo tema es mucho más serio ya que, correlacionándolo podría no ser posible resolver el problema sin crear otros supuestos significativos dentro del modelo. La recomendación es explorar si EM, EM estocástico (Gilks, Richardson y Spiegelhalter 1996) u otro método relativo a la falta de datos podría ser un medio útil para enfocar el problema de la captura selectiva.

Con independencia del modelo de captura dirigida que se emplee, es importante recordar que no todas las series de datos de captura por unidad de esfuerzo pueden constituir un buen índice de abundancia. Los desembarques de especies secundarias por unidad de esfuerzo son frecuentemente un reflejo de las preferencias del mercado y no de la abundancia. En las pesquerías multiespecíficas, los límites de captura impuestos a las especies específicas pueden producir cambios en las operaciones de la pesquería, modificar las prácticas de la captura dirigida y, por último, afectar a la coherencia de las series de datos de CPUE.

3.3 Modelos referidos a la heterogeneidad espacial en los datos de CPUE

Se ha llevado a cabo una importante tarea sobre la estandarización de los índices de CPUE en un contexto espacial referido a las especies de túnidos, especialmente para el atún rojo del sur. Los datos de palangre de esta pesquería en Japón muestran una importante reducción del área de pesca a lo largo del tiempo. Esto podría explicarse por una contracción del área donde el stock es abundante, o por efecto de las regulaciones de pesquerías (un TAC inferior), que han conducido a los pescadores a pescar únicamente en las zonas más productivas del stock. En esta pesquería se ha estandarizado la CPUE multiplicando las estimaciones de GLM para las diferentes regiones y meses por el área de la región, sumando después las diferentes regiones y sacando un promedio de los meses.

Desafortunadamente, este análisis requiere tratar con regiones donde no existe el esfuerzo de pesca, pero comprendidas, no obstante, en los caladeros históricos. Se han ensayado dos hipótesis extremas, para asumir que la abundancia en esas regiones es cero, o que iguala a la media de todas las cuadrículas en las cuales se practica la pesca. Más recientemente, se ha ensayado una tercera hipótesis utilizando una técnica de interpolación geoestadística para estimar la densidad en las áreas donde no hay pesca en un año determinado. No se ha llegado a una resolución en el debate sobre la idoneidad de ninguna de estas técnicas de interpolación. También se propuso el empleo de información sobre el hábitat para ayudar a interpolar tales densidades. A este propósito, podrían emplearse modelos GAM o mixtos.

Las tendencias en la abundancia para el atún rojo del sur son, desafortunadamente, muy sensibles a los supuestos antes señalados, especialmente durante los 10 años más recientes. Los comentarios anteriores para el atún rojo del sur, sin embargo, podrían no ser siempre aplicables a los stocks de túnidos atlánticos.

Una vez que se estiman los factores de abundancia espacial (=densidad) a partir de tales modelos, es mejor calcular la abundancia absoluta simplemente multiplicando la densidad por el tamaño de la región y sumar todas las regiones. Cuando se realiza de este modo, no obstante, se debe tener en cuenta la escala espacial de la región. Las regiones demasiado pequeñas podrían padecer problemas de agotamiento local y por tanto podrían no generar estimaciones fiables de densidad. La cooperación entre pescadores (grupos codificados) podría reducir significativamente la varianza entre barcos por el hecho de que comparten información sobre la localización de las capturas abundantes. La investigación llevada a cabo en IATTC y en ICCAT sugiere que la cooperación incrementa la eficacia de forma tan efectiva que la CPUE podría incluso incrementarse en presencia de un aumento del esfuerzo pesquero. Esto

también podría interpretarse como una búsqueda más eficaz por barcos que pertenecen a un grupo codificado en comparación con aquellos que pescan individualmente.

Comúnmente, cuando se utilizan modelos GLM para estandarizar la CPUE para obtener una heterogeneidad espacial, hay que encarar tres temas:

- La oportunidad de emplear modelos de efectos aleatorios para reducir la complejidad paramétrica (véase O'Brien y Kell, SCRS/96/173).
- Qué interacciones se deben mantener en el modelo.
- Qué diagnósticos deben ser investigados para tener la seguridad de que el modelo es adecuado para su aplicación.

En el caso de las interacciones se recomienda utilizar ensayos de probabilidad de ratio para comprobar si añadir términos extra ayuda estadísticamente al ajuste.

3.3.1 Incorporación de la dinámica de población espacial en la estandarización de la CPUE

Si se lleva a cabo la ordenación en términos espaciales (vedas espaciales) con frecuencia se justifica la incorporación de estructura espacial al modelo de evaluación de stock. Lo mismo puede decirse cuando las características del ciclo vital tienen un componente espacial, por ejemplo, cuando las migraciones guardan relación con las etapas del ciclo vital.

Medley (SCRS/00/39) muestra que se pueden desarrollar modelos para considerar explícitamente procesos como la estructura espacial y la migración junto con un modelo que incorpore la estandarización de la CPUE. Se puede mostrar que la incorporación de estos componentes en el modelo integrado puede mejorar el ajuste a los datos simulados. Al emplear estos modelos, sin embargo, se debe prestar atención a la escala espacial en la cual se crea la estandarización de la CPUE en el modelo integrado. En principio no es necesario que la escala espacial de la estandarización de la CPUE coincida con la del modelo de dinámica de poblaciones. Siempre se requerirá un cierto nivel de agregación de los datos de captura y esfuerzo (proceso previo), y que constituye una práctica general en la estandarización del esfuerzo (por ejemplo, es probable que incluso si los datos son recogidos lance por lance, se agregarán a escalas más amplias antes de que puedan ser estandarizados). Conviene recordar que cualquier serie de CPUE estandarizada obtenida a partir de un modelo integrado dependerá del modelo de población específico que se emplee. Estas series, sin embargo, pueden aún emplearse en otros análisis.

4 Recuperación del stock y métodos de evaluación

Continuar la aplicación del método de evaluación que estaba al uso en el momento en que se adoptó un objetivo específico de regeneración constituye la base para hacer un seguimiento de esa regeneración. Este proceso debería incluir análisis retrospectivos como diagnóstico para comprobar su fiabilidad.

Las siguientes discusiones sobre este tema se centraron en tres asuntos principales: 1) cuándo se considera justificado introducir un cambio en los modelos, 2) cómo se debería presentar el cambio a los gestores, y 3) cómo podemos mejorar el modelo de gestión.

4.1 Cuándo introducir el cambio

Si ya hay en vigor un plan de regeneración, ¿sería más conveniente perseverar en el actual método de evaluación destinado al asesoramiento o cambiar a un nuevo método que se percibe como una mejora? Se expresó preocupación en cuanto a que la credibilidad del proceso de evaluación podría resentirse con

un cambio en la metodología sin presentar una justificación suficiente, especialmente si posteriormente se demuestra que no era correcto. El Grupo de trabajo de ICES prefiere guardar coherencia entre un año y otro en la medida de lo posible, si bien reconoce que algunos cambios son inevitables si sólo se deben a nuevos tipos de datos. También se discutió que la ciencia puede y debe mejorar a lo largo del tiempo. Sin embargo, debido a cambios introducidos en la composición de los miembros de los grupos de trabajo, algunas veces varía la percepción de qué modelo es “el mejor”. Además, algunos miembros sugirieron que los cambios de mayor entidad presentados en una reunión de un grupo de trabajo no se utilicen en esa evaluación, sino que se demoren hasta la próxima evaluación, para de esta forma dar a los participantes la oportunidad de examinar en profundidad ese cambio. Otra preocupación se refería a que el nuevo modelo podría cambiar la percepción de abundancia absoluta. Dado que ICCAT emplea elementos de comprobación relativos, esta sea quizá la parte menos importante del problema, aunque persiste el del establecimiento de los TAC. Según la experiencia de ICES, se ha probado que es necesaria la presentación de los resultados del modelo antiguo y los del nuevo modelo, si bien no sería lo suficientemente necesario como para convencer a los gestores de la necesidad de cambiar a un nuevo modelo. En resumen, se opinó que la hipótesis cero debería ser el *status quo* y que la evidencia debería ser de peso para que se llegue a producir un cambio, es decir, el grupo de trabajo debe quedar convencido de la justificación del cambio.

4.2 *Cómo presentar los cambios*

Cuando los gestores están familiarizados con un método que ha estado vigente durante algún tiempo, es probable que estén en disposición de escuchar el asesoramiento científico. Con frecuencia, cuando se introduce un cambio, los científicos encargados de la evaluación no han logrado convencer a los gestores de que el cambio es apropiado. Esto es particularmente evidente cuando hay una votación a favor o en contra del nuevo método. Se expresó la opinión de que, en este sentido, podría servir de ayuda dotar de una estructura agil y coherente al Resumen Ejecutivo, que incluya los elementos de comprobación de la gestión, pero no así los detalles del modelo o expresiones técnicas. Sin embargo, dado que en la evaluación están implicados numerosos elementos, los cambios introducidos en la metodología deberían presentarse documentados y justificados en el informe detallado. Las razones para introducir un cambio deberán ser absolutamente nítidas, para evitar cuestiones del tipo de “¿Qué tenía de malo el modelo del año pasado?”.

4.3 *Mejoras*

En opinión del Grupo de Trabajo, el enfoque sobre procedimiento de gestión tiene cualidades deseables, pero expresó su reservas por cuanto podría ser difícil de implementar en la práctica. El enfoque sobre procedimiento de gestión se basa en estudios de simulación para evaluar los controles de gestión que manifiestan un comportamiento sólido en un amplio rango de incertidumbres clave relativas a la pesquería. Esta tarea está actualmente en desarrollo, y los progresos se darán a conocer en la próxima reunión de este Grupo de trabajo.

4.4 *Recomendaciones*

El modelo actual debería ser el de hipótesis nula, y el Grupo de trabajo debería estar lo suficientemente convencido de que cualquier método nuevo representa una mejora que podría dar paso a la posibilidad de un cambio, y preparar la suficiente documentación a tal efecto. El Resumen Ejecutivo deberá ser fluido y evitar terminología técnica, y los motivos del cambio estarán claramente explicados en el informe detallado. El enfoque sobre procedimientos de ordenación mediante simulaciones puede utilizarse para ayudar a conceder prioridades de investigación en apoyo de la gestión.

5 Otros asuntos

5.1 Examen de los documentos

El Grupo de trabajo examinó numerosos documentos que no hacían una referencia directa a los ámbitos de competencia originales, pero que, sin embargo, resultaban pertinentes con respecto al tema del Grupo de trabajo sobre Métodos. Más adelante se presenta una sinopsis de los comentarios hechos por el Grupo.

Gaertner y Fonteneau (SCRS/00/36).

Se manifestó cierta preocupación por cuanto, con frecuencia, en la actualidad se prestaba una mayor atención a la precisión de un estimador dado que a su fidelidad. En algunos casos, un simple estimador como el que aquí se presenta podría ser más sólido en términos de fidelidad que sus otras contrapartes más sofisticadas. Por supuesto, es importante, evaluar cuidadosamente los supuestos de base de cualquier método. Por ejemplo, el método por el que se aboga en este documento se basa en el equilibrio, y como tal podría tender a sobreestimar el RMS en ciertas circunstancias. Este sesgo podría ser pequeño para los stocks con una fuerte tasa de incremento intrínseca (como ocurre en el caso del listado) o quizá suavizarse por el procedimiento de aplanamiento (aproximación al equilibrio) que se presenta en el documento. La magnitud potencial del sesgo podría evaluarse a través de estudios de simulación. Sobre todo, el enfoque parece prometedor, particularmente cuando se carece de datos de esfuerzo. También podría mostrar su utilidad como norma de decisión cualitativa para los gestores y se debe alentar la realización de estudios de simulación teniendo en cuenta esta circunstancia.

Geromont y Butterworth (SCRS/00/33).

Este documento presentaba un método que permitía al modelo determinar la importancia relativa asignada a los índices de abundancia empleados para calibrar análisis de población virtual. La CBI ha empleado un enfoque similar para estimar la varianza adicional asociada a una serie temporal única y para series múltiples en Sudáfrica. El Grupo de trabajo manifestó su acuerdo con respecto a que el enfoque propuesto podría en algunos casos representar una mejora en relación a los métodos de reponderación iterativa y de varianza de entrada (por sí sola) y que valía la pena examinarlo nuevamente. Sin embargo, se expresaron reservas en cuanto a la probabilidad de que el enfoque fuera impreciso, especialmente cuando se encuentran implicadas series temporales cortas (el modelo tendría excesivos parámetros) y quizá poco exacto cuando las diversas series temporales indiquen tendencias problemáticas. Se sugirió que sería deseable utilizar el empleo de información de varianza de entrada, pero que en ese caso habría que tener en cuenta posibles sesgos en tales estimaciones de la varianza, posiblemente utilizando el método de varianza adicional sugerido.

También se sugirió que los criterios de ponderación (o exclusión) de cada índice deberían venir ampliamente determinados, externamente al modelo, incluyendo los conocimientos de los investigadores que los desarrollaron, y su posible representatividad como tendencias en la abundancia. Se deberán estudiar con cuidado los análisis que incluyan índices estadísticamente contradictorios.

Legault y Porch (SCRS/00/35).

El Grupo de trabajo lamentó que hubiera cambios demasiado frecuentes en la forma en que son ponderados los índices de abundancia. Para algunas especies, se adopta un nuevo esquema de ponderación en prácticamente cada reunión de evaluación y, en consecuencia, la credibilidad del proceso de evaluación sale dañada. El Grupo de trabajo acordó que los estudios de simulación como los que se han presentado darían un asesoramiento útil en cuanto concierne a los merecimientos de varios esquemas de ponderación. Se sugirió ampliar las simulaciones para examinar el efecto de estimar las selectividades (en contraposición a asignar los valores correctos, como se ha hecho en el documento) y el grado de

correlación entre las edades de un índice dado, necesario para suavizar la ventaja de los índices específicos de la edad sobre los índices agregados por edades. También se opinó que sería útil ampliar el alcance de las simulaciones para incluir otras características del stock y de la pesquería. Las simulaciones podrían también utilizarse para examinar la idoneidad de los ensayos de probabilidad del ratio, para determinar si las varianzas estimadas del modelo son estadísticamente justificables cuando los datos son escasos o el modelo contiene errores en su especificación.

Maury (SCRS/00/37)

El Grupo de trabajo halló que este modelo era una herramienta exploratoria útil para examinar posibles tendencias en los coeficientes de capturabilidad de diversas flotas a lo largo del tiempo, y el efecto de los cambios de percepción del RMS en la zona de pesca. Su aplicación al listado, tema que se trata en este documento, se hizo a la luz de este entendimiento, pero, sin embargo, también se indicó que podría utilizarse en el contexto de una evaluación. A este respecto, la diferencia relativamente pequeña entre los factores posteriores de Bayes (PBF, en sus siglas en inglés) asociados a distintos modelos de estructuras, se tomó como evidencia de que el modelo tenía un exceso de parámetros. Se sugirió que el número de parámetros quedara limitado a los que implican un cambio significativo en el PBF. Se expresó la preocupación de que los antecedentes empleados, en particular con respecto a la condición aleatoria de los coeficientes de capturabilidad, podrían terminar liderando los resultados, y que se debería tener cautela a la hora de realizar su selección.

McAllister et al. (SCRS/00/34)

El Grupo de trabajo acordó que el enfoque general de presentar análisis en la forma de tablas de decisión de análisis era una ayuda prometedora para tanto los científicos como los gestores. Podría ser particularmente útil para determinar las prioridades de investigación en el futuro. Se hizo la precisión técnica de que el antecedente del nivel de RMS se había desarrollado a partir de datos de últimamente habían recibido duras críticas. Podría ser adecuado refundir de nuevo el modelo para utilizar un antecedente suavizado basado en los datos de Myers (por ejemplo, Myers *et al.* 1990). En este punto, se reiteró la preocupación de que la elección de antecedentes podría tener una gran influencia sobre los resultados, en particular cuando los datos no son especialmente informativos. Esto condujo a una discusión sobre cómo determinar el nivel de conocimientos de los expertos. Este problema no pertenece sólo a los métodos Bayesianos; los parámetros fijos en los métodos frecuentistas también se basan en conocimientos de los expertos. El criterio empleado para la elección de antecedentes debe ser transparente. Se indicó que en las aplicaciones comerciales existen sistemas para analizar el conocimiento de los expertos mediante la toma de decisiones con criterios múltiples. Tales protocolos podrían simplemente evitar que el debate de la elección del caso base se traslade al de la elección de antecedentes.

5.2 Sugerencias de los participantes para el futuro Grupo de Trabajo ICCAT sobre Métodos de Evaluación

Uno de los principales temas de esta primera reunión trataba sobre el control de calidad y la transferencia de metodología entre grupos de trabajo sobre especies. Si bien en un futuro estos mismos temas podrían no merecer tal atención, las subsiguientes reuniones del Grupo de trabajo sobre Métodos deben centrarse en temas específicos, con el objetivo de que ejerzan su influencia sobre la política del SCRS. Los términos de referencia de cada reunión específica deberían tratar temas prioritarios identificados por el SCRS, si bien quizá el 20 por ciento de la reunión podría dedicarse a otros asuntos de naturaleza metodológica de interés para los participantes. Se debe elaborar un asesoramiento concreto, y no simplemente una descripción del debate.

5.3 Posible colaboración con otros organismos

Cuando existen cuestiones metodológicas comunes, podría resultar beneficioso establecer una colaboración con otras organizaciones internacionales. El IOTC, por ejemplo, ha establecido un contacto informal con otros diversos organismos internacionales de pesquerías, incluyendo a CCSBT, IATTC, SPC e ICCAT, para investigar la posibilidad de establecer una colaboración en este sentido. Podría no ser necesario e incluso no deseable reunirse con otros grupos de trabajo en aquellas ocasiones en las cuales los temas sean más específicos de ICCAT. Además, las consideraciones de orden logístico, tales como la duración de las reuniones, podrían algunas veces causar dificultades. No obstante, esta colaboración podría resultar especialmente útil cuando algunos de los grupos de trabajo incluyan a participantes que dispongan de una experiencia particular que no haya sido compartida por los otros, o cuando el tiempo disponible impida que un grupo de trabajo estudie por sí mismo un tema con el suficiente detalle. El Grupo de Trabajo recomendó que ICCAT adopte esta forma de colaboración. Para celebrar reuniones de esa naturaleza, el SCRS debería consultar con los otros organismos, con el fin de establecer una agenda.

6 Recomendaciones

Punto 2 del Orden del día - Gestión de calidad

- 1 Todos los métodos empleados para la evaluación deberán estar debidamente documentados, incluyendo el Control de Calidad a todos los niveles de contribución a los resultados.
- 2 Se deberá crear un Comité Coordinador con el fin de controlar el desarrollo de aplicaciones; será reducido, y trabajará durante los períodos inter sesiones, principalmente por correo.
- 3 Para ser aceptada durante la reunión de un Grupo de trabajo, cualquier nueva metodología propuesta deberá demostrar claramente que es una mejora real en comparación con las que se emplean tradicionalmente, y deberá ser cuidadosamente contrastada. Las propuestas de nuevos métodos deberían asegurar a la mayor brevedad que las aplicaciones cumplen los criterios de Control de Calidad.
- 4 El Control de Calidad se considerará en un sentido amplio, e incluirá tanto la recolección de datos de todos sus componentes (captura y esfuerzo, pero también talla, sexo y composición por especies) como el proceso de datos (por ejemplo, captura por clases de tallas y edad), comprendido en el marco de responsabilidad del Subcomité de Estadísticas. Las percepciones resultantes sobre la calidad de los datos deberían estar incluidas en los modelos de evaluación.

Punto 3 del Orden del día - Estandarizaciones de la CPUE

- 1 De las opciones examinadas, el ratio de captura de las especies de interés a la captura total como variable para definir la selección tendía a actuar mejor que las otras opciones, pero, en determinadas circunstancias, incluso ese método podría resultar engañoso. La solución a largo plazo parece ser la recolección de datos de esfuerzo con características detalladas. Teniendo esto presente, se debería alentar la distribución espacial de los recursos y su dinámica, incluyendo toda la información medioambiental disponible, y desarrollando nuevos tipos de modelos no lineales tales como el GAM.

Punto 4 del Orden del día - Reconstitución de stock y métodos de evaluación

Se recomienda que los métodos de evaluación que se apliquen a llevar a cabo un seguimiento de la reconstitución de stock mantengan su coherencia durante este proceso. El modelo actual debería ser el de hipótesis nula, y para que llegue a producirse un cambio, el grupo de trabajo deberá estar lo suficientemente convencido de que cualquier nuevo método representa una mejora. Si así sucede, los resultados del nuevo modelo también deberán ser sometidos a consideración a la luz del modelo antiguo.

Punto 5 del Orden del día - Otros asuntos

- 1 Las reuniones subsiguientes del Grupo de trabajo ICCAT sobre Métodos de Evaluación se deberían centrar en temas específicos, con el fin de que los resultados de las tareas del Grupo de trabajo influyan en la política del SCRS. Los términos de referencia de cada reunión específica deberán tratar temas prioritarios identificados por el SCRS.
- 2 Con el fin de promover un desarrollo internacional más amplio en temas de evaluación de stock y asuntos relacionados, el Grupo de trabajo recomienda al SCRS que establezca contacto con otras organizaciones científicas de asesoramiento de pesquerías (como IOTC, IATTC, ICES), con el propósito de establecer ámbitos de competencia de interés común para crear un Grupo de trabajo metodológico con un mayor número de miembros.

7 Adopción del informe y Clausura

El informe fue adoptado durante la reunión. Debido al interés general del SCRS por este Grupo de Trabajo, es importante que el informe haya sido traducido a los idiomas oficiales con suficiente antelación al SCRS.

En la clausura de la reunión, el presidente dio las gracias a todos los participantes por la ardua tarea llevada a cabo y por la valiosa contribución aportada.

La reunión fue clausurada.

8 Literatura citada y documentos de base

8.1 Documentos preparados para la reunión

GAERTNER, D., and A. Fonteneau. Approximate estimate of the MSY from catch data without effort information. Application to tuna fisheries. **SCRS/00/36**. Este volumen.

GEROMONT, H.F., and D.S. Butterworth. Possible extensions to the ADAPT VPA model applied to western North Atlantic bluefin tuna, addressing in particular the need to account for “additional variance”. **SCRS/00/33** Este volumen.

LASSEN, H., and H. Sparholt. ICES quality assurance policy for fish stock assessment and management advice. **SCRS/00/38**. Este volumen.

LEGAULT, C.M., and C.E. Porch. Comparisons of weighting schemes for tuned virtual population analyses. **SCRS/00/35**. Este volumen.

MAURY, O. Multi-fleet non-equilibrium production models including stock surface to estimate catchability trends and fishery dynamics in a Bayesian context. Application to the skipjack tuna fishery (*Katsuwonus pelamis*) in the Atlantic Ocean. **SCRS/00/37**. Este volumen.

MCALLISTER, M., E.A. Babcock, and E.K. Pikitch. Using Bayesian methods and decision analysis as a rational basis for dealing with conflicting stock assessment results while providing management advice on stock rebuilding. **SCRS/00/34**. Este volumen.

MEDLEY, P.A.H. Integrating CPUE standardization within stock assessment. **SCRS/00/39**. Este volumen.

RESTREPO, V.R. Proposed stock assessment quality control procedures for ICCAT. **SCRS/00/40**. Este volumen.

8.2 Documentos de base distribuidos en la reunión

ANON. Report of the bluefin tuna methodology session (Excerpts). *ICCAT Coll. Vol. Sci. Pap.* Vol. 46-1, pp 187-268.

ANON. Detailed swordfish 1999 assessment report (Excerpts). En preparación.

BROWN, C.A., and C.E. Porch. A numerical evaluation of lognormal, delta-lognormal and Poisson models for standardizing indices of abundance from West Atlantic bluefin catch per unit effort data (preliminary results). *ICCAT Coll. Vol. Sci. Pap.* Vol 46-4 . pp 476-482.

BUTTERWORTH, D.S., J.N. Ianelli, and R. Hilborn. In prep. A statistical time-series model for stock assessment of southern bluefin tuna. MARAM, University of Cape Town, Rondebosch 7701, South Africa.

BUTTERWORTH, D.S., and H.F. Geromont. In prep. Simulation testing as an approach to evaluate the reliability of assessment methods: An example involving initial consideration of the one/two stock hypothesis for North Atlantic bluefin tuna. MARAM, University of Cape Town, Rondebosch 7701, South Africa.

COOKE, J.G. A procedure for using catch-effort indices in bluefin tuna assessments (revised). *ICCAT Coll. Vol. Sci. Pap.* Vol. 46-2. pp 228-232.

COOKE, J.G., and K. Lankester. Consideration of statistical models for catch-effort indices for use in tuning VPAs. *ICCAT Coll. Vol. Sci. Pap.* Vol. 45-2. pp. 125-131

GLAZER, J. 2000. Some thoughts on the by-catch issue for West Coast hake. Marine and Coastal Management, Private Bag X2, Rogge Bay, 8012, South Africa.

GLAZER, J., and D.S. Butterworth. In prep. GLM-based standardization of the South African West Coast hake catch per unit effort series, focusing on adjustments for targeting at other species. Marine and Coastal Management, Private Bag X2, Rogge Bay, 8012, South Africa.

MEJUTO, J., and J.M. de la Serna. Standardized catch rates by age and in biomass for the North Atlantic swordfish (*Xiphias gladius*) from the Spanish longline fleet for the period 1983-1998 and bias produced by changes in the fishing strategy. *ICCAT Coll. Vol. Sci. Pap* Vol. 51 pp. 1387-1411.

- O'BRIEN, C.M., and L.T. Kell. The use of generalized linear models for the modelling of catch-effort series. I. Theory. *ICCAT Coll. Vol. Sci. Pap.* Vol. 46-4. pp. 476-482.
- PALMA, C., J. Pereira, J. Mejuto, and M. Santos. Effect of simulating targeting levels on swordfish (*Xiphias gladius*) standardised CPUE estimates, caught by the Portuguese surface longline fleet in the North Atlantic. *ICCAT Coll. Vol. Sci. Pap* Vol 51 pp. 1718-1727.
- PORCH, C.E. The implications of using the frequency of zero catches and other measures as indices of abundance. *ICCAT Coll. Vol. Sci. Pap.* Vol. 46-2. pp. 237-241
- PUNT, A.E., R.W. Leslie, and A.J. Penney. A preliminary examination of the Taiwanese longline catch and effort data (1967 to 1992) for South Atlantic albacore (*Thunnus alalunga*). *ICCAT Coll. Vol. Sci. Pap.* Vol. 43 pp. 225-245.
- TAKEUCHI, Y., and K. Yokawa. A note on methods to account targeting in CPUE standardization. *ICCAT Coll. Vol. Sci. Pap* Vol 51 pp. 2280-2286.

8.3 Otras citas

- GILKS, W.R., Richardson, S. and Spiegelhalter, D.J. 1996. Markov Chain Monte Carlo in Practice. Chapman and Hall, London. 486p.
- MYERS, R.A., W.A. Blanchard, and K.R. Thompson. 1990. Summary of North Atlantic fish recruitment 1942-1987. Can. Tech. Rep. Fish. Aquat. Sci. 1743. 108pp.

Table 1. Values for logistic model parameters assumed for simulated fish with biological characteristics similar to swordfish, blue shark, skipjack tuna, yellowfin tuna and albacore tuna.

Species	Intrinsic rate of increase (r)	Carrying Capacity (K)
Swordfish-like species	0.435	120,000
Blue shark-like species	0.1	500,000
Skipjack tuna-like species	1.0	600,000
Yellowfin tuna-like species	0.8	600,000
Bigeye tuna-like species	0.6	566,000

Table 2. Results of the targeting simulations. Table entries for each scenario correspond to the sum of squared residuals (X 1000) in the derived CPUE series from the four indices of targeting: Index **I** = Catch Target / Catch Other; Index **II** = 1 / Catch Rate Other; Index **III** = 1 / Catch Other; Index **IV** = Catch Target / Catch Total.

Case	Target species	Other species	Effort	Target species	No change in targeting No change in reporting				More targeting of "other" No change in reporting				No change in targeting Increased reporting of "other"			
					I	II	III	IV	I	II	III	IV	I	II	III	IV
1	stable	stable	constant	common	1	3	3	1	3432	98	98	2	9295	9317	9317	180
2	stable	decreasing	increasing	common	356	396	119	226	1758	2511	2016	1390	6473	6247	10924	2399
3	stable	increasing	constant	common	4943	3830	3830	1615	3970	2103	2103	2371	20919	19869	19869	2541
4	decreasing	stable	increasing	common	181	21	748	24	2278	606	419	2250	11541	8711	14106	36
5	decreasing	increasing	increasing	common	19	252	331	18	682	2933	1654	996	9776	7106	12292	412
6	decreasing	increasing	increasing	common	5030	3340	5158	381	6371	5468	4790	1647	13074	11180	13102	731
7	increasing	stable	increasing	common	2839	74	74	767	9184	888	888	2701	900	2922	2922	314
8	increasing	decreasing	increasing	common	10863	32	265	4601	41031	41239	40993	21507	863	1318	1778	170
9	increasing	increasing	increasing	common	541	2058	2058	71	10876	667	667	4855	95	2502	2502	132
10	stable	stable	constant	rare	1	3	3	1	3432	98	98	2116	9295	9317	9317	3530
11	stable	decreasing	increasing	rare	356	396	119	339	1758	2511	2016	1710	6473	6247	10924	5714
12	stable	increasing	constant	rare	4943	3830	3830	2698	3970	2103	2103	2057	20919	19869	19869	9325
13	decreasing	stable	increasing	rare	181	21	748	29	2278	606	419	21	11541	8711	14106	714
14	decreasing	increasing	increasing	rare	19	252	331	16	682	2933	1654	531	9776	7106	12292	5076
15	decreasing	increasing	increasing	rare	5030	3340	5158	3235	6371	5468	4790	4307	13074	11180	13102	7509
16	increasing	stable	increasing	rare	2839	74	74	2415	9184	888	888	7747	900	2922	2922	760
17	increasing	decreasing	increasing	rare	10863	32	265	9820	41031	41239	40993	38001	863	1318	1778	697
18	increasing	increasing	increasing	rare	541	2058	2058	404	10876	667	667	9903	95	2502	2502	77

Table 3. Results of the targeting simulations. Table entries for each scenario correspond to the ranked sum of squared residuals in the derived CPUE series from the four indices of targeting: Index **I** = Catch Target / Catch Other; Index **II** = 1 / Catch Rate Other; Index **III** = 1 / Catch Other; Index **IV** = Catch Target / Catch Total. A lower rank is indicative of a better targeting proxy.

Case	Target species	Other species	Effort	Target species	No change in targeting No change in reporting				More targeting of "other" No change in reporting				No change in targeting Increased reporting of "other"			
					I	II	III	IV	I	II	III	IV	I	II	III	IV
1	stable	stable	constant	common	2	3	3	1	4	2	2	1	2	3	3	1
2	stable	decreasing	increasing	common	3	4	1	2	2	4	3	1	3	2	4	1
3	stable	increasing	constant	common	4	2	2	1	4	1	1	3	4	2	2	1
4	decreasing	stable	increasing	common	3	1	4	2	4	2	1	3	3	2	4	1
5	decreasing	increasing	increasing	common	2	3	4	1	1	4	3	2	3	2	4	1
6	decreasing	increasing	increasing	common	3	2	4	1	4	3	2	1	3	2	4	1
7	increasing	stable	increasing	common	4	1	1	3	4	1	1	3	2	3	3	1
8	increasing	decreasing	increasing	common	4	1	2	3	3	4	2	1	2	3	4	1
9	increasing	increasing	increasing	common	2	3	3	1	4	1	1	3	1	3	3	2
10	stable	stable	constant	rare	2	3	3	1	4	1	1	3	2	3	3	1
11	stable	decreasing	increasing	rare	3	4	1	2	2	4	3	1	3	2	4	1
12	stable	increasing	constant	rare	4	2	2	1	4	2	2	1	4	2	2	1
13	decreasing	stable	increasing	rare	3	1	4	2	4	3	2	1	3	2	4	1
14	decreasing	increasing	increasing	rare	2	3	4	1	2	4	3	1	3	2	4	1
15	decreasing	increasing	increasing	rare	3	2	4	1	4	3	2	1	3	2	4	1
16	increasing	stable	increasing	rare	4	1	1	3	4	1	1	3	2	3	3	1
17	increasing	decreasing	increasing	rare	4	1	2	3	3	4	2	1	2	3	4	1
18	increasing	increasing	increasing	rare	2	3	3	1	4	1	1	3	2	3	3	1

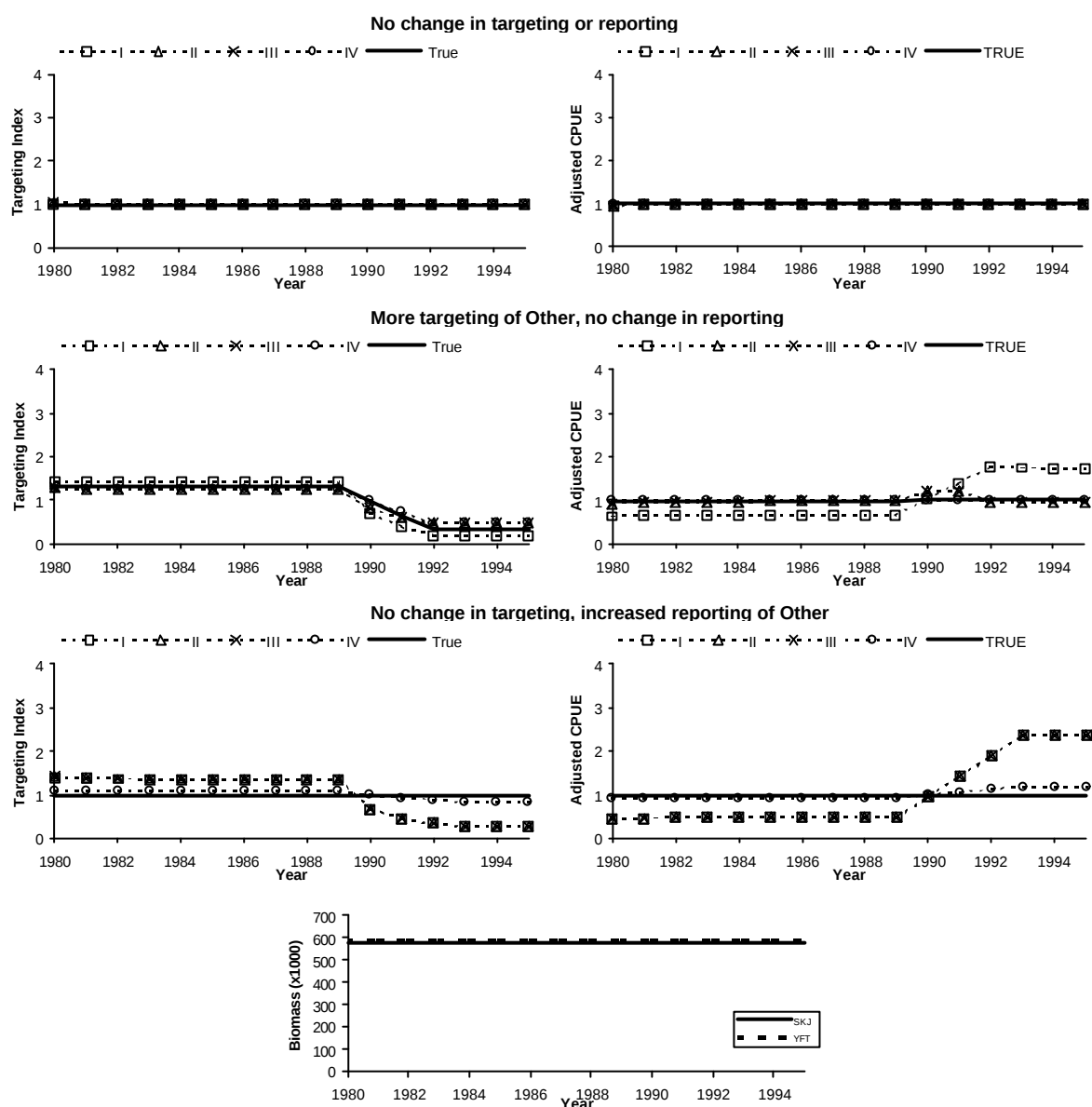


Figure 1. Sample results of the targeting simulations. The bottom panel shows the relative trends in biomass for the two hypothetical stocks. The various targeting indices (I to IV, see text) and derived CPUE indices are shown in the left- and right hand side panels, respectively, for three scenarios of changes in targeting or reporting.

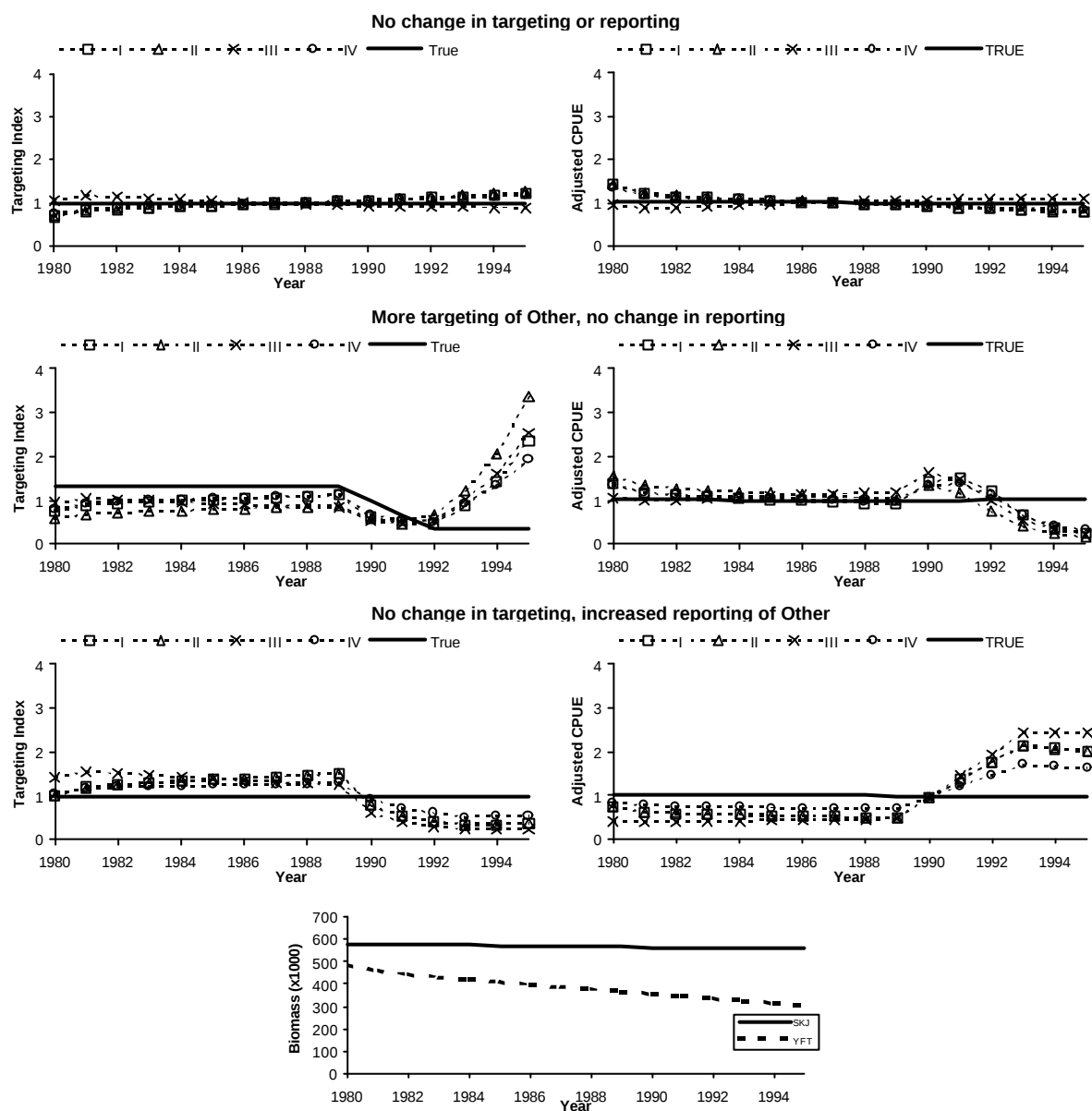


Figure 1 (cont.). Sample results of the targeting simulations. The bottom panel shows the relative trends in biomass for the two hypothetical stocks. The various targeting indices (I to IV, see text) and derived CPUE indices are shown in the left- and right hand side panels, respectively, for three scenarios of changes in targeting or reporting.

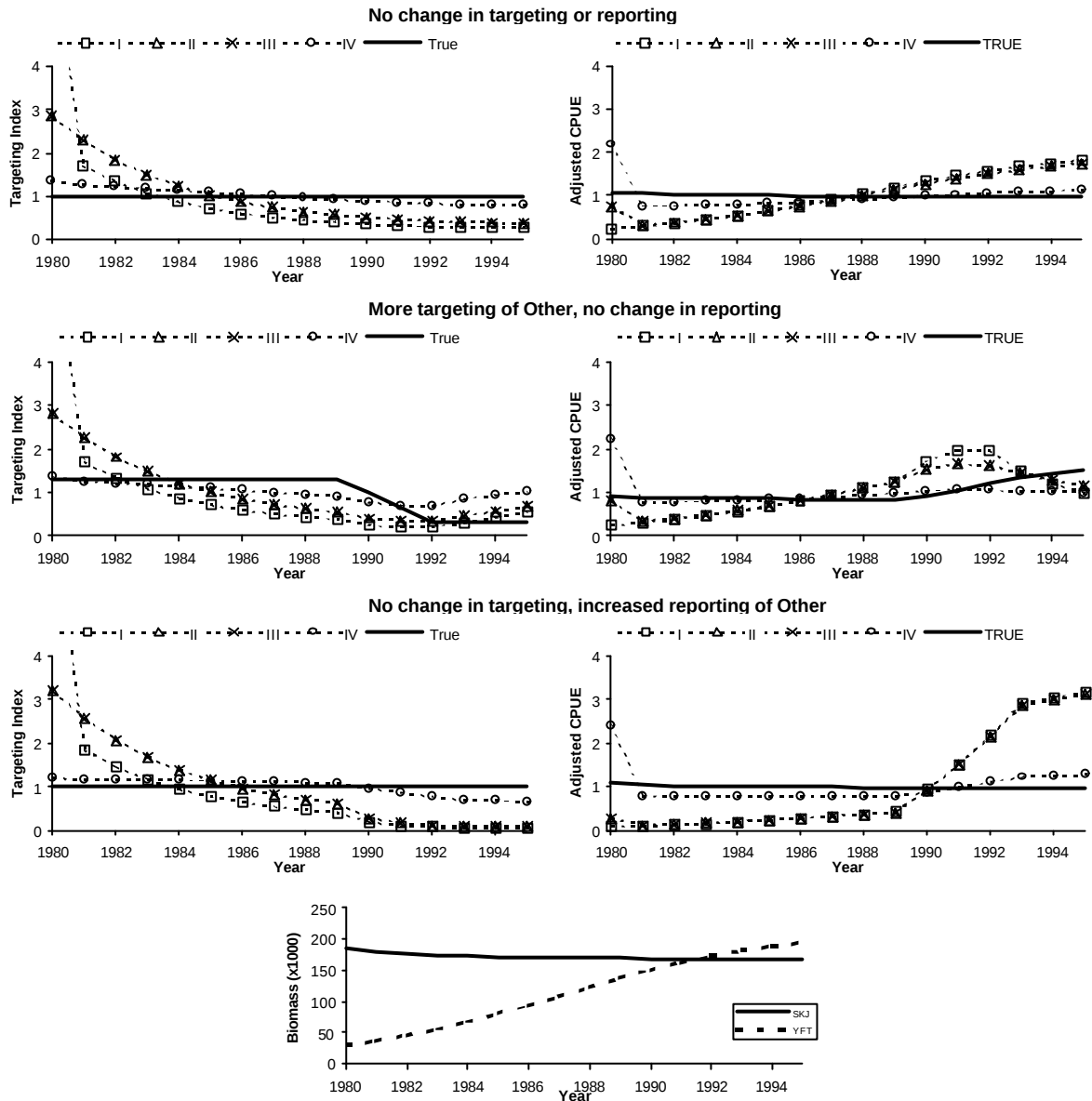


Figure 1 (cont.). Sample results of the targeting simulations. The bottom panel shows the relative trends in biomass for the two hypothetical stocks. The various targeting indices (I to IV, see text) and derived CPUE indices are shown in the left- and right hand side panels, respectively, for three scenarios of changes in targeting or reporting.

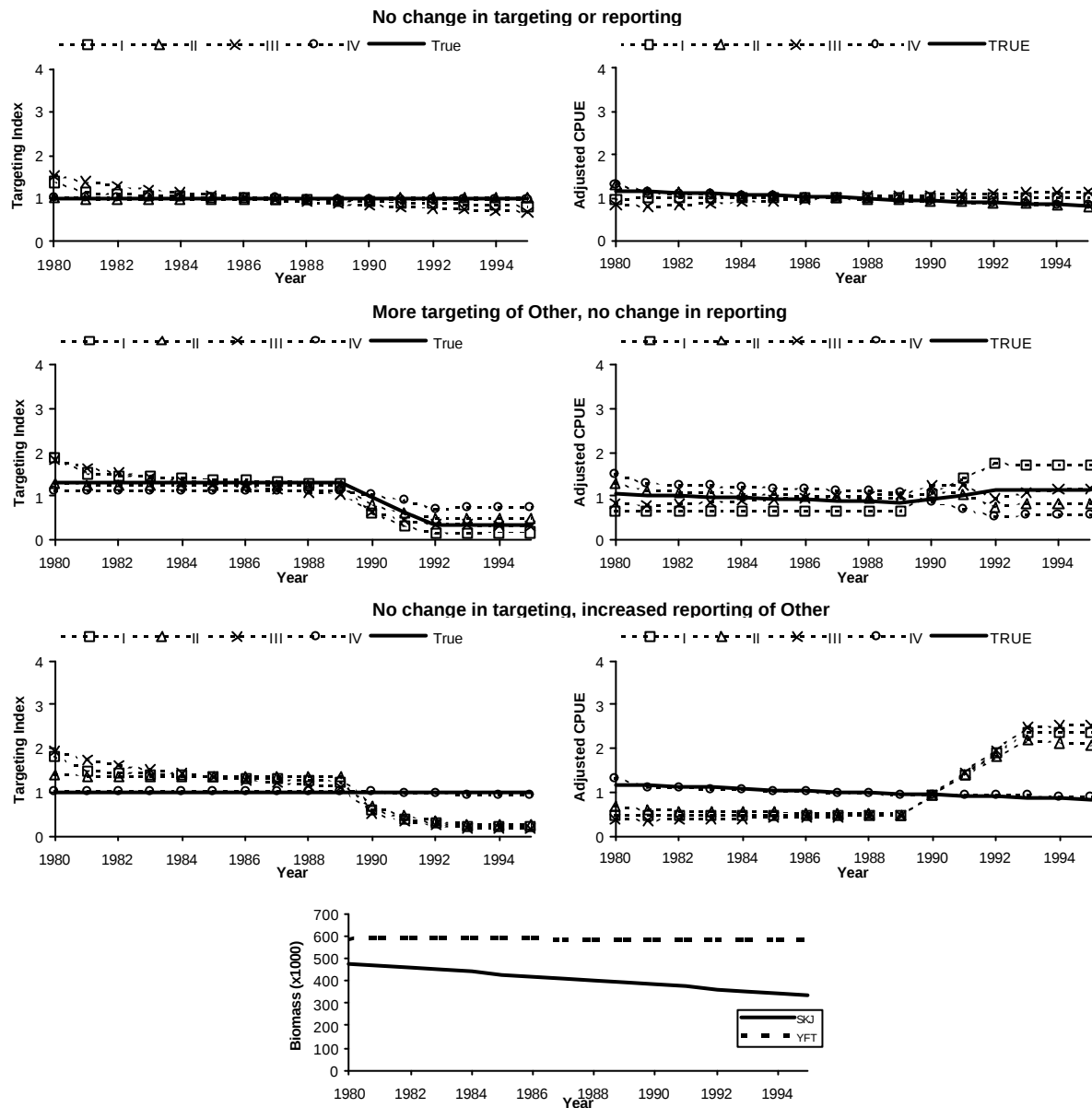


Figure 1 (cont.). Sample results of the targeting simulations. The bottom panel shows the relative trends in biomass for the two hypothetical stocks. The various targeting indices (I to IV, see text) and derived CPUE indices are shown in the left- and right hand side panels, respectively, for three scenarios of changes in targeting or reporting.

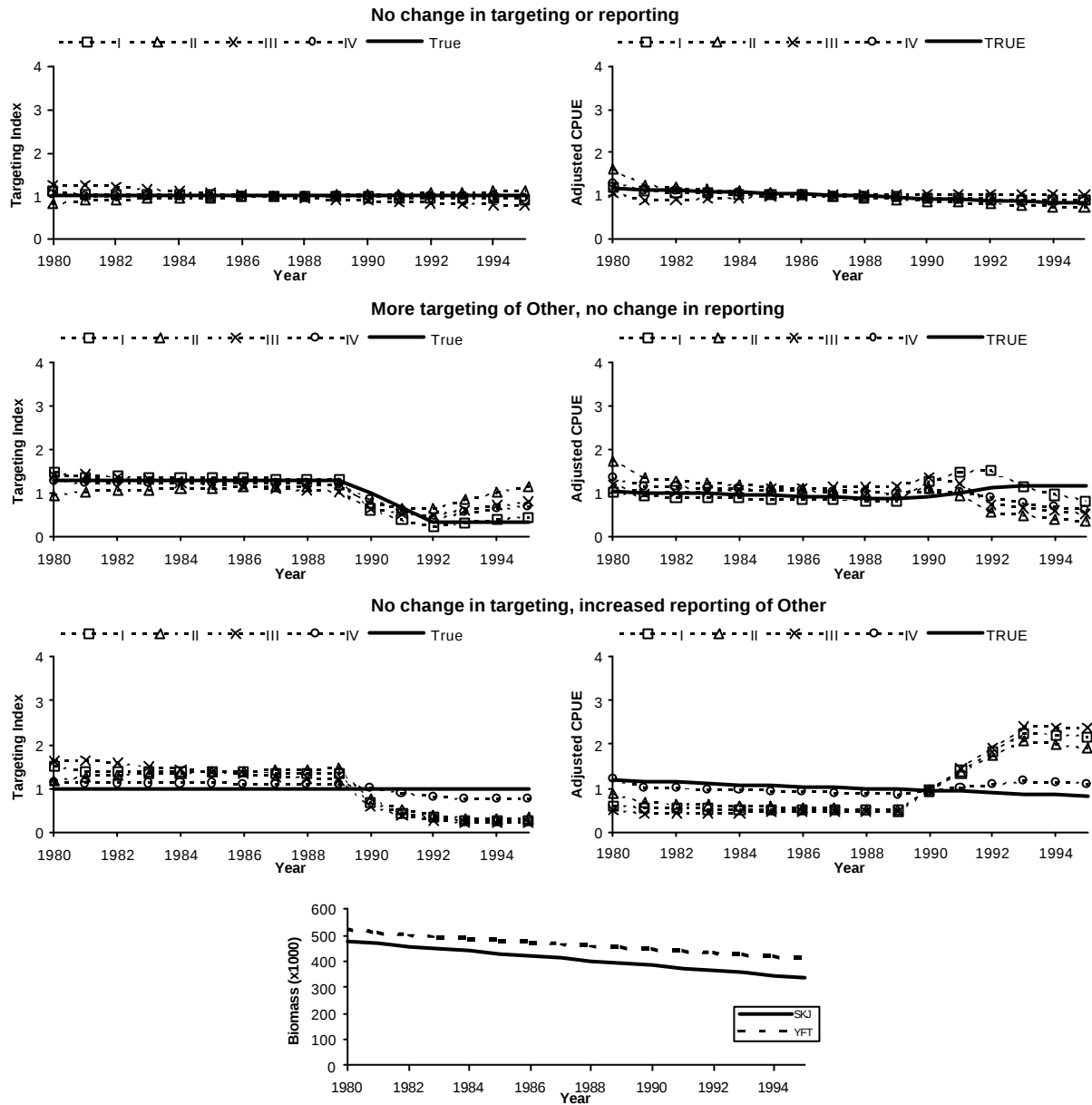


Figure 1 (cont.). Sample results of the targeting simulations. The bottom panel shows the relative trends in biomass for the two hypothetical stocks. The various targeting indices (I to IV, see text) and derived CPUE indices are shown in the left- and right hand side panels, respectively, for three scenarios of changes in targeting or reporting.

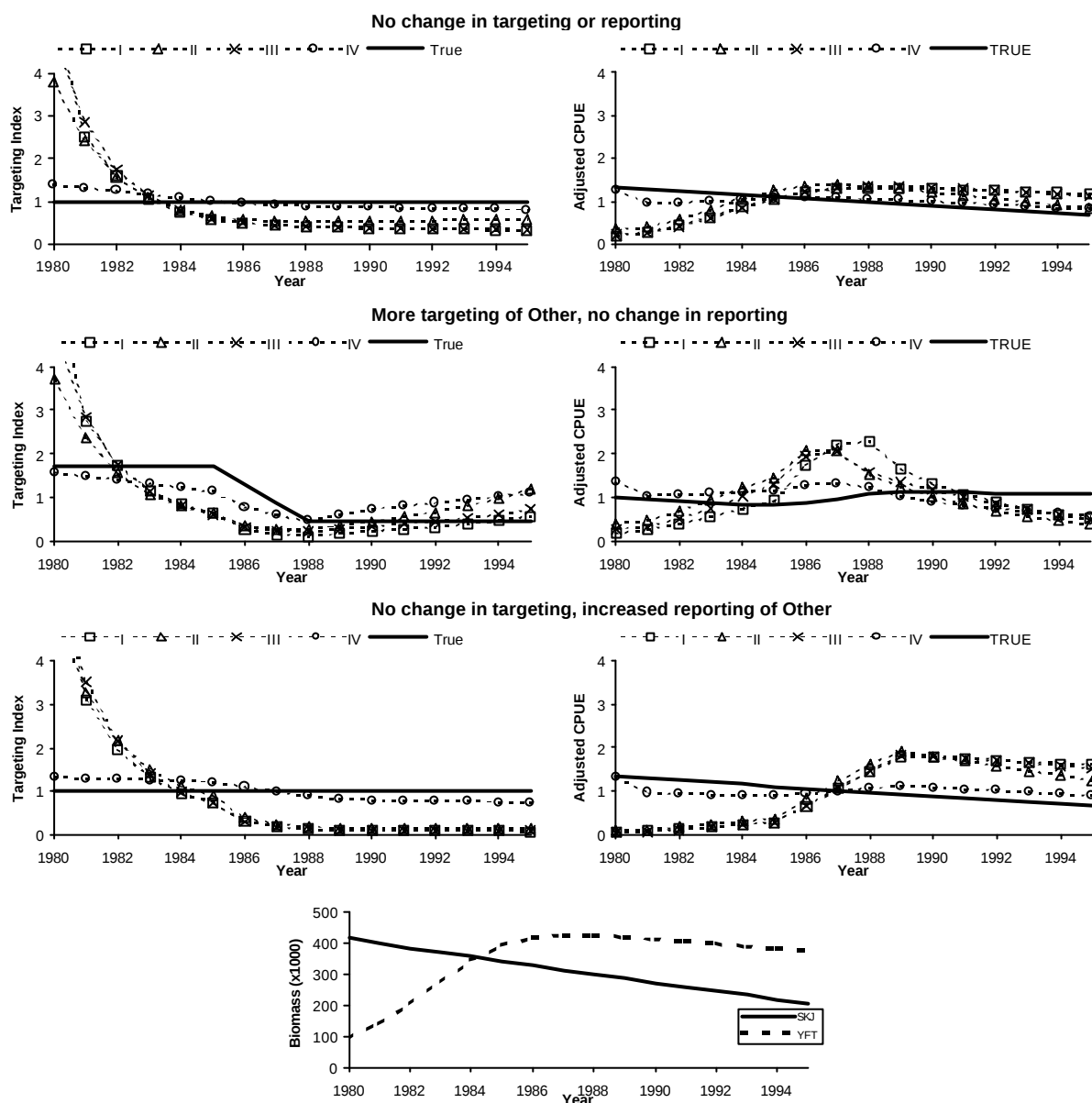


Figure 1 (cont.). Sample results of the targeting simulations. The bottompanel shows the relative trends in biomass for the two hypothetical stocks. The various targeting indices (I to IV, see text) and derived CPUE indices are shown in the left- and right hand side panels, respectively, for three scenarios of changes in targeting or reporting.

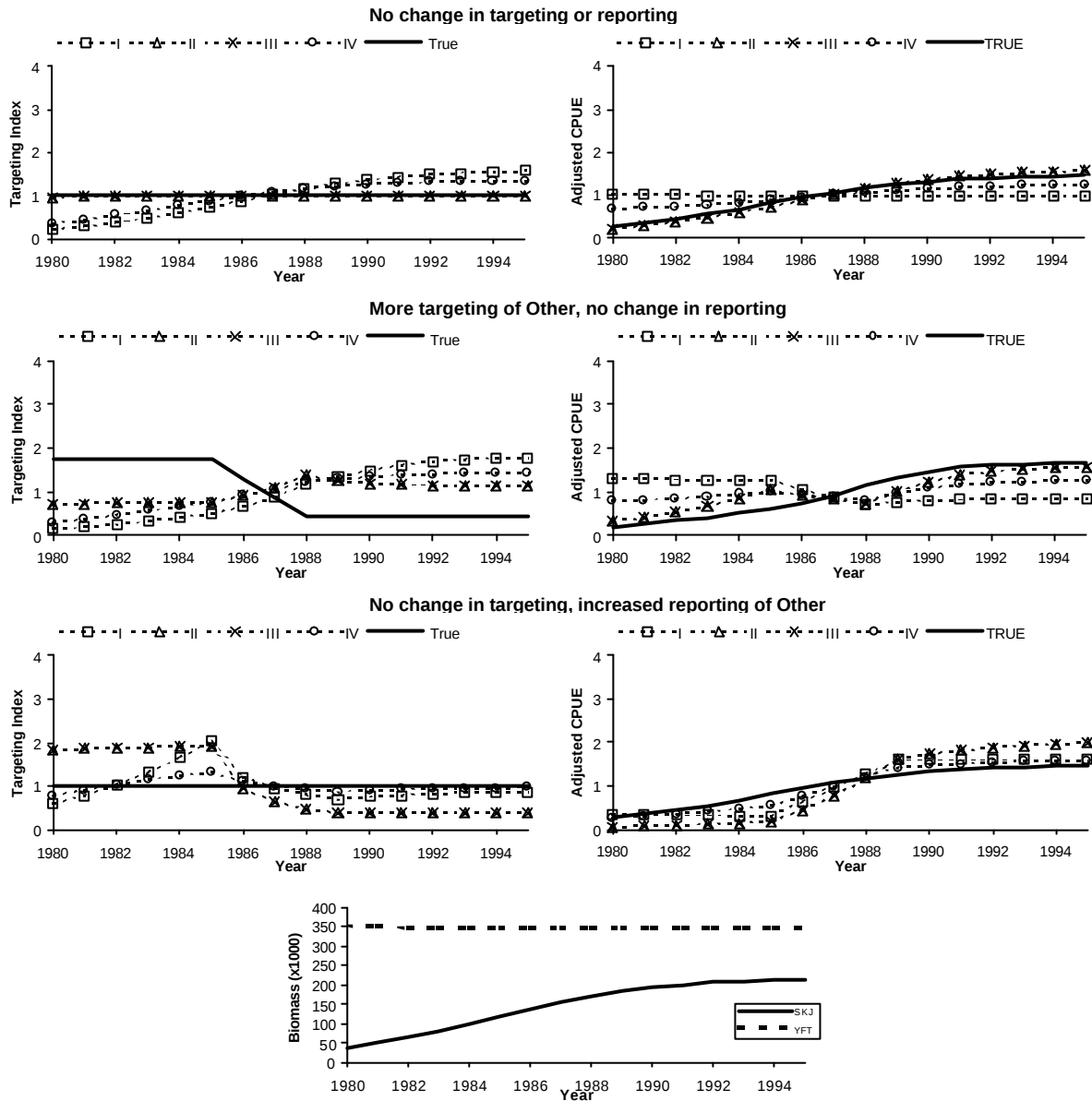


Figure 1 (cont.). Sample results of the targeting simulations. The bottom panel shows the relative trends in biomass for the two hypothetical stocks. The various targeting indices (I to IV, see text) and derived CPUE indices are shown in the left- and right hand side panels, respectively, for three scenarios of changes in targeting or reporting.

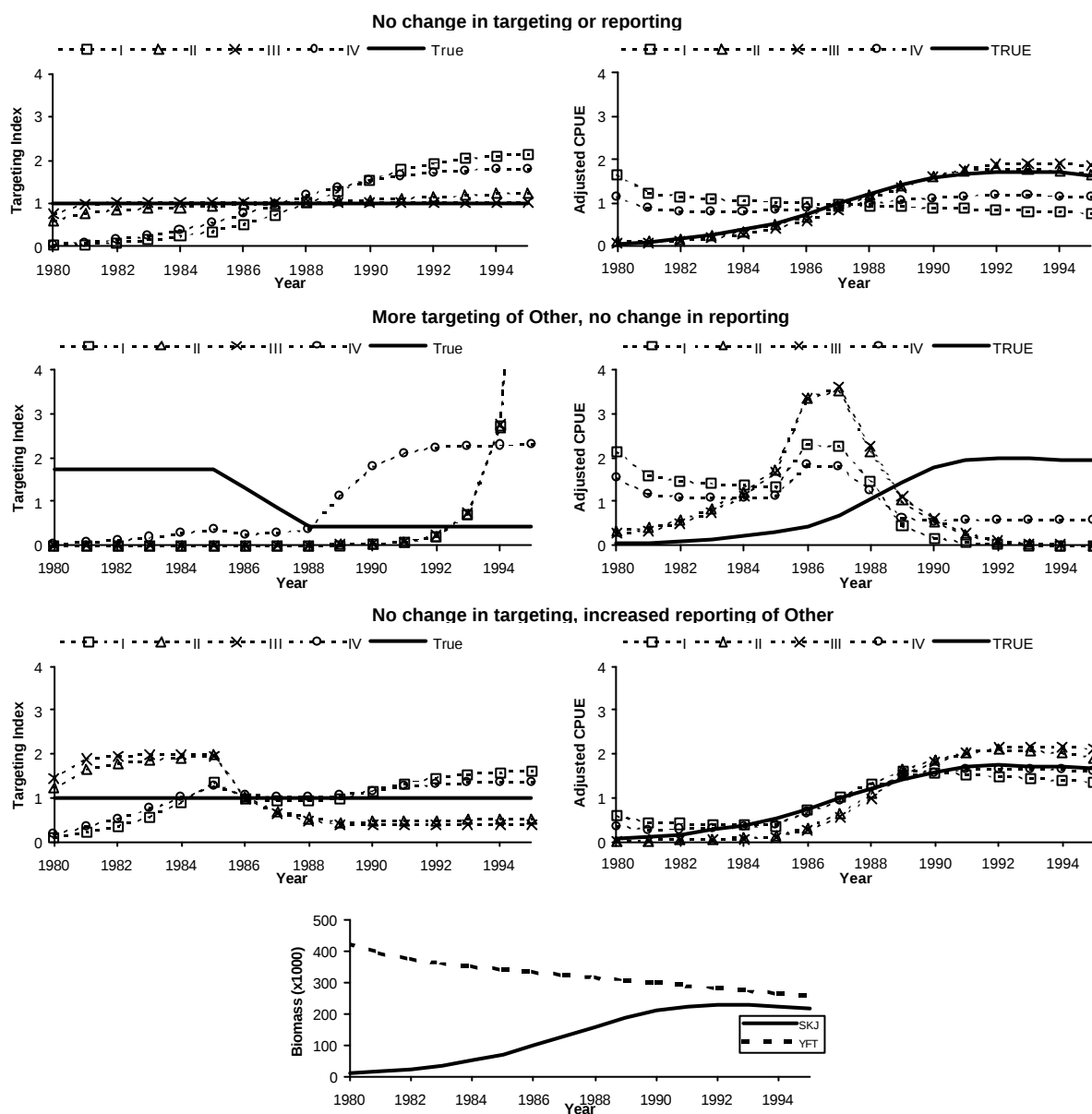


Figure 1 (cont.). Sample results of the targeting simulations. The bottom panel shows the relative trends in biomass for the two hypothetical stocks. The various targeting indices (I to IV, see text) and derived CPUE indices are shown in the left- and right hand side panels, respectively, for three scenarios of changes in targeting or reporting.

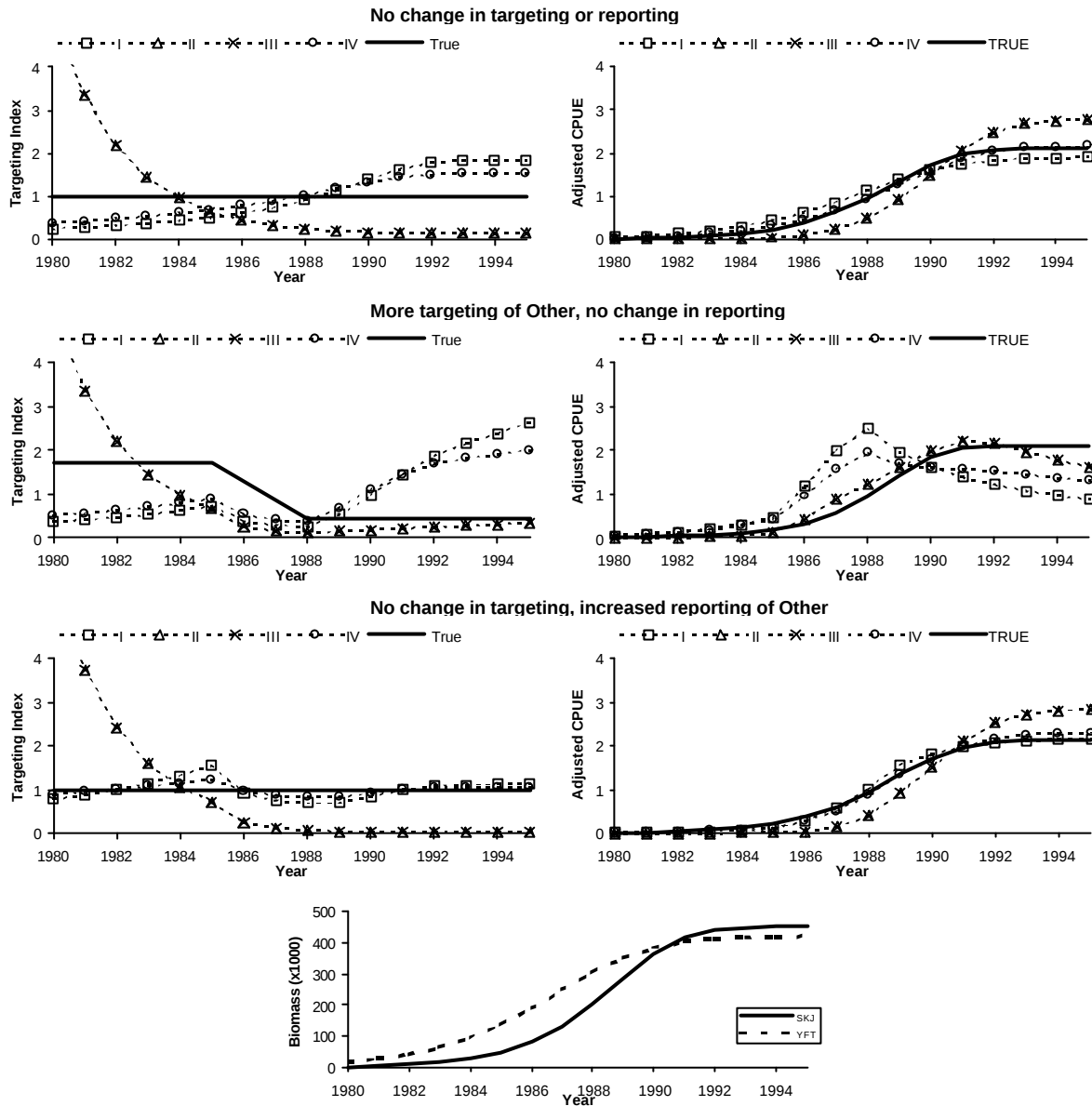


Figure 1 (cont.). Sample results of the targeting simulations. The bottom panel shows the relative trends in biomass for the two hypothetical stocks. The various targeting indices (I to IV, see text) and derived CPUE indices are shown in the left- and right hand side panels, respectively, for three scenarios of changes in targeting or reporting.

Appendix 1

ICCAT MEETING OF THE WORKING GROUP ON ASSESSMENT METHODS *(Madrid, Spain 8-11 May 2000)*

AGENDA

1. Opening, adoption of agenda and arrangements for the meeting.
2. Quality Management Systems
 - 2.1 Presentation and discussion of documents
 - 2.2 Case Studies
 - 2.3 Identification of tasks to be reviewed
 - 2.4 Development of methods review plan, including priority list
3. CPUE Standardization
 - 3.1 Presentation and discussion of documents
 - 3.2 Review of past studies and findings
 - 3.3 Testing of data sets
 - 3.4 Recommendations
4. Stock Rebuilding and Assessment Methods
 - 4.1 Presentation and discussion of documents
 - 4.2 Discussion of current practice in other fisheries fora
 - 4.3 Recommendations
5. Other Matters
6. Adoption of the Report.
7. Adjournment.

Appendice 1

ORDRE DU JOUR

- 1 Ouverture, adoption de l'ordre du jour et organisation des sessions
- 2 Systèmes de contrôle de qualité
 - 2.1 Présentation et examen des documents
 - 2.2 Études de cas
 - 2.3 Identification des tâches à réviser
 - 2.4 Élaboration du plan de révision des méthodes, y compris liste des priorités
- 3 Standardisation des CPUE
 - 3.1 Présentation et examen des documents
 - 3.2 Analyse des études antérieures et conclusions
 - 3.3 Test d'ensembles de données
 - 3.4 Recommandations
- 4 Méthodes de rétablissement des stocks et d'évaluation
 - 4.1 Présentation et examen des documents
 - 4.2 Analyse de la pratique actuelle dans d'autres enceintes de pêche
 - 4.3 Recommandations
- 5 Autres questions
- 6 Adoption du rapport
- 7 Clôture

ORDEN DEL DIA

- 1 Apertura, adopción del Orden del día y disposiciones para la reunión
- 2 Sistema de Control de calidad
 - 2.1 Presentación y discusión de los documentos
 - 2.2 Estudio de casos
 - 2.3 Identificación de las tareas a revisar
 - 2.4 Desarrollo del plan de revisión de métodos, incluyendo la lista de prioridades
- 3 Estandarización de CPUE
 - 3.1 Presentación y discusión de los documentos
 - 3.2 Examen de anteriores estudios y conclusiones
 - 3.3 Comprobación de bases de datos
 - 3.4 Recomendaciones
- 4 Recuperación del stock y Métodos de evaluación
 - 4.1 Presentación y discusión de los documentos
 - 4.2 Discusión de las prácticas actuales en otros foros de pesquerías
 - 4.3 Recomendaciones
- 5 Otros asuntos
- 6 Adopción del informe
- 7 Clausura

LIST OF PARTICIPANTS /LISTE DES PARTICIPANTS/ LISTA DE PARTICIPANTES

EUROPEAN COMMUNITY/COMMUNAUTÉ EUROPÉENNE/COMUNIDAD EUROPEA

Patterson, Kenneth

European Commission, DG Fish, Rue de la Loi, 200, BELGIUM

Tel: +322 299 2176

Fax: +322 295 5621

E-Mail: kenneth.patterson@cec.eu.int

Die, David

Cooperative Unit for Fisheries Education and Research, Rosentiel School of Marine and Atmospheric Studies, University of Miami, 4600 Rickenbacker Causeway, Miami, Florida 33149, EE.UU.

Tel: +305 361 4607

Fax: +305 361 4607

E-Mail: ddie@rsmas.miami.edu

Mejuto García, Jaime

Instituto Español de Oceanografía, Muelle de Animas, s/n, Apartado 130, 15080 - A Coruña, ESPAÑA

Tel: +34 981-205 366

Fax: +34 981-229 077

E-Mail: jaime.mejuto@co.ieo.es

Ortiz de Zárate Vidal, Victoria

Instituto Español de Oceanografía, Promontorio de San Martín, s/n, 39012 - Santander, ESPAÑA

Tel: +34 942-29 10 60

Fax: +34 942-27 50 72

E-Mail: victoria.zarate@st.ieo.es

Pallarés, Pilar

Instituto Español de Oceanografía, Corazón de María 8, 28002 - Madrid, ESPAÑA

Tel: +34 91 347 3620

Fax: +34 91 413 5597

E-Mail: pilar.pallares@md.ieo.es

Soto Ruiz, Maria

Instituto Español de Oceanografía, Corazón de María 8, 28002 - Madrid, ESPAÑA

Tel: 91 347 3620

Fax: 91 413 5597

E-Mail: maria.soto@md.ieo.es

Gaertner, Daniel

I.R.D./HEA, 911 Avenue Agropolis, B.P. 5045, 34032 Montpellier Cedex 1, FRANCE

Tel: +33-4-67 63 6981

Fax: +33-4-67 63 8778

E-Mail: gaertner@ird.fr

Maury, Olivier

I.R.D. /HEA, 911 Avenue Agropolis, B.P.5045, 34032 Montpellier Cedex 1, FRANCE

Tel: +33 4 67 63 69 82

Fax: +33 4 67 64 87 78

E-Mail: olivier.maury@mpl.ird.fr

Pianet,Renaud

I.R.D./HEA, B.P.5045, 34032 Montpellier Cedex 1, FRANCE

Tel: +33-4-67 63 69 83

Fax: +33-4-67 63 87 78

E-Mail:pianet@ird.fr

Keatinge,Michael

BIM (The Irish Seafisheries Board), Crofton Road, Dun Laoghaire, Dublin, IRELAND

Tel: +353 1 214 4278

Fax: +353 1 284 1123

E-Mail:keatinge@bim.ie

Pereira, João G.

Universidade dos Açores, Departamento de Oceanografia e Pescas, 9900 - Horta, Açores, PORTUGAL

Tel: +351-292-29 2945

Fax: +351-292-29 2659

E-Mail:pereira@horta.uac.pt

UNITED STATES/ETATS-UNIS/ESTADOS UNIDOS**Babcock,Elisabeth**

Wildlife Conservation Society, Marine Conservation Program - Bronx 200, Bronx, N.Y. 10460, EE.UU.

Tel: 00-1-718-220 2151

Fax: 00-1-718-364 4275

E-Mail:bbabcock@wcs.org

Legault,Christopher

Research Fisheries Biologist, NMFS-Southeast Fisheries Center, 75 Virginia Beach Drive, Miami, FL. 33177, EE.UU.

Tel: +1-305-361 4232

Fax: +1-305-361 4219

E-Mail:chris.legault@noaa.gov

Porch,Clay E.

Research Fisheries Biologist, NMFS-Southeast Fisheries Center, 75 Virginia Beach Drive, Miami, FL. 33177, EE.UU.

Tel: +1-305-361 4232

Fax: +1-305-361 4219

E-Mail:clay.porch@noaa.gov

Powers,Joseph E.

SCRS Chairman - Southeast Fisheries Center, 75 Virginia Beach Drive, Miami, FL 33149-1099, EE.UU.

Tel: +1-305-361 4295

Fax: +1-305-361 4278

E-Mail:joseph.powers@noaa.gov

Scott,Gerald P.

Southeast Fisheries Science Center - NMFS, 75 Virginia Beach Drive, Miami, FL. 33149-1099, EE.UU.

Tel: +1-305-361 4220

Fax: +1-305-361 4219

E-Mail:germy.scott@noaa.gov

Butterworth,Doug

Professor, Dept. Mathematics and Applied Mathematics, University of Cape Town, Rondebosch 7701, SOUTH AFRICA

Tel: +27-21-650 2343

Fax: +27-21-650 2334

E-Mail:ddl@maths.uct.ac.za

Geromont,Helena

Dept. Mathematics and Applied Mathematics, University of Cape Town, Rondebosch 7701, SOUTH AFRICA

Tel: +27 21 650 3656

Fax: +27 21 686 0477

E-Mail:helena@maths.uct.ac.za

McAllister,Murdoch

Renewable Resources Assessment Group, TH Huxley School of Environment, Engineering and Earth Science - Imperial College, 8 Prince's Gardens, London SW7 1NA, UNITED KINGDOM

Tel: 00-44-207 594 9330

Fax: 00-44-207 589 5319

E-Mail:m.mcallister@ic.ac.uk

JAPAN/JAPON/JAPÓN**Hiramatsu,Kazuhiko**

Senior Scientist, National Research Institute of Far Seas Fisheries, 7-1, 5 Chome Orido, Shimizu - Shizuoka 424-8633

JAPAN

Tel: +81-543-36 6014

Fax: +81-543-35 9642

E-Mail:hira@enyo.affrc.go.jp

Miyabe,Naozumi

National Research Institute of Far Seas Fisheries, 7-1 5 Chome Orido, Shimizu - Shizuoka 424-8633, JAPAN

Tel: +81-543-366 045

Fax: +81-543-359 642

E-Mail:miyabe@enyo.affrc.go.jp

Takeuchi,Yukio

National Research Institute of Far Seas Fisheries, 7-1, 5-Chome Orido, Shimizu-shi, Shizuoka 424-8633, JAPAN

Tel: +81-543 36 9639

Fax: +81-543-35 9642

E-Mail:yukiot@enyo.affrc.go.jp

**INTERGOVERNMENTAL ORGANIZATIONS/ORGANISMES INTERGOUVERNEMENTAUX/
ORGANISMOS INTERGUBERNAMENTALES****ICES****Lassen,Hans**

International Council for the Exploration of the Seas (ICES), Palaegade 2-4, DK-1261 Copenhagen K, DINAMARCA

Tel: +45-33 15 42 25

Fax: +45-33 93 42 15

E-Mail:hans@ices.dk

IOTC

Anganuzzi, Alejandro

Deputy Secretary, Indian Ocean Tuna Commission, P.O. Box 1011, Victoria - Mahe, Seychelles

Tel: +248 22 54 94

Fax: +248 22 43 64

E-Mail: aanganu@seychelles.net

ICCAT

Miyake, M. Peter

Restrepo, Victor

ICCAT, Corazón de María 8 - 6ª planta, 28002 Madrid

Tel: +34 91 416 56 00

Fax: +34 91 415 26 12

E-mail: info@iccat.es

(Invited external participant /Participant externe invité/ Participante externo invitado)

Medley, Paul

Sunny View, Main Street (Opposite the Village Hall), Alne YO61 1RT, UNITED KINGDOM

Tel: + 1347 838236

Fax: + 1347 838236

E-Mail: paul.medley@mail.virgin.net

LIST OF DOCUMENTS/LISTE DES DOCUMENTS/ LISTA DE DOCUMENTOS

- SCRS/00/33 Possible extensions to the ADAPT VPA model applied to western North Atlantic bluefin tuna, addressing in particular the need to account for "additional variance" - Geromont, H.F., D.S. Butterworth
- SCRS/00/34 Using Bayesian methods and decision analysis as a rational basis for dealing with conflicting stock assessment results while providing management advice on stock rebuilding - M. McAllister, M. E.A. Babcock, E.K. Pikitch
- SCRS/00/35 Comparisons of weighting schemes for tuned virtual population analyses - Legault, C.M., C.E. Porch
- SCRS/00/36 Approximate estimate of the MSY from catch data without effort information. Application to tuna fisheries - Gaertner, D., A. Fonteneau
- SCRS/00/37 Multi-fleet non-equilibrium production models including stock surface to estimate catchability trends and fishery dynamics in a Bayesian context. Application to the skipjack tuna fishery (*Katsuwonus pelamis*) in the Atlantic Ocean - Maury, O.
- SCRS/00/38 ICES Quality Assurance Policy for Fish Stock Assessment and Management Advice - Lassen, H., H. Sparholt
- SCRS/00/39 Integrating CPUE standardisation within stock assessment - Medley, P.A.H.
- SCRS/00/40 Proposed stock assessment quality control procedures for ICCAT - Restrepo, V.R.

Reference documents distributed during the meeting /Documentation de référence distribuée à la réunion/Documentación de referencia distribuida en la reunión

- Anon. Report of the bluefin tuna methodology session (Excerpts). *ICCAT Coll. Vol. Sci. Pap.* Vol. 46-1, pp 187-268.
- Anon. Detailed swordfish 1999 assessment report (Excerpts). In prep.
- Brown, C.A., and C.E. Porch. A numerical evaluation of lognormal, delta-lognormal and Poisson models for standardizing indices of abundance from West Atlantic bluefin catch per unit effort data (preliminary results). *ICCAT Coll. Vol. Sci. Pap.* Vol 46-4 . pp 476-482.
- Butterworth, D.S., J.N. Ianelli, and R. Hilborn. In prep. A statistical time-series model for stock assessment of southern bluefin tuna. MARAM, University of Cape Town, Rondebosch 7701, South Africa.
- Butterworth, D.S., and H.F. Geromont. In prep. Simulation testing as an approach to evaluate the reliability of assessment methods: An example involving initial consideration of the one/two stock hypothesis for North Atlantic bluefin tuna. MARAM, University of Cape Town, Rondebosch 7701, South Africa.

- Cooke, J.G. A procedure for using catch-effort indices in bluefin tuna assessments (revised). *ICCAT Coll. Vol. Sci. Pap.* Vol. 46-2. pp 228-232.
- Cooke, J.G., and K. Lankester. Consideration of statistical models for catch-effort indices for use in tuning VPAs. *ICCAT Coll. Vol. Sci. Pap.* Vol. 45-2. pp. 125-131
- Glazer, J. 2000. Some thoughts on the by-catch issue for West Coast hake. Marine and Coastal Management, Private Bag X2, Rogge Bay, 8012, South Africa.
- Glazer, J., and D.S. Butterworth. In prep. GLM-based standardization of the South African West Coast hake catch per unit effort series, focusing on adjustments for targeting at other species. Marine and Coastal Management, Private Bag X2, Rogge Bay, 8012, South Africa.
- Mejuto, J., and J.M. de la Serna. Standardized catch rates by age and in biomass for the North Atlantic swordfish (*Xiphias gladius*) from the Spanish longline fleet for the period 1983-1998 and bias produced by changes in the fishing strategy. SCRS/99/056.
- O'Brien, C.M., and L.T. Kell. The use of generalized linear models for the modelling of catch-effort series. I. Theory. *ICCAT Coll. Vol. Sci. Pap.* Vol. 46-4. pp. 476-482.
- Palma, C., J. Pereira, J. Mejuto, and M. Santos. Effect of simulating targeting levels on swordfish (*Xiphias gladius*) standardised CPUE estimates, caught by the Portuguese surface longline fleet in the North Atlantic. SCRS/99/129.
- Porch, C.E. The implications of using the frequency of zero catches and other measures as indices of abundance. *ICCAT Coll. Vol. Sci. Pap.* Vol. 46-2. pp. 237-241
- Punt, A.E., R.W. Leslie, and A.J. Penney. A preliminary examination of the Taiwanese longline catch and effort data (1967 to 1992) for South Atlantic albacore (*Thunnus alalunga*). *ICCAT Coll. Vol. Sci. Pap.* Vol. 43 pp. 225-245.
- Takeuchi, Y., and K. Yokawa. A note on methods to account targeting in CPUE standardization. SCRS/99/125.

Other citations /Autres textes cités/Otros trabajos citados

- Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. 1996. Markov Chain Monte Carlo in Practice. Chapman and Hall, London. 486p.
- Myers, R.A., W.A. Blanchard, and K.R. Thompson. 1990. Summary of North Atlantic fish recruitment 1942-1987. Can. Tech. Rep. Fish. Aquat. Sci. 1743. 108pp.

ASSESSMENT PROGRAM DOCUMENTATION

1. Program name

Name of the application (e.g., PROFIT)

2. Version (date) ***

Version number given by the programmer

3. Language

Computer language(s) in which the application is written (e.g., Pascal).

4. Programmer / contact person

Name, address, email of the programmer and/or persons that can be contacted with questions about the application

5. Distribution limitations

List any limitations imposed by the programmer (e.g. only executable code can be distributed)

6. Compiler needs / stand-alone

List additional software needed to run the program (e.g. C++), operating system if other than DOS/Windows, or state if the executable program does not need other software to run

7. Purpose

Brief description of what the application does (e.g., “fits a stock production model to catch and effort data making an equilibrium approximation”).

8. Description

A full description of what the application does, sufficient to allow replication. It should clearly state the major statistical and biological assumptions.

9. Required inputs

Description of the inputs needed.

10. Program outputs

Description of the program outputs.

11. Diagnostics

A list of the diagnostics produced by the program that can be used to guide modeling.

12. Other features

List of special features of the application that are not “traditional” ones (e.g., uses tagging data as auxiliary information).

13. History of method peer review

Detailed history of the review of the method, citing publications and reports.

14. Steps taken by programmer for validation

Detailed history of how the application itself (as opposed to the method) has been validated by the

programmer (e.g. by generating simulated data of known characteristics and then using the program on the simulated data)

15. Tests conducted by others

A list of the tests that persons other than the programmer (e.g., the Secretariat) have conducted.

16. Notes by iccat

Comments by a Coordinating Committee, group or person in charge of cataloguing the software, especially with reference to validation. This will essentially be the “certification” or “approval” of the software.

APPENDIX 1. Algorithm

A list of steps explaining how the program operates. This is intended to complement Item 8 and Appendix 2.

APPENDIX 2. User’s guide

A user’s guide, sufficiently detailed to allow others to run the program.

APPENDIX 3. Worked example

An example containing inputs, program run options and outputs.

APPENDIX 4. Source code

The source code, taking into account copyright restrictions. This code may or may not be distributed to persons outside the Secretariat depending on the entry on Item 5.

*** Note: Each new version of a catalogue entry should be accompanied by documentation that explains the major features associated with the change (e.g. new program options, bug fixes, etc.).

DOCUMENTATION SUR LE PROGRAMME D'ÉVALUATION

- 1 Nom du Programme**
Nom de l'application (par ex. PROFIT)
- 2 Version (date) *****
Numéro de la version donné par le programmeur
- 3 Langage**
Langage(s) informatique(s) dans lequel/lesquels l'application est écrite (par ex. Pascal).
- 4 Programmeur /personne à contacter**
Nom, adresse, e-mail du programmeur et/ou des personnes à contacter en cas de questions sur l'application
- 5 Limites de diffusion**
Répertorier toutes limites imposées par le programmeur (par ex. seuls les codes d'exécution peuvent être diffusés)
- 6 Besoins du compilateur / autonomie**
Répertorier le logiciel additionnel requis pour exécuter le programme (par ex. C++), le système d'exploitation, s'il s'avère différent de DOS/Windows, ou bien indiquer si le programme d'exécution n'a pas besoin d'un autre logiciel
- 7 Objet**
Brève description de ce que fait l'application (par ex. "ajuste un modèle de production du stock à des données de prise et d'effort en faisant une approximation d'équilibre").
- 8 Description**
Description complète de ce que fait l'application, suffisamment pour permettre la reproduction. Elle devrait indiquer clairement les principaux postulats statistiques et biologiques.
- 9 Valeurs d'entrée requises**
Description des valeurs d'entrée requises.
- 10 Valeurs de sortie du programme**
Description des valeurs de sortie du programme.
- 11 Diagnostics**
Liste des diagnostics produits par le programme qui peuvent être utilisés pour guider la modélisation.
- 12 Autres caractéristiques**
Liste des caractéristiques spéciales de l'application qui ne sont pas "traditionnelles" (par ex. utilisant des données de marquage comme information subsidiaire).

13 Historique de la méthode peer review

Historique détaillé de la méthode, avec citation des publications et des rapports.

14 Mesures prises par le programmeur aux fins de la validation

Historique détaillé de la façon dont l'application elle-même (par opposition à la méthode) a été validée par le programmeur (par ex. en créant des données simulées de caractéristiques connues et en utilisant ensuite le programme sur les données simulées)

15 Tests réalisés par d'autres personnes

Liste des tests réalisés par d'autres personnes (par ex. le Secrétariat), à l'exception du programmeur.

16 Commentaires de l'ICCAT

Commentaires d'un Comité de coordination, groupe ou personne chargé de cataloguer le logiciel, notamment en ce qui concerne la validation. Il s'agira essentiellement de la "certification" ou de l'"approbation" du logiciel.

APPENDICE 1. Algorithme

Liste des étapes expliquant le fonctionnement du programme. Celle-ci est destinée à compléter le point 8 de l'ordre du jour et l'appendice 2.

APPENDICE 2. Guide de l'utilisateur

Guide de l'utilisateur suffisamment détaillé pour permettre à d'autres personnes d'exécuter le programme.

APPENDICE 3. Exemple concret

Exemple comprenant des valeurs d'entrée, des options de passage du programme et des valeurs de sortie.

APPENDICE 4. Code source

Code source, tenant compte des restrictions en matière de droit d'auteur. Ce code peut ou peut ne pas être diffusé aux gens en dehors du Secrétariat en fonction de ce qui est spécifié au point 5 de l'ordre du jour.

***** Remarque:** Chaque nouvelle version d'une donnée de catalogue doit contenir la documentation qui explique les principales caractéristiques associées à la modification (par ex. nouvelles options du programme, réparation des virus, etc.).

DOCUMENTACIÓN PARA EL PROGRAMA DE EVALUACIÓN

- 1 Nombre del Programa**
Nombre de la aplicación (por ejemplo, PRODFIT).
- 2 Versión (fecha) *****
Número de versión asignado por el programador.
- 3 Lenguaje**
Lenguaje o lenguajes de ordenador que utiliza el programa (por ejemplo, Pascal).
- 4 Programador/Persona de contacto**
Nombre, dirección, E-mail del programador y/o personas con las que se puede establecer contacto para hacer preguntas sobre la aplicación.
- 5 Limitaciones respecto a distribución**
Cualquier limitación que haya sido impuesta por el programador (por ejemplo, solo se puede distribuir el código de ejecución).
- 6 Requisitos del compilador/Stand alone**
El software adicional necesario para ejecutar el programa (por ej. C++), sistema operativo, si no se trata del DOS/Windows, o bien aclarar que el programa de ejecución se arranca solo.
- 7 Objetivo**
Breve descripción de lo que lleva a cabo la aplicación (por ejemplo: “ajusta el modelo de producción del stock a la captura y datos de esfuerzo, realizando una aproximación equilibrada”).
- 8 Descripción**
Descripción completa de lo que lleva a cabo la aplicación, suficiente para permitir la duplicación. Se deben señalar con claridad los principales supuestos estadísticos y biológicos.
- 9 Entradas**
Descripción de las entradas necesarias.
- 10 Resultados**
Descripción de los resultados del programa.
- 11 Diagnósticos**
Lista de los diagnósticos resultantes del programa que pueden usarse como guía en la modelación.
- 12 Otras características**
Lista de las características especiales de la aplicación que no sean las “tradicionales” (por ejemplo: emplea datos demarcado como información auxiliar).
- 13 Historial del método de revisión por pares**
Historial detallado del método, citando publicaciones e informes.

- 14 Pasos dados por el programador para la validación**
Historial detallado de cómo ha sido validada la aplicación en sí (en contraposición al método) por el programador (por ejemplo, generando datos simulados de características conocidas y después, aplicando el programa a los datos simulados).
- 15 Pruebas efectuadas por otras personas**
Lista de las pruebas efectuadas por otras personas, aparte del programador (por ejemplo, la Secretaría).
- 16 Notas de ICCAT**
Comentarios del Comité de Coordinación, grupo o personas encargadas de la catalogación del software, en particular respecto a su validación. Esto será esencialmente la “certificación” o aprobación” del software.

APÉNDICE 1. Algoritmo

Lista de los pasos que explican el funcionamiento del programa. Está destinado a complementar el punto 8 y el Apéndice 2.

APÉNDICE 2. Guía del usuario

Guía destinada al usuario, con los suficientes detalles para permitir que otras personas se sirvan del programa.

APÉNDICE 3. Ejemplo

Ejemplo con entradas, opciones de ejecución del programa y resultados.

APÉNDICE 4. Código de origen

Código de origen, teniendo en cuenta las restricciones respecto a *copyright*. Este código puede, o no, distribuirse a personas ajenas a la Secretaría, dependiendo de lo estipulado en el punto 5.

***** NOTA** Cada nueva incorporación al catálogo deberá ir acompañada de documentación que explique las principales características asociadas al cambio (por ejemplo, nuevas opciones del programa, corrección de virus, etc.)

1 The problem

Catch rates are often positively correlated between species on a haul by haul basis while changes in fishing strategy where the vessel change target species will introduce a negative correlation between catch rates of different species. Furthermore, there may be area and season effects in the relative catch rates.

In a GLM framework this could be formulated

$$\log cpue_{specieshaul} = year_{species} + \text{Model for targeting and substitution} + environment_{haul} + e$$

$species = 1, 2, \dots$
 $haul = 1, 2, \dots$

There are several points in this formulation

- The catch rates all appear as dependent variables associated with noise usually quite large
- Substitution between species is modeled explicit
- The formulation includes an environment effect. This environment effect is assumed to have the same relative effect on all species and represents a positive correlation between catch rates of different species
- The environment parameter may in some cases be available based on a proxy, e.g. wind force but normally the parameter is not observed.
- There is a difference in time scale between the terms in the model formulation. Targeting is usually assumed to operate over longer time periods than do the environment influence, e.g. storms giving low catch rates usually last only days, while more fundamental price relations have a longer time perspective.

2 Model for targeting and substitution

Such models should be based on an understanding of how the fishery operate, i.e. the model will be fishery specific. In principle the model shall reflect pay off between targeting one species and loosing on other species.

The degree to which this is actually done may be inferred by studying the revenue on a haul by haul basis or trip-by-trip basis. It is certainly a distinct feature in demersal fisheries that often target several species simultaneously, e.g. the groundfish fisheries in the Northern North Sea targeting cod, haddock, whiting and to some degree plaice.

3 Example

To illustrate the idea a simple example was constructed based on simulated data. Abundance was available for two species for three years. Data for 25 hauls with 10 hauls in each of the two first years and five hauls in the last year was generated.

The target model was formulated based on abundance and price

$$T = \frac{N * p}{\sum_{species} N_{species} * p_{species}}$$

and the CPUE of a species in a haul then becomes

$$Cpue_{species,haul} = N_{species} * T_{species} * g_{haul} * e$$

with

$$e = \log N(0, s^2)$$

Pricesp are known

g was generated as a uniform random variable. The annual mean of log value = 0

This model needs a number of normalization to be fully specified. Those chosen are

- The environmental factor should reflect short term variability where catch rates of different species co-varyate therefore the average over longer time periods should be the same between years,

differences between years should be reflected as differences in abundance $\sum_{t \in year} \log g_t = 0$

- The model for targeting/substitution is subject to $\sum_{species} T_{species} = 1$. This is in this example included directly in the model.

Figure A1 shows the generated catch series for 25 hauls generated with log-normal errors with $s = 0.5$

The correct target/substitution model was used in the fit. The model is fitted by minimizing the deviance of log residuals without with equal weighting of the contributions from the CPUE of different species. The model fit was subject to the constraints that summed environment factors = 0 on an annual basis. The calculations were done in EXCEL using the SOLVER facility. Constraints were introduced through a penalty function for the summed environment parameters on an annual basis. In formulae this deviance function

$$\text{was } \sum_{species} \sum_{haul} \left(\log cpue_{species,haul} - a_{species,haul} - \log \frac{\exp(a_{species}) * p_{species,haul}}{\sum_{sp \in allspecies} \exp(a_{sp}) * p_{sp,haul}} - g_{haul} \right)^2 = \min$$

Annual abundance was estimated and are compared with their true values (**Table A1**). **Figure A2** compares the estimated environment parameters with those used to generate the data.

1 Le problème

Les taux de capture sont positivement corrélés entre les espèces à chaque remontée d'engins, tandis que des modifications de stratégie de pêche (c.-à-d. si le bateau change d'espèce-cible) introduisent une corrélation négative entre les taux de capture d'espèces différentes. En outre, des effets spatio-saisonniers peuvent intervenir dans les taux de capture relatifs. Dans un contexte de modélisation linéaire généralisée (GLM), ceci pourrait être formulé comme suit:

$$\log cpue_{specieshaul} = année_{species} + \text{modèle pour ciblage et substitution} + \text{environnement}_{haul} + \alpha$$

species = 1,2,...

Haul= 1,2,...

Cette formulation suscite plusieurs observations:

- Les taux de capture apparaissent tous comme des variables dépendantes associées au bruit habituellement assez élevé
- La substitution entre espèces est explicite au modèle
- La formulation comprend un effet sur l'environnement. Cet effet sur l'environnement est censé avoir le même effet relatif sur toutes les espèces et représente une corrélation positive entre les taux de capture de différentes espèces.
- Le paramètre de l'environnement peut dans certains cas être disponible sur la base d'un indice approchant, par ex. la force du vent, mais généralement le paramètre n'est pas observé.
- Il y a une différence dans le temps entre les termes dans la formulation du modèle. Le ciblage est censé fonctionner habituellement sur de plus longues périodes que ne le fait l'influence de l'environnement, par ex. les orages qui entraînent de faibles taux de capture ne durent en général que quelques jours, tandis que des rapports de prix plus fondamentaux ont une perspective temporelle plus longue.

2 Modèle pour le ciblage et la substitution

Ces modèles devraient se fonder sur une compréhension du fonctionnement de la pêche, c.-à-d. que le modèle sera spécifique de la pêche. En principe, le modèle va refléter le compromis entre le ciblage d'une espèce et la perte d'autres espèces.

On peut savoir dans quelle mesure cela se passe en réalité en examinant les recettes à chaque remontée d'engins ou à chaque sortie. Il s'agit certainement d'une caractéristique distincte des pêcheries démersales qui ciblent souvent plusieurs espèces simultanément, par ex. les pêcheries de poisson de fond dans le Nord de la Mer du Nord ciblent la morue, l'aiglefin, le merlan et dans une certaine mesure la plie.

3 Exemple

Pour illustrer l'idée, on a établi un exemple simple sur la base de données simulées. L'abondance était disponible pour deux espèces pendant trois ans. On a généré des données pour 25 coups de filet, soit

10 coups de filet les deux premières années, et cinq coup de filets la dernière année. Le modèle cible a

été formulé en fonction de l'abondance et du prix.
$$T = \frac{N * p}{\sum_{species} N_{species} * p_{species}}$$

et la CPUE d'une espèce dans un coup de filet devient alors

$$CPUE_{species,haul} = N_{species} * T_{species} * \tilde{a}_{haul} * \tilde{a}$$

avec
 $\tilde{a} = \log N(0, \sigma^2)$

Les prix p sont connus.

$\tilde{a}$ a été créé comme une variable aléatoire uniforme. Moyenne annuelle de la valeur logarithmique = 0

Ce modèle exige une complète spécification de la normalisation, à savoir:

- Le facteur environnemental devrait traduire la variabilité à court terme où les taux de capture de différentes espèces sont en covariance; la moyenne sur de plus longues périodes devrait ainsi être la même entre les années, les différences entre les années

devant être considérées comme des différences d'abondance $\sum_{t \in year} \log g_t = 0$

- Le modèle pour le ciblage/la substitution est soumis à $\sum_{species} T_{species} = 1$. Dans cet exemple, ceci est directement inclus dans le modèle.

La **Figure A1** indique la série de captures générée pour 25 coups de filet produite avec des erreurs log-normales avec $\sigma = 0.5$

Le modèle correct de ciblage/substitution a été utilisé dans l'ajustement. Le modèle est ajusté en minimisant la déviance des valeurs résiduelles logarithmiques avec une pondération égale des contributions de la CPUE de différentes espèces. L'ajustement du modèle a été soumis aux contraintes qui totalisaient les facteurs environnementaux = 0 sur une base annuelle. Les calculs ont été effectués dans EXCEL en utilisant l'option SOLVER. Des contraintes ont été introduites au moyen d'une fonction de pénalisation pour les paramètres environnementaux totalisés sur une base annuelle. Dans les formules, cette fonction de déviance était

$$\sum_{species} \sum_{haul} \left(\log cpue_{species,haul} - a_{species,haul} - \log \frac{\exp(a_{species}) * p_{species,haul}}{\sum_{sp \in allspecies} \exp(a_{sp}) * p_{sp,haul}} - g_{haul} \right)^2 = \min$$

L'abondance annuelle a été estimée et comparée avec ses valeurs réelles (**Tableau A1**). La **Figure A2** compare les paramètres environnementaux estimés avec ceux utilisés pour créer les données.

1 El problema

Con frecuencia, las tasas de capturas están relacionadas de forma positiva entre especies, lance por lance (haul), mientras que las modificaciones en la estrategia pesquera, cuando el barco cambia de especie objetivo, introducirá una correlación negativa entre las tasas de captura de diferentes especies. Además podría haber efectos de zona y temporada en las tasas relativas de capturas.

En el marco de un Modelo Lineal Generalizado, ésto se podría formular como sigue:

$$\log cpue_{species\ haul} = year_{species} + \text{modelo de dirección del esfuerzo y sustitución} + \text{medio ambiente}_{haul} + g$$

$species = 1, 2, \dots$

$haul = 1, 2, \dots$

Existen varios puntos en esta fórmula:

- Todas las tasas de captura figuran como variables dependientes asociadas a un ruido normalmente importante.
- La sustitución entre especies se modela de forma explícita.
- La fórmula incluye un efecto de medio ambiente. Se supone que este efecto tendrá la misma repercusión relativa sobre todas las especies y representa una correlación positiva entre las tasas de captura de diferentes especies.
- El parámetro de medio ambiente podría estar disponible en algunos casos basado en una aproximación, por ejemplo, la fuerza del viento, pero normalmente este parámetro no se observa.
- Hay una diferencia en la escala temporal entre los términos en la fórmula modelo. Se supone que la dirección del esfuerzo ejerce su influencia en períodos más largos que el medio ambiente, por ejemplo, las tormentas que producen tasas de captura bajas suelen durar tan sólo unos días, mientras que las cuestiones fundamentales de precio tienen una perspectiva más amplia.

2 Modelo para la dirección del esfuerzo y sustitución

Estos modelos deberían estar basados en una comprensión de la forma en que opera la flota, es decir, el modelo será específicamente para la flota. En principio, el modelo reflejará un balance entre la búsqueda de una especie y el abandono de otra especie.

El grado al cual esto tiene realmente lugar puede deducirse estudiando las ganancias, lance por lance o bien viaje por viaje. Es una característica de las pesquerías de demersales que con frecuencia dirigen su esfuerzo a varias especies simultáneamente, por ejemplo, las pesquerías de demersales en la zona norte del Mar del Norte que dirigen su esfuerzo al bacalao, eglefino, pescadilla y, en cierto grado a la solla.

3 Ejemplo

Para ilustrar el concepto se puso un sencillo ejemplo basado con datos simulados.

Se disponía de la abundancia de dos especies durante tres años. Se generaron datos para 25 lances, con 10 lances en cada uno de los dos primeros años y cinco lances en el último año.

El modelo de dirección del esfuerzo se formuló basándose en la abundancia y el

$$\text{precio: } T = \frac{N * p}{\sum_{\text{species}} N_{\text{species}} * p_{\text{species}}}$$

y la CPUE de una especie en un lance es entonces:

$$Cpue_{\text{species, haul}} = N_{\text{species}} * T_{\text{species}} * \tilde{a}_{\text{haul}} * g \quad \text{con}$$

$$g = \log N(0, \sigma^2)$$

Se conocen los precios p.

$\tilde{a}$ se generó como variable aleatoria uniforme. Media anual de valor $\log = 0$.

Este modelo precisa un número de normalización para estar totalmente especificado. Se han seleccionado:

- S** El factor de medio ambiente debería reflejar la variabilidad a corto plazo, donde las tasas de captura de diferentes especies presentan una covarianza, por tanto la media a lo largo de periodos de tiempo más largos debería ser la misma entre años, las diferencias entre años deberían quedar reflejadas como diferencias en la abundancia

$$\sum_t \log \tilde{a}_t = 0$$

- S** El modelo para la dirección del esfuerzo/sustitución está sujeto a $\sum_{\text{species}} T_{\text{species}} = 1$. Este es un ejemplo incluido directamente en el modelo.

La **Figura A1** presenta la serie de captura generada para 25 lances generados con errores lognormales con $\sigma = 0.5$.

En el ajuste se usó el modelo correcto de dirección del esfuerzo/sustitución. El modelo se ajusta minimizando la desviación de los residuos logarítmicos con igual ponderación de los aportes de la CPUE de diferentes especies.

El ajuste del modelo estaba sujeto a limitaciones; la suma de los factores de medio ambiente es igual a 0 en base anual. Los cálculos se hicieron con EXCEL usando SOLVER. Las limitaciones se introdujeron por medio de una función de penalización para la suma de parámetros de medio ambiente en base anual. La fórmula de esta función

$$\text{era: } \sum_{\text{species}} \sum_{\text{haul}} \left(\log c p u e_{\text{species}, \text{haul}} - a_{\text{species}, \text{haul}} - \log \frac{\exp(a_{\text{species}}) * p_{\text{species}, \text{haul}}}{\sum_{sp \in \text{allspecies}} \exp(a_{sp}) * p_{sp, \text{haul}}} - g_{\text{haul}} \right)^2 = \min$$

La abundancia anual se estimó y comparó con sus auténticos valores (**TablaA1**). La **FiguraA2** compara los parámetros de medio ambiente estimados con los usados para generar datos.

Table A1. Comparison between the estimated and true (simulated) abundance values.

SPECIES	YEAR	A B U N D A N C E ESTIMATE	ABUNDANCE USED TO GENERATE DATA
1	1	123.7	100
	2	161.8	200
	3	391.9	300
2	1	35.0	30
	2	20.6	20
	3	10.0	10

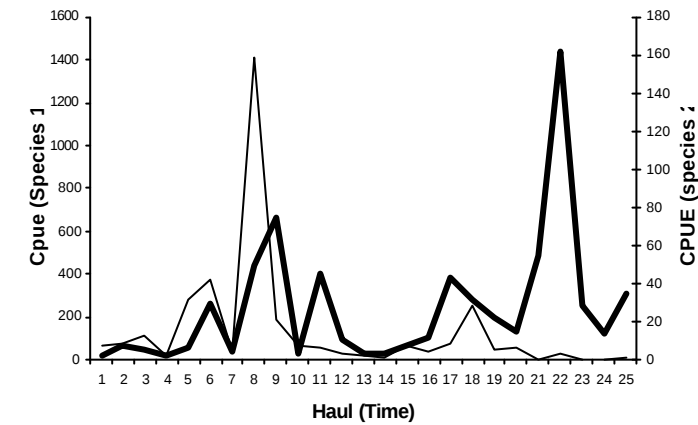


Fig A1. Catch rates for two species taken together in the same haul. Simulated data 25 haul

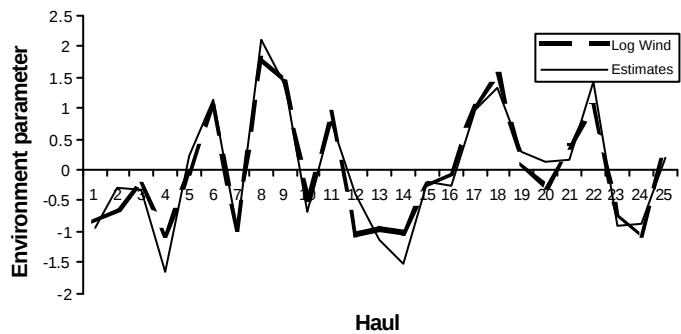


Fig A2. Estimated and "true" values of environment parameter by haul