Report on the 2004 Assessments of South Atlantic Tilefish and Snowy Grouper

NIWA Client Report: WLG2004-51
August 2004

NIWA Project: ERI05901

Report on the 2004 Assessments of South Atlantic Tilefish and Snowy Grouper

R.I.C.C. Francis

Prepared for

University of Miami

Independent System for Peer Review

NIWA Client Report: WLG2004-51
August 2004

NIWA Project: ERI05901

National Institute of Water & Atmospheric Research Ltd
301 Evans Bay Parade, Greta Point, Wellington
Private Bag 14901, Kilbirnie, Wellington, New Zealand
Phone +64-4-386 0300, Fax +64-4-386 0574
www.niwa.co.nz

Contents

Reviewed and approved for release by:

Executive Summary

The 2004 assessments of tilefish and snowy grouper in the southeast United States were reviewed as part of the SEDAR (South East Data, Assessment and Review) process. The Assessment Review Panel met 26-29 July 2004 at the Holiday Inn in Charlotte, North Carolina. The data and assessments were presented to the Panel, additional analyses were requested and carried out, and the Panel discussed the results and wrote its Consensus Report.

The data used were not strong, but were the best available, adequate for use in the assessments and, with minor exceptions, used appropriately. Given that these were first-time assessments for both stocks, the assessment techniques were sound and the results should be valuable to fishery managers. The presentation and documentation of the assessments was generally clear and detailed with only one significant exception: that the derivation of the likelihood weights was insufficiently explained.

Suggestions are given for the consideration of those charged with future assessments of these stocks.

1. BACKGROUND

This report reviews the 2004 assessments of tilefish (*Lopholatilus chamaeleonticeps*) and snowy grouper (*Epinephelus niveatus*) in the management area of the South Atlantic Fisheries Management Council, at the request of the University of Miami (see Appendix 1). In terms of recent catches, these species are the two most important of a group of eight snapper-groupers known as the South Atlantic Deep Water Complex. The author was provided with the assessment reports for both species, the report from the associated Data Workshop, and supporting documents (Appendix 2), and participated in the SEDAR 4 (South East Data, Assessment, and Review) Assessment Review Panel Workshop that considered these assessments. This workshop constituted the last of the three phases of the SEDAR 4 process, with the earlier phases being a Data Workshop (3-7 November 2003) and an Assessment Workshop (7-11 June 2004).

2. REVIEW ACTIVITIES

The Assessment Review Panel Workshop was held 26-29 July 2004 at the Holiday Inn in Charlotte, North Carolina (see Appendix 3 for the Panel membership and a list of other attendees).

The Review Panel's terms of reference, as provided by the SEDAR Coordinator (Appendix 2.1), differed somewhat from those given to the author as part of his Statement of Work (Appendix 1). The Panel followed the former, except that the last term was deleted at the instruction of the SEDAR Coordinator (i.e., the Panel was not required to compile a Stock Advisory Report).

Mike Prager gave a useful introductory talk outlining some features of U.S Fisheries Management. Doug Vaughan then discussed the data available for snowy grouper and Erik Williams presented the assessment. The panel discussed the assessment and requested some additional analyses. These were done and the results presented to the Panel (see below). The same sequence was followed for tilefish, with the data presented by Doug Vaughan and the assessment by Kyle Shertzer. The Panel then drafted their Consensus Report with input from others present.

The Panel's task was simplified because the assessments were very similar in terms of the data available, the analytical techniques, and the method of presentation.

This, and the fact that the Panel was not required to write a Stock Advisory Report, allowed the Workshop to finish a day earlier than scheduled.

2.1 Assessment structure and results

In this section I give a brief description of the two assessments in order to provide a context for the rest of this report.

In terms of data and structure, the two assessments were very similar. The observations comprised the landings by fishery, several time series of abundance indices from CPUE (catch per unit effort), and several time series of length and age frequencies (LFs and AFs). In an initial run, these observations were fitted by weighted maximum likelihood to predictions from an age-structured population model. The weights applied to each likelihood component in this run were derived subjectively after a series of preliminary model runs (not presented) which explored many alternative sets of weights. Constraints were applied to force the initial spawning stock biomass ($SSB_{initial}$) to be close to 0.9 of the virgin value (SSB_{virgin}) and to discourage extreme variation in recruitment deviations. The estimated parameters (204 for snowy grouper and 147 for tilefish) fell into five groups: virgin recruitment (1 parameter); CPUE catchabilities (3 for snowy grouper and 2 for tilefish); selectivities (20 and 20); fishing mortalities (139 and 93); and recruitment deviations (41 each). Natural mortality was modelled as age-dependent (varying as an exponential function of body weight) following Lorenzen (1996).

For some runs requested by the Panel the model was simplified substantially into what is sometimes called an age-structured production model (ASPM). This was done by increasing the likelihood weights associated with the landings to force the model to fit the landings almost exactly and making recruitment deterministic. This greatly reduced the effective number of parameters estimated (from 204 to 24 for snowy grouper and from 147 to 13 for tilefish).

In order to characterise uncertainty in the assessments a large number of Monte Carlo bootstrap (MCB) runs were done for each stock. These runs differed from the initial run in that some model inputs were randomly varied from run to run. Three types of inputs were "randomised": some of the parameters that were held fixed in the initial run (concerning natural mortality, $SSB_{initial}$, and stock-recruitment steepness), the likelihood weights (within a range of ±25%), and some of the observations (the CPUE indices). After discarding unsatisfactory runs, the 10th, 50th, and 90th percentiles of selected model outputs were calculated. The

50th percentile (i.e., the median) was treated as the best estimate and the other two percentiles were treated as indicating an approximate range of uncertainty.

The snowy grouper assessment suggested that fishing mortality first exceeded F_{msy} (the fishing mortality that will produce the maximum sustainable yield) in the mid 1970s and has fluctuated around $3F_{msy}$ since the early 1980s. This high fishing mortality rate caused the population biomass to decrease below SSB_{msy} in the early 1980s and it has continued to decline ever since. A feature of this assessment was that the MCB runs fell into two quite distinct groups: 1) a realistic group (1470 outcomes) in which population biomass was on the order of a few thousand tonnes and recent fishing mortalities were about $3F_{msy}$ and 2) an unrealistic group (846 outcomes) with very high population biomasses and very low fishing mortality. The latter group was discarded.

The tilefish assessment suggested that fishing mortality first exceeded F_{msy} in the early 1980s and has remained there since. This high fishing mortality rate caused the population biomass to decrease to near MSY levels in the mid 1980s, where it has remained ever since. Fishing mortality in recent years has exceeded F_{msy}, but the population has been maintained at B_{msy} because of above average recruitment. Only two of the 1100 MCB runs were deemed unsatisfactory and this was simply because the model failed to converge.

Some of the additional analyses requested by the Panel, and the results from these, are described briefly in the rest of this section.

2.2 Additional analyses for snowy grouper

A comparison of total observed and predicted landings showed that the latter exceeded the former by 14% overall, and by more (I calculated 33%) in the period before 1990. This degree of under-estimation of landings was not considered implausible.

A plot of estimated SSB by sex showed that males were estimated to be much more depleted than the females and that the current population was strongly dominated by females.

Attempts were made to understand the bipartite nature of the MCB runs by seeking combinations of the random components which would typically produce either satisfactory or unsatisfactory runs. No such explanation was found.

A run in which the model was forced to fit the landings almost exactly produced an output like one of the unrealistic MCB runs. The results of two ASPM runs depended on what was done with $SSB_{initial}$. When it was constrained to be equal to $0.9SSB_{virgin}$ (as in the initial run) the result was like one of the unrealistic MCB runs. When it was unconstrained it was estimated to be very low (about $0.2SSB_{virgin}$) but the estimated exploitation rates were more realistic (like those in the initial run). These runs were interpreted as showing that the population decline implied by the length composition data was clearly greater than could have been caused by the observed landings in the early years, so these landings must have been substantially under-estimated.

A model run in which natural mortality was independent of age produced results similar to those from the initial run. This suggests that although natural mortality undoubtedly varies with age it may not be important to model it thus.

2.3 Additional analyses for tilefish

An ASPM run produced a biomass trajectory that was similar to the initial run except that it led to lower biomass in the last 5-10 years. This showed the importance, in the initial run, of the positive recruitment residuals in the last few years. This is why, in the initial run, SSB is near SSB_{msy} in the final years although the fishing mortality exceeds F_{msy}.

3. FINDINGS

I was impressed by these assessments and the way they were presented to the Panel. They were mostly well documented. I particularly appreciated the inclusion of model equations and source code. Verbal descriptions of such complicated analyses are inevitably imprecise so it is good to be able to turn to the equations, or source code, for clarification of details. Presentations to the Panel were always clear and the assessment team was unfailingly helpful in response to requests for clarification or further analyses.

In the remainder of this section I first present my findings in relationship to each of the first four tasks of the Review Panel (as stated in Appendix 1) and then make some suggestions for future assessments.

3.1 Task 1: The data

The data used for both species were scientifically sound and appropriate for use in stock assessments (with minor exceptions), adequate to make useful inferences about stock status, and the best available for this purpose.

The exceptions concern the AFs for tilefish (and possibly snowy grouper) and one snowy grouper CPUE index. These exceptions are minor because these data sets had very little influence on the outcomes of the assessments. The problem with the tilefish AFs was that they appeared not to be representative of the catch (the LFs from the longline fishery seemed different from those for the fish from which otoliths were collected, though no formal statistical comparison was made). This is a common problem when AFs are estimated directly (rather than via an age-length key). It arises because such AFs are often constructed from many small samples and it is difficult to select a small sample that is random (and thus representative). Given the small sample sizes for the snowy grouper AFs it is quite possible that these data sets were not representative either. The existence of a zero in the 1992 MARMAP chevron trap CPUE index for snowy grouper was problematic because the lognormal error structure assumed for CPUE indices does not allow zeroes. This zero appears to have led to the estimation of a completely implausible selectivity for the MARMAP traps (Figure 43A) and thus a consistently poor fit to the associated AFs (Figure 39). The trap CPUE time series may not be an appropriate index for the assessment because its substantial oscillations (Figure 13) suggest that it may be indexing fluctuations in the availability of fish to this gear rather than changes in abundance.

Although the data were adequate to allow some useful inferences about stock status it should be stressed that these inferences are not strong, and there are substantial weaknesses in the data. For both stocks the assessments depended primarily on the LFs. But the stock assessment models could use the LFs only by assuming that the relationship between age and length was well known and that the associated selectivities had not changed over time (except for headboat LFs for snowy grouper). Age estimates for both species are unvalidated and uncertain, and any long-term changes in selectivities (which could be caused either by changes in gear or changes in the times and places that fishers choose to fish) would be likely to bias model estimates. The CPUE indices were not influential, but they would not be expected to be unless they covered long periods during which abundance had changed substantially. Another data weakness concerned the landings, which were known to be unreliable in the early years.

3.2 Tasks 2 and 4: Estimation of population parameters and benchmarks

I evaluated the estimation techniques used in these assessments in the knowledge that this was the first time that either stock had been fully assessed. It is not a simple task to assess a stock for the first time with models as complex as those used here. Every stock presents a different suite of problems to detect, consider, and solve. The first time it is assessed we can expect most major problems to be addressed and the general form of the model to be set. However it is normal that there should be other problems that are merely identified and, for lack of time or information, set aside to be dealt with in subsequent assessments. In this context I believe that the methods used to estimate population parameters and benchmarks in these assessments (weighted maximum likelihood and MCB runs of an age-structured population model) were adequate, appropriate, scientifically sound, and the best available.

I agree with the assessment team's conclusion that simple surplus-production models were not useful for these stocks.

3.3 Task 3: Best population parameters

With one reservation, I agree with the assessment team's decision that the best estimates of population parameters from these assessments are the median values from the MCB runs.

My reservation concerns the setting of the initial likelihood weights. The values assigned to these weights can have a profound effect on the estimated stock status so it is important that the rationale for this assignment be well documented. I did not feel that this was done in sufficient detail to allow me to judge whether or not I agreed with the chosen weights. Thus my acceptance of the conclusions of the assessments must be contingent on the assumption that I would find the weights acceptable. I should add that I have no grounds to doubt this assumption – it's just that I feel I had insufficient information to test it. Many preliminary model runs were done in setting the weights and I am not suggesting that all should have been presented. That would have swamped the Panel and not helped much. What I think *was* possible (and desirable) was that a narrative be constructed that described a sequence of decisions, with supporting reasons, leading to the accepted weights.

3.4 Suggestions for future assessments

The comments and suggestions given below are intended for the consideration of those charged with future assessments of these stocks. They should not be taken as criticisms of the current assessments. As I have said above, it is not a simple task to assess a stock for the first time, and we should not expect that all problems will be solved in the first assessment. I know that the assessment team had already identified some of the issues raised below and had flagged them for future consideration.

I first discuss issues common to both assessments and then those that were specific to just one.

3.4.1 Length and Age Frequencies

The acceptance criteria for LFs and AFs could be improved. Each LF or AF was accepted if its sample size exceeded a threshold (usually 25, sometimes 50). This doesn't make sense. A strength of maximum-likelihood estimation is that it automatically compensates for the loss of information as sample size decreases, so there is no theoretical lower limit on sample size. Acceptance criteria should be based on whether each LF or AF is representative of the catch. My suggestion is that an LF or AF should be acceptable only if it provides sufficient information to calculate an effective sample size.

How can we calculate an effective sample size for an LF (say)? By a simulation exercise in which the data are repeatedly resampled (bootstrapped) to generate a set of simulated LFs from which we can calculate the standard error (SE) of each proportion in the observed LF. The effective sample size, N_{eff}, is the number which minimises the difference between the bootstrap SEs and the theoretical values given by $[p(1\text{-}p)/N_{\mathrm{eff}}]^{0.5}$. So one requirement for acceptance is that there must be a non-trivial sample collected from each stratum of the catch. The strata must, of course, be constructed before sampling. How we might define 'non-trivial' depends on the sample structure and the nature of variability within a stratum, but one idea would be to require a minimum number of landings per stratum. The other requirement is for randomness at each stage of sampling (e.g., landings to be sampled selected at random, fish to be selected at random from the landing). Of course some judgement is necessary in deciding what is sufficiently random because the logistics of fisheries sampling usually preclude full formal randomness. However, when otoliths are taken from a subset of fish measured for an LF it is easy (and desirable) to test whether this has been done randomly by comparing lengths of otolithed fish with those in the LF, as was done for tilefish.

An example of a formal test for randomness in this context is given in Appendix 3 of Francis (2002).

It may be worth considering using length-mediated estimation for AFs (i.e, using age-length keys rather than direct estimation). Direct estimation of AFs (as used in these assessments) is very difficult because it usually requires many small samples, and the smaller the sample the harder it is to make sure it is randomly selected. The point is that age-length keys don't require that otoliths selection be random (as long as it is random within each length class, which is much easier to acheive). During the Workshop it was suggested that the degree of overlap between the length distributions of adjacent age classes for snowy grouper and tilefish precludes the use of age-length keys. I don't think this is true. However, it requires only a simple simulation experiment to determine which method produces, for a given sampling cost, the more precise AFs. Of course, it is not worth considering direct estimation of AFs unless random selection of otoliths can be assured.

3.4.2 Landings as observations

I think the way landings were modelled in these assessments could be improved. Each year's landing from a fishery was treated as an independent unbiased observation with a lognormal error distribution and a specified CV (coefficient of variation). However, the discussion of sources of error in these landings suggested to me that the primary concern was with bias. The likely direction and extent of bias was not known but it seemed probable, given its source, that it would be similar in groups of adjacent years. Thus a better model would be to divide the landings into blocks of adjacent years and assume constant bias within each block: say $L_{ij,\mathrm{obs}} = b_j L_{ij,\mathrm{true}} + e_{ij}$, where L_{ij} denotes the landing (observed or true) from the ith year in the jth block, b_j is a multiplicative bias, and the e_{ij} are the random error components. In principle we can, with sufficient information, estimate both the bias and the random error. However, I suggest that, given the data available for these assessments, we have virtually no ability to estimate the random components. Thus, a better approach would be to ignore the random components (assuming they will cancel each other out) and set $L_{ij,\mathrm{obs}} = b_j L_{ij,\mathrm{true}}$. This would substantially reduce the number of parameters to be estimated (by perhaps 137 for snowy grouper and 91 for tilefish, assuming two blocks of years) and avoid misleading the model with erroneous assumptions (the independence of errors in the assessments where the assumed model was $L_{i,\mathrm{obs}} = L_{i,\mathrm{true}} + e_i$). It also avoids the need to fabricate arbitrary CVs for the landings. Note that in the snowy grouper assessment there is strong autocorrelation in the landings residuals (Figure 31), which supports the above model of bias in blocks.

3.4.3 Length-based selectivities

It is generally believed that selectivity is much more a function of length than of age. Therefore, I think it would be better to estimate selectivities as functions of length, rather than of age. This requires the model to convert each length-based selectivity to an age-based one (using the estimated distribution of length at each age), which has two advantages. First, it avoids age-based selectivities that are implausibly steep, which was the case for almost all of those estimated in the present assessments (it is not possible for the selectivity to change greatly from one age to the next when there is a great deal of overlap in the length distributions for adjacent ages). Second, when growth differs between males and females it provides a more realistic way of modelling selectivity.

3.4.4 The desirability of being more statistical

Statistical models, like those used here, provide a powerful tool for dealing with uncertainty. They allow us to assign appropriate weights to different sources of information and they tell us how certain we can be about our inferences. In practice it is impossible to gain the full power of these models because we are unable to correctly specify all the statistical components of the model and so are often forced to add arbitrary non-statistical components. I suggest that our aim should be to minimise these non-statistical components, and in this section I suggest some ways in which I think this might be achieved for snowy grouper and tilefish.

The first thing is to avoid, as much as possible, non-statistical terms in the objective function. For example, if we treat the recruitment deviations as being lognormally distributed then the arbitrary (non-statistical) weight applied to the sum of squares of log recruitment deviations (to avoid extreme variation) is effectively an inverse variance. So why not specify it as such? There are quite a lot of published estimates of σ_R (the standard deviation of log recruitment) that can be used to provide a reasonable default value (e.g., Beddington and Cooke (1983), Myers et al (draft)). Also, given a value of σ_R, a simple simulation exercise (such as was done by Chris Legault during the Workshop) can be used to determine how much SSB can be expected to vary from year to year in an unfished population. This would allow the non-statistical constraint that was applied to SSB(initial) to be recast as a (statistical) prior distribution.

To deal with the likelihood components associated with the observations we need to discuss the nature of error. An approach that I have found useful is to write $(X_{\mathrm{obs}}\text{-}X_{\mathrm{pred}}) = (X_{\mathrm{obs}}\text{-}X_{\mathrm{true}}) + (X_{\mathrm{true}}\text{-}X_{\mathrm{pred}})$, where X_{obs} is our observation of some quantity, X_{true} is the true value of the quantity, and X_{pred} is the model's prediction of it. Thus the total error $(X_{\mathrm{obs}}\text{-}X_{\mathrm{pred}})$, which is modelled in our likelihood, is the sum of an observation error $(X_{\mathrm{obs}}\text{-}X_{\mathrm{true}})$ and what I call a process error $(X_{\mathrm{true}}\text{-}X_{\mathrm{pred}})$, this last being caused by all the simplifying assumptions (e.g., time-invariant selectivities and natural mortality) that we are forced to make in formulating our model. We can often estimate observation error outside the stock-assessment model (e.g., the CVs calculated for the CPUE indices measure observation error, as do the above-mentioned bootstrap-derived effective sample sizes for AFs and LFs). Process error is much more difficult, but becomes a bit easier if we assume, as seems reasonable, that all observations of the same type have the same sized process error. Thus, since CVs add as squares, we might say that $c_{ij,\mathrm{total}}^{2} = c_{\mathrm{process}}^{2} + c_{ij,\mathrm{observation}}^{2}$, where c_{ij} denotes a CV of the ith observation in the jth series of CPUE indices and c_{process} is the common process-error CV. This allows us to use a statistically interpretable quantity like c_{process} rather than a non-statistical likelihood weight. Of course it's still not easy to find an appropriate value for c_{process} (one approach that I've used for trawl surveys and CPUE is given in Francis et al 2003). Things don't work so easily with multinomial distributions (such as are use for LFs and AFs) but a pragmatic solution is to assume that $N_{\mathrm{total}}^{-1} = N_{\mathrm{observation}}^{-1} + N_{\mathrm{process}}^{-1}$.

Although there are still difficulties in deciding how large a process error term should be we do have an objective measure of how well we have done: by comparing the size of the residuals with that which is expected from the likelihood function. For example, with a normal or lognormal error distribution we can calculate the standard deviation of the normalised residuals, which should be about 1. Much smaller (or larger) values indicate that the total error CV is too large (or too small).

What is needed to make the MCB analysis more statistical is to devise probability distributions that best describe the uncertainty in the parameters that are being randomised. I acknowledge that this appears a daunting task but point out that these distributions are analogous to Bayesian priors, and there is an extensive literature on the problem of eliciting prior distributions. The advantage of making the MCB analysis more statistical is that it would allow a probabilistic interpretation of the MCB outputs (e.g., we could say that we are 80% confident that an estimated quantity (like SSB or MSY) lies between the 10th and 90th percentile of the MCB estimates).

There are two other issues associated with the MCB runs. First, all the observations should be randomised, and not just the CPUE. Given that the present assessments appeared to be driven by the LFs, and not much affected by the CPUE indices, it is regrettable that it was only the latter that were randomised. Second, it made no sense to me to scale the CPUE CVs to a maximum of 0.3 in randomising these observations. Given any CV, *c*, we can construct a lognormal variate Y with mean 1 and CV = c by setting $\sigma^2 = \log(1+c^2)$ and $Y = \exp(\sigma Z - 0.5\sigma^2)$, where Z is a standard normal variate.

3.4.5 Age data

There is clearly a need for validation of the ageing of both species so that we can have more confidence in the AFs and the age-length conversion matrix. This matrix is very important in an assessment in which LFs are influential. Since it is sensitive to the assumption that is made about how the variance of length at age varies with age this assumption should be checked carefully. Replicate age estimates of the same otoliths (preferably by different readers) can be used to generate an age misclassification matrix (in which the *i*th row gives the likely distribution of estimated ages for a fish of true age *i*) which can be used to modify the likelihood components associated with LFs and AFs.

3.4.6 Other general matters

The MCB analyses are a good way to replace one type of sensitivity analysis whose aim is to quantify uncertainty. Another type of sensitivity analysis which could have been useful in the Workshop would have been to rerun the initial run several times, each time dropping one type of data, thus showing the extent to which the assessments depended on each data type.

There were several small problems in both assessments, mostly in the documentation. It should be made clear that the calculation of generation time involves only female fish (I understand that this was how the calculations were made, but that was not clear to me from the reports). In fitting the von Bertalanffy equation the assumption used was clearly that the *standard deviation* of length at age was proportional to the mean length (not the *variance*, as stated). In the formula for the age-length conversion matrix the superscript 2 is misplaced. Equations should be given for the per-recruit calculations. It might be worth checking the method of fitting the maturity ogives for both species because the fitted curve is to the right of all data points for which the proportion mature is not near 0 or 1 (see Figure 5 for snowy grouper and Figure 8 for tilefish). In the tables documenting the model it might avoid confusion if a clear distinction were made between fixed parameters (e.g., growth parameters, LF sample sizes), estimated parameters (e.g., selectivity parameters, fishing mortalities), derived quantities (e.g., length at age, selectivity at age) and observations (which are characterised by having an associated likelihood component, e.g., CPUE, LFs).

3.4.7 Snowy Grouper

I think it might be useful to try some more sophisticated techniques (e.g., GAMs or tree-based regression) to seek an explanation of the unrealistic MCB runs. This may be informative. It might be worth dropping the Chevron trap CPUE index (for reasons given above). It seems a matter of some concern that more than half the catch is of immature fish. It is worth considering explicitly modelling the three categories of fish: immature, mature female, mature male (i.e., keeping track of numbers of fish by age and category)

3.4.8 Tilefish

I think it would be worthwhile to explicitly model sex (i.e., to keep track of numbers by sex, as well as by age — the assessment report stated that this was not possible because the landings and LFs were not sex-specific, but I don’t see why). As females are smaller at age than males they probably do not have the same selectivity at age as males do, so modelling selectivity as length-based would be better.

4. CONCLUSIONS

I believe that the assessments of snowy grouper and tilefish that were presented to the Panel provide information that should be very useful to fishery managers. The assessment team did a good job of dealing with the available data and constructing sound first-time assessments.

5. REFERENCES

Beddington, J.R.; Cooke, J.G. (1983). The potential yield of fish stocks. *FAO Fisheries Technical Paper 242*. 50 p.

Francis, R.I.C.C. (2002). Estimating catch at age in the Chatham Rise hoki fishery. *New Zealand Fisheries Assessment Report 2002/9*. 22 p.

Francis, R.I.C.C.; Hurst, R.J.; Renwick, J.A. (2003). Quantifying annual variation in catchability for commercial and research fishing. *Fishery Bulletin 101*: 293-304.

Lorenzen, K. (1996). The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. *Journal of Fish Biology 49*: 627-647.

Myers, R.A.; Bowen, K.G.; Zouros, I.A. (draft). Recruitment variability of fish and aquatic invertebrates. Draft paper available in pdf form from http://fish.dal.ca/~myers/papers.html.

APPENDIX 1: Statement of Work

This appendix contains the Statement of Task that formed part of the consulting agreement between the University of Miami and the author.

Consulting Agreement between the University of Miami and NIWA (Dr. Chris Francis)

General

South East Data, Assessment, and Review (SEDAR) is a joint process for stock assessment and review of the South Atlantic, Gulf of Mexico, and Caribbean Fishery Management Councils; NOAA Fisheries, SEFSC and SERO; and the Atlantic and Gulf States Marine Fisheries Commissions. SEDAR is organized around three workshops: data, assessment, and review. Input data are compiled during the data workshop, population models are developed during the assessment workshop, and an independent peer review of the data and assessment models is provided by the review workshop. The peer review panel is composed of stock assessment experts, other scientists, and representatives of council, fishing industries, and non-governmental conservation organizations. Final SEDAR documents include a stock assessment report produced by the data and assessment workshops, a review panel report evaluating the assessment (drafted during the review panel workshop), a review panel report that summarizes the peer-reviewed assessment results, and collected stock assessment documents considered in the SEDAR process.

NMFS-SEFSC requests the assistance of two assessment scientists from the CIE: one to serve as Chair and one to serve as a technical reviewer for the SEDAR 4 Review Panel that will consider assessments for two species from the South Atlantic deepwater snapper-grouper complex: tilefish and snowy grouper.

These species are within the jurisdiction of the South Atlantic Fishery Management Council and respective southeastern states. The review workshop for SEDAR 4, South Atlantic deepwater complex stock assessments, will take place at the Holiday Inn Center City, Charlotte, NC from July 26 (beginning at 2:00 pm) through July 30, 2004 (ending at 1:00 pm). Meeting materials will be forwarded electronically and in hard copy. Please contact John Carmichael (SEDAR Coordinator; 843-571-4366 or John.Carmichael@safmc.net) for additional details.

Hotel arrangements

Holiday Inn Center City, 230 N. College Street, Charlotte, NC 28202. Phone: (704) 335-5400, (800) 465-4329; Fax (704) 376-4921. Please make reservations by June 16 and to receive the ‘SEDAR Workshop’ group rate of $91.94 (including tax).

SEDAR Assessment Review Panel Tasks

The SEDAR Assessment Review Panel will evaluate the tilefish and snowy grouper stock assessments, input data, assessment methods, and model results as put forward in stock assessment reports. The Assessment Review Panel will:

1. Evaluate the adequacy and appropriateness of all data used in the assessment, and state whether or not the data are scientifically sound and the best available.
2. Evaluate the adequacy, appropriateness, and application of the methods used to estimate population parameters such as abundance, biomass, and exploitation; state whether or not the methods are scientifically sound and the best available;
3. Recommend appropriate or best estimated values of population parameters such as abundance, biomass, and exploitation.
4. Evaluate the adequacy, appropriateness, and application of the methods used to estimate population benchmarks (MSY, Fmsy, Bmsy, MSST, MFMT, etc.). State whether or not the methods are scientifically sound and the best available,
5. Recommend appropriate values for population benchmark criteria.
6. Evaluate the adequacy, appropriateness, and application of the methods used to project future population status and, if appropriate, evaluate stock rebuilding; state whether or not the methods are scientifically sound and the best available.
7. Recommend probable values of future population condition and status.
8. Develop recommendations for future research for improving data collection and the assessment.
9. Prepare a Peer Review Panel Consensus Summary summarizing the peer review panel's evaluation of the tilefish and snowy grouper stock assessments and addressing the Terms of Reference. (Drafted during the Assessment Review Panel workshop with a final report due three weeks after the workshop ends.)
10. Prepare a Stock Advisory Report summarizing the stock assessments. (Drafted during the Assessment Review Panel workshop with a final report due three weeks after the workshop ends.)

The Assessment Review Panel's primary duty is to review the assessments presented. In the course of this review, the Chair may request a reasonable number of sensitivity runs, additional details of the existing assessments, or similar items from technical staff. However, the review panel is not authorized to conduct an alternative assessment or to request an alternative assessment from the technical staff present. If the review panel finds that an assessment does not meet the standards outlined in Items 1 through 4, above, the panel will outline in its report the remedial measures that the panel proposes to rectify those shortcomings.

The Review Panel Report is a product of the overall Review Panel, and is NOT a CIE product. The CIE will not review or comment on the Panel's report, but shall be provided a courtesy copy, as described below under "Specific Tasks." The CIE product to be generated is the Chair's report, also discussed under Specific Tasks.

Specific Tasks

The CIE designee shall serve as review panelist of a SEDAR Stock Assessment Review Panel workshop for SEDAR 4, South Atlantic tilefish and snowy grouper, July 26 - 30, 2004 (See attached agenda.). The workshop panel shall review stock assessments for South Atlantic tilefish and snowy grouper in the jurisdiction of the South Atlantic Fishery Management Council and applicable southeastern states.

It is estimated that the review panelist's duties will occupy a maximum of 14 workdays; several days prior to the meeting for document review; five days at the SEDAR meeting, and several days following the meeting to ensure that final review comments on documents are provided to the Chair and to complete a CIE review report.

Roles and responsibilities:

1. Prior to the meeting the CIE reviewer shall be provided with the stock assessment reports and associated documents for South Atlantic tilefish and snowy grouper. The reviewer shall read these documents to gain an in-depth understanding of the stock assessment and the resources and information considered in the assessment.
2. During the Review Panel meeting, the reviewer shall participate, as a peer, in panel discussions on assessment validity, results, recommendations, and conclusions. The reviewer also shall participate in the development of the Peer Review Panel Consensus Summary and Stock Advisory Report;
3. Following the Review Panel meeting, the reviewer shall review and provide comments to the Panel Chair on the Peer Review Panel Consensus Summary and Stock Advisory Report.
4. No later than August 20, 2004, the reviewer shall submit a written CIE review report[1] consisting of the findings, analysis, and conclusions, addressed to the "University of Miami Independent System for Peer Review," and sent to Dr. David Sampson, via email to David.Sampson@oregonstate.edu, and to Mr. Manoj Shivlani, via email to mshivlani@rsmas.miami.edu. The report shall address points 1-4 under the above heading: SEDAR Assessment Review Panel Tasks. See Annex I for details on the report outline.

Workshop Final Reports

The Chair shall send final review workshop reports to the University of Miami Independent System for Peer Review, Dr. David Die via email to ddie@rsmas.miami.edu.

Final workshop reports (in Word or WordPerfect format and in hardcopy) shall also be sent to:

Nancy Thompson, NMFS Southeast Fisheries Science Center, 75 Virginia Beach Drive, Miami, FL 33149 (email, Nancy.Thompson@NOAA.gov)

Larry Massey, 101 Nina Drive #302, Virginia Beach, VA 23462 (email, Larry.Massey@NOAA.gov)

John Carmichael, SAFMC, One Southpark Circle, Suite 306, Charleston, SC 29407 (email, John.Carmichael@safmc.net)

Robert Mahood, South Atlantic Fishery Management Council, One Southpark Circle, Suite 306, Charleston, SC 29407 (email, Robert.Mahood@safmc.net)

[1] The written Reviewer report will undergo an internal CIE review before it is considered final. After completion, the CIE will create a PDF version of the Reviewer report that will be submitted to NMFS and the consultant.

For Additional Information or Emergency:

SEDAR contact: John Carmichael, One Southpark Circle, Suite 306, Charleston, SC 29407. Phone: 843-571-4366; cell phone (843) 224-4559. Email: John.Carmichael@safmc.net.

ANNEX I: Contents of CIE Reviewer Report

1. The reviewer report shall be prefaced with an executive summary of findings and/or recommendations.

2. The main body of the reviewer report shall consist of a background, description of review activities, summary of findings, and conclusions/recommendations.

3. The reviewer report shall also include as separate appendices the bibliography of materials provided by the Center of Independent Experts and a copy of the Statement of Work.

APPENDIX 2: Materials Provided

The author was provided with the following materials by the SEDAR Coordinator.

1. Terms of Reference and Panel Instructions for SEDAR 4 Review Workshop, Atlantic Deepwater Snapper-Grouper: Tilefish and Snowy Grouper (see Appendix 2.1)

2. SEDAR 4 Data Workshop Summary, Deep Water Complex, November 3-7, 2003

3. Assessment of Snowy Grouper (*Epinephelus niveatus*) in the South Atlantic Fishery Management Council Management Area. Section III.A of SEDAR Stock Assessment Report.

4. Assessment of Tilefish, *Lopholatilus chamaeleonticeps,* in the South Atlantic Fishery Management Council Management Area. Section III.B of SEDAR Stock Assessment Report.

5. South Atlantic Deepwater Snapper Grouper Document List. Appendix A of SEDAR Stock Assessment Report

6. AD Model Builder code for tilefish statistical catch-at-age model. Appendix B of SEDAR Stock Assessment Report

7. Documents from SEDAR4 Atlantic and Caribbean Deepwater Snapper Grouper (listed in Appendix 2.2).

8. Reference papers from SEDAR4 Atlantic and Caribbean Deepwater Snapper Grouper (listed in Appendix 2.3).

9. A CD containing items 1-4, 7 and 8 above.

APPENDIX 2.1: Terms of Reference and Instructions for the Review Panel

I. Terms of Reference

The SEDAR Assessment Review Panel will evaluate the tilefish and snowy grouper stock assessments, input data, assessment methods, and model results as put forward in stock assessment reports. The Assessment Review Panel will:

1. Evaluate the adequacy and appropriateness of all data used in the assessment and state whether or not the data are scientifically sound;

2. Evaluate the adequacy, appropriateness, and application of the methods used to estimate population parameters such as abundance, biomass, and exploitation and state whether or not the methods are scientifically sound;

3. Evaluate the adequacy, appropriateness, and application of the methods used to estimate population benchmarks (*e.g., MSY, Fmsy, Bmsy, MSST, MFMT, or their proxies*) and state whether or not the methods are scientifically sound;

4. Evaluate the adequacy, appropriateness, and application of the methods used to project future population status and, if appropriate, evaluate stock rebuilding; state whether or not the methods are scientifically sound;

5. Ensure that all available required assessment results (*as listed in the SEDAR Stock Assessment Report Outline*) are clearly and accurately presented in the Stock Assessment Report and that such results are consistent with the Panel's decisions regarding adequacy, appropriateness, and application of the data and methods;

6. Evaluate the performance of the Data and Assessment Workshops with regards to their respective Terms of Reference, and state whether or not the Terms of Reference for those previous workshops are adequately addressed in the Stock Assessment Report;

7. Review the assessment workshop's recommendations of future research for improving data collection and the assessment, and make any additional recommendations warranted;

8. Prepare a Peer Review Consensus Summary summarizing the Panel's evaluation of the tilefish and snowy grouper stock assessments and addressing each Term of Reference. (Drafted by the Panel during the Review Workshop with a final report due three weeks after the workshop ends.);

9. Prepare a Stock Advisory Report summarizing the stock assessments. (Drafted during the Assessment Review Panel workshop with a final report due three weeks after the workshop ends.).

II. Review Panel Instructions

The Assessment Review Panel is charged with reviewing the technical aspects of the presented stock assessment and making judgements regarding the assessment that are based solely upon scientific merit. At no point during the deliberations should the Review Panel consider the implications that the assessment and its results may have upon management decisions or resource users. This is not to imply in any way that such considerations are not important, but rather to acknowledge several important facts: (1) consideration of management impacts is beyond the scope of the charge to the Review Panel, (2) SEDAR specifically strives to separate management considerations from assessment decisions, (3) Review Panel participants are selected based on technical, biological, and assessment knowledge, not social and economic knowledge of a fishery, (4) consideration of social and economic consequences is specifically mandated to the Council and various Council Committees composed of experts qualified to evaluate the social and economic consequences of management actions.

The Assessment Review Panel is discouraged from holding formal votes. Decisions should be based upon the unanamious consenus of the entire panel. In the event that the Chair feels that all avenues for agreement have been exhausted and unanimous consensus is not achievable, the Chair may instruct that the majority opinion be reflected in the report and allow the minority opinon holders to prepare and submit a minority report.

The Assessment Review Panel's primary duty is to review the assessments presented. In the course of this review, the Chair may request a reasonable number of sensitivity runs, additional details of the existing assessments, or similar items from technical staff. However, the review panel is not authorized to conduct an alternative assessment nor to request an alternative assessment from the technical staff present.

If the review panel finds that an assessment does not meet the standards outlined in Items 1 through 6, above, the panel will outline in its report the remedial measures to be taken by the assessment analysts to rectify those shortcomings.

Review Panel members are expected to participate in the entire workshop from start to finish. The supporting Council's strongly discourage panel members from leaving early. Panelists should expect that the Workshop will require the entire time alloted and plan travel accordingly. To this end, workshops are scheduled for an afternoon start and early adjournment to reduce the need for weekend travel.

APPENDIX 2.2: Documents from SEDAR4 Data Workshop

#	Title	Author(s)
SEDAR4-DW-01	Indices of Abundance from Commercial Logbook Data: South Atlantic stocks	Shertzer, K.; McCarthy, K.
SEDAR4-DW-02	MRFSS Landings and Length Data Summary for the South Atlantic	Vaughan, D. S.
SEDAR4-DW-03	General Canvass Landings Statistics for the South Atlantic Region	Poffenberger, J.
SEDAR4-DW-04	Summary information on commercial fishing operations in Puerto Rico from 1969-2001 and reporting rates needed to adjust commercial landings.	Cummings, N. Matos-Caraballo, D.
SEDAR4-DW-05	Summarized reported commercial landings in Puerto Rico from 1969-2001 with specific notes on the silk snapper landing category.	Cummings, N. Matos-Caraballo, D.
SEDAR4-DW-06	Not used	
SEDAR4-DW-07	Information on the general biology of silk and queen snapper in the Caribbean.	Cummings, N
SEDAR4-DW-08	Preliminary Estimation of Reported Landings, Expansion Factors and Expanded Landings for the Commercial Fisheries of the United States Virgin Islands**.**	Valle-Esquivel, M. Diaz, G.A.
SEDAR4-DW-09	Preliminary species composition estimates of TIP samples from commercial landings in the U.S. Virgin Islands.	Diaz, G. A. ; Valle-Esquivel, M.
SEDAR4-DW-10	Standardized Catch Rates of Silk Snapper, Lutjanus vivanus, from the St. Croix .S.Virgin Islands Handline Fishery during 1984 - 1997.	Cass-Calay, S.L.; Valle-Esquivel, M.
SEDAR4-DW-11	Standardized Catch Rates of Queen Snapper, Etelis oculatus, from the St. Croix U.S. irgin Islands Handline Fishery during 1984 – 1997	Cass-Calay, S.L.; Valle-Esquivel, M.
SEDAR4-DW-12	Discard Estimates for the South Atlantic Region.	Poffenberger, J.
SEDAR4-DW-13	Size Frequency Data from the Trip Interview Program, South Atlantic Region	Poffenberger, J.
SEDAR4-DW-14	Size frequency distributions of silk snapper and queen snapper from dockside sampling of commercial landings in the U.S. VI	Diaz, G. A.; Valle-Esquivel, M.
SEDAR4-DW-15	Preliminary information on the recreational catch of silk, queen, and blackfin snapper, from 2000 through 2002 in Puerto Rico with additional notes on sand tilefish	Cummings, N.; Slater, B.; Turner, S.
SEDAR4-DW-16	Preliminary analysis of some deepwater species in the South Atlantic headboat survey data.	Williams, E.; Dixon, B.
SEDAR4-DW-17	Age, growth and reproductive biology of the blueline tilefish, *Caulolatilus microps*, along the southeastern coast of the United States, 1982-99.	Harris, P. J.; Wyanski, D.M.; Powers, P.T.
SEDAR4-DW-18	Age, growth and reproduction of tilefish, Lopholatilus chamaeleonticeps, along the southeast Atlantic coast of the United States,	Palmer, S.M.; Harris, P.J.; Powers, P. T.

	1980-87 and 1996-98.	
SEDAR4-DW-19	Deep-water species report. South Carolina and Georgia.	Low, B.
SEDAR4-DW-20	South Atlantic Snapper-Grouper Regulatory Overview	Carmichael, J.
SEDAR4-DW-21	Summary of MARMAP sampling	Anon.
SEDAR4-DW-22	Blueline tilefish life history; How to assess reef fish stocks: Excerpts from NMFS-SEFC-80	various
SEDAR4-DW-23	Preliminary size frequency information for silk, queen, and blackfin snapper from the Puerto Rico commercial fisheries from 1985 through 2002 with additional notes on sand tilefish	Cummings, N.J. Phares, P
SEDAR4-DW-24	Brief summary of SEAMAP data collected in the Caribbean Sea from 1975 to 2002	Ingram, W.
SEDAR4-DW-25	Yellowedge Grouper age-length key	Bullock & Godcharles
SEDAR4-DW-26	Estimating catches and fishing effort of the southeast united states headboat fleet, 1972-1982	Dixon & Huntsman
SEDAR4-DW-27	Trends in Catch Data and Estimated Static SPR Values for Fifteen Species of Reef Fish Landed along the Southeastern United States, February 1998.	Potts, Burton & Manooch
SEDAR4-DW-28	Trends in Catch Data and Estimated Static SPR Values for Fifteen Species of Reef Fish Landed along the Southeastern United States, February 2001.	Potts & Brennan
SEDAR4-DW-29	Description of the Southeast Fisheries Science Center's Logbook Program for Coastal Fisheries	Poffenberger, J.

APPENDIX 2.3: References from the SEDAR4 Data Workshop

Bohnsack, J. A. and A. Woodhead. 1995. Proceedings of the 1987 SEAMAP passive gear assessment workshop at Mayaguez, Puerto Rico. NOAA Tech. Mem. NMFS SEFSC 365.

Bullis, H. R. Jr. and A. C. Jones, *ed.* 1976. Proceedings: Colloquium on snapper-grouper fishery resources of the Western Central Atlantic Ocean. FL SeaGrant Report No. 17.

Bullock, L. H., M. F. Godcharles, and R. E. Crabtree. 1996. Reproduction of yellowedge grouper, *Epinephelus flavolimbatus*, from the Eastern Gulf of Mexico. Bull. Mar. Sci. 59(1) 224-228.

Erickson, D. E. and G. D. Grossman. 1986. Reproductive demography of tilefish from the South Atlantic Bight with a test for the presence of protogynous hermaphroditism. Trans. Am. Fish. Soc. 115:279-285.

Grimes, C. B. and S. C. Turner. 1999. The complex life history of tilefish *Lopholatilus chamaeleonticeps* and vulnerabilty to exploitation. Am. Fish. Soc. Symp. 23:17-26.

Grossman, G. D., M. J. Harris, and J. E. Hightower. 1985. The relationship between tilefish, *Lopholatilus chamaeleonticeps*, abundance and sediment composition off Georgia. Fish. Bull. 83(3):443-447.

Harris, M. J. and G. D. Grossman. 1985. Growth, mortality, and age composition of a lightly exploited tilefish substock off Georgia. Trans. Am. Fish. Soc. 114:837-846.

Hightower, J. E., and G. D. Grossman. 1989. Status of the tilefish, *Lopholatilus chamaeleonticeps,* fishery off South Carolina and Georgia and recommendations for management. Fish. Bull. 87:177-188.

Huntsman. G.R., Nicholson, W.R., Fox, W.W.Jr. 1982. The biological bases for reef fishery management: proceedings of a workshop help October 7-10 1980 at St. Thomas, Virgin Islands of the United States. NOAA Technical Memorandum NMFS-80.

Low, B., G. Ulrich, and F. Blum. 1982. The fishery for tilefish, *Lopholatilus chamealeonticeps*, off South Carolina and Georgia. SC Wildl. and Mar. Res. Div. Charleston, SC.

Low, R. A. Jr., G. F. Ulrich, and F. Blum. 1983. Tilefish off South Carolina and Georgia. Mar. Fish. Rev. 45(4-6)16-26.

Manooch, C. S., and D. L. Mason. 1987. Age and growth of the Warsaw grouper and black grouper from the Southeast region of the United States. Northeast Gulf Sci. 9(2):65-75.

Matheson, R. H. and G. R. Huntsman. 1984. Growth, mortality, and yield-per-recruit models for speckled hind and snowy grouper from the United States South Atlantic Bight. Trans. Am. Fish. Soc. 115:607-616.

Parker, R. O. Jr. and R. W. Mays. 1998. Southeastern U. S. deepwater reef fish assemblages, habitat characteristics, catches, and life history summaries. NOAA Tech. Report. NMFS-138.

Ross, J. L. 1982. Feeding habits of the gray tilefish *Caulolatilus microps* (Goode and Bean, 1878) from North Carolina and South Carolina waters. Bull. Mar. Sci. 32(2):448-454.

Ross, J. L. and G. R. Huntsman. 1982. Age, growth, and mortality of blueline tilefish from North Carolina and South Carolina. Trans. Am. Fish. Soc. 111:585-592.

Russel, G. M., E. J. Gutherz, and C. A. Barans. 1988. Evaluation of demersal longline gear off South Carolina and Puerto Rico with emphasis on deep-water reef fish stocks. Mar. Fish. Rev. 50(1):26-31.

Tester, P. A., C. A. Wolfe, R. L. Dixon, and G. R. Huntsman. 1983. Reef fish distributions off North Carolina and South Carolina as revealed by headboat catches. NOAA Tech. Mem. NMFS-SEFC-115.

Wyanski, D. M., D. B. White, and C. A. Barans. 2000. Growth, population age structure, and aspects of the reproductive biology of snowy grouper, *Ephinephelus niveatus*, off North Carolina and South Carolina. Fish. Bull. 98:199-218.

APPENDIX 3: Attendees at SEDAR4 Assessment Review Panel Workshop

CIE Participants

Robert Mohn	Bedford Institute of Oceanography, P.O. Box 1006, Dartmouth, N.S., CANADA B2Y 4A2 Phone: 902-426-4592. Email: mohnr@mar.dfo-mpo.gc.ca
Chris Francis	National Institute of Water and Atmospheric Research Ltd P.O. Box 14-901, Kilbirnie, Wellington, NEW ZEALAND Phone: +64-4-386 0300, Fax: +64-4-386 0574 Email: c.francis@niwa.cri.nz

Panel Members

Chris Legault	Northeast Fisheries Science Center 166 Water Street, Woods Hole, MA 02543-1026 Phone: 508-495-2025, Fax: 508-495-2258 Email: chris.legault@noaa.gov
Scott Nichols	Pascagoula Laboratory, PO Drawer 1207 Pascagoula MS 39568-1207 Phone: 228-762-4591 ext. 269, Fax: 228-769-9200 Email: scott.nichols@noaa.gov
Robert Muller	FL Fish and Wildlife Conservation Commission Fish and Wildlife Research Institute, 100 Eighth Avenue SE St Petersburg, FL 33701 Phone: 727-896-8626 ext. 4118, Fax: 727-893-1374 Email: robert.muller@fwc.state.fl.us
Doug Rader	Environmental Defense, 2500 Blue Ridge Road, Suite 330 Raleigh, NC 27607 Phone: 919-881-2601, Fax: 919-881-2607 Email: drader@environmentaldefense.org

SEDAR Coordinator

John Carmichael	SEDAR, 1 South Park Circle, Suite 306 Charleston, SC 29407 Phone: 843-571-4366, Fax: 843-769-4520 Email: john.carmichael@safmc.net

Presenters

Mike Prager	Beaufort NOAA Lab, Mike.Prager@noaa.gov
Doug Vaughn	Beaufort NOAA Lab, Doug.Vaughan@noaa.gov
Kyle Shertzer	Beaufort NOAA Lab, Kyle.Shertzer@noaa.gov
Erik Williams	Beaufort NOAA Lab, Erik.Williams@noaa.gov

Other participants

Louis Daniel	Vice-Chair, South Atlantic Fishery Mgmt. Council
John Merriner	SouthEast Fishery Science Center
Julie Weeder	South East Regional Office, NMFS
Larry Massey	SouthEast Fishery Science Center
Gerard Dinardo	Pacific Islands Fisheries Center, NMFS, Hawaii