SEDAR 27 South Atlantic and Gulf of Mexico Yellowtail Snapper Assessment Center for Independent Experts (CIE) Review Report

Paul Medley

paulahmedley@yahoo.co.uk

July 2012

Executive Summary

The yellowtail snapper assessment from the South Atlantic and Gulf of Mexico has been carried out by the National Marine Fisheries Service (NMFS) and the Florida Fish and Wildlife Conservation Commission (FFWCC) in 2012, and fully described in a series of reports (Appendix 1), which were the subject of this independent peer review. This report describes the findings, conclusions and recommendations of my review of these stock assessment reports as required in the Statement of Work (Appendix 2).

The data are sufficient for carrying out a stock assessment. While sampling has not been sufficient in all years and fleets, data collection and sampling in recent years appear to have improved.

Given the available data, the stock assessment method is sound and has been applied correctly. Development of the software to handle more types of missing data and a better explanation on how effective sample sizes have been handled would be valuable improvements.

My overall conclusion based on the reports and available data is that the stock is not overfished or undergoing overfishing.

The stock assessment could be further improved in modeling and reporting uncertainties in the assessment. Projections, in particular, should incorporate probability estimates (MCMC). A greater range of sensitivities should be considered based on their likely impact on the determination of stock status and ACLs.

In terms of research and data collection to reduce uncertainty, as well as better data particularly for monitoring discards and discard survival, I have recommended some further research on the available abundance indices, and some additions and improvements to the ASAP software.

Background

The yellowtail snapper assessment from the South Atlantic and Gulf of Mexico has been carried out by the National Marine Fisheries Service (NMFS) and the Florida Fish and Wildlife Conservation Commission (FFWCC). The last CIE review that included a yellowtail snapper assessment was conducted in August 2003 and therefore a CIE review was requested of the 2012 assessment. The yellowtail snapper assessment was originally scheduled for the SEDAR 27 in November 2011, but the analysis was incomplete at that time and therefore this review has been undertaken separately.

Yellowtail snapper was assessed for SEDAR 3 using an age-structured assessment model (Integrated Catch-at-Age) for estimating stock status with data through 2001. The Stock Status Report for Yellowtail Snapper estimated for 2001 that F was 0.17 and SSB was 4,481 metric tons, that SSB_{2001}/SSB_{MSST} was 1.78 (not overfished) and F_{2001}/F_{OY} was 0.92 (not overfishing).

A new assessment was completed and provided to the CIE reviewers on 1st June 2012. This independent peer review report describes the findings, conclusions and recommendations of my review of this stock assessment as required in the Statement of Work (Appendix 2).

Reviewer's Role in the Review Activities

The review was conducted as desk study of the stock assessment documents, which were sent as electronic files (pdf format). The documents received are listed in Appendix 1. No meeting was held and therefore there was no opportunity to request clarification of some points or any additional runs of the model with alternative configurations. However, the stock assessment software (ASAPv2) is publicly available and the base case data file was also provided, so that it was possible to carry out some alternative runs to clarify some issues. This was particularly useful to verify my understanding of the model and its behavior and to test the sensitivity of the results to some assumptions. However, such tests were not exhaustive, and the review primarily addresses the stock assessment described in the main assessment document.

Summary of Findings

1. Evaluate the data used in the assessment, addressing the following:

a) Are data decisions made by the assessment panel sound and robust?

I found no decisions made by the assessment panel concerning use of data that I disagreed with. There is always a measure of judgment in the choices made, but the decisions were sound and justified adequately in this case.

The approach to stock structure appears reasonable. Although the population and therefore assessments are not split between the Gulf and the Atlantic, allowance is made through standardizing GLMs and calculations of catch-at-age where there are sufficient data. Most Gulf catches for the commercial fishery appear to be taken from the Florida Keys, which might suggest a close relationship between at least the fishable biomass of the two areas.

No data sets were rejected and small samples were included. Outliers were removed as might be expected from morphometric data and catch and effort data. This process appeared reasonable, and would certainly make the assessment conclusions safer, since they may remove unduly influential data points.

The life history parameters seem reasonably well supported. The method used to calculate the natural mortality-at-age, weight-at-age and maturity-at-age are appropriate.

Choice of initial effective sample size (ESS) weights for the age and length compositions was appropriate and reasonable. The ESS weights were based on a square root transform of the number of samples, which would decrease the relative weight on larger samples. It should be noted that there is no need to stabilize the variance, which would be expected to increase for strata containing larger samples, since the change in variance is taken into account in the likelihood. The reason for reducing the weights on age and length compositions is that the sample size is over-estimated because age and length are not sampled independently either within or between vessels. The effect of this on sample size would likely be proportional. As a result, it would be worth considering setting the ESS as proportional to the sample size rather than applying a square root transform.

The total catch by each fleet is reasonably well estimated. The landings which form the majority of the catch are reported for the commercial vessels or estimated from intercept surveys for MFRSS and Headboats. Discards were not well estimated and depended upon small, potentially biased data. However, discards only make up a small proportion of the catch, and their impact is further reduced because the available information suggests that discard mortality is low.

Selection of trips for CPUE abundance indices was objective and applied reasonable criteria. This is a difficult area in multispecies fisheries, and while I am not convinced the solution used here is perfect, it is objective and probably represents the best available method. I suspect that the approach might mask changes in catchability, further discussed under Term of Reference 2.

The NMFS-UM Reef Visual Census (RVC) is the only, and potentially a very useful, fishery independent abundance index for this species. The basis for selecting only the later period (1998-2010) from the entire time series (1979-2010), of the NMFS-UM RVC appears reasonable. The basis for this is the lower coefficient of variation (CV) values from improvements in sample design (increased level of stratification and the inclusion of the MPAs as a stratum). The higher CVs mean that even if earlier data would be included, they would have little influence. It remains unfortunate that the index time series is so short, since this will greatly lessen its impact on the stock assessment. The NMFS-UM RVC index was not linked to any fleet and so the selectivity pattern was derived from proportions at age data.

b) Are data uncertainties acknowledged, reported, and within normal or expected levels?

All data uncertainties appear to be acknowledged and reported. Certainly data since 1993 are within normal levels, but some data are absent or sampling is low during some periods 1981-1992, particularly 1981-1983. Information is also very limited on discards and discard mortality.

The reporting of landings and the sampling of retained catch from each of the fleets (commercial, recreational, and head boat) is derived from different reporting systems and involves different intercept and sampling strategies. These differences have been accounted for, as far as possible, and should have little influence on final results.

Length and age measurements are missing for 1981-1991 in the Atlantic region. Length measurements are missing for 1982-1983 and ages for 1982-1991 in the Gulf region. There was very little information about discarded fish and no information on the size composition of discards. There is also a dependence on self-reporting of commercial discards which may lead to under-reporting. Where possible, corrections have been applied to such biases.

It is worth noting that, to fully meet this Term of Reference, a reviewer would need to know the data collection system well and not just rely on the report. From previous reviews, I have some knowledge of the data collection systems used, but it is not profound. It is therefore possible that there are additional uncertainties in the data collection of which I am not aware.

c) Are data applied properly within the assessment model?

The data are prepared appropriately using standard methods and subsequently applied properly within the assessment model. Much of the data are processed through models before being used by the final assessment model (ASAP). Data processing includes standardizing CPUE indices and estimating means or totals based on GLMs (Discards, age-at-maturity, morphometric conversions) and age length keys (ALK). These data are specifically prepared to have the form and format necessary for the stock assessment model.

The method used to estimate natural mortality at age is reasonable and appropriate. The age-based form of the natural mortality is probably more accurate than constant mortality-at-age.

The CPUE indices were derived from generalized linear models (GLM). A subset of trips were selected in each case for the CPUE index using a classification of trips based on the catch species composition, which is a reasonable approach given the available data.

Standardizing CPUE indices using GLMs serves two purposes. Firstly, it reduces random noise and bias in the indices. It can be claimed that the GLMs used do this, since they account for variation among fishing locations, crew, season and so on. Given that fishing is not random, this is a reasonable approach. Secondly, GLMs may apply a correction for catchability which will often increase as fishing becomes more efficient. The GLMs do not contain all appropriate covariates to be sure that this purpose is achieved.

As well as standardizing CPUE, a delta-lognormal GLM was used to estimate the average discard per unit effort, which was raised to the total number of discards using the total effort. The average discard per unit effort for the 2002-2010 period was also applied to the 1993-2001 period (when discards were not reported). Using GLMs for this is reasonable as it should improve estimates by using the available covariates which are complete and influence discards. It should be noted that with low discard mortality rate, errors in these estimates should have a low impact on the overall stock assessment.

d) Are input data series reliable and sufficient to support the assessment approach and findings?

The data are certainly sufficient to support the assessment approach from 1993-2010. While there is always room for improvement, sampling has covered most areas required.

Key data are absent from 1981-1993. While the decisions that have been made to cover these periods when data are absent are reasonable and justified, they still cover over a good deal of uncertainty and could have introduced bias into the assessment. Fortunately, because the problems occur in the early part of the time series, the influence they have is very small.

The main problem is that the model requires much of the information to be broken down by age. For discards and release, the data are sparse and considerable effort has been made in finding ways to fill out periods where data are absent. The choices that have been made are reasonable, but they do suggest that the model is not entirely compatible with the data.

I am not entirely convinced that all the abundance indices are measuring the same component of the stock, because the indices do not exhibit the same patterns. Yellowtail snapper tend to be limited in range, so there is a possibility for local depletion where fleets operate in different locations, which might not be picked up on the spatial scales used to separate catch and effort. However, the different lengths of the time series makes it difficult to be sure, since significant variation in abundance appears to have occurred before the commercial CPUE became available.

Conclusions and Recommendations

The data are sufficient for carrying out a stock assessment. While sampling has not been sufficient in all years and fleets, data collection and sampling in recent years appears to have improved. It is this later period which should have the greatest influence on the stock assessment.

In making decisions about stratifying data into components, statistical models should be used to help decide what stratification to use. While there may be genuine differences between strata in growth or selectivity, for small samples these differences may be swamped by observation error.

Age-length keys are essentially contingency tables which can be analyzed using simple GLMs or GAMs (ICES WKSDFD 2005). Analysis of deviance could be used to decide whether to maintain splits among areas, fleets and years, and whether data might be combined. This is not the only criteria that should be used to split data, but could be a more useful way to justify how missing data and strata with poor sampling might be covered by the available sampling.

It might also be worth considering using GAMs to estimate age-length keys rather than the pooling method used. GAMs would allow differences between areas, fleets and years to be estimated parametrically, which would use fewer parameters in estimating the catch and discard at age and remove some biases in much the same way that GLMs do for abundance indices. However, this approach would be superseded by modeling the age and length composition within the stock assessment model itself.

2. Evaluate the methods used to assess the stock, taking into account the available data.

a) Are methods scientifically sound and robust?

All the methods used in the stock assessment are scientifically sound and robust. The methods applied are standard and well-tested methods. This does not mean, however, that there are no better ways to carry out some of the analyses that capture the uncertainty better. However, these alternative methods follow the same basic model and approach as used in this assessment.

ASAPv2 applies a standard approach to modeling age-structured populations using fishery data. The data requirements are reasonably flexible in that the input data can be generated using standard data from catch monitoring, CPUE indices, surveys and biological sampling. ASAPv2 does not attempt to model all data as it is collected, however, and most input data will need to undergo some processing.

Two assessment approaches were considered. The age-length key (ALK: using age and length data) method was chosen in preference over the "direct ageing" method (DA: using only age data), which was clearly correct. The DA method would not have made use of the length composition data. While there are uncertainties over growth, excluding the length samples in a fishery which is data poor would make little sense.

As for any model, the accuracy of the assessment results will depend upon the degree to which assumptions are met. There is no reason to suppose from the available diagnostics that there is any critical breakdown in the assumptions that would invalidate the assessment.

b) Are assessment models configured properly and used consistent with standard practices?

The assessment models are configured properly and are certainly consistent with standard practice. I did not identify any significant problems, but there may be a few changes which might lead to improved model fit.

Catchability coefficients were held constant over the 30 years in the model. Due to the lack of appropriate covariates, the model does not attempt to adjust for changing catchability apart from through the selectivity blocks. This is a limitation on the data rather than the model.

The selectivity model used for all fleets is logistic, which is an appropriate default assumption, unless evidence suggests an alternative. However, the standardized residuals for the commercial, MRFSS and Headboat age composition suggest that a selectivity function allowing a decrease in selectivity for older fish may fit the data better. It may require some work to discover a function that improves the fit, and the double-sided logistic in ASAPv2 may not be appropriate. Another adjustment which might be considered is to alter the natural mortality.

Selectivity blocks were defined based on regulatory changes. This is a reasonable approach as it is the only external information available on which to base changes in selectivity. However, I suspect that further improvements might be possible, either combing selectivity blocks where there is little data to support the differences (1981-1991), or looking at additional blocks to fit to the data better (e.g. try 1991-97 for the recreational fisheries).

There were four abundance indices used. These are treated as separate indices of abundance in the model and should, generally speaking, follow the same trends. The two recreational indices, MRFSS and Headboat, are in general agreement and cover the longest period. The commercial index covers a shorter period, and while it does not contradict any trends in the recreational fishery index, the relationship between the commercial and recreational indices appears very weak. The much shorter visual survey index is most closely aligned to the commercial index, but again any relationship is weak.

Use of fleet selectivity for standardized MRFSS, Headboat and commercial indices is appropriate since the standardization model used a log-link function and therefore all effects were multiplicative. The NMFS UM RVC index standardized residual plots suggest that the selectivity used in this case is satisfactory.

There was no strong reason to exclude any index. Although differences among indices might be due to changes in catchability, targeting and differences among fishing areas, it is quite likely that the information from the combined indices is more likely to represent the overall changes in stock abundance than any single index. Furthermore, the uncertainty resulting from the differences among indices should be captured within the MCMC.

In this assessment, the choice was made to de-emphasize large numbers of samples by using the square root of the sample sizes rather than using a cap on sample sizes or proportional adjustment. In addition, the method applied an iterative re-weighting scheme to estimate effective sample sizes (ESS). In some cases, this may have resulted in poor weightings due to the lack of data in some selectivity blocks (notably discards). While iterative re-weighting is objective, the final weights are not necessarily convincing estimates of the relative variance of data sources. However, the final ESS that were used as the basis for weighting in the model were unclear.

c) Are the methods appropriate for the available data?

It is not clear that all the methods are entirely appropriate for the available data. The most important problem for the stock assessment method is the inability to cope with small sample sizes or where data are missing.

Errors are allowed in the model for catch-at-age, which means that the method overall copes with observation errors in the data from the catch monitoring, and the age and length composition sampling. This will have the tendency for the model to be more smoothed through the data compared to VPA which fits the selectivity exactly in each year.

It is not clear to me that the iterative re-weighting scheme used to estimate the ESS (multinomial weights) is appropriate where data are missing or sample sizes are small. This perhaps requires clearer justification. Some limitation is required to prevent excessive weights on small data values where they can fit the data well.

GLMs are used to reduce the index variance and correct for some changes in catchability. The GLMs do appear, particularly for Headboat and MRFSS, to have a significant impact on indices. The changes appear to make the indices more coherent. Since they are independently derived, this change would appear to be a significant improvement from the nominal indices.

The GLM does not include all appropriate covariates (e.g. technical changes to fishing) which might account for catchability change. The commercial index in particular shows an upward trend. Although an increase in abundance is quite possible during this period, it may also be partly the result of changes in catchability.

Conclusions and Recommendations

Given the available data, the stock assessment method is sound and has been applied correctly.

For models with low data, there is always a conflict between the detail required by the stock assessment and the information available from the data. This is best resolved in my opinion by modeling the data as directly as possible. While this is not yet done with abundance indices, age and length composition can be modeled directly in stock assessment software (notably Stock Synthesis 3). ASAP does do this to the extent that the proportions for the catch-at-age are modeled using a multinomial likelihood with the ESS determining the variance. However, ASAP does not model lengths directly and does not appear to deal with the missing biological sampling data.

It should not be necessary to pool data to cover strata where data are missing. Although the ability to deal with missing data of this type is implied in the ASAP documentation, it may be that when the ESS are fitted, missing data are not accounted for.

Overall, the use of ESS needs to be described more clearly in the documentation. It was unclear to me exactly what weights were used in the final fit.

Covariates related to catchability should be recorded for standardization purposes. It is suspected that catchability may have changed in ways that the standardization procedures have not accounted for. If these covariates can be identified, they should be monitored and used in standardization. Possible covariates include changes in engine size, electronic equipment (plotter, GPS), fisherman's experience, bait type, hook size and so on.

3. Evaluate the assessment findings with respect to the following:

a) Are abundance, exploitation, and biomass estimates reliable, consistent with input data and population biological characteristics, and useful to support status inferences?

The trends in abundance, exploitation and biomass estimates are consistent with the input data. Overall, the data and model support the stock status inferences.

The discards mortality, which is poorly estimated, would change the overall level of abundance. However, even with 100% discard mortality, the maximum number of discarded fish killed would be less than 6% of the total.

The changes in the abundance indices are consistent with changes in catch, with both showing broadly consistent trends. The catch-at-age data is noisier, but is not inconsistent with the changes in abundance and exploitation estimates. The population's biological characteristics suggest that the stock should be productive and it might be expected to sustain the current levels of catch.

b) Is the stock overfished? What information helps you reach this conclusion?

I do not believe that the stock is overfished. The available indicators based on data alone (mean age, catch and CPUE indices) do not suggest that the stock is overfished. There is no long term decline in age or in CPUE indices in response to increases in catch, which would indicate higher levels of exploitation. Age and CPUE indices over the available time series since 1981 suggest, if anything, the stock has been increasing in size.

The only possible way the data may indicate that the stock is overfished would be for the fishery to have maintained the stock as overfished since the start of the time series in 1981. This is not supported by the age composition of the catches or by the catch rates.

This general conclusion is supported by the stock assessment analysis. The stock assessment model applies a biological interpretation to these indices and indicates that the stock status is not overfished. Although the model requires more assumptions, it does combine all the information to determine the stock status. The model has no major inconsistencies and gives a more precise interpretation of the data.

c) Is the stock undergoing overfishing? What information helps you reach this conclusion?

The overall conclusion is that the stock is not undergoing overfishing. Given the conclusion that the stock is not overfished, any decrease in catches and increasing trend in abundance indices strongly implies that overfishing is not occurring. This appears to have been the case until 2008. However, catches in 2009 and 2010 were the highest since 1999, and therefore whether these can be sustained depends on greater precision in the stock status estimates.

According to the stock assessment model, current fishing mortality is well below the MSY reference point (or any proxy), despite catches having risen in the last few years. The catch (including discards) was 770 t compared to the MSY of 1700 t. There is no other evidence that overfishing might be occurring.

d) Is there an informative stock recruitment relationship? Is the stock recruitment curve reliable and useful for evaluation of productivity and future stock conditions?

There is a relationship between stock size and recruitment driven by the increasing trend in spawning stock biomass and increasing trend in recruitment estimated in the model. This trend has allowed the steepness parameter to be fitted. The fit is reasonable, but the trend is weak, so it is possible that the relationship will change as more data become available. Nevertheless, this is probably as well-estimated as can be obtained from stock assessments. Therefore, the stock recruitment is not only informative, but can be considered reliable and useful for the evaluation of productivity and for the projection. Despite this, it would still be good practice to use proxy reference points rather than the stock recruitment relationship, bearing in mind that the estimates of proxies would most likely be more stable over successive stock assessments.

e) Are the quantitative estimates of the status determination criteria for this stock reliable? If not, are there other indicators that may be used to inform managers about stock trends and conditions?

The quantitative estimates of the status determination criteria for this stock are reliable enough to determine the appropriate management controls. The data are certainly reliable enough to inform managers about stock trends. Overall the stock biomass has most likely been increasing and the condition of the stock is well above MSY or any suitable proxy.

The main uncertainty is the level of sustainable catch, which depends upon absolute, rather than relative, measures of abundance and productivity. This is not known precisely, and therefore some precaution is recommended than just targeting MSY.

4. Evaluate the stock projections, rebuilding timeframes, and generation times, addressing the following:

a) Are the methods consistent with accepted practices and available data?

The methods used for stock projections are consistent with accepted practices and the available data. The stock does not require rebuilding, so no methods are used for rebuilding timeframes and generation times.

Projections are built into ASAPv2, which allow the user to set future exploitation levels. They are based on running the stock assessment forward from the end of the data time series. This implies that the projections are consistent with the assessment model and should produce meaningful results.

The projections do not incorporate alternative management actions or uncertainty. In the projections presented, the current fishing mortality was applied to 2020. Projections are more useful if they are run with different management options, although the assessment scientists would need to know the range of options to be tested.

The projections do not model uncertainty. They do not include various assessment errors estimated by the retrospective analysis, the MCMC or the sensitivity analyses. They also do not include the most important effect on the projections, which is the recruitment variation. The expected recruitments are provided by the fitted Beverton and Holt stock-recruitment function only.

b) Are the methods appropriate for the assessment model and outputs?

The projection methods are integrated into the ASAP v2.0 software, which forms part of the NMFS toolbox. The projections depend on standard age structured population model and the stock recruitment relationship. The age-structured model applies the standard practice for projections of this type, with separable selectivity.

The stock-recruitment relationship will regress the projection to mean equilibrium values of population size within 5 years, and therefore projections into the medium term are unlikely to be reliable. However, this is perhaps an attribute of the knowledge of population dynamics rather than this particular model.

c) Are the results informative and robust, and useful to support inferences of probable future conditions?

The results are informative and robust. They are useful to support inferences of probable future stock conditions in the near term. They are not likely to predict stock sizes accurately beyond 5 years, but can still be used to set precautionary catch quotas and determine likely stock status up the end of this period. The projections do not cover the range of uncertainty however.

d) Are key uncertainties acknowledged, discussed, and reflected in the projection results?

The projections do not include the observation error (MCMC simulations) or recruitment deviations. The best fit estimates only are used in the projection.

Retrospective bias is considered, but is not explicit in the projections. This can occur with unrecorded changes in catchability, productivity or catch. The retrospective bias suggests that spawning stock size may be lower than estimated and that fishing mortality may be higher.

Conclusions and Recommendations

The projections are valid and provide useful information for management decision-making. However they do not account for uncertainty and they do not consider a range of management options. The stock does not require rebuilding, so no methods are used for rebuilding timeframes and generation times.

Projections should use MCMC estimates for the projections to build probability distributions for the likely outcomes. The projections should also account for retrospective bias and sensitivity analyses.

Projections should cover a range of possible management controls. Ideally these scenarios should be provided by the fisheries management based upon their policy and objectives. In the absence of such specified scenarios, alternative fishing mortalities above and below the present level (covering the historical range) should be used.

5. Consider how uncertainties in the assessment, and their potential consequences, are addressed.

- Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods
- Ensure that the implications of uncertainty in technical conclusions are clearly stated.

The technical conclusions from the report indicate that there is a very low risk of overfishing, which taking into account all information is valid. However, there could be improvements in the way that uncertainty is presented.

Uncertainties are primarily covered by the MCMC simulations which are used to map out the likelihood. Prior probabilities are not used. The MCMC covers recruitment parameters, such as steepness as well as catchability and selectivity, and therefore covers the main observation and process errors.

Figures 10.7.14-16 present the results from the analyses used to assess uncertainty. Results are not presented in tabular form. It would be useful to report maximum likelihood, MCMC median and 90% confidence limits for all values of interest (F_{2010}, SSB_{2010}, $SSB_{30\%}$, $SSB_{40\%}$, SSB_{MSY}, $SSB_{2010/MSY}$, $F_{2010/MSY}$ etc.). This would allow decision-making to take better account of risks.

Sensitivity analyses were limited to different discard mortality estimates. It appears that these were chosen on the basis that the parameter was uncertain, not the potential impact on the outcome of the assessment. From the available information, it appears that the results from the stock assessment will be very insensitive to discarding if the current estimated rates are correct. Other uncertainties that might have more impact would be alternative natural mortality values, different steepness for the stock recruitment relationship, the exclusion of early time series data due to poor sampling coverage and other selectivity blocks and functions. However, none of these scenarios would be likely to change the outcome of the assessment.

6. Consider the research recommendations provided and make any additional recommendations or prioritizations warranted.

- Clearly denote research and monitoring that could improve the reliability of, and information provided by, future assessments.

The assessment document identifies some recommendations with respect to discards, but no other research recommendations. Discards in quantity and length/age composition is uncertain for all fleets and clearly requires further research. Their influence on the current assessment appears limited, but if they have been underestimated then the effect of discarding could be not only high, but may result in a less optimistic determination of stock status. In addition, the assessment team believed that discards may be overestimated in the 1992 MRFSS data, and this might be due to outliers because the number of estimated discards was anomalously high; therefore these data should be reviewed. Further research on discarding is clearly justified.

Significant further research on this species alone may not be justified because risks of overfishing appear low. However, a number of research activities might cover several species simultaneously, which would be more cost effective. With this in mind, I suggest the following research areas:

1. Research potential covariates of vessels, crew and gear, which can be used to adjust CPUE for changes in catchability.
2. The RVC survey method could be combined with intensive experimental fishing to improve estimates of selectivity, link RVC selectivity to fleet selectivity and improve understanding of catchability. This might also be a way to research variables that affect catchability.

7. Provide guidance on key improvements in data or modeling approaches which should be considered when scheduling the next assessment.

Current sampling levels should be maintained. The current level of biological sampling, with the additional discard data, is necessary for this type of model.

The model may take better account of past poor sampling cover by either applying a more rigorous way to pool samples or ensure assessment software can be configured for missing data. The latter is clearly preferable and simpler to apply.

The ASAP software should be developed so that it can:

1. deal with missing age composition data in a better way;
2. fit not only age, but length composition and length conditional on age where appropriate (avoiding external age-length keys);
3. Incorporate MCMC and recruitment deviations into the projections.

The discards could be modeled as a decision made after capture. In ASAPv2 discards are modeled as proportions discarded which are estimated from the data. It might be possible to extend this so that discarding is modeled as function of age and/or other variables. For example, discards in the recreational fishery may depend upon what size composition has already been caught as well as the minimum size regulation, with discarding more likely if the catch within a trip already contains larger fish.

References

ICES WKSDFD (2005) Report of the Workshop on Sampling Design for Fisheries Data. ICES Advisory Committee for Fisheries Management. 1-3 February 2005, Pasajes, Spain. ICES CM 2005/ACFM: 11.

Appendix 1: Bibliography of materials provided for review

O'Hop, J., Murphy, M., and Chagaris, D. (2012) The 2012 Stock Assessment Report for Yellowtail Snapper in the South Atlantic and Gulf of Mexico. May 29, 2012. Fish and Wildlife Conservation Commission, Fish and Wildlife Research Institute, 100 Eighth Ave Southeast, St. Petersburg, Florida 33701-5020. YTS_FWC_SAR.pdf

McCarthy, K. 2011a. Commercial vertical line vessel standardized catch rates of yellowtail snapper in southern Florida, 1993-2010. National Marine Fisheries Service, Southeast Fisheries Division. Sustainable Fisheries Division Contribution SFD-2011-015. SEDAR27-RD01

McCarthy, K. 2011b. Calculated discards of yellowtail snapper from commercial vertical line fishing vessels in southern Florida. National Marine Fisheries Service, Southeast Fisheries Division. Sustainable Fisheries Division Contribution SFD-2011-016. SEDAR27-RD02

Chagaris, D. 2011a. Standardized catch rates of yellowtail snapper (*Ocyurus chrysurus*) from the headboat fishery in southeast Florida and the Florida Keys. YTS-RD03

Chagaris, D. 2011b. Standardized catch rates of yellowtail snapper (*Ocyurus chrysurus*) from the Marine Recreational Fisheries Statistics Survey in south Florida, 1981-2010. YTS-RD04

In addition, the ASAP data set used in the stock assessment was provided and the ASAP version 2.0 software was downloaded from the NOAA Fisheries Toolbox website (http://nft.nefsc.noaa.gov/).

Appendix 2: Statement of Work for Dr. Paul Medley

External Independent Peer Review by the Center for Independent Experts

SEDAR South Atlantic and Gulf of Mexico Yellowtail Snapper Assessment Review

Scope of Work and CIE Process: The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Technical Representative (COTR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide impartial and independent peer review without conflicts of interest. CIE reviewers are selected by the CIE Steering Committee and CIE Coordination Team to conduct the independent peer review of NMFS science in compliance the predetermined Terms of Reference (ToRs) of the peer review. Each CIE reviewer is contracted to deliver an independent peer review report to be approved by the CIE Steering Committee and the report is to be formatted with content requirements as specified in **Annex 1**. This SoW describes the work tasks and deliverables of the CIE reviewer for conducting an independent peer review of the following NMFS project. Further information on the CIE process can be obtained from www.ciereviews.org.

Project Description:

The yellowtail snapper assessment from the South Atlantic and Gulf of Mexico regions is a collaborative effort between the National Marine Fisheries Service (NMFS) and the Florida Fish and Wildlife Conservation Commission (FFWCC). This assessment was previously scheduled as part of the SEDAR 27 review held in November 2011, but the assessment model was not completed in time for consideration during the SEDAR 27 review. The last SEDAR review on the yellowtail snapper assessment by CIE reviewers was conducted in August 2003; therefore, a CIE review is requested of the yellowtail snapper assessment.

The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**.

Requirements for CIE Reviewers: Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. CIE reviewers shall have working knowledge and recent experience in the application of stock assessment, statistics, fisheries science, and marine biology sufficient to complete the primary task of reviewing the technical details of the methods used for the assessment. Each CIE reviewer's duties shall not exceed a maximum of 10 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review as a desk review, therefore no travel is required.

Statement of Tasks: Each CIE reviewers shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Prior to the Peer Review: Upon completion of the CIE reviewer selection by the CIE Steering Committee, the CIE shall provide the CIE reviewer information (full name, title, affiliation, country, address, email) to the COTR, who forwards this information to the NMFS Project Contact no later the date specified in the Schedule of Milestones and Deliverables. The CIE is responsible for providing the SoW and ToRs to the CIE reviewers. The NMFS Project Contact is responsible for providing the CIE reviewers with the background documents, reports, and other pertinent information. Any changes to the SoW or ToRs must be made through the COTR prior to the commencement of the peer review.

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE Lead Coordinator on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

Desk Review: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs can not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COTR and CIE Lead Coordinator.** The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements.

Contract Deliverables - Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

Specific Tasks for CIE Reviewers: The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Conduct an independent peer review in accordance with the ToRs (**Annex 2**).
3) No later than 19 June 2012, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and CIE Regional Coordinator, via email to Dr. David Sampson david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

Schedule of Milestones and Deliverables: CIE shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

21 May 2012	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact.
4 June 2012	NMFS Project Contact sends the stock assessment report and background documents to the CIE reviewers. Background documents may be sent to the CIE reviewers one week earlier.
4-15 June 2012	Each reviewer conducts an independent peer review as a desk review.
19 June 2012	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator.
3 July 2012	CIE submits the CIE independent peer review reports to the COTR.
10 July 2012	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director.

Modifications to the Statement of Work: This 'Time and Materials' task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery Management Council, and Council's SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COTR within 10 working days after receipt of all required information of the decision on changes. The COTR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

Acceptance of Deliverables: Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COTR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COTR (William Michaels, via William.Michaels@noaa.gov).

Modifications to the Statement of Work: This 'Time and Materials' task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery Management Council, and Council's SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COTR within 10 working days after receipt of all required information of the decision on changes. The COTR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

Acceptance of Deliverables: Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COTR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COTR (William Michaels, via William.Michaels@noaa.gov).

Support Personnel:

William Michaels, Program Manager, COTR
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-427-8155

Manoj Shivlani, CIE Lead Coordinator
Northern Taiga Ventures, Inc.
10600 SW 131st Court, Miami, FL 33186
shivlanim@bellsouth.net Phone: 305-383-4229

Roger W. Peretti, Executive Vice President
Northern Taiga Ventures, Inc. (NTVI)
22375 Broderick Drive, Suite 215, Sterling, VA 20166
RPerretti@ntvifederal.com Phone: 571-223-7717

Key Personnel:

NMFS Project Contact:

Julie A Neer, SEDAR Coordinator
4055 Faber Place Drive, Suite 201
North Charleston, SC 29405
Email: Julie.neer@safmc.net Phone: (843)571-4366

Annex 1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

3. The reviewer report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Statement of Work

Annex 2: Tentative Terms of Reference for the Peer Review

SEDAR South Atlantic and Gulf of Mexico Yellowtail Snapper Assessment Review

1. Evaluate the data used in the assessment, addressing the following:
   a) Are data decisions made by the assessment panel sound and robust?
   b) Are data uncertainties acknowledged, reported, and within normal or expected levels?
   c) Are data applied properly within the assessment model?
   d) Are input data series reliable and sufficient to support the assessment approach and findings?
2. Evaluate the methods used to assess the stock, taking into account the available data.
   a) Are methods scientifically sound and robust?
   b) Are assessment models configured properly and used consistent with standard practices?
   c) Are the methods appropriate for the available data?
3. Evaluate the assessment findings with respect to the following:
   a) Are abundance, exploitation, and biomass estimates reliable, consistent with input data and population biological characteristics, and useful to support status inferences?
   b) Is the stock overfished? What information helps you reach this conclusion?
   c) Is the stock undergoing overfishing? What information helps you reach this conclusion?
   d) Is there an informative stock recruitment relationship? Is the stock recruitment curve reliable and useful for evaluation of productivity and future stock conditions?
   e) Are the quantitative estimates of the status determination criteria for this stock reliable? If not, are there other indicators that may be used to inform managers about stock trends and conditions?
4. Evaluate the stock projections, rebuilding timeframes, and generation times, addressing the following:
   a) Are the methods consistent with accepted practices and available data?
   b) Are the methods appropriate for the assessment model and outputs?
   c) Are the results informative and robust, and useful to support inferences of probable future conditions?
   d) Are key uncertainties acknowledged, discussed, and reflected in the projection results?
5. Consider how uncertainties in the assessment, and their potential consequences, are addressed.

- Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods
- Ensure that the implications of uncertainty in technical conclusions are clearly stated.

6. Consider the research recommendations provided and make any additional recommendations or prioritizations warranted.
   - Clearly denote research and monitoring that could improve the reliability of, and information provided by, future assessments.
7. Provide guidance on key improvements in data or modeling approaches which should be considered when scheduling the next assessment.
8. Prepare a Peer Review Report summarizing the Reviewer's evaluation of the stock assessment and addressing each Term of Reference.