# Independent review report for the 2012 Stock Assessment of Yellowtail Snapper in the South Atlantic and Gulf of Mexico

Anders Nielsen

July 5, 2012

# Executive Summary

This report is prepared for the Center for Independent Experts (CIE). It contains an independent and impartial review of the assessment of yellowtail snapper from the South Atlantic and Gulf of Mexico, as presented in the assessment report and background documents listed in appendix A.

The data used in the assessment were well presented in the report and overall found to be suitable for carrying out an analytic assessment. The decisions regarding pre-processing of data were found to be in line with common practice.

The quantitative estimates from the presented analytical assessment model were not found to be reliable, due to the following observations: a) Large and mostly unexplained discrepancies between estimates from this assessment and estimates from the assessment model presented in the SEDAR 3 report. b) Large and consistent retrospective biases on key output quantities (spawning stock biomass, fishing mortality, and recruitment). c) Residuals showing consistent overestimation of older ages (7-12+) in the commercial data.

The main relative trends in spawning stock biomass and fishing mortality are however evaluated to be reliable, as they are common in both this assessment, all retro runs, and in the assessment in the SEDAR 3 report. On the basis of these relative trends the stock is found not to be overfished, and found not to be undergoing overfishing.

Main recommendations to improve this assessment in the future are:

- Compare thoroughly to assessment in SEDAR 3.
- Improved age sampling.
- Redefine plus group.
- Try a model with a more flexible selectivity pattern.
- Allow natural mortality to vary over time.
- Less subjective weighting of data sources.

# Background

This review was done at the request of the Center for Independent Experts (CIE). It reviews the 2012 assessment of yellowtail snapper from the South Atlantic and Gulf of Mexico, which is collaborative effort between the National Marine Fisheries Service and the Florida Fish and Wildlife Conservation Commission. The last SEDAR review on yellowtail snapper was conducted in August 2003, and therefore a CIE review is requested. This is a desk review and as such no meeting has been attended by the reviewer for this assessment.

# Descriptions of the reviewers role

This reviewer has independently read the documents provided (see appendix A) and independently authored this review report.

# Findings for each term of reference

To ensure that all terms of reference are covered, and that comments are interpreted with reference to the correct terms, the terms are listed in gray boxes with corresponding reviewer comments following.

1. Evaluate the data used in the assessment, addressing the following:
   - Are the decisions made by the assessment panel sound and robust?
   - Are data uncertainties acknowledged, reported, and within normal or expected levels?
   - Are the data applied properly within the assessment model?
   - Are input data series reliable and sufficient to support the assessment approach and findings?

This reviewer is impressed with the clarity in the report describing the data used and how data are treated to be usable in the assessment. Most of what is described is standard procedure in most assessments, and therefore as sound and robust as most assessments. Having the detail clearly stated strengthens confidence in the data treatment of this assessment. There are naturally a few issues that could be further investigated.

The natural mortality coefficient does enter the models as 'data', even if it really should be considered part of the model assumptions. It is common in analytic assessments that a single more or less arbitrary number is used across all ages and years. In this assessment it is chosen to change the natural mortality assumption to use a theoretical relationship between length and natural mortality. This is likely an improvement. To further explore this change - and its consequences - it would be useful to include selected comparative graphs between the old natural mortality setup and the new one. Such graphs could for instance compare for spawning stock biomass, Fishing mortality at reference age (age 5), estimated recruitment, and estimated selectivities.

The natural mortality coefficient is further assumed to be constant for all years in the data period, but since it is assumed to be length dependent it should be possible to obtain yearly estimates. It is noted on page 11 of section 10 that "There was no basis on which to deploy different mortality vectors by year for yellowtail snapper, ...", but no argument is stated. This reviewer can understand that using raw length based estimates for each year may be

too fluctuating, but a smoothed version could also be considered. Another validation for having a length based natural mortality constant in time would be if the length-at-age was constant over time, but a figure or table documenting this is not found in the report. This is important to consider, because wrongly assuming natural mortality to be constant over time, could lead to retrospective bias, similar to those seen in this assessment (fig. 10.7.16).

The splitting between release mortality and release survival is new to this reviewer, but seems consistently and correctly carried all the way through in the assessment. It is based on scarce data, but that is often the case with discard related information. A suitable sensitivity analysis is reported, which show the main conclusions to be robust w.r.t. this splitting.

The age-length relationship is a central part of the data processing for this assessment. It is largely well described and in line with what is used elsewhere. The age-length relationship is well described by a von Bertalanffy growth model, and the additional detailed modeling of the standard deviation around this curve is excellent. It is however evident that there is very little difference (if any) in length-at-age for ages larger than age 6 (fig. 5.11.9). The point is, that based on length it is very difficult to distinguish between ages larger than 6. This leads to questioning if including age data up to age 12+ is the right choice, or if the plus group should be set to a lower age. This question is in enforced by the problematic residual pattern seen in fig. 11.1.1 where the older ages are consistently over-predicted by the model fit.

It is slightly confusing to the reader to plot raw CPUE and modeled CPUE in the figures (8.8.10, 8.8.11, 8.8.14, 8.8.15 and corresponding figures in the background documents) without making it absolutely clear to the reader that those two sets of CPUE cannot be expected to be similar (except if the explanatory variables were identical in all years). The decision to use standardized CPUE instead of raw CPUE is supported in the corresponding background documents, and in the referenced published papers.

The data, as they are, are applied properly in the assessment model. Comments w.r.t. treatment of uncertainties are explained in the context of the assessment model, so they can be found under term of reference 2.

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

   - Are methods scientifically sound and robust?
   - Are assessment models configured properly and used consistent with standard practices?
   - Are the method appropriate for the available data?

The assessment model itself is not sufficiently described in the assessment report. There are a number of ways the description could be improved. a) There are too many misprints in equations (e.g. $\alpha$ should be $a$ in equation 10.2.2.4, missing subscript $a$ in equation 10.2.2.9,

and $p$SSB should be $p_{\mathrm{SSB}}$ in equation 10.2.2.14.). These simple misprints are obvious to stock assessment modelers, but can be very confusing to others. b) Throughout the model description there are mentions of so-called 'deviations' (as opposed to normal model parameters), but these deviations are not defined or explained anywhere in the report. Readers who are not familiar with this concept in AD Model Builder will not be able to read and interpret these estimates correctly, and furthermore it leads to the wrong parameter count at page 16 of section 10. Deviations are simply parameter vectors summing to zero, which means that a deviation vector of length $n$ only contains $n - 1$ free model parameters. c) Finally, there is no information about what probability distributions are assumed for the different data sources (e.g., is a certain index assumed to be normal or log-normally distributed). This is a serious omission, because it leaves the model description incomplete. There is no obvious way of guessing exactly what is assumed, so the (avid) reader has no other choice than to read the ASAP2 technical documentation and the ASAP2 input file supplied in appendix B.

This reviewer was confused by the statement below equation 10.2.2.4 that states "Parameters were estimated using the MS-Excel Solver ...", which is in conflict with the statement on page 16 of section 10 where a listing of the parameters estimated by the ASAP2 model includes the same parameters. The latter is taken to be correct.

Standard deviations for the different information sources are not estimated in the ASAP2 model, but specified by assigning values to coefficients of variations (CVs) and effective sample sizes (ESSs) (page 20, section 10). For the indexes a more objective iterative procedure was applied. Assigning standard deviations based on subjective judgment to different information sources has consequences. a) The relative weight given to the different data sources is derived from those standard deviations, so the risk is that the final estimates are based on misleading data, whereas valuable data are ignored. b) The confidence intervals for the final estimates of the quantities of interest (e.g. SSB, recruitment, and F) are constructed by propagating the - now assumed - standard deviations through the model. This means that assigning standard deviations to observations is equivalent to (indirectly) assigning the width of the final confidence intervals.

Exploring the likelihood surface via McMC methods is an excellent way of dealing with the non-linearity and non-normality of the model, but the subjectivity introduced by assuming standard deviations for certain data sources remains unchanged.

The graphical representation of the model's standardized residuals is very detailed, and offers great opportunity for investigating details, but the greater perspective can be somewhat lost when the residual plot for a single fleet spans 60 frames and 6 pages. The same information could be presented in a bubble-plot (similar to fig. 7.9.11) for each fleet with years on the x-axis, age on the y-axis, color of bubble indicating sign of the standardized residual, and size of bubble indicating absolute value of the standardized residual.

Judging by the standardized residual plot (fig. 11.1.1, 11.1.2, and 11.1.3) the model fit is poor. a) For the commercial data the model is consistently predicting too high proportions for the older ages (say greater than age 6), and has at least a strong tendency to predict too low proportions for the younger ages. The same pattern is seen for commercial, head

boat, and MRFSS fleets. This could indicate that the selectivity assumption is invalid. b) Standardized residuals from discard data and from the NMFS-UM RVC index show a different problem. The absolute values of the standardized residuals are relatively small for older ages and relatively large for younger ages, which could indicate that the standard deviation is not correctly specified w.r.t. ages. Part of the explanation for this is that very few observations are available for the older age classes, so this may not be too problematic, but the trend seems to be consistent in all years, and not only for the ages with low or zero observations.

The retrospective analysis (fig. 10.7.16) is also problematic. For fishing mortality every new year of data increased the previous estimates, for spawning stock biomass each new year of data decreased the previous estimates, and for recruitment each new year of data decreased the previous estimates. All of these trends were consistent in the five years considered by the retrospective analysis. This retrospective pattern is further unusual in how far back these changes persist. Consider for example the spawning stock biomass (fig. 10.7.16b) estimated with data ending in 2005 compared to values estimated with data ending in 2010. In 2005 the difference is roughly 3000mt (or ca. 30%), but ten years prior in 1995 the difference was still roughly 1700mt (or ca. 20%). Such retrospective pattern are often seen as a consequence of assuming something in the model to be constant in time, when in reality it is not.

3. Evaluate assessment findings with respect to the following:

- Are abundance, exploitation, and biomass estimates reliable, consistent with input data and population biological characteristics, and useful to support status inferences?
- Is the stock overfished? What information helps you reach this conclusion?
- Is the stock undergoing overfishing? What information helps you reach this conclusion?
- Is there an informative stock recruitment relationship? Is the stock recruitment curve reliable and useful for evaluation of productivity and future stock conditions?
- 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?

An issue, which in this reviewer's opinion deserves more attention in the assessment report, is the discrepancy between this assessment and the assessment reported in the SEDAR 3 report. It is mentioned (first paragraph in section 10.4) that this assessment resulted in "lower fishing mortality". The estimated fishing mortality in this assessment is ca. a fifth of the estimated fishing mortality in the SEDAR 3 assessment, comparing the period covered by both assessments (1981-2001). The estimated spawning stock biomass in this assessment is ca. four times the estimated spawning stock biomass in the SEDAR 3 assessment. The estimated number of recruits in this assessment is ca. two times the estimated number of recruits in the SEDAR 3 assessment. These differences are so large, that it is important that

the differences (in data and assumptions) are fully understood. The reasons mentioned in the assessment report (more data available, and length adjusted natural mortality assumption) are, in this reviewer's judgment, not likely to explain such substantial differences.

It should be noted that the assessment in the SEDAR 3 report did not show retrospective bias, and the residual plots were less problematic than for the current assessment.

It should be noted, that even with very different levels of fishing mortality, recruitment, and spawning stock biomass between the two assessments, the relative trends were similar.

The residual and retrospective analysis showed several severe mismatches between the data used and model assumptions, and there is a substantial unexplained difference in the period 1981-2001 between the SEDAR 3 assessment and the current. For these reasons this reviewer does not consider the abundance estimates, exploitation pattern, and biomass to be reliably estimated in this assessment.

Quantitative estimates based on this assessment are not considered reliable by this reviewer for the reasons previously stated.

The problem with the residuals relates to the age-specific observations, but the fit to the overall landing weights, discard weights, and indexes (fig. 10.7.2, 10.7.3, and 2.7.4) show a general ability of the model to follow the main relative trends. Similarly the retrospective analysis showed a bias, but it did not change the overall stock development trends. The overall stock development trends indicate increasing spawning stock biomass (fig. 10.7.9), decreasing fishing mortality (fig. 10.7.10), and stable recruitment (fig. 10.7.7), which indicated to this reviewer that the stock is not overfished and the stock is not undergoing overfishing.

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

- Are the methods consistent with accepted practices and available data?
- Are the methods appropriate for the assessment model and outputs?
- Are results informative and robust, and useful to support inferences of probable future conditions?
- Are key uncertainties acknowledged, discussed, and reflected in the projection results.

The stock projections carried out are deterministic and the resulting trends are very constant. Possibly stochastic projections would be more interesting. Uncertainties are not considered. In the assessment report the stock is not deemed overfished, or undergoing overfishing, so no rebuilding plan is considered.

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 the technical conclusions are clearly stated.

As already mentioned under term of reference number 5, the standard deviations for the different information sources are not estimated in the assessment model, but specified by assigning values to coefficients of variations (CVs) and effective sample sizes (ESSs) (page 20, section 10). This reviewer would prefer objective estimates based on model fit (e.g. maximum likelihood estimation). It should be noted that such estimation of variance parameters can also be problematic and sometimes lead to biased estimates, but with three fleets and one additional index, it should be possible. It is important, because all later statistical inference from the model comes from the uncertainties associated with the input data.

As is standard practice, certain parts of the input are simply considered given as known co-variates. Here for instance are natural mortality, stock weights, and proportion mature. Sensitivity to these can be visualized and discussed, but are generally well understood. The sensitivity to the splitting between release mortality and release survival was investigated.

A key assumption in this assessment is the logistic selectivity curve. If this assumption is invalid, it could partly explain the problematic residual pattern. It is also one of the few assumptions that is different between this assessment model and the assessment model used in the SEDAR 3 report. No attempt is made in the assessment report to investigate the sensitivity to this assumption in the model.

The likelihood is explored via McMC, which is an excellent way of dealing with the non-linearity and non-normality, but it does not help w.r.t. the issues mentioned above.

Finally, the uncertainty w.r.t. model choice seems especially relevant for this assessment considering the large discrepancy between this assessment and the assessment in the SEDAR 3 report. This uncertainty has not been investigated.

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 report suggest as a research recommendation to gather data on released fish (sizes, quantities, disposition at release). Data on releases are sparse, so that should be

encouraged. The authors of this assessment model should be complimented on implementing a model that can actually carry this information all the way through to the final estimates.

The assessment report suggests an investigation of possible outliers in certain parts of the data. This seems like a reasonable thing to do. In this context it is also suggested to restrict the data to 1993 to the present. As a sensitivity measure that is a valid suggestion, but maintaining a longer data series is valuable to provide a historical perspective, so correcting the outliers sounds more promising.

Recommendations are listed in the section on "conclusions and recommendations".

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

Recommendations are listed in the section on "conclusions and recommendations".

8. Prepare a Peer Review Report summarizing the Reviewer's evaluation of the stock assessment and addressing each Term of Reference.

This report.

# Conclusions and recommendations

**Compare thoroughly to assessment in SEDAR 3.** Possibly the most important issue in this assessment, and one which is only briefly mentioned in the assessment report, is the large discrepancy between the current assessment and the SEDAR 3 assessment in the period they both cover. It needs to be determined what is causing such a discrepancy before either assessment can be trusted. Currently the SEDAR 3 assessment looks more credible, to this reviewer, due to the better model diagnostics. A valid starting point for such an investigation would be to run the SEDAR 3 data with the current assessment model (or the other way around) to see if the difference is due to the data, or if it is caused by different model assumptions. Once the difference is understood for the same data subset, it should be possible to make a more informed model choice for the entire data set.

**Improved age sampling.** Look into getting more direct age sampling, for the fleets where age data are currently unavailable. The ability to age-classify based on lengths seems limited for ages larger than age 6 (fig. 5.11.9). The two obvious ways to investigate the problematic residual pattern is via the assumed selectivity, or as here suggested via age classification.

**Redefine plus group.** A related suggestion to improving the age sampling data, would be to approach the issue via the modeling approach. If in fact the problem is age-classification of the older ages, then a quick (intermediate) solution could be to set

the largest age group to for instance age seven or older (7+). This may improve the residual plot and strengthen confidence in the model. Currently the largest age group is set to age twelve or older (12+).

**Try a model with a more flexible selectivity pattern.** In this assessment separate time periods (1981-1984, 1985-1991, and 1992-2010) are used to soften the rigid assumption of multiplicative/separable fishing mortality. The periods and the number of periods are based in subjectively defined regulatory periods. Within each period the selectivity is assumed constant. If such assumptions are invalid it can cause retrospective patterns. It is suggested to try an assessment model with a more flexible fishing mortality pattern, or simply modify the ASAP2 model used to allow more flexible patterns. Fishing at age could be setup as random walks within each age group or as spline functions. AD Model Builder supports both options.

**Allow natural mortality to vary over time.** The natural mortality is assumed to be length dependent, but constant for all years (1981-2010). In the interest of narrowing down what causes the retrospective pattern in this assessment, a time varying natural mortality should be tested. Instead of using the overall length-at-age distribution to calculate the natural mortality, it could be calculated from the yearly length-at-age distribution. If yearly estimates are too fluctuating, or data limited, then blocks of, for instance, five years can be used, or the yearly estimates can be smoothed.

**Less subjective weighting of data sources.** Assigning standard deviations subjectively to the different information sources entering the assessment has a number of negative consequences. An alternative option is to use maximum likelihood estimation, which amounts to simply treating the standard deviation parameters in the model like any other model parameter. In order to make the model identifiable it will be necessary to assume that certain standard deviation parameters are constant over time and common for some age groups, but such assumptions can be validated via the residuals. Using maximum likelihood estimation is not free of problems, as the estimates will generally not be unbiased (as is also true with most of the other parameters in the model), but at least it is an objective approach.

**Improved release sampling.** Already mentioned under term of reference 6.

**Simplified residual graphics.** This is a trivial, small, but important request. Instead of, or at least in addition to, the very detailed presentation of the model's standardized residuals, it is suggested that each fleet's information be presented in a bubble-plot (similar to fig. 7.9.11) with years on the x-axis, age on the y-axis, color of bubble indicating sign of the standardized residual, and size of bubble indicating absolute value of the standardized residual. In addition to saving space, it would: a) Help to see patterns across years. b) allow compression of standard deviation assumption across years, which is currently difficult because the residual plots have different scales. c) Make it possible to follow cohort patterns.

# Appendix A: Bibliography of materials provided for review

SEDAR. 2003. SEDAR 3 — Complete stock assessment report of yellowtail snapper in the southeastern United States. SEDAR 3 Assessment Report 1. South Atlantic Fishery Management Council. Charleston SC.

SEDAR. 2012. The 2012 Stock Assessment Report for Yellowtail Snapper in the South Atlantic and Gulf of Mexico. Florida Fish and Wildlife Conservation Commission and Wildlife Research Institute. St. Petersburg, Florida.

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.

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.

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

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

# Appendix B: Copy of the CIE Statement of Work

**Attachment A**

**Statement of Work for Dr. Anders Nielsen**

**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.