




**UNITED STATES DEPARTMENT OF COMMERCE  
National Oceanic and Atmospheric  
Administration**

National Marine Fisheries Service  
Southeast Fisheries Science Center  
75 Virginia Beach Drive  
Miami, Florida 33149 U.S.A.  
(305) 361-4200 Fax: (305) 361-4499

September 19, 2017

**MEMORANDUM TO:** Gregg Waugh  
Regional Administrator, Southeast Regional Office

**FROM:** Bonnie J. Ponwith, Ph.D.  
Science Director 

**SUBJECT:** Request for evaluation of Gray Triggerfish assessment efforts

The following report is in response to the South Atlantic Fishery Management Council's request for evaluation of Gray Triggerfish assessment efforts.

cc: Jack McGovern and Rick DeVactor  
Monica Smit-Brunello  
Theo Brainerd, Erik Williams, and Larry Massey

## **Evaluation of Gray Triggerfish Assessment Efforts**

### **1. Review reasons why the Gray Triggerfish assessment in SEDAR 32 was halted and note the extent to which those issues have been resolved.**

The Gray Triggerfish assessment in SEDAR32 was halted due to issues with the age data submitted through the data workshop, which included contributions from NMFS Beaufort Lab (NMFS-B) and South Carolina Department of Natural Resources (SCDNR). Prior to SEDAR32, two age workshops were held – one in September 2011 and the second in October 2012 (see SEDAR32-DW03). Participants were from NMFS-B, SCDNR and NMFS Panama City Lab (NMFS-PC). NMFS-PC were actively ageing Gulf of Mexico Gray Triggerfish. Following the first workshop, the measure of precision in age readings between NMFS-B and SCDNR was not acceptable, and a bias in the readings was also noted. Following the second workshop, precision was improved and acceptable and there was no significant bias between the labs' readings. Based on those results, NMFS-B and SCDNR went into production ageing in preparation for SEDAR32.

While preparing for the assessment workshop of SEDAR32, the lead analyst noted discrepancies in the age data that were difficult to interpret. The size-at-age and growth rates between the fishery-dependent (NMFS-B data) and fishery-independent (SCDNR data) data sets were markedly different. The ages at recruitment to the fishery-independent gear and the fishery-dependent gears were also different, and opposite what was expected. The lead analyst requested a re-reading of a subset of samples by NMFS-B and SCDNR. The results revealed that within-lab, paired readings were consistent, but there was a highly significant bias between the two labs' readings. The age readers from each lab were identifying different features on the age structure (dorsal spine) that they were considering as annuli. The discrepancies could not be easily resolved; thus, the assessment was halted.

Following the halt of the Gray Triggerfish assessment, NMFS-B, SCDNR and NMFS-PC convened another age workshop held in November 2013 and exchanged calibration sets. The exchange of samples showed that the age readers from SCDNR were inconsistent in their age readings, while NMFS-B was consistent throughout the production ageing process. A conference call was convened between staff of both labs and included assessment analysts. The decision was made for NMFS-B to read all age samples, fishery-independent and fishery-dependent, and to use those readings for SEDAR41.

The main research recommendation cited in SEDAR32 and SEDAR41 was for an age validation study to be done on Gray Triggerfish. NMFS-B was awarded a three-year MARFIN grant to conduct this study which began in the fall of 2015. The study is coming to a close and the samples are being processed and analyzed. The results should be available in the winter of 2017.

### **2. Review the reasons why the Gray Triggerfish assessment in SEDAR 41 did not pass peer review and note the extent to which those issues have been resolved.**

Data issues and modeling issues were identified by the SEDAR 41 review panel as reasons for why the assessment did not pass review. These are outlined in detail in the SEDAR 41 review workshop report.

A major concern of the Review Panel was an error in the fishery-independent MARMAP trap age compositions that was reported to the Review Panel on the third day of the Review Workshop. Increment counts rather than calendar ages were originally provided to the Data Workshop and were used to characterize fishery-independent age compositions throughout the assessment process. This error was found and corrected at the Review Workshop and a revised model was developed, but there were changes in the model fit that could not be fully explored with the remaining time at the workshop. The error resulting in the incorrect age compositions has been resolved.

The second data issue related to inclusion of the 1990 (first year) value in the fishery-independent MARMAP trap index. This year was the lowest index value of the time series and its inclusion had a significant effect on the initialization of the model. Effects from Hurricane Hugo were brought forward at the Review Workshop as a potential mechanism affecting catchability to the survey in this year, and the Review Panel recommended a multi-species investigation of catch rates from the MARMAP survey in this year as a way to determine the legitimacy of the 1990 index value. This investigation has not been conducted to date. The Review Panel also recommended exploring gear saturation effects on the MARMAP trap index. Some work has been conducted on this issue for Gray Triggerfish and other species but a detailed investigation of gear saturation effects on standardized abundance indices for use in a Gray Triggerfish assessment has not been conducted.

The major modeling issue identified by the Review Panel was the decision to up-weight the fishery-independent MARMAP trap index. The approach recommended in the literature (Francis 2011) and used in prior SEDAR assessments is to iteratively re-weight data sources in the assessment model to achieve the best fit across the various inputs. Then if this approach does not result in an adequate fit to abundance indices, Francis (2011) recommends further up-weighting abundance indices to achieve a better fit. In the case of Gray Triggerfish, the MARMAP fishery-independent trap index was not fit well even after reweighting. Given that this is a fishery-independent index and therefore less susceptible to catchability changes and sampling biases, the Assessment Panel decided efforts should be made to improve the fit to the trap index. An approach was developed at the Assessment Workshop that up-weighted the trap index to varying degrees (2-10 fold) and evaluated the associated degradation in fit to other data sources (length and age compositions) that occurred in order to improve the index fit. A weight was chosen (6-fold) by the Assessment Panel that maximized the fit to the fishery-independent trap index with minimal loss of fit to the other data sources in the model. The Review Panel questioned the decision to up-weight the fishery-independent index and some reviewers disagreed with the need to fit the index well given the large annual CVs. Up-weighting the index also forced the fit to the suspect 1990 (initial) index value, which was also of concern to the Review Panel. Sensitivity analyses that investigated alternative weights (including no up-weighting), exclusion of the 1990 value, alternative model start dates (i.e., 1974), and inclusion or exclusion of various indices did not result in a change in terminal fishing or biomass status. The issue of how best to weight the fishery-independent trap index has not been resolved.

The Review Panel also felt that composition data may be over-weighted in the model because both length compositions and age compositions were included in the base model run. Model runs with length compositions excluded did alter the model fit, but did not result in a change in stock status.

A more general, and perhaps the major, issue noted by the Review Panel and evident throughout the assessment process was the lack of contrast in the data underlying the assessment of Gray Triggerfish. In particular, abundance indices often increased in concert with increasing landings and age compositions expanded in periods when landings (and fishing mortality) were relatively high, opposite to patterns expected if fishing was having a significant effect on the stock. This lack of contrast was particularly evident in (unsuccessful) attempts to fit an age aggregated surplus production model, which depends on strong contrast between landings and abundance indices. While the various model configurations developed throughout the assessment process as well as the suite of sensitivity analyses conducted were consistent in the conclusion (also noted by the Review Panel) that there was little evidence that Gray Triggerfish were overfished or experiencing overfishing, the model fits and the stability of the model were often sensitive to changes in inputs or model assumptions.

### **3. Compare and contrast the assessment issues experienced in the South Atlantic with the Gulf assessment efforts, and comment on how lessons learned through the Gulf efforts can help ensure a future successful South Atlantic assessment of Gray Triggerfish.**

The Gulf and South Atlantic assessments of Gray Triggerfish are similar in that they both are based on a catch-age model formulation, include age based natural mortality, reproductive output, and selectivity, model multiple recreational and commercial fishing fleets, and fit to multiple indices of abundance, age compositions, length compositions, and landings. A detailed comparison of the productivity measures and assessment assumptions is provided in the SEDAR 41 Assessment Workshop Report (Table 28).

South Atlantic and Gulf of Mexico Gray Triggerfish are different stocks and the two regions have different exploitation histories and available data sources. A major difference between the two regions is that Gray Triggerfish have been heavily exploited in the Gulf with an estimated 5-fold decline in biomass since the 1950s. Multiple assessments since 2001 indicate the Gulf stock is overfished and experiencing overfishing (SEDAR 43 Assessment History and Review). In contrast, the South Atlantic Gray Triggerfish assessment had a start year of 1988. This year was chosen primarily because Gray Triggerfish were considered undesirable prior to the 1980s and presumably had high discard rates; data were not available to quantify historical discard rates and, hence, reconstruction of historical removals (landings + discards) was not possible. There is little trend in Gray Triggerfish landings or indices since 1988 and so little trend in estimated biomass. One consequence of this difference in exploitation history is that the recruitments associated with periods of low biomass in the Gulf allowed steepness to be estimated in the assessment model; steepness was not estimable in the South Atlantic model and average recruitment was assumed over the assessment time period.

The Gulf assessment was also informed by three fishery-independent abundance indices and several fishery-dependent indices, while the South Atlantic assessment was informed by a single fishery-independent index using a single gear (traps) over a limited portion of the stock area. The fishery-independent MARMAP trap index was in conflict with other data sources in the model, particularly the landings history, which led to the need to up-weight the index in order to fit it. The Gulf abundance indices did not require up-weighting and appeared to reflect changes in landings, with periods of high landings followed by declines in abundance indices. In contrast, periods of high landings were not necessarily reflected by subsequent declines in abundance in MARMAP trap index. As a result, there appears to be better contrast in the data underlying the Gulf assessment than the South Atlantic assessment.

The criteria used to judge the utility of the South Atlantic and Gray Triggerfish assessment for management differed between the South Atlantic and Gulf review panels. The Gulf Gray Triggerfish assessment did not include a Review Workshop with CIE reviewers but, rather, was reviewed by the Gulf Scientific and Statistical Committee (SSC). Based on the Review Summary, the Gulf SSC was more amenable to using the assessment (or aspects of the assessment) for management, despite identified shortcomings in the assessment and the data. In contrast, the CIE reviewers for the South Atlantic Gray Triggerfish assessment recommended the issues outlined in the Review Panel report be addressed by the Assessment Workshop panel prior to making a status determination and recommending an ABC for the South Atlantic stock, even though the identified issues and prior sensitivity analyses had no effect on fishing or biomass status.

**4. Identify any research and monitoring needs, or additional impediments to assessing Gray Triggerfish in the South Atlantic, and comment on the efforts necessary to resolve the remaining issues to support a future assessment.**

Research recommendations from the SEDAR 41 Gray Triggerfish Data, Assessment, and Review Workshops are described in section IV of the final workshop report. The main research needs in light of the most recent SEDAR 41 assessment are (1) age validation studies and increased sampling for age compositions (near completion), (2) reliable indices of abundance, and (3) accurate landings and discard time series.

Age validation studies for Gray Triggerfish are currently in progress.

Several aspects of the historical (i.e., 1990-2010) MARMAP fishery-independent trap index should be evaluated. A multispecies investigation of Hurricane effects on the 1990 value of the MARMAP fishery-independent trap index should be conducted. A consistently low value for multiple species would support the hypothesis that the survey had low catchability in this year. More generally, environmental effects (e.g., temperature) on catchability of the trap survey for Gray Triggerfish and the effects of changes in spatial coverage (with expansion of the survey in 2010) and other potential changes should be investigated and, where possible, accounted for in the index standardization process. Finally, methods to integrate the historical trap and more recent combined trap and video survey data should be developed and evaluated.

Methods to re-construct historical removals of Gray Triggerfish should be investigated. Desirability of Gray Triggerfish appeared to increase in the 1980s but there was no information available prior to this time to reconstruct historical discards, which were presumably high. This led to the decision to use a later (1988) start date for the model which, in combination with the low 1990 index value, resulted in a low estimated initial biomass that was of concern to the reviewers.

Analytical methods to address data conflicts and data weighting within the context of integrated assessment models should be investigated. The Francis (2011) method of iterative re-weighting has been adopted in SEDAR assessments as an objective approach to balancing the model fit across the various datasets in an assessment model. However, the Francis (2011) approach also recommends giving primacy to abundance indices (up-weighting beyond that from iterative reweighting) if these are not fit well. No specific guidance on the degree of up-weighting is provided, however, and ultimately, this decision has a subjective element. This issue (further up-weighting the index) was discussed at length at the both the Assessment and Review Workshops. Ultimately, the two panels took opposing views on the decision to up-weight.

A procedure for handling data errors discovered after the Data Workshop should be developed. In theory, the data used in the assessment model are finalized by the data deadline following the Data Workshop and do not change throughout the assessment. In practice, data errors are often discovered throughout the process. In most cases, these are corrected without consequence, in some cases via an *ad hoc* working group if the error is substantial. Unfortunately, in the case of Gray Triggerfish, the error was discovered and reported to the Review Panel very late in the process (third day of the Review Workshop). Given the limited time, there was limited discussion of the nature of the error, no review of the revised data that was subsequently provided at the workshop, and limited time to fully understand the effects on the assessment model.

The objectives of the Review Workshop should be clearly communicated to the Review Panel and the analytical group. The difference between the Gulf and South Atlantic SEDAR assessment reviews for Gray Triggerfish illustrate the wide disparities that can occur in the review process. The issues raised by the SEDAR 41 Review Panel were legitimate, and some were easily resolvable (inclusion vs. exclusion of the 1990 index value, use of length compositions) while others were not (data weighting). But they were not considered in the context of the larger goal of the assessment process, which is presumably to provide advice to managers. All of the issues raised at the SEDAR 41 Review Workshop were investigated by additional model runs and sensitivity analyses and in no case, was there a change in stock status (overfishing, overfished). Hence, an alternative recommendation could have been to use the model for status determination but not setting ABC (without additional work).

In the future, paired assessments should choose species that have similar assessment histories, data and modeling issues (to the extent that these are known), and levels of political controversy. In the case of SEDAR 41, Gray Triggerfish, a species with no prior assessment and known ageing issues, was paired with Red Snapper, a controversial species with a substantial assessment history. There were comments by both Assessment and Review Workshop panelists

that there was insufficient time to fully address the Gray Triggerfish assessment given the amount of time that was dedicated to Red Snapper.

**5. Comment on the type of assessment model that would be appropriate for Gray Triggerfish and when an assessment could be attempted given the available data.**

Most of the issues with the SEDAR 41 Gray Triggerfish assessment were related to the data used to inform the assessment, not the particular assessment model structure used. As a result, a catch-age model and perhaps a production model are most appropriate for this stock. Stock assessment models rely on an effect of fishing being evident in the data (e.g., increased landings result in a decrease in abundance indices). No assessment model structure can remedy this lack of contrast in the data inputs. Hence, there is a possibility that this issue will be present in future assessments for this stock. Therefore, efforts should be made to remedy any data issues, particularly the annual age compositions, the indices of abundance (in particular the fishery-independent index), and the landings/discards history prior to undertaking another assessment. Data limited approaches are an alternative approach that can be employed, but these often require strong assumptions and, while providing information to recommend catch levels, do not typically provide information on stock status.

