

SEDAR 100 Gulf Gray Triggerfish Assessment Review

Center for Independent Experts (CIE) Independent Peer Review Report

Anders Nielsen

September 2026

Executive Summary

SEDAR 100 is the first completed benchmark assessment of Gulf gray triggerfish (*Balistes capriscus*) in more than a decade, following SEDAR 43 in 2015 and the abandonment of SEDAR 62 in 2019 over data quality concerns. It is a standard assessment, implemented in Stock Synthesis 3, and its results are intended to go forward to the Gulf Council SSC as the basis for fishing level recommendations.

The assessment is based on an extensive rebuilding of the input data. All fish aged five years and older were re-aged under a protocol externally validated by bomb radiocarbon and ages zero and one were re-evaluated. The fleet structure was consolidated to Gulf-wide fleets by gear, and a new proxy-based method was introduced for shrimp trawl bycatch. Natural mortality was raised and growth is now estimated inside the model. The model itself is formulated to rely less on age-readings and more on length measurements. The volume and quality of this work is impressive, and it addresses the issues that stopped the previous attempt. Hence, it is of limited utility to directly compare with SEDAR 43.

The methods are scientifically sound. Stock Synthesis is an appropriate and well-tested framework, the modelling priorities were clearly stated, and the configuration follows standard US practice. During the review meeting the analytical team presented the work openly and responded to a considerable number of requests for additional runs and diagnostics, several of which materially changed the panel's conclusions. The most consequential of these concerned steepness. The assessment was presented with steepness fixed at 0.99, a common pragmatic choice which effectively makes recruitment independent of stock size. For this stock, however, the likelihood profiles indicated that steepness was estimable, and the individual data components agreed on the estimate. This conclusion survived shortening the time series in two steps and removing the recreational compositions. The model with steepness estimated fits the historical data better, largely resolves a retrospective pattern, and removes a tension that had previously required the recruitment deviations to be forced to sum to zero. The review panel concluded that this is the preferred model candidate. The estimated value of about 0.42 is, however, not considered biologically plausible for this species, and the historical stock-recruitment pairs span only a narrow range of spawning biomass, so neither the estimated curve nor the fixed one should be extrapolated into projections or equilibrium calculations.

I recommend the following:

- Estimate steepness rather than fixing it at 0.99, on the grounds that it is consistently estimable here and that estimating it improves both fit and diagnostics.
- Define status determination criteria and projections so that they do not depend on the stock-recruitment curve at either value, using per-recruit quantities for the fishing mortality threshold and recent observed recruitment, rather than virgin recruitment, as the scaling constant for the biomass threshold (see ToR 5).
- Base short-term projections on the empirical distribution of recruitment over a recent, well-informed window selected by an objective criterion.

- Investigate the convergence behaviour revealed by the jitter analysis, in which only 42 of 100 runs returned the same solution, and simplify the model, if possible without loss. Until this is resolved, jitter runs must be repeated before any future run can be assumed to have found the global solution.
- Down-weight the composition data. The one-step-ahead residuals show clear positive correlation between neighbouring length classes, which the Dirichlet-multinomial cannot represent, so the composition data currently carry more weight than their information content justifies (optional, as results are consistent).
- Explore MCMC and self-simulation as alternatives to asymptotic standard errors, which rest on regularity conditions that are not demonstrated here and are implausible given the number of year-specific parameters and the evidence of identifiability problems (longer term).
- Investigate the MASE diagnostic, which in one case reports the model as 43 times worse than a naive prediction. This reviewer does not believe that to be the case, and suspects the diagnostic is misspecified, but it should be checked rather than set aside.

The assessment is a substantial improvement on what came before, and this reviewer considers the model suitable as a description of the historical dynamics of the stock, subject to the changes above. However, the results should only be trusted conditionally on the data reconstructions they rest on. For 36 of the model's 80 years the recreational removals are reconstructed rather than surveyed, and the shrimp trawl bycatch, which accounts for more than half of all fish killed in every decade from the 1940s to the 2010s, is not directly observed in any year of the time series. There are no recruitment deviations before 1981, so the early trajectory is a deterministic forward calculation driven by assumed removals. Adding gray triggerfish to the species list of the shrimp observer program is, in this reviewer's view, one of the highest priorities. Removals of this magnitude should be directly monitored. Quantities anchored on the unfished state are weakly determined, which is the reason to move the status determination criteria to a per-recruit basis.

Background

An independent peer review of the SEDAR 100 stock assessment of Gulf gray triggerfish (*Balistes capriscus*) is conducted as part of the SEDAR (Southeast Data, Assessment and Review) process. A single stock is assessed, as the Gulf population is treated as a single unit stock. This review was prepared following a review workshop held in person in Tampa, FL, on 11-14 August 2026, and is prepared for the Center for Independent Experts (CIE), following the statement of work contained in Appendix B. A list of the documents reviewed may be found in Appendix A.

The review panel consisted of Luiz Barbieri (Chair, Gulf Council SSC), Jie Cao (Gulf Council SSC), Erik Lang (Gulf Council appointee), and three CIE reviewers: Yong Chen, Nicholas Fisch and Anders Nielsen. The assessment was presented by Skyler Sagarese and Katie Siegfried of the Southeast Fisheries Science Center, and the workshop was organised by SEDAR staff.

SEDAR review workshops provide independent peer review of stock assessments prepared through SEDAR data and assessment workshops. The goal of the review is to ensure that the assessment and results presented are scientifically sound. The review panel's primary responsibility is to ensure that assessment results are based on sound science, appropriate methods, and appropriate data. During the review, the panel is allowed some flexibility to ask questions and to request additional investigations (e.g., additional sensitivities, modifying the assessment, configuration, and assumptions within reason). The assessment team gave excellent presentations on all subjects and quickly provided the many additional investigations requested by the review panel. The assessment carries operational consequences, and the review has been conducted based on that understanding.

The last completed assessment for this stock was SEDAR 43 in 2015, with data through 2013, which found the stock to be overfished and provided the basis for the rebuilding plan under which it has since been managed. A subsequent attempt, SEDAR 62, was halted near completion in 2019 because of data issues. Hence, SEDAR 100 is the first completed benchmark assessment for Gulf gray triggerfish in more than a decade.

Description of the Individual Reviewer's Role in the Review Activities

This reviewer independently read the documents provided for review as needed in preparation for this SEDAR 100 review for Gulf gray triggerfish, traveled, participated actively in physical review meetings in Tampa, FL (11-14 August 2026), reviewed each term of reference, identified key issues, suggested guidance, contributed to the joint panel summaries presented at the last day of the meeting, and independently authored this review report.

Findings by 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 (with gray highlighting) with corresponding reviewer comments following. The terms of reference below are reproduced verbatim, except that in the version submitted before the meeting ToR 2 was word for word the same as ToR 1. I inquired if that implied that a ToR was missing, but I was assured by the SEDAR coordinator that no ToR was missing. In the listing below I have removed the duplicate and shifted the following numbering accordingly.

ToR 1: Evaluation of the data used in the assessment

Evaluate the data used in the assessment, including discussion of the strengths and weaknesses of data sources and decisions. Consider the following:

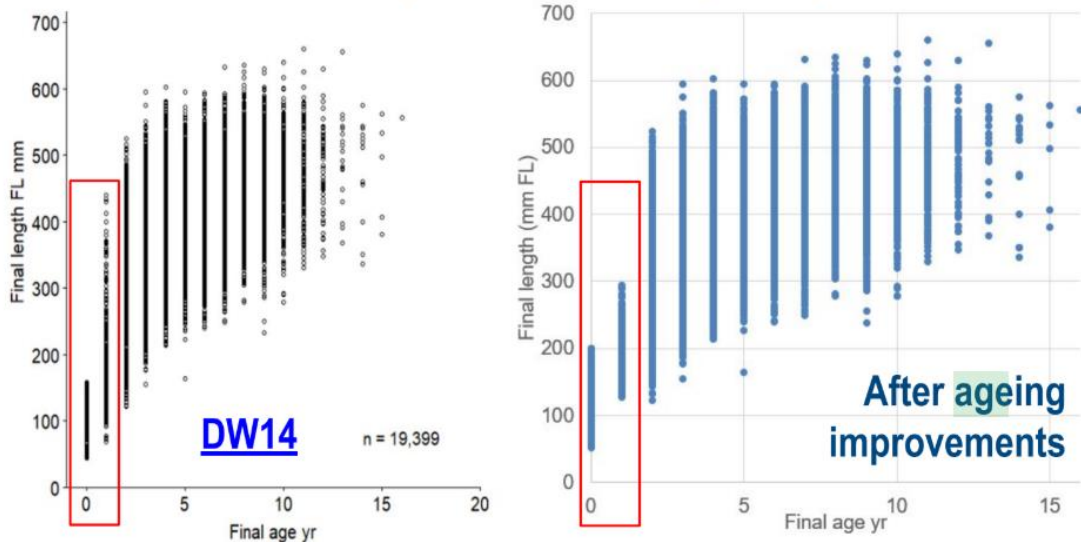
- *The context of the data availability in the region. Are the procedures appropriate for our data limitations?*
- *Are data decisions made by the Data Workshop (DW) and Assessment Workshop (AW) justified?*
- *Are data uncertainties acknowledged, reported, and within normal or expected levels?*
- *Is the appropriate model applied properly to the available data?*

- *Are input data series sufficient to support the assessment approach?*

The data used in the assessment have undergone massive revision. Data quality issues (especially concerning age determination) were the reason the previous SEDAR 62 was interrupted. An enormous amount of work has gone into this data improvement. It was surprising to see the magnitude of the changes in the data series, and it does seem reasonable to quietly wonder if the data processing has fully settled by now. Nevertheless, the data used in the assessment is sufficient to support the assessment approach, but some parts of that data are very indirectly observed (reconstructed), and hence the results should be carefully evaluated and trusted only conditionally on those data reconstructions. Hence, more emphasis should be given to the model diagnostics. Procedures are improved and appropriate and the decisions made by the DW and AW are justified, as assumptions are clearly stated. The large uncertainties are acknowledged including the increase in uncertainty back in time, but the potential biases in the reconstructed data parts are not included, except partly in sensitivity runs. The levels of uncertainty in the data for this assessment are large. The model configuration has been adjusted to rely less on age data, which makes the model more robust to problematic ageing.

Life history

Age data are the single largest improvement over SEDAR 43. All fish aged 5+ were re-aged under an improved spine protocol. The protocol is externally validated by bomb radiocarbon. In addition, age-0 and age-1 were also re-aged. This is an impressive amount of work. The maximum age (t_{max}) was maintained, but the assessment panel confirmed that it was still consistent with the re-assigned ages, which is important because the maximum age is used to re-define the fixed level of natural mortality as $M=5.4/t_{max}$ (Hamel and Cope 2022). The age-specific natural mortality values follow Lorenzen. The re-aging improved the age reading which is visible, especially at age 1, but the size-at-age remains extraordinarily variable (e.g., a 40cm fish can be anything from 2 to 15 years old) and the age data do not show any cohort signal.



Comparing length-at-age before (left) and after re-aging. Notice especially the improvement at age 1

An age-error matrix is included to acknowledge the still problematic aging, and the model is configured in a way to rely less on age observations (using conditional age-at-lengths and length-based selectivity for all fleets).

Growth is modelled as a single sex-combined von Bertalanffy curve. While sexual dimorphism may be expected it was not evident in the current data. Previous studies had shown sexual dimorphism, but those were based on the previous age readings. The asymptotic length L_{∞} recommendation has changed (dropped) from 87cm (SEDAR 9) to 59cm (SEDAR 43), to now 47cm, but the assessment panel chose to freely estimate all growth parameters and the estimate landed at around 54cm.

Maturity and fecundity were not updated. Maturity was switched from an age-based to a length-based ogive, a defensible response to the ageing problems, but it reuses SEDAR 43 parameters. Discard mortality from the recreational fleet was increased to 40% from a synthesis of recent studies. It was previously 5% in SEDAR 43, with 10% used as a sensitivity. The 40% value was also used for the commercial vertical line fleet given its similar fishing practices and no specific data.

Commercial fishery statistics

Commercial vertical line (handline/vertical line, with trap and “other” gears lumped in based on their similar size distributions) and commercial longline, which are kept separate because it captures much larger and older fish. The previous area-split was dropped, so for this assessment they are Gulf-wide. The data are derived from the accumulated landings system 1949-2024 and supplemented by state trip tickets where available (FL 1986+, LA 2000+, AL 2002+, TX 2014+, MS 2015+). There are no data before 1949, so the data are reconstructed/interpolated from 1945 to 1948 from an assumed catch of 0.01t in 1945. The reconstruction produced relatively minor adjustments compared to the SEDAR 43 values in recent years, but large changes from 1945-1960. Landings uncertainties are

assigned based on expert opinion borrowed from a different species. A Commercial Vertical Line CPUE index was developed, but excluded from the model, on the grounds that regulation changes make it unrepresentative.

Recreational fishery statistics

The recreational catches are aggregated into one fleet, which combines charter, private, and headboat, because they have similar selectivity. Similar to the commercial landings, the previous area-split was dropped. The source of the recreational catches is the Marine Recreational Information Program (MRIP) from 1981, but further back it is reconstructed from CPUE data from 1955-1980, and before that interpolated back to an assumed starting value of 1250 fish in 1945. Notice that this main data source is reconstructed for 36 out of 80 years. The recreational landings exceed commercial landings throughout the series. The difference in the recreational catch series from the SEDAR 43 series is large (often more 100%) even in the most recent years of SEDAR 43. Uncertainties were assigned from the MRIP estimation and set to a larger value for earlier years.

Shrimp bycatch

The shrimp bycatch fleet is unique. It represents the juvenile gray triggerfish caught and killed in the assessment, but the data for this fleet are not based on any observed juvenile gray triggerfish. Instead, two proxy species (Lane Snapper and Red Snapper) were selected (because they co-occur with gray triggerfish) in an analysis of SEAMAP trawl data filtered to approximate commercial shrimping grounds. Proxy bycatch from 1984 on was then estimated based on a previously approved method. Further back to 1960 the bycatch was extrapolated based on the relative trend in shrimp effort, and finally it was interpolated linearly down to an assumed value in 1947. With all of these assumptions in place, a comparison with other sources of removal shows that in most years (but not the twenties) more than half of the removals are from a fleet which was essentially never observed, or at least very indirectly observed. The use of this data source, and its effect on the estimated quantities, should be monitored closely.

Fishery-independent surveys

Two trawl surveys (SEAMAP early 1987-2008 and late 2009-2024) track juveniles and the combined video 1996-2024 tracks adults. All three supply abundance index data and length compositions. The selectivity from the SEAMAP early is also used for the shrimp bycatch. The combined video index is the only index for the adult part of the population, but it is included with two selection blocks (2006-2009 and 2020-2024).

ToR 2: Strengths and weaknesses of the assessment methods

Evaluate and discuss the strengths and weaknesses of the methods used to assess the stock, considering the available data. Consider the following:

- *Are methods scientifically sound and robust?*
- *Are priority modeling issues clearly stated and addressed?*
- *Are the methods appropriate for the available data?*

- *Are assessment models configured properly and used in a manner consistent with standard practices?*

The assessment for gray triggerfish is carried out using Stock Synthesis 3 (SS3), a stock assessment model that is among the most frequently applied in US fisheries and worldwide. As part of NOAA's modelling toolbox and with decades of application, SS3 is widely regarded as scientifically sound and robust. Modelling priorities are clearly stated and addressed by the assessment team. The model is overall configured in line with standard US practices. Remaining issues are more with regard to the quality of the data and how they were acquired and reconstructed. In this regard, the modelling framework lags behind other available software in one notable respect: SS3 lacks a more novel feature, namely random effects. Such a capability would be particularly useful for this assessment, as it would allow treating observations simply as "missing" rather than having to reconstruct them from very indirect observations when they are in fact missing.

Model framework and configuration

Stock Synthesis is an appropriate and well-tested framework. It is highly configurable and can scale from (almost) a simple biomass model to highly advanced age-based models with length observations, sex, regions, growth-platoons, and more. A further advantage is the accompanying toolbox, which facilitates the production of standardised, recognisable output and provides a comprehensive suite of diagnostics.

The model set up for Gulf gray triggerfish is a one sex, one area, age-based population model. Selectivity is length-based throughout, and growth is estimated internally from the length and conditional age-at-length compositions; natural mortality, maturity-at-length, fecundity, the length-weight relationship and the ageing error matrix are fixed inputs.

The observational data are as follows. Removals: recreational landings in numbers (1945-2024) and recreational discards in numbers (1981-2024); commercial landings in metric tons from vertical line (1945-2024) and longline (1979-2024), and commercial vertical line discards in numbers (2000-2024); and shrimp bycatch, entered not as an annual series but as a single median value assumed representative of 1946-2023, with its annual pattern supplied by a shrimp effort index. Abundance indices: from the SEAMAP trawl surveys, split into an early (1987-2008) and a late (2009-2024) series with separate catchabilities, and the combined video survey (1993-2024); plus one fishery-dependent series, headboat CPUE, truncated at 2007. Size compositions: annual length frequencies in 2 cm fork length bins from 2 to 74 cm, for the landings and discards of each fishing fleet and for the two trawl series and the video survey. Age compositions: conditional age-at-length from the three fishing fleets only (2003-2024, plus group at age 10); no survey contributes age data. Mean weight: the annual mean weight of recreational landings, fitted as its own likelihood component.

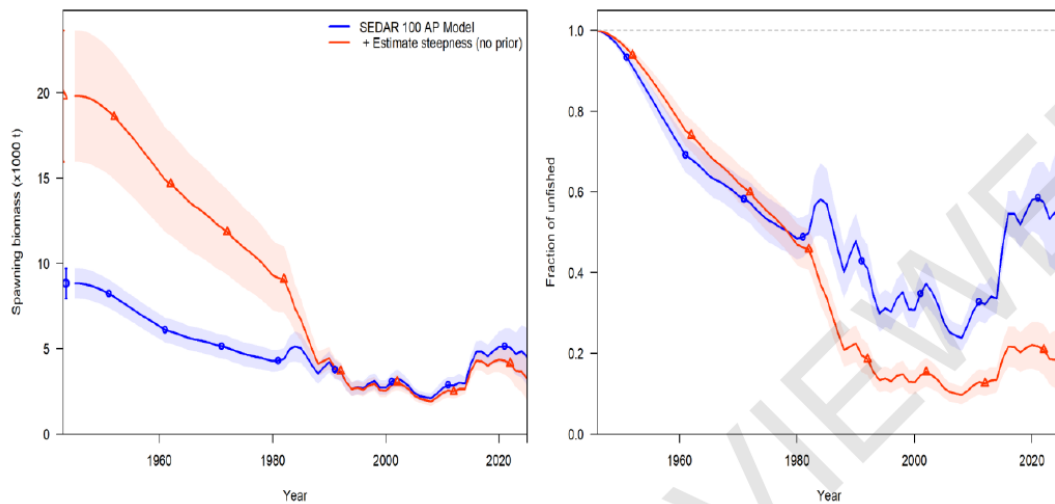
The model contains more than 260 model parameters, which are estimated, and additional parameters and penalties, which are fixed at arbitrary values. The model is complex. Based on the model diagnostics (described below) it borders on being more complex than what can be supported by the observations.

The reconstructed shrimp fleet discards are included in a bit of a complex manner. The annual pattern is not coming from the bycatch estimates, but from the shrimp effort series, which is included as an index assumed proportional to fishing mortality. The super-period catch (1,612,310 fish for 1946-2023) sets the level of removals for the period and the effort index sets the shape over time. This portion of the model has around 80 parameters, which are almost exclusively shaped by this shrimp fleet discard reconstruction. This shape matters, because there are no recruitment deviations before 1981 (because their estimation led to high CVs). Hence, for the first 36 of the model's 80 years, the entire population trajectory is a deterministic forward calculation that: a) starts from an equilibrium assumption, b) subtracts the assumed removals, and then c) replaces them via the deterministic stock-recruitment curve. Shrimp bycatch is more than 70% of the fish killed in every decade of that period. So, in the 36 years where the model has no data and no free parameters, three-quarters of the removals are supplied by a series whose annual shape is the shrimp effort index. Even more problematic, the effort index for 1945-1959 is itself interpolated, as a straight ramp, so the historical depletion trajectory is not estimated, but it is simply the scaled effort curve. The estimates for the more recent period is, luckily, fairly robust to the starting year, but all of this underlines that the inference from the first period should not be the foundation for reference points or management quantities.

The model is configured to use length-based selectivity throughout, which is a reasonable choice, because of the problematic age data. This makes the model less reliant on age data, particularly at older ages, where length distributions tend to be very similar.

Multiple parameters are included as "fixed at model estimate". This may seem like a pragmatic trick to reduce correlation among estimated parameters and stabilize estimation, but it removes those parameters from the Hessian, and hence from propagating the uncertainties forward to the estimates from the model. The result is that the final estimated standard deviations (also on the quantities of interest) become understated.

The model candidate presented by the assessment panel had steepness fixed at 0.99. This is a fairly common thing for assessments configured in Stock Synthesis. Stock Synthesis assumes a Beverton-Holt stock-recruitment function, and the ability to estimate the steepness parameter in this function depends on the contrast of the historic stock and recruitment pairs implied by the data in the model. In many cases it is not possible to estimate the steepness parameter. In these cases, a pragmatic choice is often to fix it near one (0.99), which implies that the stock-recruitment curve very rapidly increases and reaches a level where it plateaus, effectively acting as a hack to approximate a constant recruitment model regardless of stock size. The stock-recruitment model has consequences for the reconstruction of the historic time series and for MSY-based reference points. When steepness is fixed to 0.99 (which is a biologically unrealistic value), then reference points are typically calculated via SPR proxies instead.



The difference from estimating steepness on spawning stock biomass on unfished fraction

For Gulf gray triggerfish the model diagnostics (profile likelihoods) clearly indicated that the steepness could be estimated, and the different data sources were all consistent about the estimated value. This was skeptically evaluated by rerunning the diagnostics with shortened time series and without recreational compositions, but all investigations led to the same conclusion, which was that steepness could be reliably estimated. Furthermore, the model with steepness estimated had lower negative log-likelihood, almost fixed previous severe retrospective patterns, and removed a worrying tension in the recruitment deviations. (Previously it had been necessary to force the recruitment deviations to be centered around zero, and when the forcing was removed, they would switch to a different level, with the tension also showing up in the retrospective pattern for recruitment). The model with steepness estimated described the historic data better, and the resulting historic time series was also very different. Fixing steepness at 0.99 would more than double the historic estimates of SSB. The estimated steepness value, which fitted the historic period best, was around 0.42, which is also biologically unrealistic, so this value should also not be used as basis for MSY reference points. Alternatives will be considered under ToR 5 below.

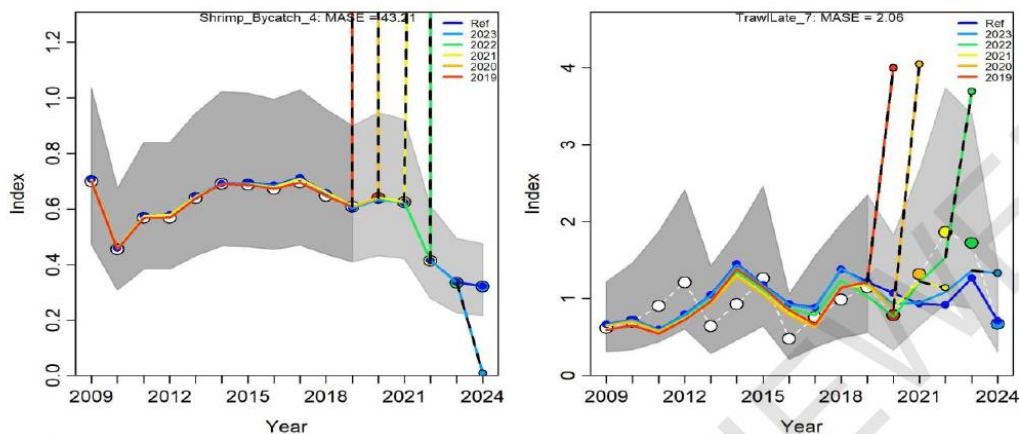
Further investigations were made into estimating the level of the natural mortality, either with fixed steepness or jointly with steepness, but the profiles were less clear for the natural mortality, and it was concluded that estimation of M was not stable enough to be recommended with the data available for Gulf gray triggerfish.

Diagnostics

The model diagnostics provided are comprehensive for the assessment model and include the following: Pearson residuals, leave-out runs, retrospective runs, jitter analysis, and profile likelihoods for key parameters. As with most practical assessment models, such model diagnostics cannot be expected to be perfect. The diagnostics are useful anyway to identify major problems and guide analysts to improved model configurations.

While not a diagnostic in itself, comparing the results to previous models for the same stock does provide some insight into the stability of the assessment model. For Gulf gray triggerfish the model results are very different from previous models. The difference originates primarily from the re-processing of the data entering the model, and a substantial amount of work was put into fixing previous known data issues, so differences are expected. Nevertheless, considering the differences from SEDAR 43 to SEDAR 100, it is reasonable to question whether the data processing has been fully settled by now.

One of the most concerning diagnostics to this reviewer is the jitter analysis. This is a purely technical investigation, which showed that the model is not converging consistently. This is concerning because it cannot be trusted that if a new year of observations is added and the model is re-run, then the model will converge to the global minimum. For the actual model presented, none of the 100 runs resulted in a more optimal (lower negative log-likelihood) solution, but only 42 runs returned the same solution. This means that no single future run can be trusted to have reached the global solution based on one set of initial values. Therefore, the jitter runs must be repeated before one can trust that a global solution has been found. This issue also affects uncertainty calculations based on the Hessian matrix, because this approach requires identifiable parameters. In short, the jitter analysis clearly indicates that the model should be simplified to be more robust. On the positive side, the resulting time series of management-relevant quantities showed similar trends, even at the different local optima.



Example of the Mean Absolute Scaled Error (MASE) diagnostics presented. Notice that the predictions are completely outside the model region

The Mean Absolute Scaled Error (MASE) diagnostics is, if correct, also very problematic. However, this reviewer does not trust that it is correctly specified. For a given time series, MASE is supposed to compare a model-based one-step prediction with the simple prediction rule of “next year is the same as this year.” A MASE value of 0.5 means that the model has half the prediction error of the simple rule; a MASE value of 1 means the two are equal; and a MASE value of 2 means that the simple rule is providing better predictions than the model (by a factor of 2). The MASE values and graphs (e.g., Fig. 99) provided for the base model indicated, in one case, that the model was 43 times worse than the simple rule. This reviewer cannot believe that to be the case. Furthermore, it is not clear from the

report or from discussions at the review meeting what exactly constitutes a “prediction” from the model in this setting. The model has several year-specific parameters, and to form a prediction, some assumption about these parameters must be made, but these assumptions are not specified.

ToR 3: Treatment of uncertainty

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

- *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.*
- *Comment on the relationship of this variability with ecosystem or climate factors and mechanisms for encompassing this into management reference points.*

The method used to evaluate uncertainties is the default in Stock Synthesis and the standard across many integrated assessments. It is recommended that the Stock Synthesis package be extended to include other standard methods, which may be better suited to capture uncertainties in such complex assessment models. The potential uncertainty originating from the extensive use of reconstructed data series is difficult to capture within the model, so it is important to consider sensitivity runs. The projections and management reference points are deterministic.

Coverage of uncertainty

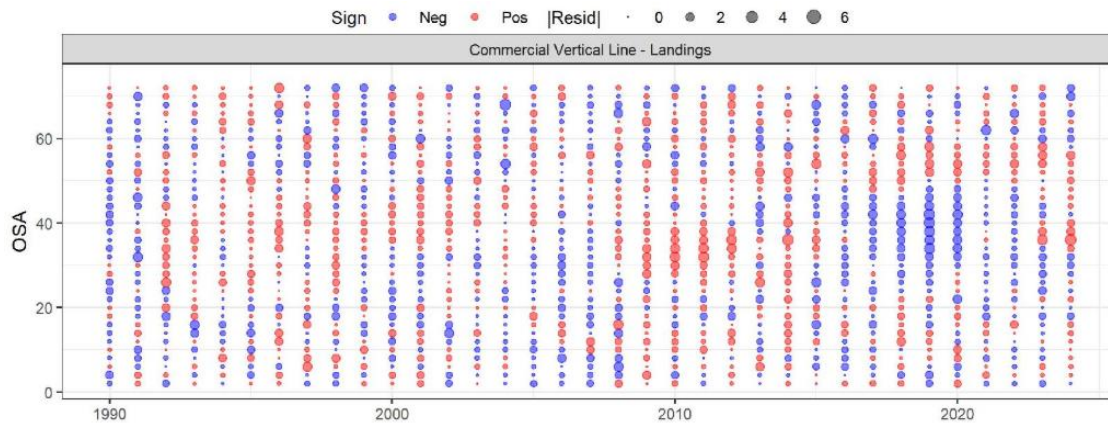
The assessment model almost exclusively quantifies the uncertainty using the asymptotic normality of the maximum likelihood estimates. The standard deviations are derived from the inverted Hessian matrix, and hence from a second-order Taylor approximation. There are a few things to note about this approach in this context.

The asymptotic normality of the maximum likelihood estimator is an asymptotic result, so it is only valid when the number of observations informing each parameter is “large”. Importantly, this is a frequentist property, meaning it describes the behaviour of the estimator across repeated hypothetical samples. The result hinges on regularity conditions, which include a well-behaved likelihood, identifiable parameters, and the true parameter vector being in the interior of the parameter space. It is not demonstrated that the asymptotic property is valid, and since there are year-specific parameters (recruitment deviations, F values for the recreational fleet, and F values for the shrimp bycatch), which increase in number every time a new year of data is added, it is implausible. Furthermore, the results of the jitter analysis (described above under ToR 2) indicate problems with the identifiability, as different initial values can result in different local minima. All of this suggests that the negative log-likelihood is not well approximated by a nice convex quadratic function.

Priors are included on many parameters (Dirichlet-multinomial parameters, recruitment deviations, and “weak” penalties to avoid bounds). While such priors are included for pragmatic reasons (e.g., to enhance stability) they do have consequences for the asymptotic result mentioned above. Technically, the estimator is no longer a maximum likelihood (ML)

estimator, but a maximum a posteriori (MAP) estimator. The MAP estimator is still asymptotically normal, but its variance is smaller than the variance of the purely data-informed ML estimator (in formulas: $\text{Var}(\theta_{\text{MAP}}) \approx (I(\theta) + P)^{-1}$, where I is the Fisher information and P is the inverse variance of the assumed prior). Note that the reported standard deviations are no longer purely frequentist standard deviations and they will always be smaller than the asymptotic standard deviations of the ML estimator (which are the square roots of the diagonal elements of $I(\theta_0)^{-1}$). Notice also, that the estimator is now asymptotically biased, as it converges to a point that is a compromise between the data and the prior (not to the true θ_0), but this bias shrinks at rate $1/n$ so it is asymptotically negligible.

Multiple key parameters are fixed (CV of young, CV of old, sigma R, steepness, M Lorenzen, maturity parameters, and more). Fixing a parameter, even at the model's own estimate, deletes a dimension from the Hessian matrix. Consequently, the estimated uncertainty, as measured by the inverse Hessian, does not account for variability in the fixed parameters and understates the true uncertainty, because the fixed parameters are not truly known. Such parameters are estimated or assumed elsewhere, and their uncertainty should propagate into the final estimates. The reported standard errors are conditional on the fixed values, and this may lead to overly narrow confidence intervals, and an underestimation of the true variability in derived quantities such as final SSB, F, and other management relevant quantities.



One-step-ahead (OSA) residuals for the compositions of the combined video survey requested by reviewers during meeting. Notice the clustering of red and blue dots indicating positive correlation

The assumed Dirichlet-multinomial does not describe the compositional observations. It is evident from the OSA residuals (see graph) that most compositions are positively correlated with neighboring length-classes, and the runs test rejects randomness. This is in contradiction with the negative correlations assumed in the Dirichlet-multinomial. Ignoring these positive correlations artificially increases the weight of the observations, as it overlooks the fact that each observation is partially explained by other observations. This effectively overstates the information content of the composition data.

Alternative methods

Alternatives to the asymptotic standard deviations, which are often used in assessments, are MCMC methods and simulation studies (a.k.a. parametric bootstrap or self-tests). MCMC techniques set up Markov chains with an equilibrium distribution, which is the posterior distribution of the joint parameter vector (assuming a uniform prior on previously prior-free parameters). This can be used to quantify the uncertainty, and does not rely on any asymptotic properties with respect to number of observations. It does require the Markov chains to be run for a long time, but that is a technical rather than a conceptual limitation. Full simulation studies are the “gold standard” for uncertainty quantification, where a large number of data sets are simulated forward from the model, then each simulated data set is used to estimate all parameters, and the distribution of the estimates are used to quantify the uncertainty. Again, this approach does not rely on any asymptotic properties with respect to number of observations. It is strongly recommended that these approaches are explored as part of a robust uncertainty quantification framework.

Sensitivities, profiles and predictive analysis

In addition to the model based asymptotic uncertainty estimates, a broad and well selected set of sensitivity runs was presented (including scenarios on recreational removals, discard mortality, bycatch in shrimp fishery, growth platoons, steepness, start year, and leaving out indices). Most sensitivity runs showed similar overall trends, but estimating steepness (and to a lesser extent using growth platoons) exhibited a considerable axis of uncertainty. Estimating steepness decreased recent SSB-ratio estimates (mainly by increasing absolute SSB in the past). Much of the review meeting was dedicated to exploring these sensitivities further.

Profile likelihoods were investigated for key fixed parameters, and for steepness the profile appeared to contradict fixing it at 0.99. The context is that it is often not possible to estimate steepness, and hence it has become standard practice to fix it at 0.99, which effectively means that the stock-recruitment relationship is just a flat horizontal line. This is not realistic, so when this fix is in effect, then reference points are determined via SPR proxies, which do not rely on the stock-recruitment relationship. For gray triggerfish, however, it appeared possible to estimate steepness. All the partial profiles (for different data parts) agreed that steepness could be estimated, and the overall estimate was around 0.42. To further validate this conclusion the data series were started later in two steps (to investigate if the ability to estimate was an artifact of earlier uncertain/reconstructed data series). Further, it was investigated if excluding the composition data from the recreational observations changed the conclusion. However, the estimation of steepness was robust to all these experiments. Furthermore, the model with estimated steepness improved the overall model fit, improved the retrospective pattern for recruitment, and resolved an important issue with tension in the model between estimated recruitment deviations and the early period where recruitment is constant. The review meeting concluded that the model where steepness is estimated is the preferred model candidate.

The projections and management reference points are deterministic, and do not directly include ecosystem or climate factors. It is listed as a research recommendation from the assessment panel to improve the understanding of how sargassum affects Gulf gray triggerfish recruitment, and to build such an index into the assessment model and model consistent projections.

ToR 4: Recommendations to improve the assessment

Provide, or comment on, recommendations to improve the assessment.

- *Consider the research recommendations provided by the Data and Assessment workshops in the context of overall improvement to the assessment and make any additional research recommendations warranted.*
- *If applicable, provide recommendations for improvement or for addressing any inadequacies identified in the data or assessment modeling. Describe these recommendations in sufficient detail for application and should be practical for short term implementation (e.g., achievable within ~6 months). List longer-term recommendations as research recommendations above.*

Recommendations from the data and assessment workshops

The research recommendations from the data and assessment workshops are valid and relevant, but were not prioritized.

Of the recommendations from the data workshop (listed in appendix D), the recommendation to add Gulf gray triggerfish to the list of species for which data are currently collected in the shrimp observer program seems to be the most important. This would provide real observations of this substantial (and in some periods the main) source of gray triggerfish removals. Relying solely on indirect proxy information from other species for these removals is potentially a major source of uncertainty and bias in the assessment of Gulf gray triggerfish.

Of the recommendations from the assessment workshop (listed in appendix E), this reviewer would again give highest priority to any effort to improve the shrimp bycatch estimates. This important source of removals needs to be estimated more directly (or truly treated as unknown). After this, more direct approaches to estimate natural mortality should be prioritized, and the efforts to improve the combined video survey are also important.

Long term research recommendations

The data situation for this assessment is difficult, because only very indirect sources of information are available for major removal sources (especially back in time). It further complicates things that the management domain is centered around virgin biomass derived quantities, which require extrapolation back in time (or an assumption about the initial depletion level if the assessment is started later). One tool that would be useful in this situation (and many others) is to **allow random effects in the assessment model**. A random effect in a statistical model is a quantity, which has a distribution, but is not observed. So, this first case is exactly a missing observation and casting the removals from the shrimp fleet as random effects would make sense, and doing so would acknowledge the

uncertainty which comes from not having observed it. The likelihood would not change. The stock dynamics would still predict the catch, the same distribution would be assumed, and a variance parameter would be assigned. The model would work by integrating over all possible values of the catch and then providing a predicted catch. Random effects would also be the natural framework to handle recruitment deviations. Again, the same model formulation would apply as is currently used, but the standard deviation for the recruit deviations σ_R would be estimable, it would not be necessary to assign periods where the deviations are fixed to zero, and projections would naturally use the estimated σ_R to describe the distribution of future recruits.

The use of random effects could be further expanded to express year-specific quantities as random processes described by a few parameters. For example, the yearly log-scale fishing mortality of a fleet could be expressed as a random walk, defined by just two parameters. Rolling out such approaches for all year-specific quantities would greatly simplify the model. The model would no longer require new parameters each time data for a new year are added, and the estimated process parameters could be used for model-consistent projections. Such approaches are already used in state-space assessment models such as SAM (State-space Assessment Model), WHAM (Woods Hole Assessment Model), SPoRC (Stochastic Population over Regional Components), and NCAM (Northern Cod Assessment Model), and they seem even more relevant here.

The next research recommendation is to **set up a simulation study**. This should likely be a short-term recommendation, but the fallout could take some time, so it is listed as a long-term recommendation. A self-simulation (simulation study) is important in order to validate all the technical aspects of the assessment model and should always be recommended. In this case the issues seen with the jitter analysis make it even more important to validate that the model gives consistent estimates, and that the asymptotic inference is valid. In case the asymptotic inference is not valid, then the simulation approach itself is the way to quantify the uncertainties.

Short term recommendations

Estimate steepness. The model version with steepness fixed at 0.99 forced a low estimate of virgin biomass, and the lack of data in the early part of the assessment gave the impression of a low historical biomass. Steepness is consistently estimable and improves model diagnostics when estimated.

Adjust projections and status determination criteria to not depend on the stock-recruitment curve (either one). A suggestion is described below (ToR 5).

Study the convergence. The jitter analysis is concerning, so understanding why the model fails to find the same solution is important. It may be helpful to study the estimated correlation between parameters or profile potentially problematic parameters. Simplifying the model to obtain robust and consistent estimates would be the optimal solution.

Study the one-step prediction ability. The failure of the MASE diagnostic should be investigated. It is most likely simply misleading, but it is worth checking.

Down-weight the compositions. The composition (OSA) residuals clearly show positive correlations. This means that the model is not accounting for the positive correlation between observations and is therefore giving them more weight than they should receive. This is an expected problem, and for assessments configured in Stock Synthesis, the standard approach is to down-weight these components (e.g., using variance inflation factors or effective sample size adjustments).

ToR 5: Evaluation of the stock projections

Evaluate the stock projections, including discussing strengths and weaknesses, and consider the following:

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

The projection method described is consistent with standard and accepted practice for assessments configured in Stock Synthesis. The review identified that fixing steepness at 0.99 was not consistent with the observations and recommended estimating steepness instead. This has consequences for the projections. If the estimated stock-recruitment relationship is trusted, then that should form the basis for projections and reference points. For Gulf gray triggerfish the estimated steepness value of 0.42 was judged to be unrealistically low for the species. It was concluded that the stock-recruitment was useful for the historic estimation, but not for projections or equilibrium considerations. In this situation, appropriate projection and status determination methods must be formulated without relying on the stock-recruitment function. It is recommended to use recent well-informed empirical recruitment estimates as the basis for short-term projections and SPR-based quantities for status determination. The assessment includes uncertainties (with some parameters fixed, see ToR 3 above), but the standard projection results are deterministic.

Methods

FMSY is the fishing mortality rate that maintains the population at a level where it can produce its Maximum Sustainable Yield. Estimating it requires a reliable understanding of the stock-recruitment relationship, which is frequently difficult to obtain. For assessments defined in Stock Synthesis this often leads to fixing the steepness parameter to 0.99. When recruitment has effectively been set constant at all stock sizes (which is unrealistic), the F30%SPR proxy is often used as a practical alternative. The specific percentage can vary and a more conservative approach might use F40%SPR as the proxy instead. The choice of which proxy to use often depends on the life history of the stock and the level of risk managers are willing to accept.

For Gulf gray triggerfish the observations in the historic period did not support fixing steepness at 0.99 and allowed it to be consistently estimated. This gave a better fit to the observations and broadly improved the model diagnostic. The fact that the estimated

stock-recruitment curve improved fitting data from the historic period is, however, not the same as trusting the estimated stock-recruitment function for future projections. The historic stock-recruitment pairs only span a fairly narrow stock-size (SSB) range, and using this estimated value for projections means extrapolating outside the range.

In this situation it is recommended to base short-term projections on the empirical recent recruitment distribution rather than the stock-recruitment relationship (which is not trusted at either steepness value). Recruitment should be drawn from a well-justified recent period (e.g., recent 10 year window to represent recent conditions), including only sufficiently estimated cohorts, and retaining observed variability. The averaging period should be selected using an objective criterion. This approach avoids extrapolating the fitted stock-recruitment curve outside the historically observed SSB range. Projections should ideally also incorporate stochastic uncertainty, present alternative recruitment scenarios, and report projection intervals.

SPR_0 (the unfished spawner per recruit), SPR_F (spawner per recruit at fishing rate F), and F_X (the fishing rate F at which SPR_F/SPR_0 is X , written for example as $F_{30\%SPR}$) are quantities computed from biology and selectivity, so computing them does not rely on knowing the stock-recruitment relationship. The maximum fishing mortality threshold can be defined as $MFMT = F_X$, and the target biomass as $B_X = X \cdot SPR_0 \cdot \bar{R}$. The average recruitment $\bar{R}$ should be defined over a recent period (to reflect recent conditions) with well informed recruitment estimates. The minimum stock size threshold is defined as $MSST = 0.5 \cdot B_X$. These proxies can be defined in the situation where the stock-recruitment relation is not judged to be trustworthy.

ToR 6: Review Workshop Summary Report

Prepare a Review Workshop Summary Report describing the Panel's evaluation of the stock assessment and addressing each Term of Reference.

This reviewer contributed to the panel's summary report (drafted the response to ToR 3) and read and discussed all remaining sections.

Appendix A: Bibliography of materials provided for review via sedarweb.org

Document	Title
SEDAR100-SAR Section I	Introduction
SEDAR100-SAR Section II	Data Workshop Report
SEDAR100-SAR Section III	Assessment Report
SEDAR100-SAR Section IV	Research Recommendations
SEDAR100-DW-01 to DW-22	Data Workshop working papers
SEDAR100-AP-01	Gray Triggerfish Post-Release Mortality
SEDAR100-RD01 to RD07	Reference documents

Appendix B: Performance Work Statement

Performance Work Statement (PWS)
National Oceanic and Atmospheric Administration (NOAA)
NOAA Fisheries
Center for Independent Experts (CIE) Program
External Independent Peer Review
Performance Work Statement (PWS)
National Oceanic and Atmospheric Administration (NOAA)
NOAA Fisheries
Center for Independent Experts (CIE) Program
External Independent Peer Review
SEDAR 100 Gulf Gray Triggerfish Assessment

August 11 - 14, 2026

Background

NOAA Fisheries is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NOAA Fisheries science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination. Specifically, science products that the agency can reasonably determine that will have, when disseminated, "*a clear and substantial impact on important public policies or private sector decisions.*" Additionally, peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards¹.

Scope

The **SouthEast Data, Assessment, and Review (SEDAR)** is the cooperative process by which stock assessment projects are conducted in NOAA Fisheries' Southeast Region. SEDAR was initiated to improve planning and coordination of stock assessment activities and to improve the quality and reliability of assessments.

¹ https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2005/m05-03.pdf

The SEDAR 100 review workshop will be a CIE assessment review of the stock assessment of Gulf gray triggerfish. The review workshop provides an independent peer review of SEDAR stock assessments. The term review is applied broadly, as the review panel may request additional analyses, error corrections and additional runs of the assessment models provided by the assessment panel. The review panel is ultimately responsible for providing advice to ensure the assessment is appropriate for use by fishery managers.

Requirements

NOAA Fisheries requests three reviewers to conduct an impartial and independent peer review in accordance with this Performance Work Statement (PWS), OMB Guidelines, and the ToRs below. Modifications to this Performance Work Statement (PWS) and ToRs cannot be made during the peer review, and the Contracting Officer's Representative (COR) and the CIE contractor shall approve any modifications prior to the peer review. All ToRs must be addressed in each reviewer's report.

The reviewers shall have a working knowledge of stock assessment modeling, statistics, fisheries science, and marine biology sufficient to complete the primary task of providing peer-review advice in compliance with the workshop Terms of Reference fisheries stock assessment. Expertise with Stock Synthesis modeling platform, and the associated model diagnostics would be helpful. Familiarity with evaluating length-and-age composition data, selectivity patterns, and model diagnostics in data-moderate or data-limited contexts is preferred.

The chair, who is in addition to the panel of three reviewers, will not be provided by the CIE. Although the chair will be participating in this review, the chair's participation (e.g., labor and travel) is not covered by this contract.

The specified format and contents of the individual peer review report are found in **Annex 1**. The Terms of Reference (ToRs) of the peer and review are included in **Annex 2**. The tentative agenda of this virtual panel review and process review meetings is included in **Annex 3**.

Tasks for Reviewers

Each CIE reviewer shall complete the following tasks in accordance with the PWS and Schedule of Milestones and Deliverables herein.

Pre-review Background Documents: Two weeks before the peer review, the NOAA Fisheries Project Contact will email the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NOAA Fisheries Project Contact will consult with the CIE on where to send documents. The CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance with the PWS scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

Documents to be provided to the CIE reviewers prior to this review include:

S100 Data Workshop Report (<https://sedarweb.org/documents/sedar-100-gulf-gray-triggerfish-data-workshop-report/>)

S100 Assessment Process Report (to be provided two weeks prior to review)

S100 Working papers and Reference documents (<https://sedarweb.org/assessments/sedar-100-gulf-triggerfish/>)

Panel Review Meeting: The CIE reviewers shall conduct the independent peer review in accordance with the PWS and ToRs and shall not serve in any other role unless specified herein. The CIE reviewers shall attend and actively participate in a professional and respectful manner as members of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein. The NOAA Fisheries Project Contact is responsible for any meeting arrangements (e.g., conference room reservations, video conferencing logistics, etc.).

Chair Summary Report: The reviewers shall assist the Chair of the meeting with contributions to the summary report.

Contract Deliverables - Independent CIE Peer Review Report: The CIE reviewers shall complete an independent peer review report in accordance with this PWS. Reviewers are not required to reach a consensus. The CIE reviewer shall complete the independent peer review according to required format and content as described in **Annex 1**. The CIE reviewer shall complete the independent peer review addressing each ToR as described in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Foreign National Security Clearance

When reviewers participate during a panel review meeting at a government facility, the NOAA Fisheries Project Contact is responsible for obtaining the Foreign National Security Clearance approval for reviewers who are non-US citizens. For this reason, the reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the Project Contact for the purpose of their security clearance, and this information shall be submitted at least two weeks in advance. For additional information, please see the following link: <https://www.commerce.gov/osy/programs/foreign-access-management>. All CIE reviewers must have a REAL ID-compliant form of identification to access Federally owned or leased facilities. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance

The place of performance shall be Tampa, FL.

Period of Performance

The CIE reviewers' duties shall not exceed **14** days to complete all required tasks.

Schedule of Milestones and Deliverables

The contractor shall complete the tasks and deliverables in accordance with the following schedule:

Within two weeks of award	Contractor secures participation of the CIE reviewers.
At least two weeks prior to the panel review meeting	Pre-review documents are provided to the reviewers.
August 11 - 14, 2026	The reviewers participate and conduct an independent peer review during the panel review meeting.
Approximately 2 weeks later	Contractor receives draft reports
Within 3 weeks of receiving draft reports	Contractor submits final reports to the Government

* The Peer Review Summary Report will not be submitted to, reviewed, or approved by the Contractor.

Applicable Performance Standards

The acceptance of the contract deliverables shall be based on three performance standards: (1) The reports shall be completed in accordance with the required formatting and content in **Annex 1**; (2) The reports shall address each ToR as specified **Annex 2**; and (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel

All travel expenses shall be reimbursable in accordance with Federal Travel Regulations ([Travel resources | GSA](#)), and all contractor travel must be approved by the COR prior to the actual travel. Any travel conducted prior to the receipt of proper written authorization from the COR will be done at the Contractor's own risk and expense. International travel is authorized for this contract. Travel is not to exceed the agreed upon price in accordance with the Price Schedule.

Confidentiality and Data Privacy

This contract may require that services contractors have access to Privacy Information. Services contractors are responsible for maintaining the confidentiality of all subjects and materials and may be required to sign and adhere to a Non-disclosure Agreement (NDA).

NOAA Fisheries Project Contact:

Katie Siegfried, Branch Chief, Gulf Fisheries Branch
NOAA Fisheries Service, Southeast Fisheries Science Center
101 Pivers Island Rd, Beaufort, NC 28516
Kate.Siegfried@noaa.gov

Annex 1: Peer Review Report Requirements

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations as specified in the ToR.
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.
 - a. Reviewer should describe in their own words the review activities completed during the panel review meeting, including a brief summary of findings, of the science, conclusions, and recommendations.
 - b. Reviewer should discuss their independent views on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.
 - c. Reviewer should elaborate on any points raised in the Summary Report that they feel might require further clarification.
 - d. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the panel summary report. The CIE independent report shall be an independent peer review of each ToRs and shall not simply repeat the contents of the summary report.
3. The reviewer report shall include the following appendices:
 - Appendix 1: Bibliography of materials provided for review
 - Appendix 2: A copy of the CIE Performance Work Statement
 - Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Annex 2: Terms of Reference for the Peer Review

SEDAR 100 Gulf Gray Triggerfish Assessment

CIE reviewers are contracted to complete their independent peer review based on the ToRs. Therefore, the CIE-NOAA Fisheries review and approval process is based on whether the CIE independent reports addressed each of the ToRs.

1. Evaluate the data used in the assessment, including discussion of the strengths and weaknesses of data sources and decisions. Consider the following:
 - The context of the data availability in the region. Are the procedures appropriate for our data limitations?
 - Are data decisions made by the Data Workshop (DW) justified?
 - Are data uncertainties acknowledged, reported, and within normal or expected levels?
 - Is the appropriate model applied properly to the available data?
 - Are input data series sufficient to support the assessment approach?
2. Evaluate the data used in the assessment, including discussion of the strengths and weaknesses of data sources and decisions. Consider the following:
 - The context of the data availability in the region. Are the procedures appropriate for our data limitations?
 - Are data decisions made by the DW justified?
 - Are data uncertainties acknowledged, reported, and within normal or expected levels?
 - Is the appropriate model applied properly to the available data?
 - Are input data series sufficient to support the assessment approach?
3. Evaluate and discuss the strengths and weaknesses of the methods used to assess the stock, considering the available data. Consider the following:
 - Are methods scientifically sound and robust?
 - Are priority modeling issues clearly stated and addressed?
 - Are the methods appropriate for the available data?
 - Are assessment models configured properly and used in a manner consistent with standard practices?
4. Consider how uncertainties are addressed in the assessment, and their potential consequences.
 - 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.
 - Comment on the relationship of this variability with ecosystem or climate factors and mechanisms for encompassing this into management reference points.
5. Provide, or comment on, recommendations to improve the assessment.
 - Consider the research recommendations provided by the Data Workshop in the context of overall improvement to the assessment and make any additional research recommendations warranted.
 - If applicable, provide recommendations for improvement or for addressing any inadequacies identified in the data or assessment modeling. Describe these

recommendations in sufficient detail for application and should be practical for short term implementation (e.g., achievable within ~6 months). List longer-term recommendations as research recommendations above.

6. Evaluate the stock projections, including discussing strengths and weaknesses, and consider the following:
 - Are the methods consistent with accepted practices and available data?
 - Are the methods appropriate for the assessment model and outputs?
 - Are the results informative and robust, and are they useful to support inferences of probable future conditions?
 - Are key uncertainties acknowledged, discussed, and reflected in the projection results?
7. Prepare a Review Workshop Summary Report describing the Panel's evaluation of the stock assessment and addressing each Term of Reference.

Reviewers must provide details on the basis of their responses to the review questions.

Annex 3: Tentative Agenda for SEDAR 100 Gulf Gray Triggerfish

August 11 – 14, 2026

Tampa, Florida

NOAA Fisheries Project Contact: Katie Siegfried

Tuesday:

8:30 a.m.	Introductions and Opening Remarks - Agenda Review, TOR, Task Assignments	Coordinator
9:00 a.m. – 12:00 p.m.	Assessment Presentations - Assessment Data & Methods - Identify additional analyses, sensitivities, corrections	Analytic Team
12:00 p.m. – 1:30 p.m.	Lunch Break	
1:30 p.m. – 4:30 p.m.	Assessment Presentations (continued) - Assessment Data & Methods - Identify additional analyses, sensitivities, corrections	Analytic Team
4:30 p.m. – 5:30 p.m.	ToR Review and Daily wrap up	Chair
5:30 p.m. – 6:00 p.m.	Public comment	Chair

Tuesday Goals: Initial presentations completed, sensitivity and base model discussion begun

Wednesday:

8:30 a.m. – 12: p.m.	Panel Discussion - Assessment Data & Methods - Identify additional analyses, sensitivities, corrections	Chair
12:00 p.m. – 1:30 p.m.	Lunch Break	
1:30 p.m. – 5:00 p.m.	Panel Discussion/Panel Work Session - Continue deliberations - Review additional analyses - Recommendations and comments	Chair
5:00 p.m. – 5:30 p.m.	ToR Review and Daily wrap up	Chair
5:30 p.m. – 6:00 p.m.	Public comment	Chair

Wednesday Goals: sensitivities and modifications identified, preferred models selected, projection approaches approved, Report drafts begun

Thursday

9:00 a.m. – 12: p.m.	Panel Discussion - Assessment Data & Methods - Identify additional analyses, sensitivities, corrections	Chair
12:00 p.m. – 1:30 p.m.	Lunch Break	
1:30 p.m. – 4:30 p.m.	Panel Discussion/Panel Work Session - Continue deliberations	Chair

	- Review additional analyses	
	- Recommendations and comments	
4:30 p.m. – 5:00 p.m.	ToR Review and Daily wrap up	Chair
5:30 p.m. – 6:00 p.m.	Public comment	Chair
Thursday Goals: sensitivities and modifications identified, preferred models selected, projection analysis reviewed, Report draft continued		
<u>Friday</u>		
8:30 a.m. – 11:30 a.m.	Panel Discussion or Work Session	Chair
	- Final sensitivities reviewed.	
	- Projections reviewed.	Chair
	- Review Reports	
11:30 a.m. – 12:00 p.m.	Public comment	Chair
12:00 p.m.		ADJOURN
Friday Goals: Complete assessment work and discussions, final base configuration available. Draft Reports reviewed.		

Appendix C: Panel Membership or other pertinent information from the panel review meeting.

Workshop time and place

The SEDAR 100 Gulf Gray Triggerfish Review Workshop was held in Tampa, FL August 11-14, 2026.

List of participants

Review Panel

Name	Affiliation
Luiz Barbieri (Chair)	GMFC SSC
Jie Cao	GMFC SSC
Yong Chen	CIE Reviewer
Nicholas Fisch	CIE Reviewer
Erik Lang	GMFC Appointee
Anders Nielsen	CIE Reviewer

Analytic Team

Name	Affiliation
Skyler Sagarese	SEFSC
Katie Siegfried	SEFSC

Staff

Name	Affiliation
Julie A Neer	SEDAR
Alyssa Clift	GMFC Staff
Lisa Hollensead	GMFC Staff
Matt Freeman	GMFC Staff
John Froeschke	GMFC Staff
Jesscia Matos	GMFC Staff
Ryan Rindone	GMFC Staff
Carrie Simmons	GMFC Staff

Council Representative

Name	Affiliation
Kevin Anson	Gulf Council

Workshop Observers via Webinar

Name	Affiliation
------	-------------

Lisa Ailloud	SEFSC
Shanae Allen	USF/FWC-FWRI
Jordan Bajema	FWC-FWRI
Anessa Carter	FWC-FWRI
Rob Cheshire	SEFSC
Manuel Coffill-Rivera	
LaTreese Denson	SEFSC
Leonardo Eguia	GFMC Staff
Alena Figueroa	FWC-FWRI
Francesca Forrestal	SEFSC
Sarah Gardiner	GFMC Staff
Challen Hyman	USF
Tiffany Hopper	TXPWD
Emily Ott	SEDAR
Chloe Ramsay	FWC-FWRI
Seth Leimgruber	FWC-FWRI
Molly Stevens	SEFSC
James Swarthout	AL DCNR
Nathan Vaughan	Vaughan Analytics
Kym Walsh	LAWLF

Appendix D: Research recommendations from the data workshop

- Adding Gray Triggerfish to the list of species for which data are currently collected in the Shrimp Observer Program.
- Outfitting the commercial shrimp fleet with modern, tamper-proof GPS tracklog devices.
- Applying electronic technologies to enable observers to collect size information on all fish species captured in fish trawls.
- Modifying the allocation of observer sampling effort to better incorporate net configuration and diurnal period.
- The potential for field experiments to improve calibrations of SEAMAP and commercial shrimp trawls.

Appendix E: Research recommendations from the assessment workshop

Natural mortality

- Explore more direct approaches to estimating natural mortality, for example mark-recapture approaches (conventional, telemetry, or close-kin).
- Explore ways to better reflect uncertainty around the mortality-at-age vector.

Reproduction

- Continue data collection for maturity and fecundity.
- Update maturity-at-age estimates to see whether the age data updates have changed our understanding of maturity.

Discard mortality

- Continue data collection from observer programs or electronic monitoring programs (e.g. Roberts and Neidig 2019 [SEDAR62-WP10]), particularly to get a better handle on discard mortality for the commercial fleets.

Shrimp bycatch

- Estimates of shrimp bycatch are highly uncertain and need further investigation once observer data allow more advanced approaches for estimation.

Selectivity and catchability

- Further investigate and quantify changes in selectivity and catchability through time to improve fit to the length and age composition.

Surveys

- Increase collection of length and age information for compositions.
- Continue to improve upon standardized indices from the video survey to account for survey design changes through time and space to ensure trends are representative of relative abundance.

Sargassum

- Improve our understanding of how sargassum affects Gray Triggerfish recruitment and which life stages are most affected.
- Develop indices (with uncertainty estimates) to enable hypothesis testing of environmental linkages within the assessment model.