SEDAR

Southeast Data, Assessment, and Review

SEDAR 41

South Atlantic Gray Triggerfish

SECTION V: Review Workshop Report

April 2016

SEDAR
4055 Faber Place Drive, Suite 201
North Charleston, SC 29405

Table of Contents

1. Introduction 3
1.1 Workshop Time and Place 3
1.2 Terms of Reference 3
1.3 List of Participants 5
1.4 Document List 7
2. Review Panel Report 14
Executive Summary 14
2.1 Statements Addressing Each ToR 14
2.2 Summary results of Analytical Requests 29

1. Introduction

1.1 Workshop Time and Place

The SEDAR 41 Review Workshop for South Atlantic Gray Triggerfish (*Balistes capriscus*) was held March 15-18, 2016 in North Charleston, SC. Review Panel members were presented all information generated throughout the Data (DW) and Assessment (AW) Workshops and webinars, and the Review Workshop (RW) Panel then developed a consensus review and analysis of the stock assessment model and inputs according to a number of SEDAR Terms of Reference.

1.2 Terms of Reference

1. Evaluate the data used in the assessment, including discussion of the strengths and weaknesses of data sources and decisions, and consider the following:
    a) Are data decisions made by the DW and AW sound and robust?
    b) Are data uncertainties acknowledged, reported, and within the normal or expected levels?
    c) Are data properly applied within the assessment model?
    d) Are data input series reliable and sufficient to support the assessment approach and findings?

2. Evaluate and discuss the strengths and weaknesses of the methods used to assess the stock, taking into account the available data, and consider the following:
    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 and consider 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 to 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, including discussing the strengths and weaknesses, and consider 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 are they useful to support inferences of probably 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.
    a) 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.
    b) Ensure that the implications of uncertainty in technical conclusions are clearly stated.

6. Consider the research recommendations provided by the Data and Assessment workshops and make any additional recommendations or prioritizations warranted.
    a) Clearly denote research and monitoring that could improve the reliability of, and information provided by, future assessments.
    b) Provide recommendations on possible ways to improve the SEDAR process.

7. Consider whether the stock assessment constitutes the best scientific information available using the following criteria as appropriate: relevance, inclusiveness, objectivity, transparency, timeliness, verification, validation, and peer review of fishery management information.

8. Compare and contrast assessment uncertainties between the Gulf of Mexico and South Atlantic stocks.

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

10. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference. Develop a list of tasks to be completed following the workshop. Complete and submit the Peer Review Summary Report in accordance with the project guidelines.

1.3 List of Participants

REVIEW WORKSHOP PANELISTS

Luiz Barbieri	Review Panel Chair	SAFMC SSC
Mike Armstrong	Reviewer	CIE
Jon Helge Volstad	Reviewer	CIE
Stephen Smith	Reviewer	CIE
Steve Cadrin	Reviewer	SAFMC SSC
Churchill Grimes	Reviewer	SAFMC SSC

ANALYTICAL REPRESENTATIVES

Kevin Craig	Lead Analyst, GTF	SEFSC Beaufort
Kate Siegfried	Lead Analyst, RS	SEFSC Beaufort
Kyle Shertzer	Assessment Team	SEFSC Beaufort
Erik Williams	Assessment Team	SEFSC Beaufort
Rob Cheshire	Assessment Team	SEFSC Beaufort
Eric Fitzpatrick	Assessment Team	SEFSC Beaufort

APPOINTED OBSERVERS

Rusty Hudson	Recreational/Commercial	FL / SFA
Robert Johnson	For-Hire	FL

APPOINTED COUNCIL REPRESENTATIVES

Zack Bowen	Council Member	SAFMC
Mark Brown	Council Member	SAFMC
Chris Conklin	Council Member	SAFMC

COUNCIL AND AGENCY STAFF

Julia Byrd	Coordinator	SEDAR
Julie O'Dell	Admin	SEDAR / SAFMC
Chip Collier	Fishery Biologist	SAMFC
Mike Errigo	Fishery Biologist	SAFMC
Nick Farmer	Fishery Biologist	SERO

WORKSHOP ATTENDEES

Joey Ballenger, SCDNR
Peter Barile, SFA
Myra Brouwer, SAFMC
John Carmichael, SAFMC
Brian Cheuvront, SAFMC
Lora Clarke, PEW

Amy Dukes, SCDNR
Jimmy Hull, FL fisherman
Julie Neer, SAFMC
Adam Nelson, FL fisherman
David Nelson, FL fisherman
Michael Nelson, FL fisherman
Paul Nelson, FL fisherman
Marcel Reichert, SCDNR
Tracey Smart, SCDNR

*Appointees marked with a * were appointed to the workshop panel but did not attend the workshop.

1.4 Document List

SEDAR 41 review workshop working papers and reference documents.

Document #	Title	Authors
Documents Prepared for the Review Workshop		
SEDAR41-RW01	Addendum to SEDAR41-DW16: Report on Life History of South Atlantic Gray Triggerfish, *Balistes capriscus*, from Fishery-Independent Sources: UPDATE on analyses of maturity, spawning fraction, and sex ratio	Kolmos et al. 2016
SEDAR41-RW02	Age structured production model (ASPM) for U.S. South Atlantic Red Snapper (*Lutjanus campechanus*)	SFB-NMFS 2016
SEDAR41-RW03	Age structured production model (ASPM) for U.S. South Atlantic Gray Triggerfish (*Balistes capriscus*)	SFB-NMFS 2016
SEDAR41-RW04	Red Snapper: Additional BAM diagnostics, analyses, and code	SFB-NMFS 2016
SEDAR41-RW05	Model Diagnostics and Source Code for SEDAR 41 Gray Triggerfish (*Balistes capriscus*) Benchmark Stock Assessment	SFB-NMFS 2016
Reference Documents		
SEDAR41-RD01	List of documents and working papers for SEDAR 32 (South Atlantic Blueline Tilefish and Gray Triggerfish) – all documents available on the SEDAR website.	SEDAR 32
SEDAR41-RD02	List of documents and working papers for SEDAR 9 (Gulf of Mexico Gray Triggerfish, Greater Amberjack, and Vermilion Snapper) – all documents available on the SEDAR website.	SEDAR 9
SEDAR41-RD03	2011 Gulf of Mexico Gray Triggerfish Update Assessment	SEDAR 2011
SEDAR41-RD04	List of documents and working papers for SEDAR 24 (South Atlantic red snapper) – all documents available on the SEDAR website.	SEDAR 24
SEDAR41-RD05	List of documents and working papers for SEDAR 31 (Gulf of Mexico red snapper) – all documents available on the SEDAR website.	SEDAR 31

SEDAR41-RD06	List of documents and working papers for SEDAR 15 (South Atlantic red snapper and greater amberjack) – all documents available on the SEDAR website.	SEDAR 15
SEDAR41-RD07	2009 Gulf of Mexico red snapper update assessment	SEDAR 2009
SEDAR41-RD08	List of documents and working papers for SEDAR 7 (Gulf of Mexico red snapper) – all documents available on the SEDAR website.	SEDAR 7
SEDAR41-RD09	SEDAR 24 South Atlantic Red Snapper: management quantities and projections requested by the SSC and SERO	NMFS - Sustainable Fisheries Branch 2010
SEDAR41-RD10	Total removals of red snapper (*Lutjanus campechanus*) in 2012 from the US South Atlantic	NMFS - Sustainable Fisheries Branch 2013
SEDAR41-RD11	Amendment 17A to the Fishery Management Plan for the Snapper Grouper Fishery of the South Atlantic Region	SAFMC 2010
SEDAR41-RD12	Amendment 28 to the Fishery Management Plan for the Snapper Grouper Fishery of the South Atlantic Region	SAFMC 2013
SEDAR41-RD13	Total removals of red snapper (*Lutjanus campechanus*) in 2013 from the U.S. South Atlantic	NMFS - Sustainable Fisheries Branch 2014
SEDAR41-RD14	South Atlantic red snapper (*Lutjanus campechanus*) monitoring in Florida for the 2012 season	Sauls et al. 2013
SEDAR41-RD15	South Atlantic red snapper (*Lutjanus campechanus*) monitoring in Florida for the 2013 season	Sauls et al. 2014
SEDAR41-RD16	A directed study of the recreational red snapper fisheries in the Gulf of Mexico along the West Florida shelf	Sauls et al. 2014
SEDAR41-RD17	Using generalized linear models to estimate selectivity from short-term recoveries of tagged red drum *Sciaenops ocellatus*: Effects of gear, fate, and regulation period	Bacheler et al. 2009
SEDAR41-RD18	Direct estimates of gear selectivity from multiple tagging experiments	Myers and Hoenig 1997

SEDAR41-RD19	Examining the utility of alternative video monitoring metrics for indexing reef fish abundance	Schobernd et al. 2014
SEDAR41-RD20	An evaluation and power analysis of fishery independent reef fish sampling in the Gulf of Mexico and U.S. South Atlantic	Conn 2011
SEDAR41-RD21	Consultant's Report: Summary of the MRFSS/MRIP Calibration Workshop	Boreman 2012
SEDAR41-RD22	2013 South Atlantic Red Snapper Annual Catch Limit and Season Length Projections	SERO 2013
SEDAR41-RD23	Southeast Reef Fish Survey Video Index Development Workshop	Bacheler and Carmichael 2014
SEDAR41-RD24	Observer Coverage of the 2010-2011 Gulf of Mexico Reef Fish Fishery	Scott-Denton and Williams
SEDAR41-RD25	Circle Hook Requirements in the Gulf of Mexico: Application in Recreational Fisheries and Effectiveness for Conservation of Reef Fishes	Sauls and Ayala 2012
SEDAR41-RD26	GADNR Marine Sportfish Carcass Recovery Project	Harrell 2013
SEDAR41-RD27	Catch Characterization and Discards within the Snapper Grouper Vertical Hook-and-Line Fishery of the South Atlantic United States	Gulf and South Atlantic Fisheries Foundation 2008
SEDAR41-RD28	A Continuation of Catch Characterization and Discards within the Snapper Grouper Vertical Hook-and-Line Fishery of the South Atlantic United States	Gulf and South Atlantic Fisheries Foundation 2010
SEDAR41-RD29	Continuation of Catch Characterization and Discards within the Snapper Grouper Vertical Hook-and-Line Fishery of the South Atlantic United States	Gulf and South Atlantic Fisheries Foundation 2013
SEDAR41-RD30	Amendment 1 and Environmental Assessment and Regulatory Impact Review to the Fishery Management Plan for the Snapper Grouper Fishery of the South Atlantic Region	SAFMC 1988
SEDAR41-RD31	Final Rule for Amendment 1 to the Fishery Management Plan for the Snapper Grouper Fishery of the South Atlantic Region	Federal Register 1989
SEDAR41-RD32	Population Structure and Genetic Diversity of Red Snapper (Lutjanus campechanus) in the U.S.	Gold and Portnoy 2013

	South Atlantic and Connectivity with Red Snapper in the Gulf of Mexico	
SEDAR41-RD33	Oogenesis and fecundity type of Gulf of Mexico gray triggerfish reflects warm water environmental and parental care	Lang and Fitzhugh 2014
SEDAR41-RD34	Depth-related Distribution of Postjuvenile Red Snapper in Southeastern U.S. Atlantic Ocean Waters: Ontogenetic Patterns and Implications for Management	Mitchell et al. 2014
SEDAR41-RD35	Gray Triggerfish Age Workshop	Potts 2013
SEDAR41-RD36	Age, Growth, and Reproduction of Gray Triggerfish *Balistes capriscus* Off the Southeastern U.S. Atlantic Coast	Kelly 2014
SEDAR41-RD37	Assessment of Genetic Stock Structure of Gray Triggerfish (*Balistes capriscus*) in U.S. Waters of the Gulf of Mexico and South Atlantic Regions	Saillant and Antoni 2014
SEDAR41-RD38	Genetic Variation of Gray Triggerfish in U.S. Waters of the Gulf of Mexico and Western Atlantic Ocean as Inferred from Mitochondrial DNA Sequences	Antoni et al. 2011
SEDAR41-RD39	Characterization of the U.S. Gulf of Mexico and South Atlantic Penaeid and Rock Shrimp Fisheries Based on Observer Data	Scott-Denton et al. 2012
SEDAR41-RD40	Does hook type influence the catch rate, size, and injury of grouper in a North Carolina commercial fishery	Bacheler and Buckel 2004
SEDAR41-RD41	Fishes associated with North Carolina shelf-edge hardbottoms and initial assessment of a proposed marine protected area	Quattrini and Ross 2006
SEDAR41-RD42	Growth of grey triggerfish, *Balistes capriscus*, based on growth checks of the dorsal spine	Ofori-Danson 1989
SEDAR41-RD43	Age Validation and Growth of Gray Triggerfish, *Balistes capriscus*, In the Northern Gulf of Mexico	Fioramonti 2012
SEDAR41-RD44	A review of the biology and fishery for Gray Triggerfish, *Balistes capriscus*, in the Gulf of Mexico	Harper and McClellan 1997

SEDAR41-RD45	Stock structure of gray triggerfish, *Balistes capriscus*, on multiple spatial scales in the Gulf of Mexico	Ingram 2001
SEDAR41-RD46	Evaluation of the Efficacy of the Current Minimum Size Regulation for Selected Reef Fish Based on Release Mortality and Fish Physiology	Burns and Brown-Peterson 2008
SEDAR41-RD47	Population Structure of Red Snapper from the Gulf of Mexico as Inferred from Analysis of Mitochondrial DNA	Gold et al. 1997
SEDAR41-RD48	Successful Discrimination Using Otolith Microchemistry Among Samples of Red Snapper *Lutjanus campechanus* from Artificial Reefs and Samples of *L.campechanus* Taken from Nearby Oil and Gas Platforms	Nowling et al. 2011
SEDAR41-RD49	Population Structure and Variation in Red Snapper (Lutjanus campechanus) from the Gulf of Mexico and Atlantic Coast of Florida as Determined from Mitochondrial DNA Control Region Sequence	Garber et al. 2003
SEDAR41-RD50	Population assessment of the red snapper from the southeastern United States	Manooch et al. 1998
SEDAR41-RD51	Otolith Microchemical Fingerprints of Age-0 Red Snapper, *Lutjanus campechanus*, from the Northern Gulf of Mexico	Patterson et al. 1998
SEDAR41-RD52	Implications of reef fish movement from unreported artificial reef sites in the northern Gulf of Mexico	Addis et al. 2013
SEDAR41-RD53	Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species	Then et al. 2014
SEDAR41-RD54	Length selectivity of commercial fish traps assessed from in situ comparisons with stereo-video: Is there evidence of sampling bias?	Langlois et al. 2015
SEDAR41-RD55	MRIP Calibration Workshop II – Final Report	Carmichael and Van Vorhees (eds.) 2015
SEDAR41-RD56	Total Removals of red snapper (*Lutjanus campechanus*) in 2014 from the U.S. South Atlantic	SEFSC 2015
SEDAR41-RD57	Assessing reproductive resilience: an example with South Atlantic red snapper *Lutjanus campechanus*	Lowerre-Barbiere et al. 2015

SEDAR41-RD58	Overview of sampling gears and standard protocols used by the Southeast Reef Fish Survey and its partners	Smart et al. 2014
SEDAR41-RD59	MRIP Transition Plan for the Fishing Effort Survey	Atlantic and Gulf Subgroup of the MRIP Transition Team 2015
SEDAR41-RD60	Technical documentation of the Beaufort Assessment Model (BAM)	Williams and Shertzer 2015
SEDAR41-RD61	Stock Assessment of Red Snapper in the Gulf of Mexico 1872-2013, with Provisional 2014 Landings: SEDAR Update Assessment	Cass-Calay et al. 2015
SEDAR41-RD62	Excerpt from the December 2013 SAFMC SEDAR Committee Minutes (pages 11-21 where SEDAR 41 ToR were discussed)	SAFMC SEDAR Committee
SEDAR41-RD63	Population structure of red snapper (*Lutjanus campechanus)* in U.S. waters of the western Atlantic Ocean and the northeastern Gulf of Mexico	Hollenbeck et al. 2015
SEDAR41-RD64	SEDAR31-AW04: The Effect of Hook Type on Red Snapper Catch	Saul and Walter 2013
SEDAR41-RD65	SEDAR31-AW12: Estimation of hook selectivity on red snapper (*Lutjanus campechanus*) during a fishery independent survey of natural reefs in the Gulf of Mexico	Pollack et al. 2013
SEDAR41-RD66	Effect of Circle Hook Size on Reef Fish Catch Rates, Species Composition, and Selectivity in the Northern Gulf of Mexico Recreational Fishery	Patterson et al. 2012
SEDAR41-RD67	Effect of trawling on juvenile red snapper (*Lutjanus campechanus*) habitat selection and life history parameters	Wells et al. 2008
SEDAR41-RD68	SEDAR24-AW05: Selectivity of red snapper in the southeast U.S. Atlantic: dome-shaped or flat topped?	SFB-SEFSC 2010
SEDAR41-RD69	Hierarchical analysis of multiple noisy abundance indices	Conn 2010
SEDAR41-RD70	Data weighting in statistical fisheries stock assessment models	Francis 2011
SEDAR41-RD71	Corrigendum to Francis 2011 paper	Francis

SEDAR41-RD72	Quantifying annual variation in catchability for commercial and research fishing	Francis et al. 2003
SEDAR41-RD73	Evolutionary assembly rules for fish life histories	Charnov et al. 2012
SEDAR41-RD74	User's Guide for ASPIC Suite, version 7: A Stock-Production Model Incorporating Covariates and auxiliary programs	Prager 2015
SEDAR41-RD75	Standing and Special Reef Fish SSC, September 2015 Meeting Summary (see pages 4-7 for SEDAR 43 review)	Gulf of Mexico Standing and Special Reef Fish SSC
SEDAR41-RD76	Standing and Special Reef Fish SSC, January 2016 Meeting Summary (see pages 2-7 for SEDAR 43 review)	Gulf of Mexico Standing and Special Reef Fish SSC
SEDAR41-RD77	SEDAR 43 Gulf of Mexico Gray Triggerfish Stock Assessment Report	SEDAR 43
SEDAR41-RD78	Review of 2014 SEDAR 31 Gulf of Mexico Red Snapper Update Assessment	Gulf of Mexico Standing and Special Reef Fish SSC
SEDAR41-RD79	Influence of soak time and fish accumulation on catches of reef fishes in a multispecies trap survey	Bacheler et al. 2013

2. Review Panel Report

Executive Summary

The Review Workshop (RW) Panel was presented outputs and results of the SEDAR 41 South Atlantic Gray Triggerfish stock assessment. The primary assessment model used was the Beaufort Assessment Model (BAM), a software package that implements a statistical catch-at-age framework. The formulation is an age-structured population model that is fit using standard statistical methods to data available from surveys and fishing fleets, such as landings, discards, indices of abundance, age compositions, and length compositions. The modeling framework is nearly identical to other common assessment packages, such Age Structure Assessment Program (ASAP) and Stock Synthesis (SS), and the programming language (AD Model Builder) is the same across all three. A secondary, surplus-production model (Stock Production Model Incorporating Covariates, ASPIC) was also explored as a way to evaluate model uncertainty and provide a comparison of assessment results. However, the low contrast in Gray Triggerfish landings and indices of abundance caused ASPIC results to be considered non-informative. Further, an error with the Chevron Trap survey age composition data used in the base configuration of the BAM model was discovered during the RW (the age compositions used at the Assessment Workshop were based on the number of annuli and the corrected data were based on calendar-year age). Based on the magnitude of changes to the data, results and model diagnostics developed from the Assessment Workshop base model as well as concerns about overfitting the CVID survey, the Review Panel felt that the proposed base model parameterization was inappropriate to provide information on Gray Triggerfish stock status or benchmarks. The Review Panel recommends that further modeling is needed to fit the corrected age data and to resolve the fit to the CVID survey (perhaps investigating a multispecies year effect in 1990) to consider possible effects from Hurricane Hugo.

2.1 Statements Addressing Each ToR

1. Evaluate the data used in the assessment, including discussion of the strengths and weaknesses of data sources and decisions, and consider the following:
    a) Are data decisions made by the DW and AW sound and robust?
    b) Are data uncertainties acknowledged, reported, and within the normal or expected levels?
    c) Are data properly applied within the assessment model?
    d) Are data input series reliable and sufficient to support the assessment approach and findings?

    General comments

    Data decisions made by the DW and AW were sound and robust. The Review Panel acknowledges the considerable efforts of the DW and AW to compile the data and evaluate their strengths and weaknesses. The development of input data and parameters for the BAM and ASPIC models required an extremely thorough compilation and

evaluation of all available data at the DW. Modifications made subsequently by the AW were fully explained.

Data uncertainties were acknowledged, reported, and were within the normal or expected levels, where this could be ascertained from information provided to the RW. Data on fishery catches and length/age compositions, and fishery-dependent and independent relative abundance indices, varied widely in coverage and quality. Complex manipulations and standardisation methods were often required to try and develop coherent time series from diverse data sources of differing designs, coverage and accuracy, and the combined data will have biases that in some cases are poorly understood especially in earlier years of the time series. All decisions made by the DW and AW in compiling data were explained and justified in detail. Data quality metrics were provided by the DW in terms of numbers of samples, CVs, or alternative plausible data series or biological parameters. These were used by the AW to weight data series in the assessment model, estimate the uncertainty in the assessment results using the Monte Carlo/bootstrap method, or to explore the sensitivity of the assessment to data decisions and uncertainty. The sensitivity analyses were carried out altering one input at a time, and did not explore the impact of combinations of adjustments.

The data were properly applied within the assessment model. Any issues with application of the data such as time periods for fitting, use of length and age data from the same sampling schemes, or weighting of data according to data quality metrics, were explored at the SEDAR-41 RW if not previously evaluated by the DW and AW.

Data input series were mostly considered reliable and sufficient to support the assessment approach and findings. Reliability and sufficiency was evaluated based on a-priori criteria where possible, supported by data quality metrics such as numbers of samples or CVs and by model fits. The assessment is supported by a well-designed fishery-independent trap survey since 1990 and a wide range of fishery-dependent data covering landings and discards. The Review Panel was mostly concerned about how the fishery-dependent and fishery-independent data were treated in the BAM assessment than with the inherent reliability of the data, leading to the decision to put the assessment on hold pending a response from the AW as a whole to suggestions made by the Review Panel (this is explained under Term of Reference 2 below).

An evaluation of the strengths and weaknesses of the data sources and decisions is given below for each type of data used.

Life history parameters

Life history data and assumptions used in the Gray Triggerfish assessment include stock structure, reproductive biology and natural mortality. The assessment was sensitive to estimates of natural mortality (M) as is generally the case, although sensitivity to trends in M could not be evaluated as there is no information on this. An age-dependent, year-

invariant estimate of M was determined by a meta-analysis approach using growth parameters and maximum observed age. Reproductive biology was included in the model by computing total annual egg production at age based on maturity, length, number of batches and batch fecundity, allowing the effect of age structure on reproductive output to be reflected in setting SSB reference points and stock status. Interannual variation in fecundity, a possible source of uncertainty, was not able to be included as historical information was not available.

Age in Gray Triggerfish is difficult to determine from hard structures, a previous impediment to developing a stock assessment. Strenuous efforts have been made to develop and validate accurate methods to determine age based on increments in dorsal spines (converted to calendar years), and those results are used in the BAM model. Gray Triggerfish age readings show a broad distribution of length at age relative to the annual growth increment in length (Fig. 1 – from presentation to SEDAR-41 RW), which in combination with selectivity assumptions affect the ability for BAM to estimate annual age compositions through fit to sample length compositions.

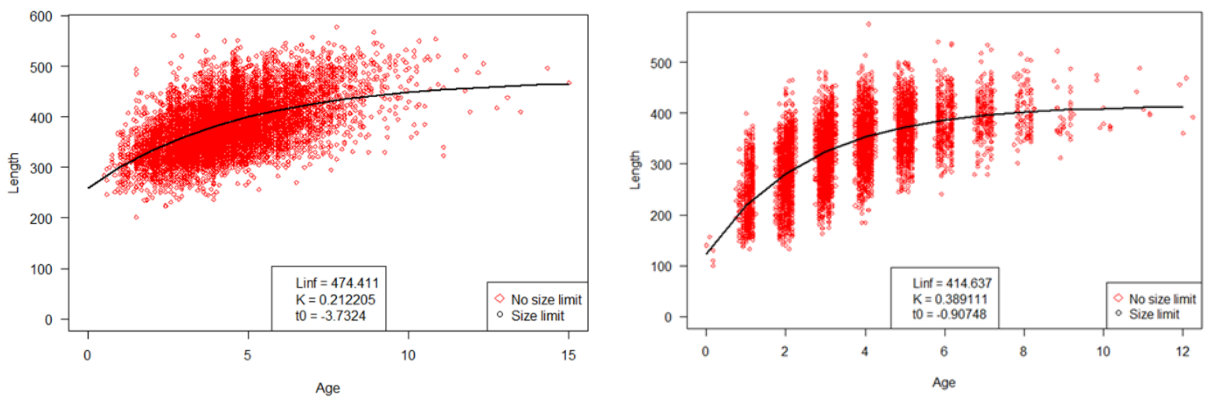

Fig. 1. Left: fishery dependent length-age data for Gray Triggerfish; right: fishery-independent (CVID trap) data.

Fishery removals

Reconstruction of a historical series of commercial and recreational fishery removals–landings and dead discards–was made back to 1988 for the assessment. This required a large number of decisions to impute missing values or to calibrate data series where design has changed, particularly for the change from the NMFS Marine Recreational Fisheries Statistics Survey (MRFSS) surveys (1981 to 2003) to the Marine Recreational Information Program (MRIP: 2004 to present), and for developing discards time series.

Landings of commercial handline fleets have improved in accuracy over time, and the DW proposed CVs that could be used for MCB uncertainty analysis in the assessment. Recreational landings of headboats are estimated from the Southeast Region Headboat

Survey (SRHS) log book scheme which has improved in quality over time due to introduction of mandatory reporting in 1996 and improved logbook supply from 2008 onwards. Private boat and charter boat landings since early 1980s were estimated from MRFSS/MRIP, which has a robust and peer-reviewed statistical design with improved design and precision over time, and for which CVs are estimated directly based on efficient estimators.

Discards estimates are inherently less reliable than landings for both the commercial and recreational fleets, and for the commercial handline fleet involved extrapolating observations for 2001-2011 to other years back to 1988. Separate discards estimates for the open and closed seasons since 2012 were made for this fishery based on effort. Recreational head boat discards estimates are available from SRHS log books since 2004, but for previous years back to 1988 are inferred using MRIP charter boat data adjusted using ratios of SRHS to MRIP estimates for 2004-2013. Gray Triggerfish discards estimates from SRHS and MRIP are self-reported and are not verified. All these uncertainties and data manipulations introduce error in the time series.

Length and age compositions

For Red Snapper, the SEDAR-41 AW used age composition data in preference to length composition data in BAM where both data exist, but for Gray Triggerfish the AW fitted both length and age compositions for headboats and the CVID survey which will result in some over-weighting of composition data. Length compositions from 1988 onwards were fitted for landings of commercial lines, headboats and from 1990 for the CVID survey. Headboat discards length frequencies were fitted from 2005 onwards. Age compositions for the commercial handline fleet and headboat landings were fitted mainly from the 2000s onwards and for the CVID survey from 1990 onwards. The CVID age data for Gray Triggerfish were found towards the end of the review meeting to have not been converted to calendar ages. Revised data were provided, and the assessment was rerun. (*see ToR's 2 and 3 for elaboration on this*)

The age compositions used at the Assessment Workshop were based on the number of annuli and the corrected data were based on calendar-year age. The change in estimates of age composition were substantial for 1991 to 2007, in which the corrected age compositions are approximately one year older. The change in estimates was less substantial after 2007. The difference in magnitude of the change results from a change in sampling protocol. From 1991 to 2007, sampling protocol was based on a target number of age samples per length bin, and the sample requirements were typically collected early in the sampling season, requiring conversion to calendar year by adding one to the annuli count an age. Sampling after 2007 was based on a random sampling design. Thus, post 2007, fish retained for age determination were distributed throughout the entire sampling season, requiring less of a correction.

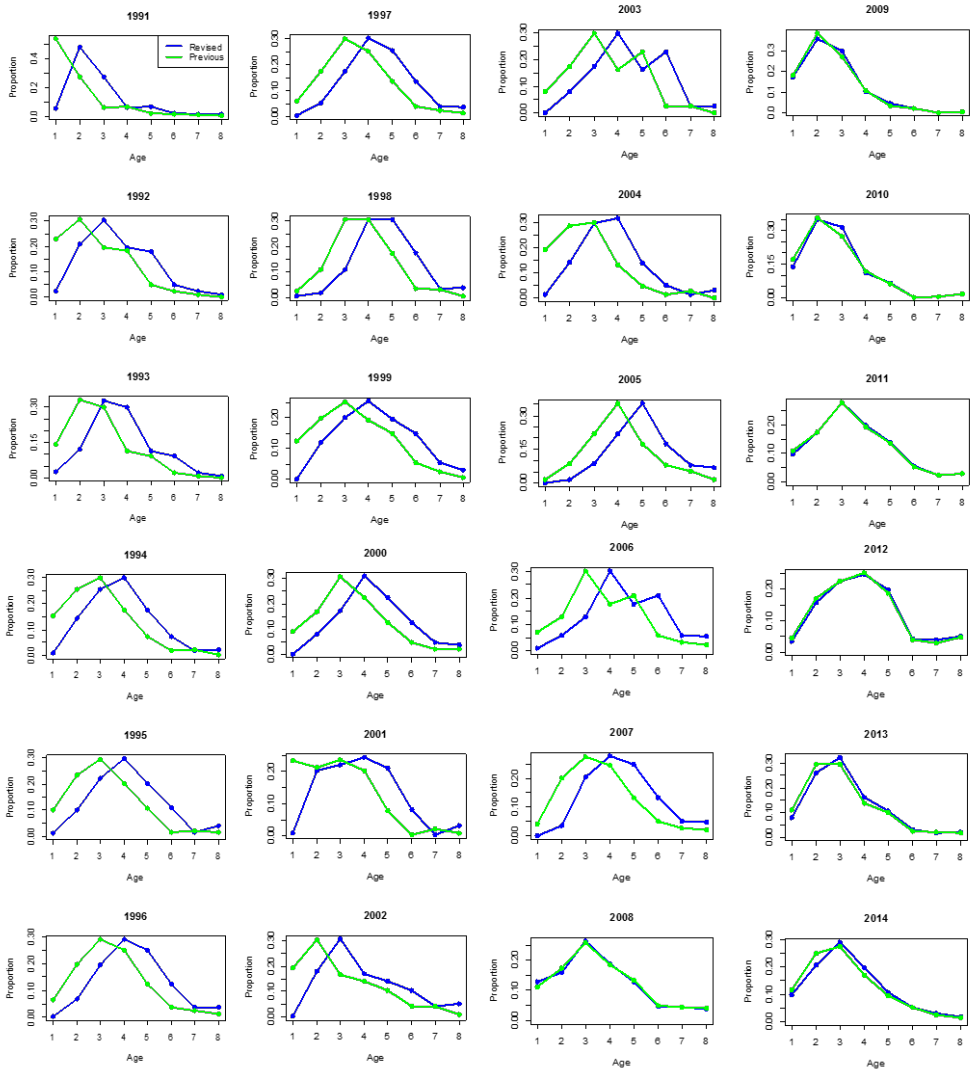

Figure 2. Age composition of the Chevron Video Trap Survey from the Assessment Workshop (green) and corrected (blue).

The Review Panel also expressed concerns regarding the broad length at age distributions relative to annual growth increment for many of the age classes making up a large portion of fishery-dependent and fishery-independent data (see Fig. 1 above), which will affect the ability of BAM to estimate annual age compositions through fit to length composition sample data. The Review Panel requested a sensitivity run of BAM with length data omitted where age data were available. This resulted in a deterioration in fit to some age composition data suggesting that the sampling for age, given the age error matrix, was inadequate for those years or could also reflect the correlations between length and age data collected from the same samples and the use of length data to weight the age compositions in each length class.

Relative abundance indices

The Review Panel considers the rationale for including CPUE data from three fisheries-dependent surveys and the two fisheries-independent series from the Chevron Trap Survey and the combined CVID trap/video survey in the BAM stock assessment model to be reasonable. The combination of trap/video survey indices from 2010 is clearly supported since the video camera is mounted on the trap, and hence cannot be considered independent. The standardized index for the fisheries-independent survey based on a zero-inflated model accounts for yearly shifts in sampling distributions relative to covariates that affect catch rates for Grey Triggerfish, and is restricted to depths from 10 to 94 m where Grey Triggerfish has been captured in any of the monitoring programs.

The various sources of systematic errors (e.g., spatial coverage, selectivity, trap saturation) and random errors (e.g., sample sizes) in each individual series are well documented. The established growth curves based on different data sources suggest that larger fish have lower probability to be captured in the CVID trap/video survey than by recreational fisheries.

Since Gray Triggerfish is unlikely to be the target for most fishing trips, the fishery-dependent CPUE series has potential to provide robust indices of abundance during periods of stable management measures that affect fishing behavior. However, the management measures applied to red snapper fishing since 2010 are likely to have caused a shift in targeting to other species which could also potentially cause a shift in catchability or selectivity for Gray Triggerfish.

The application of the fishery-dependent relative abundance data in the assessment model follows standard practice and appears sound. However, since the CPUE indices of abundance partly cover different depths/areas it should be noted that they do not individually cover the entire stock. The inclusion of these indices with equal weight in the model could therefore cause bias. A combination of the CPUE series external to the model based on their spatial/depth coverage is an alternative that may be explored in future assessments.

The use of the CVID chevron trap and video survey in the BAM model for Gray Triggerfish was considered in detail at the Review Workshop meeting due to the Assessment Panel's decision to up-weight the series by a factor of 6 to ensure a good fit to the index. This fishery-independent index covers center of geographic distribution of the stock (North Carolina – North Florida), full depth range, and extends over nearly the entire time series (1990-2014). However, the first year (1990) of the CVID chevron trap survey were conducted after hurricane Hugo, and may have experienced drastic lower catching efficiency due to strong habitat disturbances. The Review Panel recommends that in future assessments the inclusion of this first year be reconsidered.

The results of the stock assessment modeling depend on the relative weights assigned to different data sets. However, there is no consensus amongst practitioners as to the best

approach to data weighting. This stock assessment follows the common practice of weighting compositional catch data and abundance indices in two stages. The input data are first assigned relative weights before the model is run, and then iteratively weighted during a model run to improve model fit. Ideally, stage one weighting would use information about sample sizes (primary sampling units, and lower level sample sizes) and the way in which the data were collected (i.e., multi-stage survey designs), through calculated precision and effective sample sizes (Francis 2011; Pennington and Vølstad 1994). In particular, abundance indices by cohorts are likely to have different precision due to differences in the number of primary sampling units (e.g., trips, or trap-sets) where the cohorts are caught (Aanes and Vølstad 2015). In general, the multi-stage sampling can introduce complex correlation structures among cohorts, and drastically reduce the effective sample sizes for estimating compositions, and indices of cohorts (Aanes and Vølstad 2015). This would allow different weighting to each data point. The annual CV's for the abundance indices from the fisheries-independent surveys are computed using bootstrapping (SEDAR41-DW52) but the weight assigned to the input data to the BAM model is chosen ad-hoc. In particular the six-fold up-weighting of the fishery-independent abundance indices (across cohorts) in the base model is poorly justified. The abundance indices appear to be assigned fixed stage one weights. However, since sample sizes vary over the years and across cohorts, a fixed CV could cause bias.

An estimate of the variance of the fishery-independent indices based only on the between trap variability will underestimate the true variance of the abundance indices if catchability varies over time. Pennington and Godø (1995) estimated the actual variance of survey abundance indices by cross-calibrating independent VPA estimates and survey catch per tow indices. The CPUE series could potentially be used to cross-calibrate the fishery-independent indices in a time series analysis.

Finally, the Review Panel recommends that future assessments investigate the possibility of a gear saturation effect for the Chevron Trap survey gear. Under the standard soak time of 90 mins Bacheler *et al*. (2013) have shown that catch rates of Gray Triggerfish tapers off once a moderate number of total individuals were already caught in a trap.

2. Evaluate and discuss the strengths and weaknesses of the methods used to assess the stock, taking into account the available data, and consider the following:
   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?

The Review Panel concluded that this ToR was not met.

The base configuration of the Beaufort Assessment Model (BAM) from the Assessment Workshop was revised with corrected age compositions of the Chevron Trap survey

during the Review Workshop. Although the determination of stock status was not influenced by the correction, results from the original base model run and from the corrected base model were somewhat different. The Review Panel requested and received two revised models to resolve apparent difficulties in fitting to the survey and associated estimates of abundance in the first year of the assessment series. Based on the magnitude of changes to the data, results and model diagnostics from the Assessment Workshop base model as well as concerns about overfitting the survey, the Review Panel recommends that further modeling is needed to model the corrected data appropriately.

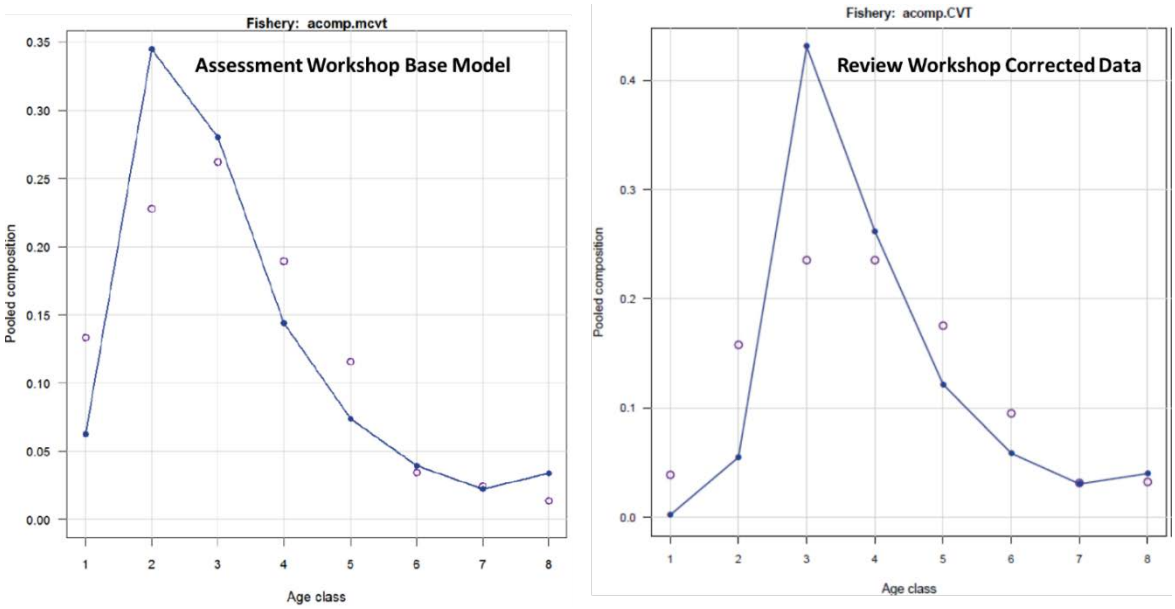

Figure 3. Observed and predicted age composition of the Chevron Video Trap Survey from the Assessment Workshop based model (left) and the corrected data and revised model from the Review Workshop (right).

In addition to the data corrections and model diagnostics, the Review Panel was concerned that the up-weighted survey (up-weighted 6-fold) in the Assessment Workshop base model was overfitting to the survey (i.e., fitting to survey measurement error rather than signals of abundance trends). The estimates of abundance at age in the first year of the assessment were the lowest in the assessment time series, despite being relatively early in the exploitation history. The low abundance estimates are produced by closely fitting to the first year of the trap survey, which was the lowest in the survey time series. An alternative with no up-weighting of the survey (sensitivity run 6) produced estimates of abundance in the first year of the assessment that were similar to the rest of the time series.

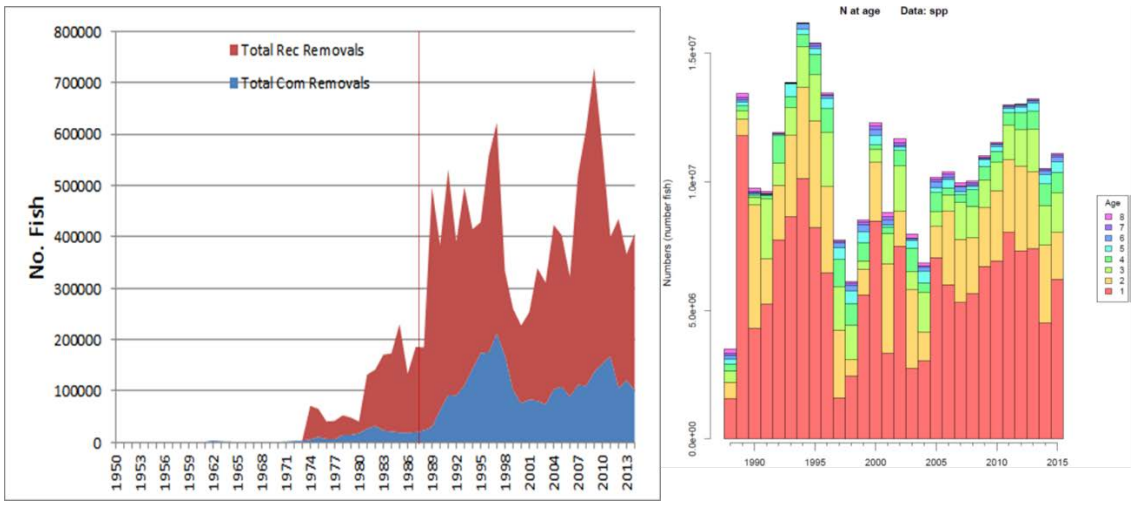

Figure 4. Exploitation history and abundance at age estimated from the corrected base model.

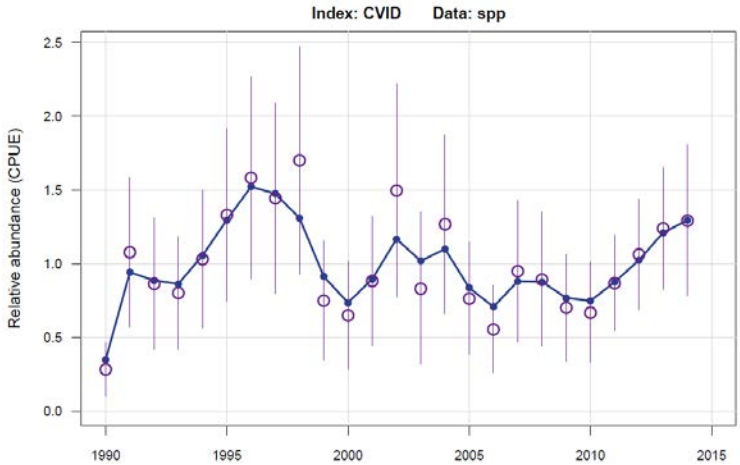

Figure 5. Observed and predicted Chevron Video Trap Survey indices from the corrected base model.

The Review Panel requested alternative analyses to help it understand the results. An alternative BAM configuration with a starting year of 1974 estimated a series of low recruitment to explain the low survey index in 1990. The Review Panel was also concerned that the need to up-weight the survey may result from using composition samples twice (as age compositions and length compositions). An exploratory analysis that removed length compositions for fleets with age compositions, with no up-weighting of the survey, still did not fit the survey well.

The extremely low estimates of abundance in the first year of the assessment may result from an unusual survey observation in the first year of the survey, rather than overfitting

the entire survey series. An exploratory analysis that removed the 1990 survey observation produced estimates of abundance in the first year of the assessment that were similar to the rest of the time series. The Chevron trap survey began in 1988, but the protocol was being refined in 1988 and 1990. There have been no changes to the design of the survey since 1990. However, Hurricane Hugo was 7-8 months prior to the 1990 survey. A study of Jamaican reef fish found changes in abundance, behavior, and distribution a year after Hurricane Allen (Kaufman 1983).

3. Evaluate the assessment findings and consider the following:

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

The base model estimated very low levels of abundance in the initial years of 1988 and 1989 at a time when exploitation was expected to be quite low. Sensitivity runs determined that a combination of fitting the model closely to the low 1990 CVID point by using 6 times up-weighting and the assumed selectivity for the CVID resulted in low abundance and recruitment in the first two years. Fitting the CVID without up-weighting essentially resulted in no appreciable trend over the time series although the fit was contained within the confidence intervals for all the survey points. An additional issue was identified with inclusion of both age and length compositions in the fitting process which was explored after the CVID age compositions had been corrected for errors discovered late in the week. Additional runs to establish a base case with the corrected age compositions and removing length compositions when age compositions resulted in poor fits to the Headboat and CVID age compositions. At this point, the review panel concluded that there wasn't enough time left in the meeting to establish a base case for gray triggerfish and the assessment panel needed to review the findings to date and work with the assessment team to develop a new base case.

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

Without an accepted base case, the review panel was unable to determine if the stock was overfished with respect the standard reference points. However, based on the information presented the review panel could say that there was no evidence for a decline in abundance or biomass at this time.

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

Without an accepted base case, the review panel was unable to determine if overfishing was occurring with respect the standard reference points. Based on the information presented to the review panel there was no evidence that current levels of removals have resulted in overfishing.

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

The stock recruitment curve was not informative as there was little evidence for low recruitment at low stock size. Inference was based on setting steepness to 0.99 and mean annual recruitment was assumed. Lognormal deviations around the mean were estimated in the model.

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?

Without a reliable base case, quantitative estimates of status determination were not available.

4. Evaluate the stock projections, including discussing the strengths and weaknesses, and consider 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 are they useful to support inferences of probably future conditions?
    d) Are key uncertainties acknowledged, discussed, and reflected in the projection results?

Since the base BAM model for Gray Triggerfish was not accepted by the Review Panel projections results were only reviewed in terms of the methodological approaches used—i.e., projections results were not considered as providing plausible scenarios and, therefore, were not investigated in detail.

The projection method used is consistent with those used widely in SEDAR assessments based on statistical models such as BAM and Stock Synthesis, and is consistent with the available data. Further the method described for the stochastic projections that extended the Monte Carlo/ Bootstrap (MCB) fits of the assessment model with added stochasticity in recruitment, and hence the propagation of uncertainty from the assessment into the projection period is internally consistent.

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

b) Ensure that the implications of uncertainty in technical conclusions are clearly stated.

Like it was described above for ToR 4, because the base BAM model for Gray Triggerfish was not accepted by the Review Panel assessment uncertainty results were only reviewed in terms of the methodological approaches used. Further, the Review Panel is concerned that many of the reported uncertainties on quantities of interest are a consequence of the assumed (and fixed) observation variance parameters. No clear evidence of the appropriateness of these assumed values has been presented.

Because of the large number of parameters in BAM a thorough evaluation of convergence and model sensitivity is necessary, but difficult. Uncertainties in the assessment were explored through (1) a mixed Monte Carlo and bootstrap (MCB) analysis to quantify random errors in the assessment output; (2) sensitivity analysis around the base BAM run; and (3) the use of alternative assessment models.

The MCB runs included ranges of values of natural mortality, discard mortality and fecundity at age agreed by the assessment working group, together with bootstrap selection of data using well-justified error distributions

The sensitivity analyses were used to explore a wide range of data decisions, model assumptions and model configurations to examine the robustness of stock status determination. The model was run for a plausible range of values for each factor. The Review Panel noted that the sensitivity testing by alternating one factor at a time, although commonly done, may not fully reflect the uncertainty in model outputs from a complex model such as BAM with a large number of parameters where many are likely to be correlated (e.g., Saltelli and Annoni (2010). Global sensitivity analysis (Saltelli et al. 2008) may be used to untangle the contribution of single factors/parameters and interactions between parameters to the overall variability in model output. Anderson et al. (2011) provide an excellent overview of the literature, and many examples of applications of global sensitivity analysis to Integrated Assessment Models in climate research, and some of these are likely to be applicable to the BAM model.

Model uncertainty was mainly explored by running an alternative Stock Production Model Incorporating Covariates (ASPIC software of Prager, Version 7.03, 2005) that relies on length-age aggregated catch and CPUE indices, with no compositional catch being included. However, since the Assessment Panel concluded that none of the ASPIC runs produced during the Assessment Workshop produced plausible results and reasonable model diagnostics the use of this alternative modeling approach could not be properly evaluated for Gray Triggerfish in SEDAR 41. The Review Panel recommends that this approach be further explored at the next Gray Triggerfish assessment.

6. Consider the research recommendations provided by the Data and Assessment workshops and make any additional recommendations or prioritizations warranted.

    a) Clearly denote research and monitoring that could improve the reliability of, and information provided by, future assessments.
    b) Provide recommendations on possible ways to improve the SEDAR process.

- Increased fishery independent information, in particular reliable indices of abundance and age compositions.
- Increased age sampling and evaluation of ageing error over the stock area and from all fleets, particularly the general recreational fleet.
- More research to better understand the life history of Gray Triggerfish is needed, including natural mortality, maturity, and reproductive potential, particularly for the youngest ages.
- The effects of environmental variation on the changes in recruitment or survivorship of Gray Triggerfish.

7. Consider whether the stock assessment constitutes the best scientific information available using the following criteria as appropriate: relevance, inclusiveness, objectivity, transparency, timeliness, verification, validation, and peer review of fishery management information.

    The Review Panel concluded that, as configured, the SEDAR 41 Gray Triggerfish stock assessment model could not be considered the best scientific information available.

8. Compare and contrast assessment uncertainties between the Gulf of Mexico and South Atlantic stocks.

    It was not possible to complete this ToR since the SEDAR 41 stock assessment could not be successfully completed (i.e., many of the assessment uncertainties could not be fully evaluated).

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

    The extremely low estimates of abundance in the first year of the assessment may result from an unusual survey observation in the first year of the survey, rather than overfitting the entire survey series. An exploratory analysis that removed the 1990 survey observation produced estimates of abundance in the first year of the assessment that were similar to the rest of the time series. The Chevron trap survey began in 1988, but the protocol was being refined in 1988 and 1990. There have been no changes to the design of the survey since 1990. However, Hurricane Hugo was 7-8 months prior to the 1990

survey. A study of Jamaican reef fish found changes in abundance, behavior, and distribution a year after Hurricane Allen (Kaufman 1983).

The Review Panel recommends that further modeling is needed to fit the corrected age data and to resolve the fit to the CVID survey (perhaps investigating a multispecies year effect in 1990) to consider possible effects from Hurricane Hugo and a justification for removing the 1990 survey observation.

10. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference. Develop a list of tasks to be completed following the workshop. Complete and submit the Peer Review Summary Report in accordance with the project guidelines.

    This report constitutes the Review Panel's summary evaluation of the stock assessment and discussion of the Terms of Reference. The Review Panel will complete edits to its report and submit a final document to the SEDAR program for inclusion in the full set of documents associated with SEDAR 41.

References

Aanes, S., and Vølstad, J. H. 2015. Efficient statistical estimators and sampling strategies for estimating the age composition of fish. Canadian Journal of Fisheries and Aquatic Sciences, 72(6): 938-953.

Anderson, B., Emanuele Borgonovo, Marzio Galeotti and Roberto Roson. 2012. Uncertainty in Integrated Assessment Modelling: Can Global Sensitivity Analysis Be of Help? Working Paper Series. Working Paper 52. ISSN 1973-0381. IEFE - The Center for Research on Energy and Environmental Economics and Policy at Bocconi University. This paper can be downloaded at www.iefe.unibocconi.it

Bacheler, N. M., Bartolino, V., & Reichert, M. J. 2013. Influence of soak time and fish accumulation on catches of reef fishes in a multispecies trap survey. Fishery Bulletin, 111(3): 218-232.

Francis, Chris RIC, Rosemary J. Hurst, and James A. Renwick. 2003. Quantifying annual variation in catchability for commercial and research fishing. Fishery Bulletin 101: 293-304.

Francis, RIC Chris. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68: 1124-1138.

Kaufman, L.S. 1983. Effects of Hurricane Allen on reef fish assemblages near Discovery Bay, Jamaica. Coral Reefs 2: 43-47.

Mitchell, W. A., Kellison, G. T., Bacheler, N. M., Potts, J. C., Schobernd, C. M., & Hale, L. F. 2014. Depth-related distribution of post-juvenile red snapper in southeastern US Atlantic Ocean waters: ontogenetic patterns and implications for management. Marine and Coastal Fisheries, 6(1): 142-155.

Pennington, Michael, and Olav Rune Godø. 1995. Measuring the effect of changes in catchability on the variance of marine survey abundance indices. Fisheries research 23: 301-310.

Pennington, M., and Vølstad, J. H. 1994. Assessing the effect of intra-haul correlation and variable density on estimates of population characteristics from marine surveys. Biometrics, 725-732.

Stephens, A., and MacCall, A. 2004. A multispecies approach to sub-setting logbook data for purposes of estimating CPUE. Fisheries Research, 70: 299-310.

2.2 Summary Results of Analytical Requests