Mid-Atlantic Fishery Management Council
800 North State Street, Suite 201, Dover, DE 19901
Phone: 302-674-2331 | FAX: 302-674-5399 | www.mafmc.org
P. Weston Townsend, Chairman | Michael P. Luisi, Vice Chairman
Christopher M. Moore, Ph.D., Executive Director

MEMORANDUM

Date: January 22, 2024

To: Council

From: Brandon Muffley and Julia Beaty, Council staff

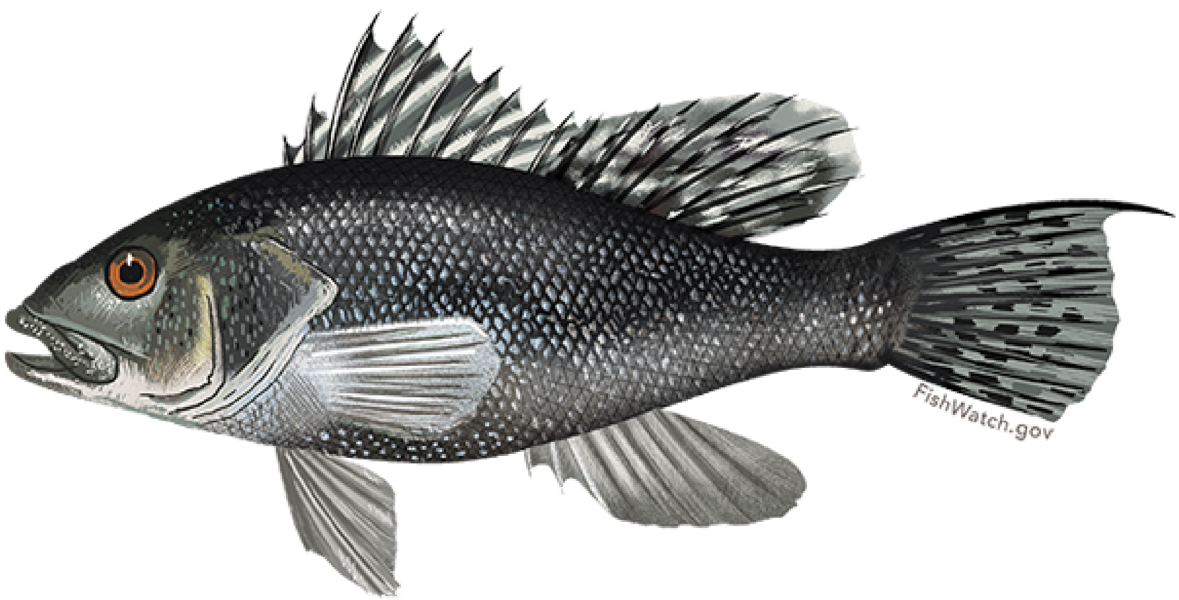
Subject: Meeting Materials – Black Sea Bass Research Track Stock Assessment

On Wednesday, February 7, 2024, the Northeast Fishery Science Center (NEFSC) will provide the Mid-Atlantic Fishery Management Council (Council) with an overview of the recently completed research track stock assessment and peer review for black sea bass. The newly approved stock assessment will be updated with information through 2023 as part of the management track peer review in June 2024. Results from the June management track review will be used to inform management and future catch specifications.

Materials listed below are provided for Council consideration of this agenda item.

- 2023 Report of the Black Sea Bass Research Track Working Group Report – Executive Summary
  - The full working group report can be found at: Stock Assessment Support Information (SASINF) Search Tool
- Summary Report of the Black Sea Bass Research Track Stock Assessment Peer Review

Report of the Black Sea Bass *(Centropristis striata*) Research Track Stock Assessment Working Group



November 20, 2023

TABLE OF CONTENTS

PARTICIPANTS

Working Group

NAME	AFFILIATION	E-MAIL
Kiersten Curti	NEFSC	kiersten.curti@noaa.gov
Anna Mercer	NEFSC	anna.mercer@noaa.gov
Julia Beaty	MAFMC	jbeaty@mafmc.org
Gavin Fay	UMASSD-SMAST	gfay@umassd.edu
Marissa McMahan	Manomet	mmcmahan@manomet.org
Jason McNamee	RIDEM	jason.mcnamee@dem.ri.gov
Tim Miller	NEFSC	timothy.j.miller@noaa.gov
Sam Truesdell	NEFSC/MADMF	samuel.truesdell@noaa.gov
Ricky Tabendera	NEFSC/HCRI	ricky.tabandera@gmail.com

Contributors and Participants in Working Group Meetings

NAME	AFFILIATION	E-MAIL
Tracey Bauer	ASMFC	tbauer@asmfc.org
Gregory Bopp	NEFSC	gregory.bopp@noaa.gov
Jason Boucher	NEFSC	jason.boucher@noaa.gov
Krystina Braid	Stonybrook	krystina.braid@stonybrook.edu
Russ Brown	NEFSC	russell.brown@noaa.gov
Jeffrey Brust	NJDEP	Jeffrey.Brust@dep.nj.gov
Michael Celestino	NJDEP	mike.celestino@dep.nj.gov
Lisa Chong	MSU	chonglis@msu.edu
Steve Doctor	MDDNR	steve.doctor@maryland.gov
Greg DiDomenico	Lunds	gDiDomenico@lundsfish.com
Alexander Dunn	NEFSC	alexander.dunn@noaa.gov
Alexa Galvan	VMRC	alexa.galvan@mrc.virginia.gov

Remy Gatins	NEU	remygatinsa@gmail.com
Alex Hansell	NEFSC	alex.hansell@noaa.gov
Andy Jones	NEFSC	andrew.jones@noaa.gov
Scott Large	NEFSC	scott.large@noaa.gov
Bobby Lemeiux	Industry	bobbylemieux10@gmail.com
Chris Legault	NEFSC	chris.legault@noaa.gov
Emily Liljestrand	NEFSC	emily.liljestrand@noaa.gov
John Maniscalco	NYDEC	john.maniscalco@dec.ny.gov
Nichola Meserve	MADMF	nichola.meserve@mass.gov
Andie Painten	SMAST	apainten@umassd.edu
Willow Patten	NCDENR	willow.patten@ncdenr.gov
Maria Cristina Perez	SMAST	mperez12@umassd.edu
Michele Traver	NEFSC	michele.traver@noaa.gov
Abigail Tyrell	NEFSC	abigail.tyrell@noaa.gov
Hannah Verkamp	CFRF	hverkamp@cfrfoundation.org
Matthew Vincent	NOAA	matthew.vincent@noaa.gov
John Wiedenmann	Rutgers	john.wiedenmann@rutgers.edu
Rich Wong	DEDFW	Richard.Wong@delaware.gov
Paula Fratantoni	NEFSC	paula.fratantoni@noaa.gov

LIST OF ACRONYMS

ASAP	Age structured assessment program
ASM	At sea monitoring
B_{MSY}	Biomass at maximum sustainable yield
$B_{threshold}$	Biomass level defining an overfished state
ChesMMAP	Chesapeake Bay Multispecies Monitoring and Assessment Program
CPA	Catch per angler
CPUE	Catch per unit effort
CV	Coefficient of variation
F	Fishing mortality rate
F_{40}	Fully selected F achieving 40% of unfished spawning biomass per recruit
F_{MSY}	Fishing mortality rate at maximum sustainable yield
M	Natural mortality rate
MRIP	Marine Recreational Information Program
MSY	Maximum sustainable yield
NAFO	Northwest Atlantic Fisheries Organization
NEAMAP	Northeast Area Monitoring and Assessment Program
NEFOP	Northeast Fisheries Observer Program
NEFSC	Northeast Fisheries Science Center
NMFS	National Marine Fisheries Service
NOAA	National Oceanic and Atmospheric Administration
NRCC	Northeast Regional Coordinating Council
OSA	One-step-ahead
SARC	Stock assessment review committee

SAW	Stock assessment workshop
SDC	Status determination criteria
SE	Standard error
SPR	Spawning biomass per recruit
SS	Stock synthesis modeling approach
SSB	Spawning stock biomass
SSC	Mid-Atlantic Fishery Management Council's Scientific and Statistical Committee
TOR	Term of reference
VAST	Vector Autoregressive Spatio-Temporal models
VIMS	Virginia Institute of Marine Science
VTR	Vessel trip report
WHAM	Woods Hole Assessment Model
YOY	Young of year

EXECUTIVE SUMMARY

The working group was formed in July 2021 and met over the following two years to address its terms of reference (TORs). This report represents consensus of the working group and includes contributions from working group members and participants.

TOR1: Ecosystem and Climate Influences

"*Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under impacted TORs*."

The working group explored several avenues for integrating ecosystem considerations in the black sea bass stock assessment, which are described in TOR 1: Ecosystem and Climate Influences and in the Truesdell & Curti 2023b, Hansell & Curti 2023, Tabendera et al. 2023, McMahan & Tabendera 2023, McNamee 2023, and Mercer et al. 2023 working papers. In an effort to recognize the impact that climate change has on the biology of black sea bass, the working group evaluated and implemented time varying growth and maturity, developed new age-length keys that are regionally and seasonally specific (Truesdell & Curti 2023b working paper), and conducted spatiotemporal modeling with environmental covariates (Hansell & Curti 2023 working paper). The working group also evaluated ecosystem influences on black sea bass, which included a literature review and development of oceanographic indicators for black sea bass recruitment and mixing rates between regions (Tabendera et al. 2023 working paper). After careful consideration, the working group moved forward with integrating a bottom temperature covariate on recruitment in the stock assessment model. In addition, the working group explored black sea bass food habits and empirical approaches for estimating natural mortality, which suggested maintaining natural mortality at 0.4 (McMahan & Tabendera 2023 and McNamee 2023 working papers). Finally, the working group made a significant effort to gather ecological and fishery knowledge from black sea bass stakeholders through public events and one-on-one conversations. The information gleaned from this effort was critical for sense checking the data inputs and model outputs of the black sea bass stock assessment, and also contributed to the development of novel standardized catch per unit effort (CPUE) indices from the commercial trawl fleet (Mercer et al. 2023 working paper).

TOR2: Fishery Data

"*Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data*."

The working group's analysis of black sea bass fishery data and discard mortality are described in TOR 2: Fishery Data and in the Beaty et al. 2023, Curti et al. 2023a, Curti et al. 2023b, Truesdell & Curti 2023a, and Verkamp et al. 2023 working papers. For the commercial component of the black sea bass fishery, the primary gears used are otter trawls, pots, and handlines (Curti et al. 2023a working paper). Over the commercial catch time series (1989-2021), trawl gears accounted for 45% of the commercial landings, pots and traps represented 41%, handlines accounted for 10% and other gears comprised the remaining 5%. Total commercial landings averaged approximately 1,240 mt through 2007, decreased to an average of 739 mt between 2008-2012 due to quota regulations, and generally increased from 2013 onward to a time series maximum of 2,013 mt in 2021 due to both population and regulatory changes. Over the course of the time series, the proportion of commercial landings that came from the northern region generally increased from an average of 24% through 2000 to a maximum of 83% in 2018.

Black sea bass commercial landings are distributed from Cape Hatteras to Cape Cod, with a concentration of landings inshore (<30m) representing the summer fishery, and a concentration of landings offshore representing the winter fishery (Curti et al. 2023b working paper). The spatial distribution of black sea bass commercial landings has changed over time, with the highest landings shifting from the waters off of Virginia, Delaware, and New Jersey in early years (1994-2005) to the waters off of New York, Rhode Island, and Massachusetts in recent years (2006-2021). The total commercial landings from the continental shelf south of New York and Rhode Island has also increased in recent years (2016-2021), potentially reflecting increased availability in these areas.

Commercial landings by market category varied over time. Landings prior to 2000 were primarily small and medium fish, and landings since 2010 have been primarily large and jumbo individuals. Annual length samples were combined across gears to permit length expansions by region, semester and market category. The primary differences in size composition among gears were accounted for by completing catch expansions separately for each market category. Region, year and semester-specific age-length keys were applied to expanded commercial landings-at-length to estimate commercial landings-at-age for each region (Truesdell & Curti 2023b). Landings-at-age in the northern and southern regions showed an expansion in the age structure over the time series with ages 6^+ becoming more prevalent from approximately 2000 onward.

Commercial discards were estimated by gear type for bottom trawl, gillnet, handline, pots/traps and scallop gears. Total annual commercial dead discards in the north averaged approximately 28 mt through 2000, increased to an average of 86 mt in the 2000s, and then increased substantially during the 2010s to a maximum of 918 mt in 2017. Total annual commercial dead discards in the south generally varied without trend over the 1989-2021 time series and averaged 66 mt. Across both regions, bottom trawls were generally the greatest source of discards, though scallop gear and pots/traps were also dominant in some years. The spatial distribution of discarded catch from observed commercial trips is greatest on the outer continental shelf. In recent years (2015-2021), total observed discards have increased in nearshore waters south of Rhode Island and Massachusetts as well as offshore around Hudson Canyon.

Discard length expansions were completed for each region, semester, year and gear type. Discard length composition data were obtained from the Northeast Fisheries Science Center (NEFSC) Northeast Fisheries Observer and At Sea Monitoring programs, and the Commercial Fisheries Research Foundation (Verkamp et al. 2023). Resulting expanded discards-at-length showed an increase over the time series in the maximum length in both regions and an increase in the median discarded length in the northern region. The same age-length keys used for commercial landings were also applied to expanded commercial discards-at-length to estimate commercial discards-at-age for each region. Similar to the trends in landings, discards-at-age in both the northern and southern regions showed an expansion in the age structure over the time series with ages 6^+ becoming more prevalent from approximately 2000 onward.

Trends in total commercial catch varied by region. In the northern region, total commercial catch averaged approximately 450 mt through 2010 but then increased to a maximum of 2,346 mt in 2017 and averaged approximately 1,850 mt since 2017. In the southern region, total commercial catch averaged approximately 940 mt through 2005, decreased during the late 2000s and has averaged 450 mt since 2010. Across regions, the majority of commercial catch is landed, but the proportion of the catch that is discarded has increased since 2010, especially in the northern region.

After extensive literature review and analysis (Beaty et al. 2023 working paper), the working group decided to assume 15% discard mortality for handlines, pots and traps and 100% discard mortality for trawl, gillnet and scallop gears.

The primary source of recreational catch data, including annual weight and catch-at-age for both harvest and discard, is NOAA's Marine Recreational Information Program (MRIP) which provides estimates back to 1981. The MRIP program estimates quantities and coefficients of variation (CVs) for harvest weight and discards in numbers via angler interviews and observations on retained fish which occur primarily at shore-side fishing locations. Recreational harvest and dead releases substantially increased in the northern region beginning in approximately 2010; prior to 2010 harvest and releases generally increased but at a modest rate (Truesdell & Curti 2023a working paper).

Recreational fishing effort for black sea bass from party/charter vessels is largely concentrated in nearshore waters from Cape Hatteras to Cape Cod in water depths less than 30 meters. Since 2005, the number of black sea bass trips in Long Island Sound and Southern New England has increased. The distribution of recreational fishing effort has also expanded in deeper waters across the continental shelf in recent years (2015-2021; Curti et al. 2023b working paper).

The size composition for total recreational catch was limited to fish larger than 10 cm and included very few fish larger than approximately 55 cm. Median size of recreational harvest increased over time in both the north and the south and the median size of recreational discards also increased though not as dramatically. Large cohorts were not evident by eye in the length compositions, but after they were converted to ages these year classes, especially 2011 and 2015 in the northern region, were evident in the age compositions (Truesdell & Curti 2023a working paper).

TOR3: Survey Data

"*Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data*."

The working group examined numerous fishery-independent surveys as potential indices of index black sea bass relative abundance, which are described in TOR 3 Survey Data and in the Truesdell & Curti 2023c, Hansell & Curti 2023, Painten et al. 2023, Brust et al. 2023, and Jones et al. 2023 working papers. In the northern region, the surveys explored included: the NEFSC, Northeast Area Monitoring and Assessment Program (NEAMAP), Massachusetts, Rhode Island and Connecticut Long Island Sound spring and fall bottom trawl surveys, the Massachusetts and Rhode Island ventless trap surveys, and the New York Peconic Bay trawl survey (Truesdell & Curti 2023c working paper). In the southern region, the surveys explored included: the NEFSC winter, spring and fall surveys; the NEAMAP spring and fall surveys, the New Jersey bottom trawl survey, the Delaware trawl survey and the Maryland trawl survey. The working group considered incorporating each of the surveys in three ways: using the data directly as a stratified or geometric mean (depending on the survey design), standardizing the indices using generalized linear models, and compiling an aggregate index using a spatiotemporal model (i.e., VAST). After fully vetting each option, the working group decided to move forward with Vector Autoregressive Spatio-Temporal models (VAST) indices to account for time-varying catchability among surveys and the small geographic footprint (and potentially changing availability) of the state surveys in comparison to the range of the stock.

Seasonal VAST models were used to produce both aggregated and age-based distribution and abundance estimates (Hansell & Curti 2023 working paper). VAST model results suggest that black sea bass center of gravity has shifted northeast in the southern region and that their range has expanded poleward. VAST model results further suggest that relative abundance has increased in the northern region and remained stable in the southern region.

In addition to trawl survey indices, the working group also considered a ventless trap survey index (Painten et al. 2023 working paper). The ventless trap survey time series, however, was limited in length and, thus, the working group did not prioritize the inclusion of this index in model runs.

The working group also developed and considered two fishery-dependent indices of abundance: recreational catch per angler (CPA) and commercial CPUE. Black sea bass stock assessments since 2017 have included an abundance index based on recreational CPA. This index provides broad spatial and temporal coverage that is difficult to achieve with federal and state-run fishery independent surveys. After reviewing diagnostics and comparing trends to other possible indices of abundance, the working group decided to include the recreational CPA index in the stock assessment model (Brust et al. 2023 working paper).

In an effort to explore the utility of fine-scale fishery dependent data from the commercial fleet to the black sea bass stock assessment, the working group developed standardized commercial CPUE indices for bottom trawl gear (Jones et al. 2023 working paper). To do this, the working group combined data sets from two fine-scale fishery dependent collection programs: 1) the NEFSC's Study Fleet Program, and 2) the Northeast Fisheries Observer Program. The standardized CPUE indices largely followed the trends of the survey and recreational fishery indices, and provided complementary information about trends in the black sea bass stock. Though the commercial CPUE indices from this effort are not included in any model runs, they are useful as a qualitative 'sense checking' comparison.

TOR4: Stock Size and Fishing Mortality

"*Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and* *(b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied*."

The working group developed two stock assessment models that are described in TOR 4: Stock Size and Fishing Mortality and in the Miller et al. 2023, Miller 2023, and Fay et al. 2023 working papers. The proposed base model uses a multi-stock, multi-region extension of the Woods Hole Assessment Model (WHAM) R package (Multi-WHAM refers to this extension of WHAM) to simultaneously model the northern and southern regions of the stock and movement of fish originating in northern region (see Stock Structure and Spatial Partitioning section for a description of the regions, Miller et al. 2023 and Miller 2023 working papers). Recreational CPA and spring VAST aggregate indices for the northern and southern regions along with corresponding age composition data were used to inform the model. Catch and associated age composition data for regional recreational and commercial fleets were also used. The model also includes effects of a winter bottom temperature covariate on recruitment in the northern region. Process errors in the latent bottom temperature covariate, recruitment, survival, and selectivity of some fleets and indices are estimated as random effects. The working group arrived at the proposed base model from analyzing more than 30 different fits of Multi-WHAM to different sets of observations. The proposed base model exhibits negligible retrospective patterns in fishing mortality or spawning stock biomass (SSB) for either region and one step ahead (OSA) residuals appear adequate for most of the data components.

WHAM outputs indicate that SSB in the northern region averaged approximately 1,300 mt through 2005, beyond which it steadily increased to a maximum of almost 16,300 mt in 2016 and has averaged approximately 13,400 mt since 2017. This consistent and sustained increase in the northern SSB was largely driven by strong 2011 and 2015 year classes. In contrast, SSB in the southern region averaged approximately 3,800 mt before increasing to a peak of 11,200 mt in 2002 as strong 1998, and especially 1999, year classes moved through the population. SSB in the south then decreased back to an average of 4,300 mt through the late 2000s and early 2010s and then steadily increased during the last eight years of the time series to approximately 7,500 mt in 2021. Stock-wide SSB across the northern and southern regions combined was estimated at 22,630 mt in 2021.

Recruitment estimates indicated that year class strength varied substantially between the two regions. In the north, the 2011 and 2015 year classes were the biggest recruitment events of the time series. In the southern region, these year classes were both above the time-series average, but were not of the magnitude observed in the north. In contrast, in the south the largest recruitment events occurred during the beginning of the time series with the 1994 and 1999 year classes. Stock-wide recruitment across the northern and southern regions combined was estimated at 35.2 million in 2021, 95% of the 1989-2021 time series average.

Fully-selected fishing mortality rates have been similar for both regions, ranging across the time series from 0.44-1.31 in the north and 0.24-1.70 in the south. Over the time series, fishing mortality in the north largely varied without trend and averaged 0.71. In the southern region, however, fishing mortality was generally higher during the beginning of the time series, averaging 1.19 through 1997, declined during the late 1990s and has averaged 0.40 since 2001. Fleet-specific fishing mortality rates indicate notable differences between regions, where the southern recreational fishing mortality exhibited the largest fishing mortality of the four fishing fleets through the late 1990s and then generally decreased during the 2000s to an average of 0.24 since 2011. In contrast, fishing mortality rates for the recreational fleet in the north have trended from the lowest of the four fleets during the 1990s, averaging 0.21, to the highest fleet-specific rates since 2009, averaging 0.49. Fully-selected total fishing mortality across all regional fleets was estimated at 1.12 in 2021.

A stock synthesis (SS) modeling approach produced similar results, suggesting that the results are robust to a range of data and model decisions (Fay et al. 2023 working paper). The SS model, however, exhibits strong retrospective patterns in both fishing mortality and SSB.

TOR5: Status Determination Criteria

"*Update or redefine status determination criteria (SDC; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based* *estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs*."

As described in TOR 5 Status Determination Criteria, the Multi-WHAM package was used to develop biological reference points based the most recent 5-year average of age-specific maturity, SSB weight, catch weight, fleet selectivity, and natural mortality estimates to calculate $F_{40\%}$, and the average annual recruitment for years after 1999 to estimate SSB at $F_{40\%}$. The average of recruitments after 1999 for each region were used to weight the region-specific equilibrium spawning biomass per recruit (SPR) estimates to determine the stock-wide unfished SPR and the fishing mortality at 40% of this unfished value. The total estimated fully selected fishing mortality that achieved 40% of unfished SPR was $F_{40} = 1.03$ and values for the north and south were 0.71 and 0.32, respectively. The percentages of unfished SPR for the northern and southern regions were 39% and 41%, respectively. The estimated total equilibrium SSB at F_{40} was 12,491 mt, and for the northern and southern regions, estimates were 6,474 and 6,017 mt, respectively. In 2021, there is a 0.71 probability of $F>F_{40}$ and SSB >0.5 SSB(F_{40}), a 0.29 probability of $F<F_{40}$ and SSB>0.5SSB(F_{40}), and a negligible probability of SSB<0.5 SSB(F_{40}).

The objective of this research track is to develop the assessment and projection methodology that will be used in subsequent management track assessments. As such, stock status recommendations are not part of the research track Terms of Reference and the results from this research track assessment will not be used directly in management. Instead, this research track assessment will inform a management track assessment scheduled for June 2024. The 2024 management track assessment will provide updated estimates of stock status using data through 2023 and will be used to inform management measures for 2025-2026.

TOR6: Projection Methods

"*Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of* *resulting projections considering the effects of uncertainty and sensitivity to projection assumptions*."

The objective of this research track TOR is to develop the projection methodology that will be used in subsequent management track assessments. The working group used WHAM to configure short-term (2022-2024) projections, as described in TOR6 Projection Methods. Following the methods used to estimate reference points under prevailing conditions (TOR5), region-specific average annual recruitment estimates for years after 1999 and the most recent 5-year average of age-specific maturity, SSB weight (by region), catch weight (by fleet), fleet selectivity (by fleet), and natural mortality estimates (by region) were used to conduct short-term projections. Models for random effects on the bottom temperature covariate, recruitment, and survival were used to predict bottom temperature and abundance-at-age in the projection years. Given that this is a research track stock assessment with a focus on methodology, these projection results will not be used directly in management. A management track assessment scheduled for June 2024 will provide updated projections using data through 2023 and will be used to inform management measures for 2025-2026.

TOR7: Research Recommendations

"*Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 2 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations*."

This working group reviewed and prioritized previous and new research recommendations, as described in TOR7 Research Recommendations. High priority research topics include 1)

Movement rates and cues, including research to quantify movement between the northern and southern regions and research on environmental drivers of this movement, 2) Role of varying recruitment and strong year classes in stock dynamics, including drivers of recruitment, 3) Development of reliable indices of abundance beyond existing surveys, 4) Enhanced port sampling or similar programs to bolster the data that support estimation of fishery length and age compositions, and 5) Metrics for measuring recruitment as a response variable to environmental indicators. Medium priority research topics include 1) Environmental drivers of recruitment, 2) Expanded fishery-independent abundance indices, 3) Use of industry study fleet data, 4) Discard mortality rates, particularly for gear types for which there has been limited or no new recent research, 5) Methods for filling bottom temperature data gaps for use as an environmental indicator, including consideration of new data sources and analytical products, 6) Development of a commercial CPUE index, 7) Socioeconomic drivers of recreational and commercial fishing for black sea bass and associated species, 8) Impacts of expansion into the northern range of the stock on fishing behavior, 9) Food web interactions and impacts on stock productivity and 10) Incorporation of a fall VAST index, and 11) Scaling recreational catch CVs. Other research priorities include 1) Further evaluation of the two region structure of the model, 2) Spatial patterns in growth, recruitment, and mortality, 3) Quantification of range expansion, 4) Habitat use and seasonal changes, 5) Sex change, sex ratios, and spawning dynamics, 6) Natural mortality, 7) Precision and uncertainty in discard estimates, and 8) Exploring separate age-length keys by semester, region, and fishery/survey after 2008 when more data are available.

TOR8: Backup Assessment Approach

"*Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment*."

As described in TOR8 Backup Assessment Approach, the working group recommended that if the proposed Multi-WHAM assessment approach does not meet peer review standards, a simpler

WHAM configuration that emulates ASAP (i.e. model with only fixed effects) is used as the backup approach. This fixed-effects ASAP-like WHAM model would still integrate biological, catch, age composition and index information, and therefore, is considered a more informative contingency plan than a purely empirical approach. Following standard practice, a retrospective adjustment would be applied to the terminal year estimates if the rho-adjusted values fall outside of the 90% confidence intervals of the original values.

Summary Report of the Black Sea Bass Research Track Stock Assessment Peer Review

December 5 - 7, 2023
Northeast Fisheries Science Center, Woods Hole, Massachusetts

Report prepared by Panel Members:
Olaf Jensen (Chair), MAFMC SSC
Jean-Jacques Maguire, independent contractor with CIE
Sven Kupschus, independent contractor with CIE
Joel Rice, independent contractor with CIE

Introduction

The Northeast Region Coordinating Council (NRCC)[1] has developed an enhanced stock assessment process to improve the quality of assessments. The process involves two tracks of assessment work: 1) a management track that includes routine updates of previously approved assessment methods to support regular management actions (e.g., annual catch limits), and 2) a research track that allows comprehensive research and development of improved assessments on a stock-by-stock or topical basis. The research track assessment process allows for a more thorough review of information available and for the evaluation of different assessment approaches than would be possible in a standard stock assessment process where the results are immediately used for management advice. This Panel reviewed the Research Track Assessment for the northern stock of black sea bass.

The previous stock assessment for the northern stock of black sea bass (BSB) (https://www.google.com/url?q=https://repository.library.noaa.gov/view/noaa/39406&sa=D&source=docs&ust=1702049662893310&usg=AOvVaw3_x9gT-g1DXYlR1OSKQ1Au) was based on a two independent region-specific Age-Structured Assessment Program (ASAP) models with the division between the northern and southern *stock components* occurring roughly at Hudson Canyon. A separate southern *stock* of black sea bass south of Cape Hatteras, NC is assessed and managed separately and was not the focus of this Research Track assessment. The Black Sea Bass Research Track Working Group (WG) opted to maintain the two-region approach with the same regions but developed new fishery-dependent and fishery-independent indices of relative abundance, tested environmental covariates of recruitment, and explored two modeling frameworks: a multi-region extension of the Woods Hole Assessment Model ("multi-WHAM") and Stock Synthesis (SS).

[1] Atlantic States Marine Fisheries Commission (ASFMC), Greater Atlantic Regional Fisheries Office (GARFO), Mid-Atlantic Fishery Management Council (MAFMC), New England Fishery Management Council (NEFMC), and Northeast Fisheries Science Center (NEFSC).

The work of the WG has been reviewed by the Black Sea Bass Research Track Peer Review Panel that met via Webex from December 5-7, 2023. The Panel included three independent scientists selected by the Center for Independent Experts (CIE): Jean-Jacques Maguire (independent contractor and member of the Scientific and Statistical Committee of the New England Fisheries Management Council), Sven Kupschus (European Commission Joint Research Center, Italy) and Joel Rice (Joel Rice Consulting, USA). The Panel was chaired by Olaf Jensen (University of Wisconsin - Madison and member of the Scientific and Statistical Committee of the Mid-Atlantic Fisheries Management Council).

The Working Group Assessment Report and 18 supporting Working Papers were made available to the panel on the data portal (https://apps-nefsc.fisheries.noaa.gov/saw/sasi/sasi_report_options.php) on November 14, 2023. The Panel was also given access to the GitHub repositories used by the WG where they could access model code, data input files, and model outputs including figures and tables. Individual Panel Members and the Chair took the lead in providing first drafts of various sections of the report, but the entire Panel is responsible for the whole report. Prior to the meeting, members of the Panel met with Michele Traver (NEFSC's Stock Assessment Workshop Process Lead), Larry Alade (Chief, NEFSC Population Dynamics Branch) and Alexander Dunn (Communications Specialist, NEFSC Population Dynamics Branch) to review and discuss the meeting agenda, reporting requirements, meeting logistics and the overall process.

Presentations made by WG members during the review panel are listed in the agenda (Appendix 2) and available as PDFs on the data portal. Other WG members were present and answered questions from the review panel and contributed to the discussions on various topics. Jessica Blaylock, Toni Chute, Giovanni Gianesin, Brian Linton, and Emily Liljestrand acted as rapporteurs throughout the meeting (see Appendix 4 for materials provided and Appendix 5 for meeting attendees). The WG was chaired by Anna Mercer (NEFSC) and included staff from NOAA Fisheries, academia, a non-governmental organization, and state fishery management agencies. Terms of Reference for the WG are provided in Appendix 1.

Panel members and the Chair drafted this Summary Report in a Google Doc. The Panel Chair compiled and edited this Summary Report with assistance (by correspondence) from the CIE Panelists, before submission of a draft report to the WG. The scope of the WG review of the draft was limited to suggesting corrections for errors of fact or requesting that Panel recommendations be clarified. Additionally, each of the CIE Panelists will submit their separate reviewer's reports to the CIE.

The Panel agreed that all nine TORs had been met: TORs 1-3 and 5-9 ***fully met*** and TOR 4 ***partially met***. The Panel agrees that the new assessment framework proposed by the WG (multi-WHAM) is a significant advance from the previous ASAP models and is an acceptable basis for providing management advice, including estimating biological reference points (BRPs) and making projections. However, the Panel recommends conducting additional sensitivity runs (described under TOR 4 and 7) before deciding on a final model configuration for use in the management track assessment.

The Panel's evaluation of the WG's response to the nine TORs is provided below and key recommendations are summarized under TOR 7.

Evaluation of the Terms of Reference for Black Sea Bass

1. Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under impacted TORs.

The Panel agreed that this TOR has been **fully met**.

Considerable work was accomplished by the WG under this TOR and important advancements have been made not only in describing ecosystem change, but also in incorporating quantitative links between ecosystem change and stock dynamics in the assessment model. Two specific accomplishments under this TOR stand out: (1) development of a bottom temperature index and including it in the assessment model to help predict recruitment to the northern region, and (2) developing a time series of annually varying biological reference points that model changes in stock productivity without having to specify the mechanistic basis for these changes. This second accomplishment was not explicitly framed by the WG as belonging to TOR 1, but it substantially advances the underlying goal of incorporating ecosystem change into the stock assessment.

Work under this TOR began with a hypothesis driven exploration of relationships between the marine environment and different aspects of BSB life history. The WG then narrowed in on the relationship between bottom temperature and winter distribution of BSB across the continental shelf. A bottom temperature index was created from a new temperature reanalysis product (du Pontavice et al. 2023) based on an oceanographic model of bottom temperature across the Northeast US shelf. This index was initially tested as a predictor of recruitment through comparison of the recruitment deviations from the 2021 ASAP models for BSB. The strong correlation among these variables led the WG to include bottom temperature as a linear predictor of recruitment in the base multi-WHAM assessment model. The WG conducted a sensitivity run of the multi-WHAM model without this temperature-recruitment relationship and estimated a similar recruitment time series. However, the temperature-recruitment relationship is influential in projecting recruitment and provides a potential link for future incorporation of bottom temperature projections from oceanographic forecasting models.

The WG conducted additional analyses in an attempt to develop environmental indices that could be used as a predictor of mixing between the two regions. Black sea bass have undergone a pronounced northeastward expansion of their spatial distribution over the past 40 years (Bell et al. 2015). The WG considered the possibility that mixing rates may be higher when the centers of gravity in the northern and southern region are closer. A second hypothesis related to winter shelf water volume (SWV) and the seasonal offshore migration of BSB. Based on observations from Miller et al. (2016), the WG considered the possibility that in winters with higher SWV, BSB must travel farther offshore to escape this colder water and the potential offshore winter mixing between the northern and southern stock components is greater. The WG's analyses did not support using either of these relationships in the stock assessment.

The multi-WHAM model itself also allows for incorporation of environmental change into predictions of stock dynamics without explicitly specifying the underlying mechanistic basis. The WG's proposed model includes random effects on recruitment and survival, selectivity, and on the indices of relative abundance. Random effects on recruitment, specified as an autoregressive process, allow for estimation of recruitment trends (and interannual variation) without specifying the environmental driver(s) of recruitment. Similarly, the WG's use of dynamic biological reference point calculation provides a mechanism by which stock status determination and management advice can reflect apparent stock productivity changes (in this case, apparent increases in productivity) without the need to develop explicit environmental covariates of productivity. Black sea bass appear to be among the climate change "winners" (i.e., species whose productivity has increased with warming, Free et al. 2019) and the new assessment framework developed by the WG provides a mechanism to incorporate such change into management advice. Dynamic reference points will, however, present some additional challenges in communicating management advice as they represent an additional source of uncertainty in projections.

In addition to the ecosystem indicator work described above, the WG conducted several additional analyses under this TOR, including: (1) an update of age-length keys used to account for changes in somatic growth, (2) key informant interviews (n=16) with commercial and recreational fishing industry stakeholders, and (3) a comprehensive evaluation of approaches to estimating natural mortality (M) external to the assessment model.

The stakeholder interviews were useful for identifying factors that may have caused changes in catch per unit of effort (CPUE) or selectivity. These interviews generally corroborated estimated changes in selectivity from the assessment model and some of the ecological and ecosystem processes considered in the WHAM model. For example, the age of fully selected fish estimated by the model increased through time for the recreational fleet in the northern region, which is consistent with increasing minimum length limits discussed by recreational anglers.

A new tool for estimating M from life history, taxonomic, and environmental factors (Cope and Hamel (2022) was applied to data for BSB. The value of M used in previous assessments (0.4) was near the center of the distribution of plausible values generated by the Cope and Hamel (2022) tool and the WG concluded that there was insufficient information to justify a change in M from 0.4. The panel noted that not all of the approaches used in the Cope and Hamel tool are equally accurate and future work on this topic should consider alternative weighting methods for arriving at a point estimate of M. For example, Then et al. (2014) reviewed many of these approaches and found that the cross-validation error of methods based on maximum age was approximately half that of methods based on growth model parameters. The panel also recommended additional sensitivity runs of the multi-WHAM model with alternative plausible values of M given that M is relatively poorly estimated and is often an influential fixed value in stock assessment models. This recommendation is discussed in more detail under TOR 4.

2. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

The Panel agreed that this TOR has been **fully met**.

The WG comprehensively addressed the TOR in its work. The panel particularly endorses the efforts to maintain the maximum contrast in cohort signal whenever possible and the approach of automation of procedures for effective and consistent application during the management track process. The panel notes that the WG's efforts appeared focused towards the application by the Multi-WHAM assessment model and that for other model applications other options may have been possible with different risks and benefits. The WG conducted an extensive analysis of the available commercial catch information for the BSB stock. Limited reliable age composition data were available before 1989. Data from 1989 onwards demonstrated contrast between the northern and southern stock components as well as the ability to identify large and small cohorts (particularly the large 2011 and 2015 cohorts in the northern region) consistently as already suggested by the previous assessment process using ASAP. For this reason the WG focused its effort on maintaining these aspects of the data in order to support a more modern stock assessment method aimed at dealing with some of the shortcomings (e.g., strong retrospective patterns) of the previous ASAP models. The panel agrees with the focus on maintaining the cohort structure in the ALK for use in Multi-WHAM.

Age-Length-Keys

The WG developed a single two-area, bi-seasonal, conditional age length key from all data sources. While ideally one would retain fleet specific age information, paired age / length samples were sparse at the beginning of the time series and numbers for gear, area and season year combinations were too low to provide reasonable age compositions. The WG prioritized cohort consistency by developing an all fleets ALK and aggregating fish > 35cm across seasons as seasonal growth differences were small by this size. This leaves the length structure to account for differences in fleet selectivities.

The panel felt this was an appropriate treatment of the data and agreed with the WG that the risk of fleets targeting specific ages within a mixed length distribution in a population with good overlap between age distributions is very small.

Where less than ca 250 age-length pairs were available borrowing of sample information from 'proximal' samples was implemented in the order of region, semester, and region and semester to preserve cohort and spatial structure in the assessment input information.

Length samples from the commercial sector

The WG hoped to maintain fleet specific age compositions in order to be able to model the selectivities independently. However, historically this was not possible due to a lack of available length samples. The commercial catch is sampled by market category and the lack of consistent coverage of categories made raising of those catches problematic. In recent data with comprehensive temporal coverage for all fleets, it was found that the length distributions between market categories varied considerably more than the variation among gears within the market categories. Category therefore served as a more reasonable proxy of selectivity. Therefore, samples and landings were combined across gears for raising, resulting in a single commercial fleet.

The panel supported this decision to combine across gears as it focuses on retaining the contrast in cohort structure, while having a minimal effect on the model accuracy. First the available evidence provided does not indicate substantial spatial separation in distribution of ages past the recruitment age and up to age 3 (the last modeled age of selectivity). The panel concluded there should be sufficient information in the length composition data for the model to be able to cope with the assumption of a single commercial fleet.

Commercial landings

Landings data were treated as census data, but unfortunately the location information has a different resolution than the region division at the Hudson Canyon. Statistical units spanning the Canyon were therefore assigned to either the south or the north region.

Commercial discards

Commercial discards were assigned 100% mortality for trawl and gillnet fleets and 15% mortality for pots/traps and handlines. Both the occurrence (due to regulation and economics) and the data availability of discards has increased in the time series.

Sampling data from observer programs also increased in recent years. The same alk aggregation procedure / prioritization was followed as for the retained portion of the catch but often greater levels of aggregation / borrowing was required to reach the minimum sample thresholds. meaning aggregation was necessary over greater numbers of domains for a larger number of area season year combinations.

The panel felt reassured by the consistency of the cohort structure of the data raised in this way suggesting the aggregation had little impact on modeled population dynamics, but a common ALK was used which may then provide a false sense of reliability. Nevertheless, the panel felt raising of the discard biomass to the length structure was a sensible method of raising the data.

Recreational retained catch

Recreational length compositions and their uncertainty for retained fish were taken predominantly from the MRIP intercept survey with some minor supplementation from other sources. Largely following the design-based estimates associated with the sampling design for the retained component. The panel noted that, in the northern region of the stock in particular, a large portion of the fishing mortality originates from the recreational fishery.

Recreational dead discards

Because recreational monitoring activities are largely shore-based, information on the discarded component is available only as total numbers released and not in weight nor at-length. The WG concluded that the best estimate for the discard length composition was the observer data. These data are available only since 2004 and only from the headboat (party) sector. Although this represents a relatively small proportion of the total releases the WG made the assumption that the for-hire (head boat samples collected at sea) were representative of all recreational releases. There are differences between the head boat sector and other recreational sectors, but the WG felt that there were no compelling reasons to suspect differences among recreational sectors in discard length composition. Recreational data are only available at the resolution of the state so the Hudson Canyon boundary was implemented only approximately with states assigned to either the north or south subregion despite recognition that anglers (particularly in NY and NJ) sometimes fish in one region but land in the other.

Recreational release estimates are provided as individuals at length, whereas the input to the WHAM model currently requires aggregate removals in weight and compositional data (proportion of numbers-at-age). Therefore, an all length-weight data combined length weight relationship was used for the conversion.

3. Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data.

The Panel agreed that this TOR has been **fully met**.

The WG addressed all aspects of the TOR. The panel notes that the majority of documentation and presentations focused on the evaluation of the WG proposed assessment methodology (Multi-WHAM) and that evaluation of other methods (e.g. ASAP, SS3) was limited and largely restricted to TOR 4 through the assessment diagnostics.

Data from 10 fishery-independent surveys covering the stock area were available covering spring-fall and north-south components of the stock although it is noted that the distribution of the surveys in these strata is not even. Not all surveys provided associated age information and where lacking these were imputed from length distributions using the general ALK. A number of covariates were included in standardization models both GLMs (individual surveys) and VAST (single index). The WG decided early on to stay with the resolution of the previous ASAP process so surveys were grouped to provide spring and fall indices in the south and in the north. The VAST models used the data combined over regions to estimate abundance, but the results were subsequently split between north and south for assessment purposes.

Most of the presentation and discussions focused on the development of the VAST model as the WG decided that this was their preferred method for incorporating all of the surveys into the WHAM model. In addition, it was discovered just prior to the review that the fall VAST index had been incorrectly adjusted for the presence of 0-age fish so that this was not reviewed.

VAST is now widely used in the US and elsewhere to standardize indices across multiple surveys and as such has been extensively reviewed in general so the approach was accepted by the review panel as an appropriate method to reduce conflicts among indices in the assessment model. However, relatively little information in terms of model diagnostics was presented at the panel review and the index is used as an age-based index and the age-specific spatial results were provided during the review which made an independent evaluation difficult. Most of the evaluation is based on the consistency runs where the different index formulations were compared in the ASAP and multi-WHAM models (TOR 4), after which further model runs focused on the VAST index.

Although the general application of VAST is at least statistically sound there are some concerns in its direct application here. These are:

1) The inclusion of environmental covariates in the model is not entirely clear. While a lot of emphasis was placed on the center of gravity of the population for which inclusion of temperature may be appropriate, the purpose of the index is to inform the model regarding abundance. Here temperature should only be used to account for variation in conditions sampled (due to random sampling) not in the systematic change in the conditions as we might expect from climate change. Having accounted for temperature differences in the index it then seems inconsistent to look for these changes in the assessment model. While the index is based on predictions rather than the year effect in the model which uses temperature as a covariate, the results are dependent on the suitability of the temperature fit and the models ability to predict the temperature at the node points.

2) The treatment of the different surveys appears from the results to act mainly through a single scalar as opposed to age specific ones, although the panel was told this is implemented in the VAST application. Therefore, potential differences in selectivity between the different surveys may be underestimated and with the strong weighting by area the large offshore surveys would then present biased indices of the age structure. For the spring survey this is less of an issue as most individuals are found offshore, but the index from the fall survey, which occurs while BSB are migrating offshore, will likely suffer significantly from this issue. However, this could not be tested since although planned, an error was discovered so the correct data was not available to the panel. When a disproportionate part of the population is located in one or more areas, surface area alone is an inappropriate weighting metric so should not be applied without considering per area densities.

3) Density distributions by age for the two VAST indices do not show clear interannual shifts in the spatial distribution plots by age (provided to the panel during the m and surprisingly little segregation between ages, but do track cohorts reasonably well in scale across the different years, particularly in the spring survey. This suggests either the proposed large environmental impacts of temperature and shelf water volume which were implied by the WG from the raw data were overemphasized, the VAST implementation was

unintentionally able to assign this variance to covariates other than the spatial realm or VAST was too constrained to be able to follow the differences in distribution between years. Addressing this last possibility presumably would require an interaction term between the year and spatial effects.

A more in depth analysis of the VAST model developed, particularly with regards to the impact of the various data sources and covariate effects would have helped the panel better understand the suitability of the application for the intended purpose. Few diagnostics were included of the VAST models themselves and evaluation was mainly restricted to a comparison of the stock dynamics (SSB, F and recruitment) derived from the WHAM model in comparison with the ASAP model previously used.

Recreational Catch per Angler (CPA)

The WG revised the methodology for a previously available Recreational CPA index (used since 2016 in the assessment process), to reduce potential for hyperstability. Much of the focus was therefore on the identification of trips that could plausibly have caught BSB. The Jaccard method previously employed was evaluated against a number of different methods aiming to increase robustness with regards to ecosystem processes such as prevalence of other species and the northern range expansion of BSB. The corrected indices provided very similar results in terms of the standardized guild composition but the log-odds ratio method was eventually preferred due to the greater resolution on the appropriate cut-off values for targeted versus not targeted trips and visual inspection of the diagnostics.

The catch (retained plus all discarded individuals) from the recreational monitoring programs was used to assign catches to the identified effort and these were modeled by a GLM with effects of Year + State + Wave (season) + Mode (shore, private boat and party boat) + Area (N-S nested within state). The WG noted that confidence intervals (CIs, 95th percentile) were estimated via bootstrapping using 500 iterations for each region. The resulting CI of the index was extremely tight, i.e. close to the mean, presumably due to the large sample numbers. The multi-WHAM model adjusted for this perceived underestimation of the recreational CPA coefficient of variation (CV) by estimating a scaling factor for this CV.

The panel considered the change in effort estimation positive and justified, but had the usual concerns of recreational CPUE indices in general being susceptible to hyperstability. The concerns were somewhat alleviated by the consistency of patterns in the rec CPA index with other indices. As in other data sources the contrast in the data in the northern area is large and may mask finer scale hyperstability issues as abundance in the area reaches a plateau. The panel recommends that the management track process continue to examine the Recreational CPA index when updated annually for signs of hyperstability which can arise from a wide variety of factors, many of which cannot be simply addressed through better processing/estimation of the index.

Commercial CPUE index:

The WG commendably explored the development of a commercial CPUE index for the research track review. While the index is not extensively included in the assessment model exploration it does represent an approach to balance the weighting in the assessment between the recreational and commercial fisheries, potentially helping to improve the information on the differences in the selectivities either through shared selectivities with the fleets or through development into an age based index.

The estimation of effort follows a similar procedure to the recreational index although it is noted that the uncorrected Jaccard index is still used here. Commercial CPUE is derived by haul from the NMFS Northeast Fisheries Observer Program (NEFOP) and the Study Fleet Program. A number of covariates relating to haul are provided by these sampling programs and other environmental and socio economic parameters are added post-hoc through the available covariates for the purpose of standardization.

The standardization method applied is a generalized additive model (GAM) using a Tweedy distribution applying splines, with location modeled in two dimensions (*s(Latitude,Longitude)*). While the variables used seem relevant to the standardization, there is considerable collinearity in these variables which may reduce the effectiveness of the standardization to remove bias and in fact can introduce biases. The splines are poorly informed at the terminal ends of the range (for example depth) which means they are rather susceptible (less certain) as they essentially represent extrapolations. Future data may therefore considerably alter the effects and may readjust the index over time in subsequent assessment updates.

The panel felt the efforts made in developing the index were highly informative and strongly support further development for future assessments. While the WG determined that the resulting index is currently not suitable for inclusion in the assessment process, it can provide qualitative information for the development and treatment of the catch data in the assessment as well as introduce a greater understanding of the drivers of the commercial fishery to evaluate the suitability of the final assessment approach. Particularly the fine spatial scale of the fishery catches is a valuable asset which could have been more advantageously used in the assessment development.

4. Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and (b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied.

The Panel concluded that ToR 4 had been **partially** met.

The resulting assessment is accepted for use in subsequent management track processes subject to the recommendations under ToR 7 (below) being addressed.

The WG had established a rationale for input and parameter selection that was clearly described, researched and documented. What was not shown was how sensitive or fragile the model was to the selection of the inputs (CPUE and parameter values). The WG analyzed state and federal survey data, recreational catch per angler (Rec CPA), and compiled an aggregated VAST index of abundance. Based on this analysis the WG selected what were perceived to be the indices of abundance that most likely represented the true stock dynamics (VAST and Rec CPA). The WG did not sufficiently explore how sensitive the final Multi-WHAM model is to the inclusion of either of the indices (i.e. a 'leave one out' run). Similarly, the choice of the value for natural mortality was consistent with previous assessments and logical based on the analysis presented, but the impact of this parameterization on the stock status and trend was not explored for the WG's preferred multi-WHAM model as it was for the SS3 model.

At a minimum a limited exploration of the structural uncertainty with respect to the WG selected inputs (indices of abundance) and parameterization should be explored and presented so that the resulting effect on status determination could be evaluated. For example, Punt et al. (2021) noted that natural mortality rates are often considered to be among the most important parameters in a stock assessment, but they are also among the most difficult parameters to estimate using commonly available data. As reported in Table 1 run 13 of the Miller et al. (2023) WG paper, the multi-WHAM run that attempted to estimate natural mortality did not converge. The panel notes that it was difficult to discern how robust or sensitive the model was to this parameterization of M.

The **Panel recommends** that the WG conducts sensitivity analyses including: (1) an exploration of alternative parameterizations for natural mortality (e.g. different age-independent constant values, or age-dependent M), (2) profiles of the initial fishing mortality (i.e. initial depletion) and, (3) an evaluation of which individual surveys should be included in the VAST index by comparing WHAM estimates (e.g., biomass time series) from the proposed run with individual fishery independent surveys. Surveys that do not appear to accurately reflect changes in stock size through this analysis should not be included in the VAST index.

The multi-WHAM framework and application of multi-WHAM for assessment of the BSB stock was presented through the relevant section in the main report under Tor 4, as well as multiple working papers (Miller 2023, Miller et al. 2023) along with a helpful and comprehensive presentation to the Panel meeting. The Panel appreciated the extensive description of research that had gone into the assessment formulation and testing.

Model fit diagnostics that were presented included a jitter analysis, one step ahead (OSA) residuals retrospective patterns analysis, self tests and mean absolute scaled error (MASE) as described in Kell et al. (2021). The diagnostics indicated that the proposed base run is likely appropriate for developing a status determination, pending the outcome of the recommended additional sensitivity runs. Note that in this research track assessment a status determination is not requested/required.

The choice of a temporal change in selectivity from several fully selected ages to primarily the oldest ages for the northern recreational fleet corresponds well with the regulatory changes that repeatedly increased size limit over time. The survey data were aggregated via VAST indices and the working group stated that this accounts for changes in catchability in those fleets over time and should be used in the base model. Panel members inquired as to whether individual state and federal trawl survey indices may better track individual portions of the population, and the working group stated that due to the interactions of the limited geographic footprint of many of the surveys with the black sea bass seasonal migration patterns the VAST estimates were perceived to be a better choice. A comparison of the last run to use the individual state and federal trawl survey indices (bridging run 7) and a model run with the aggregate VAST survey and other model improvements (run 34) shows broadly similar trajectories and scale but some divergent trends for the north after 2014 (Figure 1).

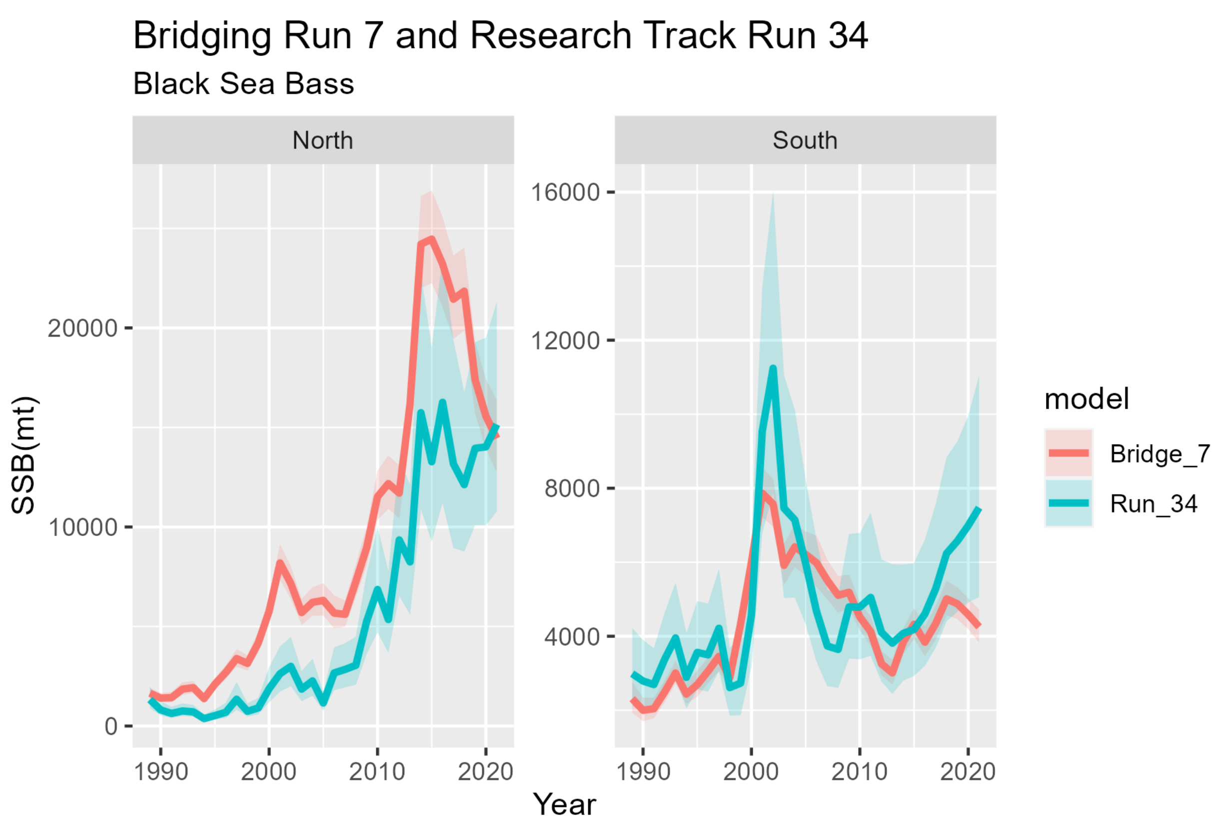


Figure 1. A comparison of using individual survey indices in the Bridge Run 7 (Bridge_7) and the proposed base case from the research track (Run_34). Shaded areas indicate a 95% confidence interval. Note that Bridge Run 7 did not estimate a scalar on the CV of the Rec CPA.

The Panel agrees that the new assessment framework proposed by the WG (multi-WHAM) is a significant advance from the previous ASAP model and is an acceptable basis for providing management advice, including estimating biological reference points (BRPs) and making projections. However, the Panel recommends conducting the additional sensitivity runs described above before deciding on a final model configuration for use in the management track assessment.

5. Update or redefine status determination criteria (SDC; point estimates or proxies for BMSY, $B_{THRESHOLD}$, FMSY and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs.

The Panel agreed that this ToR was **fully met**.

The previous biological reference points for black sea bass are from the 2021 Management Track Assessment. The previous BSB assessment reference points were calculated using the non-parametric yield and SSB per recruit long-term projection approach (NEFSC 2021). That assessment concluded that the black sea bass stock was not overfished and overfishing was not occurring in 2019 relative to the updated biological reference points. The reference points are $F_{40\%}$ as the proxy for F_{MSY}, and the corresponding $SSB_{40\%}$ as the proxy for the SSB_{MSY} biomass target.

The approach used by the WG for the Multi-WHAM base model reference points was based on the most recent 5-year average of age-specific maturity, catch weight, fleet selectivity, and natural mortality estimates to calculate $F_{40\%}$, along with the average annual recruitment for years after 1999 to estimate SSB at $F_{40\%}$ (Miller et al. 2023). Based on this approach, the stock-wide F40% is based on a stock-wide unfished SPR that represents a weighted average of the region-specific unfished SPR estimates. The WG report notes that multi-WHAM considers productivity to vary over time and provides "annual estimates of SPR-based reference points that use the annual inputs to the per-recruit calculations for F at a specified percentage of unfished spawning biomass per recruit. Annual estimates of $F_{40\%}$ and SSB at $F_{40\%}$ are provided as well as the status of annual F and SSB estimates relative to these reference points." This differs from the previous assessment in that the approach used with Multi-WHAM the stock-wide F40% is based on a stock-wide unfished SPR that represents a weighted average of the region-specific unfished SPR estimates as opposed to the previous assessments where a stock-wide F40% was based on the average of the region-specific F40% estimates.

The WG report notes that "Total SSB across regions has been above the annual SSB (F40%) reference points since 2014, and the combined fully selected fishing mortality has been near (either slightly above or slightly below) the annual F40% reference point since 2011" (Miller et al. 2023). Consistent with the past assessment (2021, Figure 2) the current model shows a general increasing trend in $SSB/SSBF_{40\%}$, along with a general decrease in $F/F_{40\%}$ over the temporal domain of the model (Figures 3 and 4). In contrast to the previous ASAP model, the proposed (2023) base case shows a fluctuating but relatively stable population since 2014, in contrast to the 2021 (previous) assessment which showed the population experiencing a steep decline in SSB in the years following 2014.

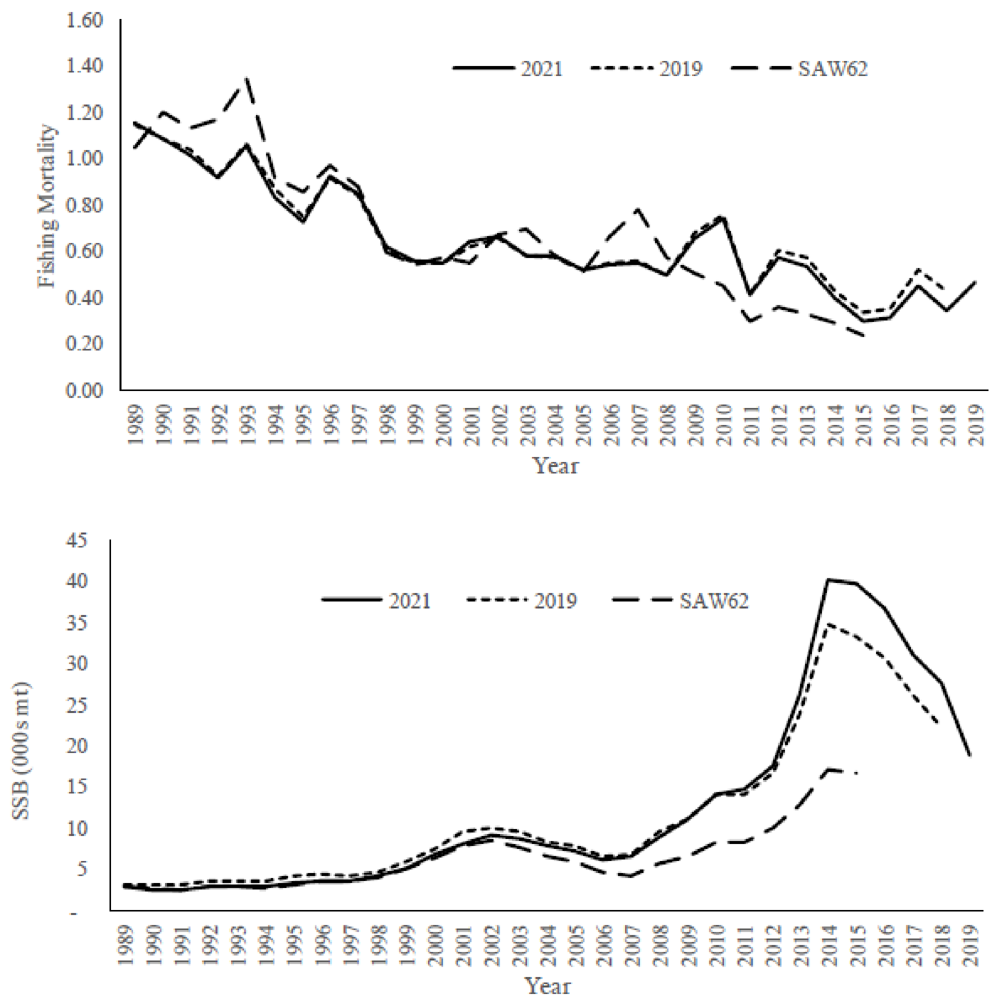


Figure A5. Historical retrospective of the 2016 (SAW 62; NEFSC 2017), 2019 and 2021 (Operational Assessment) stock assessments of black sea bass. The heavy solid lines are the 2021 Operational Assessment estimates. SAW62 did not include revised MRIP estimates.

Figure 2. Figure A5 from the 2021 Operational Assessment of BSB.

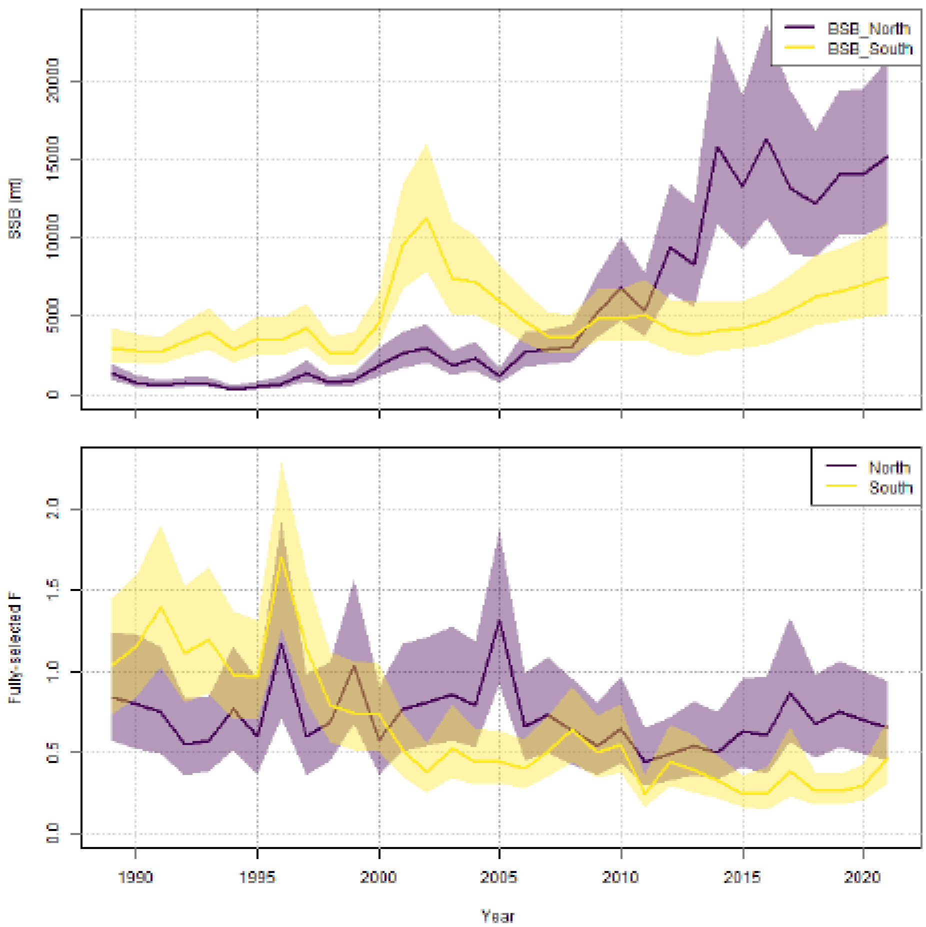


Figure 4.2: Estimated spawning stock biomass (top) and fully-selected fishing mortality (bottom) for 1989-2021 in the northern (purple line) and southern (yellow line) region. Polygons represent 95% confidence intervals.

Figure 3. Working Group assessment report Figure 4.2

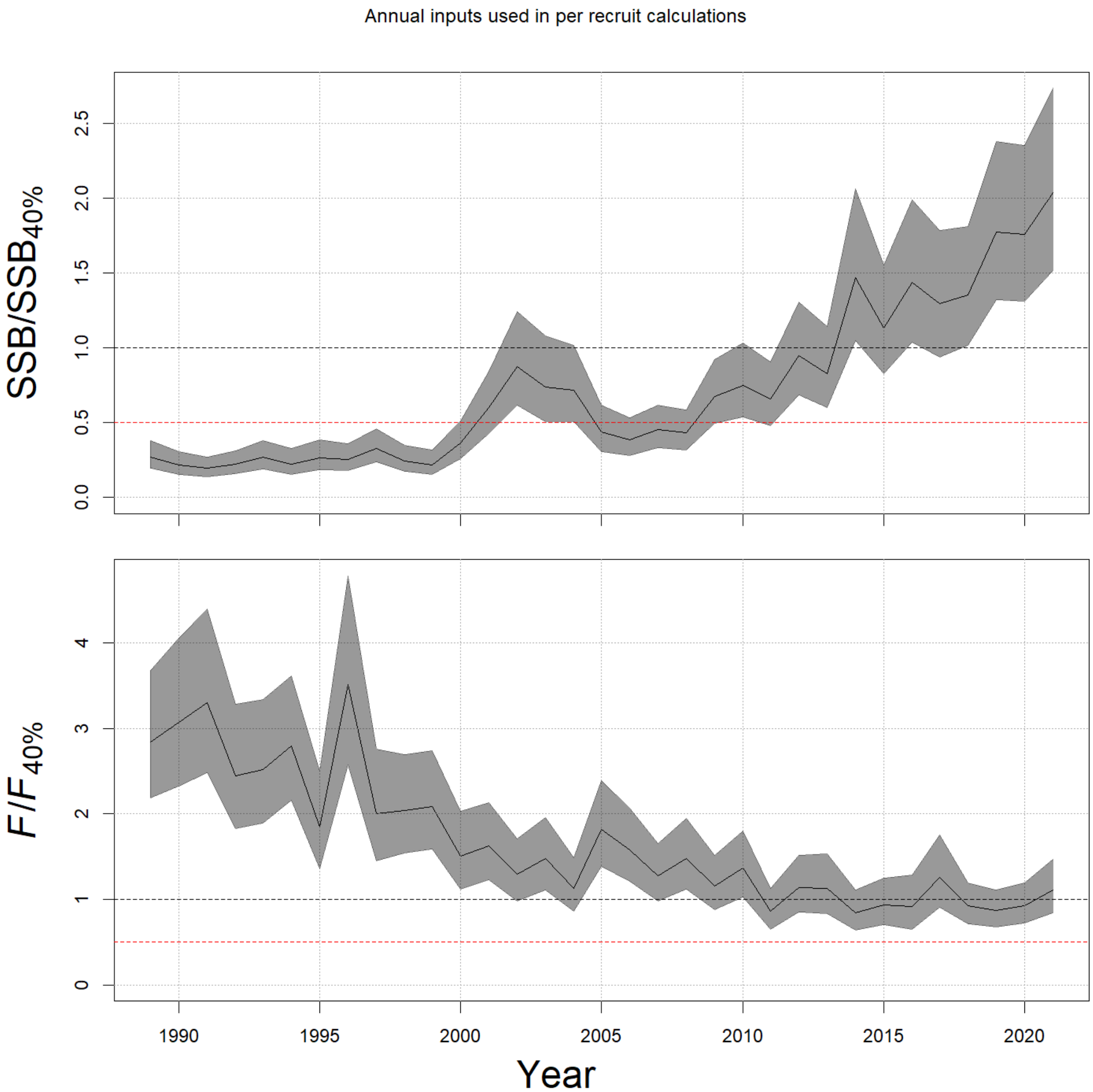


Figure 4. 2023 WG assessment report Figure 5.1. Status of total spawning stock biomass (top) and total fully-selected fishing mortality rates (bottom) relative to annual reference point estimates for 1989-2021. Gray polygon represents 95% confidence intervals.

The WG assessment report (Miller et al. 2023) noted that:

> "The objective of this research track is to develop the assessment and projection methodology that will be used in subsequent management track assessments. As such, stock status recommendations are not part of the research track Terms of Reference and the results from this research track assessment will not be used directly in management. Instead, this research track assessment will inform a management track assessment

scheduled for June 2024. The 2024 management track assessment will provide updated estimates of stock status using data through 2023 and will be used to inform management measures for 2025-2026."

6. Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of resulting projections considering the effects of uncertainty and sensitivity to projection assumptions.

The Panel agreed that this ToR was **met** for Black Sea Bass.

The WG recommended that the suggested assessment model framework for Black Sea Bass, Multi-WHAM, which can do short-term projections internally, should be used for short term projections based on the proposed candidate model run. The assumptions of recruitment, growth, maturity, natural mortality, and selectivity used to make stochastic projections of stock size and catches for 2022-2024 use the same approach as used for the definition of reference points under ToR 5. Models for random effects on the bottom temperature covariate, recruitment, and survival are used to predict bottom temperature and abundance-at-age in the projection years. Region-specific average annual recruitment estimates for years after 1999 and the most recent 5-year average of age-specific maturity, SSB weight (by region), catch weight (by fleet), fleet selectivity (by fleet), and natural mortality estimates (by region) were used to conduct short-term projections. Random effects in the projections revert to the mean after a few years. The panel has no recommendation to change the approach suggested by the WG.

7. Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 2 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.

The Panel agreed that this ToR has been **fully met**.

The WG thoroughly reviewed previous recommendations and updated their status. The WG also made new research recommendations classified as High priority, Medium priority and Low priority. The Panel suggests below a slightly different version of the prioritization of new research recommendations, including those from the panel, as: i) necessary for the management track, ii) high priority, iii) medium/long term and iv) low priority. Within each category, the order of the recommendations represents the Panel's suggestions.

Necessary for management track

Conduct sensitivity analyses including:

1) an exploration of alternative parameterizations for natural mortality (e.g. different age-independent constant values, or age-dependent M)

2) profiles of the initial fishing mortality (i.e. initial depletion)

3) an evaluation of which individual surveys should be included in the VAST index by comparing WHAM estimates (e.g., biomass time series) from the proposed run with individual fishery independent surveys. Surveys that do not appear to accurately reflect changes in stock size through this analysis should not be included in the VAST index.

High Priority

1. Examine the updated CPA indices for signs of hyperstability which can arise from a wide variety of factors, many of which cannot be simply addressed through better processing/estimation of the recreational CPA index.
2. Conduct additional research on scaling the recreational catch CVs to improve confidence in these data and the resulting CPA indices.
3. Further consider the development of a commercial CPUE index. The index reviewed by the WG includes data from a broad area, it can account for socioeconomic drivers of catch, and can be a useful tool for understanding changes in abundance and fisheries operations.
4. Develop a method to fully utilize all available fishery-dependent size data (e.g. from the Commercial Fisheries Research Foundation's black sea bass research fleet) even if it does not include market categories.
5. The WG developed dynamic reference points as output from the assessment. While current stock status has a relatively clear interpretation, the aim is for managers to maintain good stock status. With dynamic reference points, future stock status can systematically change without change in conventional estimates of MSY as interpreted by managers. This presents a challenge of trying to hit a moving target without knowledge of speed and direction of the target. The WG should provide managers with guidance on how to interpret this information to maintain a healthy stock.

Medium/long term

1. Evaluate the impact to the assessment model outputs of enhanced or diminished port sampling in the future to evaluate impacts of changes to data streams that support estimation of fishery length and age compositions.
2. Further consideration of the appropriate metrics for measuring recruitment as a response variable to environmental indicators.
3. Additional research into environmental drivers of recruitment.
4. Explore ways to fill gaps in bottom temperature data for use as an environmental indicator, including consideration of new data sources and analytical products.
5. Examine guidelines for integrating fishery-dependent indices in assessments developed by ICCAT to determine whether they could be useful for the BSB assessment. (https://www.iccat.int/Documents/CVSP/CV074_2017/n_2/CV074020404.pdf).

Low priority

1. Further evaluation of the socioeconomic drivers of recreational and commercial fishing for black sea bass and associated species.
2. Further evaluation of how expansion into the northern range of the stock may impact fishing behavior.
3. Explore separating age-length keys by semester, region, and fishery/survey after 2008 when more data are available.

8. Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment.

The Panel agreed that this ToR was **fully met**.

The Index-Based Research Track Working Group simulation-tested the performance of several empirical Index Based Methods (IBMs) (NEFSC 2020, Legault et al. 2023) and concluded that empirical methods such as Ismooth did not perform better than statistical catch-at-age models that required retrospective adjustment (e.g., the previous ASAP model used in the 2021 BSB assessment). The WG recommended that if the proposed multi-WHAM assessment model is rejected, an alternative simpler multi-WHAM model without random effects parameterized to mimic the previously accepted ASAP model should be used with a retrospective adjustment applied to the terminal year estimates of F and SSB. The Panel agrees that the proposed multi-WHAM model is acceptable (after evaluation of sensitivity runs recommended under TOR 4 are conducted) and that the alternative ASAP-like multi-WHAM model is likely to present worse diagnostics and performance than the proposed multi-WHAM model.

9. Identify and consider any additional stock specific analyses or investigations that are critical for this assessment and warrant peer review, and develop additional TOR(s)* to address as needed.

No additional TORs were developed by the WG.

References

Bell, R.J., Richardson, D.E., Hare, J.A., Lynch, P.D. and Fratantoni, P.S., 2015. Disentangling the effects of climate, abundance, and size on the distribution of marine fish: an example based on four stocks from the Northeast US shelf. ICES Journal of Marine Science, 72(5), pp.1311-1322.

Cope, JM, Hamel, OS. 2022. Upgrading from M version 0.2: An application-based method for practical estimation, evaluation and uncertainty characterization of natural mortality. Fisheries Research 256

Free, C.M., Thorson, J.T., Pinsky, M.L., Oken, K.L., Wiedenmann, J., Jensen, O.P. 2019. Impacts of historical warming on marine fisheries production. Science. 363: 979-983.

Kell, L.T., Sharma, R., Kitakado, T., Winker, H., Mosqueira, I., Cardinale, M., Fu, D., 2021. Validation of stock assessment methods: Is it me or my model talking? ICES Journal of Marine Science 78, 2244–2255. https://doi.org/10.1093/icesjms/fsab104

Legault, C.M., Wiedenmann, J., Deroba, J.J., Fay, G., Miller, T.J., Brooks, E.N., Bell, R.J., Langan, J.A., Cournane, J.M., Jones, A.W., Muffley, B. 2023. Data Rich but Model Resistant: An Evaluation of Data-Limited Methods to Manage Fisheries with Failed Age-based Stock Assessments. Can. J. Fish. Aquat. Sci. 80: 27–42.

Miller AS, Shepherd GR, Fratantoni PS (2016) Offshore Habitat Preference of Overwintering Juvenile and Adult Black Sea Bass, *Centropristis striata*, and the Relationship to Year-Class Success. PLoS ONE 11(1): e0147627. doi:10.1371/journal.pone.0147627

du Pontavice, Hubert, Zhuomin Chen, and Vincent S. Saba. "A high-resolution ocean bottom temperature product for the northeast US continental shelf marine ecosystem." Progress in Oceanography (2023): 102948

Punt, A.E., Castillo-Jordán, C., Hamel, O.S., Cope, J.M., Maunder, M.N. and Ianelli, J.N., 2021. Consequences of error in natural mortality and its estimation in stock assessment models. *Fisheries Research*, *233*, p.105759.

Then, A.Y., Hoenig, J.M., Hall, N.G., Hewitt, D.A. 2015. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. ICES Journal of Marine Science, 72(1), pp.82-92.

Appendix 1 - Terms of Reference for Black Sea Bass Research Track Stock Assessment

1. Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under impacted TORs.

2. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

3. Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data.

4. Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and (b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied.

5. Update or redefine status determination criteria (SDC; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs.

6. Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of resulting projections considering the effects of uncertainty and sensitivity to projection assumptions.

7. Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 2 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.

8. Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment.

Appendix 2 – Initial agenda for Black Sea Bass Research Track Assessment Peer Review meeting, December 5-7, 2023.

Tuesday, December 5, 2023

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:15 a.m.	Welcome/Logistics Introductions/Agenda/ Conduct of Meeting	Michele Traver, Assessment Process Lead Larry Alade, Acting PopDy Branch Chief Olaf Jensen, Panel Chair	
9:15 a.m. - 9:45 a.m.	Introduction/Executive Summary	Anna Mercer (WG chair)/Kiersten Curti (assessment lead)	Biology, movement, management overview, flag areas of major progress in the RT (new data sources, indices, M exploration, discard mortality exploration, new model, ESP, etc)
9:45 a.m. - 10:30 a.m.	Term of Reference (TOR) #2	Julia Beaty, Kiersten Curti	Discard Mortality, Commercial catch CFRF Research Fleet data
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:30 a.m.	TOR #2 cont.	Kiersten Curti, Sam Truesdell, Julia Beaty	Recreational catch
11:30 a.m. - 12:00 p.m.	Discussion/Summary	Review Panel	
12:00 p.m. - 12:15 p.m.	Public Comment	Public	

Time	Topic	Presenter(s)	Notes
12:15 p.m. - 1:15 p.m.	Lunch		
1:15 p.m. - 2:45 p.m.	TOR #3	Kiersten Curti, Sam Truesdell, Alex Hansell	NEFSC BTS, NEAMAP, State Surveys, Ventless Trap Survey, VAST indices
2:45 p.m. - 3:00 p.m.	Break		
3:00 p.m. - 3:45 p.m.	TOR #3 cont.	Jeff Brust, Andy Jones	Recreational CPA and Commercial CPUE
3:45 p.m. - 4:00 p.m.	Discussion/Summary	Review Panel	
4:00 p.m. - 4:15 p.m.	Public Comment	Public	
4:15 p.m.	Adjourn		

Wednesday, December 6, 2023

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:05 a.m.	Welcome/Logistics Introductions/Agenda	Michele Traver, Assessment Process Lead Olaf Jensen, Panel Chair	
9:05 a.m. - 10:30 a.m.	TOR #1	Scott Large, Kiersten Curti, Jason McNamee, Anna Mercer	Time varying growth and maturity, Spatiotemporal modeling, Ecosystem indicators, Trophic ecology, Natural Mortality, Stakeholder engagement
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 12:45 p.m.	TOR #4	Tim Miller Kiersten Curti	WHAM
12:45 p.m. - 1:45 p.m.	Lunch		
1:45 p.m. - 2:45 p.m.	TOR #5	Tim Miller Kiersten Curti	Reference Points
2:45 p.m. - 3:30 p.m.	TOR #6	Tim Miller, Kiersten Curti	Projections
3:30 p.m. - 4:00 p.m.	Discussion/Summary	Review Panel	
4:00 p.m. - 4:15 p.m.	Public Comment	Public	
4:15 p.m.	Adjourn		

Thursday, December 7, 2023

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:05 a.m.	Welcome/Logistics Introductions/Agenda	Michele Traver, Assessment Process Lead Olaf Jensen, Panel Chair	
9:05 a.m. - 10:15 a.m.	TOR #4 cont'	Gavin Fay Jason McNamee	SS
10:15 a.m. - 10:30 a.m.	Break		
10:30 a.m. - 10:45 a.m.	TOR #8	Kiersten Curti	Summarize WHAM recommended model; Alternative Assessment Approach
10:45 a.m. - 11:30 a.m.	TOR #7	Julia Beaty	Research Recommendations
11:30 a.m. - 12:00 p.m.	Discussion/Summary	Panel	
12:00 p.m. - 12:15 p.m.	Public Comment	Public	
12:15 p.m. - 1:15 p.m	Lunch		
1:15 p.m. - 4:00 p.m.	Report writing	Panel	
4:00 p.m.	Adjourn		

Appendix 3 - Performance Work Statement (PWS) - Center for Independent Experts (CIE) Program – Black Sea Bass Research Track Peer Review

Background
The National Marine Fisheries Service (NMFS) 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). NMFS 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, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards .

Scope

The Research Track Peer Review meeting is a formal, multiple-day meeting of stock assessment experts who serve as a panel to peer-review tabled stock assessments and models. The research track peer review is the cornerstone of the Northeast Region Coordinating Council stock assessment process, which includes assessment development, and report preparation (which is done by Working Groups or Atlantic States Marine Fisheries Commission (ASMFC) technical committees), assessment peer review (by the peer review panel), public presentations, and document publication. The results of this peer review will be incorporated into future management track assessments, which serve as the basis for developing fishery management recommendations.

The purpose of this meeting will be to provide an external peer review of the black sea bass stock. The requirements for the peer review follow. This Performance Work Statement (PWS) also includes: Annex 1: TORs for the research track, which are the responsibility of the analysts; Annex 2: a draft meeting agenda; Annex 3: Individual Independent Review Report Requirements; and Annex 4: Peer Reviewer Summary Report Requirements.

Requirements

NMFS requires three reviewers under this contract (i.e. subject to CIE standards for reviewers) to participate in the panel review. The chair, who is in addition to the three reviewers, will be provided by either the New England or Mid-Atlantic Fishery Management Council's Science and Statistical Committee; although the chair will be participating in this review, the chair's participation (i.e. labor and travel) is not covered by this contract.

Each reviewer will write an individual review report in accordance with the PWS, OMB Guidelines, and the TORs below. Modifications to the PWS and TORs cannot be made during the peer review, and any PWS or TORs modifications prior to the peer review shall be approved by the Contracting Officer's Representative (COR) and the CIE contractor. All TORs must be addressed in each reviewer's report. The reviewers shall have working knowledge and recent experience in the use and application of index-based, age-based, and state-space stock assessment models, including familiarity with retrospective patterns, model diagnostics from various population models, and how catch advice is provided from stock assessment models. In addition, knowledge and experience with simulation analyses is helpful.

Tasks for Reviewers

- Review the background materials and reports prior to the review meeting
  - o Two weeks before the peer review, the Assessment Process Lead will electronically disseminate all necessary background information and reports to the CIE reviewers for the peer review.
- Attend and participate in the panel review meeting
  - o The meeting will consist of presentations by NMFS and other scientists, stock assessment authors and others to facilitate the review, to provide any additional information required by the reviewers, and to answer any questions from reviewers
- Conduct an independent peer review in accordance with the requirements specified in this PWS and TORs, in adherence with the required formatting and content guidelines.
- Reviewers are not required to reach a consensus. Individual reviewer perspectives should be provided in their individual reports, and any lack of consensus should be clearly described in the panel's summary report.
- Each reviewer shall assist the Peer Review Panel Chair with contributions to the Peer Review Panel's Summary Report.
- Deliver individual Independent Reviewer Reports to NMFS according to the specified milestone dates.
- This report should explain whether each research track Term of Reference was or was not completed successfully during the peer review meeting, using the criteria specified below in the "Tasks for Peer Review Panel."
- If any existing Biological Reference Points (BRP) or their proxies are considered inappropriate, the Independent Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRPs are the best available at this time.
- During the meeting, additional questions that were not in the Terms of Reference but that are directly related to the assessments and research topics may be raised. Comments on these questions should be included in a separate section at the end of the Independent Report produced by each reviewer.
- The Independent Report can also be used to provide greater detail than the Peer Reviewer Summary Report on specific stock assessment Terms of Reference or on additional questions raised during the meeting.

Tasks for Review panel

- During the peer review meeting, the panel is to determine whether each research track Term of Reference (TOR) was or was not completed successfully. To make this determination, panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. Criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. If alternative assessment models and model assumptions are presented, evaluate their strengths and weaknesses and then recommend which, if any, scientific approach should be adopted. Where possible, the Peer Review Panel chair shall identify or facilitate agreement among the reviewers for each research track TOR.

- If the panel rejects any of the current BRP or BRP proxies (for BMSY and FMSY and MSY), the panel should explain why those particular BRPs or proxies are not suitable, and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing BRPs or BRP proxies are the best available at this time.
- Each reviewer shall complete the tasks in accordance with the PWS and Schedule of Milestones and Deliverables below.

Tasks for Peer Review Panel chair and reviewers combined:
Review the Report of Black Sea Bass Research Track Working Group.

The Peer Review Panel Chair, with the assistance from the reviewers, will write the Peer Reviewer Summary Report. Each reviewer and the chair will discuss whether they hold similar views on each research track Term of Reference and whether their opinions can be summarized into a single conclusion for all or only for some of the Terms of Reference of the peer review meeting. For terms where a similar view can be reached, the Peer Reviewer Summary Report will contain a summary of such opinions.

The chair's objective during this Peer Reviewer Summary Report development process will be to identify or facilitate the finding of an agreement rather than forcing the panel to reach an agreement. Again, the CIE reviewers are not required to reach a consensus. The chair will take the lead in editing and completing this report. The chair may express their opinion on each research track Term of Reference, either as part of the group opinion, or as a separate minority opinion. The Peer Reviewer Summary Report will not be submitted, reviewed, or approved by the Contractor.

Place of Performance
The place of performance shall be remote, via WebEx video conferencing.

Period of Performance
The period of performance shall be from the time of award through January, 2024. Each reviewer's 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 2 weeks of award	Contractor selects and confirms reviewers
Approximately 2 weeks later	Contractor provides the pre-review documents to the reviewers
December 5 - 7, 2023	Panel review meeting

Approximately 2 weeks later	Contractor receives draft reports
Within 2 weeks of receiving draft reports	Contractor submits final reports to the Government

* The Peer Reviewer 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
(2) The reports shall address each TOR as specified
(3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
No travel is necessary, as this meeting is being held remotely.

Restricted or Limited Use of Data
The contractors may be required to sign and adhere to a non-disclosure agreement.

NMFS Project Contact
Michele Traver, NEFSC Assessment Process Lead
Northeast Fisheries Science Center
166 Water Street, Woods Hole, MA 02543
Michele.Traver@noaa.gov

Annex 1. Generic Research Track Terms of Reference

1. Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under impacted TORs.
2. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.
3. Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data.
4. Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and (b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied.

5. Update or redefine status determination criteria (SDC; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs.
6. Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of resulting projections considering the effects of uncertainty and sensitivity to projection assumptions.
7. Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 2 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.
8. Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment.
9. Identify and consider any additional stock specific analyses or investigations that are critical for this assessment and warrant peer review, and develop additional TOR(s)* to address as needed.

Research Track TORs:

General Clarification of Terms that may be
Used in the Research Track Terms of Reference

Guidance to Peer Review Panels about "Number of Models to include in the Peer Reviewer Report":

In general, for any TOR in which one or more models are explored by the Working Group, give a detailed presentation of the "best" model, including inputs, outputs, diagnostics of model adequacy, and sensitivity analyses that evaluate robustness of model results to the assumptions. In less detail, describe other models that were evaluated by the Working Group and explain their strengths, weaknesses and results in relation to the "best" model. If selection of a "best" model is not possible, present alternative models in detail, and summarize the relative utility each model, including a comparison of results. It should be highlighted whether any models represent a minority opinion.

On "Acceptable Biological Catch" (DOC Nat. Stand. Guidelines. Fed. Reg., v. 74, no. 11, 1-16-2009):

Acceptable biological catch (ABC) is a level of a stock or stock complex's annual catch that accounts for the scientific uncertainty in the estimate of Overfishing Limit (OFL) and any other scientific uncertainty…" (p. 3208) [In other words, OFL ≥ ABC.]

ABC for overfished stocks. For overfished stocks and stock complexes, a rebuilding ABC must be set to reflect the annual catch that is consistent with the schedule of fishing mortality rates in the rebuilding plan. (p. 3209)

NMFS expects that in most cases ABC will be reduced from OFL to reduce the probability that overfishing might occur in a year. (p. 3180)

ABC refers to a level of ''catch'' that is ''acceptable'' given the ''biological'' characteristics of the stock or stock complex. As such, Optimal Yield (OY) does not equate with ABC. The specification of OY is required to consider a variety of factors, including social and economic factors, and the protection of marine ecosystems, which are not part of the ABC concept. (p. 3189)

On "Vulnerability" (DOC Natl. Stand. Guidelines. Fed. Reg., v. 74, no. 11, 1-16-2009):

"Vulnerability. A stock's vulnerability is a combination of its productivity, which depends upon its life history characteristics, and its susceptibility to the fishery. Productivity refers to the capacity of the stock to produce Maximum Sustainable Yield (MSY) and to recover if the population is depleted, and susceptibility is the potential for the stock to be impacted by the fishery, which includes direct captures, as well as indirect impacts to the fishery (e.g., loss of habitat quality)." (p. 3205)

Participation among members of a Research Track Working Group:

Anyone participating in peer review meetings that will be running or presenting results from an assessment model is expected to supply the source code, a compiled executable, an input file with the proposed configuration, and a detailed model description in advance of the model meeting. Source code for NOAA Toolbox programs is available on request. These measures allow transparency and a fair evaluation of differences that emerge between models.

Annex 2. Draft Review Meeting Agenda
{Final Meeting agenda to be provided at time of award}

Black Sea Bass Track Assessment Peer Review Meeting

December 5 – 7, 2023

For Details, Please see the following link: https://www.fisheries.noaa.gov/event/black-sea-bass-2023-research-track-peer-review

Annex 3. Individual Independent Peer Reviewer Report Requirements

1. The independent Peer Reviewer report shall be prefaced with an Executive Summary providing a concise summary of whether they accept or reject the work that they reviewed, with an explanation of their decision (strengths, weaknesses of the analyses, etc.).

2. The report must contain a background section, description of the individual reviewers' roles 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. The independent report shall be an independent peer review, and shall not simply repeat the contents of the Peer Reviewer Summary Report.

a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including a concise summary of whether they accept or reject the work that they reviewed, and explain their decisions (strengths, weaknesses of the analyses, etc.), conclusions, and recommendations.

b. Reviewers should discuss their independent views on each TOR even if these were consistent with those of other panelists, but especially where there were divergent views.

c. Reviewers should elaborate on any points raised in the Peer Reviewer Summary Report that they believe might require further clarification.

d. The report may include recommendations on how to improve future assessments.

3. The report shall include the following appendices:

Appendix 1: Bibliography of materials provided for review
Appendix 2: A copy of this Performance Work Statement
Appendix 3: Panel membership or other pertinent information from the panel review meeting.

Annex 4. Peer Reviewer Summary Report Requirements

1. The main body of the report shall consist of an introduction prepared by the Research Track Peer Review Panel chair that will include the background and a review of activities and comments on the appropriateness of the process in reaching the goals of the peer review meeting. Following the introduction, for each assessment /research topic reviewed, the report should address whether or not each Term of Reference of the Research Track Working Group was completed successfully. For each Term of Reference, the Peer Reviewer Summary Report should state why that Term of Reference was or was not completed successfully. It should also include whether they accept or reject the work that they reviewed, with an explanation of their decision (strengths, weaknesses of the analyses, etc.)

To make this determination, the peer review panel chair and reviewers should consider whether or not the work provides a scientifically credible basis for developing fishery management advice. If the reviewers and peer review panel chair do not reach an agreement on a Term of Reference, the report should explain why. It is permissible to express majority as well as minority opinions.

The report may include recommendations on how to improve future assessments.

2. If any existing Biological Reference Points (BRPs) or BRP proxies are considered inappropriate, include recommendations and justification for alternatives. If such alternatives cannot be identified, then indicate that the existing BRPs or BRP proxies are the best available at this time.

3. The report shall also include the bibliography of all materials provided during the peer review meeting, and relevant papers cited in the Peer Reviewer Summary Report, along with a copy of the CIE Performance Work Statement.

The report shall also include as a separate appendix the assessment Terms of Reference used for the peer review meeting, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panel advice.

Appendix 4 - Materials provided or referenced during the Black Sea Bass Research Track Stock Assessment Peer Review meeting

Working papers and presentations were available on a NEFSC website (https://apps-nefsc.fisheries.noaa.gov/saw/sasi.php) by selecting the species and year of assessment.

Working Papers and Background Documentation:

2023_BSB_UNIT_BackLit_2016.SAW62.NEFSC.CRD.17-03.pdf
2023_BSB_UNIT_BackLit_Cope_Hamel_2022.pdf
2023_BSB_UNIT_BackLit_Miller.et.al.2016.PlosONE.pdf
2023_BSB_UNIT_BackLit_Moser.Shepherd.2009.JNWAFS.pdf
2023_BSB_UNIT_ReadMe_Document_V2_12_2_2023.pdf
2023_BSB_UNIT_WP_Beatyetal2023_DiscardMortality.pdf
2023_BSB_UNIT_WP_Brustetal2023_RecCPA.pdf
2023_BSB_UNIT_WP_Curtietal_a_2023_CommercialCatch.pdf
2023_BSB_UNIT_WP_Curtietal_b_2023_SpatialDistribution.pdf
2023_BSB_UNIT_WP_Fayetal2023_StockSynthesisApp.pdf
2023_BSB_UNIT_WP_Hansell_Curti2023_VAST_V2.pdf
2023_BSB_UNIT_WP_Jones_Mercer2023_CommCPUE.pdf
2023_BSB_UNIT_WP_McMahan_Tabenderal2023_FoodHabits.pdf
2023_BSB_UNIT_WP_McNamee2023_NaturalMortality.pdf

2023_BSB_UNIT_WP_Merceretal2023_StakeholderKnowledge.pdf
2023_BSB_UNIT_WP_Miller2023_Multi-WHAM.pdf
2023_BSB_UNIT_WP_Milleretal2023_WHAM.pdf
2023_BSB_UNIT_WP_Paintenetal2023_VentlessTrapSurvey.pdf
2023_BSB_UNIT_WP_Tabenderaetal2023_EcosystemConsiderations.pdf
2023_BSB_UNIT_WP_Truesdell_Curti_a_2023_RecreationalCatch.pdf
2023_BSB_UNIT_WP_Truesdell_Curti_b_2023_AgeLengthKeys.pdf
2023_BSB_UNIT_WP_Truesdell_Curti_c_2023_Surveys.pdf
2023_BSB_UNIT_WP_Verkampetal2023_CFRFResearchFleet.pdf

Presentations

2023_BSB_UNIT_ppt_Intro.pdf
2023_BSB_UNIT_ppt_TOR1_agelengthkeys.pdf
2023_BSB_UNIT_ppt_TOR1_biology.pdf
2023_BSB_UNIT_ppt_TOR1_ecosystemindicators.pdf
2023_BSB_UNIT_ppt_TOR1_naturalmortality.pdf
2023_BSB_UNIT_ppt_TOR1_stakeholderknowledge.pdf
2023_BSB_UNIT_ppt_TOR2_commercialdata.pdf
2023_BSB_UNIT_ppt_TOR2_discardmortality.pdf
2023_BSB_UNIT_ppt_TOR2_recreationaldata.pdf
2023_BSB_UNIT_ppt_TOR2_totalfisherycatch.pdf
2023_BSB_UNIT_ppt_TOR3_VAST.pdf
2023_BSB_UNIT_ppt_TOR3_VAST_revisedindexplots.pdf
2023_BSB_UNIT_ppt_TOR3_commercialCPUE.pdf
2023_BSB_UNIT_ppt_TOR3_recreationalCPA.pdf
2023_BSB_UNIT_ppt_TOR3_surveyindices.pdf
2023_BSB_UNIT_ppt_TOR4-6_stocksynthesis.pdf
2023_BSB_UNIT_ppt_TOR4_MultiWHAM.pdf
2023_BSB_UNIT_ppt_TOR4_WHAMforBSB_V2.pdf
2023_BSB_UNIT_ppt_TOR5-6_WHAMreferencepoints_projections.pdf
2023_BSB_UNIT_ppt_TOR7_researchrecommendations.pdf
2023_BSB_UNIT_ppt_TOR8_backupapproach_V2.pdf

Appendix 5 - Meeting attendees at the Black Sea Bass Research Track Stock Assessment Peer Review meeting

Black Sea Bass Research Track Peer Review Attendance
December 5-7, 2023

GARFO - Greater Atlantic Regional Fisheries Office
MAFMC - Mid Atlantic Fisheries Management Council
NEFSC - Northeast Fisheries Science Center

NJDEP - New Jersey Department of Environmental Protection
NYSDEC - New York State Department of Environmental Conservation
RIDEM - Rhode Island Department of Environmental Management
SMAST - University of Massachusetts School of Marine Science and Technology
~~~~~~

Olaf Jensen - Chair
Sven Kupschus - CIE Panel
JJ Maguire - CIE Panel
Joel Rice - CIE Panel

Larry Alade - NEFSC, Acting Population Dynamics Branch Chief
Michele Traver - NEFSC, Assessment Process Lead

Abby Tyrell - NEFSC
Adelle Molina - Stony Brook University
Aleksandra Bavdaz - SensFish
Alex Dunn - NEFSC
Alex Hansell - NEFSC
Alicia Miller - NEFSC
Andy Jones - NEFSC
Amanda Hart - NEFSC
Anna Mercer - NEFSC
Brandon Muffley - MAFMC Staff
Brian Linton - NEFSC
Charles Adams - NEFSC
Charles Perretti - NEFSC
Chengxue Li - NEFSC
Chris Legault - NEFSC
Dave McElroy - NEFSC
Elizabeth Soranno - Commercial Fisheries Research Foundation
Emily Keiley - GARFO
Emily Liljestrand - NEFSC
Gary Shepherd - former NEFSC employee
Gavin Fay - SMAST
Giovanni Gianesin - NEFSC
Greg DiDomenico - Lund’s Fisheries
Hannah Verkamp - Commercial Fisheries Research Foundation
Jason Boucher - NEFSC
Jason McNamee - RIDEM
Jeffrey Brust - NJDEP
Jeff Kaelin - Lund's Fisheries
Jessica Blaylock - NEFSC
John Maniscalco - NYSDEC
Joseph Beneventine - Recreational fishing industry
Julia Beaty - MAFMC
Kate Wilke - The Nature Conservancy
~~~~~~

Kathy Sosebee - NEFSC
Kiersten Curti - NEFSC
Laura Solinger - NEFSC
Marissa McMahan - Manomet
Mary Kate Munley - NEFSC
Meghan Lapp - SeaFreeze Ltd.
Mike Celestino - NJDEP
Mike Simpkins - NEFSC
Olaf Ormseth - Independent contractor
Remy Gatins - Northeastern University
Sarah Salois - NEFSC
Sam Truesdell - NEFSC
Scott Large - NEFSC
Sefatia Romeo Theken - Deputy Commissioner, MA Department of Fish and Game
Stephanie Owen - NEFSC
Steve Cadrin - SMAST
Steve Cannizzo - New York Recreational & For-Hire Fishing Alliance
Steve Doctor - Maryland Fisheries Service Department of Natural Resources
Steve Witthuhn - Rec. Captain, Top Hook Charters
Tara Trinko Lake - NEFSC
Tim Miller - NEFSC
Toni Chute - NEFSC
Tony Wood - NEFSC
Tracey Bauer - North Carolina DMF