Report on the

SEDAR 64 Yellowtail Snapper Assessment Review

Prepared for:
The Center for Independent Experts

Dr Kevin Stokes
Stokes.net.nz Ltd
59 Jubilee Rd
Khandallah
Wellington 6035
New Zealand
Ph: +64 (04) 973 7305
E-mail: kevin@stokes.net.nz

EXECUTIVE SUMMARY

The report must be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

The South East Data, Assessment, and Review (SEDAR) 64 for yellowtail snapper took place at the Courtyard by Marriott, St Petersburg, Fl from 24th to 26th February 2020. The meeting was well organized and administered and was open and transparent. The Review Workshop (RW) Panel was able to reach consensus on all issues and the individual inputs to the RW Panel Report was provided according to the schedule for compilation by the Chair. This report is an individual report that largely reflects the Panel Report although with some additions and minor departures.

The Southeastern US yellowtail snapper stock assessment was carefully considered and conducted and generally well reported. SEDAR 64 documentation is clear, and materials provided on the SEDAR website are comprehensive. The Analytical Team is clearly well on top of the many issues and provided excellent materials and presentations to the RW, as well as responding well to requests during the RW Meeting.

The methods used (SS3 and ASAP) are standard in the US and SEDAR region. The base case model reasonably fits to the available data and sensitivity and robustness tests indicate it is a sufficient basis to inform status determination and other management advice.

The model has been well explored given data and time availability but is hampered by lack of data beyond 2017. The assessment base case appears to be reliable as a basis for informing decision making. And represents the best scientific information available. Against traditional single species standards, the fishery appears neither to be overfished nor experiencing overfishing.

BACKGROUND

*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.*

Yellowtail Snapper

Yellowtail snapper (*Ocyurus chrysurus*), a *Lutjanid*, is a reef-associated species distributed from Brazil in the south as far as Massachusetts in the north, and through the Caribbean and Gulf of Mexico. The species is caught throughout its range but predominantly in Brazil, Mexico and the United States, where the main concentration is in the Florida Keys and the southeast coast of Florida where suitable reef habitat is available. Genetic studies (as reported to the Review Workshop, Data Inputs) suggest stock differentiation between South Florida and the Caribbean (both east and west), as well as between other areas (e.g. Belize and Brazil). There seems to be a reasonable basis to treat the US fish as comprising a single stock for assessment purposes, with a focus on the Florida Keys and southeast Florida, from which the large majority of catches (circa 96%) are taken and throughout which area yellowtail snapper have access to reef habitat.

The species generally is reef-associated, with juveniles found in seagrass beds and mangroves but migrating to coral reefs. Breeding is on the reef edges. As noted in the Review Meeting in discussion on the USA stock, the migration of juveniles from seagrass beds or mangroves to coral reefs has been widely observed in the Caribbean but very little in Florida. The Reef Fish Visual Census (RVC) counts juveniles (i.e., <190mm) and adults (i.e., >190mm) and it was noted in discussion that the juvenile RVC sees fish at sizes that would in other areas be expected in seagrass beds or mangroves rather than on reefs. The general life cycle for the US stock is therefore inferred from but not fully consistent with observations for other regions. Once settled on reefs, adults have restricted movement, of the order of 1km.

The maximum age of yellowtail snapper in the defined stock assessment area (Florida) is 20 years, with the large majority of fish (circa 90%) being aged 2 to 6 years old, circa 60% aged 2 and 3, and few fish aged over 16. Within Florida, older fish tend to be distributed in the Keys, with a higher proportion of young (2 years old) on the southeast coast. Northwards, however, off the Carolinas, fish have been aged up to 28 years old. From discussion, it seems these fish are not regarded as part of the breeding population, being "strays" that have been transported away from areas with suitable reef habitat. Similarly, fish on the southeast coast may be "overflow" from the Keys which is the center of gravity of the population. The area north of Florida is not included in the stock assessment. How to treat the northern fish (with higher observed maximum age) in estimating growth and natural mortality is an area of debate. Variations in age and growth suggest considerable plasticity in life histories depending on environmental conditions. Based on histological data, 50% maturity is at age 1.7, reaching 100% by age 4.

Commercial and recreational fisheries for yellowtail snapper in the assessment area have a long history with data available from multiple sources for an extended period. While commercial data are available from 1962, recreational catch information (from multiple sources) is only available from 1980 onwards. The starting year for the stock assessment, however, is 1992, chosen because prior to then data were deemed to be of questionable quality, available age composition data for the different fleets became more consistent, and some indices of abundance first became available (SEDAR 64 Assessment Report, 3.2.6.9). Commercial catches were of the order of 500mt per year in the early 1980s and increased to near 1,000mt by 1993/94. Catches since then slowly decreased to near 500mt in 2007 but have since trended upwards to circa 1,000mt per year in recent years. Retained recreational catches are dominated by private sources (as estimated using MRIP) and are estimated to have approached or exceeded 4,000mt in some years (1981, 1982, 1984, 1989, 1991) but have varied greatly. In recent years (2013-2017) catches have been in the rage 1,300-1,900mt per year.

Commercial discards from 2013 to 2017 are estimated to be in the range 23.5 to 59.2mt. Recreational discard estimates are measured in thousands of fish, rather than in mt; from 2013 to 2017, estimates vary from about 1,500 thousand (i.e. 1.5 million) to 5,000 (i.e. 5 million). The median length of discarded fish is circa 200mm, well below the minimum TL of 305mm, with weight circa 0.15kg, implying recreational discard tonnages in the range 225-750mt. Fisheries generally are prosecuted using chumming and lines.

While the stock is treated as a single US stock for assessment purposes, with the concentration of catch in the Florida Keys, management is guided by both SAFMC and GMFMC FMPs, as well as by States, notably Florida. The GMFMC has a minimum size limit and a recreational bag limit in place which are consistent with SAFMC measures (i.e. 305mm minimum TL limit and 10 fish bag limit). The GMFMC does not currently define management reference points but these are under development. The SAFMC in 1999 defined an MSY-proxy of 30%SPR and an OY-proxy of40%SPR, and in 2014 modified the MSST as 75%SSBmsy. The GMFMC sets a combined commercial and recreation ACL while the SAFMC sets separate ACLs and has additionally applied commercial closures. The current GMFMC combined ACL is 901,125lbs (409mt) while the SAFMC commercial ACL is 1,596,510lbs (724mt) and the recreational ACL is 1,440,990lbs (654mt). Stock assessments in 2003 (SEDAR 3) and 2012 (SEDAR 27A) both found that the population of yellowtail snapper was not overfished and not experiencing overfishing.

The SEDAR 64 review was originally scheduled for 2019 but was delayed for administrative reasons. Data for the benchmark stock assessment, however, have not been updated and though taking place in 2020, the stock assessment only considers data up to and including 2017.

REVIEW PROCESS

*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.*

*a. **Reviewers must describe in their own words the review activities completed during the panel review meeting**, including a brief summary of findings, of the science, conclusions, and recommendations.*
d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

The South East Data, Assessment, and Review (SEDAR) 64 for yellowtail snapper took place at the Courtyard by Marriott, St Petersburg, Fl from 24th to 26th February 2020. Terms of Reference (ToR) for the stock reviews are given in Appendix 2, Annex 2.

Participants in the review are listed in Appendix 3. The SEDAR Panel comprised a GSMFC SSC appointed Chair (Powers), and reviewer (Lorenzen), two SAFMC reviewers (Schueller and Sharov), and three CIE reviewers (Maguire, Stephenson and Stokes). Notification of the meeting and dissemination of papers followed closely the schedule laid out in the CIE Statement of Work (see Appendix 2). Materials were provided using the SEDAR website (see Appendix 1). Public comments were able to be made using the SEDAR website and opportunity was provided during the review meeting; none was made. Overall, administration of the review was sound.

I have noted elsewhere in reviews that it would be beneficial to have presentations posted in advance rather than during the meeting. It can be difficult for analytical teams to have everything prepared too far in advance, but even late draft presentations would be helpful to reviewers and pre-meeting posting on the SEDAR

website is encouraged. It would also be very useful for presentations always to have numbered pages/slides to assist discussions. For SEDAR 64, presentations were all posted during the first session.

The SEDAR Panel was tasked with providing inputs to the Panel report as outlined by the Chair during the opening session; specific contributions were agreed during the meeting with all Panel members contributing. The Chair assumed overall responsibility for the Panel report with draft sections, in bullet form, due from members by the start of the final day of the meeting in order to allow feedback during the final session.

The meeting followed the general outline of the draft agenda (Appendix 2) but with sufficient flexibility to allow necessary responses from the STAT. In my opinion, the meeting was well run and Panelists, Analytical Team members, and the public were afforded proper opportunities for input and comment. I am not aware of any problems with notification of the meetings and interpret from the presence of stakeholder representatives, and lack of complaint, that notification was appropriate. All participants were able to participate throughout the meeting and opportunity was explicitly and regularly given by the chair for input. Many participants other than Panelists and Analytical Team members contributed usefully to discussion and I believe that all were provided appropriate opportunity for involvement both during the Panel meeting and during extra-mural discussions. Enough time was provided to look in reasonable detail at data inputs and modeling decisions and to contemplate assessment outputs. I am confident that the SEDAR 64 resulted in informed and reasonable conclusions.

REVIEWER'S ROLE IN THE REVIEW ACTIVITIES

The role of the reviewer is set out in the CIE Statement of Work, Attachment A, attached here in Appendix 2, Attachment A. CIE reviewers are tasked with producing an independent report to the CIE. As part of the stock specific ToR (ToR 9), the reviewers are additionally tasked with contributing to a Peer Review Summary Report for the review.

In addition to conducting necessary pre-review preparations, including the review of background material and papers provided in advance by the SEDAR project and a conference call, I (Stokes) participated in all discussions and contributed bullet points on ToR 3 to the draft Summary Report, provided to the Chair overnight on 25th February following agreement with other panelists. With other Panelists I discussed all ToR responses prior to the Panel providing feedback in the closing session. Along with other Panel members I responded to the Chair to enable production of the report according to schedule.

SUMMARY OF FINDINGS BY STOCK

*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.*

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 summary report that they believe might require further clarification.
e. The report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The report shall represent the peer review of each TOR, and shall not simply repeat the contents of the summary report.

YELLOWTAIL SNAPPER

ToR 1 *Evaluate the data used in the assessment, including discussion of the strengths and weaknesses of data sources and decisions, and consider the following:*

a) *Are data decisions made by the DW and AW sound and robust?*
b) *Are data uncertainties acknowledged, reported, and within normal or expected levels?*
c) *Are data applied properly within the assessment model?*
d) *Are input data series reliable and sufficient to support the assessment approach and findings?*

Maturity and fecundity data are carried over from previous assessments and were not discussed in detail. A key area of discussion during the Review Meeting was appropriate use of data to determine natural mortality, with some concern that an observed maximum age of 28 in Georgia waters implied a much lower overall natural mortality than the observed maximum age of 20 in Florida waters. The DW and AW clearly considered these issues in some detail. The DW made a recommendation to apply the Hoenig and Lorenzen methods using the entire observed data set to determine the maximum age but the AW discussed and decided to use Florida data only but also to explore consistency of natural mortality estimation with estimates derived from growth modelling (following Jensen). Using available data, the AW estimated a maximum age of 14 (using the Hoenig method) would be consistent – less than the 20 year maximum observed in Florida waters and much less than the 28 year maximum observed in Georgia. The AW adopted natural mortality estimates using Florida-only data. In my view, the issue of stock structure and dynamics is moot but the assessment as structured to inform management is appropriate, including the use of data from Florida waters; the use of Florida-only maximum age data is consistent with this.

Other key data decisions made by the DW and AW relate to the start year of the assessment and to discard mortality rates. The rationale for starting in 1992 relates entirely to data availability, consistency and quality and is appropriate, especially in combination with sensitivity testing to the earlier (1980) start date. It should be noted that the final data year is 2017; though this is unfortunate, the reasons for lack of 2018 data availability are beyond the scope of DW and AW decision-making.

Discard rates as live releases from the recreational fishery (measured by MRIP) are very high (see *Background* section, above), especially of small, including juvenile, fish. Commercial discard rates are very small in comparison. In assessment terms, the MRIP discards are the only potential concern. During the Review Meeting, the discard mortality rate was discussed, and it was noted that previously a rate of 30% had been used based on other snapper species, but not yellowtail snapper. At SEDAR 27A a rate of 10-11% had been used and this is the basis for the currently assumed 10% for all fleets though with sensitivity testing using 15% for commercial and higher (20 and 30%) for recreational fleets. The total amount of discarding of small fish by the recreational fleet suggests some incongruity between the minimum TL and other management measures - but this is not an assessment issue. Understanding the discard mortality when the discarding rate is high, however, is an assessment issue because it potentially has a large effect on SPR calculations and ACL determination. The DW and AW have made a pragmatic decision to maintain the assumption of 10% discard mortality with judicious use of sensitivity testing to ensure robustness of advice given uncertainties.

The assessment is unusual (fortunate) in having a range of fishery-dependent and -independent surveys which appear to be reasonably consistent in trend. Standard statistical approaches to estimating associated CVs have been applied. All surveys could be the subject of specific review, and the MRIP, of course, has been much considered elsewhere. The Reef Fish Visual Census (RVC) are well described in working papers included in the SEDAR 64 papers; they use point survey methods which are well-established though it is unclear how accurate identification of fish to 1cm bins is feasible or how much error this might create in defining ages, especially given the highly smeared distribution of lengths at age apparent in the age-length key (see below).

During the Review Meeting, the issues of age validation and reading history were briefly considered. As noted during discussion, the issue of ageing precision was addressed by the DW, with working papers made available. The clear indication is of easy-to-read otoliths and high precision age readings by experienced age readers, readings are therefore verified though only possibly validated depending on the interpretation of readily distinguishable annuli as annual markers. The ALK reflects the high degree of smearing of lengths for each age, reflecting highly plastic growth.

In

FL (cm)	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	20
10	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.40	0.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	0.33	0.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	0.00	0.86	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	0.12	0.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16	0.32	0.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17	0.27	0.73	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18	0.04	0.78	0.15	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
19	0.00	0.70	0.27	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
20	0.00	0.63	0.28	0.07	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
21	0.00	0.51	0.35	0.07	0.03	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	0.00	0.31	0.42	0.14	0.08	0.02	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
23	0.00	0.20	0.46	0.19	0.07	0.04	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
24	0.00	0.10	0.49	0.25	0.09	0.05	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25	0.00	0.08	0.48	0.26	0.11	0.04	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
26	0.00	0.05	0.47	0.27	0.12	0.05	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
27	0.00	0.03	0.46	0.28	0.13	0.06	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
28	0.00	0.01	0.40	0.29	0.15	0.07	0.03	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
29	0.00	0.01	0.35	0.32	0.17	0.08	0.04	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
30	0.00	0.00	0.28	0.31	0.19	0.10	0.05	0.03	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
31	0.00	0.00	0.23	0.35	0.18	0.11	0.07	0.03	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
32	0.00	0.00	0.17	0.37	0.21	0.11	0.07	0.03	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
33	0.00	0.00	0.13	0.38	0.23	0.12	0.07	0.03	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
34	0.00	0.00	0.08	0.41	0.23	0.13	0.07	0.04	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
35	0.00	0.00	0.05	0.39	0.27	0.13	0.07	0.04	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
36	0.00	0.00	0.03	0.30	0.30	0.16	0.10	0.05	0.02	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
37	0.00	0.00	0.02	0.25	0.32	0.19	0.10	0.05	0.03	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
38	0.00	0.00	0.02	0.20	0.31	0.17	0.13	0.08	0.04	0.03	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
39	0.00	0.00	0.02	0.17	0.31	0.21	0.11	0.07	0.06	0.03	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
40	0.00	0.00	0.01	0.07	0.24	0.22	0.16	0.15	0.06	0.04	0.03	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
41	0.00	0.00	0.01	0.05	0.21	0.24	0.16	0.12	0.10	0.03	0.03	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
42	0.00	0.00	0.01	0.03	0.16	0.24	0.20	0.14	0.07	0.08	0.05	0.02	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00
43	0.00	0.00	0.01	0.02	0.07	0.18	0.20	0.15	0.15	0.07	0.07	0.03	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
44	0.00	0.00	0.01	0.02	0.05	0.16	0.16	0.17	0.16	0.09	0.03	0.07	0.02	0.01	0.02	0.01	0.00	0.00	0.00	0.00
45	0.00	0.00	0.01	0.01	0.09	0.09	0.16	0.11	0.13	0.14	0.09	0.07	0.03	0.01	0.03	0.01	0.01	0.00	0.00	0.00
46	0.00	0.00	0.00	0.01	0.03	0.08	0.16	0.15	0.13	0.11	0.08	0.06	0.08	0.04	0.04	0.01	0.01	0.00	0.00	0.00
47	0.00	0.00	0.00	0.00	0.03	0.07	0.19	0.17	0.12	0.13	0.08	0.06	0.02	0.04	0.03	0.03	0.01	0.00	0.01	0.00
48	0.00	0.00	0.00	0.00	0.01	0.06	0.06	0.09	0.14	0.09	0.08	0.14	0.13	0.09	0.03	0.03	0.03	0.01	0.01	0.00
49	0.00	0.00	0.00	0.03	0.03	0.06	0.00	0.15	0.09	0.12	0.06	0.12	0.06	0.15	0.06	0.00	0.00	0.03	0.03	0.00
50	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.09	0.11	0.09	0.13	0.16	0.11	0.09	0.09	0.00	0.09	0.00	0.00	0.00
51	0.00	0.00	0.00	0.00	0.06	0.00	0.13	0.13	0.19	0.00	0.25	0.00	0.00	0.13	0.06	0.06	0.00	0.00	0.00	0.00
52	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.09	0.27	0.00	0.00	0.09	0.00	0.36	0.00	0.09	0.00	0.00	0.00	0.00
53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.17	0.50	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00
54	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.00	0.00	0.00	0.00	0.00	0.00	0.33	0.00	0.00	0.00
55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.50	0.00	0.00	0.00	0.00	0.00
57	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33	0.33	0.00	0.00	0.33	0.00	0.00	0.00	0.00	0.00
58	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

response to the specific ToR a-d:

a) Are data decisions made by the DW and AW sound and robust?

Yes, decisions by both the DW and AW are sound and robust.

b) Are data uncertainties acknowledged, reported, and within normal or expected levels?

Yes, data uncertainties are within normal or expected levels given the life histories of reef fish species (e.g., protracted spawning, smearing of year-class signals) which lead to uncertainties in growth estimations and inferences on natural mortality.

c) Are data applied properly within the assessment model?

This ToR is confounded somewhat with ToR 2. Yes, the data are used and appropriately considered in the SS3 stock assessment model. The approach taken is to present all data to the model with no forced weighting and to apply standard tuning approaches to balance process and observation errors for all data sets. Given all indices show consistent trends and do not conflict with composition data, this is not problematic in practice. However, consideration by the DW of each index and its relative merits would be helpful in future in case external weighting decisions are required. The larger issue perhaps is that the model is fitting to both length

and age with growth in the model; the highly smeared ALK means that age is poorly determined and it may be that a length and growth only model would be more appropriate.

d) Are input data series reliable and sufficient to support the assessment approach and findings?

Yes, the data are sufficient to support the assessment approaches, status determinations, and subsequent ACL determinations.

ToR 2 *Evaluate and discuss the strengths and weaknesses of the methods used to assess the stock, taking into account the available data, and considering the following:*

a) Are methods scientifically sound and robust?

b) Are assessment models configured properly and consistent with standard practices?

c) Are the methods appropriate for the available data?

The main stock assessment method used Stock Synthesis 3 (SS3; Methot and Wetzel, 2013). SS3 is able to integrate several sources of information and can be applied even where data are limited. It is extensively used in the US for a wide range of stocks. It is a statistical method that uses the available data in a raw form as appropriate in a single analysis. This allows for consistency in assumptions and permits the uncertainty associated with all data sources within any model run to be propagated to final model outputs (Maunder and Punt, 2013).

As a modelling framework, SS3 is sound and robust. However, SS3 requires skilled usage by analysts and specific applications need to be judged on merit and in relation to the assessment requirements. The DW and AW carefully considered and developed data inputs for the stock assessment and the analytical team, in limited time, has systematically and thoughtfully conducted the assessment, with the AW Report providing somewhat restrained but sufficient details to provide confidence that the assessment is sound and robust.

Integrated assessments such as SS3 are able to include a variety of datasets simultaneously and provide diagnostics which may reveal conflicts between datasets and influences on estimated parameters. Conducting a stock assessment is an iterative process involving weighting of datasets and assumptions with regard to diagnostics, and exploration of alternative models/hypotheses. Use of SS3 allows not just for the consistent propagation of uncertainty within a model run but also allows exploration of between model uncertainty and the robustness of status determination and management advice. The analytical team tested for robustness by conducting sensitivity tests (varying assumptions and data use), jack-knife analyses (systematic removal of abundance indices), parametric bootstrapping (revealing possible tension between length composition and conditional age-at-length data), jitter analyses (to test base case model convergence), retrospective analysis (to test for systematic bias in current and recent year estimates), and application of MCMC (to test if the base case approximated medians differ from the fully integrated estimates.

The assessment team also used ASAP, another widely used stock assessment modelling framework. ASAP is an age-based assessment framework and cannot accommodate length data; it therefore does not estimate growth and cannot fit selectivity as a function of length. It is less flexible than SS3 but can be used as a check of the consistency of results with SS3 and is an important step in the assessment for continuity reasons, having been used as the main assessment tool in the previous assessment (SEDAR 27A). It is less sensitive than SS3 to small changes in data, parameters, and constraints. ASAP produced results broadly similar to those from SS3 for fishing mortality, biomass, and recruitment, providing further confidence in the robustness of the SS3 application.

Taken as a whole, the methods are sound and robust, are configured properly and consistent with standard practice, and are appropriate for the available data. The assessment is new, using SS3 for the first time for

Yellowtail snapper and while the base case is a credible and robust basis for status determination and management advice, there is potential for improvement in the future. Apart form lack of up-to-date data (not a model issue) and data issues (ToR 1) the model could benefit from further exploration of the use of length data, length-based selectivity and growth estimation, etc. Notwithstanding the wide range of length at age, given the life history of Yellowtail snapper and ontogenetic movement, age-based selectivity might be a sufficient assumption for assessment purposes. SS3 can use length data but it doesn't have to.

ToR 3 *Evaluate the assessment findings and consider the following:*

a) Are population estimates (model output – e.g. abundance, exploitation, biomass) reliable, consistent with input data and population biological characteristics, and useful to support status inferences?

As noted in the Background section, Yellowtail snapper is distributed across a wide geographic range though with distinct stocks in different areas. The USA stock is treated as a single, closed stock found along the entire USA GoM coastline, around Florida and along the Atlantic coastline where it has been found as far north as Massachusetts. However, the fishery is largely centered around southern and southeastern Florida, especially in the Florida Keys. Yellowtail snapper in USA waters are fast growing and long-lived but with highly plastic life history depending on environmental conditions - growth rates and longevity are highly variable across regions even within Florida and there is large variation in size at age. This plastic life history creates problems in defining appropriate parameters in stock assessment, or appropriate data to present to model.

Stock assessment models account for approximately 96% of the catch of yellowtail snapper from the GoM and SE Atlantic, with data presented to the model only from the Florida Keys and southeast Florida. The base case stock assessment model includes landings and discards split by fleet (commercial, Head boat, and recreational), fishery-dependent and -independent indices, and age and length compositions.

The base case stock assessment is implemented using SS3; all data are presented to the model with no forced weighting and the model is tuned using standard procedures and subjected to a variety of sensitivity and robustness tests. The assessment provides reliable estimates of abundance, biomass, and exploitation, consistent with input data and population biological parameters which can be used to infer status and inform management based on proxy reference points.

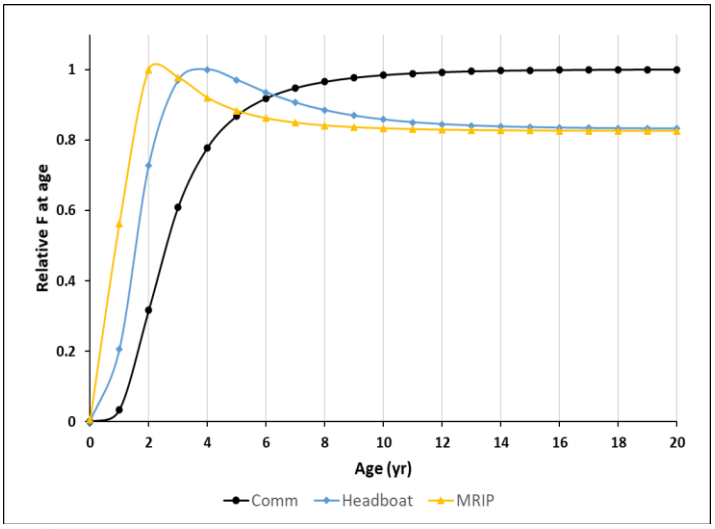
b) Is the stock overfished? What information helps you reach this conclusion?

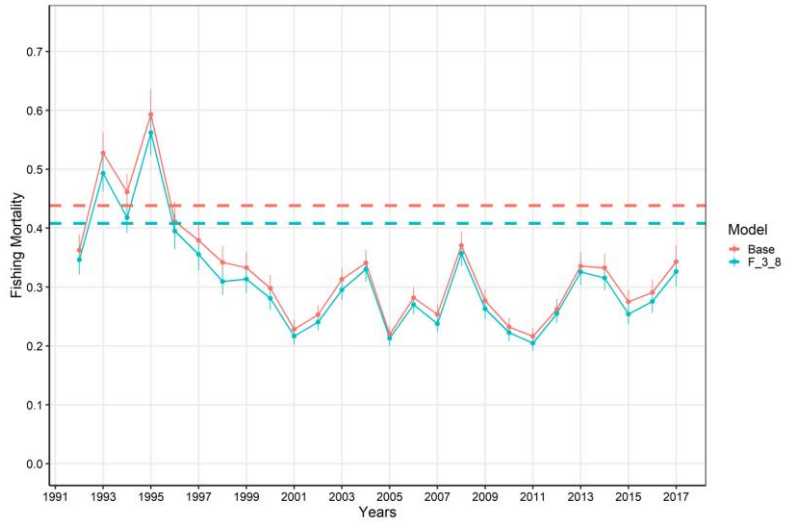
The stock is not overfished. The base case assessment estimates $SSB_{F30\%SPR}$ as 1,904 mt and $SSB_{current}$ as 3,223 mt. Sensitivity tests considering alternative selectivity, natural mortality, and steepness indicate the status determination (not overfished) is robust. For values of steepness between 0.7 and 0.9 (the base case estimate is 0.81) there is no effective variation in status determination.

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

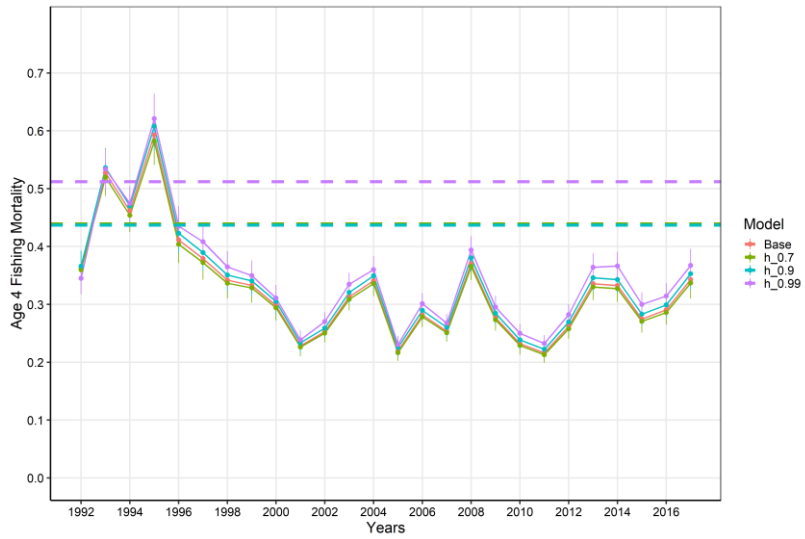
The stock is not undergoing overfishing. The base case assessment estimates $F_{30\%SPR}$ as 0.44 and $F_{current}$ as 0.29. Sensitivity tests considering alternative selectivity, natural mortality, and steepness indicate the status determination (not undergoing overfishing) is robust.

It is notable that the exploitation pattern varies considerably by fishery (see below for base case selectivity by fishery). During the Review Meeting, estimates of F by fishery was considered (peak F for MRIP is at age 2, for recreational headboats at age 4, and reaches maximum for commercial fisheries at age 8+. The figures below shows fishing mortality by year estimated in the base case SS3 assessment for age 4 (base; as used in developing status determination) and averaged for ages 3-8. The use of age 4 fishing mortality as a basis for determining overfishing status appears to be robust.



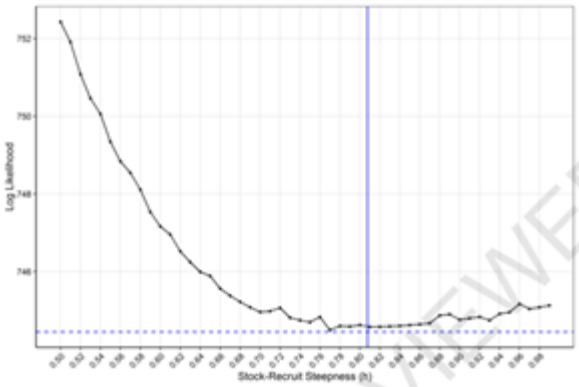


Sensitivity to fixed stock-recruitment steepness is low (see below) with no effect on status determination (or OY).

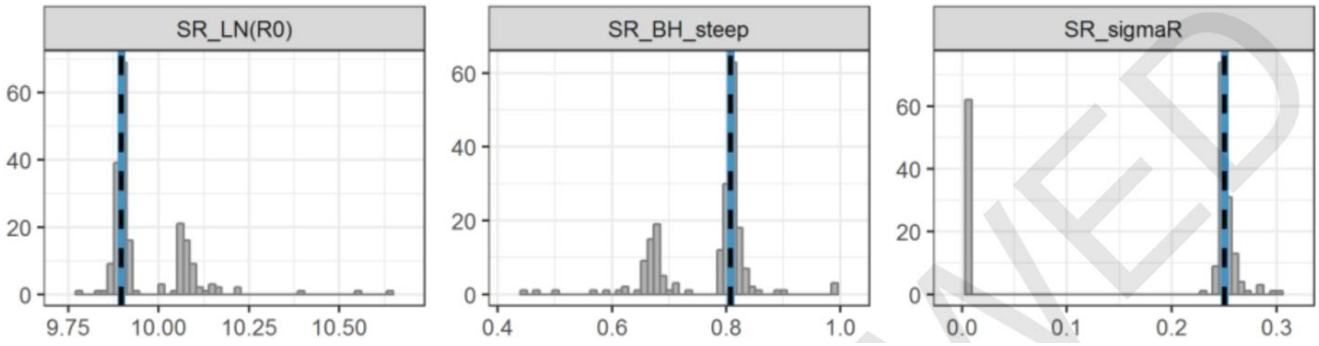


d) Is there an informative stock recruitment relationship? Is the stock recruitment curve reliable and useful for evaluation of productivity and future stock conditions?

A Beverton-Holt stock recruitment model was implemented within the SS3 framework and an estimate of steepness was obtained (0.81 in the base case). However, the likelihood profile (see below), the jitter analysis (see below), the fact that the stock has not been reduced to levels where strong density dependence might occur, and the fact that steepness is confounded with natural mortality (and selectivity) all suggest that the stock-recruitment relationship is not informative for defining future productivity. The review Panel therefore recommended that 30% SPR continue to be used as an MSY proxy.



The jitter analysis produced bi-modal distributions for R0, steepness, and sigma R with the left-hand side mode suggesting constant recruitment, i.e. sigma R = 0. This suggests that the stock-recruitment parameters are not well defined.



e) Are the quantitative estimates of the status determination criteria for this stock reliable? If not, are there other indicators that may be used to inform managers about stock trends and conditions?

See also ToR 5 on uncertainty.

All indices and composition data (from 1992-2017) have been put into the model with no forced weighting to give preference to a particular data set. The tuning of process and observation errors given the data presented to the base case model has followed standard (Francis, 2011) approaches with some testing of alternative composition weighting options and jack-knifing of abundance indices. Assessment sensitivity to maximum age (affecting growth function determination), and other inputs/assumptions (see above) was tested.

Resulting quantitative, management-related estimates are reliable to the extent that the stock definition is appropriate, and data from the Florida Keys and SE Florida are representative of the USA stock as a whole.

ToR 4 *Evaluate the stock projections, including discussing strengths and weaknesses, and consider the following:*

a) Are the methods consistent with accepted practices and available data?

ToR for the AW are specific for the case when the stock is neither overfished nor experiencing overfishing (i.e., project at F=Fcurrent, F=Fmsy, F=0.75Fmsy, and at equilibrium yield. The analytical team made projections from the base case assessment for all four scenarios as required, using standard projection methods tested and included in SS3 and applied using common assumptions on recruitment – fixed S-R parameters as estimated in the base case assessment in projection years except the first (2018) for which the most recent three-year average was used. This is all consistent with accepted practices and available data.

b) Are the methods appropriate for the assessment model and outputs?

Yes, see ToR 4a.

c) Are the results informative and robust, and useful to support inferences of probable future conditions?

Projections are shown for the base case only and recruitments drawn from the stock-recruit relationship span only a limited range. While recruitment fluctuations are not large compared to many species the values drawn for projections are likely not a good representation of future fluctuations. Nevertheless, they provide useful guidance on likely trends in spawning biomass under different exploitation levels. Fcurrent (0.295) is very close to 0.75F30%SPR and the projections are therefore similar, suggesting a fairly stable but possibly slowly decreasing stock. At the higher F levels projected, especially at equilibrium yield, the projections suggest a steeper decline.

With projections starting in 2018, it would be useful perhaps to consider any information available on recruitment to test recruitment assumptions in the projections. The only obvious source of information would be the RVC indices, especially for juveniles. However, it is notable in the DW working documents that the indices show little fluctuation, implying recruitment signals are primarily drawn for composition data. It is of note also that RVC indices generally are trending upwards, inconsistent with recruitment estimates for recent years from the base case assessment.

d) Are key uncertainties acknowledged, discussed, and reflected in the projection results?

The projections are wholly deterministic. Given time constraints on the analytical team (as noted e.g. at the AW Report section 3.3.10) this is understandable; the team should be commended for ensuring a full assessment and report with all components was completed in the time available. Though deterministic, uncertainties are acknowledged and discussed but not explicitly taken into account in the projections other than by providing confidence intervals on projected quantities.

ToR 5 *Consider how uncertainties in the assessment, and their potential consequences, are addressed.*

- *Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods*
- *Ensure that the implications of uncertainty in technical conclusions are clearly stated*

The assessment adopts standard approaches to model fitting/tuning given estimated or assumed errors in data inputs. Estimates of derived parameter uncertainty also follow standard methods coded within SS3. In addition, the AW has reported on estimates of uncertainty based on bootstrapping and MCMC. It is not clear that the bootstrapping analysis can be used but the MCMC is informative.

At sections 3.3.13 and 3.5, the AW Report reasonably describes uncertainty in the assessment and in estimates of reference points. Consideration of uncertainty for the base case run in a broad sense, including robustness,

utilizes a range of likelihood profiling, jitter, sensitivity, retrospective, bootstrap, and jack-knife analyses. These methods are all appropriate for exploration of uncertainties related to data inputs, model assumptions, and observation error, and model fitting (along with standard diagnostics). The base case model was also run using MCMC.

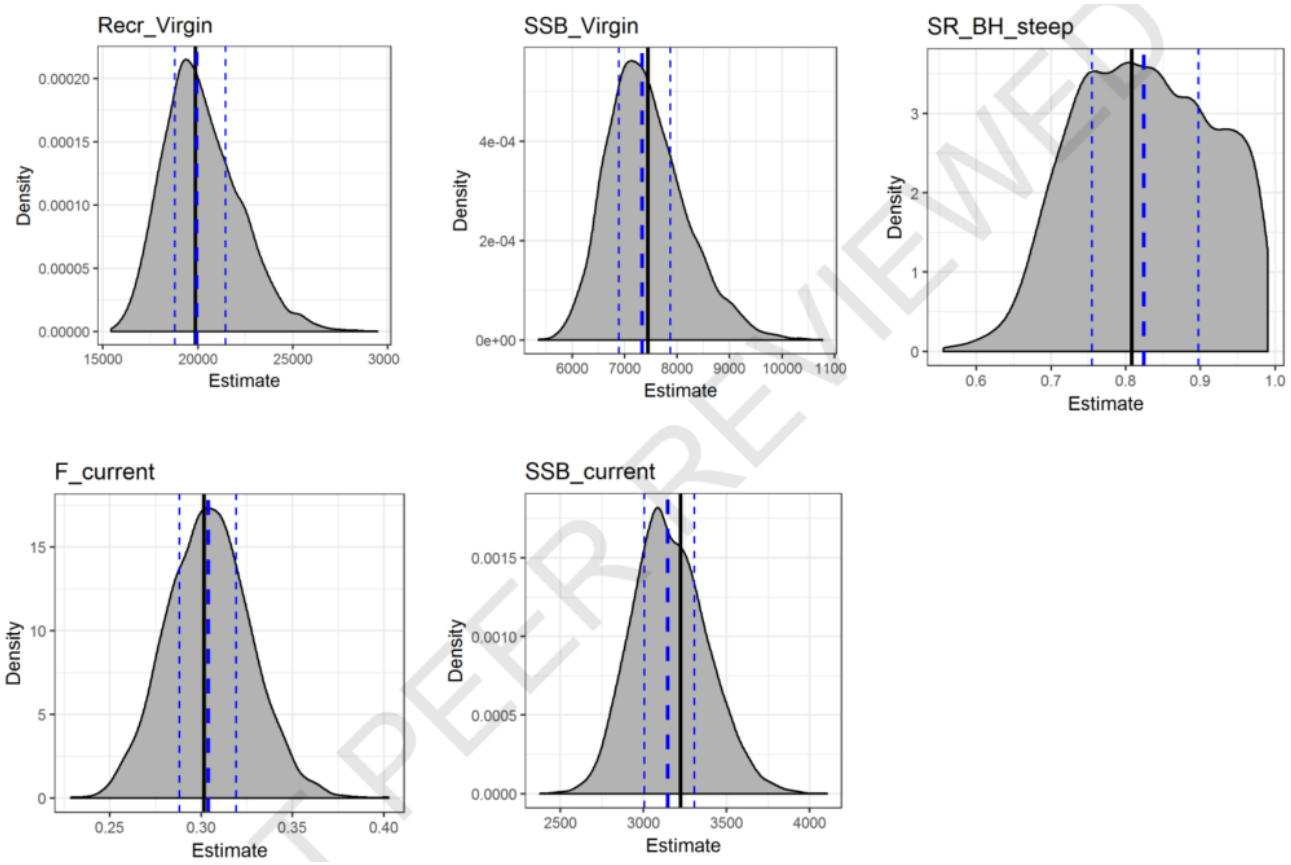
No analytical consideration was given to issues related to the assumption of a single, homogenous stock cf a known wider distribution, variation in age and size structure between Florida Keys and southeast Florida.

Jitter analyses essentially test for the robustness of the assessment; does it actually converge? The analyses provide confidence that it is a global solution. Likelihood profiling examines parameter identifiability and can help guide whether to estimate a parameter or fit to fixed values. Retrospective analyses provide confidence that there is not systematic under- or over-estimation caused by model structure or assumptions or the underlying data. The jack-knife analyses test for differences in information between abundance indices. All of these are generally carried out without any re-tuning of the model and typically re-use tuning settings from the base case. To do otherwise would be too time-consuming though for jack-knife testing it would be appropriate. The same applies to sensitivity testing but depending on what is being tested, re-tuning might be a fairer test of sensitivity. Notwithstanding, all analyses generally provide confidence in the base case as robust, removing uncertainty as to its applicability.

The description of parametric bootstrapping in the AW Report is a little hard to follow and was not discussed at length during the RW Meeting. The AW Report states that the bootstrap analyses demonstrated sensitivity to effective sample sizes of conditional age-at-length data, possibly due to conflict between length composition and the conditional age-at-length data. Given the reported technical difficulties with the bootstrapping, it is in fact hard to draw conclusions about uncertainty associated with the base case estimates of uncertainty.

MCMC analyses were carried out by the analytical team. Standard statistics were used to judge convergence as well as visual inspection. The MCMC provides a reasonable estimation of uncertainty associated with the base case model estimates and derived parameters (though not of structural uncertainty). The plot below shows the posterior distribution of selected parameters and derived quantities of management relevance. In the plots, the grey distributions are derived from MCMC and the black line shows the base case median estimate. The blue lines show the MCMC interquartile range.

The RW can accept or reject a stock assessment but does not lead to a new assessment or base case. In passing, it is worth noting (see ToR 3) that there is no information on stock-recruit steepness (h). This can be seen readily from the likelihood profile (ToR 3) as well as simple sensitivity tests at different h values (as presented during the RW Meeting). It can also be seen in the wide posterior from the MCMC.



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

- *Clearly denote research and monitoring that could improve the reliability of, and information provided by, future assessments*

The DW and AW made a number of research recommendations related to age and growth, length composition data, commercial discards, headboat landings and discards, and fishery independent data. No stock assessment specific recommendations were made.

Age and growth

A range of specific recommendations are made by the DW and AW relating to sampling outside of Florida and to different treatments of data. These are all relevant to improving understanding of stock structure and dynamics and are worthwhile.

Length composition data

This recommendation by the DW and AW is the length counterpart to the preceding recommendation and seem highly relevant.

Commercial discards

The recommendation notes there is high uncertainty about commercial discards and suggest an exploration of collection and analysis methods to improve precision. If the current estimates are an indication, then commercial discards are at a low level and comprise a negligible fraction of all catches. While more precise

estimates are always welcome, higher priority should be given to recreational landings and discards which are much higher and potentially impactful on estimates of status and potential yield.

Headboat landings and discards
There is considerable uncertainty associated with headboat landings and discards and as for commercial discards, better information would be welcome. However, as for commercial discards, the apparent scale of catches and discards compared to MRP is very small.

Fishery independent data
The DW and AW make recommendations for age sampling from fishery independent sources and for seagrass sampling of <100mm Yellowtail snapper. As limited rationales are given for these recommendations it is hard to comment. It is unclear how the first recommendation differs from the recommendation on age and growth and a clearer recommendation would be helpful. The second recommendation is aimed at recruitment estimation, but it is unclear if this would in fact be useful given current data include information on 0-year-old fish.

Additional comments
As noted during the RW Meeting, more analysis and synthesizing of existing biological information pertaining to spatial stock structure and dynamics is needed. The stock occurs at the edge of the species' range and cursory analysis suggests that movement, migration, and life history plasticity are important factors in stock dynamics. In addition to further sampling and exploration on age and growth, which could provide insight to stock structure and dynamics, further investigation of stock structure, spawning areas, larval transport and juvenile/adult movement using methods such as otolith microchemistry or stable isotopes would be useful.

While the DW and AW suggest more/better sampling/estimation of commercial and headboat fisheries, no recommendation is made with respect to MRIP estimated discards or of discard mortality. Given the apparent scale and annual variability in MRIP discards, and potential to impact on status and potential yield estimates, this seems surprising.

As noted at ToR 1, ageing appears to be verified and consistent, but validation is not guaranteed. Age validation studies to test whether growth checks are laid down consistently throughout the area of distribution and sampling (which includes tropical and temperate habitats) and reflect annual increments could be useful in this respect.

- *Provide recommendations on possible ways to improve the SEDAR process*

See comments under Review Process.

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

The criteria listed in the ToR are extensive but not explicitly defined, making it difficult to respond definitively. In a recent review of science processes for the IPHC (Stokes, 2019) an alternative set of defined criteria was adopted, based on a general framework developed and used for New Zealand fisheries quality assurance, which itself drew on practices elsewhere. The New Zealand Research and Science Information Standard (MPI, 2011) defines key principles for science information quality, noting that it relates to i) relevance, ii) integrity, iii) objectivity and iv) reliability. The standard further notes that the primary, internationally accepted mechanism for evaluating the quality of research and science information is v) peer review and, as such, peer review is

both a principle and a mechanism. The explicit definitions of the key principles and mechanism assist in making commenting more precise.

This is not to suggest that the exact framework should be adopted by SEDAR, but it would be helpful to have explicit definitions. It is notable, for example, that the Review Panel comment on *Relevance* refers just to the stock assessment technique while the definition adopted for the IPHC review speaks to the relevance of the science to fisheries management questions being addressed. Similarly, the Review Panel comment on *Objectivity* relates to recognition of shortcomings while the definition adopted for the IPHC review speaks much more widely to important aspects of the process. Some clarification of definitions would assist Review Panels respond to the SEDAR ToR in a consistent and helpful way.

Notwithstanding these comments:

Appropriate: The Assessment framework used, SS3, is an appropriate tool to provide relevant advice to management processes. The use of alternative models and exploration of sensitivities to data and assumptions is appropriate.

Relevance: The stock assessment is highly relevant to management needs, providing estimates of stock status and potential catches, and identifying uncertainties. However, there is some concern that the stock definition and population structure modelled may not be a good representation of the stock. There is further concern that for administrative reasons, only data up to and including 2017 were available for analysis, potentially undermining the relevance of the assessment for advisory purposes.

Inclusiveness: The SEDAR process (see Peer Review, below) is open through multiple steps with apparent good opportunity for all interested parties to be included.

Objectivity: The stock assessment analyses appear to be unbiased with no assumptions made on the basis of risk aversion and no explicit favoring of specific data sets. The data used similarly appear to be based on objective processes and are not biased during analyses. Data shortcomings are recognized and considered during analyses.

Transparency: As noted by the review Panel, *The assessment model was subjected to various adjustments and the differences in the outputs were candidly explained*. However, while the stock assessment report offers a clear explanation of what was done, it arguably could provide more insight as to why some analytical decisions were made. Given the timing of the review and time available for preparation of the report this should not be construed as criticism of the analytical team. More generally, in terms of SEDAR process, communication is good and inclusiveness is provided for; the process is transparent.

Timeliness. In terms of timeliness for the review process, data inputs and stock assessment (SS3) input files were supplied about two weeks before the meeting. The Assessment model outputs were supplied two days before. It would be ideal if outputs and presentations could be provided slightly earlier.

[NB re Verification and Validation. Validation is concerned with determining whether a conceptual model such as a stock assessment is an accurate representation of the system under study, while verification is the process whereby a program or process is determined to perform as intended (e.g., the software used to fit a stock assessment model).]

Verification: Data available for stock assessment were analyzed using SS3 (Methot and Wetzel, 2013) and also ASAP, the model used in the previous SEDAR 27a assessment. The software packages are widely used and have the subject of extensive testing/verification.

Validation: The validity of the stock assessment depends on i) if the underlying stock structure assumed (a simple dynamic pool model) sufficiently represents the spatial dynamics and ontogeny of yellowtail snapper; ii) whether life history assumptions are appropriate; and ultimately iii) management requirements of the stock assessment and whether the models used can be used to provide appropriate advice. There are some outstanding questions (see ToR 8) about stock definition *per se*, appropriateness of a simple dynamic pool model, and use of Florida-only data. However, stock assessment models are not used to provide accurate representations of reality so much as to inform management in a consistent and useful way. The stock assessment using SS3 utilizes available data and makes standard assumptions and is designed to provide necessary input to the management process. Robustness of resulting advice is tested through i) use of the ASAP model; ii) sensitivity and other tests.

Peer Review: The SEDAR process includes multiple steps where collegiate and stakeholder review is possible, culminating in the Review Workshop which includes multiple CIE and SSC appointed reviewers. All processes are open and transparent. All data, assumptions, analyses and conclusions are considered.

ToR 8 *Provide suggestions on key improvements in data or modeling approaches that should be considered when scheduling the next assessment.*

The key issues are data availability and timeliness. The new assessment (SEDAR 64) in 2020 uses data only up to 2017. As data processing catches up and data for Yellowtail snapper become available on a more normal schedule, a new assessment might be scheduled. Timing should be in relation to the last year of data (2017) in the assessment, not to the year of SEDAR 64 (2020).

While the next assessment could be based on the same stock structure and dynamic assumptions, data to explore alternative structures would be helpful. Research recommendations (see ToR 6) could be helpful in this regard. If projects collecting data relevant to stock structure are underway, this could be relevant to scheduling of the next stock assessment.

Similarly, if age validation work is undertaken, then any changes from current interpretation could indicate the need for sooner stock assessment (if status determination were potentially impacted) or later (if data processing requirements resulted).

ToR 9 *Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference. Develop a list of tasks to be completed following the workshop. Complete and submit the Peer Review Summary Report in accordance with the project guidelines.*

Completed. See sections on *Review Process* and *Reviewer's Role in Review Activities*.

GENERAL CONCLUSIONS AND RECOMMENDATIONS

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 Southeastern US yellowtail snapper stock assessment was carefully considered and conducted and generally well reported. SEDAR 64 documentation is clear, and materials provided on the SEDAR website are comprehensive. The Analytical Team is clearly well on top of the many issues and provided excellent materials and presentations to the RW, as well as responding well to requests during the RW Meeting.

The methods used (SS3 and ASAP) are standard in the US and SEDAR region. The base case model reasonably fits to the available data and sensitivity and robustness tests indicate it is a sufficient basis to inform status determination and other management advice.

The model assumes a single US stock with simple dynamics and uses data only from Florida. Evidence suggests a discrete US stock but it is known that fish are distributed west across the US Gulf of Mexico and east to international waters and north along the US eastern seaboard as far north as Massachusetts. Catches of Yellowtail snapper are predominantly in Florida waters (96%) with the greatest proportion from the Florida Keys. Even within Florida (Keys cf southeast) life histories are plastic. Further north, beyond the core area, life histories appear different, with maximum age considerably higher. This indicates potential structural uncertainty associated with the stock assessment.

Nevertheless, given 96% of Yellowtail catches come from Florida waters, use of Florida-only data and treatment as a single stock is likely sufficient to provide robust management advice.

For the Yellowtail snapper stock, landings data for the commercial fleet are good and while discards are not precisely estimated they are at a low level. Recreational landings are not well estimated for headboats but MRIP landings are sound. Discard estimation for recreational fisheries is imprecise and the scale of discards associated with MRIP fisheries is large enough potentially to impact on future status determination. Age and other biological data are restricted to Florida waters, especially in the core fishery areas around the Keys. Age and other biological data from other areas would be useful to help identify variations that might be accounted for in future assessments. The major uncertainty currently relates to stock structure and dynamics and the appropriateness of the Florida-only single stock. Ageing is verified but not validated though there is no specific reason to doubt interpretation. However, with very large variation in length by age, possibly also spatially, validating ageing and sampling across a wider area could help in future modelling.

The model has been well explored given data and time availability but is hampered by lack of data beyond 2017. The assessment base case appears to be reliable as a basis for informing decision making. Against traditional single species standards, the fishery appears neither to be overfished nor experiencing overfishing.

APPENDIX 1

BIBLIOGRAPHY

Prior to the Workshop, all materials were provided *via* the SEDAR website: http://sedarweb.org/sedar-64.

During the workshop multiple presentations were given, and additional materials were provided on request, including further background documents and presentations as well as responses to Panel requests. All files were made available on the SEDAR website in a timely fashion using an open Wi-Fi connection throughout the meeting. All materials remain publicly available on the website.

Additional references are:

- Maunder, M.N., and Punt, A. E. (2013). A review of integrated analysis in fisheries stock assessment, Fisheries Research , 2013, vol. 142, pp 61-74
- Methot, R.D. and C. Wetzel (2013) Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research 142:86-99. https://www.researchgate.net/journal/0165-7836_Fisheries_Research
- MPI (2011) Ministry for Primary Industries Research and science information standard for New Zealand fisheries. 31 pp. https://www.mpi.govt.nz/dmsdocument/3692-research-and-science-information-standard-for-new-zealand-fisheries
- Stokes (2019) https://iphc.int/uploads/pdf/priphc/independent-peer-review-0909.pdf

APPENDIX 2

Performance Work Statement (PWS)
National Oceanic and Atmospheric Administration (NOAA)
National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program
External Independent Peer Review

SEDAR 64 Yellowtail Snapper Assessment 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. (http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf). Further information on the CIE program may be obtained from www.ciereviews.org.

Scope

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

SEDAR 64 will be a compilation of data, an assessment of the stock, and CIE assessment review conducted for S.E. U.S. yellowtail snapper. The review workshop provides an independent peer review of SEDAR stock assessments. The term review is applied broadly, as the review panel may request additional analyses, error corrections and sensitivity runs of the assessment models provided by the assessment panel. The review panel is ultimately responsible for ensuring that the best possible assessment is provided through the SEDAR process. The stock assessed through SEDAR 64 is within the jurisdiction of the Gulf of Mexico and South Atlantic Fisheries Management Councils.

The specified format and contents of the individual peer review reports are found in **Annex 1**. The Terms of Reference (TORs) of the peer review are listed in **Annex 2**. Lastly, the tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements
NMFS requires three (3) reviewers to conduct an impartial and independent peer review in accordance with the Performance Work Statement (PWS), OMB guidelines, and the TORs below. The reviewers shall have a working knowledge in stock assessment, statistics, fisheries science, and marine biology sufficient to complete the primary task of providing peer-review advice in compliance with the workshop Terms of Reference fisheries stock assessment.

Tasks for Reviewers
1) Two weeks before the peer review, the NMFS Project Contacts will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contacts will consult with the contractor on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the PWS scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

Working papers, reference documents, and the Data Workshop and Assessment Process Reports will be available on the SEDAR website: http://sedarweb.org/sedar-64

2) Attend and participate in the panel review meeting. The meeting will consist of presentations by NOAA and other scientists, stock assessment authors and others to facilitate the review, to answer any questions from the reviewers, and to provide any additional information required by the reviewers.

3) After the review meeting, reviewers shall conduct an independent peer review report in accordance with the requirements specified in this PWS, OMB guidelines, and TORs, in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus.

4) Each reviewer should assist the Chair of the meeting with contributions to the summary report.

5) Deliver their reports to the Government according to the specified milestones dates.

Foreign National Security Clearance
When reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for reviewers who are non-US citizens. For this reason, the reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project

Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/ and http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration- system.html. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance
The place of performance shall be at the contractor's facilities, and in St. Petersburg, FL.

Period of Performance
The period of performance shall be from the time of award through April 2020. Each CIE 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 two weeks of award	Contractor selects and confirms reviewers
2 weeks prior to the panel review	Contractor provides the pre-review documents to the reviewers
February 25-27, 2020	Panel review meeting
Approximately 3 weeks later	Contractor receives draft reports
Within 2 weeks of receiving draft reports	Contractor submits final reports to the Government

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; and (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract. Travel is not to exceed $12,000.

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

Project Contacts:

Larry Massey – NMFS Project Contact
150 Du Rhu Drive, Mobile, AL 36608
(386) 561-7080
larry.massey@noaa.gov

Julie Neer - SEDAR Coordinator
SEDAR Coordinator
Science and Statistics Program
South Atlantic Fishery Management Council
4055 Faber Place Drive, Suite 201
North Charleston, SC 29405
Julie.Neer@safmc.net

Annex 1: Peer Review Report Requirements

1. The report must be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

2. The 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.

   a. Reviewers must describe in their own words the review activities completed during the panel review meeting, including a brief summary of findings, of the science, conclusions, and recommendations.

   b. 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 summary report that they believe might require further clarification.

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The report shall represent the peer review of each TOR, and shall not simply repeat the contents of the summary report.

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 2: Terms of Reference for the Peer Review

SEDAR 64 Yellowtail Snapper Assessment Review

1. Evaluate the data used in the assessment, including discussion of the strengths and weaknesses of data sources and decisions, and consider the following:
   a) Are data decisions made by the DW and AW sound and robust?
   b) Are data uncertainties acknowledged, reported, and within normal or expected levels?
   c) Are data applied properly within the assessment model?
   d) Are input data series reliable and sufficient to support the assessment approach and findings?
2. Evaluate and discuss the strengths and weaknesses of the methods used to assess the stock, taking into account the available data, and considering the following:
   a) Are methods scientifically sound and robust?
   b) Are assessment models configured properly and consistent with standard practices?
   c) Are the methods appropriate for the available data?
3. Evaluate the assessment findings and consider the following:
   a) Are population estimates (model output – e.g. abundance, exploitation, biomass) reliable, consistent with input data and population biological characteristics, and useful to support status inferences?
   b) Is the stock overfished? What information helps you reach this conclusion?
   c) Is the stock undergoing overfishing? What information helps you reach this conclusion?
   d) Is there an informative stock recruitment relationship? Is the stock recruitment curve reliable and useful for evaluation of productivity and future stock conditions?
   e) Are the quantitative estimates of the status determination criteria for this stock reliable? If not, are there other indicators that may be used to inform managers about stock trends and conditions?
4. Evaluate the stock projections, including discussing strengths and weaknesses, and consider the following:
   a) Are the methods consistent with accepted practices and available data?
   b) Are the methods appropriate for the assessment model and outputs?
   c) Are the results informative and robust, and useful to support inferences of probable future conditions?
   d) Are key uncertainties acknowledged, discussed, and reflected in the projection results?
5. Consider how uncertainties in the assessment, and their potential consequences, are addressed.
   - Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods

- Ensure that the implications of uncertainty in technical conclusions are clearly stated

6. Consider the research recommendations provided by the Data and Assessment workshops and make any additional recommendations or prioritizations warranted.
   - Clearly denote research and monitoring that could improve the reliability of, and information provided by, future assessments
   - Provide recommendations on possible ways to improve the SEDAR process
7. Consider whether the stock assessment constitutes the best scientific information available using the following criteria as appropriate: relevance, inclusiveness, objectivity, transparency, timeliness, verification, validation, and peer review of fishery management information.
8. Provide suggestions on key improvements in data or modeling approaches that should be considered when scheduling the next assessment.
9. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference. Develop a list of tasks to be completed following the workshop. Complete and submit the Peer Review Summary Report in accordance with the project guidelines.

Annex 3: Agenda - SEDAR 64 Yellowtail Snapper Assessment Review

February 25-27, 2020
Saint Petersburg, Florida

Tuesday:

9:00 a.m.	**Introductions and Opening Remarks** *- Agenda Review, TOR, Task Assignments*	**Coordinator**
9:30 a.m. – 11:30 a.m.	**Assessment Presentations** **-** *Assessment Data & Methods* *- Identify additional analyses, sensitivities, corrections*	**Analytic Team**
11:30 a.m. – 1:00 p.m.	**Lunch Break**	
1:00 p.m. – 6:00 p.m.	**Assessment Presentations (continued)** **-** *Assessment Data & Methods* *- Identify additional analyses, sensitivities, corrections*	**Analytic Team**
6:00 p.m. – 6:30 p.m.	**Public comment**	**Chair**

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

Wednesday:

8:00 a.m. – 11:30 a.m.	**Panel Discussion** **-** *Assessment Data & Methods* *- Identify additional analyses, sensitivities, corrections*	**Chair**
11:30 a.m. – 1:00 p.m.	**Lunch Break**	
1:00 p.m. – 6:00 p.m.	**Panel Discussion/Panel Work Session** *- Continue deliberations* *- Review additional analyses* *- Recommendations and comments*	**Chair**

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

Thursday:

8:00 a.m. – 11:30 a.m.	**Panel Discussion** *- Final sensitivities reviewed.* *- Projections reviewed.*	**Chair** **Chair**
11:30 a.m. – 1:00 p.m.	**Lunch Break**	
1:00 p.m. – 5:30 p.m.	**Panel Discussion or Work Session** - *Review Reports*	**Chair**
5:30 p.m. – 6:00 p.m.	**Public comment**	**Chair**
6:00 p.m.	**ADJOURN**	

Thursday Goals: Complete assessment work and discussions, final results available. Draft Reports reviewed.

APPENDIX 3
PERTINENT INFORMATION FROM THE REVIEW

LIST OF PARTICIPANTS

Workshop Panel

Joseph Powers (Chair) ..... GMFMC SSC
Kai Lorenzen ..... GMFMC SSC
J.J. Maguire ..... CIE
Amy Schueller ..... SAFMC SSC
Alexei Sharov ..... SAFMC SSC
Peter Stephenson ..... CIE
Kevin Stokes ..... CIE

Analytic Team

Shanae Allen, Co-Lead Analyst ..... FWRI, St. Petersburg
Chris Swanson, Co-Lead Analyst ..... FWRI, St. Petersburg

Appointed Observers

Ed Walker ..... GMFMC AP

Attendees

Dustin Addis ..... FL FWC, St. Petersburg
Luiz Barbieri ..... FL FWC, St. Petersburg
Martha Guyas ..... FL FWC, GMFMC Rep, Tallahassee
Jessica McCawley ..... FL FWC, SAFMC Rep, Tallahassee
Bob Muller ..... FWRI, St. Petersburg
Joseph Munyanderaro ..... FWRI, St. Petersburg
Joe O'hop ..... FWRI, St. Petersburg

Staff

Julie Neer ..... SEDAR
Mike Errigo ..... SAFMC Staff
Ryan Rindone ..... GMFMC Staff
Charlotte Schiaffo ..... GMFMC Staff