Independent Peer Review Report on the SouthEast Data and Stock Assessment Review (SEDAR) 64 Florida Yellowtail Snapper

St. Petersburg, Florida
24 February – 26 February 2020

Prepared for Center of Independent Experts

By

Peter Stephenson
18 b La Grange St, Innaloo,
WA 6018, Australia
Email peterstephenson48@gmail.com

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 Review Meeting for SouthEast Data, Assessment, and Review (SEDAR64) met on February 24- 26, 2020 to examine the stock assessment of yellowtail snapper (*Ocyurus chryurus*).

The Assessment Team presented a comprehensive report on the model structure, data inputs, and detailed model outputs. Requests for further analyses were dealt with promptly and the Assessment Team demonstrated they were very familiar with the chosen assessment software, Stock Synthesis 3. Participants at the meeting provided valuable comments on the previous assessment and the details of the fishery.

The Review Panel discussed many aspects of the yellowtail snapper assessment, including the extent of the assessment area (Florida Keys and South East Florida), data inputs, model structure, values allocated to parameters, stock status, and future projections. On the basis of these deliberations, the Review Panel found that the base model assessment was sound and it was endorsed. The Review Panel members provided input into the Panel Report which was compiled by the Chair and reconciled between the members.

I addressed all the Terms of Reference and concluded that the assessment was scientifically sound and able to provide good evidence that the stock was not overfished nor undergoing overfishing. A summary of my recommendations for the yellowtail snapper assessment are:

- The assessment software chosen for this assessment, SS3, provided reasonably stable parameter estimates and reliable measures of the derived quantities needed to determine stock status. The use of SS3 should continue, but the ASAP model should also be used to provide comparative results.
- The source of juvenile fish should be determined.
- The degree of mixing of adult fish between the assessment area and NE Florida/North Carolina and also Gulf of Mexico should be investigated.
- Some otolith collection resources should be re-directed from the Florida Keys to areas where the sample numbers are low, for example NE Florida.
- The ageing process should be examined to see if the large age range in each length class is realistic.
- The selectivity functions should be examined with the view to using simpler formulations in cases where this would make sense.
- If catches increase outside the current yellowtail snapper assessment area, there may be good reason to have a multi-area assessment, depending on the degree of mixing of adult yellowtail snapper.
- The iterative calculation of effective sample size should be routinely performed when the model structure or input values are altered.
- If the model area needs to be extended to encompass NE Florida/North Carolina, then it would be appropriate to revise the allocated natural mortality.
- The maximum age, used to determine natural mortality, should not necessarily be revised when outlier old fish are discovered.
- The steepness parameter of the stock-recruitment relationship should be fixed or tightly bounded.
- Consideration should be given to reducing the influence of the age composition data in the assessment model.
- If there is *a priori* knowledge of the reliability of each abundance index, then this could be used to weight them in the assessment.
- When the data in the SS3 input files are not identical to that in the Data Workshop reports, an explanation should be either documented or mentioned at the Review Workshop.

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.

The review workshop for the SouthEast Data, Assessment and Review (SEDAR 64) on yellowtail snapper (*Ocyurus chryurus*) was held at the Courtyard Hotel in Saint Petersburg, Florida from 24 February to 26 February, 2020. The meeting venue, in the hotel where the panellists were staying, was very convenient. The meeting was efficiently organized and all the technical equipment worked flawlessly.

The members of the Review Panel (Appendix 3) had diverse backgrounds and all made major contributions to the review process with their comments and questions. There were many other attendees at the meeting (Appendix 3) who made a significant contribution through their insightful comments.

The Data Workshop Reports, background materials, and the Assessment Report (Appendix 1), together with the inputs for the SS3 assessment model, were available three weeks before the meeting, giving the Review Panel ample time to become familiar with the materials required for the review.

At the Review Meeting, the Assessment Team confidently presented a detailed description of the model outputs and responded quickly to all questions from the review panel, demonstrating their competence in using the software package Stock Synthesis 3.30.13.

This report represents the independent review of Peter Stephenson in accordance with the guidelines stated in the Performance Work Statement shown in Appendix 2.

Review Process

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 panellists, 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.

Prior to the workshop I read all the Data Workshop (DW) reports supplied for this review, including the documents containing information not used in the assessment. I carefully reviewed the input files for the Stock Synthesis 3 base model (starter, data, control, and forecast files) and checked how these matched the information in the DW files. I also made myself familiar with the model structure but did not run the SS3 model.

I travelled to St. Petersburg, Florida, and actively participated in discussion on the details of the assessment results and also the additional analyses requested by the Review Panel. Through these discussions, I was satisfied that I could make a determination on the reliability of the assessment to determine the stock status of yellowtail snapper.

I have addressed each of the Terms of Reference listed below with my comments.

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?

The Assessment group decided to restrict the assessment to the area of Florida Keys and South East Florida. This pragmatic decision was based on two things

- 96% of the catch comes from these two areas.
- Limited information (e.g., tagging data in the Keys) indicates there is little movement of yellowtail snapper from the Florida Keys to the more northerly areas of NE Florida and North Carolina, nor into the areas of NW Florida and the Gulf of Mexico.

There was considerable discussion on the issue and I believe the use of a one area model is appropriate with the current distribution of landings.

There is a need to determine the larval source of yellowtail snapper outside the Florida Keys. Although there are suitable seagrass areas and mangroves in NE Florida and North Carolina, juveniles have not been observed. It is also important to know the degree of mixing of adult yellowtail snapper between the Florida Keys, where the juveniles are known to occur, and NW Florida, NE Florida, and further north into North Carolina.

Landings. There are estimated landings for the Commercial Fleet (1950-2017), Headboat Fleet (1981-2017), and the Recreational Fleet (Marine Recreational Information Program: MRIP; 1981-2017). At the Assessment Review Meeting, it was decided that the current assessment base model should use data from 1992 to 2017. The CVs were only known for the MRIP fleet but acceptable choices were made to allocate standard errors to the landings of the other two fleets.

At the request of the Review Panel, a model run from 1981 to 2017 was compared to the base model (1992-2017). The resulting stock status, indicated by spawning stock biomass and fishing mortality, was not substantially different. Unfortunately, the 2018 landings were not available for the current assessment. Figure 1 is an excellent overview of the data sources available for the assessment.

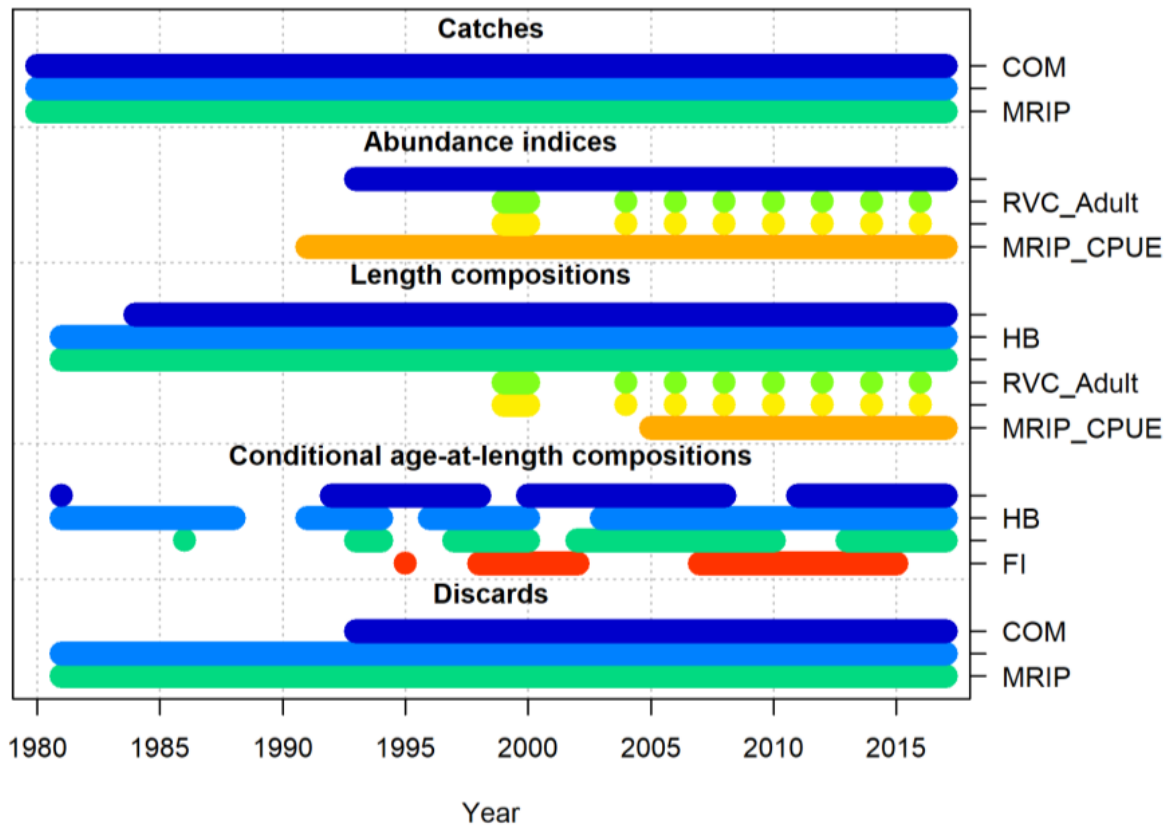


Figure 1. Data streams available by year for the SEDAR 64 Southeastern U.S. Yellowtail Snapper stock assessment with the starting year selected as 1992.

Discards. Discard rates were available for the Commercial Fleet, Headboat Fleet, and the MRIP Fleet for varying and limited time periods. CVs were available for Commercial, and MRIP but not the Headboat fleet. Sensible decisions were made for the allocation of standard errors.

Indices of Abundance. There are numerous abundance indices described in the document list. The methods used to determine the indices are well described and the statistical methods used are sound. At a Population Abundance Workshop (PAW) the indices were evaluated and four were chosen for the assessment: Commercial Index (pounds/hook hour), MRIP (Marine Recreational Information Program) Index (number of fish/trip), and the Reef Video Census-Adult Index, and the Reef Video Census-Juvenile Index. At the review workshop, it would have been good if there had been some discussion of the reasons for choosing these four indices and excluding the others.

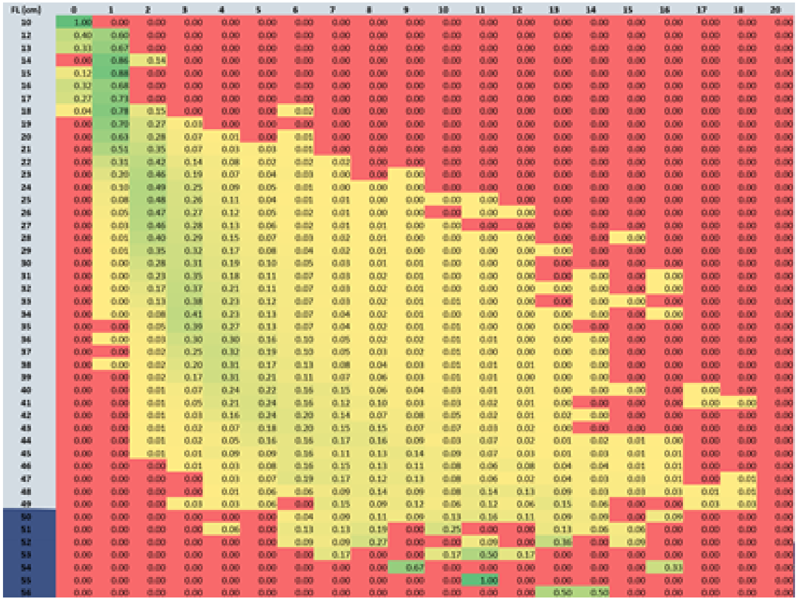
I noticed that the indices reported in the Data Workshop Report for the RVC did not match those in SS3 input file, data.dat, whereas those for the MRIP did match. This was because bias correction $\sqrt{log_e(1+\text{CV}^2)}$ was required in the former but not the latter. The understanding of the data inputs is vital to understanding the model results and I feel a brief explanation at the Review Meeting or in the RVC documentation would have been useful.

I believe the Review Panel would have benefited from some discussion on the assessments team's a priori notion of the relative reliability of the four different indices. If the members of the PAW believe the CVs of the indices reflect the appropriate weighting in the assessment model, then the current setting of "extra standard error" equal to zero is appropriate. I believe there is a case for considering weighting of the indices in the SS3 control file in future assessments.

In summary, the graphs of the selected indices appear consistent, having similar trends over time and I believe they support the assessment approach. In addition, the jack-knife analysis supports this belief.

Age composition data. The assessment model was constructed with ages from 0 to 20 with ages over 11 being allocated to a 12+ group. The parameters of the von Bertalanffy growth curve, estimated in the base model assessment, resulted in L_∞ being quite low and k being high. This generated considerable discussion with the Review Panel who requested a model run with data extended beyond Florida Keys-SE Florida, The results of the model run, showed a larger value of L_∞ and lower k. This was expected as the requested model run included older fish collected outside the Keys-SE Florida area.

The Review Panel also requested to see the Yellowtail Snapper age-length key (Figure 2). This showed a large range of ages for each length class, especially for lengths above the legal size.



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	0.01	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	0.01	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	0.05	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	0.05	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

Figure 2. Yellowtail snapper age-length-key

For me, this raised a number of questions.

1. *Is there sexual dimorphism in growth*? The Assessment Team assured me the answer was no.

2. *Are the yellowtail snapper otoliths difficult to read*? Typically, many species found in low latitude waters have otoliths which are difficult to read. Several members of the Florida laboratory assured me that this species is not difficult to read. In a quality control ageing exercise, where 1197 otoliths were viewed by six ageing staff, it was clear that some are much more consistent than others. There were some cases where the ring count for individual fish varied by three annuli from the agreed number. It should be investigated whether improved consistency is possible. Statisticians have explained that counting annuli using "Machine Learning" could improve consistency. I personally have my doubts about that as reading annuli is too much of an art.

3. *Are the ageing protocols satisfactory*? The documentation indicates that accepted ageing protocols are being observed. There are six staff doing fish ageing and each otolith has the annuli counted twice, generally by two different people, followed by reconciliation. There is a reference collection of 100 otoliths which is used to ensure consistency between readers and to ensure counts do not drift. I am unsure of the current practice, but some otoliths in the Reference Collection need to be changed periodically because after a few of viewings, the readers will remember the number of annuli they allocated to a particular fish, especially for difficult to read otoliths.

4. *Are the annuli annual growth rings*? Marginal Increment Analysis was used verify that the counts represent annual growth rings. I have no problem accepting that the annuli are annual growth rings.

5. *Are the age composition data useful for the stock assessment*? The Review Panel discussed the idea that the length-composition are a far more reliable data set than the age-composition, and suggested the assessment should possibly be conducted with only the length-composition data in conjunction with an age-length key. The large variation in ages for each length class still gives you a problem. I believe that age-

composition should be retained but considerable effort should go into investigation into the wide range of age-at-length.

Natural Mortality. The assignment of a value for natural mortality is difficult in many stock assessments. The assessment team explored values of natural mortality derived from growth parameters and found the results to be unreasonable. As mentioned previously, the growth parameters for this species are highly uncertain, as indicated by the age-length-key, and there is large variation in growth parameters with different data sets.

The value selected will specify the resilience of the modelled stock to fishing and can greatly affect conclusions drawn about stock status. The Assessment Team chose the method of Hoenig (1983). With the maximum age of 20 years used in the assessment area of Florida Keys-SE Florida, the value of natural mortality is 0.223. This method for assigning a value for natural mortality is easy to understand and has been used in a multitude of assessments and I agree with the choice in this assessment. In addition, the assumption is made that natural mortality is related to fish length and M at age is determined using the estimated von Bertalanffy parameters and the method of Lorenzen (2005). This method is commonly used in Florida assessments and I consider this to be acceptable.

There was discussion by the panel whether the maximum age of 20 is appropriate, given that there are a number of fish of age 23 and 24, and indeed one of age 28. The Review Panel requested that the model be run with a value of M=0.16 corresponding to a maximum age of 28. As expected, the age-4 fishing mortality was higher and the spawning stock biomass was lower. As the threshold and target levels also change, the stock status was not substantially different.

The older fish (age 23 and 24) did not occur the assessment area of the fishery and I believe that with the current model structure, the choice of maximum age is appropriate. In the future, if there are significant landings outside Florida Keys-SE Florida, and the model structure is changed to a multi-area model, the maximum age and the corresponding assignment of a value for M should be re-visited. In addition, I believe that maximum age should not be changed whenever a really old fish is discovered. From the 45,000 yellowfin snapper aged, a maximum age of 23 or 24 is appropriate and I believe it does not need to be revised when an outlier (e.g., 28 years) is discovered.

Selectivity. The assessment team chose a flat-top selectivity for the Commercial data, a dome-shaped selectivity for MRIP, and dome selectivity for the discards. Unfortunately, there was little discussion by the Assessment Team of the logic used for this initial formulation.

I would generally choose a flat-top selectivity unless it was expected to be otherwise, for example if there were maximum size regulations, if the large fish were known to migrate out of the fishing area and are not available, or there was gear selectivity. In the chosen assessment area, there is little evidence of any of these being the case. It is possible that the larger fish move out of the assessment area and end up in NE Florida or further north, but the tagging data from the Florida Keys does not suggest this.

The assessment model estimated values for all the selectivity parameters, albeit several with very high standard deviations. For example,

length-composition-top_logit-HB estimate: -12.63, standard deviation 83.80,

length-composition-top_logit-MRIP: estimate 14.64 standard deviation 55.4,

length-composition-descend-MRIP estimate -10.57, standard deviation 88.56.

The Review Panel asked the Assessment Team to do a model run with flat top selectivity for MRIP. The age-4 fishing mortality was higher, as expected, because more age-4 fish are being selected. The spawning stock biomass was considerably lower but the threshold level also changed considerably. As a result of this exercise, I believe the base model gives a satisfactory result, but in future assessments, I would start with the flat-top being the default unless you expect it to be otherwise.

Fishing Mortality. The Assessment Team used age-4 Fishing Mortality for determination of stock status. The Review Panel were interested in the fishing mortality for all ages. The plot provided showed that for the commercial fleet, the F values flatten out at age 6, for Headboat the F values peak at age 4, and for

MRIP the *F* values peak at age 2 and flatten out at about age 8. From this, I conclude that the model Base model of using age-4 fishing mortality is well justified.

The Assessment Team did a catch-curve analysis to estimate fishing mortality using the Chapman and Robson method. The age data came from the assessment area, ages used were 4 to 20. The result was a fishing mortality of 0.35. There is good agreement between this result and the base model fishing mortality of the assessment model. This is further evidence that the assessment model is producing results that are sound and useful for determining stock status.

Recruitment Deviations. These were handled in the usual manner with an included bias adjustment calculated in the SS3 software. When the effective sample size was calculated, the model suggested bias adjustment changed slightly. The Assessment Team used a parsimonious approach and did not make a further bias adjustment. I believe this is an acceptable approach, as the recruitment deviations are likely to show little change.

Retrospective Analysis. This is used to see if there are systematic changes in derived quantities, like spawning biomass, which could lead to errors in determining the stock status. If there are temporal changes in input parameters that are not catered for in the assessment, then these systematic changes may occur. When data is removed for 1, 2, 3 …7 years, there is a slight systematic decrease in spawning biomass except for removal of 7 years, where there is an increase. For age-4 fishing mortality there were no clearly apparent systematic changes. For age-0 recruits there were some patterns, especially when 6 years of data were removed. The values of Mohn's Rho for spawning biomass was -0.04 (acceptable), for recruitment was -0.10 (not great), and for age-4 fishing mortality was 0.06 (acceptable). The retrospective analysis does not indicate serious problems with the model structure.

Discard Mortality. Changes to Commercial discard mortality from the base model of 10% to 15% resulted in little difference to the model fit. Increasing the MRIP discard mortality to 20% worsened the model fit and increasing MRIP discard mortality to 30%, marginally improved the model fit. In both cases, the result was an increase of MRIP fishing mortality, slightly increased current spawning stock biomass, and slightly increased MSST. The base model 10% discard mortality appears satisfactory.

Age-composition. The base model assumes the age-composition data conforms to a multinomial distribution. This distribution is widely used and easily understood, and I believe is a good choice for incorporating the data in the assessment. For comparison, a Dirichlet distribution was used. The latter resulted in a poor fit to the data and it was rightly rejected for this assessment.

Effective Sample Size. The Assessment Team used iterative calculation of effective sample size for all length composition and age-at-length data using the procedure built into SS3. This is common practice and essential to obtaining a sensible weighting between the various components in the likelihood function.

When the Review Panel requested additional analyses, such as fixing the value of steepness, changing natural mortality, using flat-top selectivity for MRIP, the iterative calculation of effective sample size was not re-done. The review Panel then asked for information on the effect of re-weighting when the model structure is changed. The Assessment Team supplied the results of the Jack-knife analysis (Figure 3) showing the change in the effective sample size when the re-calculation was done after the removal of each index. The results generally show a 5%-8% change, with the greatest change in effective sample size being 27% when MRIP_CPUE is removed.

Weighting Type	Fleet/Survey	Base Run	Remove MRIP CPUE	Remove Comm CPUE	Remove RVC Adult	Remove RVC Juvenile
Length Comps	Com	4.36	5.56	4.02	4.27	4.41
	HB	1.03	1.00	1.05	1.01	1.05
	MRIP	1.50	1.32	1.51	1.45	1.53
	RVC Adult	0.48	0.52	0.49		0.44
	RVC Juv	0.92	1.00	1.05	0.98	
	MRIP_CPUE	6.73		6.67	6.74	7.89
Age @ Length	Com	0.18	0.17	0.17	0.17	0.17
	HB	0.30	0.24	0.29	0.27	0.30
	MRIP	0.14	0.14	0.13	0.14	0.15
	FI	0.16	0.11	0.10	0.10	0.17

Figure 3. Comparison of the Francis Weights for each Jack-knife run.

I believe that the re-calculation of effective sample size should be performed after any changes are made to the model structure, input data, or input parameters. This can be tedious, but I think that a simple procedure could be written to automate this.

In summary, the data decisions made in the Data and Assessment Workshops were sound and led to an assessment which was stable and able to provide advice on the status of the stock. The decisions made on the model structure were explained and the uncertainties acknowledged. As mentioned above, discussion on the reason for including the four chosen indices, and an indication of their relative importance in the assessment would have been valuable.

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?

The tool for the assessment was Stock Synthesis V3.30.13 (denoted SS3). This is used worldwide and has proved to be reliable in many stock assessments. The stock assessment team demonstrated that they had a thorough understanding of SS3 and although the model had convergence problems for some input data situations, these were recognised and handled well by the team. In the SEDAR64 assessment, I consider the assessment method robust.

The assessment team conducted various model runs to compare the results using the previous assessment software ASAP. Model runs with the data for the previous assessment were conducted using ASAP and SS3. The runs with the current data were also conducted in ASAP and SS3. The comparisons showed generally similar results for derived quantities required for determining stock status, that is fishing mortality and spawning biomass. This provides further evidence that the chosen assessment software is a reliable tool for synthesising the available data to produce the quantities needed to determine future catch levels.

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

The model set up was consistent with common practice and I could easily follow the well annotated information in the SS3 data and control files which were supplied by the Assessment Team.

The parameter estimates and derived quantities were generally stable for various input changes (jitter analysis, jack-knife analysis, and alternative model configurations suggested by the Review Panel). This indicated the model was well configured and nearly always gave results in which the stock status was not greatly different from the base model.

Eleven out of the 85 estimated parameters are highly correlated (correlation >0.7). Ideally, you would want to change the model formulation to reduce the number of correlated parameters, but this is not always possible. Of the eleven, four relate to size dome shaped selectivity, which is not un-expected. This has already been discussed, but to re-iterate, if possible, I would use the simple selectivity, a flat-top, unless there were logical reasons for doing otherwise (e.g., discards). The von Bertalanffy growth parameters are highly correlated, which is expected and there is not much you can do about that with this model formulation. The Beverton-Holt steepness parameter, *h*, is correlated with the initial number in the population, which typically occurs. This has been discussed and the panel suggested, and I agree, that steepness should be fixed at say 0.8. Unfortunately, allocating values to parameters causes the variability of the derived quantities of interest to be underestimated.

The model configuration using length composition and also age-at-length is common practice and a good strategy. The large range of ages in each length bin, especially after the legal size, is a cause for concern. Apart from determining the effective sample size, I believe it is worth investigating the impact of having some artificial weighting of the length and age-at-length data sets.

The model base model is configured properly and produced sound results.

c) Are the methods appropriate for the available data?

SS3 is well suited for analysis of the data available in this fishery. It is flexible and allows for many data sets and configuration options. The Assessment Team are very familiar with the workings of this software and incorporated the data appropriately. SS3 is a valuable tool and I believe it should be used for future assessments. I believe ASAP should also be used so that comparisons can be made between the quantities of interest derived from both software packages.

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?

The choice of a single area model for the Florida Keys-SE Florida is justified at the present time as nearly all the landings (96%) come from the chosen assessment area. If landings increase into NE Florida and further north in the future, more information will be required on spawning areas, adult movement, and possible differences in growth.

The assessment model includes landings and discards by fleet, fishery independent, and fishery dependent abundance indices, length-composition, and length-age-composition data. The stock assessment model, SS3, is well suited for integrating this type of data and it was well configured by the assessment team. The basic diagnostics for an assessment model, such as plots of model fits and residual plots, gave satisfactory results indicating the model is an informative assessment tool.

The Assessment Team used a number of standard procedures to test whether the assessment model was consistent. Jitter analysis, whereby starting values were varied, resulted in model convergence in most cases. A jack-knife analysis where indices were removed, one at a time, from the input data, produced estimates of abundance, exploitation, and spawning biomass which varied considerably from the base model, especially the removal of the MRIP index which produced larger estimates of spawning biomass. In all cases stock status was essentially unchanged. Thus, the population estimates are generally consistent with the trajectory of the indices.

The model estimates of derived parameters such as abundance, exploitation, and spawning biomass, are reliable and useful to support stock 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?

The stock is not overfished. The standard accepted for this fishery for assessing spawning biomass is the Minimum Stock Size Threshold, which is $0.75*SSB_{F30\%SPR}$. The Spawner-Per-Recruit estimates, based on data for this species, I consider to be reliable. The current estimated spawning stock biomass is well above the accepted Minimum Stock Size Threshold. In addition, the graph of the model estimated spawning stock biomass has been increasing over the last 10 years.

The stock is not experiencing overfishing. The accepted standard is the Maximum Fishing Mortality Threshold which is $F_{30\%SPR}$. The SPR estimates are reliable, I believe, and the value of $F_{current}$ (the GM of F on age-4 fish from 2015-2017) is below the accepted standard. Although the fishing mortality has increased in the last 2 years, it is not outside the range of values that the model estimated for the last decade.

The Review Panel requested model runs (e.g., change steepness, change natural mortality, alter selectivity of MRIP) did not substantially alter the comparison of the quantities of interest to the threshold levels. This gives me further confidence in the reliability of the reported stock status.

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

The Assessment Team chose the Beverton-Holt stock-recruitment relationship. This is a widely used and is an acceptable choice for this assessment. The steepness parameter had an estimated value of h=0.808, there being some informative data on the increasing portion of the curve. The likelihood profile was very flat over the values h=0.75 to h=0.9, meaning that the stock-recruitment curve fits the data equally well over a range of values of steepness.

At the request of the Review Panel, the Assessment Team did model runs with steepness fixed at h=0.7 and h=0.9. The results showed little change in the recruitment from the base model where h=0.808. A run with h=0.99 was conducted, but the model did not converge and the result should be ignored. The spawning biomass showed little difference being highest for h=0.7 and lowest for h=0.9. The age-4 fishing mortality varied little from the base model.

My experience is that estimation of steepness is problematic as is the case in this assessment and I believe the Assessment Team should not estimate steepness but rather assign it a value, say h=0.8, or h=0.808, or estimate it with tight bounds.

The stock-recruitment relationship selected is an increasing function and gives a reasonable fit to the data for lower stock sizes. The problem is, it gives an equally satisfactory fit for a quite large range of values of steepness, which is quite common. I believe it is not necessary to explore alternative stock recruitment relationships, as this formulation provides satisfactory information on stock status. In addition, the optimum yield for the values of steepness from 0.7-0.9 were almost the same.

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?

The assessment team used a simple model structure (one area, one sex) and conducted the assessment using standard approaches. They explored many alternatives for the input parameters (steepness, natural mortality, discard mortality), various data input formulations (different time series of landing data, leaving out indices), and the results indicated quite similar conclusions for the level of age-4 fishing mortality and spawning biomass compared to the adopted threshold levels. The Optimum Yield and the Fishing Mortality at Optimal Yield can be reliably estimated using this assessment model.

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?

The Assessment Team used a standard approach whereby the 5 year projection, 2018-2022, was based on the average F for 2015-2017, $F_{30\%SPR}$, and $0.75F_{30\%SPR}$. The approach used is standard in the region and consistent with accepted practice.

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

The projections were implemented in the software package SS3 and have been used in many assessments and are reliable. The recruitment for the projections is calculated from a Beverton-Holt stock recruitment relationship and will not reflect the variability that is likely to occur.

The decline in the projected spawning biomass appears contrary to what the current trend suggests. This is because the recruitment is based on the current average levels of recruitment and do not take into account the strong recruitment of 2011 to 2014. However, the method used is appropriate and the results are consistent with the approach adopted.

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

The results produced used a standard and much used procedure and are robust. The results give a somewhat conservative view of the future stock size, but are useful to provide information for determination of stock status and future catch levels.

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

The uncertainties presented consist of asymptotic confidence intervals for the derived quantities and are rather narrow. The method used was discussed and the limitations of the approach used are acknowledged. In the future, methods for obtaining more realistic uncertainty can be explored, but for the current assessment, the approach is satisfactory.

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

a) Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods.

The uncertainty was explored thoroughly using standard methods of sensitivity runs, likelihood profiles, retrospective analysis, parametric bootstrapping, and Monte Carlo Markov Chain runs. All of these methods are standard procedures and the results were valuable to understanding the degree of uncertainty of the outputs.

The bootstrap runs are generally useful for exploring uncertainty but the results presented showed that for 445 of the 500 runs, at least one parameter estimate was close to its bound. Consequently, it was difficult to interpret the uncertainty using this procedure. There is work to be done for the next assessment to improve the usefulness of this procedure.

The likelihood profile is valuable to investigate the stability of the parameter estimates. This procedure will show how reliably the model finds a global minimum for the negative log-likelihood, shows how flat the negative log-likelihood function is near the parameter estimate, and how often the model fails to find the minimum. The procedure indicated that the SS3 model showed stable estimation of the parameters profiled, but as mentioned earlier the steepness parameter has a very flat likelihood profile. The initial fishing mortality for the Commercial, MRIP, and Headboat fleets was not very stable, missing the minimum on a number of occasions. Although this procedure indicated some instability of some parameters, in general it produces reliable estimates for the parameters. I think the likelihood profile procedure should be extended to more parameters to get a better overall picture of the model convergence.

The MCMC procedure is valuable for checking the reliability of asymptotic confidence intervals for derived parameters. In addition, it is valuable for generation of non-symmetric confidence intervals which can be especially valuable for cases when the derived parameter, for example spawning biomass, decreases to low levels. In this fishery, where clearly the stock is not overfished nor is overfishing occurring, the asymptotic confidence intervals are reliable and satisfactory.

The many results presented to the Review Panel were informative and indicated the SS3 model was useful for evaluating the sensitivity of the model to different input parameters and model configurations. When the Review Panel asked for model runs different to those in the presentation, in several cases, these sensitivity investigations had already been done and were presented in a very timely manner.

In future review workshops, it would be good if the Assessment Team presented more information on more of the analyses conducted. There was little discussion on the reason for flat-top or domed selectivity. If alternative options were investigated by the assessment team, then these should have been presented to the Review Panel.

There was considerable discussion on the weighting of the composition-data and the indices. I believe there is merit in determining a sensible weighting of the various data inputs according to their perceived relative reliability. This could possibly be considered in the next assessment.

These thorough investigations show that the assessment model performed well with selected model structure: one area, one sex, pooled data for the Florida Keys-SE Florida, and the time period 1992-2017. This structure is justified for the current assessment.

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

The assessment team did a thorough job in exploring the uncertainties and pointing out the implications for the fishing mortality and spawning biomass trajectories. I am confident that these derived quantities are suitable for making decisions about the stock status and optimal yield.

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

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

The current model is structured as a one area, one sex model. The area used for the assessment was the Florida Keys-SE Florida. If future landings expand into other areas where yellowtail snapper occur, it will be important to collect the following further information.

- Where are the spawning areas, if any, outside the current assessment area?
- What is the degree of movement of adults between the current assessment area and NW Florida, S. Carolina, and the Gulf of Mexico?
- Are there biological differences between yellowtail snapper in the assessment area and those further north?

More detail is provided on this in TOR 8.

b) Provide recommendations on possible ways to improve the SEDAR process

The process for SEDAR64 was excellent. Having the meeting at the hotel was an excellent idea. The technical side of things worked perfectly, and the support staff did an excellent job. The Data Workshop and Assessment Workshop documents, and the SS3 files used for the assessment, were available well ahead of the Review Workshop.

The Assessment Team members were well prepared, had an excellent understanding of the SS3 assessment package, and their presentations were very professional. All queries by the Review Panel were answered in a timely manner.

Here are a few very minor points.

- In the Data Workshop reports it would be an improvement if a consistent format was used. Figure number always below the graph: e.g., S64_DW_1.
- In each of the DW reports, the authors say “**Please cite this document as”.** It would be good if the document list was in the format suggested by the authors rather than three columns with the authors in the right hand column.
- In the PowerPoint slides, it would be an improvement if the slides were numbered.

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.

Appropriate: The assessment model used for this assessment was SS3. It is more versatile that ASAP and is used in many countries for stock assessments and is robust and reliable. It is well suited for providing the outputs necessary to determine stock status and future catch levels.

Relevance. The SS3 assessment tool is highly appropriate for synthesis of the data sets available for yellowtail snapper.

Inclusiveness: At the review meeting, there were representatives from various agencies, including a commercial fisher. All participants were invited to comment on several occasions during the workshop.

Objectivity: The model outputs are based on the best available data inputs, and the limitations of these inputs are recognized and acknowledged.

Transparency: The assessment team produced many diagnostics which indicated the short-comings of the parameter estimation and convergence. When the Review Panel asked about the procedures used in the model, the assessment team were always open and candid in their explanations.

Timeliness. The Data and Assessment information and the SS3 input files were supplied well before the meeting. The previous assessment was done a decade ago. I believe the next assessment should be done in a timelier manner.

Verification: The assessment software ASAP was used in the previous assessment (SEDAR 27a). The current data were analyzed in SS3 and also ASAP with the data setup and the parameters used matching as closely as possible in the two models. The trajectories of the derived parameters, spawning biomass and fishing mortality showed similar trajectories. I believe SS3 is a good choice for the yellowtail snapper assessment, but the ASAP software should also be used in the future to provide verification of the results.

Validation: Comparisons were made between the outputs of the SS3 and ASAP model, using both the data of the previous assessment and the current data. Although each model had its strengths, SS3 is widely used, very flexible and is an excellent choice for future stock assessments in this fishery.

Peer Review: At the Data and Assessment Workshops, the data sources were reviewed and decisions were made on the start year for the model, the extent of the model area, what indices would be used in the assessment, and the method of determining natural mortality. The process used is extensive and valuable.

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

The decision to use a single area model which encompassed the area of the Florida Keys-SE Florida was justified as nearly all the landings of yellowtail snapper come from this area. It is important to get information on the sources of juvenile fish in a much wider area (NE Florida, North Carolina, and west of Florida in the Gulf of Mexico).

The degree of mixing of adult fish should be investigated. I recommend the use of existing otoliths and stable isotope analysis. This can be done rapidly and is a fairly inexpensive way to determine the extent to which adult fish are moving between the assessment area and NE Florida/North Carolina and also Gulf of Mexico. This procedure is likely to detect stock separation over a distance of 200 nm – 300 nm if it exists.

The age-length key shows a large range of ages in each length class. The Marginal increment analysis, I feel, rules out annuli not being annual growth rings. Florida FWC staff assured me that yellowtail snapper otoliths are not difficult to read. There was a large deviation from the agreed age especially from some readers (a deviation of 3 annuli for an agreed count of 7) during the quality control ageing of 1,197 otoliths. There should be an investigation into ways of improving consistency.

There is Reference Collection of 100 otoliths which are probably used periodically to help readers get their eye in before reading yellowtail snapper. There is no mention of whether the Reference Collection is static. I suggest a portion of otoliths in the Reference Collection should be replaced periodically because with a static collection, readers will remember the count they gave previously to difficult otoliths.

The collection of otoliths is focused in the Florida Keys, where I believe the number of otoliths collected is excessive (1827 collected in 2016) but in NE Florida to North Carolina it ranged from 1 to 21 collected annually. I suggest that some resources for otolith collection in the Keys be re-directed to areas where the numbers are very low, like NE Florida.

The shape of the selectivity functions are a concern for me. When MRIP selectivity was changed from dome shaped to flat-top, the components of the –LL showed a large change in the age-composition component and little change in the other components. With this change to selectivity, the fishing mortality was vertically displaced higher and closer to the thresh-hold level and the spawning biomass was lower, rising less rapidly in recent years, and was closer to the thresh-hold level. My recommendation is that the shape of the selectivity function should be chosen by the Review Panel and not by the assessment model fit, especially if the age data is influential in the model fit.

There was concern by the Review Panel, and I agree, that more investigation is required into the merit of the current use of age-at-length data in the model. The length-composition data is obviously more reliable and it could be worth artificially weighting these two data sets, as well as the indices, and comparing the model outcomes.

The iterative determination of the effective sample size should be re-calculated, whenever the inputs or model structure is changed. This was done for the jack-knife analysis, but not, I believe, for the ad hoc requests of the Review Panel. In some cases, the re-weighting could result in significant changes to the derived parameters. I believe, a simple procedure could be written to perform the 4 or 5 re-weighting runs automatically.

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.

A Panel Report has been prepared and approved by the Review Panel.

Appendix 1. SEDAR 64 South-eastern US Yellowtail Snapper Document List

Documents Prepared for the Data Workshop

S64-DW-01	Campbell, Matthew D., Kevin R. Rademacher, Michael Hendon, Paul Felts, Brandi Noble, Ryan Caillouet, Joseph Salisbury, and John Moser. 2018. SEAMAP Reef Fish Video Survey: Relative Indices of Abundance of Yellowtail Snapper. SEDAR64-DW-01. SEDAR, North Charleston, SC. 27 pp.
S64-DW-02	Swanson, Christopher E. 2019. A model-based index of Yellowtail Snapper, *Ocyurus chrysurus*, in the Dry Tortugas using Reef Fish Visual Census data from 1999-2016. SEDAR64-DW-02. SEDAR, North Charleston, SC. 11 pp.
S64-DW-03	Swanson, Christopher E., Kerry Flaherty-Walia, and Alejandro Acosta. 2019. Juvenile Yellowtail Snapper, Ocyurus chrysurus, collected from short-term fisheries-independent surveys in Florida Bay and the Florida Keys from 1994 – 2003. SEDAR64-DW-03. SEDAR, North Charleston, SC. 19 pp.
S64-DW-04	Swanson, Christopher E., and Robert G. Muller. 2019. A model-based index of Yellowtail Snapper, Ocyurus chrysurus, for the Florida Reef Tract from Card Sound through the Florida Keys using Reef Fish Visual Census data from 1997-2016. SEDAR64-DW-04. SEDAR, North Charleston, SC. 11 pp.
S64-DW-05	Herbig, Jennifer, Jeffrey Renchen, Alejandro Acosta. 2019. Fisheries-independent data for Yellowtail Snapper (Ocyurus chrysurus) from reef-fish visual surveys in the Florida Keys and Dry Tortugas, 1999-2016. SEDAR64-DW-05. SEDAR, North Charleston, SC. 40 pp.
S64-DW-06	Swanson, Christopher E. 2019. A model-based index of Yellowtail Snapper, Ocyurus chrysurus, for the Northern Florida Reef Tract from Government Cut through Martin County using Reef Fish Visual Census data from 2012-2016. SEDAR64-DW-06. SEDAR, North Charleston, SC. 10 pp.
S64-DW-07	Carroll, Jessica, Kristen Rynerson, Brittany Barbara 2019. Accuracy and precision of Yellowtail Snapper (Ocyurus chrysurus) age determination. SEDAR64-DW-07. SEDAR, North Charleston, SC. 11 pp.
S64-DW-08	Herbig, Jennifer, Alejandro Acosta, Ariel Wile. 2019. Abundance and Distribution of Juvenile Yellowtail Snapper in Nearshore Seagrass Habitat in the Middle Florida Keys. SEDAR64-DW-08. SEDAR, North Charleston, SC. 11 pp.
S64-DW-09	Herdter, Liz. 2019. Standardized Catch Rates of Yellowtail Snapper (*Ocyurus chrysurus*) from the Marine Recreational Information Program (MRIP) in Southeast Florida and the Florida Keys, 1981-2017. SEDAR64-DW-09. SEDAR, North Charleston, SC. 41 pp.
S64-DW-10	Allen, Shanae Liz Herdter, and Kelly Fitzpatrick. 2019. Overview of the Southeast Region Headboat Survey and Data Related to Yellowtail Snapper (*Ocyurus chrysurus*). SEDAR64-DW-10. SEDAR, North Charleston, SC. 25 pp.
S64-DW-11	Herdter, Liz and Allen, Shanae. 2019. Standardized Catch Rates of Yellowtail Snapper (*Ocyurus chrysurus*) from the U.S. Headboat Fishery in Southeast Florida and the Florida Keys, 1981-2017. SEDAR64-DW-11. SEDAR, North Charleston, SC. 44 pp.
S64-DW-12	Matter, Vivian M. and Richard C. Jones. 2019. Recreational Survey Data for Southeast Yellowtail Snapper. SEDAR64-DW-12. SEDAR, North Charleston, SC. 23 pp.

S64-DW-13 Brown, Steve and Chris Bradshaw. 2019. Historical Commercial Fishery Landings of Yellowtail Snapper in Florida and the Southeastern U.S. SEDAR64-DW-13. SEDAR, North Charleston, SC. 16 pp.

S64-DW-14 Bradshaw, Chris and Steve Brown. 2019. Length frequency distributions for yellowtail snapper collected by TIPS in the Southeast from 1984 to 2017. SEDAR64-DW-14. SEDAR, North Charleston, SC. 13 pp.

S64-DW-15 Atkinson, Sarina F., Kevin J. McCarthy, Allison C. Shideler. 2019. Length distribution and release discard mortality for southeastern yellowtail snapper. SEDAR64-DW-15. SEDAR, North Charleston, SC. 6 pp.

S64-DW-16 Lazarre, D. 2019. A Summary of Observer Data Related to the Size Distribution and Release Condition of Yellowtail Snapper from Recreational Fishery Surveys in Florida. SEDAR64-DW-16. SEDAR, North Charleston, SC. 30 pp.

S64-DW-17 Scyphers, S. and K. Furman. 2019. Social Dimensions of the Recreational Fishery for Yellowtail Snapper (*Ocyurus chrysurus*) in Florida. SEDAR64-DW-17. SEDAR, North Charleston, SC. 7 pp.

S64-DW-18 McCarthy, Kevin and Jose Diaz. 2019. Calculated discards of yellowtail snapper from commercial vertical line fishing vessels in southern Florida. SEDAR64-DW-18. SEDAR, North Charleston, SC. 15 pp.

Documents Prepared for the Assessment Process

S64_AP_01 Allen, S.D. 2019. Weighted Length Compositions for U.S. Yellowtail Snapper (*Ocyurus chrysurus*) from 1981-2017. SEDAR64-AP-01. SEDAR, North Charleston, SC. 36 pp.

S64_Assess_Report SEDAR 64 Southeastern US Yellowtail Snapper SECTION III: Assessment Process Report. 2020. SEDAR. 178 pp.

Reference Documents

S64-RD01 Coral Reef Conservation Program (CRCP) Local Action Strategy (LAS) Project 3B. "Southeast Florida Coral Reef Fishery-Independent Baseline Assessment" - 2012-2013 Interim Report.

S4-RD02 Florida Fish and Wildlife Conservation Commission. Implementing the Dry Tortugas National Park Research Natural Area Science Plan - The 10-Year Report.

S64-RD03 Jennifer L Herbig, Jessica A Keller, Danielle Morley, Kristen Walter, Paul Barbera, Alejandro Acosta. Examining movement patterns of yellowtail snapper, *Ocyurus chrysurus*, in the Dry Tortugas, Florida.

S64-RD04 SAFMC Snapper Grouper Advisory Panel. Yellowtail Snapper Fishery Performance Report.

S64-RD05 Francesca C. Forrestal, M. Danielle McDonald, Georgianna Burress and David J. Die. Reflex impairment and physiology as predictors of delayed mortality in recreationally caught yellowtail snapper (*Ocyurus chrysurus*).

S64-RD06 Claudine T. Bartels and Karole L. Ferguson. Preliminary Observations of Abundance and Distribution of Settlement-Stage Snappers in Shallow, Nearshore Seagrass Beds in the Middle Florida Keys.

S64-RD07 William F. Loftus. *Lutjanus Ambiguus* (Poey), a Natural Intergeneric Hybrid of *Ocyurus Chrysurus* (Bloch) and *Lutjanus Synagris* (Linnaeus).

S64-RD08 M. L. Domeier and M. E. Clarke. A Laboratory Produced Hybrid Between *Lutjanus Synagris* and *Ocyurus Chrysurus* and a Probable Hybrid Between *L. Griseus* and *0. Chrysurus* (Perciformes: Lutjanidae).

S64-RD09 Beverly Sauls and Oscar Ayala. A Survey to Characterize Harvest and Regulatory Discards in the Offshore Recreational Charter Fishery off the Atlantic Coast of Florida.

S64-RD10 Kerry E. Flaherty-Walia, Brett Pittinger, Theodore S. Switzer, Sean F. Keenan. Seagrass Habitats as Nurseries for Reef-Associated Fish: Evidence from Fish Assemblages in and Adjacent to a Recently Established No-Take Marine Reserve in Dry Tortugas National Park, Florida, USA.

S64-RD11 A. Acosta, C. Bartels, J. Colvocoresses, and M. F. D. Greenwood. Fish assemblages in seagrass habitats of the Florida Keys, Florida: spatial and temporal characteristics.

S64-RD12 Kyle Dettloff and Vivian Matter. Model-estimated conversion factors for calibrating Coastal Household Telephone Survey (CHTS) charterboat catch and effort.

estimates with For Hire Survey (FHS) estimates in the Atlantic and Gulf of Mexico with application to red grouper and greater amberjack.

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-nationalregistration-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.
   a) Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods
   b) Ensure that the implications of uncertainty in technical conclusions are clearly stated
6. Consider the research recommendations provided by the Data and Assessment workshops and make any additional recommendations or prioritizations warranted.
   a) Clearly denote research and monitoring that could improve the reliability of, and information provided by, future assessments
   b) Provide recommendations on possible ways to improve the SEDAR process
7. Consider whether the stock assessment constitutes the best scientific information available using the following criteria as appropriate: relevance, inclusiveness, objectivity, transparency, timeliness, verification, validation, and peer review of fishery management information.
8. 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 24-26, 2020. Saint Petersburg, Florida

Monday:

9:00 a.m. Introductions and Opening Remarks Coordinator

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

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

Tuesday:

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

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

Wednesday:

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

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

Appendix 3. List of Participants

Review Workshop Panel

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

Analytic Team

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

Attendees

Ed Walker	Appointed Observer	GMFMC AP

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	Co-ordinator	SEDAR
Mike Errigo		SAFMC
Ryan Rindon		GMFMC
Charlotte Schiaffo		GMFMC