CENTER FOR INDEPENDENT EXPERTS (CIE) REVIEW OF SEDAR 46: U.S. CARIBBEAN DATA-LIMITED SPECIES

Dr Cathy Dichmont

Cathy Dichmont Consulting
cathydichmont@gmail.com

Contact Details

Name: Dr Catherine M. Dichmont
Email: cathydichmont@gmail.com

Review of SEDAR 46: U.S. Caribbean data-limited species

CONTENTS

1 EXECUTIVE SUMMARY

The review workshop for the Caribbean data-limited species took place in Miami, Florida from February 23-25 2016. In attendance were review panel members Drs Apostolaki, Ault, Dichmont, Hoenig, Medley, and chair Vicinte with the stock assessment team and other scientists involved in the stock assessment, managers and industry. The review was undertaken in a very co-operative light with requests for additional work met, including undertaking extensive sensitivity tests. The team are thanked for a very constructive meeting.

The review investigated the body of work that applies a Management Strategy Evaluation (MSE) using the DLMtool authored by Dr Carruthers to six species-island units in the Caribbean. Extensive sensitivity tests were undertaken that investigated the effect of different operating model and management model settings. Together with those undertaken during the review, this body of work is very comprehensive.

The species-island unit selection was based on landings, availability of length data and basic life history parameters, and the relative importance of the species to each of Puerto Rico, St Thomas and St Croix islands. One of the species chosen was more data limited and selected as a test case. The approach used to select species-island units is supported.

Extensive work was undertaken to create a comprehensive database of information (data and life history parameters) that was required to undertake the Management Strategy Evaluation. A detailed description of the uncertainty inherent in the data was provided. Given the data limited status of these species-island units, the best available data and life history parameters were used in the analysis. Further work in this area is recommended, focusing resources on aspects, such as growth parameters for St Thomas queen triggerfish and appropriate parameter values for the length at full selection.

The operating model was poorly described in the DW/RW report, but the use of the DLMtool is appropriate and has been correctly applied to the cases. The performance tests of the operating model were of a limited nature and focused on whether runs have converged. Further tests were undertaken during the workshop and these highlight areas for future work.

The DLMtool does not allow uncertainty in parameters (beyond uniform distributions) to be input in a sophisticated manner. Correlations between parameters are ignored. The result is that a greater region of uncertainty is selected, which could include parameter combinations that are unreasonable. However, it should be pointed out that this approach is more precautionary, since this is likely to result in more performance measures being excluded from the candidate list. These tests would also emphasise highly robust performance measures. Some changes to the DLMtool are suggested, although it is acknowledged that the author of the toolkit is not part of the assessment team.

The search of possible Management Procedures was extensive given what was available in the toolkit; however, these Management Procedures were not further adapted to the species-island units. Many of the Management Procedures have been applied in quite different fisheries elsewhere. Thus, further refinement of a smaller set of Performance Measures should be undertaken to tune them to the local circumstances.

A series of Management Procedures was tested, amongst these were:

a) two delay-difference Management Procedures (that cannot be implemented in practice given the uncertainty in the catch and catch rate index) were included to show how a more data rich Management Procedure would perform, and
b) a perfect knowledge Management Procedure, which is a very useful standard for comparison.

The presently applied Management Procedure was also tested during the workshop and showed that it does not fall within the critical range of acceptable Management Procedure performance.

The mean length Management Procedure was initially undertaken as a separate process outside the toolkit, which meant the results were difficult to integrate. However, during the workshop this integration was completed, which was of great benefit. Integrating this application within the toolkit is supported and recommended.

Extensive testing of different selectivity functions, life history parameters and whether key parameters were biased or imprecise was undertaken. There is an impressive amount of tests to highlight both operating model and management procedure sensitivities, which greatly helps final selection of candidate Performance Measures. However, the biased and imprecise tests were not appropriately set up. An upward bias was applied to parameters that would normally have been correlated. Furthermore, it is preferred that full sensitivity tests are undertaken on the biased and imprecise settings (i.e. it is the base model), rather than the unbiased and precise Operating Model settings. Real life sensitivity tests were extensive and useful. These highlighted the difference between the Performance Measures that were applied within the MSE and the smaller subset of the "real world" Performance Measures.

Four Performance Measures were described. These emphasised long term performance, and are standard ones used in many MSEs. They were appropriately aligned with the management requirements. During the workshop more transitional short term yield performance was investigated, but further work in this area would be useful.

The results were summarised prior to the workshop, but did not really fulfil all that was needed to select candidate Management Procedures. During the workshop, the team were able to reduce the list of candidate Management Procedures. These were based on the Performance Measure cut-offs, whether the management procedures showed important sensitivities in the tests, the uncertainty in the real world OFL distribution and whether the data the Management Procedure uses is in fact reasonably reliable. The final list is supported. However, further Performance Measures should be developed that would capture this approach in a more formal way.

Research recommendations are made that particularly emphasise changes to the DLMtool's handling of variance and bias, amongst others.

Comments on the workshop process are made, including that an additional workshop between the data and review workshops be undertaken in the future. It is suggested that there be a data workshop where the input data and parameters are selected (with the MSE and real world Performance Measures in mind), a second workshop where the Operating Model and Management Procedures are selected and run, and a final review workshop where the results are reviewed.

All Terms of References were met and the candidate Management Procedures selected as part of the review process can be used for management.

The assessment team did a great deal of original work in a short time period and responded fully to panel requests for additional analyses; they are to be congratulated and thanked for a constructive and open review.

2 BACKGROUND

The review workshop for the Caribbean data-limited species took place in Miami, Florida from February 23-25 2016. In attendance were review panel members Drs Apostolaki, Ault, Dichmont, Hoenig, Medley, and chair Vicinte with the stock assessment team, and other scientists involved in the stock assessment and management. The review was undertaken in a very co-operative light with requests for additional work met, including undertaking extensive further sensitivity tests. The team are thanked for a very constructive meeting.

The panel members were provided with material pertaining to the Management Strategy Evaluations, the Data/Assessment Review Workshop, the mean length-based estimator and several relevant items of background material (see page 58 for a list of documents provided). Presentations on the data and Management Strategy Evaluation were provided during the review and were comprehensive and informative. Public comment was also provided by two fishers who gave valuable background, and addressed comments and requests from the panel.

More detailed tests were undertaken and provided during the review process on request. These included refining the MSE, tests and final list of candidate Management Procedures that could be used for management.

It should be noted that all the tables and figures are from the Addenda provided after the workshop, or derived from these.

3 DESCRIPTION OF THE INDIVIDUAL REVIEWER'S ROLE IN THE REVIEW ACTIVITIES

A set of documents was provided to the reviewers prior to the workshop (see Appendix 2). These included three key documents on the Data/Assessment Review Workshop, the DLMtool Management Strategy Evaluation work, and the mean length estimator. The final agenda and Statement of Works are attached in Appendix 1. Workshop attendees are listed in Appendix 2.

The review workshop was held from 23-25 February 2016 in Miami, Florida. In attendance were review panel members Drs Cathy Dichmont, John Hoenig, Jerald Ault, Vance Vicinte (chair), Paul Medley, Panayiota Apostolaki, stock assessment authors and other scientists involved in the stock assessment.

A series of presentations by the assessors was provided (see Appendix 1).

The panel provided an overview of their findings on the last day of the workshop. A panel report was contributed to according to the timelines.

The Terms of Reference of the review were:

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 the data-limited methods scientifically sound and robust?
    b. Are the methods appropriate given the available data?
    c. Are the data-limited models configured properly and used in a manner consistent with standard practices?
    d. Are the quantitative estimates produced reliable? Does the method produce management metrics (e.g. MSY, ABC, ACL) or other indicators (e.g. trends in F or Z, probability of overfishing) that may be used to inform managers about stock trends and conditions?
3. 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.
4. 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 future assessments.
    b. Provide recommendations on possible ways to improve the SEDAR process.
5. 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.
6. Provide guidance on key improvements in data or modeling approaches that should be considered when scheduling the next assessment.
7. CIE Reviewers May contribute to a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference.

4 SUMMARY OF FINDINGS FOR EACH TOR IN WHICH THE WEAKNESSES AND STRENGTHS ARE DESCRIBED

4.1 BACKGROUND

Below is a response to each Term of Reference. In these, the term "assessment" is interpreted as relating to the Management Procedure. The term "assessment" is being used to determine whether the Management Procedure results in overfishing or a stock being overfished over the long term (whereas a stock assessment implies determination of stock status, which is not an available option in most of these data limited analyses).

Commonly used terms and their abbreviations in this and the reviewed reports are provided in Table 1.

Table 1. Commonly used terms and abbreviations

Term	Abbreviation
Management Strategy Evaluation	MSE
Puerto Rico	PR
St. Thomas	STT
St. Croix	STX
Catch per Unit Effort	CPUE
Total Allowable Catch	TAC
Overfishing Limit	OFL
Data Workshop	DW
Assessment Workshop	AW

4.2 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?

4.2.1 ToR conclusion

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

Yes, since they do not attempt to undertake a stock assessment in the real world case, but rather use the data within the context of the MSE or as a sub-set for the real world Management Procedures.

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

Yes, the strengths and weaknesses of the data are acknowledged.

C) Are data applied properly within the assessment model?

Yes, these are used with the appropriate caution in the data limited approaches.

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

Yes, for the final candidate set of Management Procedures that were selected.

4.2.2 Background

The Data/Assessment process applied a triage approach to the data, and includes choosing the species-island units that would be selected for the MSE, the data and parameters to be applied for a) the Operating Model, b) the Management Procedure within the MSE, and c) the Management Procedures with real world data. These have been undertaken with care, although further work on the Operating Model process is described in Terms of Reference 2 and 6.

The data for the Caribbean are acknowledged to be of a quality and quantity that would not sustain a credible traditional stock assessment, although processes are underway to address this, for example, using independent surveys. As a result, the data triage concentrated on the key basic data that would be required for the MSE. The data and associated parameters were suitable for simulation testing purposes such as an MSE, but these data would not be able to support a stock assessment that requires, at the least, a reliable catch and index of abundance time series. The acknowledged shortcomings of the data restricted the choice of Management Procedures available in the real world and these limitations will remain for some time.

4.2.3 Species-island unit selection

Six species-island units were chosen for the MSE tests. Five of these were seen as being less data limited than the others, whereas the sixth was seen as an example of a species that may provide challenges for the MSE. These six were chosen primarily based on their importance in terms of landings in each of PR, STT and STX, and availability of other data such as length. The choice of the species-island unit is supported, particularly adding a species that is much more data limited.

For each species-island unit, the dominant gear type data were input into the MSE and real world tests. This choice is appropriate so that a more common selectivity, etc., applies to the data. The DLMtool can only apply one index of abundance, which further emphasised the importance of the correct choice of a single gear type.

4.2.4 Data triage

Within the DW/RW report and during the review, the variability in the catch quality over time and between species-island units have been highlighted. These weaknesses were correctly applied when the candidate Management Procedures for OFL calculations were chosen.

4.2.4.1 Data and life history parameters (general)

The most applicable life history parameters were selected and these are mostly appropriate (see comment on length at full selection). In many cases, the variances are not well described in the reports. If it is the case that the original study did not provide the variances, then the study authors may be approached for the original data. This process has been successfully applied elsewhere as this has allowed the original data to be re-analysed with the associated uncertainty for use in the Operating Model. In the MSE, uniform distributions are used for the priors, so maximum and minimum values are input as a set variance from the mean.

Recommendation 2. Investigate the possibility of obtaining the original data for life history studies where the uncertainty is not well described.

Although in many cases the data are marginal, these are adequate for many of the Management Procedures in the real world and for tests within the MSE (as these, for example, only use recent data which are more accurate).

Given the context, the team assembled the best scientific data available and addressed the uncertainties in the MSE. They also developed Management Procedures that were robust to these uncertainties, which is the best approach for data limited fisheries.

It is important that data and parameter reviews emphasise the needs of the MSE, which are to describe the mean response and associated variances, possible sensitivities and other uncertainties, including bias of the Operating Model, Management Procedure and real world cases.

Recommendation 3. The overall approach is sound, but in the future, should focus on what is needed for the MSE and real world Management Procedures. As such, the assessors should concentrate on key data used by the Management Procedures e.g. CPUE, length frequency, key life history parameters.

Recommendation 4. The focus should also be on refining the data and parameters for the key Management Procedures that are in the final list of Management Procedures and perhaps those that almost made this list.

Recommendation 5. The MSE provides strong indications to which data and parameters the Management Procedures are most sensitive. Future work should concentrate on these.

4.2.4.2 Growth

The most representative parameter values were used in the Operating Model based on a triage of available data and their applicability. However, a meta-analysis of, for example, growth parameters and natural mortality at the family level from globally applicable or the Caribbean regions' data may be more informative to provide parameter value uncertainty and correlations. For example, meta-analyses undertaken in Hawaii provided very useful input to their simulation study to calculate OFLs. This work also provides guidance to possible biases and uncertainty values.

Recommendation 6. Greater emphasis should be given on more realistic input values for biases and the sensitivity of the MSE to these in future iterations, especially when considering correlations in parameters.

Recommendation 7. Undertake a meta-analysis of key parameters such as growth and natural mortality to calculate their variances and potential biases in the Operating Model.

The usefulness of applying an age-based Operating Model that assumes smooth growth for a lobster species was questioned; however, several studies have shown that these can be a good approximation, especially with large uncertainties and is therefore correctly used for the spiny lobster applications.

4.2.4.3 Catch

Perhaps the greatest weakness in several of the species-island units are the catch data (as acknowledged by the assessors). Several expansion factors are applied over time to allow for biases in the data over time. Also, re-allocating unspecified species to recent species-specific categories, for example, parrotfish provided challenges. For parrotfish, a more appropriate percentage value should be applied as the one used was lower than that indicated by the data.

Sensitivity tests during the review were undertaken to explore an increase in the catch of +30% for each of the species island units, i.e. how bias in the catch could impact the TAC recommendations. As expected, this bias generally affects the delay difference Management Procedures more than the others in absolute numbers (Figure 1 and Figure 2); however, for the other Management Procedures there are quite species-specific results. For PR Hogfish and Yellowtail Snapper, the +30% catch impact was an increase in TAC proportion of about 1.1-1.5 for all the Management Procedures. Unusually, the delay difference "DD" impact is a decrease in TAC for STT Queen Triggerfish.

Recommendation 8. Review the re-allocation factor applied to unspecific parrotfish.

In some respect these uncertainties were also addressed in the observation error model of the Operating Model, especially for the biased sensitivity tests, and sensitivity tests changing the *CV_Cat* input value. Furthermore, many of the Management Procedures use the most recent catches, which in many cases are more accurate.

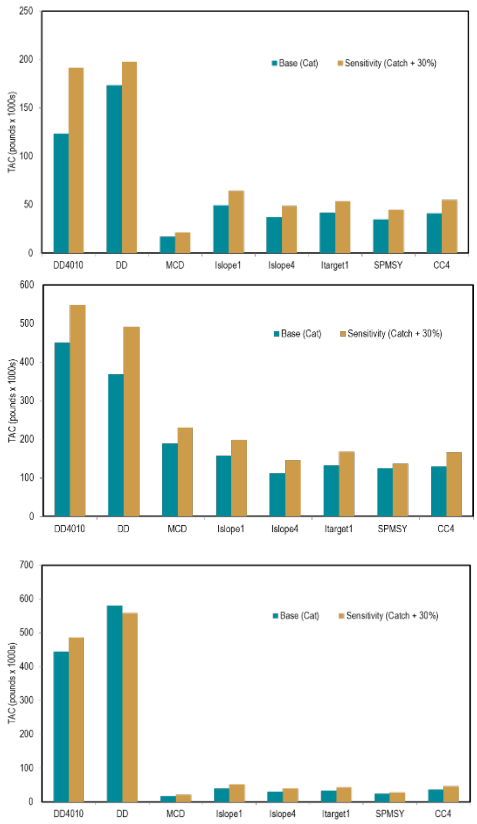

Figure 1. TAC sensitivity to +30% bias in catch. Top: PR Hogfish, Middle: PR Yellowtail Snapper, Bottom: STT Queen Triggerfish, Source: Cummings and Skyler SEDAR 46 Addendum

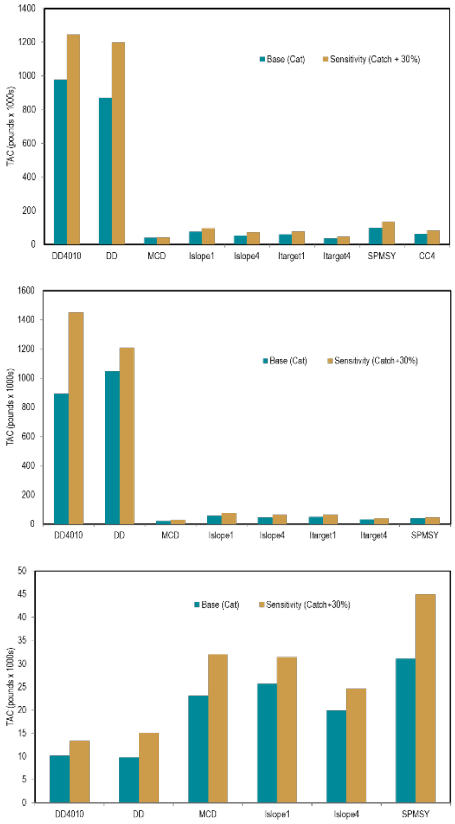

Figure 2. TAC sensitivity to +30% bias in catch. Top: STT Spiny Lobster, Middle: STT Spiny Lobster, Bottom: STX Stoplight Parrotfish. Source: Cummings and Skyler SEDAR 46 Addendum

4.2.4.4 CPUE

Catch rates used in the Management Procedures were not standardised. The standardisation attempted was not seen as successful, with only a few factors available with result that the standardised and unstandardised indices were essentially the same. The decision to use unstandardised data is therefore supported; however, more work should be undertaken on this process given that the *Islope* and similar Management Procedures tended to outperform others and these rely on the indices (although only over a short time series, so are likely to be less affected by standardisation issues). Furthermore, a sub-set of the fishers' data over the time series (that is known to be more reliable) can also be used for standardisation as a reference set.

Recommendation 9. Undertake more in-depth analyses on the standardisation of CPUE data.

Recommendation 10. Include sub-sampling of the data to create a more reliable reference set to compare with the base data CPUE set.

4.2.4.5 Length at full selection

The values of the length at full selection in the operating (LFS) was chosen as the 95th percentile of the (all years combined) cumulative length-frequency distribution for the main species and key gear combination. The plots of these show that the chosen lengths at full selection are often at the extreme end of the distribution with few animals of the population fully available to the gear. The norm is to either estimate the selectivity within a population model (probably only realistic in a data rich situation and therefore not applicable to these species-island units) or to use the mode (as

used in the mean length approach). Given that one is trying to encompass all uncertainties, to assume that a large part of the population is not available to fishing is not precautionary, given the unknowns. The 95th percentile and mode values of the length frequency distributions are provided in Table 2. The difference between them is important for species-island units such as Puerto Rico (PR) hogfish and yellowtail snapper, and STT triggerfish. The use of the mode would highlight changes when dome shaped selectivity relative to knife edged selectivity is tested – presently there is too little of the population available for full selection for the dome shaped test to be effective.

Table 2. The 95th percentile and mode values of the length frequency distributions Source: Cummings and Skyler SEDAR 46 Addendum

Species-island unit	SEDAR 46 AW LFS (95th Percentile) (MSE range)	SEDAR 46 RW LFS (Mode) (MSE range)
PR_Hog dive	**544 (2.68, 3.62)**	**280 (1.38, 1.86)**
PR_YT handline	**406 (1.11, 1.25)**	**280 (0.98, 1.33)**
STT_QT trap	386 (1.56, 2.11)	300 (1.21, 1.64)
STT_SL trap	133 (1.76, 2.38)	100 (1.32, 1.79)
STX_SL dive	120 (1.59, 2.15)	90 (1.19, 1.61)
STX_stop dive	338 (1.43, 1.94)	270 (1.15, 1.55)

Sensitivity tests of the choice of length at full selection were undertaken after the RW. These showed that it impacted on the choice of candidate Management Procedures where some were removed from the list and others were included (Table 3). For PR Hogfish, the delay difference model produced errors which should be further investigated as they may reflect an important cause. In both the PR cases, the performance measures are reasonably to very sensitive (e.g. PNOF for FMSYref) to this change. For STT Queen Triggerfish, the impact is less especially for the imprecise and biased Management Procedure list. In general, SPMSY seems quite sensitive to this assumption, and to a lesser extent EDCAC.

Table 3. Changes to the Candidate Management Procedure list after using the mode as the length at full selection. Derived from Source: Cummings and Skyler SEDAR 46 Addendum

Species-island unit	Precise-unbiased changes	Imprecise and biased changes
PR Hogfish	Errors: DD, DD4010 Drop: SPMSY, BK, SPR40_ML Add: YPR	DD, DD4010 SPMSY Ltarrget1
PR Yellowtail Snapper	Errors: None Drop: Fratio, IT10, ITM, SPMSY Add: None	None SPMSY None
STT Queen Triggerfish	Errors: None Drop: EDCAC, SPMSY Add: None	None SPMSY None
STT Spiny Lobster	Errors: None Drop: EDCAC, SPMSY Add: None	None DD, SPMSY DCAC, DCAC_40
STX Spiny Lobster	Errors: None Drop: SPMSY Add: CC4	None SPMSY None
STX Stoplight Parrotfish	Errors: None Drop: SPMSY Add: None	None SPMSY DD

Recommendation 11. It is recommended that the more precautionary modal length be chosen and that sensitivity to this Operating Model parameter be tested. This should include the dome shaped selectivity tests.

Recommendation 12. Further work should be undertaken to obtain the most appropriate selectivity function for the species-island units

4.2.4.6 Selectivity: PR Hogfish

Discussions during the workshop including input from industry highlighted that PR Hogfish selectivity may be dome shaped – a test that was not undertaken prior to the RW. Industry described that it was common practice to avoid catching larger animals as customers preferred plate-sized fish. MSE tests were undertaken by changing this assumption (Table 4). These results show only low levels of sensitivity to this assumption, except for the Management Procedure SPMSY, which falls out of the candidate list of Management Procedures with dome shaped selectivity. This test, however, should remain standard, given the information provided.

Table 4. Sensitivity to the assumption of dome-shaped selectivity for PR Hogfish for the base (left panel) and an alternative Operating Model assuming high-dome selectivity (right panel) for all MPs meeting the S46 DW/AW performance criteria. Results are shown for both the base observation model (precise, unbiased – top two tables) and also the alternative observation model (imprecise, biased – bottom two tables). Source: Cummings and Sagarese SEDAR 46 Addendum

15%LH, **asymptotic**, precise, unbiased				
MP	PNOF	B50	LTY	AAVY
FMSYref	95.2	98.6	100.0	100.0
EDCAC	57.9	96.9	97.4	58.4
MCD	79.0	98.2	96.6	75.8
Fratio	61.8	94.9	96.0	52.0
BK	79.0	95.1	93.5	59.2
DCAC4010	92.0	98.6	91.5	68.4
Islope1	55.6	82.2	83.0	96.2
Islope4	57.3	82.1	80.2	96.2
IT10	69.1	91.5	79.8	98.8
ITM	68.8	91.0	78.3	98.8
IT5	67.0	88.5	77.4	97.0
LstepCC1	59.4	83.3	74.2	96.2
LstepCC4	59.1	83.2	74.1	96.2
SPMSY	80.5	92.1	63.8	98.2
CC4	73.9	92.0	30.4	100.0
Itarget1	78.4	94.9	26.3	100.0
Ltarget4	92.6	97.5	2.4	99.8

15%LH, **high-dome**, precise, unbiased				
MP	PNOF	B50	LTY	VY
FMSYref	96.7	98.4	99.4	100.0
MCD	83.1	98.1	95.8	74.2
EDCAC	58.9	97.5	95.4	67.4
YPR	56.3	91.0	92.7	53.6
Fratio	58.2	92.7	92.1	56.4
DCAC4010	93.0	98.5	90.1	72.0
BK	78.3	95.5	89.7	56.8
Islope1	53.7	82.4	80.5	97.2
IT5	65.5	88.1	79.9	98.2
ITM	69.6	91.0	78.8	99.0
IT10	70.2	91.7	78.4	99.0
Islope4	56.2	82.2	76.9	97.0
LstepCC1	58.0	83.7	71.6	98.0
LstepCC4	57.9	83.7	70.8	97.6
Itarget1	72.4	93.0	31.8	100.0
CC4	70.9	90.1	28.7	100.0
Ltarget4	90.6	96.8	3.1	100.0

15%LH, **asymptotic**, imprecise, unbiased				
MP	PNOF	B50	LTY	VY
FMSYref	95.5	97.3	100.0	100.0
Islope1	55.2	76.8	71.2	93.4
Islope4	56.3	76.8	65.0	93.0
LstepCC4	57.9	78.1	61.8	95.2
LstepCC1	57.9	78.3	59.8	95.6
SPMSY	75.2	88.2	59.8	97.0
IT5	64.1	82.5	57.2	96.8
IT10	67.1	85.5	56.4	98.2
ITM	66.9	85.0	54.5	97.8
Itarget1	67.9	83.8	28.3	71.2
CC4	62.1	79.6	26.9	84.2
Ltarget4	77.9	89.2	17.2	65.6

15%LH, **high-dome**, imprecise, unbiased				
MP	PNOF	B50	LTY	VY
FMSYref	96.4	98.2	99.8	100.0
Islope1	57.3	79.7	63.9	91.4
IT5	62.3	83.3	62.3	96.6
Islope4	59.2	79.7	61.5	94.8
ITM	62.7	83.9	61.1	97.0
LstepCC4	61.3	81.7	59.8	96.8
LstepCC1	61.4	81.4	59.6	96.6
IT10	63.1	84.4	59.2	97.4
Itarget1	68.2	86.1	26.7	73.2
CC4	63.8	82.0	25.2	83.4
Ltarget4	79.3	91.3	16.3	64.0

4.3 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 the data-limited methods scientifically sound and robust? b) Are the methods appropriate given the available data? c) Are the data-limited models configured properly and used in a manner consistent with standard practices? d) Are the quantitative estimates produced reliable? Does the method produce management metrics (e.g. MSY, ABC, ACL) or other indicators (e.g. trends in F or Z, probability of overfishing) that may be used to inform managers about stock trends and conditions?

4.3.1 ToR conclusion

A) Are the data-limited methods scientifically sound and robust?

Yes, the DLMtool has been peer reviewed and its use for this purpose is appropriate.

Yes, the management procedure methods have also been peer reviewed.

B) Are the methods appropriate given the available data?

Yes, this is a data limited set of fisheries and this toolkit and Management Procedures have been created for this specific purpose.

C) Are the data-limited models configured properly and used in a manner consistent with standard practices?

Yes, there are no substantial issues, although there are future refinements suggested.

D) Are the quantitative estimates produced reliable? Does the method produce management metrics (e.g. MSY, ABC, ACL) or other indicators (e.g. trends in F or Z, probability of overfishing) that may be used to inform managers about stock trends and conditions?

Yes, within the context of data limited approaches, the assessors have available the necessary information from which to produce OFLs.

4.3.2 MSE Performance Measures

The toolkit includes a set of Performance Measures for comparing the different Management Procedures. The four Performance Measures are appropriate, being a) the probability of not overfishing ≥50% (PNOF), b) probability of being overfished (defined here as being below half of the biomass at Maximum Sustainable Yield (Bmsy) (B50)), c) a Performance Measure that captures the variability of the catches or the degree of catch stability – average annual variability in yield that has remained within 15% of the previous year's TAC ≥ 50% of the time (AAVY), and d) long term yield calculated from the last 10 years of the projections (LTY). Unlike many other MSEs elsewhere in the world, there are no target reference point Performance Measures. The overfished and overfishing Performance Measures would be seen as limit reference point Performance Measures, so define the outside OFL space. For example, a Performance Measure that is almost half the time below half Bmsy would still fall within the acceptable Management Procedure group given the cut off measures applied to the Performance Measures. This should be much more clearly highlighted.

Recommendation 13. Highlight that the initial cut off values for the performance measures would include management procedures that had almost 50% of the simulations result in overfishing or the stock being overfished.

Recommendation 14. Highlight that the catch from the Management Procedure is treated in the MSE as the final implemented ABC, yet tended to be set up as the OFL.

Good Management Procedure diagnostics were provided, including LTY, PNOF, AAVY and B50 trade-off plots, as well as value of information plots which highlight which Operating Model parameters most influenced LTY. This information, together with additional sensitivity tests (next ToR 3), were extremely helpful in reducing the full set of Management Procedures to a smaller set of candidate Management Procedures.

The report does not provide the actual mathematical formulae for calculating some of these Performance Measures. Providing them is good practice since there are several ways of calculating these, not all of them appropriate.

Recommendation 15. Future reports should provide the actual mathematical formulae for calculating the Performance Measures.

4.3.3 DLMtool Management Procedures

The data limited MSE tool developed by Carruthers and co-workers (Carruthers *et al.* 2014; Carruthers & Hordyk 2015) were applied to the six species-island units. The DLMtool is an R package that contains the normal MSE components of an operating model and data-limited management procedures. The MSE facilitates sensitivity testing of uncertainties in the data and life history parameters/assumptions. A sub-set of the available Management Procedures was utilised – these have been used elsewhere in the world. The toolkit and many of the data limited management procedures are published in the peer reviewed literature (Carruthers *et al.* 2014; Geromont & Butterworth 2014).

4.3.3.1 Selection of Management Procedures for the MSE tests

A subset of the DLMtool Management Procedures was tested for these cases – these were selected from a candidate set based on whether they apply, whether they provide a good reference set and how a more data-rich method would perform. The Management Procedures all require a different mix of data, thereby being variously sensitive to the species-specific pros and cons of the data and parameters. For example, the catch only methods require a catch series and information on depletion, whereas the index-based method mostly needs a recent index of abundance, in this case, CPUE. The mean length estimation method, modified to a Management Procedure (Huynh 2016), uses length and growth data. In addition, a reference Management Procedure that assumed perfect information was included. The selection of Management Procedures for testing in the MSE is supported.

Each MP has its own implied target or reference point, but these do not necessarily match each other. The OFL from the reference method (FMSYref) would be the "best" OFL that could be selected given the uncertainties in the Operating Models. A constant catch (CC4) method uses only recent historical catch data, but is set up to reduce the catch to 70% of the average of the most recent 5 years. The surplus production MSE (SPMSY) Management Procedure requires some data rich parameters of the Schaefer form surplus production parameters and, therefore, a catch and index of abundance series (or input depletion) and is based on an MSY assumption. The index-based Management Procedures used catch rate data either through using the slope of the recent catch rate (Islope1 and the more conservative Islope4) or the slope with an assigned target catch rate (Itarget1; the more conservative Itarget4; the differently formulated IT5 and ITM) to maintain constant CPUE or a target CPUE. The next set of Management Procedures were depletion based, i.e. needed some indication of what is the depletion or assumed a value (the Depletion-Corrected Average Catch – DCAC, DCAC_40 and DCAC4010). DCAC4010 has the additional assumption that a traditional 40:10 hockey stick harvest control rule applies. The abundance based methods have been used extensively in data limited applications and include the Fmsy/M ratio (Fratio), the Beddington

and Kirkwood life history method (BK) and Yield-per-Recruit Analysis (YPR). These all assume various Fmsy proxies. Finally, the data moderate approaches using the delay difference model (refs) (DD) and the DD with a 40:10 rule (DD4010) were applied even though these would not be applicable in the real world.

Recommendation 16. More emphasis should be given to the fact that each Management Procedure has implied target levels and these are not consistent across Management Procedures

It should be noted that the mathematical description of IT5 and ITM are not fully specified and this should be corrected. These should still be multiplied by a scalar e.g. catch.

Recommendation 17. Correct the formulae for IT5 and ITM in future reports. This does not affect the results as they are correctly implemented in the DLMtool.

Not all the Management Procedures tested can be used in real life, but the choice of Management Procedures was appropriate and highlighted the sensitivities to data and Management Procedure choice (and, therefore, underlying assumptions). As a result, the choice of these Management Procedures is supported.

4.3.3.2 Current Management Procedure

During the workshop, the existing Management Procedures ("CFMC_NoSD" with no standard deviation and "CFMC" with a standard deviation of 0.1) were tested for PR Yellowtail Snapper. Three additional Management procedures were tested, being a) median catch over the most recent 3 years (MCThree); (b) median catch over the most recent 10 years (MCTen); and (c) the third highest catch over the entire time series (THC). The current Management Procedure, whether with the SD version or not, does not conform to the cut-off rule of being above 50% PNOF nor is it above 50% B50 (Figure 3) for both the precise, unbiased and imprecise, biased tests. These results show that the current procedure is likely to place the resource at risk. In terms of PNOF, not one of the constant catch procedures are above the 50% PNOF cut-off and most are below the PB cut-off.

Recommendation 18. The current Management Procedure is likely to place the resource at biological risk. Further investigate this result for management purposes

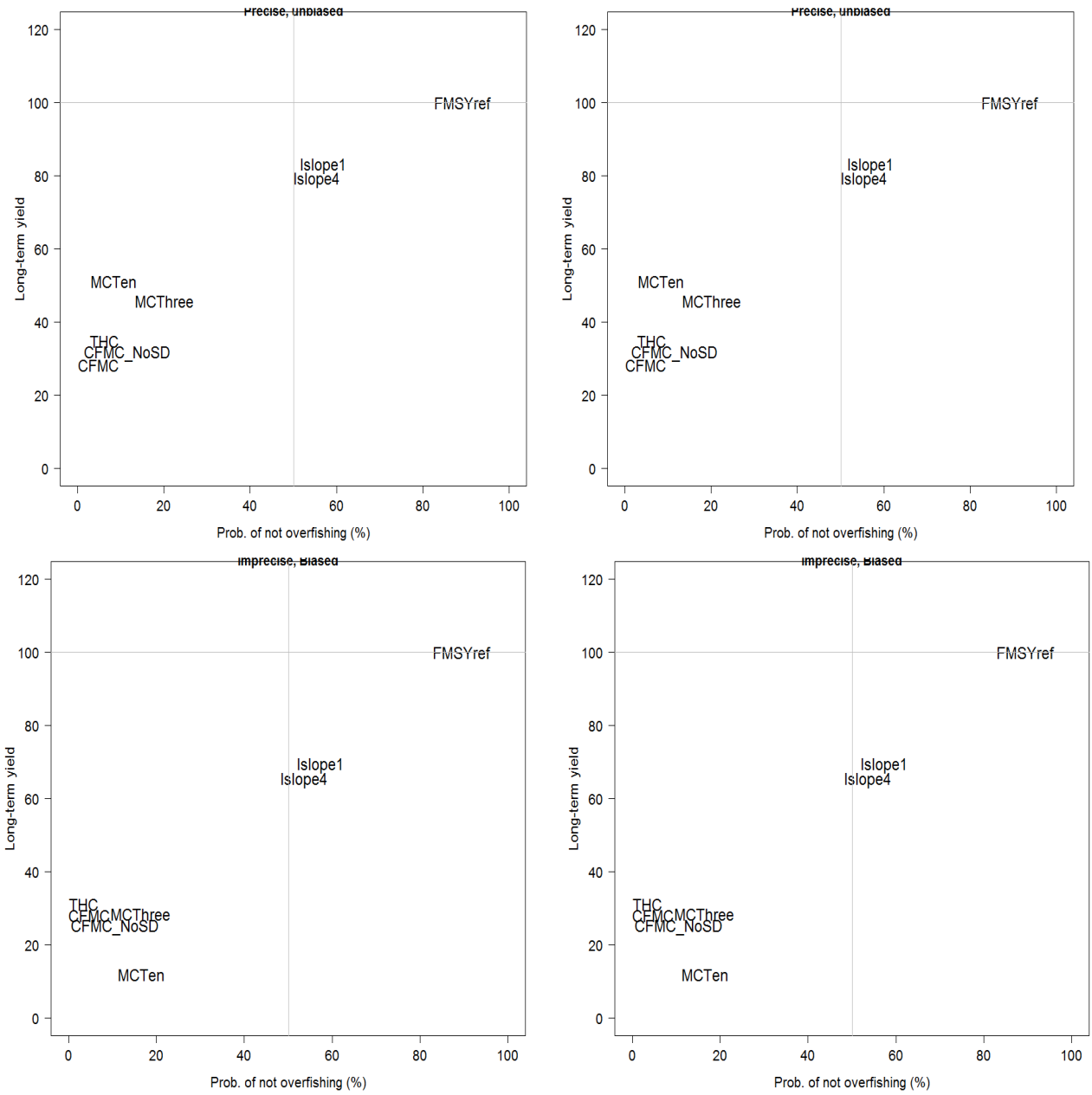

Figure 3. MSE results for PR Yellowtail Snapper for the current and related constant catch Management Procedures. Source: Cummings and Skyler SEDAR 46 Addendum

This result of the use of constant catch Management Procedures has also been shown to be poor in several other studies (Carruthers *et al.* 2014; Geromont & Butterworth 2014), i.e. these do not perform well relative to other Management Procedures and do not satisfy the Performance Measure criteria. The inclusion of the current Management Procedure should be a standard part of the tested Management Procedures.

Recommendation 19. Include the existing Management Procedure in the set of Management Procedures tested. These show how the existing Management Procedures compare to other candidate Management Procedures and also with respect to the chosen Performance Measures.

4.3.3.3 Short-term Performance Measures

Short-term transitional Performance Measures were not tested prior to the workshop. An additional Performance Measure, Short Term Yield (STY) was added after the workshop. Often, there are large transitional characteristics of implementing a Management Procedure that are important considerations in the selection of the candidate Management Procedures. For PR Hogfish (Table 5), for example, the STY using DD4010 is much lower than DD despite the DD4010 ultimately achieving the higher LTY compared to DD. None of the Management Procedures outperform the FMSYref Management Procedure in terms of STY.

Recommendation 20. Short Term Yield and other transitional Performance Measures should be further developed and used for future MSEs.

Generally, there is no clear relationship between LTY and STY relative to the FMSYref Management Procedure (Figure 4). Although the range of STY values are higher compared to LTY as expected.

Table 5. Performance metrics for probability of not overfishing (PNOF), long term yield (during last 10 years of the 40 year projection = LTY), and short term yield (STY; during first 10 years of the 40 year projection) for the six species-island units evaluated in S46. All performance statistics are relative to the performance of the FMSYref MP. Source: Cummings and Sagarese SEDAR 46 Addendum

Puerto Rico hogfish				Puerto Rico yellowtail snapper				St. Thomas queen triggerfish			
MP	PNOF	LTY	STY	MP	PNOF	LTY	STY	MP	PNOF	LTY	STY
DD4010	93.2	99.0	51.2	DD4010	75.3	97.5	77.3	DD	80.6	95.9	61.8
DD	76.8	98.9	75.0	DD	55.7	96.7	90.7	DD4010	93.8	92.3	48.7
MCD	79.0	96.8	68.2	MCD	71.6	94.0	83.4	MCD	79.5	91.2	58.3
Fratio	61.8	95.3	66.3	Fratio	59.5	93.9	79.2	Fratio	59.2	89.5	62.3
BK	79.0	93.6	54.0	Islope1	60.9	81.9	89.5	Islope1	58.0	88.9	64.4
Islope1	55.6	83.6	83.5	Islope4	61.1	79.9	89.1	Islope4	60.5	84.0	75.1
Islope4	57.3	80.9	82.5	SPMSY	72.6	60.9	72.0	Itarget1	52.9	78.8	73.8
SPMSY	80.5	63.6	69.1	CC4	77.6	34.2	82.6	CC4	58.2	65.1	66.6
CC4	73.9	31.7	73.1	Itarget1	87.3	25.3	77.9	SPMSY	80.1	58.7	64.7
Itarget1	78.4	29.2	72.4								

St. Thomas spiny lobster				St. Croix spiny lobster				St. Croix stoplight parrotfish			
MP	PNOF	LTY	STY	MP	PNOF	LTY	STY	MP	PNOF	LTY	STY
Fratio	56.1	84.9	71.6	DD	66.9	84.4	66.8	DD	89.0	85.9	35.9
MCD	66.5	83.7	65.9	MCD	66.3	84.3	64.9	MCD	82.3	79.8	35.7
DD	67.7	83.4	68.1	DD4010	75.9	80.0	56.7	Fratio	57.8	79.5	47.1
DD4010	77.1	80.3	56.8	Islope1	67.9	66.8	64.0	DD4010	96.7	78.3	30.6
Islope1	65.6	67.1	65.2	Islope4	68.7	63.2	63.8	Islope1	59.8	73.0	50.6
Islope4	66.3	62.4	64.9	CC4	54.7	60.3	60.4	Islope4	64.2	62.4	48.5
Itarget1	60.1	57.8	58.4	Itarget1	59.9	58.9	57.3	SPMSY	81.0	34.2	33.8
CC4	56.1	56.6	61.6	SPMSY	66.8	50.4	58.6	Itarget4	97.9	0.0	3.4
SPMSY	67.5	46.9	59.5	Itarget4	98.8	0.0	16.1				
Itarget4	98.8	0.2	16.2								

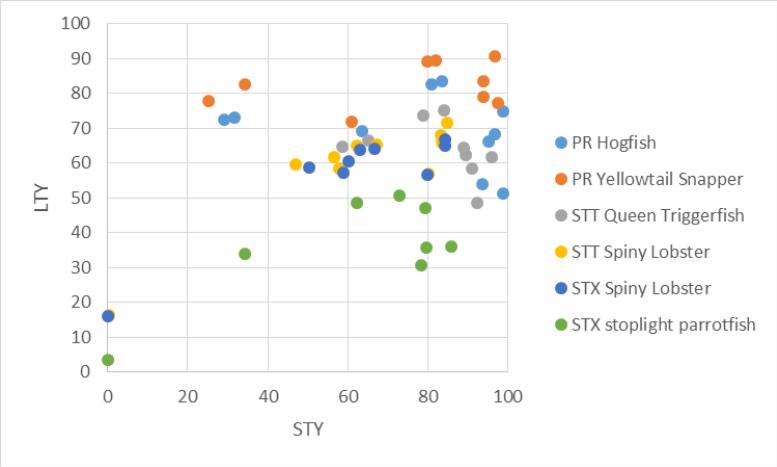

Figure 4. Long Term Yield (LTY) versus Short Term Yield (STY) for the six species-island units. Based on Source: Cummings and Skyler SEDAR 46 Addendum

4.3.4 Mean length Management Procedure

Mean length estimators (Hoenig 1983; Ault & Ehrhardt 1991; Ault, Smith & Bohnsack 2005; Gedamke & Hoenig 2006; Ault *et al.* 2008) have been used extensively especially for tropical fisheries. These approaches have also been both reviewed in CIE processes and within the journal review process. For the most part the applied method for the selected species-island units is based on that of Gedamke and Hoenig (2006) and was not an integral part of the DLMtool. The time dynamic (non-equilibrium) version of the mean length estimator was applied, i.e. it successively estimates changes in fishing mortality in the time series assuming up to three changes in mortality over time. During the workshop the present mean length form was added to the toolkit as a draft and it would be a good addition if this becomes standard for the freely available toolkit. AIC was used to select between the different iterations with the more complex model only selected if the reduction in AIC was >4 units. The method assumes continuous and constant recruitment, deterministic growth and stepwise changes in total mortality. It uses several proxies for *Fmsy* such as $F_{0.1}$, $F_{SPR=30\%}$ and $F_{SPR=40\%}$ - all internationally recognised proxies. Like the other MSE tests, the tests used the precise and unbiased results. The same Performance Measures were also used in the tests. The conclusions drawn from the mean length estimator are robust except for STT queen triggerfish as discussed below.

The modal lengths from length histograms were used to derive the value of the length at full selection, *Lc*. Knife edged selectivity was assumed above *Lc*. This is different from the values used in the DLMtool MSE which was a much higher value, derived from 95th percentile. The use of the modal length is much more appropriate.

TACs were calculated for the six species-island units using two reference points, $F_{0.1}$ (the fishing mortality rate at which the slope of the yield per recruit curve is 10% of that at the origin) and $F_{SPR\%}$ (the fishing mortality rate that reduces the spawning biomass-per-recruit to a certain set percentage from the unfished population). These are reasonable reference points as they are nationally and internationally used.

4.3.4.1 Performance Measures, including Short-Term Yield

The final set of Performance Metrics of the Management Procedures from the MSEs of the six Caribbean species-island units including both short and long-term measures are provided in Table 6, being: the probability of not overfishing (PNOF), the probability of biomass above half B_{MSY} (B50), the probability of achieving long term yield (LTY) and short term yield (STY), and the probability of

annual variability in yield to remain within 15% (AAVY). Base stock and fleet dynamics were considered with unbiased and biased observation dynamics. These show that the $F_{0.1}$ reference point for the biased set approaches perform well in terms of achieving the PNOF, B50 and AAVY cut-offs for all species-island units, whereas this was not the case for the $F_{30\%}$ reference point. The unbiased results are more variable with respect to $F_{0.1}$.

Table 6. Performance Measures of the mean length Management Procedure adding Short-Term Yield (STY). Source Huynh SEDAR 46 Addendum

	Unbiased Observation					Biased Observation				
MP	PNOF	B50	LTY	STY	AAVY	PNOF	B50	LTY	STY	AAVY
Yellowtail snapper										
FMSYref	91	98	100	79	100	90	98	100	79	100
YPR_ML*	54	73	60	45	95	71	81	38	36	94
SPR30_ML	15	41	53	72	96	40	55	42	59	94
SPR40_ML*	40	62	57	56	94	59	71	39	45	96
Hogfish										
FMSYref	96	96	100	58	100	96	97	100	60	100
YPR_ML*	70	84	63	13	78	79	87	34	16	85
SPR30_ML	24	48	42	51	82	44	59	37	42	87
SPR40_ML*	49	70	58	29	77	62	74	36	29	86
Queen triggerfish										
FMSYref	93	97	96	66	100	94	98	95	67	100
YPR_ML*	68	83	57	20	92	75	87	32	20	94
SPR30_ML	23	50	52	52	96	38	59	45	46	97
SPR40_ML*	46	68	62	36	93	57	74	39	35	96
Spiny lobster STT										
FMSYref	73	94	85	65	100	70	93	85	62	100
YPR_ML*	25	56	32	59	99	52	72	25	41	97
SPR30_ML	2	26	12	68	91	12	36	16	63	93
SPR40_ML	7	36	19	68	93	26	48	18	57	95
Spiny lobster STX										
FMSYref	72	91	79	52	100	72	90	78	54	100
YPR_ML*	35	63	35	48	99	55	74	27	36	99
SPR30_ML	3	35	19	62	96	15	44	23	57	96
SPR40_ML	9	43	25	60	97	29	55	26	50	97
Stoplight parrotfish										
FMSYref	86	97	97	71	100	86	96	97	73	100
YPR_ML*	52	72	60	36	96	67	81	52	29	98
SPR30_ML	14	40	59	62	98	26	49	73	56	98
SPR40_ML	31	56	64	51	97	47	66	64	43	98

* Indicates the MPs which met management criteria (PNOF > 50%, B50 > 50%, and AAVY > 50%) from the MSEs using the biased observations.

The TAC ranges from some these candidate Management Procedures are quite wide, especially for the SPR40_ML Management Procedure. They are more precise for Queen Triggerfish and Stoplight Parrotfish.

Table 7. Distribution summaries of the TAC for the candidate Management Procedures taken from the biased and imprecise observations set for the six species-island units. Source Huynh SEDAR 46 Addendum

	Quantile (x 1000 pounds)				
MP	Min	25%	Median	75%	Max
Yellowtail snapper					
YPR_ML	31.9	109.1	166.0	241.2	734.1
SPR40_ML	29.6	114.1	176.2	293.7	1506.3
Hogfish					
YPR_ML	4.2	26.1	40.2	72.8	890.7
SPR40_ML	9.1	31.2	50.7	86.9	1575.2
Queen triggerfish					
YPR_ML	4.7	13.0	18.9	30.2	189.3
SPR40_ML	6.3	18.0	30.2	56.1	1611.0
STT Spiny lobster					
YPR_ML	14.2	45.5	69.4	111.5	3726.5
STX Spiny lobster					
YPR_ML	0.9	9.8	17.5	32.6	2607.2
Stoplight parrotfish					
YPR_ML	1.0	3.8	5.6	8.2	27.7

4.3.4.2 Protogyny: PR Hogfish

The mean length Performance Measure is the only approach that can be modified to include protogyny in hogfish. An additional sensitivity test was undertaken during the review workshop where the approach was updated to include sex change in Hogfish. The mean of two depth-specific sex ratio ogives was used in the analysis since the Trip Interview Program database indicated these animals are caught at both depths. The Reference points (Table 8) are affected by this assumption, especially if only those for females are considered. Brooks *et al.* (2008) simulation tested the impact of protogyny on different reference points. The results were sensitive to life history parameters and other assumptions such as selectivity, but found that for moderate values of the steepness of the fertilisation function, it is best to calculate biological reference points of both sexes as a default, especially when the degree of sperm limitation is not known. The situation is complex in that Brooks *et al.* (2008) also suggested the use of female-specific reference points if fertilization rates are high, but at low fertilization rates, male reference points performed better when the timing of sex change is static. While this information remains less known for this species, the assessment team's suggested use of both sexes for mean length reference points are supported.

Recommendation 21. Use both sexes for mean length Management Procedure reference points as a default until more information on fertilisation rates and sex change timing is known.

Table 8. Spawning Potential Ratio (SPR) reference points for PR Hogfish from the spawning stock biomass of both sexes, males only, and females only. Source Huynh SEDAR 46 Addendum

Reference point	Both	Males	Females
F30%	0.17	0.12	0.55
F40%	0.12	0.09	0.37

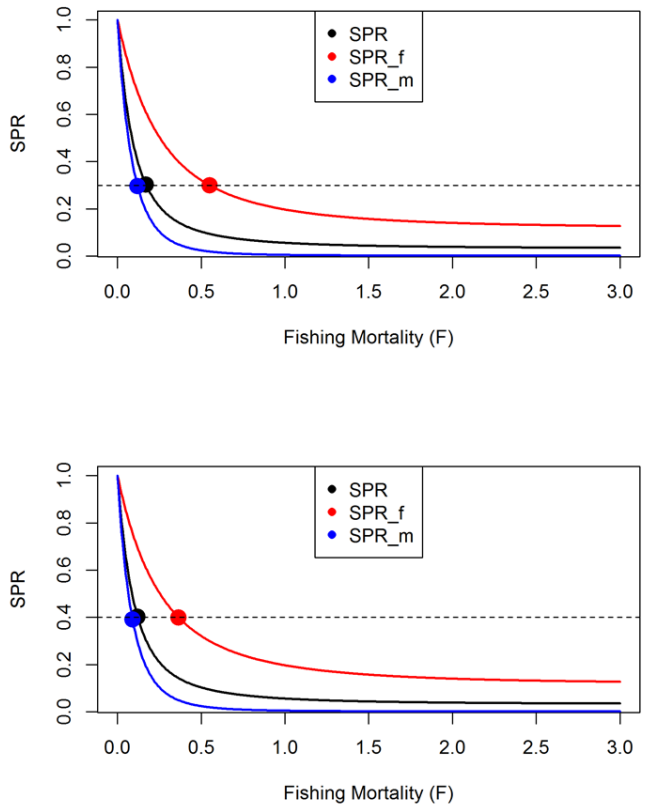

Figure 5. The Spawning Potential Ratio (SPR) curves for PR Hogfish against fishing mortality for both sexes (black), females only (red) and males only (blue). Top figure uses the SPR=30% and bottom SPR=40% reference points.

4.3.4.3 Dome-shaped selectivity versus growth: STT Queen Triggerfish

Sensitivity to the length at full selection was also tested prior to the workshop. If the total mortality (*Z*) is sensitive to the parameter it shows that selectivity may be dome shaped or the asymptotic length (L_∞) is miss-specified. This was the case for STT queen trigger fish where the *Z* values ranged from 1.15 to 1.70 as the Lc values were changed. During the workshop, these tests for queen triggerfish were expanded to include L_∞ so as to identify whether the issue is dome shaped selectivity or an overestimate of L_∞. The new results demonstrated that the total mortality values were sensitive to both values.

Estimated mortality rates F_{recent} and ratios of $F_{0.1}$ and F_{recent} in instantaneous and annual exploitation rates for St. Thomas queen triggerfish from a grid of values for L_c and L_∞ are provided in Table 9. The results show that there is sensitivity in the F_{recent} values to both L_c and L_∞, especially for L_∞ values larger than 415mm. This suggests dome-shaped selectivity. This sensitivity continues for F_{ratio} and μ_{ratio}, but is more subtle. The OFL is more robust to changes in L_c for a fixed L_∞, than changes to L_∞ for a fixed L_c. This means that dome-shaped selectivity is less of an issue (but

potentially still important) for the mean length Management Procedure for STT Queen Triggerfish, than growth rate parameters. As a result, a priority would be to undertake a growth study for STT queen triggerfish.

Recommendation 22. Undertake a growth study for STT queen triggerfish as the mean length estimator has highlighted a strong sensitivity to these parameters.

Table 9. Estimated recent fishing mortality rates, F_{recent}, F_{ratio} ($F_{0.1}/F_{recent}$) and μ_{ratio} (the annual exploitation rate equivalent of F_{ratio}). Source Huynh SEDAR 46 Addendum

	Linf			
Lc	**415**	**500**	**605.3**	**700**
	F_{recent}			
280	0.03	0.35	0.74	1.09
300	0.04	0.42	0.89	1.31
320	0.04	0.50	1.08	1.59
340	0.02	0.58	1.27	1.89
360	-0.04	0.62	1.44	2.18
	F_{ratio}			
280	11.67	0.83	0.31	0.17
300	10.25	0.69	0.28	0.18
320	11.25	0.70	0.27	0.14
340	25.50	0.60	0.23	0.12
360	-	0.66	0.20	0.11
	u_{ratio}			
280	9.99	0.85	0.39	0.25
300	8.58	0.73	0.38	0.28
320	9.24	0.75	0.38	0.26
340	20.18	0.67	0.35	0.24
360	-	0.73	0.33	0.23

4.3.5 Calculation of TACs using real world data

In addition to the MSE tests, the Management Procedures were tested using real world data. This means that only Management Procedures that could be undertaken in the real world would be highlighted here. These were appropriately implemented. Sensitivity tests are discussed in ToR 3.

4.4 ToR 3. 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.

4.4.1 ToR conclusion

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.

An extensive amount of sensitivity tests was undertaken. These sampled a large range of variance and bias. Uncertainty in the parameters was also addressed, especially as a part of the Operating Model. These have been undertaken, to a large extent, appropriately to reflect the extent of the uncertainty. They also capture most of the sources of uncertainty. These tests were a significant contribution to narrowing down the selection of candidate Management Procedures for OFL use. Some suggested changes to the DLMtool with regard to how uncertainty is handled are discussed below. In addition, changes to how the biased and imprecise sensitivity tests should be approached in the future are described.

Only convergence tests to the Operating Model were undertaken. Time series analyses were investigated during and after the RW, showing their importance to Management Procedure selection and MSE performance.

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

The implications of the uncertainty are clearly stated and carried into the Management Procedure candidate list.

4.4.2 DLMtool operating model – uncertainty

An issue for this review was that the operating model specifications were not described in the workshop papers and were an additional appendix to the Carruthers *et al.* (2014) paper which was in an obscure location. Although it was available after a search, the importance of the operating model set up was underplayed in the work.

Recommendation 23. Future iterations of this toolkit for these applications or similar should emphasise the operating model set up and its specifications.

4.4.2.1 Variance

Unlike data rich MSEs, where the Operating Models are highly conditioned (i.e. they are similar to an assessment model), the DLMtool Operating Model is provided with extensive input parameters and their associated uncertainty to simulate an age-based population. The Operating Model samples parameter settings where upper and lower bounds have been provided and assumes a uniform distribution, i.e. they are uninformative priors. Also, these priors are assumed to act independently of each other; an aspect which is discussed further below. The toolkit does not include certain aspects such as time- and age-varying mortality, and hermaphroditism. Despite these potential weaknesses, the toolkit is an appropriate tool to apply to these data limited applications.

Recommendation 24. The DLMtool is an appropriate tool to test data limited management procedures and should, therefore, continue to be applied to these species-island units until more data rich methods become available.

The input variance was assumed to capture reasonable ranges of life history stages and fishery dynamics. However, a characteristic of the toolkit is the way it samples uncertainty variables – these are added for most parameters and it samples from these without consideration of the inherent correlation structure between several of these parameters. For example, it is well known that several of the growth and natural mortality parameters are correlated, indeed natural mortality can be calculated from maximum age (Hoenig 1983; Hewitt & Hoenig 2005) or growth parameters (Pauly 1980). This means that a much larger range of uncertainty was sampled than was necessary for the MSE and that, in some instances, a set of parameters would describe an animal that does not exist. There are two consequences of this set up: a) the model takes longer to run than it needs to, and b) the uncertainty is likely to be over-estimated. The former is an issue as this influences the number of sensitivity tests that can be undertaken when time is limited, whereas the latter is less of an issue since one tends to prefer methods that over-estimate uncertainty in data-limited situations.

However, this is an important aspect of the DLMtool that requires further refinements. Despite this weakness of the DLMtool, its use and application to these species-island units have been appropriate.

Recommendation 25. Work with the DLMtool authors to better sample the available population parameters considering their inherent correlation structure.

All the distributions assume a uniform distribution, but in some cases there may be more information available than is assumed. Some flexibility on how variance, bias, and correlates are added in the toolkit would be preferable. Of course, the toolkit is developed by another group and additions to the toolkit can only be suggested by the users.

Recommendation 26. In the future, the Operating Model should be a key focus, where the parameter correlations, biases and gradients selected can be considered.

The impact of correlations on the growth parameters was tested for the mean length Management Procedure and is discussed below.

4.4.2.2 Mean length sensitivity tests to correlation in growth parameters

Sensitivity analysis of life history parameters to correlation in the von Bertalanffy growth parameters was also undertaken for PR Yellowtail Snapper and Hogfish using a Monte Carlo procedure. A variance-covariance matrix was calculated from the CVs from the life history studies. Scatterplot matrices are shown using the $F_{0.1}$ and $F_{40\%}$ reference points (Figure 6 to Figure 9). These show that the reference points are less sensitive to correlated growth parameters, whereas those for F_{recent} and F_{ratio} (and therefore also the estimator) are sensitive.

Recommendation 27. Undertake further analyses of the effect of correlated parameters on the mean length estimator.

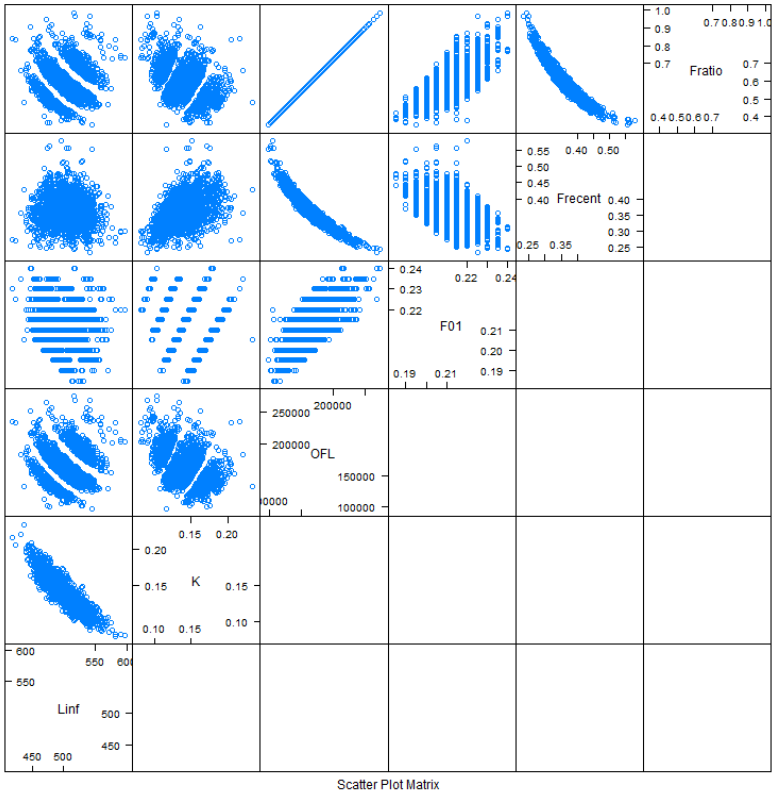

Figure 6. Scatterplot matrix of the sensitivity of the OFL to L_∞ and K for PR Yellowtail Snapper using $F_{0.1}$ as the reference point. Source Huynh SEDAR 46 Addendum

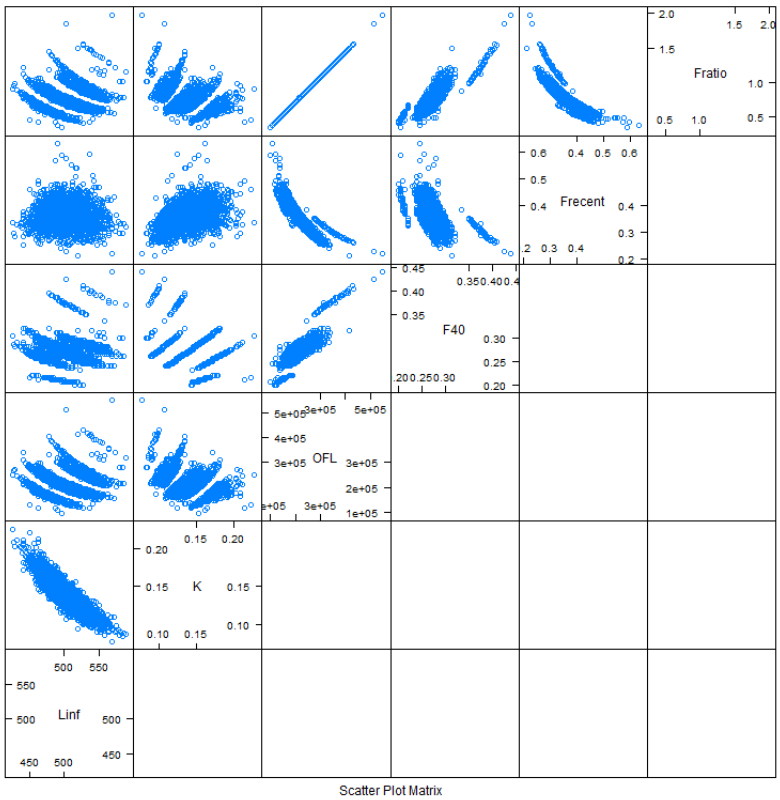

Figure 7. Scatterplot matrix of the sensitivity of the OFL to L_∞ and K for PR Yellowtail Snapper using $F_{40\%}$ as the reference point. Source Huynh SEDAR 46 Addendum

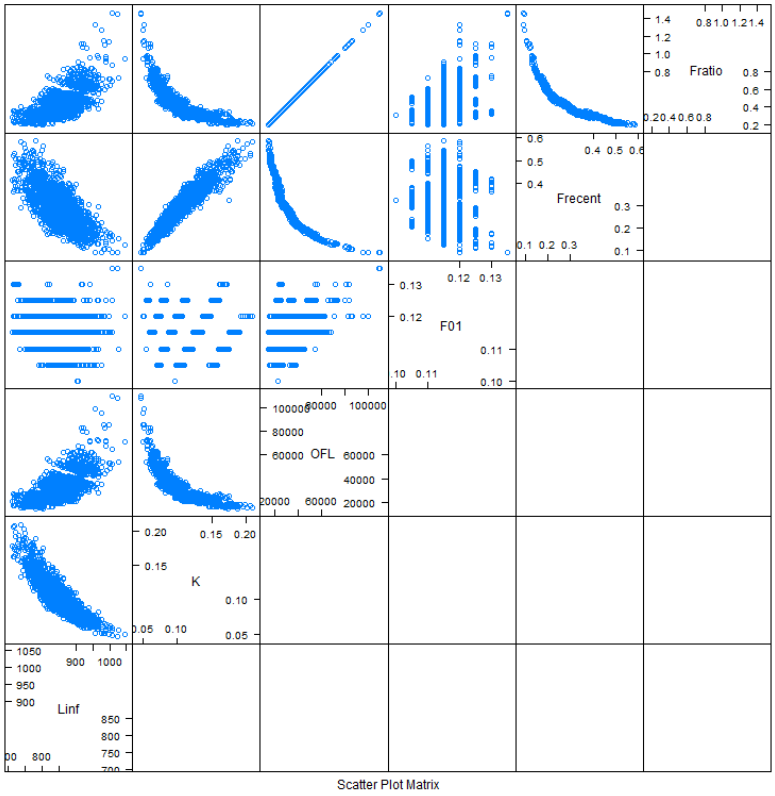

Figure 8. Scatterplot matrix of the sensitivity of the OFL to L_∞ and K for PR Hogfish using $F_{0.1}$ as the reference point. Source Huynh SEDAR 46 Addendum

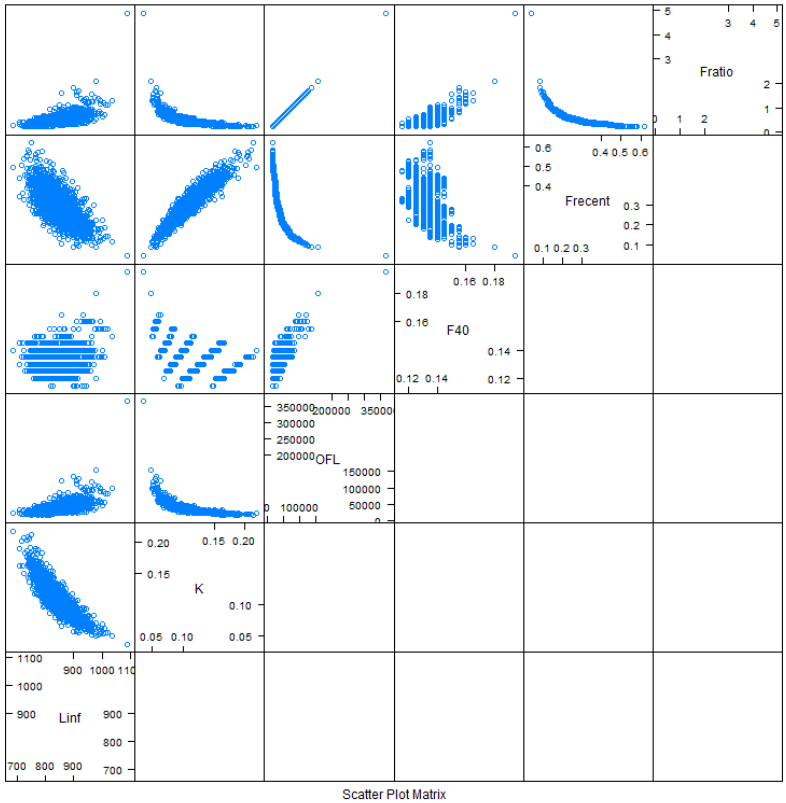

Figure 9. Scatterplot matrix of the sensitivity of the OFL to L_∞ and K for PR Hogfish using $F_{40\%}$ as the reference point. Source Huynh SEDAR 46 Addendum

When the OFL estimation procedure in the DLMtool was included in the correlated stochasticity in the growth parameters while all else was kept at the default of independent uncertainty, the OFL mostly showed narrower confidence intervals (as one would expect) (Table 10).

These analyses support the stated need to include these correlations in the DLMtool and also support the view that the range of candidate Performance Measures is likely to change.

Recommendation 28. Undertake further analyses of the effect of correlated parameters on the mean length estimator outside and within the DLMtool.

Table 10. Summary of the OFL distribution for PR Yellowtail Snapper and Hogfish using independent or correlated (-0.9) growth parameters. Source Huynh SEDAR 46 Addendum

	Quantile (x 1000 pounds)				
	Min	25%	Median	75%	Max
Yellowtail snapper					
F0.1					
Independent Linf/K	31.9	109.1	166.0	241.2	734.1
Correlated Linf/K	30.6	113.3	153.2	218.7	595.4
F40%					
Independent Linf/K	29.6	114.1	176.2	293.7	1506.3
Correlated Linf/K	29.1	136.8	196.7	279.4	891.8
Hogfish					
F0.1					
Independent Linf/K	4.2	26.1	40.2	72.8	890.7
Correlated Linf/K	8.4	30.1	43.1	62.5	1884.7
F40%					
Independent Linf/K	9.1	31.2	50.7	86.9	1575.2
Correlated Linf/K	13.8	36.7	51.0	83.8	717.5

4.4.2.3 Bias

The choice of biased values is incongruous (they are all in the same direction), which means that the bias and imprecise sensitivity tests are not appropriately undertaken.

Recommendation 29. Review the bias values input in the MSE.

4.4.2.4 MSE Diagnostics: time series results

Only one diagnostic was tested prior to the workshop – whether the model had converged, i.e. that enough replicates and simulations were undertaken for the Performance Measures that were developed prior to the workshop. For all species-island units, these clearly had converged. There may actually be a case for undertaking fewer simulations although this is not essential. This choice of sample sizes is supported.

However, no further Operating Model tests were undertaken prior to the review. Time series of the data should be more closely investigated than was undertaken prior to the RW. On request from the panel, time series plots of median and percentile envelopes, and also individual simulation runs were provided. Both are needed, the one set emphasises the overall trend whereas the other highlights the individual journey, which more closely simulates actual events.

An MSE run for only one simulation and one repetition for the PR Yellowtail Snapper was undertaken post RW for the candidate Management Procedures (Figure 10). These highlight that the execution of CC4 tends to produce a sequence of declining TACs. This result is not surprising given

the Management Procedure formula and that it was originally a constant catch approach. This Management Procedure should be implemented with caution over a long period. These results also show that the observation error around the catch in the MSE is quite high. The Fratio and MCD particularly highlight very large changes to TACs (as also reflected in the AAVY Performance Measure).

Recommendation 30. Undertake single simulation and repetition runs to investigate Management Procedure performance over time.

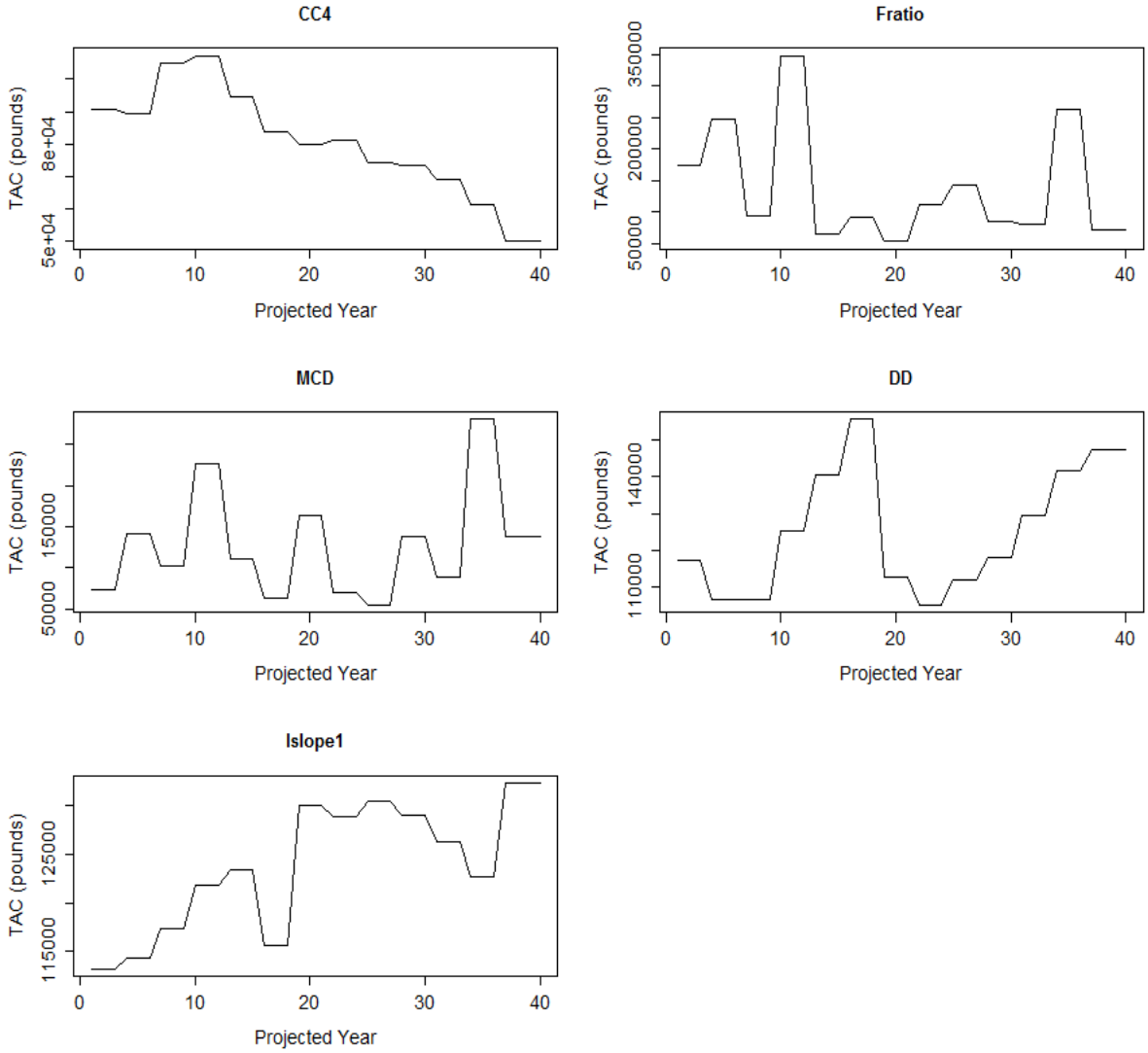

Figure 10. MSE run for only 1 simulation and 1 repetition for the PR Yellowtail Snapper showing TAC over the projected time period for the candidate Management Procedures. Source Cummings and Sagarese SEDAR 46 Addendum

In addition, the time series plots of the MSEs using all the simulations and repetitions were requested and provided during and after the RW. These time series again highlight the jagged Fratio and MCD performance over time, whereas Islope has much smoother TAC transitions. Although some of the runs in CC4 are different, the general trend is a steady downward decline in TACs over time.

Recommendation 31. Investigate time series plots of operating and management model outputs, highlighting medians and percentile envelopes, and also, individual simulation runs.

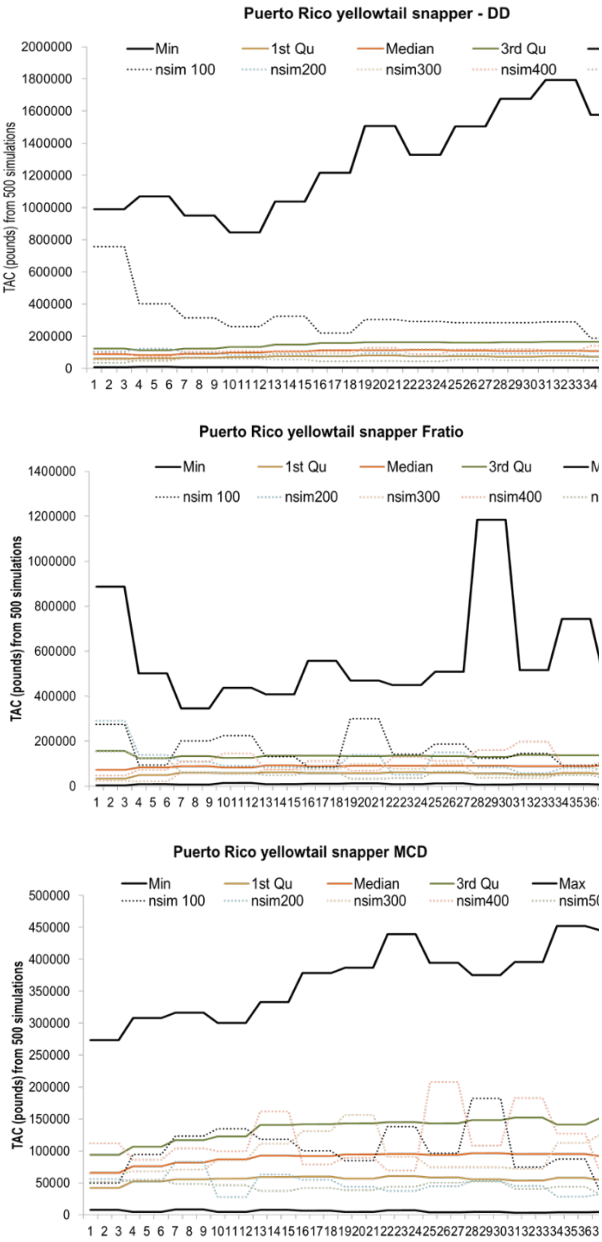

Figure 11. TAC statistics from within the MSE for selected MPs (DD, Fratio, MCD, Islope1, and CC4) for PR Yellowtail Snapper. These five MPs were selected for this examination as they frequently met the performance criteria across all species-island units. Summary statistics (minimum, 25% percentile [1st Qu], median, 75th percentile [3rd Qu], maximum) are provided for the TACs (y axis) by MP along with 5 individual simulations within each MP (i.e., the 100th [nsim 100], 200th [nsim200], 300th nsim300], 400th [nsim400], and 500th simulations [nsim500]) for the 40 year projection period (x axis). Source Cummings and Sagarese SEDAR 46 Addendum

Recommendation 32. Undertake a dedicated workshop on the Operating Model parameters and how these are set up. Allow for correlations between the parameters and focus on the input biases so that the bias and imprecise Operating Model can be used as the base Operating Model rather than the alternative.

The time series results do highlight that the projections are not properly constrained. Some of the projections appear to be highly unlikely through, for example, producing catches well above those previously observed. Here Management Procedure performance is difficult to assess since these robustness tests are in some projections excessive. Several options are available, e.g. rejection probabilities or joint probability densities. These issues may be related to the unsophisticated manner in which the DLMtool addresses uncertainty as a whole.

Recommendation 33. Projections need to be more constrained to reflect possible scenarios through better conditioning of the operating model on the available past observations, adding a rejection probability to outcomes or improving the joint parameter probability densities.

4.4.3 Operating Model sensitivity tests

Extensive sensitivity tests were undertaken. Those pertaining to the Operating Model are:

- Life history parameter uncertainty and biases (15% or 5% CV; biased and unbiased)
- Selectivity (high or moderate dome, or asymptotic)
- Data input to the Management Procedure uncertainty and biases (precise and imprecise; biased or unbiased)

The choice of these tests are supported.

In all the species-island units, the base and alternative(s) Operating Models that address specific sensitivities within the underlying population structure were tested. However, the base observation was that the data provided by the observation module within the Operating Model were precise and unbiased. Much greater emphasis should be placed on the biased and imprecise set and should in the future be the base rather than alternative set. Sensitivity tests and choice of candidate Performance Measures concentrated on unbiased and precise Operating Model settings, which is not supported. Given the extent of the information limitations, an assumption would be that the information is biased and imprecise.

Recommendation 34. The base set should assume that the biased and imprecise setting is the norm, given the data limited nature of the fisheries selected

Recommendation 35. The biases should be selected with greater caution than a fairly standard increase in CV for all key parameters and data without consideration of correlations in parameters.

Value of Information plots in the Operating Model plots are provided. These identify the parameters in the Operating Model that are correlated with utility. These show the value of perfect information and contribute to emphasising information that is most beneficial to know well.

For all the Performance Measures, only the average values are shown. The norm is to also show the upper and lower percentiles which give the uncertainty around these Performance Measures.

Recommendation 36. Provide percentiles for the Performance Measures in the sensitivity tests.

4.4.4 Real world data TACs

The distributions of the OFLs for each Performance Measure using real world data are provided. These are an important source of information for removing or selecting candidate Management Procedures and its continued use for this purpose is supported. In most cases, the delay difference Management Procedures showed great OFL uncertainty.

Sensitivity tests were undertaken of some real world parameters relevant to the Management Procedure. The figures are informative, but should include the Management Procedures that are not affected by the tested parameter for completeness, i.e. the sensitivity tests highlight sensitivity but not insensitivity which should also be included.

Recommendation 37. The figures showing the results of sensitivity tests to assumptions (e.g. Figure 3.3.1.8) should also show the Management Procedures that do not use these assumptions, highlighting their insensitivity to these assumptions.

Further sensitivity tests were undertaken on the real world data cases to assess sensitivity of the TAC calculations to data input values and their associated CVs. Some combinations e.g. Fratio and CV cat are also provided. These again highlight the importance of varying input parameters and are useful to select the final candidate list of Management Procedures and also the value of obtaining better information in the future.

4.4.5 Implementation

4.4.5.1 OFL versus ABC

The MSE is implementing the value as the final TAC set in the process. The results, therefore, assume that no further uncertainties were added to the catch calculated from the Management Procedure and this should be noted. In other words, the Management Procedure was assumed to deliver the ABC/TAC, yet was set up as an OFL. Usually, the OFL would have additional uncertainties included to set the final ABC.

Recommendation 38. Highlight that the implementation of the Management Procedures treated the TACs as the final ABC.

4.4.5.2 Multi-year TAC sensitivity

In all tests, the Management Procedures were implemented every 3 years. Sensitivity to this time period should be undertaken as this assumption can have a substantial impact on the results. It can be undertaken in two separate ways, both have their different merits. The first is to use the Management Procedure every 3 years, but to include that this is not implemented correctly. This scenario could be modelled by adding implementation uncertainty to the MSE (of course this option is presently unavailable in the DLMtool). The second option is to modify the existing Performance Measures to test different constant periods, which are accurately implemented.

Recommendation 39. Test the sensitivity of the Management Procedures to different TAC change periods, both through implementation uncertainty and using different values of input periods.

4.4.5.3 Implementation uncertainty

Presently, the DLMtool does not include implementation uncertainty. This inclusion of the effect of human behaviour and management systems on the implementation of a TAC is a crucial component

of evaluating management systems (Fulton *et al.* 2011) and it would be appropriate for this component to be added to the toolkit, if possible.

Recommendation 40. Suggest that implementation uncertainty be added to the DLMtool.

4.4.6 Final selection of candidate Management Procedures for OFL calculation

In the DW/AW report, the usefulness of the different sensitivity tests to narrow down the candidate Management Procedures is underplayed. During the review workshop, further work was undertaken on this aspect and an approach to selecting a smaller set of candidate Management Procedures was provided. The final findings are appropriate.

The assessment team as an Addendum provided the process of producing the final list of candidate Management Procedures and which Management Procedures can be used for management for the precise and unbiased (Table 11) and the imprecise and unbiased observation error model (Table 12). It should be noted that these latter do not include the current Management Procedure as it is not precautionary and does not reach the cut-offs applied. Some additional suggested categories are also included by the reviewer, although these do not affect the final Management Procedure results (Table 12). This approach is supported; however, another approach would be to develop Performance Measures that will help with the selection process, e.g. sensitivity, OFL range, availability and reliability of real world data, etc. These will allow the approach to be more open and transparent (although some judgement calls are generally still likely).

Recommendation 41. Develop further Performance Measures that could be used to refine the Performance Measures to a smaller set of candidate Management Procedures.

The final selection of candidate Management Procedures is supported. It is not surprising, given the uncertainties, that the slope class of Management Procedures (which tend to work on relative changes rather than absolute values) are more robust than many of the others. The mean length Management Procedure is also useful and can be applied to PR Hogfish and Yellowtail Snapper.

However, none of the chosen Management Procedures were "tuned" to their specific case. The original Management Procedures were from very different situations and were originally tested in MSEs outside of the toolkit (e.g. Geromont and Butterworth (2014)). Usually, each Management Procedure chosen is more specifically tuned to a specific case. For example, the CC4 is set at 70% of the most recent 5-year catches. It assumes that the catches are too high and that the past five years are better than, say, 4 years. Similar issues apply to each Management Procedure.

Recommendation 42. The candidate Management Procedures are appropriately chosen. However, future work should concentrate on how best to tune the settings of these (and any new Management Procedures) for the particular cases to which they are applied, since these Management Procedures tend to come from quite different cases in other parts of the world and were subjected themselves to tuning to their specific cases.

Table 11. Identification with relevant support for the exclusion of Management Procedures for further use in recommending catch levels based on the precise and unbiased observation model assumption and revised length at full selection results. A strikethrough indicates exclusion and an asterisk could apply if a lower LTY of less than 50% could be applied. Bold red are additions by reviewer. Source Cummings and Sagarese SEDAR 46 Addendum

Acceptance Issue	**PR Hogfish**	**PR Yellowtail Snapper**	**STT Queen Triggerfish**	**STT Spiny Lobster**	**STX Spiny Lobster**	**STX Stoplight Parrotfish**	**Research Recommendations**
Data quality							
Depletion uncertain	~~MCD~~	~~MCD~~	~~MCD~~	~~MCD~~	~~MCD~~	~~MCD~~	Convene expert team to develop estimates of depletion, explore Productivity-Susceptibility Analysis (see NMFS 2011)
Current Abundance uncertain	~~Fratio, YPR~~		~~Fratio~~	~~Fratio~~	~~Fratio~~	~~Fratio~~	Convene expert team to develop estimates of current abundance using better estimates of F (e.g., from mean length approaches)
Catch and effort history uncertain	~~**DD, DD4010**~~	~~**DD, DD4010**~~	~~**DD, DD4010**~~			~~**DD, DD4010**~~	
Life history							Convene workshop to characterize LH demographics and uncertainty estimates
Uncertain maximum Age and/or Mort			~~DD, DD4010~~	~~DD, DD4010~~	~~DD, DD4010~~		
Protogyny							
Uncertain growth Parameters			~~YPR_ML~~			~~DD, DD4010, YPR_ML~~	
Index of abundance restricted						~~Islope1, Islope4, Itarget4~~	Develop statistically robust fishery-independent surveys
Unrealistic results							
Catch recommendations exceeding or near largest observed catches		~~DD, DD4010~~	~~DD, DD4010~~	~~DD, DD4010~~	~~DD, DD4010~~		Further investigation into discard estimates, catch reporting and verification
OFL uncertain	~~**DD, DD4010**~~	~~**DD, DD4010**~~	~~**DD, DD4010**~~	~~**DD, DD4010**~~	~~**DD, DD4010**~~	~~**DD, DD4010**~~	
Unacceptable performance in MSE							

Long-term yield < 50% relative to FMSYref	~~Itarget1, CC4~~	~~Itarget1, CC4~~	~~Itarget1, CC4~~	~~Islope1*, Islope4, Itarget1, Itarget4, CC4~~	~~Islope1, Islope4, Itarget1*, Itarget4, CC4*~~	~~Itarget4~~	Convene methods workshop to develop framework for assessing data limited stocks (e.g., NMFS 2011)

Table 12. Identification with relevant support for the exclusion of Management Procedures for further use in recommending catch levels based on the imprecise and biased observation model assumption and revised length at full selection results. A strikethrough indicates exclusion and an asterisk could apply if a lower LTY of less than 50% could be applied. Bold red are additions by reviewer. Source Cummings and Sagarese SEDAR 46 Addendum

Acceptance Issue	PR Hogfish	PR Yellowtail Snapper	STT Queen Triggerfish	STT Spiny Lobster	STX Spiny Lobster	STX Stoplight Parrotfish	Research Recommendations
Data quality							
Depletion uncertain							Convene expert team to develop estimates of depletion, explore Productivity-Susceptibility Analysis (see NMFS 2011)
Current Abundance uncertain							Convene expert team to develop estimates of current abundance using better estimates of F (e.g., from mean length approaches)
Catch and effort time series uncertain	~~**DD, DD4010**~~	~~**DD, DD4010**~~	~~**DD, DD4010**~~			~~**DD, DD4010**~~	
Life history							Convene workshop to characterize LH demographics and uncertainty estimates
Uncertain maximum Age and/or Mort			~~DD~~		~~DD~~		
Protogyny							
Uncertain growth parameters			~~YPR_ML, SPR40_ML~~			~~DD, YPR_ML~~	
Index of abundance restricted						~~Islope1, Islope4, Itarget1~~	Develop statistically robust fishery-independent surveys
Concerns over catch						~~CC4~~	Revisit landings
Unrealistic results							
Catch recommendations exceeding or near largest observed catches			~~DD~~		~~DD~~		Further investigation into discard estimates, catch reporting and verification
OFL uncertain	~~**DD, DD4010**~~	~~**DD, DD4010**~~	~~**DD, DD4010**~~	~~**DD, DD4010**~~	~~**DD, DD4010**~~	~~**DD, DD4010**~~	
Unacceptable performance in MSE							

Long-term yield < 50% relative to FMSYref	~~Islope1*, Islope4*, Itarget1, CC4, YPR_ML, SPR40_ML *~~	~~Itarget1, CC4, YPR_ML*, SPR40_ML*~~	~~Islope1*, Islope4*, Itarget1, CC4~~	~~Islope1*, Islope4*, YPR_ML*, Itarget1, CC4~~	~~Islope1*, Islope4*, YPR_ML*, Itarget1, CC4~~	~~Islope1, Islope4, CC4, Itarget1~~	Convene methods workshop to develop framework for assessing data limited stocks (e.g., NMFS 2011)

Table 13. Potential methods for providing catch recommendations based on sufficiency and quality of data, model assumptions, and performance metrics for the MSE using the revised length at full selection assumption and assuming precise and unbiased data inputs within the observation model. - Indicates no recommendations made. Source Cummings and Sagarese SEDAR 46 Addendum

Recommended methods	PR Hogfish	PR Yellowtail Snapper	STT Queen Triggerfish	STT Spiny Lobster[1]	STX Spiny Lobster[2]	STX Stoplight Parrotfish
Index-based	Islope1, Islope4	Islope1, Islope4	Islope1, Islope4	Islope1	Itarget1	-
Catch-based	-	-	-	-	CC4	-
Length-based	YPR_ML	YPR_ML	-	-	-	-

[1] Note that for St. Thomas spiny lobster, the inclusion criterion for long-term yield of 50%, as used by the analysts, would result in no recommended MPs. Islope1 could be applied if a lower relative long-term yield would be acceptable.

[2] Note that for St. Croix spiny lobster, the inclusion criterion for long-term yield of 50%, as used by the analysts, would result in no recommended MPs. The relative long-term yields for applicable methods ranking from highest to lowest are: CC4 (36%), Itarget1 (35.8%), Islope1 (31.8%), Islope4 (30.3%), and Itarget4 (0%). CC4 and Itarget1 could be applied if a lower relative long-term yield would be acceptable.

Table 14. Potential methods for setting catch recommendations based on sufficiency and quality of data, model assumptions, and performance metrics for the MSE using the revised LFS assumption and assuming imprecise and biased data inputs within the observation model. - Indicates no recommendations made. Source Cummings and Sagarese SEDAR 46 Addendum

Recommended methods	PR Hogfish[1]	PR Yellowtail Snapper[2]	STT Queen Triggerfish[3]	STT Spiny Lobster[4]	STX Spiny Lobster[5]	STX Stoplight Parrotfish
Index-based	Islope1, Islope4	Islope1, Islope4	Islope1, Islope4	Islope1, Islope4	Islope4, Islope1	-
Length-based	SPR40_ML	YPR_ML, SPR40_ML	-	YPR_ML	YPR_ML	-

[1] Note that for Puerto Rico hogfish, the inclusion criterion for long-term yield of 50%, as used by the analysts, would result in no recommended MPs. The relative long-term yields for applicable methods ranking from highest to lowest are: Islope1 (39.8%), SPR40_ML (36%), Islope4 (35.9%), YPR_ML (34%), Itarget1 (12%), and CC4 (9%). Islope1, Islope4 and SPR40_ML could be applied if a lower relative long-term yield would be acceptable.

[2] Note that for Puerto Rico yellowtail snapper, if a lower relative long-term yield would be acceptable, both SPR40_ML (39%) and YPR_ML (38%) could be applied.

[3] Note that for St. Thomas queen triggerfish, the inclusion criterion for long-term yield of 50%, as used by the analysts, would result in no recommended MPs. The relative long-term yields for applicable methods from highest to lowest are: Islope1 (48.9%), Islope4 (49.2%), CC4 (29.9%), and Itarget1 (29.1). Islope1 and Islope4 could be applied if a lower relative long-term yield would be acceptable.

[4] Note that for St. Thomas spiny lobster, the inclusion criterion for long-term yield of 50%, as used by the analysts, would result in no recommended MPs. The relative long-term yields for applicable methods from highest to lowest include Islope1 (29.8%), Islope4 (26.8%), YPR_ML (25%), CC4 (18.5%), and Itarget1 (16.8%). Islope1, Islope4 and YPR_ML could be applied if a lower relative long-term yield would be acceptable.

[5] *Note that for St. Croix spiny lobster, the inclusion criterion for long-term yield of 50%, as used by the analysts, would result in no recommended MPs. The relative long-term yields for applicable methods from highest to lowest include YPR_ML (27%), Islope4 (25.8%), Islope1 (25.6%), CC4 (20.2%), and Itarget1 (18.3%). YPR_ML, Islope4 and Islope1 could be applied if a lower relative long-term yield would be acceptable.*

4.5 ToR 4. 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 future assessments. • Provide recommendations on possible ways to improve the SEDAR process.

Several of the documents and presentations provided research recommendations, and strength and weaknesses of the various approaches. Reviewer recommendations are in Section 5.2.

4.5.1 Data/assessment review workshop recommendations

Research recommendations were provided in each sub-section of the report but also as an overall section in itself. Comments are made against the original recommendations (quoted in italics).

1. *A statistical review of existing fishery independent surveys to identify an optimum sampling design for development of fishery independent abundance indices. Fishery independent surveys can contribute critical information regarding trends in stock abundance, which can be applied in relatively simple management procedures.*
2. *Develop indices of abundance for spiny lobster using all available data since 1970s with focus on a fishery independent survey.*

Independent surveys are seen as an important dataset for fisheries management, but usually are also very expensive. If the resources are available, then independent surveys are important. However, it would be worth using the DLMtool to investigate how comprehensive the survey needs to be and also what level of CV would make the survey index less effective than other indices.

3. *Investigate more justifiable estimates of stock depletion (Dep) and deletion over time (Dt), such as through Productivity-Susceptibility Analysis (e.g. Cope at al. 2015) or using methods such as mean length estimators.*
4. *Investigate more justifiable estimates of current stock abundance.*

Depletion and estimates of current stock abundance are notoriously difficult to obtain without a data rich stock assessment model. To my knowledge the Productivity-Susceptibility Analysis is not a tool that would provide these values, but the mean length would be able to contribute. The reality is that with data limited methods, this particular gap is common and is why other data limited management strategies are also used, for example ones that rely on a relative index of abundance such as *Islope*. In terms of the MSE, it would be important to test the Management Procedures' performance under different levels of depletion as a sensitivity test. This allows one to determine which Management Procedure is most robust to stock status.

Recommendation 43. Undertake a sensitivity test in the MSE using different levels of depletion in the Operating Model.

5. *Enhanced catch at length by gear sampling is needed to better inform selectivity at age.*

Many of the Management Procedures that performed well required knowledge of selectivity or length frequency. Good catch at length data are essential and this research priority is supported.

6. *Investigate fleet dynamics to more capture fishery dynamics.*

This does become a higher priority as the MSE gets more refined. There are several fleet dynamic models available that make some broad assumptions, but many MSEs develop fishery specific fleet models. Furthermore, fleet dynamics become more important as TACs are set that impact on the fishery and fisher behaviour changes with different management options. The importance of this research recommendation would depend on basic tests on the sensitivity of the MSE results to the fleet model.

7. *Identify target catch or index levels which could be used in conjunction with catch and index time series.*
8. *Identify target length levels which could be used in conjunction with catch and a length frequency series.*

These are essential as this review has already highlighted that each tested Management Procedure had different implied target levels. Usually, one would expect the Management Procedure to be more tuned than they presently are. This is a high priority and the recommendation is supported.

9. *Develop a weighting scheme for length composition and multiple gear fisheries reflective of the stock.*

A stated above, length data are important to many of the Management Procedures and this recommendation is supported.

10. *Consider organising species into species complexes for assessment based on life history, market characteristics, and vulnerability. This could help streamline the stock assessment process in a data-limited context.*

This ties in very well with a recommendation of doing life history reviews at family level as was done in Hawaii. This recommendation is supported and is a high priority.

11. *Conduct a transparent review process of data-limited methods used to set ACLs similar to the workshop conducted by the Pacific Fishery Management Council (PFMC) SSC (NMFS 2011).*
    a. *What are the data requirements and assumptions of the method?*
    b. *What are the conditions under which the method is applicable?*
    c. *Is the method correct from a technical perspective?*
    d. *How robust are model results to departures from model assumptions and atypical data inputs?*
    e. *Incorporation of model uncertainty*
    f. *Identify process to evaluate model results that incorporates objectivity, transparency (i.e., simulation/ management strategy evaluation)*
    g. *What level of review is appropriate for assessments conducted using the method?*

It is essential that more time is allocated to the needs of the MSE as described in the review text. In the review process it is recommended that there is an additional workshop that will allow many of the above to be undertaken. Supported.

12. *Draw from SEDAR 46 DLMtool sensitivity results to identify data collection priorities.*

Also recommended in the review text. There are extremely useful results to inform priorities. High priority.

13. *Convene a workshop to review demographic data for all species in US Caribbean.*

As above, supported but the work should be in the context of the MSE needs. It should also undertake meta-analyses to best draw from a range of studies, and consider undertaking the parameter setting at group level such as family, rather than species by species.

4.5.2 DLMtool

14. *Revisions of the DLMtool software to enhance the model functionality to allow multiple indices of abundance.*

Would be very useful, but as yet not essential for the purpose, given the MSE process is at an early stage. Multiple indices would add complications such as data weighting that may be difficult to resolve in the short term.

15. *Revision of DLMtool software to allow age varying M.*

Again this would be useful but not the most important change in the MSE that is required for this application.

16. *Allow for implementation error of the harvest control rule (e.g. TAC overages) within the implementation model in the MSE.*

This is essential as the MSE without any implementation errors is incongruous given these are usually applied to data limited fisheries where perfect implementation is usually harder to achieve. High priority.

17. *Idealized observation models for catch composition data - currently DLMtool simulates catch-composition data from the simulated catch composition data via a multinomial distribution and some effective sample size. This observation model may be unrealistically well-behaved and favour those approaches that use these data. Harvest control rules may be integrated into data-limited Management Procedures.*

Although this is an issue, the current practice is often an approach used in other MSEs. Important but not highly, compared to other changes.

18. *Harvest control rules – in the version of the DLMtool applied in SEDAR 46, harvest control rules (e.g., the 40-10 rule) must be written into a data-limited Management Procedure. There is currently no ability to do a factorial comparison of say 4 harvest control rules using 3 Management Procedures (the user must describe all 12 combinations). The reason for this is that it would require further subclasses. For example, the 40-10 rule may be appropriate or output of DBSRA but it would not be appropriate for some of the simple management procedures such as DynF that already incorporate throttling of TAC recommendations according to stock depletion.*

This would be quite a big change, but an important addition to the model. This aspect would also be high priority when a Management Procedure is applied to a depleted stock. This can of course be manually coded, so a medium priority compared to other recommendations.

19. *Implementation error – In this edition of the DLMtool there is no implementation error. The only imperfection between management recommendation and the simulated TAC comes in the form of the MaxF argument that limits the maximum fishing mortality rate on any given age-class in the operating model. The default is 0.8 which is high for all but the shortest living fish species.*

As above, this is a high priority. Some more specific recommendations are provided in the review against the terms of references as to what additional options should be included in the implementation error module.

4.6 ToR 5. 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 candidate management procedures and MSE constitutes the best scientific information:

a) *Relevance* – it uses the most relevant data and life history parameters possible for data limited methods.
b) *Inclusive* – the analyses were comprehensive although there are suggested changes to the sensitivity tests provided.
c) *Objectivity* – the method uses a tool developed outside the system and has been parameterised with great care.
d) *Transparency* – the settings are clear; however, the description of the Operating Model, although available, was not immediately obvious. This does not change the overall conclusions.
e) *Timeliness* – the work was undertaken under time stress. Some additional steps in the process are recommended to allow for more verification and validation of the MSE settings and tests.
f) *Verification* – the work was undertaken with comprehensive sensitivity tests. There is some question as to which model was chosen as the base model, but overall the test selected the most appropriate Management Procedure at this time.
g) *Validation* – it is difficult to validate a data limited approach. The input values were carefully chosen and extensive sensitivity tests were undertaken. Perhaps the mean length estimator could be used more to inform the Operating Model. Care should be taken on how the mean length estimator is then implemented as a Management Procedure.
h) *Peer review* – the DLMtool and Management Procedures have been peer reviewed in the scientific literature. The mean length estimator has been extensively used on tropical fish species, passed panel peer reviews, published in the scientific literature and extensively simulation tested.

4.7 ToR 6. Provide guidance on key improvements in data or modeling approaches that should be considered when scheduling the next assessment.

The approach was undertaken over a very short time period. It required selection of:

1. the species-island units,
2. the appropriate data and parameters for both the Operating Model and Management Procedures,
3. Operating Model and Management Procedure MSE sensitivity tests,
4. Performance measures,
5. Real world input values,
6. Real world sensitivity tests;

And there was a time consuming process of undertaking the MSE and real world tests, collating the information, evaluating the results and writing the report. In the present process, there was a data/assessment workshop, several webinars and the assessment review workshop.

Although the MSE will get more streamlined over time, the norm elsewhere is to spend as much time on the Operating Model as one does on the Management Procedures. A more appropriate approach would be to undertake an additional workshop between the data and review workshops. This means that there would be a data workshop where the input data and parameters are selected (with the MSE and real world Performance Measures in mind), a second workshop where the Operating Model and Management Procedures are selected and run, and a final review workshop where the overall results as investigated and reviewed.

The second workshop would be able to investigate more closely the Operating Model and Management Procedure behaviour and whether the Management Procedures can be more "tuned" to the specific case.

The species-island unit selection should either be known prior to the data workshop or developed as the first step in the workshop.

Recommendation 44. The process should be changed so that an additional operating model and management procedure workshop is added between the data and review workshops. This will allow more time to investigate the Operating Model and Management Procedure behaviour and separated the data work with that of the MSE and all its tests.

The Terms of Reference for data limited approaches need refining. Rather than being orientated towards assessment language these should be rewritten to initially address Operating Model settings, then the Management Procedures and then finally the selection of candidate Management Procedures. This would more closely resemble the data-limited Management Strategy Evaluation process.

Integration of the mean length Management Procedure was underway during the process, but results were provided that were still a mix of separate and combined results. Further refining of joint work is required so that all Management Procedures are being assessed within the same framework and with the same settings/assumptions.

4.8 ToR 7. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference.

Contributions to panel summary have been provided, including comments on the draft report. The final panel summary report was provided.

5 Conclusions and Recommendations in accordance with the ToRs.

5.1 Conclusions

The review investigated the body of work that applies a Management Strategy Evaluation (MSE) using the DLMTool authored by Dr Carruthers to six species-island units in the Caribbean. Extensive sensitivity tests were undertaken that investigated the effect of different operating model and management model settings. Together with those undertaken, this body of work is very comprehensive.

All Terms of References were met and the candidate Management Procedures selected as part of the review process can be used for management. The present Management Procedures do not meet the requirements using the cut offs of the Performance Measures.

The assessment team did a great deal of original work in a short time period and responded fully to panel requests for additional analyses; they are to be congratulated and thanked for a constructive and open review.

5.2 RECOMMENDATIONS

5.2.1 High priority recommendations

The following recommendations are classified as a high priority:

- The DLMTool approaches uncertainty and biases in a basic way. Lack of correlation structure in the parameter values and biases mean that projections can include populations and catch values that are not realistic, which can greatly affect Management Procedure performance and therefore selection,
- A more precautionary length at full selection such as the mode of the catch length distribution should be used as the base model (as opposed to the initial 95th percentile of the cumulative length distribution),
- Future workshop should concentrate on the data and information required for the MSE and delete Management Procedures at an earlier stage if these are unrealistic given the available information,
- Some thought should be given to the ultimate vision of what is to be achieved – whether to emphasise a few species-island units or to spread the approach to others. This would affect how further work is undertaken,
- More transparency is required on how the present model implements the catches from the Management Procedure, e.g. as an ABC and without implementation error,
- The existing Management Procedure should be included in the tests as a reference,
- Short-term transitional Performance Measures should be further developed,
- Undertake a growth study for STT queen triggerfish if the mean length Management Procedure is an attractive alternative,
- More emphasis should be made of the time series performance measures beyond convergence,
- Include implementation uncertainty (with regard to the difference in implemented catch and actual catch, the ability to change the TAC as required by the Management Procedure etc.),
- Develop a greater set of Performance Measures to help reduce the full set of Management Procedures based on the approach already used, and
- An additional workshop should be added between the DW and the RW to emphasise Operating Model set up and Management Procedure selection for the tests.

5.2.2 Full list of recommendations

Recommendation 1. Investigate the possibility of obtaining the original data for life history studies where the uncertainty is not well described. 10

Recommendation 2. The overall approach is sound, but in the future, should focus on what is needed for the MSE and real world Management Procedures. As such, the assessors should

6 REVIEW PROCESS

The review was undertaken in a very constructive light with extensive requests met. The assessors should be complimented on their open and positive approach to any comments. The background material was very comprehensive and useful, apart from the description of the operating model (already commented) on. Sensitivity tests were comprehensive and extensive.

The Terms of References were a little out of context given the data limited approaches applied and that no assessment was undertaken (except for the delay difference model in the MSE). For example, ToR uses the word “assessment” which the panel agreed to interpret as being a determination whether the Management Procedure is sustainable (overfishing and overfished) over the long term or not, whereas stock assessment implies stock status (which is not an option in these analyses). It is better to use the MSE words more specifically, i.e. Management Procedures. Furthermore, parameters are used for both the operating model, Management Procedures and real world situations. The Terms of Reference should probably ask the reviewer to be more specific with regard to all these components.

As stated above, the approach undertaken here occurred over a very short time period. It required selection of:

7. the species-island units,
8. the appropriate data and parameters for both the Operating Model and Management Procedures,
9. Operating Model and Management Procedure MSE sensitivity tests,
10. Performance measures,
11. Real world input values, and

12. Real world sensitivity tests

Thereafter, the time consuming process of undertaking the MSE and real world tests, collating the information, evaluating the results and writing the report had to be done. In the present process, there was a data/assessment workshop, several webinars and the assessment review workshop. Although this process will get more streamlined as one undertakes more MSEs for these species-island groups and learn from others, the norm elsewhere is to spend as much time on the Operating Model as one does on the Management Procedures. Changes are suggested in Terms of Reference 6.

7 APPENDIX 1: CIE STATEMENT OF WORK

Statement of Work

External Independent Peer Review by the Center for Independent Experts

SEDAR 46 US Caribbean Data Limited Species Assessment Review Workshop

Scope of Work and CIE Process: The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Technical Representative (COTR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide impartial and independent peer review without conflicts of interest. CIE reviewers are selected by the CIE Steering Committee and CIE Coordination Team to conduct the independent peer review of NMFS science in compliance the predetermined Terms of Reference (ToRs) of the peer review. Each CIE reviewer is contracted to deliver an independent peer review report to be approved by the CIE Steering Committee and the report is to be formatted with content requirements as specified in **Annex 1**. This SoW describes the work tasks and deliverables of the CIE reviewer for conducting an independent peer review of the following NMFS project. Further information on the CIE process can be obtained from www.ciereviews.org.

Project Description: SEDAR 46 will be a compilation of data, an assessment of the stocks, and CIE assessment review conducted for Caribbean Data-limited Species. 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 stocks assessed through SEDAR 46 are within the jurisdiction of the Caribbean Fisheries Management Council and the territorial waters of Puerto Rico and the U.S. Virgin Islands. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements for CIE Reviewers: Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. CIE reviewers shall have working knowledge expertise 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. Experience with data-limited assessment methods would be preferred. Each CIE reviewer's duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled in **Miami, Florida during February 23-25, 2016**.

Statement of Tasks: Each CIE reviewers shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Prior to the Peer Review: Upon completion of the CIE reviewer selection by the CIE Steering Committee, the CIE shall provide the CIE reviewer information (full name, title, affiliation, country, address, email) to the COTR, who forwards this information to the NMFS Project Contact no later the date specified in the Schedule of Milestones and Deliverables. The CIE is responsible for providing the SoW and ToRs to the CIE reviewers. The NMFS Project Contact is responsible for providing the CIE reviewers with the background documents, reports, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact is also responsible for providing the Chair a copy of the SoW in advance of the panel review meeting. Any changes to the SoW or ToRs must be made through the COTR prior to the commencement of the peer review.

Foreign National Security Clearance: When CIE 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 CIE reviewers who are non-US citizens. For this reason, the CIE 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/ http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact 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 Contact will consult with the CIE Lead Coordinator 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 SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

Panel Review Meeting: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs cannot be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COTR and CIE Lead Coordinator.** Each CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein. The NMFS Project Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact is responsible for ensuring that the Chair understands the contractual role of the CIE reviewers as specified herein. The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

CIE reviewers shall conduct an impartial and independent peer review of the assessment in accordance with the SoW and ToRs herein.

A description of the SEDAR Review process can be found in the SEDAR Policies and Procedures document:
http://sedarweb.org/docs/page/A6-SEDARPoliciesandProcedures_June2014_0.pdf

The CIE reviewers may contribute to a Summary Report of the Review Workshop produced by the Workshop Panel. .

Contract Deliverables - Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

Other Tasks – Contribution to Summary Report: Each CIE reviewer may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. Each CIE reviewer is not required to reach a consensus, and should provide a brief summary of the reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

Specific Tasks for CIE Reviewers: The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate during panel review meeting in **Miami, FL during February 23-25, 2016,** as specified herein, and conduct an independent peer review in accordance with the ToRs (**Annex 2**).
3) No later than March 10, 2016, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Dr. Manoj Shivlani, CIE Lead Coordinator, via email to mshivlani@ntvifederal.net, and Dr. David Sampson, CIE Regional Coordinator, via email to david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

Schedule of Milestones and Deliverables: CIE shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

January 12, 2016	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
February 9, 2016	NMFS Project Contact sends the CIE Reviewers the pre-review documents
February 23-25, 2016	Each reviewer participates and conducts an independent peer review during the panel review meeting
March 10, 2016	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
March 24, 2016	CIE submits CIE independent peer review reports to the COTR
March 31, 2016	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: This 'Time and Materials' task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery

Management Council, and Council's SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COTR within 10 working days after receipt of all required information of the decision on changes. The COTR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

Acceptance of Deliverables: Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COTR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COTR (Allen Shimada, allen.shimada@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COTR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:
(1) The CIE report shall completed with the format and content in accordance with **Annex 1**,
(2) The CIE report shall address each ToR as specified in **Annex 2**,
(3) The CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

Distribution of Approved Deliverables: Upon acceptance by the COTR, the CIE Lead Coordinator shall send via e-mail the final CIE reports in *.PDF format to the COTR. The COTR will distribute the CIE reports to the NMFS Project Contact and Center Director.

Support Personnel:

Allen Shimada
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
Allen Shimada@noaa.gov Phone: 301-427-8174

Manoj Shivlani, CIE Lead Coordinator
NTVI Communications, Inc.
10600 SW 131st Court, Miami, FL 33186
mshivlani@ntvifederal.com Phone: 305-968-7136

Key Personnel:

NMFS Project Contact:

Julie A Neer
SEDAR Coordinator
4055 Faber Place Drive, Suite 201
North Charleston, SC 29405
(843) 571-4366
julie.neer@safmc.net

Annex 1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent report shall 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 main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including providing 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, and especially where there were divergent views.

c. Reviewers should elaborate on any points raised in the Summary Report that they feel 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 CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The CIE independent report shall be an independent peer review of each ToRs, and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include the following appendices:

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

Annex 2: Terms of Reference for the Peer Review

SEDAR 46 US Caribbean Data Limited Species Assessment Review Workshop

1. *Evaluation, findings and recommendations of data collection operations and survey design*
2. *Evaluation and recommendations of data quality*
3. *Evaluation of strengths and weaknesses, and recommendations of analytic methodologies*
4. *Evaluation and recommendations of model assumptions, estimates, and uncertainty*
5. *Evaluation, findings, and recommendations of result interpretation and conclusions*
6. *Determine whether the the science reviewed is considered to be the best scientific information available.*
7. *Recommendations for further improvements*
8. *Brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations*

Annex 3: Tentative Agenda
SEDAR 46 US Caribbean Data Limited Species Assessment Review Workshop

Miami, Florida
February 23-25, 2016

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

Wednesday
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*
6:00 p.m. – 6:30 p.m. Public comment 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 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

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

8 APPENDIX 2: PANEL MEMBERSHIP, LIST OF PARTICIPANTS AND DOCUMENTS

8.1 PANEL MEMBERSHIP

Drs Cathy Dichmont, John Hoenig, Jerald Ault, Vance Vicinte (chair), Paul Medley, Panayiota Apostolaki.

8.2 LIST OF PARTICIPANTS

Workshop Panel

Vance P. Vicinte, Chair ..... Chair, SSC
Panayiota Apostolaki ..... CIE Reviewer
Jerald S. Ault ..... Council Appointee
Cathy Dichmont ..... CIE Reviewer
John Hoenig ..... SSC
Paul Medley ..... CIE Reviewer

Analytic Representation

Nancie Cummings ..... SEFSC, Miami
Adyan Rios ..... SEFSC, Miami
Skyler Sagarese ..... SEFSC, Miami

Appointed Observers

Winston Ledee ..... Industry Representative – St. Thomas
Roberto Silva ..... Industry Representative – Puerto Rico

Observers

Molly Adams ..... Univ. of Miami
Meaghan Bryan ..... SEFSC, Miami
Quang Huynh ..... VIMS
Jeff Isley ..... SEFSC, Miami
Bill Harford ..... RSMAS/SEFSC
Michael Larkin ..... SERO
Vivian Matter ..... SEFSC, Miami
Daniel Matos-Caraballo ..... PRDNER
Kevin McCarthy ..... SEFSC, Miami
Clay Porch ..... SEFSC, Miami

Staff

Julie Neer ..... SEDAR
Julie O'Dell ..... SAFMC Staff
Kate Quigley ..... CFMC Staff

8.3 LIST OF REVIEW WORKSHOP WORKING PAPERS AND DOCUMENTS

Documents Prepared for the Review Workshop			
SEDAR46-RW-01	Estimating total mortality rates and calculating overfishing limits from length observations for six U.S. Caribbean stocks	Quang C. Huynh	14 Jan 2016
SEDAR46-RW-02	Management strategy evaluations for mean length-based management procedures using DLMtool	Quang C. Huynh	22 Feb 2016
SEDAR46-RW-03	An alternative approach to setting annual catch limits for data-limited fisheries: Use of the DLMtool and mean length estimator for six US Caribbean stocks	Nancie Cummings, Skyler Sagarese and Quang C. Huynh	22 Feb 2016

Reference Documents Submitted during the Review Workshop		
SEDAR46-RD04	Evaluating methods for setting catch limits in data-limited fisheries	Thomas R. Carruthers, André E. Punt, Carl J. Walters, Alec MacCall, Murdoch K. McAllister, Edward J. Dick, Jason Cope
SEDAR46-RD05	Evaluating methods for setting catch limits in data-limited fisheries: Supplemental Appendix A	Thomas R. Carruthers, André E. Punt, Carl J. Walters, Alec MacCall, Murdoch K. McAllister, Edward J. Dick, Jason Cope
SEDAR46-RD06	DLMtool: Data-Limited Methods Toolkit (v2.1.1)	Tom Carruthers and Adrian Hordyk
SEDAR46-RD07	Length-based assessment of sustainability benchmarks for coral reef fishes in Puerto Rico	Jerald S. Ault, Steven G. Smith, Jiangang Luo, Mark E. Monaco, and Richard S. Appeldoorn
SEDAR46-RD08	Data Limited Techniques for Tier 4 Stocks: An alternative approach to setting harvest control rules using closed loop simulations for management strategy evaluation	Jason McNamee, Gavin Fay, and Steven Cadrin
SEDAR46-RD09	Application of Data-Poor Harvest Control Rules to Atlantic Mackerel	John Wiedenmann
SEDAR46-RD10	September 2015 Mid-Atlantic SSC Meeting Report – Black Sea Bass Review	Mid-Atlantic SSC
SEDAR46-RD11	Stock assessment of protogynous fish: evaluating measures of spawning biomass used to estimate biological reference points	Elizabeth N. Brooks, Kyle W. Shertzer, Todd Gedamke, and Douglas S. Vaughan

9 APPENDIX 3: REFERENCES

Ault, J.S. & Ehrhardt, N.M. (1991) Correction to the Beverton and Holt Z-Estimator for Truncated Catch Length-Frequency Distributions. *Fishbyte***,** 37-39.

Ault, J.S., Smith, S.G. & Bohnsack, J.A. (2005) Evaluation of average length as an estimator of exploitation status for the Florida coral-reef fish community. *ICES Journal of Marine Science,* **62,** 417-423.

Ault, J.S., Smith, S.G., Luo, J., Monaco, M.E. & Appeldoorn, R.S. (2008) Length-based assessment of sustainability benchmarks for coral reef fishes in Puerto Rico. *Environmental Conservation,* **35,** 221–231.

Brooks, L., Shertzer, K.W., Gedamke, T. & Vaughan, D.S. (2008) Stock assessment of protogynous fish: evaluating measures of spawning biomass used to estimate biological reference points. *Fisheries Bulletin,* **106,** 12-23.

Carruthers, T.R. & Hordyk, A. (2015) DLMtool: Data-Limited Methods Toolkit. *SEDAR 46* (ed. A.a.R. South-east Data), pp. 40. University of British Columbia, Vancouver, Canada.

Carruthers, T.R., Punt, A.E., Walters, C.J., MacCall, A., McAllister, M.K., Dick, E.J. & Cope, J. (2014) Evaluating methods for setting catch limits in data-limited fisheries. *Fisheries Research,* **153,** 48-68.

Fulton, E.A., Smith, A.D.M., Smith, D.C. & van Putten, I.E. (2011) Human behaviour: the key source of uncertainty in fisheries management. *Fish and Fisheries,* **12,** 2-17.

Gedamke, T. & Hoenig, J.M. (2006) Estimating Mortality from Mean Length Data in Nonequilibrium Situations, with Application to the Assessment of Goosefish. *Transactions of the American Fisheries Society,* **135,** 476–487.

Geromont, H.F. & Butterworth, D.S. (2014) Complex assessments or simple management procedures for efficient fisheries management: a comparative study. *ICES Journal of Marine Science*.

Hewitt, D.A. & Hoenig, J.M. (2005) Comparison of two approaches for estimating natural mortality based on longevity. *Fisheries Bulletin,* **103,** 433-437.

Hoenig, J.M. (1983) Empirical use of longevity data to estimate mortality-rates. *Fishery Bulletin,* **81,** 898-903.

Huynh, Q. (2016) Estimating natural mortality rate and overfishing imits from length observations for six U.S. Caribbean stocks. . *SEDAR 46*, pp. 19 pp. NMFS, Miami, U.S.A.

Pauly, D. (1980) On the intterrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. *Journal du Conseil International pour l'Exploration de la Mer,* **39,** 175-192.