Table of Contents

Table of Contents .................................................................................................................. 2
1. INTRODUCTION ............................................................................................................... 2
   1.1 WORKSHOP TIME AND PLACE .................................................................................. 2
   1.2 TERMS OF REFERENCE .............................................................................................. 2
   1.3 LIST OF PARTICIPANTS .............................................................................................. 3
   1.4 LIST OF REVIEW WORKSHOP WORKING PAPERS AND DOCUMENTS .......... 4
2. REVIEW PANEL REPORT ................................................................................................. 5

1. INTRODUCTION

1.1 WORKSHOP TIME AND PLACE

The SEDAR 46 Review Workshop was held February 23-25, 2016 in Miami, Florida.

1.2 TERMS OF REFERENCE

1. Evaluate the data used in the assessment, including discussion of the strengths and weaknesses of data sources and decisions, and consider the following:
   a) Are data decisions made by the DW and AW sound and robust?
   b) Are data uncertainties acknowledged, reported, and within 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.
• 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. 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.

4. 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. Prepare a Peer Review Summary summarizing the Panel’s evaluation of the stock assessment and addressing each Term of Reference.

1.3 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
1.4 LIST OF REVIEW WORKSHOP WORKING PAPERS AND DOCUMENTS

<table>
<thead>
<tr>
<th>Documents Prepared for the Review Workshop</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SEDAR46-RW-01</td>
<td>Estimating total mortality rates and calculating overfishing limits from length observations for six U.S. Caribbean stocks</td>
</tr>
<tr>
<td>SEDAR46-RW-02</td>
<td>Management strategy evaluations for mean length-based management procedures using DLMtool</td>
</tr>
<tr>
<td>SEDAR46-RW-03</td>
<td>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</td>
</tr>
</tbody>
</table>

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, |
| SEDAR46-RD06 | DLMtool: Data-Limited Methods Toolkit (v2.1.1) | Jason Cope |
| SEDAR46-RD07 | Length-based assessment of sustainability benchmarks for coral reef fishes in Puerto Rico | Tom Carruthers and Adrian Hordyk |
| 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 | Jerald S. Ault, Steven G. Smith, Jiangang Luo, Mark E. Monaco, and Richard S. Appeldoorn |
| SEDAR46-RD09 | Application of Data-Poor Harvest Control Rules to Atlantic Mackerel | Jason McNamee, Gavin Fay, and Steven Cadrin |
| 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 |

2. REVIEW PANEL REPORT


PANEL REPORT

SEDAR 46 U.S. Caribbean Data-Limited Species Assessment Terms of Reference (ToR)

Vance Vincente (Chair), Panayiota Apostolaki, Jerald S. Ault, Catherine M. Dichmont, John M. Hoenig, and Paul A.H. Medley

Panel Overview

The overall Data Limited Method (DLM) approach presented to the SEDAR 46 Review Panel appeared appropriate, as was the general method of selecting species for these assessments. The DLM methods presented at SEDAR 46 appeared to have been applied correctly and the analysts successfully came up with a set of candidate Management Procedures (MPs). However, the Panel felt strongly that the analysts still need to refine their approach based on several principal suggestions, such as:

- The analysis should develop a more sophisticated approach to developing and assigning parameter and input data variances in the Operating Model (OM). For example, reference to the actual variance-covariance relationships for growth models would be far more appropriate than simply selecting growth parameters independently from uniform distributions.
- There is need to examine the numerical performance of the OM in much more detail. Some of the estimation procedures and MPs appeared grossly ad hoc.
- We recommend tuning of the candidate MPs to the specific species-island unit cases that are to be examined.

We are generally satisfied that the new candidate MPs outperformed the current MP, and furthermore, that the current MPs failed to meet the performance criteria used to evaluate all of the other MPs. The Panel also noted that, with regard to the current MP, ad hoc assignment of “averaged catches over the past x years” as the target ACL has no theoretical or empirical basis for selection. This type of ACL designation makes no specific reference to the actual exploitation rate required to achieve those catches or the status of the stock under that catch regime. The findings of the assessment are appropriate to guide management discussions and provide enough evidence that the candidate MPs could be used for setting annual catch limits.

Finally, the Panel agreed that the assessment team did a great deal of original work in the process of development of their SEDAR 46 analyses and presentations, and in addition, responded fully to every panel request made for additional clarifying analyses. The assessment team is congratulated by the Panel for a job very well done!
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? and, (d) Are input data series reliable and sufficient to support the assessment approach and findings?

The ToR uses “assessment” – a word we interpret as being a determination of whether the MP is sustainable over the long term or not; whereas, stock assessment implies stock status as being sustainable or non-sustainable at a particular time (which is not an option in these analyses).

The panel supports the assessors that, overall, the data were constrained such that they were not appropriate for use in a conventional stock assessment model, but they were adequate if used with the appropriate MPs (Note: the analysts did not undertake a stock assessment per se, but used the data to set up the needed parameters for the OM and MP, and then used a small subset of the MPs). Broadly the analytical approach was good, given the available data. The assessors clearly acknowledged the weaknesses of these analyses and assembled the best scientific data available. They also evaluated and addressed uncertainties in a Management Strategy Evaluation (MSE). The Data Workshop (DW) chose the test species-island units well: 5 with the best data and the 6th unit was perhaps the worst data available to them (but could still have enough data to apply an MP).

In general, variances were applied in the OM assuming uniform distributions which is questionable given the inherent Gaussian variance-covariance relationships for many of the parameters, such as life history.

- Specifically, the mean responses of the life history (LH) parameters are generally appropriate for use in the OM and MPs, with some exceptions. The mean lengths are the most reliable data sources, whereas the catch data quality was highly variable. There was no apparent statistical substance for expansion factors that have been applied to these data, and which prevented computation of specific estimate variances. Again, there were largely no variance estimates provided for these types of data, and as a result the CV was chosen in an somewhat ad hoc manner, simply because the estimates were not derived from a design-based approach. The data design and collection is weak in many respects. Unfortunately, there was not much the assessment team could do about this. Current data systems put serious limitations on broader-based analyses for a while.

Recommendations

- Concentrate future efforts on key data for these analyses. These would include robust measures of CPUE, catch and fishery-independent length frequency distributions, accurate and precise LH parameters for the entire range of key fisheries. In that regard, there should be a refined focus on specific MPs.
- The currently used MP (averaged catches over several recent years) performed poorly in comparison to a number of alternative methodologies considered in this workshop.
Future efforts could be greatly improved by refining and clarifying a range of certain data and model inputs. These actions would result in a smaller solution space in OM and key MPs.

- The feedback control should be used as an incentive to get better data, which for example a constant catch approach tends not to do.
- \( L_c \), length at first capture, was set up incorrectly in the model. We recommend using a default as either the mode of the size-frequency distribution, or perhaps even a smaller size to account for variability of length at ages. Use the mode when there is presumed knife-edged selectivity. At the same time, a test for dome-shaped selectivity would be important.

**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?**

**Responses:**

(A) Are the data-limited methods scientifically sound and robust?
Yes, the DLM Toolkit and the management procedures have been peer reviewed.

(B) Are the methods appropriate given the available data?
Yes, this is a data limited set of fisheries and this toolbox 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 (after further within-workshop runs were undertaken), although there are suggested future refinements to the DLM Toolbox and management procedure process.

(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 (i.e., one would not expect many of the above list to be estimated by these methods), they produce the necessary information from which to produce an overfishing limit (OFL).
Summary:
The DLM tool developed by Caruthers and co-workers (Carruthers et al. 2014; Carruthers & Hordyk 2015) were applied to the six species-island units. The DLM tool is an R package that contains the normal MSE components of an OM and data-limited management procedures. The MSE facilitates simulation testing of uncertainties and biases in the data and life history parameters/assumptions. A sub-set of the available MPs was utilised. These have been used elsewhere in the world. This toolkit and management procedures are freely available and the different components of the model have been peer reviewed through the journal publications process, for example Gedamke and Hoenig (2006); Carruthers et al. (2014); Geromont and Butterworth (2014).

A subset of the DLM tool MPs was tested for these cases – these were selected from a set based on whether they apply, whether they provide a good reference set and how a more data rich method would perform. The MPs all require a different mix of information, 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 MP (Huynh 2016) was tested externally to the MSE (although it was integrated during the workshop). In addition, a reference MP that assumed perfect information was included. Not all the MPs tested can be used in actuality, but the tested selection of MPs was appropriate and highlighted the sensitivities to data and underlying assumptions.

During the workshop, the MPs currently being used were tested. These do not perform well relative to the other candidate MPs and mostly do not satisfy the Performance Measure criteria as applied to the other MPs.

The toolbox includes a set of Performance Measures for comparing the different management procedures. The four Performance Measures chosen by the Assessment Panel are appropriate, being (a) the probability of not overfishing, (b) probability of being overfished, (c) the inter-annual catch variability, and (d) the long term yield. 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 and so define the outside extreme of OFL space. The MSE is implementing the Management Procedure value as the final TAC set in the process i.e. the MP was assumed to deliver the ABC/TAC, yet the Performance Measure cut offs were set up so as to conform to an OFL. Care should therefore be taken with final MP choice.

Good MP diagnostics were provided. This information, together with additional sensitivity tests were extremely helpful in reducing the full set of MPs to a sub-set of candidate MPs. However, short-term transitional Performance Measures were not tested prior to the workshop. Work during and after the review shows that these transitional statistics are important.
In addition to the MSE tests, the MPs were tested using real world data. This means that only MPs that could be undertaken in the real world would be highlighted here. These were appropriately implemented.

The operating model in the DLM tool needs further refining. The DLM tool OM is provided with extensive input parameters and their associated uncertainty to simulate an age-based population. The OM samples parameter settings where upper and lower bounds have been provided and assumed a uniform distribution, i.e., they are uninformative priors. Also, these priors are assumed to act independently of each other, which means that uncertainty is over-specified and some life-history combinations may not describe the species-island unit. The toolkit also does not include implementation uncertainty which means that it assumed that all management decisions are implemented without any error. Despite these potential weaknesses, the toolbox is an appropriate tool to apply to these data limited applications.

Only convergence statistics of the OM were investigated – guidance as to further review of the OM performance was provided during the workshop and these were provided during the review process. These highlight a few inconsistent behaviours and the value of these tests.

**ToR 3: Consider how uncertainties in the assessment, and their potential consequences, are addressed.**

Uncertainties in an assessment can arise from multiple sources: problems of data quality, the need to make assumptions in the assessment, and uncertainty about model formulation and stock population dynamics. Some of these problems can be handled individually, e.g., providing a range of possible values for an input parameter to the assessment model, fitting alternative models and using model diagnostics to look for problems in model formulation. The assessment team did all this in their pilot data-limited analyses of Caribbean fisheries.

The potential consequences of uncertainties in the assessment can be, and was, studied through MSE simulations. The assessment team relied heavily on this approach to choose MPs and evaluate their likely performance under different assumptions about uncertainty in the data and its sources (i.e. error/data inaccuracy or bias). In order to do this, they had to specify an OM for each stock to simulate what the stock might do if managed a certain way. The OM generates observations on the stock which are then fed into the assessment and management model to generate an impact on the stock. The OM is then updated, providing new observations that are fed into the assessment and management models, and so forth.

The review team felt that the use of MSE simulations was appropriate. This was a new initiative for the assessment team and it involved evaluation of a large number of scenarios. For the most part the MSE were handled very well.
Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty in the population, data sources, and assessment methods.

The review team felt a critical element of MSE is the construction of the OM model. The assessment team made a good effort to develop realistic operating models; however, these models should be substantially refined in the future assessments. In particular, there was some confusion about how selectivity in the fishery was modeled and this requires further investigation.

The assessment team devoted a great deal of attention to the unbiased and precise data scenario, and treated this as the base case. They then considered numerous variations on this theme by introducing biases and imprecision in various places. The review Panel questioned the choice of the unbiased and precise data as the base case and considered whether biased and imprecise data scenario should be the base case since this SEDAR is envisioned as a template for data poor stock assessments. Unbiased and precise data should only be considered if there is reason to believe that this is a plausible scenario for a particular data-poor stock.

The assessment team initially focused on specific metrics from the MSE simulations. The review team expressed the opinion that there are additional performance metrics that should be considered, specifically related to model diagnostics and transitional characteristics of the model. The assessment team responded by providing additional metrics.

The review team is satisfied that the assessment team evaluated the significant sources of uncertainty in the population, data sources, and assessment methods through a careful evaluation of each source of data, and through a combination of management strategy evaluation simulations, sensitivity analyses, and examination of model diagnostics.

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

The review panel felt the implications of uncertainty were clearly stated. However, the basis of those conclusions depends very much on the MSE simulations, and the nature of those simulations, i.e., the construction of the operating models and the alternative scenarios, needs to be documented fully in a technical report.
ToR 4: Consider the research recommendations provided by the Data and Assessment workshops and make any additional recommendations or prioritizations warranted. These recommendations should: (a) clearly denote research and monitoring that could improve the reliability of future assessments; and, (b) provide recommendations on possible ways to improve the SEDAR process.

This section considers the research recommendations initially provided by the DW and AW that were then considered by the SEDAR 46 Review Panel. The Review Panel generally supported the recommendations from the DW and AW, and those from the assessment team. However, the Review Panel extended these recommendations as outlined below. Recommendations fell into two general categories: (1) data; and, (2) model.

Data

One of the fuzziest aspects of the data-limited process was how exactly data reliability was qualified or quantified. We discovered that fishery data precision (e.g., coefficient of variation, CV) was not able to be determined from the current fishery catch sampling methodologies that are employed in the Caribbean. While this was probably a topic of conversation at the DW, there was insufficient discussion of these critical issues in the SEDAR 46 DW/AW report (AW). There needs to be a solid focus on data design strategies as the data-limited process moves forward in the region to establish ACLs for a range of species presently not under consideration.

Thus, two aspects of model inputs must be addressed: (1) life history demographics; and, (2) fishery-dependent data (size-structured catch and fishing effort). Research into what defines the “best” demographic parameters for DLM model inputs, for example, most accurate and precise growth (length-at-age) curve, maximum age (i.e., natural mortality rate), size at first capture (selectivity ogive), size at first sexual maturity (maturation ogive), etc. There seemed to be insufficient attention to these issues in the workshop, and arbitrary (non-estimated) CVs were applied to data inputs. Perhaps the number one priority is to refine the life history demographic parameters identified by the DW across the region, and to improve accuracy and precision of those basic data. This strategy would likely be facilitated by a workshop of technical experts convened, in the near future, to review and analyse existing life history demographic data for all relevant exploited species in the U.S. Caribbean, Southeast U.S. and Gulf of Mexico. When joint parameter variance-covariance is not available, how will estimates of uncertainty for life history demographic parameters, for example, be provided? This would include quantitative justifications for error variances and CVs.

A focus on design-based strategies for ensuring collection of accurate and precise fisheries-dependent commercial and recreational data should be advanced in the region. This would greatly improve fishery-dependent mean (and variance) estimates of landings, discards and the effort required to obtain them. The sampling protocols must be optimized to ensure representative sampling across size-age spectra over time and space. If precise estimates were obtained in the most recent years, then a data-limited analysis could identify current
exploitation rates and resource sustainability. In addition, it makes sense to conduct a statistical review, analysis and optimal sampling design of complimentary fishery-independent surveys as these could provide extremely important spatially-integral, accurate and precise information on exploitation effects by measuring what is left in the water after fishing has occurred.

More work must be done on evaluation of species selection criteria. The adequacy of the choice of species suitable for these pilot species analyses was generally successful. However, a couple of those species provided little guidance on model performance. These analyses revealed issues in three areas: (1) appropriate models and benchmarks; (2) reliable life history demographic data; and, (3) adequate fishery-dependent data.

Model
A review of appropriate data-limited methods should be conducted as soon as possible, under the auspices of SEDAR, to allow evaluation of which methods should really be used in the DLM process for evaluation. Such a technical review would consider: (1) model theoretical basis and assumptions; (2) data requirements; (3) robustness of model to departures from assumptions and data requirements; and, (4) model responses (i.e., biases) to model uncertainty. This would include a systematic analysis of the sources of variability and how they influence OM dynamics. This was nearly impossible to discern in the way that the materials were presented at SEDAR 46, which was no fault of the analysts.

Some of the model estimates produced during SEDAR 46 were very troubling due to either: (1) application of an inappropriate or an inapplicable model(s) or MP; and/or, (2) very wide ranges of error variances, while unknown, that were applied to the input data. As a result, some MPs produced forecasts of unrealistic catch levels, suggesting that their usefulness is highly dubious. Not surprisingly, when appropriate variances and covariances were applied, the median of the output distribution do not change, but the range of model output metrics were substantially reduced. Nevertheless, that did not lead to any material change in the findings of the assessment with regard to MPs that performed better. The argument that this tested the MPs with greater uncertainty and therefore could still be used as a test of robustness was only partially accepted by the review panel.

While this AW was an examination of the potential efficacy of the approach due to its “newness”, and the fact that it was 3rd party application not fully controlled by the analytical team, we believe that in future workshops the analysts should more clearly specify what is desired as an outcome of model simulations, so that the simulations can be more finely tuned to answer specific questions. Generally, feasibility and limitations of MPs to real world applications is largely determined by data sufficiency and model adequacy. Additionally, there was no guarantee that the sampling algorithms in the OM reflected reality, and to some extent particular methodologies were difficult to assess given the information available to the Panel. In general, the AW would have run more smoothly if more attention were paid to the accuracy and precision of the basic data, and adherence to the assumptions required by the applicable MPs.
A better description and explanation of what is actually going on in the DLMtool OM at the outset would have been useful and clarifying to the Review Panel. As it was, application of methodologies at times appeared quite ad hoc, particularly as related to application of means, variances and coefficients of variations of model parameters. The parameters were treated as independent random variables, when we know they are dependent. But this is in fact the DLMtool default as it tries to cover a very wide range of uncertainties. There were a number of unclear definitions, such as “model stability”, which roughly translated to how many simulation runs were required for an input level of variation where for some unspecified reason, all model parameters seemed to be varying simultaneously. This would suggest that some further attention to model sensitivity is highly warranted. Concepts as straightforward as the number of required model runs to achieve stationarity were not well substantiated.

The apparent uncertainty in both data and models for U.S. Caribbean species suggests caution when selecting MPs intended to provide management advice. Selection of a particular MP for providing catch allocation strategies for management should consider: (1) MP sensitivity to parameters; (2) satisfying model assumptions; and, (3) information quality.

**Recommendations**

More precise and clearer descriptions and rationales for model thresholds and benchmarks used in the DLM process are needed. Analyses presented at the AW focused heavily on fishery yields (i.e., catches) which made it difficult to discern the rationale for what constituted a particular preferred choice of the MPs. A broader perspective might be entertained when setting OFLs and other appropriate benchmarks. This would likely include yield risks as they relate, in addition, to benchmarks specific to both economic and ecological risks. Adherence to this philosophy would require that model thresholds are set at more conservative resource use levels than are presently considered, and this in turn would avoid theoretical searches of infeasible or impractical model decision space. It is probably not useful to go too far into the weeds in trying to assess the full complexity of a fishery at first, rather the assessment needs to focus on distinguishing sustainable from non-sustainable rates of exploitation, and then identify the appropriate annual catches required to sustain the resource(s). If multiple MPs or a subset of tools are used, then some consideration must be given to model averaging. It would appear from the AW that many of the proposed estimation methods and MPs are non-starters from the outset. This seems an opportune time to conduct a thorough analysis of DLMtool efficacy. The Panel feels that the approaches presented could have broad potential for use in the Caribbean, but still require deeper, more thoughtful consideration to determine what avenues of application allow one to achieve the greatest utility of the tool.
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 assessment, both the process and findings, represents the best scientific knowledge about the stocks and their exploitation that is currently available. As this approach has been used for the first time for these stocks and given the data-limited nature of the stocks, there is clearly, additional work that needs to be done both on the data and model side to refine the approach. However, this assessment constitutes an improvement over previous approaches and has successfully made progress with developing scientific advice to support management.

The MSE approach that the assessment has adopted to overcome challenges associated with data-limited species is a relevant and transparent way to assess the performance of different management procedures. It has been used widely to assess fisheries and management approaches and its strengths and weaknesses are well documented (Holland 2010, Butterworth and Punt, 1999). The software used (DLMtool) and many of the MPs have been peer-reviewed and includes a wide range of methods that can be applied to data-poor species so, its choice is appropriate and relevant.

The assessment has made use of biological information compiled from a range of relevant studies to inform the selection of the model parameter values. It also considered both fishery dependent and fishery independent information to describe exploitation and made use of indirect ways to improve accuracy (expert knowledge, expansion factors, etc.). This reflects the team’s efforts to include all relevant information to respond to the knowledge gaps that characterize the stocks assessed. However, despite the considerable work done, the accuracy of input data remains low; the Panel has recommended additional research to improve the accuracy of data and/or the overall robustness of the analysis (see previous ToR).

The main source of information to define the values of biological parameters was peer-reviewed papers and both life history and fisheries information was reviewed as part of the Data and Assessment workshops which included scientists, fisheries experts, and fishermen so, there has been a good level of scrutiny. Furthermore, the assessment team conducted additional analysis to address key issues identified during the review meeting and those findings also increased the robustness of the overall approach.

Although some improvements have been recommended, the assessment captures the uncertainty in input parameters well and the metrics used were appropriate to reflect the level of uncertainty in the results.

The mean length estimator which was used in previous assessments was also one of the Management Procedures (MPs) included in this analysis and that maintains continuity. The
assessment also used a range of other methods to test their performance for the 6 species-island units and provide preliminary TAC distributions and that offers additional assurance and a comprehensive picture of the assessment options available.

As this was not a conventional stock assessment per se, it was not possible to produce all the management metrics that are often calculated in conventional assessments (e.g. \( B_{\text{MSY}} \)) and those that were calculated (i.e. TAC) were characterized by high uncertainty. However, the criteria used to assess the performance of different MPs and presentation of the outcomes were relevant and provided objective and robust insight that can inform management decisions.

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

The following is a list of key improvements for further development of the assessment methodology. These should lead to improvements in identifying the best performing Management Procedures (MPs). This recognises that the approach is a work-in-progress, and further development is desirable for Caribbean fisheries.

The main recommendation for the data is to ensure that sufficient data are collected to apply the data-limited MPs selected. Reasonably precise data, such as total catches and length sampling, are required to implement the recommended MPs. Current results suggest that TIP collection of species composition, lengths and CPUE will be most important to monitor these fisheries, while total catch data will be needed to implement the catch limits. For the modeling, the following recommendations are made for the next assessment.

- Strong correlations between parameters, notably \( L_{\infty} \) and \( K \), and the length-weight parameters \( a \) and \( b \), should be accounted for in the parameter density functions. Joint parameter probability density functions should result in projections that are less variable than currently simulated. For example, strong correlations that are known to occur between \( L_{\infty} \) and \( K \) could be parameterised in a bivariate normal, rather than treating these parameters as independent. This should provide better performance measures for identifying the best MPs.

- Projections need to be more constrained to reflect possible scenarios. Currently, some projections used to assess MPs would appear to be highly unlikely (e.g. projecting catches much higher than any previously observed). While it is important to measure MP robustness, and noting the performance measures themselves are robust to uncertainty, including excessive highly unlikely projections as part of that assessment could still distort the apparent performance of MPs. Improved parameter selection might be achieved by conditioning the operating model on the available past observations, adding a rejection probability to outcomes or improving the joint parameter probability densities as above.
• If the data or information that are required for an MP are not available or not reliable, the MP should be rejected at an earlier stage in the assessment. Including these MPs in performance reports, while assuming the information they require is known, may give a misleading impression from the results.

• The performance measure of Short Term Yield requested during the review should be used to evaluate MPs. More generally, performance indicators for the MPs should cover all the requirements of MPs as they are identified, so that MPs can be rejected based on performance criteria rather than for additional external reasons.

• MSE projection diagnostics should be routinely reported. The Review Panel requested example individual TAC projections. In addition, the range of key statistics from the MSE, such as TAC, biomass, mean length and fishing mortality, would be useful for review to check the projections are valid.

• The selectivity parameter, Lc, should be set to the mode of the observed length frequency by default. This would correspond to the point of full selection when assuming knife-edge selectivity.

• The simulated data for the management procedures should, by default, reflect the properties of real data (i.e. be imprecise and biased).

• Natural mortality estimates obtained from size dependent on age information should follow typical procedures (e.g., Then et al. 2015), but probably should not sample uniform-random around the mean of the probability distribution, since of maximum age means that animals live to no less than that particular age.

For the longer term, and not necessarily for the next assessment, the method might be enhanced, particularly for Caribbean fisheries.

• It should be possible to supply parameter vectors to the DLMtool operating model rather than parameters for parametric probability density functions.

• Alternative operating models to cover different life history characteristics should be provided. Specifically, sex differentiation in growth, and protogynous or protandrous hermaphroditism could be covered.

• The method to obtain a sufficiently precise estimate of the performance indicators should be made more efficient. The current number of projections is more than sufficient, but makes the assessments time consuming. The projection length and number of simulations should be tested to ensure they are as efficient as possible but sufficient for their use. This could be achieved by a statistical test for convergence at the start of simulations rather than relying on graphical output. Unless there is a need to contrast replicates (a random draw of time dependent parameters) with simulations (a random draw of all parameters), only simulations may be required, which again could increase the analysis efficiency.
• Although data limited methods provide an important transitional solution to management for sustainability of these presently “data-limited” fisheries, longer-term objectives should focus on improved accuracy and precision of the basic fishery catch-and-effort and length-structured abundance data and key demographic parameters (i.e., lifetime growth, lifespan and mortality, recruitment indices, etc.) and associated biological and economic information for the fisheries being assessed. The DLM tool provides a starting point for a “value of information” analysis that could be used to help identify priority research.

References