Memo

To: John Boreman, Chair, Mid-Atlantic Fisheries Management Council SSC

From: Thomas Miller, Vice-Chair, Mid-Atlantic Fisheries Management Council SSC

Date: September 12, 2015

Re: Review of McNamee et al "Data Limited Techniques for Tier 4 Stocks…."

The SSC Sub-committee appointed to review the McNamee et al. white paper on data limited (DL) approaches to establish catch advice for level IV stocks, with particular reference to Black Sea Bass, met at the Doubletree by Hilton Hotel at BWI Airport on September 10, 2015

The sub-committee comprised Thomas Miller (Sub-committee chair, MAFMC SSC & Chesapeake Biological Laboratory), Olaf Jensen (MAFMC SSC & Rutgers Univ.), John Wiedenmann (Rutgers Univ.), Katie Drew (ASMFC). Others attending the sub-committee meeting included Rick Robbins (MAFMC Chair), Rich Seagraves (MAFMC Council staff) and Kiley Dancy (MAFMC staff).and Kirby Rootes-Murdy (ASMFC).

Jason McNamee (RIDFW) and Gavin Fay (U. Mass-Dartmouth) presented the results of their application of the Caruthers (2015)[1] Data Limited Method tool (DLMtool) approach to establishing management reference points to provide catch advice for Black Sea Bass to the sub-committee.

The sub-committee reached conclusions on all six assigned Terms of Reference (ToR). The sub-committee want to thank the authors of the white paper for their open and constructive approach to discussions.

Background

The McNamee et al. white paper used the Caruthers (2015)[1] DLMtool box in R to develop reference points. DLMtool evaluates the performance of 47 different fishery management procedures in an operating model which is parameterized to represent a particular species defined by a suite of biological and fisheries related parameters. Many of the 47 different management procedures are alternative "flavors" of the same approach, only with slightly different parameterizations. The selected management procedures are evaluated against a set of user defined performance measures in a closed loop management strategy evaluation (MSE) that projects a population forward under a defined management procedure by sampling from distributions of biological, fishery and observation processes. The MSE assumes perfect implementation of each management procedure. From the output of the MSE, the management procedures that are

[1] Carruthers, T. (2015). DLMtool: Data-Limited Methods Toolkit (v1.35). An R-package. We note that the package has been updated and is now Carruthers and Hordyk (2015) DLMtool: Data-Limited Methods Toolkit (v2.0)

determined to perform “best” are identified. The values of these “best” management procedures are then estimated based on the real data. In our evaluation, the sub-committee will use the term management procedure for policies tested in the closed-loop MSE evaluation and reserve the term reference point for the specific values of the management procedures selected as “best” performing derived using empirical, observed data.

The McNamee et al. white paper applied the DLMtool box approach to Black Sea Bass. McNamee et al used the probability of overfishing < 0.3, the probability that the biomass will be less than 10% of the B_{MSY} < 0.2, and the relative yield should be > 0.5 as performance measures. The closed loop MSE evaluation was undertaken and a suite of “best” management policies identified. The reference points derived from these best management procedures were then estimated for Black Sea Bass using data from 1982-2014.

ToR 1: Are the DLM procedures included in the analysis suitable for the specification of reference points for black sea bass? What nature of reference point do each of the DLM procedures provide (i.e., OFL, ABC, etc.)?

The sub-committee concluded that the McNamee et al. white paper conflated the management procedures in the DLMtool box with references points that NOAA and the MAFMC risk policy would consider both overfishing limit (OFL) and allowable biological catch (ABC) reference points. The sub-committee notes that the mixing of these two categories of reference points leads to confusion in the application of the DLMtool to MAFMC managed species.

However, the sub-committee was able to separate management procedures that represented OFL-based approaches and those that provided ABC type reference points. The sub-committee recommend that if this approach is to be used for determining catch advice for other level IV species, clear separation should be maintained between the approaches that lead to OFL-based advice and those that lead to ABC-based advice. OFL-based reference points should all provide estimates of the same quantity from different approaches, and therefore should be of similar magnitude. Thus, ensemble based approaches could be used to provide a more reliable estimate of OFL. ABC-based reference points are all trying to estimate a sustainable catch level, but there is less reason to expect that individual estimates should be similar to each other. Whether ensemble-based approaches to estimating ABC reference points are appropriate has not been evaluated. An additional advantage of maintaining the distinction between OFL- and ABC-based reference points is that the two categories of reference points may provide an additional empirical check on the reliability of each because OFL estimates should be greater than the ABC estimates.

In the McNamee et al. application of the DLMtool box to Black Sea Bass, initial model runs often failed to converge or produced biologically infeasible estimates (e.g., negative values) for the approaches that generated OFL-based performance measures (e.g., DB-SRA) during the MSE evaluation step. McNamee et al. determined this behavior resulted from values of the steepness of the stock-recruitment function. A constrained and low value of h led to improved performance. However, the authors subsequently dropped from consideration many OFL-type management procedures because of concerns over the reliability of data used to parameterize them as reference points. As a result, by implementation but not by planning, most but not all of the reference points brought forward by McNamee et al. were ABC-based reference points.

ToR 2: Does the management strategy evaluation adequately represent the life history, population dynamics, and management of black sea bass?

There are two general approaches to undertaking MSEs for managed species. In the first approach, the operating model which represents the "true" biology, exploitation pattern and management of the population is tailored to represent reality as accurately as possible. In the second approach, a generalized operating model is developed which represents a broad universe of possible dynamics. The McNamee et al white paper uses this second approach. As a consequence, the operating model used in the McNamee et al. white paper includes neither the unusual life history, i.e., protogynous hermaphroditism, nor can it fully account for the range of hypotheses regarding population structure.

However, the sub-committee concluded that because of the broad spectrum of dynamics considered by the operating model, and because Black Sea Bass is a data poor, level IV species, it was unrealistic to require the operating model to explicitly include all features of the biology and exploitation of the species being assessed. But, as a consequence, the subcommittee also recognized that the limited realism that can be demanded of the operating model also means that managers should not expect extreme precision from reference points developed by the DLMtool box. Accordingly, managers should exercise caution in applying estimates that are "aggressive" or high when compared among ABC-based reference points.

The sub-committee concluded that the McNamee et al. model represented the biology and exploitation of Black Sea Bass adequately, and could be reliable for management purposes.

ToR 3: Are selected DLM methods parameterized correctly?

The sub-committee concluded that the McNamee et al model was appropriately parameterized, and included reasonable estimates of means and variances for most parameters in the population dynamics, exploitation and observation models.

The sub-committee was concerned over the low value of the steepness, h, of the stock-recruitment relationship. The authors told us that higher, and a wide variability of estimates of h were initially used in the MSE, but that many runs of the MSE failed when these higher, and more realistic values of h were used. Accordingly, the authors chose to tightly constrain h to low values to ensure convergence of model outputs. We return to concerns over this parameter and the analysts approach in our recommendations for ToR 6.

The sub-committee was also concerned about the relatively low values of the coefficients of variation for the fishery independent survey that was used in the modeling. The sub-committee was concerned that the low CVs used (CV ~0.2 – 0.4) mean that the fishery-independent surveys implied in the operating model are likely less variable than the real fishery-independent surveys. We return to concerns over this parameter and the consequence of this decision in our recommendations for ToR 6.

ToR 4: Are model selection procedures robust (i.e., are the methods recommended expected to perform well over a range of alternative states of nature)?

A wide range of diagnostics are available in DLMtool and were evaluated by McNamee et al. These diagnostics suggest, that with the exception of the level of depletion in the population that is assumed at the beginning of the exploitation period considered (i.e., 1982), the performance of the different management procedures were generally robust to the parameterization of the operating model.

The sub-committee noted that only three performance measures were considered in evaluating management policies. The sub-committee believe that additional and important performance measures, such as the inter-annual variation in population abundance, the inter-annual variation in catch advice, and the number of years in which fisheries closed should be considered. Moreover, the sub-committee noted that some of the performance measures that were used were not well defined in the white paper and required discussion with the authors, e.g., probability of overfishing, to fully clarify. The sub-committee also recommended that the authors more fully consider the performance of management procedures over alternative states of nature (e.g., flat and increasing effort). As an example of this issue, the sub-committee noted that the ABC recommended by McNamee et al. performed well under the broad "flat" effort scenario, but did not perform well in the "increasing" effort scenario.

The sub-committee recommend especially that a performance measure based on inter-annual variability be developed as a tool for selecting the management policies to be used to develop reference points.

The sub-committee concluded that the justification for recommending the reference point selected (e.g., GB-slope) was equivocal; other reference points appeared to perform similarly well and were more robust against alternative states of nature. The subcommittee concluded that application of a single method to estimate reference points may be less reliable than an approach that incorporates an ensemble of methods.

ToR 5: Do the model and data selected for its application to derive an ABC recommendation constitute best available science?

The sub-committee concluded that the methods developed by McNamee et al. using the DLMtool box are preferred to the current SSC constant catch-based approach. We draw this conclusion because the McNamee et al. approach allows the performance of alternative management procedures to be evaluated relative to each other, whereas the performance of the current SSC constant-catch approach remains unknown.

The sub-committee concluded, based on the evidence presented in the McNamee et al white paper, that three methods used to estimate reference points provide a reasonable foundation for providing an ABC for Black Sea Bass. The three methods are dynamic, catch-based procedures that combine an estimate of recent catch and a weighted estimate of the slope of the fishery-independent survey indices over the recent period. These approaches were selected because they are adaptive and rely on data that are routinely estimated and believed to be reliable. The approaches are:

1. GB-slope. This method is based on that recommended by Geromont and Butterworth (2015)[2]. In this particular application of the method, the most recent years catch is multiplied by the slope of the last five years of the NEFSC spring fishery-independent survey data. The slope is estimated by simple regression, and a stochastic draw is made based on the mean ± SD of the regression slope. Inter-annual variability in the resulting ABC is constrained to be ± 20% of the current years catch. The equation for GB-slope is

$$ABC_{t+1} = \begin{array}{c} catch_t * (1 + \lambda \cdot s_t) \\ if \ \frac{ABC_{t+1}}{catch_t} > 1.2 \rightarrow \ 1.2 * catch_t \\ if \ \frac{ABC_{t+1}}{catch_t} < 0.8 \rightarrow 0.8 * catch_t \end{array}$$

where λ is 1, and s_t is the slope of the NEFSC Spring survey over the preceding five years. In DLMtool the slope is calculated on untransformed index values, whereas Geromont and Butterworth recommend using log-transformed index values.

GB-slope was selected by McNamee et al as the best candidate for a reference point. However, GB-slope did not perform well in simulations that evaluated a pattern of increasing effort.

The value of the reference point estimated from GB-slope for 2016 is 3,797 MT. The sub-committee notes that this level of ABC represents approximately the 90th percentile of the observed catch time series – a value that may be considered too aggressive given the assumptions in the operating model.

As an example of the performance of the GB-slope reference point, the sub-committee calculated how the ABC based on GB-slope would have varied over time (Fig. 1). The sub-committee is concerned that this index may be too volatile from a socio-economic stance.

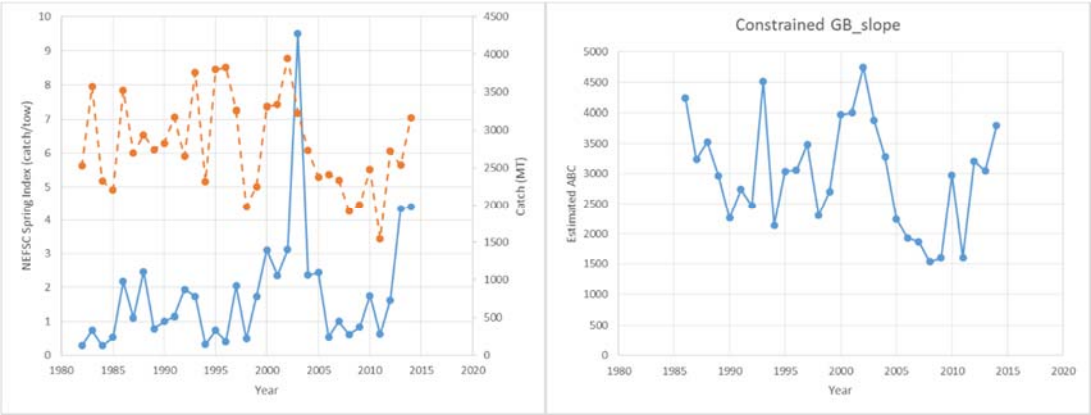

Figure 1. Time series of catch and the NEFSC Spring index (left panel) and the resulting GB-slope ABC proxy calculated from the algorithm used in DLMtool.

[2] Geromont, H. F. and D. S. Butterworth. 2015. Generic management procedures for data-poor fisheries: forecasting with few data. ICES Journal of Marine Science 72:251-261

2. iSlope1. This method is based on that recommended by Geromont and Butterworth (2015). It is very similar to GB-slope, except that the estimated slope of the fishery-independent survey is down weighted by λ=0.4. Similar constraints are applied in iSlope1 (i.e., ± 20% of the current catch) as for GB-slope.

   iSlope1 was not selected by McNamee et al, but the sub-committee noted that its performance was very similar to that of the preferred GB-slope method. Moreover, iSlope1 performed well on both the base runs and the runs simulating increasing effort.

   The value of the reference point estimated from iSlope1 for 2016 is 2,710 MT. This value is close to the average catch-based approach currently used by the SSC

3. SBT1. This is a third approach to estimating ABC that combines an estimate of the current catch and the slope of the fishery-independent survey time series.

   SBT1 was not selected by McNamee et al, but the sub-committee noted that its performance was very similar to that of the preferred GB-slope method. SBT1 performed well on the "flat effort" runs but not on the runs simulating increasing effort.

   The value of the reference point estimated from SBT1 for 2016 is 3,091 MT.

The sub-committee conclude that the most defensible approach to estimating the ABC for Black Sea Bass would be to use an ensemble of these three approaches. The sub-committee recommend the SSC consider ABCs based on one of two ensemble approaches

Simple average – this method uses the simple average of the three slope-based approaches. The value for 2016 is 3,199 MT.

Weighted average – this method increases the weight of the iSlope1 estimate because it is the only method that performed well in both the "flat effort" and "increasing effort" scenarios considered by McNamee et al. Specifically, we weighted estimates by the number of "states of nature" tables (e.g., Tables 2 and 3) in which the method was a preferred method. Thus, in this instance iSlope1 would receive twice the weighting of GB-slope and SBT1. For 2016, this weighted average is 3,077 MT.

ToR 6. What improvements to the analysis are recommended?

The sub-committee offers the following suggestions for future exploration of this approach for Black Sea Bass and for other level IV species. While this additional work would strengthen the analysis, the subcommittee agrees that the analysis as presented is acceptable for management use.

1. Explore a wider range of performance measures, in particular the sub-committee recommend considerations of performance measures that index inter-annual variability and that evaluate the magnitude of overfishing and overfished conditions not just their frequency of occurrence.
2. Re-run the MSE with the methods that caused the analyst to constrain steepness removed and re-evaluate performance. This should remove the need to constrain h.

The performance of all the management policies should be re-evaluated to determine if the slope-based reference points remain a preferred set.

3. Re-run the MSE with higher values for the CV of the fishery independent survey to ensure that the survey indices generated in the operating model have characteristics appropriate for the Black Sea Bass survey used to estimate reference points.
4. Explore the performance of slope based reference point methods when smoothed relative abundance indices are used.
5. Explore approaches to developing ensembles of management policies (e.g., what policies to include and how to weight them) and quantify the performance of these ensembles in the closed loop MSE phase of DLMtool.
6. Include catch-curve based management policies. These policies were not included in the current analyses because of concerns over the reliability of the age and length composition data available – however their performance should be evaluated.
7. Explore the consequence of imprecision in the implementation of management procedures on their performance. The current applications assumes that each MSE will be implemented perfectly which is unrealistic given experience.
8. DLMtool provides plots of each reference point and a Monte Carlo estimate of uncertainty. However, it would be helpful to managers to provide a time series of the annual estimates of each reference point by either plotting the time series of each management procedure in the MSE (e.g., Fig 1), or by plotting the point estimates of the reference points retrospectively. The sub-committee noted that the retrospective estimation will not represent the actual values that would have been used.

The SSC recognizes the strong work completed by McNamee et al. Indeed the sub-committee wishes to suggest that the SSC consider recommending that the Council seek opportunities to support further work on the DLMtool approach to parallel the work on a new analytical assessment for Black Sea Bass as insurance should that new assessment fail to pass peer-review.