

Table 1: Metrics considered when setting recreational measures under each option in this Draft Addendum/Framework. Primary metrics determine which harvest control rule bin a stock is in; secondary metrics are only used if, through the evaluation of the primary metrics, the stock stays in the current bin. Metrics considered through accountability measures may differ from those shown below. See section 3.1 for more details on the options.

Option	Metrics used to set measures					Measures are pre-determined	Expected number of sets pre-determined measures	Measures specified for 1 or 2 years
	Expected harvest*	Biomass compared to target level (B/B _{MSY})	Fishing mortality compared to threshold level (F/F _{MSY})	Recent recruitment	Biomass trend			
No action	Primary					No	N/A	1
Percent change	Primary	Primary				No	N/A	2
Fishery score	Primary**	Primary**	Primary**	Primary**		Yes	4	2
Biological reference point	Only when $F > F_{MSY}$	Primary	Primary	Secondary	Secondary	Yes	13	2
Biomass based matrix		Primary			Primary	Yes	6	2

*Expected harvest refers to expected harvest under status quo measures compared to the upcoming year(s)' RHL and could be based on past MRIP estimates, including consideration of confidence intervals for those estimates, or a model-based estimate of harvest, including considerations related to uncertainty in that estimate.

**As described in the Draft Addendum, the fishery score metrics may not be weighted evenly. The Monitoring/Technical Committees will recommend the appropriate weight for each metric. These weights can be modified through the specifications process.

Percent Change Option

Alternative considers future RHL, recent MRIP time-series average estimate, and the relationship of Biomass to Bmsy to determine what percent change should occur for management measures. Percent changes provide similar consideration for reductions and liberalizations.

1



RHL compared to MRIP estimate

Determine if the RHL for the upcoming management period is above, below, or within the confidence interval of the most recent MRIP time-series estimate.



2



Compare Biomass to Bmsy

Compare the Biomass estimate from the stock assessment to the biological reference point (Bmsy). Biomass categories are as follows:

- 150% above Bmsy
- between 100 and 150% Bmsy
- less than 100% of Bmsy



3



Find percent change in measures

The RHL and Bmsy comparison determines the appropriate management response. Measures will either be liberalized, restricted, or status quo. There are three different percentages by which measures can be liberalized or reduced.



4



Set Management Measures

Management measures are based on pre-defined % changes from the status quo.



Fishery Score Option

This infographic explains how the Fishery Score will function to select the management measure bin for summer flounder, scup, black sea bass, and bluefish

STEP 1

Stock Assessment Results

An updated stock assessment is completed and approved for management use.

STEP 2

Calculate Fishery Score Metrics

Fishing mortality, biomass, recruitment, and fishery performance metrics are drawn from the stock assessment and recent MRIP estimates.

STEP 5

Adapt New Measures if Needed

If the Fishery Score caused the stock to move from one bin to another, then the new pre-determined management measures will be applied. If the stock remains within the same bin, measures will remain the same.

STEP 3

Use Formula to Calculate Fishery Score

Fishery Score metrics are entered in the Fishery Score formula to produce a value ranging from 1 to 5. On this scale, 1 is the lowest possible score and 5 is the highest possible score.

STEP 4

Determine Management Step Based on Fishery Score

Based on the calculated Fishery Score, the stock is placed into one of four bins. Each bin has an associated level of concern, stock status, and a pre-determined set of management measures.

Fishery Score bins and the associated stock status, fishery performance outlook, and measures that are associated with each bin.

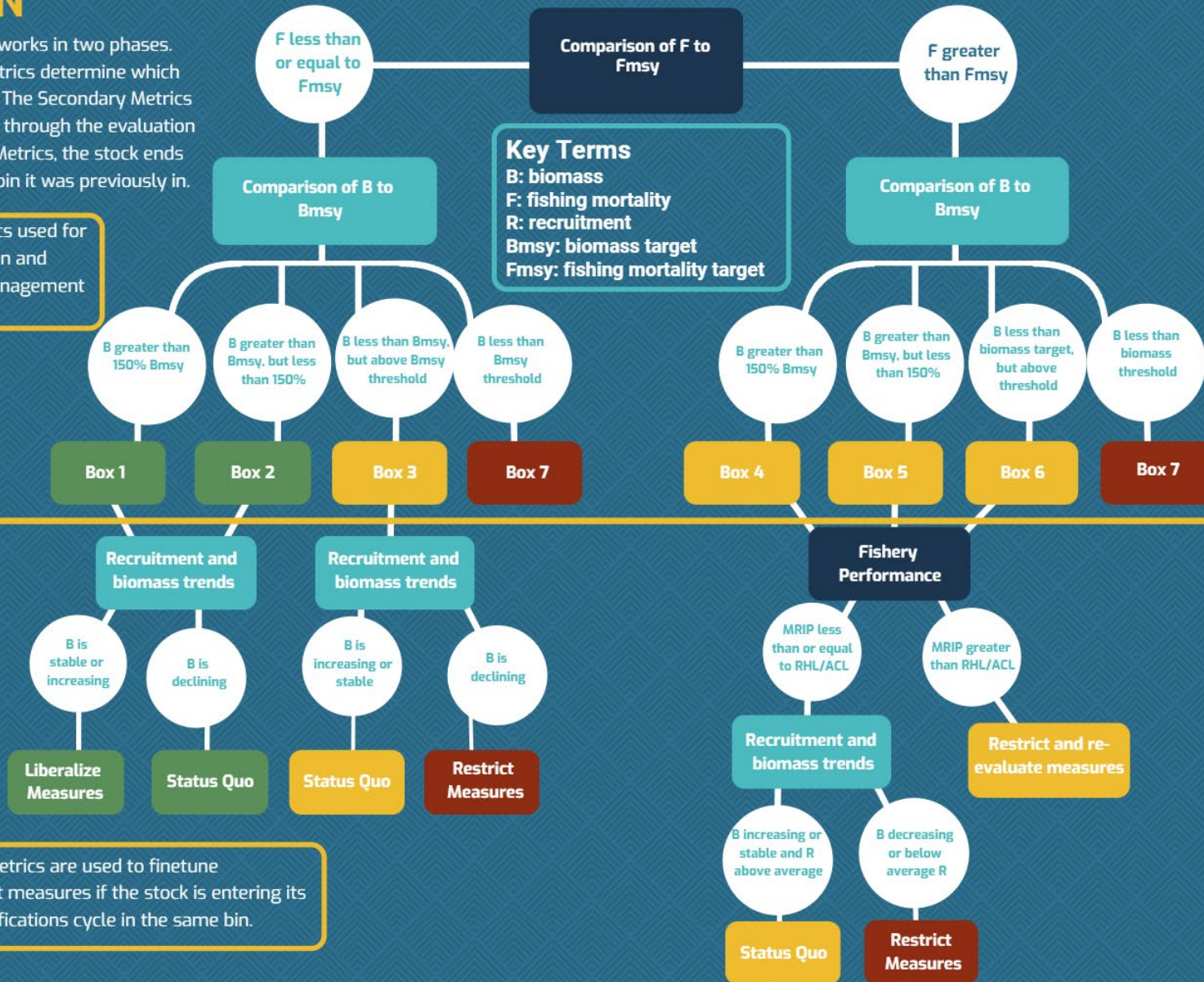
Bin	Fishery Score	Stock Status and Fishery Performance Outlook	Measures
1	4-5	Good	Most Liberal
2	3-3.99	Moderate	Liberal
3	2-2.99	Poor	Restrictive
4	1-1.99	Very Poor	Most Restrictive

BIOLOGICAL REFERENCE POINT OPTION

The BRP option works in two phases. The Primary Metrics determine which bin a stock is in. The Secondary Metrics are only used if, through the evaluation of the Primary Metrics, the stock ends up in the same bin it was previously in.

Primary Metrics used for determining bin and associated management measures

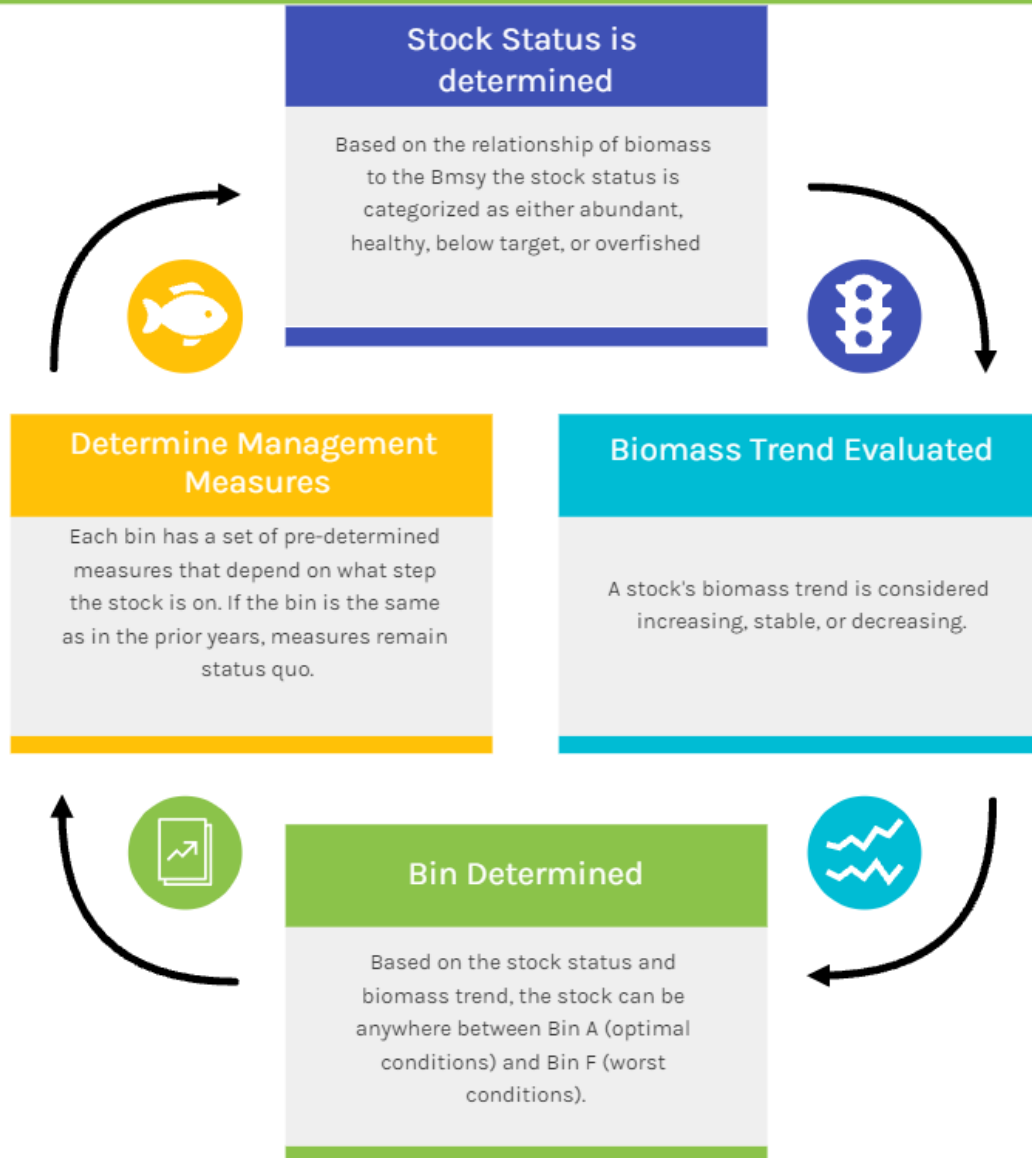
Key Terms
 B: biomass
 F: fishing mortality
 R: recruitment
 Bmsy: biomass target
 Fmsy: fishing mortality target



Secondary Metrics are used to finetune management measures if the stock is entering its second specifications cycle in the same bin.

Biomass Based Matrix Approach

Defines bin conditions based on two factors: stock status (i.e., biomass relative to Bmsy or proxy) and the most recent trend in biomass (increasing, stable, or decreasing). These parameters create a three-by-four matrix to determine which step is appropriate.



Recreational management measure matrix under the Biomass Based Matrix Approach

		Biomass Trend		
		Increasing	Stable	Decreasing
Stock Status	Abundant	Bin A		
	Healthy	Bin A	Bin B	
	Below Target	Bin C	Bin D	
	Overfished	Bin E	Bin F	

Mid-Atlantic Fishery Management Council

Sub-Group of the Scientific and Statistical Committee

Peer Review Report of Recreational Fishery Models

Jorge Holzer^{1,2}, Yan Jiao^{1,3}, and Cynthia Jones^{1,4}

¹Mid-Atlantic Fishery Management Council Scientific and Statistical Committee

²University of Maryland, Department of Agricultural and Resource Economics

³Virginia Tech, Department of Fish and Wildlife Conservation

⁴Old Dominion University, Department of Ocean and Earth Sciences

Executive Summary

The Mid-Atlantic Fishery Management Council (Council) convened a peer review panel consisting of members of the Council's Science and Statistical Committee (SSC) to review two potential recreational management models¹. A Recreational Fleet Dynamics Model (referred to in the report as RFDM) was developed by Dr. Jason McNamee (Rhode Island Dept. of Environmental Management, RIDEM) and collaborators Corinne Trusedale (RIDEM, Division of Marine Fisheries) and Savannah Lewis (Atlantic States Marine Fisheries Commission, ASMFC) for summer flounder and black sea bass. A Recreational Economic Demand Model (referred to in the report as REDM) was developed by Andrew (Lou) Carr-Harris (NMFS Northeast Fisheries Science Center) for summer flounder.

These two models are being considered for use by the Council's Fishery Management Action Team (FMAT) and the ASMFC's Plan Development Team (PDT) in the development and analyses of alternatives for the Council and ASMFC Recreational Reform Initiative². The potential use of these models would be part of the development of a Harvest Control Rule currently being considered as one component of the Recreational Reform action. The goal of the peer review was to help identify the potential utility, benefits, uncertainties, and limitations of each model for use by the FMAT/PDT during the Harvest Control Rule development and to provide any guidance as to whether these models represent an improvement to the current

¹ Dr. Lee Anderson from the SSC participated in the peer review meeting.

² For more information about the Recreational Reform Initiative, please see: <https://www.mafmc.org/actions/recreational-reform-initiative>

process and methods used by the Council and ASMFC technical groups to set recreational measures.

The peer review meeting was held on September 20, 2021 from 10:00 a.m. - 4:00 p.m. via webinar and was open to the public to listen in and ask questions. The agenda, meeting materials, and presentations can be found on the peer review meeting page at:

<https://www.mafmc.org/council-events/2021/ssc-peer-review-panel-sept20>.

Summary of Key Peer Review Conclusions and Recommendations

- Both models rely on MRIP data, though to a different degree, and are therefore subject to the limitations and uncertainties stemming from these MRIP data.
- The REDM is a simulation model which relies on the quality of the 2010 angler choice experiment data and population dynamics model. While the model has been properly specified and is sound, below we present some recommendations that may prove useful.
- The RFDM model selection process that the team has adopted is unclear and the model specifications need to be revised when considering space, time, population size and regulation variables. Also, it may be worthwhile to consider the correlation between harvest and discard when specifying and estimating the models.
- The review panel recommends that, upon implementing the revisions described below, both models be considered by the management/technical teams. For example, the performance of the RFDM could be benchmarked against that of the REDM for a couple of years. After the relevant improvements to the model coming from that process are implemented, the RFDM may be ready for use in fisheries for which bioeconomic models are unavailable due to a lack of angler preference survey data.

The peer review panel would like to thank Dr. McNamee, Dr. Carr-Harris, Ms. Trusedale, and Ms. Lewis for their commendable work and effort to develop these recreational models and for engaging in an open dialogue to address all of the questions asked by the panel. We would also like to thank Julia Beaty (Council staff) for her very valuable and informative presentation on the current process to set recreational management measures and an overview of the Recreational Reform Initiative.

Response to the Terms of Reference

The peer review panel addressed each of the Terms of Reference (*italics*) provided by the Mid-Atlantic Fishery Management Council and consensus responses (standard font) are provided below. Individual panel member reports that address the same Terms of Reference are provided as Appendix 1-3.

1. *Are the theoretical and statistical model specifications consistent with professional standards?*

a. Was the model's design and specification clearly described?

Both RFDM and REDM models are well described in the background documents provided to this panel. While the description of the model is fairly clear in both cases, it would be useful to more clearly spell out the role that MRIP data play in each model. As far as the panel understands, the catch data comes from catch-per-unit effort obtained from the on-site survey of anglers (AP AIS) wherein a survey agent often is able to examine the landings and the mail survey for effort (FES) wherein the household self-reports the number of trips that have been taken in each two-month wave. Additionally, discard information is obtained from anglers on site (AP AIS) who self-report the species and number of discards. When discards are high, uncertainty is present that is not accounted for in the modeling framework.

The RFDM model uses MRIP data disaggregated to the year, state and wave levels. The RFDM models are a set of regression models to estimate harvest and discard by fitting to the MRIP data. The explanatory variables considered include year, wave, space, regulation variables such as bag and size limits and number of days the season is open, SSB, and interaction terms. The name of the RFDM may be revised to reflect what was done since the model does not attempt to capture behavior by individual anglers or by the fleets.

The REDM model is based on data from a 2010 angler choice experiment survey, MRIP data, and the stock assessment results from a statistical catch-at-age model. The 2010 choice experiment survey provides data to estimate anglers' preferences and predict behavior under different regulations, fish caught and fish release across 4 survey regions: ME-NY, NJ, DE/MD, VA/NC. The anglers' estimated preferences are then coupled with a biological submodule that uses population projections from the most recent stock assessment. The model is currently simulated to match the number of summer flounder directed trips in 2019. In turn, that simulation results in a number of so-called choice occasions (i.e. each simulated instance in which an angler must decide whether to go fishing or to do something else). When projecting next year's recreational harvest, previous year's number of simulated choice occasions, recreational selectivity and catch per unit of effort were used. The REDM relies on data reported through surveying anglers on site. The assumption is that the data are representative of the general population of anglers. However, the data are likely overrepresent the most avid anglers. Below are suggestions on how to address this bias (see ToR #3b on page 8).

b. Are the underlying data sufficient to derive model estimates?

Both models treated MRIP data as true observations, so the model results can only be interpreted as such. Any bias in the MRIP data will be carried through in both models. These considerations notwithstanding, the MRIP data are the most complete time-series on recreational effort and harvest coastwide. In addition, the REDM also uses information collected from a choice experiment survey administered as a follow-up to the Access Point Angler Intercept Survey (AP AIS) conducted in 2010.

As mentioned earlier, the REDM selects anglers that were encountered in the on-site APAIS survey, and avid anglers are over-represented. Thus, a correction must be made to address the avidity bias that exists in the APAIS survey. Additionally, since the choice experiment survey was conducted in 2010, consideration might be given to conducting another choice experiment survey to reflect the current angler choices and preferences.

The population level projected recreational harvest relies on the number of calibrated choice occasions and recreational selectivity. The peer review panel suggests the role of the number of calibrated choice occasions may be evaluated in two ways: 1) calibrate the model using each of the individual past 5-7 years of data to see whether the resulting calibrated choice occasions are similar to the number currently used; 2) use the same number of the calibrated choice occasions from 2019 and the corresponding regulations for previous years to generate the number of trips in these years, and then compare them with the MRIP observed number of trips for a given year.

2. *How does the scale at which the model is operating (coast, regional, or state; wave or annual; fishing mode) affect the results?*

The RFDM model uses MRIP data disaggregated to the level of year, state and wave. As such, this model has the potential to provide wave-specific, year-specific, and state-specific harvest and discard estimates. However, it is worth highlighting that the MRIP survey was designed to have the lowest variance for species of interest when aggregated at the largest scale of region and year. When these data are disaggregated to state and wave, the variance increases resulting in wide confidence intervals. This may undermine the ability of the model to provide guidance for safe regulations that sustain the stock.

The REDM model is currently specified at the year and subregion levels (ME-NY, NJ, DE-MD, VA-NC) because the 2010 choice experiment survey was conducted at that subregion level. The model can readily be specified at the wave level by specifying wave-specific catch-at-length and number of fish caught per trip distributions. Disaggregation at the state level -to capture heterogeneity of anglers' preferences- would ideally entail an update of the angler preference survey. Alternatively, and given that the 2010 survey collected information on respondents' demographics, the utility function could be specified as a function of demographic characteristics (e.g., through the opt-out), which would then allow the simulation model to use state-level anglers' characteristics to predict impacts of management changes on effort, harvest, and welfare at the state level.

a. *How does data availability, uncertainty, and variability affect model results, interpretation, and application?*

Both models rely on MRIP data, so if MRIP overestimates/underestimates the recreational catch and discards, both models will be impacted. In the REDM model, the calibration data come from MRIP in the same year. In prediction years, catch-at-length is derived from recreational

selectivity, calculated from the calibration year, and projected population numbers at length. On the other hand, the RFDM relies exclusively on MRIP data for the estimation of the policy impacts on harvest and discards. The MRIP catch data includes the CPUE of landed fish which may have been observed by the survey agents but also the CPUE of discarded fish that is self-reported. These self-reported data may exhibit digit bias and can also be misidentified. Digit bias occurs when anglers don't keep track of regulatory discards or catch many of these fish and tend to estimate the number of discards, usually by stating common numbers, say 5, 10, 15 but not actually directly having counted discards. When discards are a small proportion of landings this may not result in much concern. For species with a large proportion of discards such as bluefish, when slot limits or other size limits are imposed, this issue is a cause of concern in the RFDM, which relies on self-reported discards.

b. What key assumptions affect the underlying statistical analysis and interpretation of the results? Were these assumptions and relevant uncertainties identified and characterized?

Both models rely, though to a different degree, on MRIP data. The REDM also uses a choice experiment survey to estimate anglers' preferences. The less reliable the data used are, the larger the uncertainty will be around the predictions that the models generate. All the relevant caveats regarding MRIP data apply here. Additionally, in the case of the choice experiment survey, the possibility that the population of respondents may not be representative of the general angler population (e.g., due to response bias or avidity bias) should be considered. As discussed in this report, there are alternative ways to address this concern.

Regarding the models' assumptions:

The RFDM assumes that the harvest and discard components of catch for a given species are independent and specified as separate and independent equations, which may have an impact on the uncertainty bounds around the predictions of the reduced-form model. Accounting for correlation of the error term across equations, which is likely to exist since both equations are dealing essentially with the data from the same fishing trips, may increase efficiency, and thus reduce the uncertainty bounds around the predictions. Moreover, as stressed earlier, the assumption of this model that, whatever the management measures for black sea bass, there will be zero effect on the harvest and discards of summer flounder, seems untenable given that these species are typically caught together in the same trips.

Additionally, in the RFDM, the model selection process is based on AIC and p -values but is not well described and not consistently employed. Of concern is the fact that the RFDM uses different sets of policy variables (e.g., bag limit, size limit, length of the season) in the harvest and discard equations. These management measures impact the fishing trip and should therefore be included in both equations (is it reasonable, for example, to assume that closing the season for summer flounder would only impact harvest but leave discards unchanged?). Moreover, the

partial effects of the bag limit and size limits have counterintuitive signs in some of the specifications presented, particularly for black sea bass. The selection of interaction terms does not seem reasonable either. Additionally, the year effect is not correctly specified (treated as numerical but should be categorical). In sum, the team may want to consider revising their model selection approach and how some of the variables are treated.

The REDM explicitly uses anglers' preferences to determine how different sets of management measures, through their impact to keep or release, will impact anglers' effort and welfare. Thus, estimating anglers' preferences correctly is important for the performance of this model. The authors of the REDM have carefully specified an indirect utility model that accounts for angler heterogeneity, and the estimated coefficients have the expected signs and result in willingness-to-pay for kept and released fish that are consistent with similar estimates in the literature.

Regarding the models' predictions:

The team that developed the RFDM presented, in Figures 8, 9, 14 and 15 of the background documentation within-sample predictions for the entire coast (rather than at the state level). Absent out-of-sample predictions at the state level, the review panel is unable to assess the ability of the RFDM to predict the impact of management changes. Likewise, it was suggested that the team shows what level of harvest and discards the model would project if the fishery were closed (either through a zero-bag limit or a zero-day season). This is relevant, as a fishery closure could be a management option and the models should be able to predict zero harvest for a complete fishery closure.

The author behind the REDM presented predictions at the state level for 2019. While the prediction of impacts on harvest and discards for the region were very close to the actual outcome, as expected predictions at the state level sometimes overestimated and others underestimated these impacts. It is suggested the author predicts additional years (i.e., out-of-sample predictions) to further assess model performance. Likewise, it may be worthwhile exploring how changing the calibration year (i.e., the baseline number of choice occasions) may impact the model's ability to predict policy impacts on harvest and releases.

3. *Is the model appropriate for estimating and predicting the impacts of bag, size, and season limits on recreational catch or harvest? Are the methods in the Recreational Economic Demand Model appropriate for estimating changes in recreational effort or fishing demand?*

a. *Does the modeling approach represent an improvement over current methods used to estimate impacts of management measures?*

Yes, both methods, when revised, have improvements over the current methods used to estimate the impacts of management measures. They both provide methods to evaluate changes of single or multiple factors simultaneously either based on statistical relationships (RFDM) or based on a

simulated process model (REDM). The models' ability for dealing with more than one regulation change should be better than the current method.

The RFDM is built on appropriate statistical methods to evaluate harvest and discard changes in response to a combination of alternative policies. It uses a general additive model (GAM) and has penalties for overfitting. It provides a model-based approach to evaluating impacts of regulations on harvest and discards. Separate models were built for harvest and discard for summer flounder and black sea bass, as though harvest and discard are independent from each other. Upon revision and further refinements, this model would provide a statistical evaluation of proposed harvest control rules that are currently done in an ad hoc manner. Moreover, multiple regulation changes could be evaluated simultaneously.

The REDM uses appropriate methods for estimating changes in the recreational effort and welfare by simulating scenarios with alternative management regulations. This model is built on well-established, peer-reviewed methods for economic utility models. This model combines MRIP-based data with the results from an economic choice survey in a simulation framework to evaluate alternate regulatory scenarios for the harvest, release, and likelihood of taking a fishing trip for summer flounder and black sea bass and alternate harvest target species. It links the behavioral and biological components and is designed to estimate changes in recreational effort, fishing demand, and angler welfare.

- b. What are the strengths/limitations of the modeling approach for informing management measures, especially at the regional, state, wave, or mode level? Are there specific recreational fishing measures for which use of the model would not be recommended?*

The RFDM model is constructed to inform management measures at the regional, state, and wave level based on the past calibrated MRIP records. However, the current model selection and model construction have problems and need to be revised before being used to inform management measures. Its strengths are that it uses the recalibrated MRIP time series and the model can be used to evaluate its efficacy based on how well it reflects the outcomes of historic regulations.

One RFDM weakness, as currently configured, is that the model includes Wave as a model component that is smoothed. Because Wave is actually a categorical variable, this is an inappropriate specification, and it should be used as a categorical variable without smoothing. Depending on the number of points added by smoothing, the variance associated with this model component may be underestimated. The model is currently fit to all the available data, but a better practice is to fit the model to a portion of the data and to test for fit against the remaining data portion. Another potential weakness is that harvest and discard models are independent, whereas these quantities are not independent of one another. Moreover, because summer flounder and black sea bass are often caught together, there is a good motivation to also link species.

The REDM's strengths are that it uses choice simulations specific to regional, state and wave tailored to targeted species in these areas and times and based on regulations for bag and size limits. The model can project future behavioral responses to regulatory modifications based on past years' performance, under the reasonable assumption that behavior and preferences won't show radical change over short time periods. It is powerful because it is based on MRIP access-site interviews wherein catches were observed but also on a subsequent choice survey of these same anglers.

The REDM's weakness, as with any model relying on economic add-on surveys taken on site, is that it doesn't sample the full frame of marine recreational anglers in these regions. In relying on the on-site contacts of anglers to whom surveys were subsequently sent, it over-samples avid anglers from the entire population of marine anglers. Anglers who fish more frequently have a higher probability of being sampled. While MRIP provides the correct estimate of harvest and discard, it doesn't adequately represent the regulatory preferences of the full marine-angling community. This can be corrected by weighting avidity frequencies available through the Fishing Effort Survey. Alternatively, anglers' preferences (i.e. the opt-out) may be specified as a function of demographics (since the 2010 choice experiment survey collected this information from respondents). Moreover, there was considerable self-selection and non-response (~68%) to the choice survey, that should be addressed, if possible. While 2019 simulation estimates approximate MRIP catch at the regional level, harvest or discard estimates of summer flounder at the state level exhibited different degrees of discrepancy with the actual data: harvest for New Jersey, Connecticut, Massachusetts, and Maryland (which jointly contributed 55% of summer flounder recreational landings in 2019) are predicted with less than 5% error, but the discrepancy is larger for Delaware, New York, North Carolina, Rhode Island and Virginia. Lack of fit was also seen for black sea bass for some states, though the model predicts well the harvest in New York (whose contribution to the total black sea bass harvest in 2019 was 36%). The model also assumes 100% compliance to regulations and the peer review suggests the authors consider incorporating noncompliance behavior into the model once reliable estimates of noncompliance become available.

- c. *What are the implications of using the model to predict future catch/harvest based on historical data? Are there limits on the magnitude of change in catch/harvest or stock status beyond which use of the model would not be recommended?*

The RFD model can also provide estimates of uncertainty about its predictions. However, as input MRIP data are disaggregated to year-state-wave estimates, the smaller unit survey sample sizes upon which the predictions are based will increase uncertainty. It is expected that the model will provide the most precise estimates at higher levels of aggregation. The background documentation provided to the panel, however, does not include out-of-sample predictions to assess the predictive power of the model beyond the coastwide aggregation.

In recalibrating the MRIP dataset, most species effort data converge to the old MRFSS data that relied on telephone surveys for effort before telephone surveys became unreliable. This is not true for bluefish and this species could be problematic. The issue of convergence can provide a guide for use with appropriate species to apply this model.

The REDM is based on a long time series of harvest and discard estimates from the MRIP and as the model is revised its performance can be calibrated against this time series using the scenario of appropriate state year-specific regulations. The panel has suggested modifications that may improve fit, such as correcting for avidity bias. The panel also wonders how the number of calibrated choice occasions (i.e., currently obtained by calibrating the model to match the number of trips in 2019) will affect projected effort and harvest.

- d. Can the modeling approach support development of multi-year bag, size, and season limits? If so, what criteria should be applied or developed to assess the reliability of the multi-year projections?*

The RFDM could prove valuable in providing guidance for multi-year bag, size, and season regulations upon further model revision and development as suggested in this document. Further development that links the RFDM and the REDM could provide valuable guidance that encompasses not only predictions of harvest and discard under regulatory scenarios but could also include measures of angler participation. This would be an important advance to management.

The REDM has the potential of providing guidance on the selection of multi-year bag, size and season limits upon revision. When the model was calibrated against 2019 estimated harvests and discards, it showed appreciable differences for some of the states. Upon revision, the model fitting is likely to improve and prediction error to decrease to better inform managers of the uncertainty of predictions.

- 4. Provide guidance for the following future model use considerations:*

- a. Could the model be modified to incorporate other species (e.g., scup, bluefish)?*

Yes, both models are set up to be modified and incorporate other species; however, there are likely species-specific data considerations depending upon the model. Given the existing data availability, both models could readily be applied to scup but may require additional data and/or analysis for bluefish. For example, the 2010 choice experiment survey did not include bluefish. This information underlies the angler preference estimation in the REDM and a new survey that includes bluefish would need to be conducted. In addition, the recalibrated MRIP data for bluefish show an increasing trend in discards that is now equal to the recreational harvest. Discard estimates are generated from self-reported information and are therefore more uncertain. Given the interaction between harvest and discards on a fishing trip, model parameterization and estimations should consider these trends and uncertainties in the underlying MRIP data.

b. Could future model runs be conducted by other individuals (e.g., Council/ASMFC staff or Monitoring/Technical Committee members) without major modifications?

Yes, both models are currently constructed to allow other technical staff/members to run the models without major modifications. Since most technical staff/members do not have an economic background, the REDM may require some additional training to fully understand bioeconomic models and stated preference techniques. The peer review panel also notes that full documentation as to how both models were revised and/or addressed peer review recommendations is needed prior to other technical staff/members running the models.

c. How easily could the model be updated with additional years of data or additional variables?

Both models can easily be updated with additional years of data and additional variables. The REDM may take longer to update and may necessitate additional or updated surveys to obtain information on changes in angler preferences, particularly as species distribution and availability changes.

Mid-Atlantic Fishery Management Council Sub-Group of the Scientific and Statistical Committee

Peer Review Report of Recreational Fishery Models

October 10, 2021

Individual Peer Review Report:
Dr. Jorge Holzer, University of Maryland

In addition to the comprehensive consensus report developed by the peer review panel, each member developed an individual report with detailed responses to each Term of Reference (*italics*) provided by the Mid-Atlantic Fishery Management Council and the responses (standard font) are provided below.

Response to the Terms of Reference

- 1) *Are the theoretical and statistical model specifications consistent with professional standards?*
 - a. *Was the model's design and specification clearly described?*

“The recreational Fishery Fleet Dynamics Model” (referred to below as Reduced-Form Model)

The title of the paper describing this approach is misleading as the model does not attempt to capture behavior by individual anglers’ or by the fleet of charter and party boats. Absent a module explicitly modeling fleet dynamics, it was suggested the name given to this model be revisited to better reflect what the approach is doing.

The model was well-described, but the model selection process (the specification of the equations finally selected as the preferred model) is unclear and seems somewhat ad hoc. In other words, it is unclear how the authors arrive at their preferred specifications in Tables 5-8. The reviewers highlighted the fact that selecting the models based on the Akaike information criterion (AIC) is not a good strategy in this context, given that the differences in AIC between the models considered are immaterial. Likewise, some of the claims regarding the partial effects, particularly those corresponding to the policy variables (i.e., bag and size limits) are not substantiated by the model results or by the explanations provided during discussion. In particular, the counterintuitive effects of the bag and size limits in the black sea bass harvest model (they have the opposite effect of the expected effect of these policies), are concerning. It is recommended that the authors look at alternative specifications that ensure the partial effects of the policy variables are of the expected sign.

“Recreational Fluke MSE Economic Modeling Overview” (referred to below as Structural Bioeconomic Model)

The model design and specification were clearly described. A detailed exposition of the two components of the model: i) the estimation of anglers' preferences module, and ii) the fishery simulation module, was provided by the author during the panel review presentation.

b. *Are the underlying data sufficient to derive model estimates?*

While more and better data is always welcomed, each model relies on the amount of data that allows it to derive empirical estimates. Importantly, the data requirements and capabilities of the two models are very different. The reduced-form model is essentially a curve fitting exercise which seeks to predict harvest and discards under different policy scenarios using only MRIP data. As such, that model is unable to predict changes in effort or angler welfare. The structural bioeconomic model, on the other hand, explicitly models the angler's behavioral response to alternative regulations. Thus, this model can predict not only changes in harvest and discards, but also changes in effort level and angler satisfaction (i.e., welfare). This is important as it would allow the Council to choose combinations of management measures that, conditional on achieving the conservation goals, optimize the economic efficiency of the fishery. The difference in capabilities between these two models, however, come at a cost, namely, data requirements and model complexity, which are higher for the structural model. Additional points raised during the peer review are discussed below:

The recreational Fishery Fleet Dynamics Model" (Reduced-Form Model)

The model currently assumes that, for each species, the harvest and discards equations are independent. It was noted that this is not necessarily the case as harvest and discards for a given species essentially correspond to the same trips and anglers. In these circumstances, the error terms of the two equations may be correlated. As such, joint estimation of the system comprising the two equations may result in more efficient estimates. In turn, efficiency will be important when deriving uncertainty bounds around the predictions that will be used by the Council to study the effects of management changes. Furthermore, since anglers typically catch summer flounder and black sea bass together and is the total number of kept and released fish of both species (summer flounder and black sea bass) that determines angler satisfaction, it is expected that changes in management affecting one of the species may have an impact on the harvest and discards of the other species through the effect on effort. Thus, if feasible, it may be worthwhile exploring the possibility of estimating the entire system of equations jointly.

Additionally, the peer review committee stressed the fact that the policy variables included as explanatory variables in the harvest and discard equations should be the same in both equations. The rationale is simply that those management measures regulate the fishing trip, and therefore, impact both harvest and discards. In other words, it makes little sense, for example, to expect a change in the number of days the season is open, to affect only the harvest of black sea bass and not the discards as well. Thus, the review panel suggested the authors explore alternative specifications with the same policy explanatory variables in both equations, the harvest and discard equations.

“Recreational Fluke MSE Economic Modeling Overview” (Structural Bioeconomic Model)

The panel highlighted the importance for the overall performance of the model of getting anglers’ preferences right. These preferences are estimated using the data from the choice experiment survey. Thus, modelers should spend time exploring credible alternative specifications of the indirect utility. The author of this model has done a nice job, but he may want to consider exploring other specifications. On a related topic, avidity bias may play a role here as survey respondents are typically more avid than the average angler. However, since the survey collected avidity and other demographic information, it was suggested that one way to address the possibility of avidity bias is to model the opt-out option in terms of avidity and other demographic information of respondents. Then, in the simulations, the opt-out can be adjusted to the relevant population by using that population demographic characteristics.

- 2) *How does the scale at which the model is operating (coast, regional, or state; wave or annual; fishing mode) affect the results?*
 - a. *How does data availability, uncertainty, and variability affect model results, interpretation, and application?*
 - b. *What key assumptions affect the underlying statistical analysis and interpretation of the results? Were these assumptions and relevant uncertainties identified and characterized?*

Both models naturally rely on data, and thus unavailability of data would undermine or prevent their use. Moreover, the less reliable the data used are, the larger will be the uncertainty around the predictions that the models generate. An advantage of the bioeconomic model over the reduced-form model, however, is that it explicitly characterizes the trade-offs faced by anglers and their expected behavioral response. In these circumstances, model results are easier to interpret intuitively. This feature may be important in discriminating between plausible and implausible outcomes when analyzing predictions, especially when data is scarce.

As in the discussion of ToR1, the assumption of independence of the harvest and discards equations for a given species, may have an impact on the uncertainty bounds around the predictions of the reduced-form model. Accounting for correlation of the error term across equations may increase efficiency. Moreover, as stressed earlier, the assumption of this model that, whatever the management measures for black sea bass, will not affect harvest and discards of summer flounder seems untenable given that these species are typically caught together in the same trips.

As for the resolution of the models, the structural bioeconomic model can predict the impact of management changes at the regional and coast level, and the wave or annual level. During the presentation, results were shown at the regional and coast levels for year 2019. On the other hand, the reduced-form model has the potential to predict the impact of management changes at the state, regional and coast levels, and wave or annual levels. The models do not currently provide predictions disaggregated by fishing mode.

The author of the structural bioeconomic model showed state-level predictions for 2019. It was suggested by the panel that the author presents out-of-sample predictions (i.e., for years prior to 2019). The authors of the reduced-form model, on the other hand, showed in-sample predictions at the entire coast level in Figures 8, 9, 14 and 15. The authors were asked to provide out-of-sample predictions at the state level (drop some data, re-estimate the model, compare the prediction of the newly estimated model for the period of data dropped with the actual data, and calculate the square prediction error; when this process is repeated many times, the mean squared prediction error can be used for model selection)¹, which is the level at which bag and size levels are typically set. Absent these out-of-sample predictions at the state level, it is not possible to assess how well the models predict the impact of changes in policy on harvest and discards. In this same vein, it was requested that the authors of the reduced-form model show what level of harvest and discards the model would project if the Council closed the fishery (either through a zero-bag limit or a zero-day season). This is relevant, as a fishery closure should always be in the regulator's tools box and the models should be able to predict zero harvest for a complete fishery closure. As shown during the presentation, the structural bioeconomic model can predict zero harvest associated with a fishery closure, as expected.

3) *Is the model appropriate for estimating and predicting the impacts of bag, size, and season limits on recreational catch or harvest? Are the methods in the Recreational Economic Demand Model appropriate for estimating changes in recreational effort or fishing demand?*

a. *Does the modeling approach represent an improvement over current methods used to estimate impacts of management measures?*

Yes, both models represent an improvement over the current methods as they bring structure and statistical methods to the analysis of alternative policies on harvest and discards. Regarding the economic demand model, it is a structural model that links the behavioral and biological components and is designed to estimate changes in recreational effort, fishing demand, and angler welfare.

b. *What are the strengths/limitations of the modeling approach for informing management measures especially at the regional, state, wave, or mode level? Are there specific recreational fishing measures for which use of the model would not be recommended?*

The structural bioeconomic model is a more powerful model as it uses anglers' preferences to characterize the effort response to changes in regulation, and from that response predicts harvest, discards, and anglers' welfare. The reduced-form model is unable to characterize the trade-off anglers face and therefore is unable to predict changes in effort and anglers' welfare. However, as indicated earlier, this model requires less data and can be updated much more quickly than the

¹ *k*-fold cross validation has been suggested as a model selection algorithm robust to overfitting since at least Stone (1974). Stone, M. (1974). "Cross-Validatory Choice and Assessment of Statistical Predictions." In: Journal of the Royal Statistical Society. Series B (Methodological) 36(2), pp. 111–147.

structural model. Thus, if the authors of the reduced-form model can show specifications that provide good out-of-sample predictions at the state level (as requested by the peer review panel), then the model could be used for analyzing fisheries for which there is currently no survey data available to populate a structural model. Moreover, it was suggested that the structural model applied to black sea bass and summer flounder can be used as a benchmark for the reduced-form model. Under this strategy, that model could be improved to try to match the predictions of the bioeconomic model, and after that it could be used in fisheries for which there is not enough data to develop a structural bioeconomic model.

- c. *What are the implications of using the model to predict future catch/harvest based on historical data? Are there limits on the magnitude of change in catch/harvest or stock status beyond which use of the model would not be recommended?*

As explained earlier, it is unclear how either model predicts out-of-sample. In the case of the reduced-form model, it is also unclear how it performs in predicting unusual years at the state level. When both teams provide the corresponding predictions, it will be possible to assess the models' predicting power. However, as highlighted above, the reduced-form model seems unlikely to predict a fishery closure satisfactorily, and the partial effects of the bag and size limits currently have counterintuitive signs. From the evidence provided to the review panel, the structural bioeconomic model seems better equipped to provide good predictions at the state level.

- d. *Can the modeling approach support development of multi-year bag, size, and season limits? If so, what criteria should be applied or developed to assess the reliability of the multi-year projections?*

While both models could produce multi-year predictions to inform the setting of multi-year management measures, this strategy is not recommended at this stage. In the view of the review panel, the models should be first used to predict changes year by year first, and only after satisfactory performance should they be used to recommend multi-year management measures.

4) *Provide guidance for the following future model use considerations:*

- a. *Could the model be modified to incorporate other species (e.g., scup, bluefish)?*

For the case of scup, the answer is yes for both models. However, the choice experiment survey that underlies the angler's preferences estimation in the structural model, which was conducted in 2010, does not include bluefish. Including bluefish into this model would require a new survey, which would require time to design and conduct. On the other hand, the reduced-form model can easily and readily incorporate new species, including bluefish (especially under the current assumption of independence of harvest and discards across species).

- b. *Could future model runs be conducted by other individuals (e.g., Council/ASMFC staff or Monitoring/Technical Committee members) without major modifications?*

The answer is yes for both models, but the structural model is more complex and requires understanding of bioeconomic models and of stated preference techniques (i.e., random utility

models). As such, it would take longer to train a new person to run and update this model.

- c. *How easily could the model be updated with additional years of data or additional variables?*

Both models can readily be updated with more data and variables, but it would take longer time to update the structural bioeconomic model than the reduced-form model.

Mid-Atlantic Fishery Management Council Sub-Group of the Scientific and Statistical Committee

Peer Review Report of Recreational Fishery Models

October 9, 2021

Individual Peer Review Report:
Dr. Yan Jiao, Virginia Tech University

In addition to the comprehensive consensus report developed by the peer review panel, each member developed an individual report with detailed responses to each Term of Reference (*italics*) provided by the Mid-Atlantic Fishery Management Council and the responses (standard font) are provided below.

Response to the Terms of Reference

- 1) *Are the theoretical and statistical model specifications consistent with professional standards?*
 - a. *Was the model's design and specification clearly described?*

The design and specifications of both the Recreational Fleet Dynamics Model (RFDM) and the Recreational Economic Demand Model (REDM) are clearly described. Both models used MRIP data and both models treated MRIP data as true observations.

The RFDM models are based on the MRIP data disaggregated to the level of year, state and wave. The RFDM models are a set of regression models to estimate harvest and discard by fitting to the MRIP data and the variables considered in the models include YEAR, regulation variables SSB and some interaction terms.

The REDM model is based on data from a 2010 angler choice experiment survey, the MRIP data, and the stock assessment results from a statistical catch-at-age model. The 2010 angler choice experiment survey provides data to evaluate the angler behavior under different regulations, fish caught and fish release across 4 survey regions (ME-NY, NJ, DE/MD, VA/NC). The angler's choice estimated given regulations conditions is integrated in the population projection based on the most recent stock assessment specification with the recreational harvest simulated based on a process model of the angler's choice. When projecting next year's recreational harvest previous year's # of simulated choice occasions, recreational selectivity and catch per unit of effort were used.

- b. *Are the underlying data sufficient to derive model estimates?*

Both models treated MRIP data as true observations, so the model results can only be interpreted as such. Any bias in the MRIP will be carried on in both models.

The RFDM model may be revised to consider estimating trips under various regulation situations and population sizes. Such estimate may be used to provide input for the REDM model.

The angler choice survey was done in 2010. The council may consider another up-to-date survey to reflect the angler choice and may also look into the stakeholder types and preferences without responses.

The population projected recreational harvest relies on the # of simulated choice occasions and recreational selectivity. Both of them may be verified by comparing the past 5-7 year data to see whether using results from previous years are robust or not especially the # of simulated choice occasions which directly decide the # of trips in the projected year.

2) How does the scale at which the model is operating (coast, regional, or state; wave or annual; fishing mode) affect the results?

The RFDM models are based on the MRIP data disaggregated to the level of year, state and wave. I feel the scale that the RFDM is operating is appropriate based on its purposes. This model can provide wave-specific, year-specific and state-specific harvest and discard estimates and can meet the model for the need of both monitoring and recreational regulation considerations.

The REDM model is at the time step of 2 months and the state-specific results are reported but the angler preference survey is based on 4 coastal survey regions. It does not function to provide suggestion on the monitoring and do function to simulate the potential changes in harvest given regulation changes although I have concerns on how the # of trips is simulated.

a. How does data availability, uncertainty, and variability affect model results, interpretation, and application?

Both models treat MRIP as true observations, so if MRIP overestimates the recreational catch and discard, both models will do so correspondingly. The REDM model only considered the surveys from the anglers responded. These concerns are understandable given the data availability but their influence on the recreational harvest and discard prediction may be explored through sensitivity analysis and extra add-on surveys such as a new angler preference survey. Both models can provide probabilistic estimates of the results of interests.

b. What key assumptions affect the underlying statistical analysis and interpretation of the results? Were these assumptions and relevant uncertainties identified and characterized?

In the RFDM model, the variable selection process is based on AIC and p value but are not well described and the results don't seem to match what was used in variable selection. The Year effect is treated as linear which is of high concern both statistically and biologically. The effects

of bag limit and size limit do not seem to be reasonable in some cases also. The selection of interaction terms does not seem reasonable also. Overall, the team may consider revising how the variables are treated or considered, and selected in the models. After the year effect and interaction terms are better considered, the effect of the policy considerations may make sense. The model performance may be evaluated through both model fitting and model prediction. The scale of the data is in the state level, so some correlations between harvest and discard may be hidden but it may be worthwhile to investigate the performance of modeling harvest and discard together through multivariate regressions.

The RFDM model may compare the pattern of the effect of year and the fishable biomass. Usually, the year effect is to function the change of the population size. If the year-specific population size can replace the year effect then the model can be used to predict future year's recreational harvest given population size and regulation variables.

The REDM model is in the scale of year and regions. Sensitivity runs may be done to evaluate the influence of the combination of survey regions. The results provided to compare the model projected versus the MRIP observed is at the state level, and there are a couple of states with much higher differences. Exploration of the reasons that cause such large differences is important for the application of this model and management purposes. The REDM model also simulates the number of choice occasions so that the derived # of trips matches the MRIP estimated # of trips. The description of the process is clear but it is unclear how the use of the # of the simulated choice occasion will influence the year to be projected. Such uncertainty may be evaluated through more than one calibration since the population model can be from the stock assessment results.

Potential alternative approaches may be considered by the REDM: 1) simulate the work for multiple years with MRIP estimates in the past to evaluate the uncertainty of the # of choice occasions; 2) find an alternative external approach to predict the potential # of trips. This may be combined with the RFDM model idea by investigating the # of trips from each state given the alternative regulations, fishable population size, wave and co-occur economically valuable species, etc. The uncertainty of the # of calibrated choice occasions may be evaluated through 2 ways: 1) calibrate based on the bioeconomic model in the past 5-7 year data to see whether the resulted calibrated choice occasions are similar; 2) use the same # of the calibrated choice occasions from 2019 to generate the # of trips in the past years given their population and regulations through the integrated bioeconomic model, and compare them with the MRIP observed # of trips.

The REDM may also scale down the temporal scale of the angler choice experiment from year to season or waves to expand its potential application in fisheries monitoring (see Julia's presentation and related document). The RFDM model clearly demonstrated wave effect in both harvest and discard, and likely # of trips although no studies or presented in the provided document.

3) *Is the model appropriate for estimating and predicting the impacts of bag, size, and season limits on recreational catch or harvest? Are the methods in the Recreational Economic Demand Model appropriate for estimating changes in recreational effort or fishing demand?*

Both models have the potential for estimating and predicting the impact of the bag, size and season limits on recreational harvest or catch with further revision or verification (see TORs 1 and 2).

The REDM model is appropriate for estimating changes in the recreational effort of fishing demand after the uncertainty on how the use of the # of the simulated choice occasion will influence the recreational effort for the year to be projected, how the use of past one or two years' selectivity in population project and estimated recreational effort.

- a. *Does the modeling approach represent an improvement over current methods used to estimate impacts of management measures?*

Yes, both methods have improvements over the current methods used to estimate the impacts of management measures. They both provide methods to evaluate changes of single or multiple factors simultaneously either based on statistical relationships (RFDM) or based on a simulated process model (REDM). The advantages for dealing with more than one regulation changes should be better than the current method.

- b. *What are the strengths/limitations of the modeling approach for informing management measures, especially at the regional, state, wave, or mode level? Are there specific recreational fishing measures for which use of the model would not be recommended?*

The RFDM model has the function to inform management measures at the regional, state, wave level based on the past MRIP records. However, the current model selection and model construction have problems and need to be revised before being used for informing management measures.

The REDM model has the function to inform management measures at the regional and year and wave level based on an angler choice survey and a forward projecting stochastic catch-at-age model with parameters from the catch-at-age stock assessment. The angler choice model did not consider wave differences in angler preference and combined angler behaviors in 4 regions, so these scales reflected in their ability to be used in the fisheries management measures.

- c. *What are the implications of using the model to predict future catch/harvest based on historical data? Are there limits on the magnitude of change in catch/harvest or stock status beyond which use of the model would not be recommended?*

The year effect needs to be further considered in the RFDM model and a linear relationship is not acceptable which likely influences the effect of the other regulation/policy variables. The year effect may be compared with the change of the effect of catchable biomass to see whether the size of the stock or catchable size of the stock can replace the year effect. If the stock size can replace the year effect then the use of the model to predict future catch/harvest based on historical data is possible. If the year effect can't be replaced, some assumptions well adjusted may be used when doing future predictions. The model performance may be compared based on both model fitting and prediction.

The REDM model is designed for estimating changes in the recreational effort of fishing demand. Extra uncertainty evaluation on the use of the # of the calibrated choice occasion may be explored and addressed when projecting the recreational effort for the year to be projected. Questions on how the use of past one or two years' selectivity in population project and estimated recreational effort may vary worth to be explored to better use this model.

The range of the population size for the historical data may be clarified. If the future year stock size is out of the range of the historical stock size, the models may be used with caution.

- d. Can the modeling approach support development of multi-year bag, size, and season limits? If so, what criteria should be applied or developed to assess the reliability of the multi-year projections?*

Both models have the potential. The models should provide fitting error and prediction error or uncertainty based on historical multi-year population size, recreational catch, regulations, etc.

4) *Provide guidance for the following future model use considerations:*

- a. Could the model be modified to incorporate other species (e.g., scup, bluefish)?*

Both models can be modified to incorporate other species. A new angler preference survey may be needed to incorporate new species and to provide up to date angler preference.

- b. Could future model runs be conducted by other individuals (e.g., Council/ASMFC staff or Monitoring/Technical Committee members) without major modifications?*

Both models can be conducted by other individuals without major modifications. The REDM is a simulation model once the estimation of angler choice preference is done, so should be handled reasonably.

- c. How easily could the model be updated with additional years of data or additional variables?*

Both models should be easily updated with additional years of data. The RFDM model should be easily updated with additional variables also; the REDM model may need some moderate level of modification if additional variables are included which requires revising both the angler choice preference analysis and the simulation of the projected recreational effort and harvest.

Mid-Atlantic Fishery Management Council Sub-Group of the Scientific and Statistical Committee

Peer Review Report of Recreational Fishery Models

October 11, 2021

Individual Peer Review Report:
Dr. Cynthia M. Jones, Old Dominion University

In addition to the comprehensive consensus report developed by the peer review panel, each member developed an individual report with detailed responses to each Term of Reference (*italics*) provided by the Mid-Atlantic Fishery Management Council and the responses (standard font) are provided below.

Response to the Terms of Reference

- 1) *Are the theoretical and statistical model specifications consistent with professional standards?*
 - a. *Was the model's design and specification clearly described?*

Both models had good documentation and relied on peer-reviewed papers as the basis of their construction. Like any model construction, it is difficult to write a fully complete description and this is what lead to some of our questions.

The Recreational Fleet Dynamics Model (RFDM) spelled out the equations they used in clear fashion. This model is still under development and will benefit from further revision. In a general additive model framework (GAM), it is comprised of a set of independent polynomial regression equations for harvest and for discards of summer flounder and black sea bass. It provides simulated estimates of harvest and discard under proposed regulatory changes. The model equations include year, state, wave, recruitment, bag and size limits, spawning stock biomass (SSB) and interaction terms. The depiction of the model would be better served with an exposition of the assumptions that underlie the use of Marine Recreational Information Program (MRIP) data that form the foundation of the model. MRIP has two components, the Access Point Angler Intercept Survey (APAIS) conducted on site to estimate catch- and discards-per-angler trip and the Fishing Effort Survey (FES), a mail survey to estimate the number of angler trips. Combined, they produce estimates of total harvest and discards.

The Recreational Economic Demand Model (REDM) is more complete and is built on two peer-reviewed papers presented to the panel. It relies on an economic choice model undertaken in 2010 and statistical catch-age models developed for stocks in 2019. Using the preferences that were evaluated from the 2010 survey, it simulates angler preferences in 2019. The model simulates angler choice, based on 2010 survey responses, to alternative bag and size limits under

regulatory changes for 2019 MRIP data. The Economic choice model of 2010 relied on data reported through surveying anglers on site as part of the APAIS. The assumption of the REDM is that this survey provided a random draw from the population of anglers and it did not. It over-represented the most avid anglers. (I will discuss this more below). Depending on the use of the model, this should be made clear in the model exposition and the implications should be stated for how it is best used.

Both models rely in part on MRIP estimates of harvest and discard data and take these data to be true representations of catch and harvest. Although these MRIP data are the best available science, they also have limitations that impact models and model formulation as I discuss below.

b. Are the underlying data sufficient to derive model estimates?

The MRIP data that are used for both models are the most complete time-series of data coastwide that are available not only for summer flounder and black seas bass, but for other predominantly recreational species. There are additional surveys such as the American Littoral Society's tagging study that captures data on harvests and discards of summer flounder and black sea bass that might also be used in conjunction with the MRIP data in future model development. The MRIP surveys were designed to have the lowest variance for species of interest when aggregated at the largest scale of region and year. The FES is conducted at the region, state and two-month wave levels. The APAIS is also conducted at region, state and wave levels but also can be evaluated at finer scale of localities and smaller time frames. When both of these data are disaggregated to state and wave, the variance increases. Depending on the species and wave, the variance can be quite large and depends on the sample size that the particular state had available – some states augment the MRIP survey with more sampling. When management uses confidence intervals as suggested in the August 2, 2021 memo on Harvest Control Rules, one has concern that disaggregation will result in wide confidence bands that would provide less guidance for safe regulations that sustain the stock.

The RFDM also uses statistical catch-at-age stock assessments for both species. Such assessments are well vetted but also rely on MRIP data as their basis. These data often have high variance, especially when evaluated at finer scale.

Because the REDM used the 2010 choice survey to select anglers who were encountered in the on-site survey, avid anglers are over-represented. While there is no problem using the CPUE data obtained from an on-site survey for expansion estimates of catch and discard, they do pose issues when used for economic expansions where the assumption is made that anglers are randomly selected from all angling households. When used for an economic survey, a correction must be made for the avidity bias that exists in the APAIS survey. If the goal is to determine what the general population of anglers values, the current REDM model specification lacks this correction. If the goal of the model is to evaluate what the most active anglers value, then this correction may not have to be made. Nonetheless this issue in these data needs to be explicitly addressed and stated. The panel discussed weighting avidity based on demographics as one approach to correcting avidity. The correct demographic weighting should be available in the NMFS Fishing Effort Survey (FES). Because the economic choice model was conducted in 2010, the model also assumes that angler preferences have no changed over the ensuing decade, even given changes in regulations and angler demographics.

- 2) *How does the scale at which the model is operating (coast, regional, or state; wave or annual; fishing mode) affect the results?*
- a. *How does data availability, uncertainty, and variability affect model results, interpretation, and application?*

The MRIP survey is structured as a stratified and nested design. The APAIS is stratified by state. Nested within state are wave and within wave there is a probabilistic draw on day-work shift and access points that make up the sampling frame. It is a design that is the most variable at the lowest level and variance decreases at the highest level of aggregation. The FES is also a weighted probability survey done by mail, drawn on wave within state, with greater sampling in coastal areas and supplemented from the states' list frame of marine angler license holders. These surveys are complex and I've oversimplified for expository sake.

The RFDM operates on disaggregated data by year, state, and wave based on combined APAIS and FES estimate of harvest and discard. When querying how regulatory changes will affect harvest and discard this is appropriate because regulations are made at the state level.

The REDM is based on year and four regions (ME-NY, NJ, DE/MD, VA/NC) but also can provide output at the state level. At the regional level, the model performed well. I noted in the final report that the model performs less well at the state level, where there can be wide discrepancies between model predictions and actual occurrences. For the 2019 simulation, harvest or discard estimates of summer flounder for New Jersey, Connecticut, Massachusetts, and Maryland are predicted well, but the discrepancy is larger for Delaware, New York, North Carolina, Rhode Island and Virginia. Lack of fit was also seen for black sea bass for some states. These discrepancies need to be resolved because regulatory action is taken at the state level.

- b. *What key assumptions affect the underlying statistical analysis and interpretation of the results? Were these assumptions and relevant uncertainties identified and characterized?*

The importance of stating the assumptions cannot be overemphasized as noted in the previous TOR. While there are assumptions in constructing the model processes that need to be more clearly addressed, the elephant in the room are all the assumptions that emanate from the use of MRIP calibrated data. These data demand a clear understanding of their implicit biases- avidity, self reporting, non-response- that will influence the outcome of the best designed model. These are the only data available across time and region that have been based on statistically valid sampling protocols. However, no large-scale sampling can be done without the full understanding of population and sampling frames.

The 2010 choice survey upon which the REDM is dependent must evaluate the anticipated effect of avidity bias on the model's ability to represent the entire population of anglers, or state clearly that its results favor avid anglers predominantly. This model also had a substantial proportion of non-response that is typically of economic surveys done through add-on mail surveys. Were the 2010 respondents representative of all anglers or were they different than the larger population?

One way to query this might be to evaluate the demographics of the respondents to the nonrespondents based on APAIS data from 2010. The REDM also assumes that the angler preferences have remained the same over a decade. Were the model used for other species, preferences may change.

Likewise, the RFDM relies on the validity of self-reported discards. Typically, these data show digit-bias when discards aren't valued or when there are many of them (replying to the survey agent that there were 5 or 10 discards rather than 6 and 9) and anglers may over- or under-report discards depending on the current regulations. This may be minor or not and is difficult to assess. It is most important in fisheries where there are many discards such as for bluefish.

3) *Is the model appropriate for estimating and predicting the impacts of bag, size, and season limits on recreational catch or harvest? Are the methods in the Recreational Economic Demand Model appropriate for estimating changes in recreational effort or fishing demand?*

a. *Does the modeling approach represent an improvement over current methods used to estimate impacts of management measures?*

Both models have value for management, upon revision and if their limitations are accounted for in management decisions. I would anticipate that they will have real value when they are used together. This would be a major improvement over the ad hoc approaches that are used now. The models would predict the impact of multiple regulations on harvest and discards, and angler welfare.

Currently, the RFMD model uses smoothing to render categorical data (year and wave) for inclusion as continuous variables, which they are not. The smoothing adds data points to the dependent variables and this increases the degrees of freedom that the model uses. Although the model is penalized for overfitting, a modeling approach that uses these dependent variables appropriately as categorical variates may result in increased variance and decreased degrees of freedom. As such the estimates and predictions could be optimistic and greater certainty assumed.

Currently as configured the REDM is predicting the valuation and response to management regulations of the avid angling community, not the general population of anglers. This is a problem that can be remedied and will improve the value of the model. Because it is based on a choice model conducted in 2010, it would be valuable for the survey to be repeated, if feasible. Although choices may still be the same a decade later, this assumption should be tested. Moreover, when applied to different species, the 2010 survey may not reflected the choices of those anglers, especially if there are differences in fishing effort by season because of species availability.

b. *What are the strengths/limitations of the modeling approach for informing management measures, especially at the regional, state, wave, or mode level? Are there specific recreational fishing measures for which use of the model would not be recommended?*

Both models have much to recommend them.

The RFDM model uses long time series of MRIP data and should provide greater certainty for large regional and yearly predictions but less certainty when used to predict state and wave predictions. States that add more sampling events to MRIP data will usually have less uncertainty depending on species spatial and temporal distributions. The statistical catch-at-age models use the aggregate data to provide predictions and so the uncertainty at state and wave level may be underestimated at the level where regulations will be promulgated. I noted that during the presentation that the model results presented at our meeting include only the private boat mode. Depending on the species, shore and for hire may also be important sources of data, especially for species such as bluefish.

The REDM performed well at the region level. It has the promise of simulating bag and size limit regulations of projected harvest, discard, and angler satisfaction. The model is subject to biases present in the data acquisition and so it will provide better insights where the sampling is adjusted for avidity or where avidity is less of an issue.

- c. What are the implications of using the model to predict future catch/harvest based on historical data? Are there limits on the magnitude of change in catch/harvest or stock status beyond which use of the model would not be recommended?*

The RFDM uses a long time series of revised MRIP estimates of landings and discards. Effort data for black sea bass and summer flounder used to develop the estimates converge well with previous MRFSS telephone survey estimates of effort. However, this is not true of all species for which this model may be applied in the future. For example, bluefish data do not converge through time in the calibrations and the impact of this on predictions is hard to ascertain.

The REDM relies on the stated preferences of anglers in 2010. I would anticipate that it simulates the regional preferences well when close to that period. However, if angler preferences have changed in the ensuing decade because of regulatory changes, species availability or abundance, the model may not predict well the impact of future regulatory measures. One indicator might be available in the MRIP demography data which might show whether the demographics of the marine angling community have changed. If climate change has altered fish distributions or angler behavior, then the choice preferences from 2010 may not as accurately reflect current angler choices. As offshore waters warm, we may also see a redistribution of effort to cooler waves which is currently not a focus of the model predictions.

- d. Can the modeling approach support development of multi-year bag, size, and season limits? If so, what criteria should be applied or developed to assess the reliability of the multi-year projections?*

Both models have the potential. The models should provide fitting error and prediction error or uncertainty based on historical multi-year population size, recreational catch, regulations, etc. to be more useful. The limitation and strengths discussed above will influence the accuracy of multi-year predictions.

4) *Provide guidance for the following future model use considerations:*

a. *Could the model be modified to incorporate other species (e.g., scup, bluefish)?*

Although these models may be useful for scup, there are considerable concerns with their application to bluefish. The discards for bluefish have been steadily increasing and now equal the harvest. Moreover, the discard numbers and sizes are all self-reported data which can be much less certain. It is less clear if avidity bias is as much of a concern for summer flounder, but this has not yet been ascertained.

b. *Could future model runs be conducted by other individuals (e.g., Council/ASMFC staff or Monitoring/Technical Committee members) without major modifications?*

Both models when revised and fully documented should be able to be run by council and ASFMC staff, many of whom are well qualified to do so.

c. *How easily could the model be updated with additional years of data or additional variables?*

Both models, upon revision, should be easily updated. The concern for the REDM is that the choice survey may be outdated and should probably be done each decade or so. This will add additional expense to updating the model. Add-on surveys to APAIS require good survey sampling practice (initial mailing, reminder mailings, a check on non-respondents).



Atlantic States Marine Fisheries Commission

1050 N. Highland Street • Suite 200A-N • Arlington, VA 22201
703.842.0740 • 703.842.0741 (fax) • www.asmf.org

MEMORANDUM

TO: ISFMP Policy Board
FROM: Executive Committee
DATE: October 9, 2021
SUBJECT: Tasks to address Concerns with Conservation Equivalency

The Executive Committee raised questions and concerns regarding the use of conservation equivalency in Commission FMPs. The Committee tasked a subgroup to create a list of tasks for the Management and Science Committee to address general concerns that have been raised either through the Executive Committee or species management boards, e.g. Atlantic striped bass. The subgroup develop the following list of tasks for the MSC to address.

1. Develop a way to better characterize and address uncertainty of conservation equivalency (CE) proposals, for example: Develop a buffer to account for uncertainty
 - a. Should stock status be accounted for when establishing buffers (stock status steps/tiers, control rule)
 - b. Don't want a buffer to be overly burdensome on "fringe states" - should the buffer apply differently to the fringe states?
2. Develop a retrospective analysis to see how well CE performed including the coastwide measure for comparison
 - a. This could help inform the above buffer
 - b. Consider harvest vs total removals consistent with FMP
3. For species and measures that are harder to evaluate equivalency, should CE be allowed (some measures are non-quantifiable)? Should there be bounds on CE or is anything allowed unless specifically excluded by the FMP? Should FMPs allow the mix of CW measures and CE measures within the same management process? E.g. last round of SB measures
4. Data Standards:
 - a. Are there minimum data standards for CE or a required level of review of the data sets used if not within the bounds of the minimum data standards?
 - b. Should things that cannot be quantified be permitted under CE?
5. Should there be a time limit on CE programs (set number of years, assessment cycle, etc.)?
6. Should stock status impact CE? If so how? Example, if a stock is declared overfished/overfishing is occurring then should CE be re-evaluated?

M21-118