CENTER FOR INDEPENDENT EXPERTS (CIE) EXTERNAL PEER REVIEW OF SEDAR 104 Dolphinfish

July 28th to 31st 2026, Charleston, SC

Individual Peer Review Report prepared for the Center for Independent Experts by

Dr. Daniel Howell, Institute of Marine Research, Norway

daniel.howell@hi.no

Table of Contents

Executive summary 3
Background 4
Review Activities 4
Introduction 5
Specific comments on each ToR 5
Appendix 1: Bibliography of materials provided for review 17
Appendix 2: Copy of the Performance Work Statement 18
Appendix 3: Panel membership 28

Executive summary

The Atlantic dolphinfish stock is not currently one which is managed using a stock assessment model, and there is no viable prospect of an accepted assessment model in the foreseeable future. Data is limited and the stock has high levels of uncertainty around key life-history parameters, stock levels, and migration (on both large and local scales). In many ways this stock, with high M, high migration, rapid growth and maturity, high recreational fisheries, and overall poor data represents as hard a challenge for constructing an assessment model as could be found. As a result, stock status information or reference points are not available. The Management Strategy Evaluation (MSE) process in this case, therefore, revolved around finding ways of producing scientific advice which is robust to the uncertainties in order to underpin management in the absence of traditional stock status reference points, rather than providing the stock status reference points.

The existing management is largely static and non-responsive to changes in stock levels or distribution. Although the management includes trip and bag limits as well as total allowable catch, these are limits which are rarely met in practice. Therefore, the current fishery is essentially open access, with a minimum size in some areas as the only management measure having a currently meaningful impact in most years.

The question to be answered in this review is therefore: is the MSE system developed of sufficient quantity to be able to support better management than the current system, and the judgement of this review is that the answer to that is clearly "yes, this MSE system represents a step forward and forms the best currently available scientific basis for informing current management". However, what this review exercise did not do is to assess which specific Management Plans (MPs) would be recommended for management. Evaluations of some candidate MPs were presented, and highlighted that refining the trip limits would be unlikely to significantly impact management. A wider range of MPs, as well as refined diagnostics are needed to identify which (if any) of the potential MPs meet the management and stakeholder goals.

The work reviewed here is obviously a first round, and future work should concentrate on improving the MSE framework for future revisions of the management system. Experience gained while using the outcomes of the MSE should be incorporated into a future revision of the MSE system to improve the Operating Model used here. There are likely other data available which were not suitable for use in direct model tuning, often at a local scale (e.g., local CPUE or local data on catches of large/old fish). These should be used as a "sanity check" of the model results after the review is complete (comparing local outcomes of the modelling tool to the data). This analysis should feed into future development of the MSE and can be used to relate overall model results to highly local fishing data in communication with stakeholders.

It cannot be stressed enough that the scenarios represented by the Operating Models used in this MSE analysis have been chosen to bracket the ranges of plausible uncertainty in order to test the robustness of the MPs. This is not a stock assessment, and none of the Operating Models are designed to represent a "most plausible projection of future dynamics" and absolutely should not be used for stock status or projecting stock and catches in this manner.

It should be highlighted that, at present, the majority of catches are occurring outside US waters. Management would obviously be improved if an international agreement on data collection and sharing and collective management could be reached.

Background

The SEDAR 104 Dolphinfish MSE was convened on July 28th to 31st 2026 in Charleston, SC. The review was to evaluate the MSE framework for testing Management Procedures (MPs) for dolphinfish (mahi mahi), and to evaluate the proposed candidate MPs.

The review evaluated the MSE framework as appropriate for evaluating MPs for use in managing the US dolphinfish fishery and was appropriate given the difficulties and poor data basis for modelling this stock. The review also suggested areas where the future development of the MSE framework could focus. Some MPs were presented and evaluated, but a wider range of MPs (and some improved diagnostics) would be needed to recommend MPs to managers and stakeholders.

Review Activities

The focus of the review was to review the design of the MSE framework, the range of values chosen to parameterize the system, the choice of candidate Operating Models (OM) and robustness tests. In addition, the review evaluated the candidate MPs evaluated by the MSE system. The review explicitly did not consider the specific annual advice. The review focused on the process of deriving advice from available indices, rather than the advice produced. At no point did the review attempt to find a "most plausible" OM which could be used for projection or stock status determination.

Prior to the in-person meeting, the Panel was provided with written materials to review describing the single species model and the multispecies and ecosystem models considered during the review. There was a preliminary web meeting to present an overview of the work and to allow for reviewer requests which could be addressed prior to the physical meeting commencing.

This reviewer is experienced in running stock assessments and MSEs, as well as incorporating ecosystem reality in management advice. He also has experience of multispecies modelling, and of work to incorporate more ecosystem realism in advice. He has coding skills in R and C++, however he has not attempted to run the MSE framework presented here.

Introduction

The rationale for this work is a desire to improve on the current management regime for Dolphinfish in US east coast waters. Management is currently based on size and bag limits, and an overall catch limit set as the 3rd highest reported catch in 1994-2007. There is relatively little scientific evidence underlying this management regime, and the catch limit has only taken effect in one year.

The stock is highly migratory and is highly productive with multiple recruitment events within a year and high growth rates and early maturity. The data on the stock is both limited and uncertain. There are high uncertainties around the key life history parameters for the stock (e.g., M, growth). A large fraction of the catches is taken outside US waters, and there are difficulties with accuracy of reported catch data, and lack of length sampling. Within US waters, the majority of the catches are taken by the recreational sector, which has limited sampling. As a result of these factors, there is no accepted analytic assessment for this stock, and no realistic possibility of one in the near future.

The suggested approach here is to construct a Management Strategy Evaluation (MSE) framework, with a suite (termed here a "reference set") of Operating Models (OMs) to account for the plausible uncertainty around the stock, and then a range of more extreme Robustness Tests. This framework was then used to test a range of Management Procedures (MPs), each based on an empirical time series.

The OMs are conditioned to match a spatial index generated from the Pelagic Longline fleet, using the VAST program. This fleet covers close to the whole distribution of the population and mostly catches dolphinfish as a bycatch (which somewhat reduces the difficulties involved in using commercial data as a survey index). This is the only tuning data used apart from catches and length compositions, and therefore the only data governing overall trends and spatial distribution. This is a heavy burden to place on a rather uncertain index, but in the absence of other data this is the best available.

The model is able to match the time-varying distribution of the stock according to the VAST index. However, the model is not able to produce realistic movements of the fish, because many different versions of migration could produce the same state data of the distributions. Provided no area is depleted by fishing within one or two seasons, then the ability to model the distribution is likely sufficient. Much more data (for example, extensive tagging data) would be needed to model the actual migrations.

Specific comments on each ToR

The comments below focus on the specific ToRs. It should be stressed that the reviewer endorses the MSE process as the basis for management advice. The following comments

focus on areas where there is room for improvement for future development. This focus should not be taken as a negative verdict.

1. **Evaluate the types and ranges of uncertainty used in the MSE. Explicitly consider uncertainty in data, life history, process error, nonstationarity (i.e., environmental variability and oceanographic changes), and if appropriate, fishery dynamics.**

All of the life history processes are uncertain, so the best that can be done here is to choose plausible values or ranges, and to justify those choices. There are likely multiple different specific values for the ranges which could be chosen, and the focus is on overall robustness of the MSEs to the uncertainty rather than identifying specific values for any of these parameters. In particular, M is difficult to estimate, and using a range as here is appropriate. The chosen range is reasonably justified, although further robustness tests could be conducted. For example, it is possible that M and K could be correlated (currently assumed independent) and the implications of this could be tested through a robustness test.

The recruitment relationship and "resilience" (ability to recover from low stock sizes) has no data to inform it, and therefore this cannot be parameterised. Choosing to use the range of resilience from other dolphinfish, avoiding exactly 1 (0.7-0.95), with robustness tests, is appropriate. Exactly 1 (i.e., full recruitment down to a single spawning fish) is physically unrealistic and causes difficulties with the numerical simulations, so setting an upper bound (as here) is necessary.

Modelling the growth is difficult, with different estimates available. The choice here is likely the best that can be achieved with the current data, but there is strong recommendation to do more age reading (samples per length category) in order to be able to get better data for a growth curve (and hence refine the M range).

The seasonal distribution of the fish matters a lot here, both in terms of the data and the performance of the local fisheries. The VAST index, despite the problems of using a CPUE index, gives a moderately good basis for the scope and variability of the distribution. However, there is not good information on the dynamics of that migration (as opposed to the spatial distributions), and in the absence of large-scale tagging there is not likely to be any. Therefore, heavy fishing in any one area runs an unquantifiable risk of inducing local depletion, which is not well accounted for in the MSE simulations.

a. **Were selected axes of uncertainty relevant to the stock and fishery and appropriately characterized? If relevant, were uncertainties highlighted by managers and stakeholders included?**

The majority of axes of uncertainty were appropriate to the uncertainties in the relevant parameters in most cases. The parameters identified were key sources of uncertainty (M, resilience, catches, future recruitment, migration). The ranges for M, resilience, and future recruitment were reasonably well justified. The uncertainty around catches should be extended from "unreported and discards" to also include the uncertainties in the international catches. These are not well constrained, and the catches constitute the largest part of the fishery. Therefore, this is the data which has the most potential to have an impact on the results.

One source of uncertainty which was not adequately addressed is the migration. The model is able to match the spatial distributions by season from the VAST index, but lacks data to realistically identify the actual migration routes. There is no tuning data to inform these migration routes, and therefore little that could be realistically investigated. One suggestion for an improvement here is to start with a weak prior with a default viscosity set to 1, i.e., no movement at all. Then the model will estimate the *minimum* amount of migration necessary to match the distribution data, and this parsimony should reduce the possibility of spurious artifacts. The current migration model is overly detailed given the limited data for tuning. In the future, it would be more useful to simplify the migration model and instead focus on the uncertainties in the VAST seasonal distribution estimates. This could be done by creating multiple plausible VAST models (with different assumptions) or a range of uncertainty around the VAST results, and then investigating how much impact this uncertainty has on the performance of the MPs.

b. Did levels of uncertainty for each axis meaningfully and reasonably encompass plausible uncertainty?

The levels of uncertainty were well justified, and (with the caveats of the previous comments around the international catches and the distribution uncertainties), the reviewer endorses the choices made here. Given the uncertainties, it is likely that there are alternative ranges which could have been chosen, but the selections here seem as reasonable as any.

c. Were uncertainties appropriately classified into reference (primary) or robustness (secondary) set and sufficiently explained?

Except as noted above, the classification into reference set was appropriate, and the reference set was well defined. The robustness set was derived by using expert judgement to define scenarios to meet specific stakeholder concerns, and then presented to stakeholders for approval. Then, the robustness set was extended at the review meeting. This iterative process, building on stakeholder and expert knowledge, is an appropriate way to produce a robustness set – although this happened prior to the review, so this document cannot comment directly on how the process was conducted.

The distinction between the reference set and the robustness set was clear in the meeting, although it will be important in communication that the distinction is made clear, with the robustness set representing a stress test of the MPs (rather than necessarily a set of likely outcomes). It is also important to communicate that these sets are designed to bracket the range of possibilities and do not represent an equivalent to a stock assessment.

2. Evaluate the strengths and weaknesses of methods used to simulate the stock.

The key weakness is obviously the limited knowledge around the stock status and dynamics, especially including the following: stock size, mortality, migration pattern and variability, level below which recruitment is impaired, and international catch levels. This implies difficulties in estimating the "best" value for each of these factors, but also makes it difficult to estimate the uncertainty around these estimates. All these comments here are therefore based around finding an appropriate way of simulating these unknown life history parameters and uncertainties.

The method employed here, of characterizing "plausible" uncertainty ranges and testing MPs against that range is appropriate, and represents the best (and likely only) viable approach here, and represents a real step forward in managing this fishery. The ranges of uncertainty were often difficult to pin down (given the lack of data), but the ranges chosen here were justified as well as could be done given the data.

One issue is that the uncertainty estimates from the VAST index seem unrealistically low for at least part of the index. The lowest uncertainty values are well below those which would be expected from a dedicated high-quality survey, and much too low for a CPUE index based on bycatch in a widely distributed fishery. This may have resulted in the OMs being tuned incorrectly (as the low uncertainty levels imply that the model should fit the data closely). It is important to revise this, with a plausible lower bound set on the uncertainty when tuning the OMs.

a. Are methods to generate operating models scientifically sound?

As repeatedly mentioned, there is a very high degree of uncertainty here. The methods used to address this are appropriate, choosing a range of plausible values for different key parameters and testing the MPs against that uncertainty. The work here has attempted to realistically characterize these uncertainties and to focus work on those which appear to have the highest impact on the results.

One advantage from the life history of the dolphinfish is that the generation time is so short that the initial population and catch assumptions at the start of the OM hindcast are not relevant, as all the initial fish will die off very rapidly. This removes one key uncertainty which affects MSEs of longer-lived species.

Of the remaining uncertainties, there are several which seem more feasible to address than others. Re-focussing the migration modelling to focus more on spatial distributions (which have data to support them) than actual migration (which does not) would enable a better focusing of efforts. Any improvements on the data on the international fishery would have an important impact on the accuracy of the simulations. Better age data would help refine both growth and mortality estimates.

b. Did methods meet the needs of stakeholders and/or research questions (i.e., appropriate spatial resolution, time-step, ability to measure key performance statistics)?

This review is only partially able to answer this ToR, as there is a lack of knowledge of the "stakeholder needs". The methods here seem largely appropriate for the specific issues mentioned in this ToR. The time step is seasonal, which is likely as fine as could be achieved given the data. This is probably adequate for most management needs, though may be too coarse to match all stakeholder needs. The spatial resolution is also likely as fine as could be achieved, although it is likely too coarse for stakeholder goals for local fisheries. Given the speed of the movement of fish and the limited data, finer scale modelling would be problematic.

c. If data were unavailable to parameterize the model, were the assumptions used to overcome this limitation reasonable and/or multiple assumptions included as an axis of uncertainty?

There was always some "data", even if of low quality or by borrowing strength from similar species, for the Operating Model settings. The methods used here to generate the plausible ranges of uncertainty are appropriate.

One kind of scenario testing conducted here was to create scenarios which were more extreme than could be expected in reality (for example on strong trends in movement). If such an overly strong scenario shows minimal impact on the outcomes of the management plan, then more realistic shifts are unlikely to have impact. Where the "extreme scenario" does show significant impacts, then this highlights where effort is required on the more realistic scenarios. This approach is endorsed by the reviewer.

d. Was plausibility weighting of reference operating models appropriate and sufficiently justified?

In principle, one might wish to give greater weight to more realistic OMs – although in practice it is difficult to find ways of doing this in an objective manner. However, this is not an issue for this stock. The overall knowledge of the dolphinfish life history parameters, history, and fishery is clearly limited and uncertain. Therefore, identifying (and quantifying) the likelihood of different scenarios is not realistic, and the chosen approach of giving all scenarios equal weight is appropriate.

e. **Does the robustness of the operating model grid allow the management procedure to fail?**

In principle, yes, the framework would allow for failures of MPs. However, the MPs evaluated here all had minimal impacts on the performance of the fishery, and therefore these did not "fail". It would, therefore, be a good idea to run extreme scenarios to identify what could cause the MPs to fail, both to test the system but also to identify what warning signals could be identified. In the same spirit, it would be useful to identify specific goals (more large fish, higher catch rates, etc.) and then identify what MPs (if any) would achieve these goals (even if those MPs were too extreme to be implementable). This can then inform if it is possible for management to achieve these goals at all, and if it could be possible with viably implementable MPs.

3. Evaluate operating model conditioning.

a. **Did conditioning the operating model on available data result in simulated data with reasonable statistical properties relative to what would be available to the management procedure?**

The results presented in the review did mostly indicate that the OMs resulted in "plausible" dynamics. There were some issues. One is the OMs were not able to track extreme values in the data. However, given only having a single tuning data series, it is not easy to identify if this was a problem with the OM, or simply due to the noise in the data (which should not be followed). In addition, the forecast equivalent to the VAST index was not presented in the review, so it is not possible to comment on the degree to which the simulations realistically replicated the index properties. In addition, we saw evaluations of seasonal recruitments, whereas the annual total recruitments might be a) easier to interpret, and b) more consequential to stock dynamics.

4. Evaluate MSE projections.

a. **Did the operating models project data that would be available to the management procedure?**

Yes, given that there was only a single tuning index available. Note that there are issues around timing for management, with the delays between collecting data and implementing management actions being rather long compared to the life span of the fish – with management actions on a particular generation being likely based on data from the previous generation. The MSE tool could be used to test the performance of data available on a tighter timescale.

b. **Were projection assumptions reasonable and appropriate for the stock (e.g., decisions about future parameter estimates and uncertainty; resampling approaches to generate future data; future nonstationarity, productivity, etc.)?**

The projection assumptions seemed largely reasonable, and the framework was sufficiently flexible to incorporate specific concerns as they are identified. At present there are no explicit environmental drivers (there are proxies, but not explicit drivers). These could potentially be included in further rounds of development, assuming the data is available to parameterise these.

5. Evaluate performance metrics.

As mentioned above, this is a stock with very high levels of uncertainty, which includes high levels of uncertainty in the estimate of uncertainty and hence risk. This is not a criticism of the approach here, rather the MSE is designed to provide the best available advice in the face of this uncertainty. However, it means that the metrics need to take account of this.

a. **Do performance metrics include all legally mandated management objectives as outlined in the Magnuson-Stevens Fishery Conservation and Management Act (i.e., to include status determination criteria with suitable risk tolerance to measure probability of overfishing and overfished status; and rebuild overfished stock in accordance with rebuilding requirements)?**

This reviewer lacks the knowledge to identify if "all" legally mandated management objectives have been met. The specific points mentioned in this ToR (status determination with uncertainty, overfishing and overfished status, potential rebuilding plans) are covered in the performance metrics.

b. Did performance metrics sufficiently capture additional manager and/or stakeholder objectives to achieve optimum yield, if relevant (i.e., to include desired management objectives such as: total catch, catch rate, catch size, fishing opportunity/season length, etc.)?

Again, this reviewer cannot comment on objectives beyond the specific quantities mentioned in this review. Many of these desired management outcomes are largely driven by stock dynamics, rather than management actions. In general, management actions can give a worse performance than nature allows, but cannot improve on this.

Given the dynamics of the stock, there is no scope for management to even out catches per year (except by imposing an artificially low catch limit). Total yield could be impacted by recruitment or growth overfishing, although there is no evidence that this is already

occurring (given the fecundity and rapid growth of the fish), and therefore no realistic scope for management to improve on this (except potentially in the catches of the largest fish).

The metrics presented covered most of the stated goals; however, there were several presented goals (opportunity to fish, abundance of large fish) which were not well covered by the existing metrics. New metrics specifically addressing these would be needed in order to evaluate MP performance in these respects.

6. Evaluate candidate management procedures.

The point of a MP is to give advice on how the catch levels should be adjusted to account for variations in stock size. This review focused on the MSE tool for examining MPs and examined an initial set of candidate MPs. We did not evaluate a full range of MPs and therefore are not recommending a specific set of MPs as the basis of management.

a. Did the candidate management procedures only utilize data available for the stock in practice?

The different Management Plans here all use data which is available for the stock and is likely to continue to be available as a basis for management.

The review examined alternative potential indices, but none of these were considered as currently viable as a basis for management. Environmental drivers (for example region of preferred temperature being concentrated in or out of US waters or the El Nino and Southern Oscillation "ENSO") are potentially viable predictors of stock size or fish catchability. However, these are not included in the dynamics of the OM population, and the response between stock and index cannot be simulated using the system as it is developed here. The recreational data, MRIP data, is not a viable management index: there is little signal in the index and no effort standardization over time, and therefore it does not represent a usable CPUE. The Puerto Rico index is a potential candidate, provided future data can be guaranteed, as it represents a more tightly controlled "sampling" than the other potential alternate data series, and is potentially "upstream" on the migration routes entering the US waters. However, more scientific analysis (e.g., looking at technological creep) and confirmation that the index will continue to be available, would be needed before considering basing advice on this index.

However, even though they are not suitable for direct use in model tuning, some of these additional data sets could be used as external "sanity checks" on the model outputs.

b. Did the management procedure explore management tools and controls that are applicable to the stock (e.g., appropriate management procedure archetype)?

Only a limited number of MPs were evaluated. The set of MPs focussed heavily on the impact of different trip limits on the number of retained individuals, and the results indicate that this is not an effective management procedure. Except at very low levels, the trip limit had no or minimal impact on the fishery (although it may be effective in avoiding potential creep from recreational to effectively commercial fishing, so these results should not be taken as recommending the removal of the existing limits).

Therefore, a wider range of options needs to be examined, in order to bracket what is a) possible to achieve with extreme management options, and b) what is possible to achieve with more realistic management options. These need to be integrated with the previous point about refining the metrics to better reflect stakeholder goals.

c. Were top-performing management procedures tuned to minimally achieve all satisficing criteria?

The MPs presented did achieve these basic criteria. However, as mentioned above, it should be noted that none of these had any real impact on the fishery, and a wider range needs to be examined.

d. If empirical management procedure, does the MSE include a mechanism to periodically update stock status?

The system here does not provide stock status. Without being able to create an analytic stock assessment, there is no realistic prospect of producing stock status. It is important to stress this, and that none of the individual OMs represent a "stock assessment proxy".

However, there is a possibility to gain at least qualitative information on stock trajectory (increasing, stable, declining), though not status, based on the VAST index. There is also a possibility to highlight over- or under-fishing by comparing actual fishing levels with the fishing levels recommended from an adopted MP.

7. Provide or comment on recommendations to improve the MSE.

Given the high uncertainty levels around this stock, the MSE conducted here should be considered a first round in an ongoing process. The work presented here forms an appropriate basis for management advice in the short and medium term. It is recommended that experience gained while applying any MP resulting from this review should be used to inform any future MSE work. There are two key priority recommendations here, and then a list of more detailed potential additions.

The key priority is to go through the VAST index. First, the uncertainties need to be revised, with a plausible lower bound set on the uncertainties and then re-tune the OMs. Work should also be done to evaluate how this is responding to environmental changes (the El Nino and Southern Oscillation or "ENSO", or the local concentration of suitable temperatures on the US coast).

The other priority is to expand the diagnostics set to meet the stakeholder goals (for example, fraction of large fish) and then run a wider range of MP scenarios to identify what the maximum results of management could be. This can then be used to inform a tighter range of "realistically implementable" MPs.

Further specific suggestions for improvements are listed below.

Model specific

- *Metrics*
  There are several suggestions for improved metrics in this review, on fishing opportunity and abundance of large fish in the catches (which does not really show up in the existing average size metric). The experience gained in using the system to inform advice should be used to refine the metrics as appropriate.

- *Stock migration and distribution*

Less focus on viscosity (since there is no data to inform this), more on distribution uncertainty.

- *Robustness checks around the "stress test direction" OMs*

The OMs have been chosen to focus on the uncertainty ranges which pose the highest risk to the stock or fishery. Occasionally, there can be unexpected outcomes, where a value in the perceived "safer" direction can lead to higher risks. For example, higher productivity would typically be expected to reduce the risk of stock collapse, but variability around a highly productive stock which supports a large fishery can have the opposite effect. Such concerns can be tested through robustness testing.

- *Testing of future dynamics*

The results presented here did not fully test the statistical performance of the forecast (compared to the hindcast data). There was no forecast "VAST index equivalent", and annual recruitment could be more useful than seasonal recruitment (since variations in seasonal recruitment can cancel out). Both of these should be easy to implement.

- *Initial conditions*

This is a recommendation not to do any further work on the initial conditions in the OMs. For many stocks the initial conditions can be important in the performance of

the modelling. However, the short lifespan and high turnover mean that this is not an issue for this stock.

Data specific

The data for this stock is very limited, and obviously any improvement here would improve the knowledge around the stock.

- *Work on the estimates of the international catch in tonnes*

In addition to the high level of uncertainty, there appears to be an issue with errors in at least some of the data. Improving the quality of this data, and especially better estimates of any recent and ongoing trend, would likely be the single largest improvement to the MSE framework.

- *Revisit the estimation of the discards across the fishery.*

The estimate of unreported catches from the international fleet has remained almost constant over the last 20 years, during which time the international catches have increased by around 30%. There is a clear mismatch here, which suggests that the estimate needs revision. In addition, there is no estimate of discards from the US recreational fleet.

- *Better age data*

More age data (perhaps a specified number of samples per age category) would give better information on the growth, and on bounding the mortality estimates. Better age data would be the single most important contribution to refining the life history parameters in the simulations.

- *Need better sampling from the recreational fisheries.*

The MSE was initially conditioned using very low sample sizes, although this was improved during the review. More investigations of the length composition would be useful to try and identify spatial structure. In addition, it could be valuable to use new sampling methodology in the future to find cost effective ways of improving the database (for example, asking fishers to record lengths at sea rather than relying on port sampling).

- *Focus any data improvements on the recreational catches, especially in the south and more length sampling in the north*

Recreational catches in the south dominate the US catches, and therefore data improvements here will produce large improvements. There are very few length observations in the north, which make identifying a length selectivity in north difficult. Improving this would aid especially in modelling what MPs could improve the catches

of large fish. The commercial fleet is much better sampled, but constitutes less than 1% of the catches, so there is less to be gained by focussing here.

- *Investigate spatial size structure*

The analysis conducted here assumes that there is no spatial size structure. If there are differences (either north-south or inshore-offshore), then this could invalidate at least part of the simulations conducted here.

- *Additional data sources*

Are there additional data sources (for example regional CPUE estimates) which could be used as either tuning series or for ground truthing the outputs of the OMs?

8. Prepare a Review Workshop Summary Report describing the Panel's evaluation of the MSE and addressing each TOR.

Such a report has been produced. This review report endorses the conclusions in that summary report.

Appendix 1: Bibliography of materials provided for review

Documents and presentations for the MSE were made available at: https://blue-matter.github.io/mahiMSE/ and https://sedarweb.org/assessments/sedar-104-dolphin-management-strategy-evaluation/

The main results and background information were presented in: https://sedarweb.org/documents/sedar-104-atlantic-dolphinfish-management-strategy-evaluation-assessment-process-report/

More detailed background information was presented in: Section1_S104_for_RW_v2

Appendix 2: Copy of the Performance Work Statement

Performance Work Statement (PWS)
National Oceanic and Atmospheric Administration (NOAA)
NOAA Fisheries
Center for Independent Experts (CIE) Program
External Independent Peer Review

SEDAR 104 Atlantic Dolphinfish Management Strategy Evaluation Review

July 28-31, 2026

Background

The NOAA Fisheries is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NOAA Fisheries science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards[1].

Scope

The Southeast Fisheries Science Center (SEFSC), in collaboration with the South Atlantic Fishery Management Council (SAFMC), the Southeast Regional Office (SERO), Blue Matter Science, and others, have developed a full-stakeholder management strategy

[1] https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2005/m05-03.pdf

evaluation (MSE) for dolphin (e.g., dolphinfish, mahi, dorado) in the US Atlantic. The objective of the MSE is to develop empirical (indicator-based) candidate management procedures (MPs) for dolphin in the US Atlantic that meet legally mandated (e.g., Magnuson-Stevens Fishery Conservation and Management Act) and regionally-varying stakeholder-defined management objectives, such that they may achieve optimum yield. Dolphin in the US Atlantic are managed by SAFMC. The MSE is intended to be used to inform dolphin management, such that top performing candidate MPs will be proposed for use as Amendment alternatives to the Dolphin Wahoo Fishery Management Plan and be voted on for implementation by the SAFMC.

Before an MP may be enacted, the MSE used to derive candidate MPs must be deemed best scientific information available. The review panel is requested to review the scientific validity of the dolphin MSE. Unlike stock assessments (designed to provide a best estimate of stock status and for which diagnostics are of increased importance), MSE operating models are discrete hypotheses designed to encompass plausible uncertainty and are conditioned to ensure projected data are simulated with realistic uncertainty and variability.

The specified format and contents of the individual peer review reports are found in **Annex 1**. The Terms of Reference (ToR) of the peer review are listed in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3.**

Requirements

NOAA Fisheries requires 3 reviewers to conduct an impartial and independent peer review in accordance with the Performance Work Statement (PWS), OMB guidelines, and the ToRs below. The reviewers shall have expertise in management strategy evaluation, quantitative population dynamics, and fisheries simulation analyses.

The chair, who is in addition to the three reviewers, will not be provided by the CIE. Although the chair will be participating in this review, the chair's participation (e.g., labor and travel) is not covered by this contract.

Each reviewer will write an individual review report in accordance with the PWS, OMB Guidelines, and the TORs below. Modifications to the PWS and TORs cannot be made during the peer review, and any PWS or TORs modifications prior to the peer review shall be approved by the Contracting Officer's Representative (COR) and the CIE contractor. All TORs must be addressed in each reviewer's report.

Tasks for Reviewers

1) Pre-review Background Documents: Review the following background materials and reports prior to the review:

Two weeks before the peer review, the NOAA Fisheries Project Contact will send by electronic mail or make available at an FTP site to the CIE reviewer all necessary

background information and reports for the peer review. The CIE reviewer shall read all documents in preparation for the peer review, to include:

- Trial specifications document
- Github repository: https://github.com/Blue-Matter/DolphinMSE (note currently set to private, will be made public or shared with reviewers prior to review.)
- Shiny app: (html to be developed)

2) Attend and participate in an in-person review meeting. The meeting will consist of presentations by NOAA and other scientists, stock assessment authors and others to facilitate the review, to answer any questions from the reviewers, and to provide any additional information required by the reviewers.

3) After the review meeting, reviewers shall conduct an independent peer review report in accordance with the requirements specified in this PWS, OMB guidelines, and ToRs, in adherence with the required formatting and content guidelines. Reviewers are not required to reach a consensus.

4) Each reviewer shall assist the Chair of the meeting with contributions to the summary report.

5) Deliver their reports to the Government according to the specified milestones dates.

Foreign National Security Clearance

When reviewers participate during a panel review meeting at a government facility, the NOAA Fisheries Project Contact is responsible for obtaining the Foreign National Security Clearance approval for reviewers who are non-US citizens. For this reason, the reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days in advance. For additional information, please see the following link: https://www.commerce.gov/osy/programs/foreign-access-management. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance

The place of performance shall be in Charleston, SC.

Period of Performance

Each reviewer's duties shall not exceed **14** days to complete all required tasks.

Schedule of Milestones and Deliverables: The contractor shall complete the tasks and deliverables in accordance with the following schedule.

Schedule	Milestones and Deliverables
Within 2 weeks of award	Contractor selects and confirms reviewers
Approximately 2 weeks prior to the review	Contractor provides the pre-review documents to the reviewers
July 28-31, 2026	Panel review meeting
Approximately 2 weeks later	Contractor receives draft reports
Within 3 weeks of receiving draft reports	Contractor submits final reports to the Government

* The Peer Review Summary Report will not be submitted to, reviewed, or approved by the Contractor.

Applicable Performance Standards

The acceptance of the contract deliverables shall be based on three performance standards:

(1) The reports shall be completed in accordance with the required formatting and content (2) The reports shall address each ToR as specified (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Confidentiality and Data Privacy

This contract may require that services contractors have access to Privacy Information. Services contractors are responsible for maintaining the confidentiality of all subjects and materials and may be required to sign and adhere to a Non-disclosure Agreement (NDA).

Travel

All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (Travel resources | GSA), and all contractor travel must be approved by the COR prior to the actual travel. Any travel conducted prior to the receipt of proper written authorization from the COR will be done at the Contractor's own risk and expense. International travel is authorized for this contract. Travel is not to exceed $14,000.

NOAA Fisheries Project Contact

Cassidy Peterson
Management Strategy Evaluation Specialist
Southeast Fisheries Science Center
101 Pivers Island Road, Beaufort, NC 28516-9722
cassidy.peterson@noaa.gov

Annex 1: Peer Review Report Requirements

The independent Peer Reviewer report shall be prefaced with an Executive Summary providing a concise summary of whether they accept or reject the work that they reviewed, with an explanation of their decision (strengths, weaknesses of the analyses, etc.).

2. The report must contain a background section, description of the individual reviewers' roles in the review activities, summary of findings for each ToR in which the weaknesses and strengths are described, and conclusions and recommendations in accordance with the TORs.

a. Reviewers must describe in their own words the review activities completed during the panel review meeting, including a brief summary of findings, of the science, conclusions, and recommendations.

b. Reviewers should discuss their independent views on each ToR even if these were consistent with those of other panelists, but especially where there were divergent views.

c. Reviewers should elaborate on any points raised in the summary report that they believe might require further clarification.

d. Reviewers shall provide a critique of the agency review process, including suggestions for improvements of both process and products.

e. The report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The report shall represent the peer review of each ToR, and shall not simply repeat the contents of the summary report.

3. The report shall include the following appendices:

Appendix 1: Bibliography of materials provided for review

Appendix 2: A copy of this Performance Work Statement

Appendix 3: Panel membership or other pertinent information from the panel review meeting.

Annex 2: Terms of Reference for the Peer Review

SEDAR 104 Atlantic Dolphinfish Management Strategy Evaluation Review

July 28-31, 2026

CIE reviewers are contracted to complete their independent peer review based on the ToRs. Therefore, the CIE-NOAA Fisheries review and approval process is based on whether the CIE independent reports addressed each ToRs.

1. Evaluate the types and ranges of uncertainty used in the MSE. Explicitly consider uncertainty in data, life history, process error, nonstationarity (i.e., environmental variability and oceanographic changes), and if appropriate, fishery dynamics.
    a. Were selected axes of uncertainty relevant to the stock and fishery and appropriately characterized? If relevant, were uncertainties highlighted by managers and stakeholders included?
    b. Did levels of uncertainty for each axis meaningfully and reasonably encompass plausible uncertainty?
    c. Were uncertainties appropriately classified into reference (primary) or robustness (secondary) set and sufficiently explained?
2. Evaluate the strengths and weaknesses of methods used to simulate the stock.
    a. Are methods to generate operating models scientifically sound?
    b. Did methods meet the needs of stakeholders and/or research questions (i.e., appropriate spatial resolution, time-step, ability to measure key performance statistics)?
    c. If data were unavailable to parameterize the model, were the assumptions used to overcome this limitation reasonable and/or multiple assumptions included as an axis of uncertainty?
    d. Was plausibility weighting of reference operating models appropriate and sufficiently justified?
    e. Does the robustness of the operating model grid allow the management procedure to fail?
3. Evaluate operating model conditioning.

a. Did conditioning the operating model on available data result in simulated data with reasonable statistical properties relative to what would be available to the management procedure?

4. Evaluate MSE projections.

    a. Did the operating models project data that would be available to the management procedure?

    b. Were projection assumptions reasonable and appropriate for the stock (e.g., decisions about future parameter estimates and uncertainty; resampling approaches to generate future data; future nonstationarity, productivity, etc.)?

5. Evaluate performance metrics.

    a. Do performance metrics include all legally mandated management objectives as outlined in the Magnuson-Stevens Fishery Conservation and Management Act (i.e., to include status determination criteria with suitable risk tolerance to measure probability of overfishing and overfished status; and rebuild overfished stock in accordance with rebuilding requirements)?

    b. Did performance metrics sufficiently capture additional manager and/or stakeholder objectives to achieve optimum yield, if relevant (i.e., to include desired management objectives such as: total catch, catch rate, catch size, fishing opportunity/season length, etc.)?

6. Evaluate candidate management procedures.

    a. Did the candidate management procedures only utilize data available for the stock in practice?

    b. Did the management procedure explore management tools and controls that are applicable to the stock (e.g., appropriate management procedure archetype)?

    c. Were top-performing management procedures tuned to minimally achieve all satisficing criteria?

    d. If empirical management procedure, does the MSE include a mechanism to periodically update stock status?

7. Provide or comment on recommendations to improve the MSE.

8. Prepare a Review Workshop Summary Report describing the Panel's evaluation of the MSE and addressing each TOR.

Annex 3: Tentative Agenda

SEDAR 104 Atlantic Dolphinfish Management Strategy Evaluation Review Workshop

July 28-31, 2026

Charleston, SC

Tuesday

8:30 am – 12:00 pm	**Introductions and Opening Remarks**	**Coordinator**
	- Agenda Review, TOR, Task Assignments	
	Project Presentations	**MSE Team**
	- *Clarifying questions*	
12:00 pm – 1:30 pm	**Lunch Break**	
1:30 pm – 5:00 pm	**Project Presentations (continued)**	**MSE Team**
	- *Clarifying questions*	
5:00 pm – 5:30 pm	**Public Comment**	**Chair**

Tuesday Goals: Part of initial presentations completed, requests for additional analysis identified.

Wednesday

8:30 am – 12:00 pm	**Project Presentations**	**MSE Team**
	- *Clarifying questions*	
12:00 pm – 1:30 pm	**Lunch Break**	
1:30 pm – 5:00 pm	**Panel Discussions**	**Chair**
	- MSE Team and Review Panel	
5:00 pm – 5:30 pm	**Public Comment**	**Chair**

***Wednesday Goals*:** Initial presentations completed. Summary report drafts begun.

Thursday

8:30 am – 12:00 pm	**Project Presentations**	**MSE Team**
	- *Clarifying questions*	
12:00 pm – 1:30 pm	**Lunch Break**	
1:30 pm – 5:00 pm	**Panel Discussions**	**Chair**
	- MSE Team and Review Panel	
5:00 pm – 5:30 pm	**Public Comment**	**Chair**

***Thursday Goals*:** Presentations completed, Summary report drafts continued.

Friday

8:30 am – 12:00 pm	**Panel Deliberations or Work Session**	**Chair**
	- *Recommendations and comments*.	
	- *Report writing*	

***Friday Goals*:** Finalize discussions, draft summary report reviewed.

Appendix 3: Panel membership

SEDAR 104 Review Panel Membership		
Name	Email	Affiliation
Daniel Howell	daniel.howell@hi.no	Institute of Marine Research, Norway; CIE
Ernesto Jardim	ernesto.jardim@gmail.com	Independent Consultant, CIE
Gary Melvin	gary.d.melvin@gmail.com	Independent Consultant, CIE
Jeffrey Buckel	jabuckel@ncsu.edu	North Carolina State University
Kelsey Roberts	kroberts6@umassd.edu	University of Massachusetts Dartmouth
Matt Damiano (Chair)	matt.damiano@deq.nc.gov	North Carolina Division of Marine Fisheries