Independent Peer Review of NMFS Study and Report:

A Management-focused Population Viability Analysis for North Atlantic Right Whales

Authors: Michael C. Runge, Daniel W. Linden, Jeffrey A. Hostetler, Diane L. Borggaard, Lance P. Garrison, Amy R. Knowlton, Véronique Lesage, Rob Williams, Richard M. Pace, III

Prepared for the Center for Independent Experts (CIE)

by

Wayne M Getz
wgetz@berkeley.edu

October 2022

Table of Contents

Executive Summary

The stated purpose of this report in response to the charge given to the authors by the National Marine Fisheries Service (NMFS) in setting up a Population Evaluation Tool (PET) subgroup/team is to describe:

the development of population viability analysis tool that can project the trajectory of the North Atlantic Right Whale (NARW) population under a variety of scenarios and report on a range of scenario analysis that will allow NMFS to characterize the NARW extinction risk, taking into account current and future threats, and will allow inquiry into how much improvement to present-day mortality and reproduction schedules is needed to improve population trajectories

The authors should be commended on their extensive analyses and detailed report that meets the above charge in terms of what I would refer to as their initial phase response. The report, in my opinion, credibly demonstrates that if the recently enacted U.S. regulations have the effect of reducing range wide entanglement risk by 25%, these regulations will likely turn around the current expected 50% decline in the NARW population over the next 100 years. They also credibly show that this kind and level of risk mitigation will substantially reduce, though not eliminate, the risk of the population dropping below 50 females (which some conservation biologist regard as the red line for classifying the population as now quasi-extinct—i.e., in need of extraordinary measures to protect and rebuild population numbers).

Although the proverbial "many ways to skin a cat" can be aptly applied to the various decisions the PET team have made in approaching their charge, in my opinion the current report falls short in four important ways. I make this assessment in the context of viewing this report as Phase 1 of a long-term study that is reported to be vetted for the implementation of new phases every five years; as well as being fully cognizant of the current limitations with data, with management options available to mitigate mortality risks, and with the many different approaches that can be taken to building appropriate models to fulfill the assigned PET team task. The report, in my opinion, falls short in failing:

1. To provide a proper mathematical description of the model and details of the implemented analyses;
2. To incorporate spatiotemporal structure in their current model (they mention inclusion of this structure in the future) because this structure is so central to NARW population management;
3. To provide the operational details of a scheme to be implemented in the next phase for updating model structure and identifying gaps in the data, as the PET team continues to develop the tools needed to address the most pressing NARW management questions at hand;
4. To provide the code needed to run their model and to discuss code verification and model validation procedures that were presumably undertaken at some point in the model development and analysis process.

Background

NMFS Mandate

The National Marine Fisheries Service (NMFS) has set up a Population Evaluation Tool (PET) subgroup (hereafter referred to as the PET team), and an allied Decision Support Tool (DST) implementation process, for management of the North Atlantic Right Whale (NARW) in response to the mandate NMFS have been given by Congress to conserve, protect, and manage our nation's marine living resources under the umbrellas of the Magnuson-Stevens Fishery Conservation and Management Act, the Endangered Species Act, and the Marine Mammal Protection act. The specific charge NMFS gave to PET in July 2018 in the context of NMFS's responsibility to manage the NARW population was

to develop a population viability analysis or other assessment tool that will allow the agency to characterize the NARW extinction risk, taking into account current and future threats, and will allow inquiry into how much improvement to present-day mortality and reproduction schedules is needed to improve population trajectories.

More specifically, the US recovery plan for the NARW, last revised in 2005, and as paraphrased here from the opening section of the reviewed report, is that NARWs may be considered for reclassification from endangered to threatened when all the following recovery criteria have been met:

1. The population ecology and vital rates of NARWs are indicative of an increasing population;
2. The population has increased at an average per annum rate ≥ 2% for a period of 35 years;
3. NARW has no known threats that limit its current growth potential;
4. Under current and projected conditions, the NARW population has ≤1% of dropping below quasi-extinction in the next 100 years.

NMFS, in terms of fulfilling its mandate to implement management regulations that protect the NARW population, needs to assess the quality of the analyses undertaken by the PET team. Thus, NMFS seeks scientific peer review of the PET team's work through independent review. The critique provided here is part of such a review, organized under the purview of the Center for Independent Experts (CIE). This critique, along with two others conducted completely independently of one another, is based on the following materials supplied to the reviewers by NMFS in September 2022, listed in Appendix 1.

In addition, publications consulted that are not listed in the reference section of the report under review, are listed in the references section at the end of this critique.

Critique Context

My critique is undertaken, and my comments are made in the context of this report representing Phase I of a PET team component that, it appears, will unfold in five-year phases linked to an expected regular five-year review of the status of NARWs. Thus, my comments are not just in the context of evaluating the report as a stand-alone study, but as a reporting milestone in a much longer process that is linked to the conservation of the NARW population under the NMFS mandate. For this reason, I believe, this report should have contained more details regarding future studies than were reported in the concluding ***9. Future Directions*** section of this report. Thus, I will include some discussion of what I expected to see.

I think it must also be acknowledged at the outset that no right way exists to carry out a population study of the type contained in the report under review. There are certainly incorrect procedures that can be followed—but the authors of this report are very experienced practitioners in their field and all their analyses appear to me to be legitimate. But otherwise, it is an issue of what might be the most appropriate way to address the questions at hand or carry out the tasks assigned. In this context, the merits of various approaches can be vigorously debated, and any comments that I make are made in the spirit of such a debate. Thus, my criticisms are ultimately debating points that are made from my own experiences in studying somewhat different types of systems (fisheries and terrestrial wildlife management; behavioral, movement, and disease ecology) than the NARW population that is the subject of the report under review.

Biological population modeling presents a rather challenging problem given the hierarchy of complexities that can be included such as genetics, physiology, behavior, demography, food web interactions, ecosystems and impinging anthropogenic processes at these various levels. Thus, many different tacks are possible to obtaining an appropriate model for addressing the questions at hand (Larsen 2016, Getz 2018). The scope of the analysis undertaken is considerable, with one tack not necessarily being demonstrably superior to another. The authors have done a laudable job in combining various methods of analysis, so the comments that follow are not meant to criticize the authors for their accomplishments to date—only to help them identify ways that I think will help meet them move forward in their assigned tasks.

Response to Questions

In the *terms of reference* provided in Annex 2 of the Performance Work Statement (see Appendix 2), the reviewers were asked to address three questions, numbered Q1 to Q3 below. In this section I will directly respond as best I can to these three questions (questions in bold type, my response in roman font) and then, in the following section, I will comment on several additional issues that I think are important to my review.

Q1. Based on the scientific information and analyses presented, does this report consider all of the best available data and represent an appropriate approach? If not, please indicate what information or analysis is missing and if possible, provide sources. When considering this question, please keep in mind the context in which the model was developed as provided in the model documentation. The model is not designed to consider all factors that may impact the population.

The authors listed the following data sets that they used for model parameter estimation:

1. Sightings data: a NARW Consortium list of sightings, curated by the New England Aquarium**.**

2. Carcass recovery data: a Northeast Fisheries Center aggregation of recovery reports gathered and maintained by multiple marine mammal stranding networks situated along the Atlantic coasts of US and Canada.
3. Sightings history data**:** histories were constructed from summer (1 April to 30 September, years unspecified; location and years unspecified in this section of the report, presumably summer feeding grounds) and winter (1 December to 30 March; southern calving grounds, years not specified in this section of the report) surveys.
4. Prey data: a mix of prey availability indices and biomass estimates for the Eastern Gulf of Maine (GOM, Georges Bank, and southwestern Gulf of St Lawrence (GSL) for different ranges of years for the different regions that were reduced to relative abundance measures over the period 1986-2019 for the GOM and GSL feeding grounds (Fig. 4 in the report).

Since I have not worked on any marine mammal systems myself, I am unaware of any other data that the authors could have used for their model fitting and analyses. It does strike me, however, that a much clearer exposition of the current state of the data could be made and, I would think, warrants a study of its own on whether a more coherent curation and aggregation of NARW demography data is needed and can be implemented during the subsequent phases of the PET team's work. In particular, it is not clear to me exactly which data were used to estimate which parameters in the model (this could be more clearly indicated in supplemental Table S10), particularly with regard to the overlap or complementarity of the Sightings (set 1.) and Sightings histories (set 3.) data. Also there appears to be no splitting of data to be used in model estimation versus model validation; and, in fact there is no discussion of model validation anywhere in the report (as I elaborate in my discussion below of Issue 4).

Q2. Are the baseline scenarios and use of demographic rates during 2010–2019 as the reference for most of the demographic processes appropriate for the analysis? If not, please indicate what considerations are missing and whether/why other periods should be used.

I think the various baseline scenarios are obvious and legitimate ones to undertake. So many different options exist that, ultimately, the most appropriate choice of scenarios to study are directly related to the questions that need to be address (e.g., see Alcamo and Henrichs, 2008; Kosow and Gaßner, 2008). I think this report does a credible and thorough job regarding its scenario analyses, the only suggestion that I have here is that I would have liked to see in the sensitivity analysis a greater focus on a probability of extinction than a demographic rate metric (Fig. 20 in the report). After all, in the context of protecting a species, the primary focus should not be demographic rate maximization but on minimizing quasi-extinction probability. The two metrics are certainly linked but a demographic growth rate with a lower mean that has relatively low variance may yield a lower extinction probability than one with a higher mean that has higher variance.

Q3. In general, are the scientific conclusions in the reports sound and interpreted appropriately from the information? Have the sources of uncertainty and caveats in the analyses been adequately described? If not, please indicate why not and if possible, provide sources of information on which to rely.

Every computational model of a system as complex as the demography of NARW, with its small population size, complex spatiotemporal dynamics, in an environment undergoing rapid global change in climate and anthropogenic sources of physiological stress and mortality related to fishing and shipping activities, and considerable sound and water pollution, can be heavily criticized for its simplifying assumptions, omission of processes, and methods of analysis. In this context, every reviewer lives in a glass house with respect to his/her/their own work, so it is a question of how a review can be most constructive without being unnecessarily critical. From this perspective, the analysts have done as good a job as can be expected, given the available data, team resource and time constraints, and given that the team decided to omit spatiotemporal structure at this time from their analysis. As will become clear below, I do not agree with the decision to omit this structure at this phase of the study, but it is not clear either that if they had included some spatiotemporal structure that they would have reached more salient conclusions given the current state of available data. The primary reason for including spatiotemporal structure now is to increase the value of the model as a management tool in the future and to encourage the collection of data in the future that contains more spatiotemporal information, particularly with respect to marine shipping activities, and anticipated movement activities of NARW individuals in the context of global change.

The best one can hope for from the kind of analysis undertaken in the report under review is to greatly increase our insights into how vulnerable the population is to extinction (or quasi-extinction in the context of a species-specific definition of a minimal viable wild population configuration) on how this vulnerability can be mitigated through relevant management actions (whether or not such actions can be easily implemented at this time). In this sense, the team has performed well and the conclusions from their analyses are, in my opinion, sound and supportable. As hinted to in some of my comments thus far, I think more accurate presentation of their model should have been provided along with a clearer presentation of their methods, and a deeper analysis of the robustness of results. All this would have required additional time, effort, and resources, and the actual results so obtained may have provided no more insight than those presented in the report as it currently stands. In future, however, the inclusion of spatiotemporal structure will greatly enhance the utility of the model as a tool for mitigating the negative impacts of humans and global change on NARW population, provided appropriate spatiotemporally structured data are collected.

Additional Comments

<u>Issue 1. Mathematical exposition</u>: The report falls short in providing a proper mathematical description of the model used and details of the implemented analyses

Given the complexity of the different elements that go into the analyses presented in this current PET team report, it is tempting for authors of reports of this scope to cut corners when presenting the details of the models and how various analyses were undertaken. In my opinion, the authors have succumbed to this temptation—somewhat understandably to cut down on the workload. This, however, has resulted in a “jury-rigged” presentation of the model that remains unclear in parts and impossible for others to check the computations of the presented results (e.g., estimates of parameters, scenario simulations), as well as of the details of the methods used to estimate parameters.

Model description. The authors have implemented an individually based model (IBM) that "… accounts for age- and stage-specific survival and reproductive rates, the effects of severe injury from entanglement or vessel strike, and future changes in prey availability and accessibility." They depict the demography structure of this model in Fig. 1 of their report. Even though their model is an IBM, they then formulate a discrete time, compartmental systems model, based on the structured depicted in Fig. 1. This model is expressed as a set of deterministic state transition equations numbered 1-8, and auxiliary equations 9-10 to implement the assumption of an equal sex ratio. The details of how to compute the survival parameters in the 8 transition equations are then respectively elaborated in equations 11-15 (survival) in terms of mortality hazard rate functions h (with super and subscripts indexing type of hazard, particular individual, and time—I will drop references to this scripting below) that in turn depend on mortality parameters α (with super and subscript indexing pertaining to various factors such as age, hazard type, and reproductive state/stage) and injury state indices W (subscripted to denote individual and time). Noise terms are also added to the h function equations, signaling that numerical computations involving these equations will have to be implemented as Monte Carlo simulations to represent solutions as distributions with means and standard deviations. The reproduction sub model is expressed as a logit equation that expresses the probability of reproduction in terms of various sub- and superscripted parameters that include an indicator function denoting whether an individual has been severely wounded in the current year and thus less likely to successfully reproduce, the effect of prey location, as well as past prey levels on reproduction, and a noise term that characterizes a stochastic environment. The injury state indices, W (with various subscripts) are themselves expressed in terms of severe injury functions ψ (with subscripts) from both entanglement and vessel strikes that depend on the hazard functions h (with indexing super and subscripts). Finally, the authors present a prey sub model in equations 24-27 using some rather odd notational conventions (words must be in roman rather than italics to avoid possible confusion with a string of parameters multiplied together) and loose concepts (e.g., rolling average—such averages can be taken in many ways depending on how we discount past values).

This would all be well and mostly good (not entirely good because the presentation is sloppy, e.g.: i) the equations are not presented in a clean, coherent manner; ii.) words such as "prey" are written in italics when they should be written in roman font; iii.) the vector in equation 17 has 10 dashes above β_5 so I am not sure what is being represented), except for the fact that ***the equations should be for the IBM model*** that they ultimately implement! The authors later provide some text to describe how they used the compartmental model description in their report to implement an individual-based computation, but this description is incomplete and deficient in parts. For example, in the time loop that makes random Bernoulli draws for injury, mortality and reproduction (Fig. 3 of the report), is reproduction considered before or after mortality? Also, when one has competing mortality rates, the draws are multinomial rather than binomial (i.e., Bernoulli), as described in Getz et al., (2021, 2022). In addition, it is not clear to me what initial conditions were used in the simulations. For example, did the initial conditions include uncertainty (it seems uncertainty was not included here according to text on page 44 of the report—see minor comments below)?

Parameter estimation. From Table S1, I count 30 parameters that have been estimated using various approaches. It is not clear to me from the report which of these were estimated from which data sets. This of course could be indicated in the Table S1. In Section 5, the authors make it clear that they estimated mortality parameters while treating reproductive parameters as attributes and vice versa. However, they do not go into details how they picked the value for that attribute parameter values while fitting the remaining parameters. When it comes to parameter estimation, simultaneously fitting more than a half dozen parameters at once presents considerable challenges. The authors need to report on how they overcame these challenges and discuss the robustness of their results (Gábor and Banga, 2015). For example, how difficult was it to get convergence? Did they converge to the same solutions from various starting conditions, and so on? These details need to be presented in a supplementary file that makes explicit how they obtained their fits and how robust these were.

<u>Issue 2. Spatiotemporal structure</u>: The spatiotemporal structure is critical to assessing management and modeling NARW demography

The most direct anthropogenic effects on NARW demography are the mortalities due to entanglement and vessel strikes (see Fig. 21 in the report under review). As the authors express in Section 8.4.4 "… the risk of vessel strikes is influenced by the spatiotemporal overlap between whales and vessel traffic." They then go on to say: "It is unknown whether or not the current movement and residency patterns of NARW will persist into the future." Finally, eleven lines below this they write: "The PET model is not designed to capture these complex spatial and temporal dynamics, but it is rather intended to evaluate the net or cumulative effects of changes in vessel strike mortality rate on the NARW population." Thus, the authors are fully aware of the importance of spatiotemporal structure and under global change exposure of NARW to the heavier gear and stronger ropes is that occurring (see the last sentence on page 14 of their report). However, they declined to incorporate spatial structure in their first phase analysis on the grounds that data are not currently available to support an analysis that includes spatial structure. I can appreciate this point of view, but it limits that authors ability to address the question: "If we had sufficient data to account for the most important spatial processes including mapping out the areas where entanglement and strikes occur as a function of the seasonal movement of NARWs, how much more effective could the manipulation of fishing and shipping activities be than if we ignore spatiotemporal structure, taking into account the economic impacts of such manipulations?" I believe addressing such questions lies at the core of whether it is going to be possible to develop fishing and shipping management policies that are economically viable for the fishery and shipping industries while ensuring the robust recovery of the NARW population. The sooner these structures are added to the model and embedded in a model adequacy assessment as articulated next the better it will be for obtaining the needed policies to ensure NARW survival.

Finally, in considering spatial structure some attention should likely be paid to how the current range differs from the historic range so that we can come to some understanding of why the present-day feeding grounds represent only the southern margin of the pre-whaling feeding grounds that occupied much of the Northwest Atlantic sector (Greene and Pershing 2004) and how the current range may change in the future.

Issue 3. Model adequacy assessment: A scheme should be in place as soon as possible for identifying structural deficiencies in the model and gaps in the data

In their abstract the authors write that they view their models "… as a living tool, that can be improved, adapted, and extended as new data, new methods, and new questions arise." In concluding Sections 9.3 and 9.4, they discuss how their model may achieve this. The discussion, however, is rather vague with phrases such as "Improving these models may include changes to data collection …" and "… empirical analyses associated with entanglement and vessel strike could also be coupled with development of threat scenarios that are more specific and mechanistic." At this point in time, I would have liked to see a much more detailed plan of how the team intends to increase the utility of the tools that they are developing for population evaluation, and especially in the context of spatiotemporal model structures that would enhance management analyses but would require data that has considerably more spatiotemporal structure than currently available. This should be done in the context of specific questions or problems to be solved (e.g., see Figure from Getz et al. 2018 reproduced below or Grimm et al. 2014) that if adequately addressed or solved would greatly mitigate anthropogenic sources of mortality. The most important questions to address require that the PET and decisions support tool (DST) teams work together to formulate these questions. The authors acknowledge in Section 9.2 that "One possible way the utility of the NARW population evaluation tool could be expanded is by carefully linking it to the Decision Support Tool developed by NOAA as part of the 2021 Atlantic Large Whale Take Reduction Rule (86 FR 51970)." I believe that the authors should have made a much stronger statement than this at this time along the lines of: "*To realize the utility of the NARW population evaluation tool it is absolutely essential that it be linked* to the Decision Support Tool developed by NOAA …"

In the context of data deficiencies that need to be rectified to obtain models better suited to answering the questions at hand, data on prey availability are woefully inadequate. This is particularly important in view of Harcourt et al.'s (2019) review of the population status of the three extant right whale populations in which they comment that "*Recent reproductive declines in NARW appear linked to changing food resources. While we know some large-scale movement patterns for NARW and a few SRW populations, we know little of mesoscale movements. For NPRW and some SRW populations, even broad-scale movements are poorly understood. In the face of climate change, can methodological advances help identify Eubalaena distributional and migratory responses.*" The same can be said about data relating to variation in hazard rates due to the spatiotemporal structure of NARW seasonal movements and the obvious spatial structure of shipping and fishing activities. By having spatial structure within the model at this time, the demand for spatial data may be made more compelling through scenario studies that consider the importance of accounting for spatial structure in management analyses.

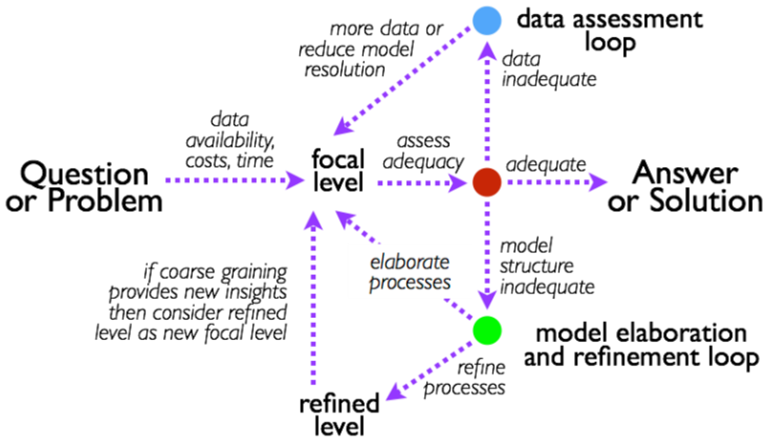

Figure from Getz et al, 2018. *Assessing model adequacy. The process of assessing model adequacy to provide an answer to a question or a solution to a problem begins at the focal level at which the model was initially formulated. Adequacy is assessed through evaluation of the relative benefits of new data and modification of the model structure (red circle). The decision might be to collect more data (blue circle) or elaborate the model (green circle) or both. The process is iterative and may lead to an increase or decrease in the structural complexity of the model.*

<u>Issue 4. Code verification and model validation and performance evaluation</u>: Confidence in the validity of the model, the veracity of the results, and utility of performance would be greatly enhanced with a fuller discussion of these issues in the PET team's phase 2 report (presumably for evaluation in five years)

Beyond careful documentation of the model, also needed are provision and verification of the code (does the code faithfully execute the mathematical formulation of the model), and validation of the model using independent data sets to assess its performance in predicting outcomes (Schuwirth et al., 2019). In particular, it is important for decision makers to have confidence that a model is a sufficiently good representation of the system being managed and have confidence that the solutions provided do indeed adequately address questions at hand. This requires models to be thoroughly vetted in terms of evaluating their utility and validating their output. The report as it stands: 1.) does not make the code available for perusal by other users; 2.) contains no discussion of any efforts made to verify that the code implements the model itself (which is not even possible when the mathematical description of the model, as discussed under Issue 1, is incomplete); 3.) completely ignores the issue of model validation; and 4.) falls short of a comprehensive discussion of the extent to which the utility of the model is hampered by its lack of spatiotemporal considerations, as discussed under Issue 3. I would thus recommend that in the next phase of the PET team's activities that they take the whole issue of "evaludation" (Augusiak et al., 2014) more seriously. If the authors are interested, they could do this by following the TRACE (TRAnsparent and Comprehensive Ecological modelling documentation) procedure, as outlined in Grimm et al. (2014).

Minor comments at identified places in the report

Page 10, second point from top: While I agree with the authors' neglect of density dependent process at this stage of the analysis, two issue should be examined more closely: 1.) if resources are limiting, as it occurs in the calving probability function graphed in Fig. 7 of the report, then location depletion of resources may cause individuals to move more often when individuals are in larger than smaller feeding groups, hence inducing some density dependent effects (Getz, 1996; Zurell et al., 2015). Also, what about the issue of finding mates at low densities and also possible inbreeding depression when population numbers are small (both inducing an Allee effect; e.g., see Gascoigne et al., 2009 and Wittmann et al., 2018)?

Page 25, first paragraph of section 4.1.1.: I am perplexed by the statement "… calf survival does not depend on survival of its mother in either the IBM …" Is this a reasonable assumption? Surely not.

Page 28, text after eq. 14.: the approach taking in the mortality submodel is to assume that entanglements and strikes from previous years does not affect the current health of an individual. However, in an IBM it is simple enough to keep a record of past entanglements and strikes, so should the authors consider the effects of multiple strikes over multiple years on mortality? I tend to think it is not necessary, but worth raising the issue here.

Pages 32 (bottom) and 33 (top): The authors approach to density dependence is extremely abrupt and could be softened (e.g., see Getz, 1996), though the effect should be rather minor.

Pages 35, bottom 2 lines: Outright extinction is not 1 individual but rather no potential future breeding pair. Since this is an IBM that contains the relevant information, the authors can be more precise about outright extinction conditions. On a related matter, I much prefer Eq. 30 as a measure of extinction to any particular quasi-extinction condition.

Page 37, "The mortality and reproduction analyses both used multistate capture-recapture …": Did the capture-recapture methods used by the team to estimate parameters account for spatial heterogeneity (e.g., using the approach of McDonald, T.L. and Amstrup, S.C., 2001 or extensions to this as reviewed by Tourani, 2022)

Page 37, Sec 5.1.1. Precisely, how much was multiple sightings used to reduce the probability of missing severe wounds?

Page 40, start of third paragraph: It seems that a mathematical description of the relationship between true and observed states is needed.

Page 40, "We used a model selection ...": Provide specifics or a citation.

Page 41, "Imputed wounds ...": How were these estimates made?

Page 42, Section 5.3.2: If survival and reproduction estimates were produced in two ways, a comparison of the two sets of values obtained should be listed and some discussion of why one set rather than another used.

Page 44, "We used the results …": Not sure what this means. Provide reference to exact method used.

Page 45, "Therefore we considered the threat of prey limitation …": Could you have evaluated the effect of this assumption by comparing the two scenarios (and making this part of an adaptive management analysis—e.g., see Dutra et al. 2015)"?

Page 48, "Fourth, the mission half or two-thirds of the NARW…": This point just reinforces the importance of including spatial structure in the model.

Page 48, Section 6.4: As I have already mentioned, I would have preferred to see a sensitivity analysis with respect to some extinction measure or with respect to ***Expected minimum population size*** measure.

Page 67, Section 8.4.5: The study of this shift needs to be made a priority in the next phase of the PET team's work.

Page 70, "… use the model to estimate potential biological removal rates …": It would have been good to see a more detail discussion of this with reference to the recent work of Punt et al (2020).

Conclusions and Recommendations

Ultimately, it is a qualitative rather than quantitative understanding provided by the kinds of models and analysis presented in this report that motivates actions most likely to preserve the species of concern. This understanding is then used to prescribe possible recovery plans. In the light of possible recovery plans laid out in this report, the future work of the PET team is arguably best accomplished by their models being used, as the team articulate in their concluding section, to derive recovery metrics that are estimated under the best current analytical methods available to the PET team to meet recovery criteria. Thus, for example, the PET team's analyses could estimate what combination of survival rates, reproductive rates, entanglement rates, vessel-strike rates, and so on, are needed to achieve these recovery requirements. Such analyses, of course, would have to be undertaken in collaboration with the Decision Support Tool team who would have to assess which of various management actions it can most plausibly implement. In short, it is not just about the models, which themselves may be quite poor predictors of future stock levels, but how these models can guide monitoring and be incorporated in an adaptive response process that is agile, risk averse, and has the confidence of the polity in terms of being able to influence positive actions to preserve the NARW population.

To improve the stated usefulness of tools that the PET team are developing and the analyses they will undertake in the next 5-year phase, I have the following recommendations that, to a considerable extent, the authors are already thinking of incorporating into their future studies (as articulated in their final “Future Directions” section of the report). My list here is what I deem to be particularly important at this point in time:

The PET team should

- Provide a proper mathematical description of all models and analytical methods used, or reference methods and software packages in a way that leaves no doubt about the computations that were undertaken: in particular, they should provide a more accurate and transparent representation of the variables used to characterize the state of individuals life-histories, spatio-temporal locations, and current physical/physiological well-being.

- Provide all data and code used in a form that others can then use to run any of the computations discussed in the report. I don’t really expect anyone to try to repeat anything more than one or two selected parts of their study, but some researchers may want to implement some of the methods on their system and check they are doing so correctly by repeating the relevant analysis described in this report to ensure they get the same results.

- Provide a more complete description of an integrated scheme for carrying out sensitivity, model-adequacy, and adaptive management analyses, to better inform identification of the most pertinent directions in new data collection and information gathering as the PET process moves forward over the next decades: i.e., procedures designed to identify i.) parts of the model that need to be either changed or elaborated, ii.) gaps in data that need to be filled, iii.) what kinds of data are needed to better address the questions and issues at hand, and iv) what questions and issues have become the most pressing to address.

- Immediately begin to examine the utility of adding spatiotemporal structure to the model in the context of getting a better handle on anticipating i.) changes in the ecological consequences for whale demography and in the movement patterns of whales in the context of global change, and ii.) spatiotemporal aspects of mortality risk reduction with appropriate handles in the model for investigating risk mitigation through the management of shipping activities.

- Formalize the relationship between the activities of the Population Evaluation Tool and the Decision Support Tool groups or, at least better articulate how these two groups will interact.

Additional References

Alcamo, J. and Henrichs, T., 2008. Chapter two towards guidelines for environmental scenario analysis. *Developments in integrated environmental assessment*, *2*, pp.13-35.

Augusiak, J., Van den Brink, P.J. and Grimm, V., 2014. Merging validation and evaluation of ecological models to 'evaludation': A review of terminology and a practical approach. *Ecological Modelling*, *280*, pp.117-128.

Dutra LX, Thebaud O, Boschetti F, Smith AD, Dichmont CM. Key issues and drivers affecting coastal and marine resource decisions: Participatory management strategy evaluation to support adaptive management. Ocean & Coastal Management. 2015;116:382-95.

Gábor, A. and Banga, J.R., 2015. Robust and efficient parameter estimation in dynamic models of biological systems. *BMC systems biology*, *9*(1), pp.1-25.

Gascoigne, J., Berec, L., Gregory, S. and Courchamp, F., 2009. Dangerously few liaisons: a review of mate-finding Allee effects. *Population Ecology*, *51*(3), pp.355-372.

Getz, W.M., 1996. A hypothesis regarding the abruptness of density dependence and the growth rate of populations. *Ecology*, *77*(7), pp.2014-2026.

Getz, W.M., Marshall, C.R., Carlson, C.J., Giuggioli, L., Ryan, S.J., Romañach, S.S., Boettiger, C., Chamberlain, S.D., Larsen, L., D'Odorico, P. and O'Sullivan, D., 2018. Making ecological models adequate. *Ecology letters*, 21(2), pp.153-166.

Getz, W.M., Salter, R., Luisa Vissat, L., Koopman, J.S. and Simon, C.P., 2021. A runtime alterable epidemic model with genetic drift, waning immunity and vaccinations. *Journal of the Royal Society Interface*, *18*(184), p.20210648.

Getz, W.M., Salter, R. and Vissat, L.L., 2022. Simulation applications to support teaching and research in epidemiological dynamics. *BMC medical education*, *22*(1), pp.1-32.

Greene, C.H. and Pershing, A.J., 2004. Climate and the conservation biology of North Atlantic right whales: the right whale at the wrong time?. *Frontiers in Ecology and the Environment*, *2*(1), pp.29-34.

Grimm, V., Augusiak, J., Focks, A., Frank, B.M., Gabsi, F., Johnston, A.S., Liu, C., Martin, B.T., Meli, M., Radchuk, V. and Thorbek, P., 2014. Towards better modelling and decision support: documenting model development, testing, and analysis using TRACE. *Ecological modelling*, *280*, pp.129-139.

Kosow, H. and Gaßner, R., 2008. *Methods of future and scenario analysis: overview, assessment, and selection criteria* (Vol. 39, p. 133). DEU.

Larsen, L.G., Eppinga, M.B., Passalacqua, P., Getz, W.M., Rose, K.A. and Liang, M., 2016. Appropriate complexity landscape modeling. *Earth-science reviews*, 160, pp.111-130.

McDonald, T.L. and Amstrup, S.C., 2001. Estimation of population size using open capture-recapture models. *Journal of Agricultural, Biological, and Environmental Statistics*, *6*(2), pp.206-220.

Punt AE, Siple M, Francis TB, Hammond PS, Heinemann D, Long KJ, Moore JE, Sepúlveda M, Reeves RR, Sigurðsson GM, Vikingsson G. Robustness of potential biological removal to monitoring, environmental, and management uncertainties. ICES Journal of Marine Science. 2020 Dec;77(7-8):2491-507.

Schuwirth, N., Borgwardt, F., Domisch, S., Friedrichs, M., Kattwinkel, M., Kneis, D., Kuemmerlen, M., Langhans, S.D., Martínez-López, J. and Vermeiren, P., 2019. How to make ecological models useful for environmental management. *Ecological Modelling*, *411*, p.108784.

Tourani, M., 2022. A review of spatial capture–recapture: Ecological insights, limitations, and prospects. *Ecology and Evolution*, *12*(1), p.e8468.

Wittmann, M.J., Stuis, H. and Metzler, D., 2018. Genetic Allee effects and their interaction with ecological Allee effects. *Journal of Animal Ecology*, *87*(1), pp.11-23

Zurell, D., Eggers, U., Kaatz, M., Rotics, S., Sapir, N., Wikelski, M., Nathan, R. and Jeltsch, F., 2015. Individual-based modelling of resource competition to predict density-dependent population dynamics: a case study with white storks. *Oikos*, *124*(3), pp.319-330.

Appendix 1: Bibliography of Materials Provided for Review

The following materials were supplied by NMFS for the review.

1. The draft report dated August 2022: ***A Management-focused Population Viability Analysis for North Atlantic Right Whales*** (authors: Michael C. Runge, Daniel W. Linden, Jeffrey A. Hostetler, Diane L. Borggaard, Lance P. Garrison, Amy R. Knowlton, Véronique Lesage, Rob Williams, Richard M. Pace, III)

2. The following two papers as background documents.

   <u>Pace et al. 2017</u>: Pace III RM, Corkeron PJ, Kraus SD. *State–space mark–recapture estimates reveal a recent decline in abundance of North Atlantic right whales.* Ecology and Evolution. 2017 Nov;7(21):8730-41.

   <u>Pace et al. 2020</u>: Pace III RM, Williams R, Kraus SD, Knowlton AR, Pettis HM. *Cryptic mortality of North Atlantic right whales*. Conservation Science and Practice. 2021 Feb;3(2): e346.

3. The following unpublished reports.

   <u>MMSA 2022</u>: Anonymous. *Draft US Atlantic and Gulf of Mexico Marine Mammal Stock Assessments 2021* plus updated section on *North Atlantic Right Whale (Eubalaena glacialis): Western Atlantic Stock.*

   <u>Linden et al. 2022</u>: Linden DW, JA Hostetler, RM Pace III, LP Garrison, AR Knowlton, V Lesage, MC Rung, and R. Williams. Multistate capture-recapture models to estimate mortality and reproduction in North Atlantic right whales (1990-2019).

Appendix 2: CIE Statement of Work

Performance Work Statement (PWS)
National Oceanic and Atmospheric Administration (NOAA)
National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program

External Independent Peer Review

North Atlantic Right Whale Population Viability Analysis

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act (ESA), and Marine Mammal Protection Act (MMPA) to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS 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]. Further information on the Center for Independent Experts (CIE) program may be obtained from www.ciereviews.org.

Scope

NMFS Greater Atlantic Region established the Population Evaluation Tool Subgroup under the North Atlantic Right Whale (NARW) Recovery Plan U.S. Implementation Team to assist NMFS in the implementation of the North Atlantic Right Whale Recovery Plan. The intention was to bring together the diversity of expertise most appropriate to develop a population viability analysis (PVA) for NARW. The Population Evaluation Tool Subgroup[2] consists of appropriate

---

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

[2] PET Subgroup Members: Dr. Richard Pace, Chair, NOAA Fisheries, Northeast Fisheries Science Center; Dr. Michael Runge, U.S. Geological Survey; Dr. Lance Garrison, NOAA Fisheries, Southeast Fisheries Science Center; Dr. Jeffrey Hostetler, U.S. Fish and Wildlife Service; Amy Knowlton, New England Aquarium; Dr. Veronique Lesage, Fisheries and Oceans Canada; Dr. Daniel Linden, NOAA Fisheries, Greater Atlantic Regional Fisheries Office; Dr. Rob Williams, ORCA

experts in integrated population models and/or population viability analyses. The need for a PVA was highlighted most recently in NOAA Fisheries' 5-year reviews for NARW (August 2012 and October 2017), required under the ESA to ensure that the listing classification of the species is accurate. The objective of the Population Evaluation Tool Subgroup is to develop a population viability analysis that will allow the agency to characterize the North Atlantic right whale extinction risk, taking into account current and future threats. This modeling effort is underway and a final report is expected in 2022 which will help identify demographic benchmarks useful to inform management and gaps in research.

NMFS is required to use the best available scientific and commercial data in making determinations and decisions under the ESA and MMPA. Given the importance of this effort and likely use in management discussions under the ESA and/or MMPA, it is critical that the PVA be based on the best available science and be statistically sound. Therefore, the CIE reviewers will conduct a peer review of the scientific information and approach in the North Atlantic right whale PVA based on the Terms of Reference (TORs) referenced below. Given the public interest, it will be important for NMFS to have a transparent and independent review process of the model used in future considerations to further the recovery of right whales.

The specified format and contents of the individual peer review reports are found in Annex 1. The Terms of Reference (TORs) of the peer review are listed in Annex 2.

Requirements

NMFS requires three (3) reviewers to conduct an impartial and independent peer review in accordance with the PWS, OMB guidelines, and the TORs below. The reviewers shall have working knowledge and recent experience in one or more of the following: (1) wildlife population modeling; (2) population viability analyses; and/or (3) quantitative ecology. In addition, experience with large whale science is helpful, though not required. Each CIE reviewer's duties shall not exceed a maximum of 10 days to complete all work tasks of the peer review described herein.

Tasks for Reviewers

Each CIE reviewer shall complete the following tasks in accordance with the PWS and Schedule of Milestones and Deliverables herein.

> **1)** Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send by electronic mail or make available at an FTP site to the CIE reviewer all necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE on where to send documents. The CIE reviewer shall read all documents in preparation for the peer review, for example:
>
> Pace III, R.M., P.J. Cockeron, S. D. Krause. 2017. State-space mark-recapture estimates reveal a recent decline in abundance of North Atlantic right whales. Ecology and Evolution. 7:8730-8741 . DOI: 10.1002/ece3.3406

Pace, RM, III, R. Williams, S.D. Kraus, A.R. Knowlton, H.M. Pettis. 2021. Cryptic mortality of North Atlantic right whales. Conservation Science and Practice. https://doi.org/10.1111/csp2.346

NMFS, 2021. North Atlantic right whale *(Eubalaena glacialis*). Draft U.S. Atlantic and Gulf of Mexico Marine Mammal Stock Assessments 2021. Pages 22-48. https://media.fisheries.noaa.gov/2021-10/Draft%202021%20NE%26SE%20SARs.pdf

2) Webinar: Additionally, approximately two weeks prior to the peer review, the CIE reviewers will participate in a webinar with the NMFS Project Contact and Population Evaluation Tool Subgroup members to address any clarifications that the reviewers may have regarding the ToRs or the review process. The NMFS Project Contact will provide the information for the arrangements for this webinar.

3) Desk Review: Each CIE reviewer shall conduct the independent peer review in accordance with the PWS and TORs, and shall not serve in any other role unless specified herein. 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.

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

Place of Performance

Each CIE reviewer shall conduct an independent peer review as a desk review, therefore no travel is required.

Period of Performance

The period of performance shall be from the time of award through October 31, 2022. The CIE reviewers' duties shall not exceed 10 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.

Within two weeks of award	Contractor selects and confirms reviewers
No later than two weeks prior to the review	Contractor provides the pre-review documents to the reviewers

August 2022	Each reviewer conducts an independent peer review as a desk review
Within two weeks after review	Contractor receives draft reports
Within two weeks of receiving draft reports	Contractor submits final reports to the Government

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; and (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel

Since this is a desk review travel is neither required nor authorized for this contract.

Restricted or Limited Use of Data

The contractors may be required to sign and adhere to a non-disclosure agreement.

Project Contact:

Diane Borggaard
diane.borggaard@noaa.gov
NMFS, Greater Atlantic Region
55 Great Republic Drive, Gloucester, MA 01930

Annex 1: Peer Review Report Requirements

1. The report must be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether or not the science reviewed is the best scientific information available.

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

3. The reviewer report shall include the following appendices:

    1. Appendix 1: Bibliography of materials provided for review

    2. Appendix 2: A copy of the CIE Performance Work Statement

Annex 2: Terms of Reference for the Peer Review

The reviewers will provide input on the following questions:

1. Based on the scientific information and analyses presented, does this report consider all of the best available data and represent an appropriate approach? If not, please indicate what information or analysis is missing and if possible, provide sources. When considering this question, please keep in mind the context in which the model was developed as provided in the model documentation. The model is not designed to consider all factors that may impact the population.

2. Are the baseline scenarios and use of demographic rates during 2010–2019 as the reference for most of the demographic processes appropriate for the analysis? If not, please indicate what considerations are missing and whether/why other periods should be used.

3. In general, are the scientific conclusions in the reports sound and interpreted appropriately from the information? Have the sources of uncertainty and caveats in the analyses been adequately described? If not, please indicate why not and if possible, provide sources of information on which to rely.