Center for Independent Experts Independent Peer Reviews of the North Atlantic Right Whale Population Viability Analysis

Trevor A. Branch
School of Aquatic and Fishery Sciences
University of Washington

Wayne M. Getz
College of Natural Resources
University of California at Berkeley

Andrew J. Read
Duke University Marine Laboratory

October 2022

Prepared by the Center for Independent Experts
Under Contract to
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Center for Independent Experts Program

Center for Independent Experts (CIE) Independent Peer Review of North Atlantic Right Whale Population Viability Analysis

Dr. Trevor A. Branch, Professor
School of Aquatic and Fishery Sciences
University of Washington

October 2022

Executive summary

The suite of models used to conduct the population viability analysis of North Atlantic right whales use modern modeling methods, are based on the best available data, and include all major factors known to affect this population. The models fit well to available data, and represent the best available science to predict risks to this population, and estimate the impact on the population of changes to prey availability, entanglement with fishing gear, ship strikes, and noise. The framework presented should be able to adjust to a variety of current and future threats, new data, and updates to management. Additional scenarios would improve the validation of the model, as follows: (1) Conduct and present diagnostics to demonstrate that the model fits are not biased near the end of the period of available data. (2) Include a regime shifts scenario that is able to mimic the periods of good and bad years for North Atlantic right whales, with some probability of future increases in the population that match past observed rates of increase. (3) Use a consistent period of time to model calving rates and injury and mortality rates. (4) Include a scenario where future calving rates are not driven by an uncertain relation between the prey abundance index but are resampled from past observed calving rates. (5) For simulations into the future, initialize the population based on individual probabilities of being alive at the end of the historical model, rather than multinomial draws.

None of these recommendations are likely to change the overall conclusions of the viability analysis paper: North Atlantic right whales are currently declining precipitously; this decline is largely due to entanglement injuries, with contributing factors including ship strikes and poor prey availability and abundance; and these declines can be reversed if entanglement is reduced by 25-50%.

Background

North Atlantic right whales are a small, endangered, population of whales whose numbers were recovering in abundance since protection from whaling in 1935. Abundance increased steadily from from 270 in 1990 to 483 in 2010, before a prolonged period of decline to the present abundance of 368 in 2019 (NOAA 2021). The recent period of decline appears to be driven by a combination of low calf production (likely due to a prolonged period of poor prey availability) and high mortality due mostly to entanglement with fishing gear, with deaths from ship strikes an additional contributing factor. In the population viability analysis reviewed here, the authors (Runge et al. 2022) have created an individual-based population model that can be projected forward to predict the probability of the population falling below threshold abundance levels (quasi-extinction), measured in terms of "proven" females, i.e. those observed with calves. They examine the effect of different levels of gear entanglement, ship strikes, and prey availability on the viability of North Atlantic right whales.

Description of individual reviewer's role in the review activities

I am one of three reviewers chosen by the Center for Independent Experts to conduct a desktop review of the population viability analysis for North Atlantic right whales (Runge et al. 2022). I am a professor at the University of Washington who works on assessments of fisheries and large whale populations, among other topics.

Major points

1. The assessment model used (Pace et al. 2017, Pace et al. 2021) and population viability analysis (Runge et al. 2022) employ modern methods, include comprehensive sets of data, and include

all factors known to affect this population. The models fit well to the available data, and represent the **best available science** for conducting a population viability analysis of North Atlantic right whales.

2. **Retrospective diagnostics** could be provided to assess any bias in the final years of the model. Mark-recapture models sometimes estimate survival that is biased low in the last few years of the model. For example, an earlier model of this population (Fujiwara & Caswell 2001) estimated mother survival of 0.63-0.78 during 1990-1995, while the most recent model estimates adult female survival to be >0.95 during 1990-1995 (Pace et al. 2017). This bias does not appear to affect the current model, however, as can be seen by comparing abundance estimates for 2010-2015 (Pace et al. 2017) with those for 2010-2015 in the latest version (Runge et al. 2022). A formal comparison of estimates of abundance from each successive annual model across common years, would be a good model validation check.

3. None of the model scenarios appear capable of reproducing the observed trends in the population between 1990 and 2010, since they are focused on projecting 2010s conditions into the future. During 1990-2010, the population grew from 270 to 483 (Pace et al. 2017) at a rate of 2.95% per year, at a rate allowing it to double in size every 24 years. Similar increases are likely in the 1970s and 1980s (e.g., Caswell et al. 1999). However, the projection model predicts zero probability of the population doubling in size even in a longer period of 35 years. Furthermore the 1990-2010 period also included a number of years with low estimated survival, low calf production, and high numbers of reported deaths (Caswell et al. 1999, Fujiwara & Caswell 2001), similar to that estimated since 2010. To fill this gap, I recommend including a **“Regime shifts” scenario**, which models the environment as alternating periods of good times and bad times. It appears that there have been four regimes since 1990: early 1990s (good), late 1990s (bad), 2000s (good), and 2010s (bad), thus to mimic this kind of interplay, regimes could be modeled by picking a random duration (e.g. 5-15 years), and quality (good or bad) in the proportions observed during 1990-2019 (roughly 2/3rd good years, 1/3rd bad years). In the good times, prey abundance is high and situated geographically as it was in the earlier period, resulting in high calving rates and crucially also having North Atlantic right whales shifted in space resulting in low levels of mortality from entanglement and ship strikes. Bad and good regimes need not alternate, thus if 2-3 good periods were strung together, the population would have a reasonable probability of doubling in size in less than 35 years.

4. In the baseline model (section 6.1), **inconsistent time periods** are used for prey-influenced calving rates (2010-19) and for injury and mortality rates (2014-19). This is a critical assumption because estimated mortality jumps sharply in 2014 from low rates (Figure 5), and 2017 (17 reported deaths) is the only year since 1990 with more than 7 reported deaths (Pace et al. 2021). Manufactured rope breaking strength increased in the mid-1990s (Knowlton et al. 2016), and a regime shift in food availability and right whale distribution occurred in 2010, resulting in a shift in right whale distribution away from existing grounds in 2010, and into more northerly areas in 2015 (Meyer-Gutbrod et al. 2021). Although there is certainly a case to be made for choosing different time periods that respectively include low average calving rates and high average mortality rates, the underlying causes of both are most likely the regime shift in 2010, and I therefore recommend that a consistent time period (2010-19) should be used to calculate both average prey-influenced calving rates, and average injury and mortality rates.A clear plot is

needed in the paper of the effects of the five key scenarios on historical and projected proven female survival and calving rates, including uncertainty. The five scenarios (six if the regimes model above is added) are the baseline, 25% and 50% reduction in entanglement, 25% reduction in ship strike risk, and prey levels at the 1990-2009 levels.

5. The model that predicts calving rates in the future relies heavily on the relationship between the *Calanus* prey abundance index and observed calving rates (Figures 7-8 in Runge et al. 2022). However, comparing the model predictions for calving rates from this relationship (Figure 8) with the actual observed calving rates (Figure 6, Pace et al. 2017) reveals considerably different patterns of calving rates over time. Notably, calving rates were generally high from 1990-1997, and low in 1998-2000, but in the Calanus predictions in Runge et al. (2022), calving rates are uniformly low during 1990-2000 — as low as in 2010-2019. For a start, Figures 7-8 should include the observed data on both plots so that the reader can better evaluate whether the model predictions are realistic. Secondly, it would be worth running a scenario based on empirical resampling of calving rates (i.e., random draws of calving rates) during different regimes (1990-1997, 1998-2000, 2001-2011, 2012-2019), rather than relying entirely on the relationship between the Calanus index and calving rates. In fisheries, identified relationships between environmental conditions and recruitment are often illusory when used to predict outcomes in future decades (Myers 1998).

6. The method used to initialize individuals in each class in year 0 of the forward simulations needs revision (Section 5.4), although this is unlikely to change the predicted outcomes of the model. It is described as "our retrospective reproduction model does terminate the time series with an estimate (1 or 0) of the alive state of each individual. We summed the terminal alive states according to the 18 classes of animals and derived proportions for each sex, age, and stage class. The posterior distribution of these proportions served as a multinomial probability from which a random draw of size N_hat (the posterior estimate from the state space model) was used to initialize the number of individuals in each class in year 0." I believe this is incorrect, but the distinction is subtle and tricky to explain. The key is that uncertainty resides at the individual level—the probability that each individual is alive or dead at the end of the retrospective time period (2019) and start of the simulations. To simplify, consider the case where there are 4 individuals in 2 classes (the sexes), with these probabilities of being alive: male 1.000, male 1.000, female 1.000, female 0.5. A multinomial draw would sample 3 or 4 individuals with probability 2/3.5 male and 1.5/3.5 female, and might result in a population of 2M2F, 4M0F, 1M3F, 2M1F (actual random multinomial draws). Clearly this is far more variable than the truth which should consist only of 50% probability of 2M1F and 50% probability of 2M2F. Therefore, instead, each starting point should be defined based on a random Bernoulli (binomial with n=1) draw for each individual in the population, with probability of success being the posterior probability of that individual being alive. Or, even more simply, each forward simulation could be based on one posterior draw of the individuals from the retrospective reproduction model, which would maintain the covariances among parameters in creating a starting population.

Minor points

p. 2 The first two paragraphs are somewhat duplicative. I suggest defining proven females (females previously detected with calves) and quasi-extinction (risk of falling below 50 proven females), and then starting the second paragraph with "We also explored scenarios that only partially remove threats, including the degree of entanglement risk…"

p. 14 first para: the estimates of female survival in these two papers are opposite, with Fujiwara & Caswell (1999) finding declining and low female survival, while Pace et al. (2017) found high female survival with no declining trend. A comment is needed here to clarify which scenario is currently thought to be most plausible and/or the reference to the outdated results should be removed.

p. 24 Model Implementation. A brief description should be added here to describe which computer language the model is implemented in, how long the runs take, and how many iterations of the model are run to estimate uncertainty in the baseline case.

p. 35 Equation 29. The subscript t is used for years throughout, but here changed to the Greek letter tau τ. Instead, minimum biomass should be t between 0 and τ with N_t subscript. The same issue is true in many equations in sections 4.5.1-4.5.4.

p. 36 The definition of the probability of decline (4.5.4) is ambiguous. For the model runs, there are i = 1000 replicates in each of t = 100 years, but the definition is missing for subscript i. It would be better defined along these lines: "For each model replicate i, the minimum abundance $N_{\min,i}$ within the 100 years is found, and the highest decline calculated as $(N_0–N_{\min})/N_0$. Then the proportion of these iterations that are greater than 0.3, 0.5, and 0.8 is calculated."

p. 36-37 The Recovery Potential Assessment (RPA) criteria (median time to reach 1000 mature individuals) need some clarification to explain what happens in scenarios when the median trajectory never reaches 1000 mature individuals. Perhaps "…when the median simulation did not exceed 1000 mature individuals within 100 years, we instead report the proportion of replicates that did exceed this threshold."

p. 36-37 For median time to reach 1000 mature individuals, please clarify how this is calculated: (a) for each model replicate calculate the year it exceeds 1000, then find the median (what to do when it never exceeds 1000?); or (b) for each year calculate whether median abundance is above 1000, and report the first such year.

Figure 4: caption should include the units and mention that this is in log-scale: "(log prey availability in mg dry weight.m^{-2})".

p. 39 "We found strong correlations between calving rate and the *Calanus* indices from eGOM and swGSL". The strength of the correlation (r^2) should be reported and Figure 7-8 should show the data and the model fit, since this has crucial bearing on predicted calving rates during 2010-19 and for the alternative run basing this on *Calanus* indices for 1990-2010.

p. 40 Section 5.2.2. Mortality rate estimates and confidence intervals are reported as hazard rate model parameter values, which are difficult to interpret. For example, 1.474 for vessel strike mortality appears to translate to <10% mortality (Figure 5). I recommend converting all the values in this paragraph to survival rates per year.

p. 43 Figures 7 & 8 need to include 0 on the y-axis. Additionally, both figures should include the actual data, not just the model predictions based on prey indices, so that the reader can judge how accurate the model predictions are.

p. 46 Section 6.3.1 "The risk reduction under full implementation was estimated at 90% for adults and 60% for juveniles. The weak-rope scenario was implemented at 50%..." This statement is unclear. Is the weak-rope scenario 50% risk reduction for adults and juveniles, or 45% and 30% respectively (50% of the risk reduction)?

p. 46 Section 6.3.2 No explanation is given for the choice of -0.3% and +0.7% per year for changes in future vessel strike mortality, or for why the declines are smaller than the increases.

p. 47 Section 6.3.3 Prey availability scenarios are assumed to be drawn randomly from either 1990-2019 ("steady") or 2010-2019 ("decline"). However, it is more likely that conditions come in regimes of good and bad conditions.

p. 47 Section 6.3.3 It is not clear to me whether the model includes the full uncertainty associated with the prey index values, and the full uncertainty associated with the relationship between the prey index and calving success. Values should include this uncertainty and not be taken from the median estimates in Figures 7 and 9.

p. 48 Section 6.3.4 Given the absence of any evidence that noise currently impacts right whale feeding or survival, I would suggest paring down the noise scenarios from six to three, and reducing the range of impacts to -10%, 0%, +10%.

p. 48 Section 6.4. The sensitivity analysis is worth doing, but it is not statistically valid to look for significant regressions when fitting to model outputs. The analysis can be retained but references to "statistically significant" outcomes should be removed throughout.

p. 50 Figure 10. Remove the gap on the y-axis between zero and the plotted values, especially important since this gap obscures population sizes close to extinction.

p. 52 Figure 11. Background colors should go from blue to white to red, with white showing steady state. It also took time to understand that the background color is simple math (birth rate minus death rate), which would be clearer with a dashed 1:1 line representing zero population growth.

p. 53 Figure 12. The y-axis label should be changed to "Probability of quasi-extinction", and "Threshold" renamed to "Proven female threshold".

p. 55 Figure 13. Include zero on the y-axis to accurately represent the decline in population size over time.

p. 56 Table 2. One of the threats is listed as "Prey", but should be named "Low prey" or "Lack of prey".

p. 57 Figures 14 and 15 should have the same y-axis so that the magnitude of threats are directly comparable.

p. 60 Section 7.3.4 Prey accessibility is modeled as ranging from -30% to +30% of the baseline scenario, which does not represent a reasonable range of outcomes. The baseline scenario (2010-19) is the worst observed, while 1990-2010 is the best observed, therefore the models should range from 30% below the worst (2010-19) to 30% above the best (1990-2010) conditions.

p. 60 Section 7.4 sensitivity analyses. More explanation is needed here. Only symbols are referred to in the text and figures (referencing Table S1 for definitions). Instead, for parameters deemed sensitive, the results should include the definition in plain English, how they impact population growth rates, and why they are important in the model (or submodels).

p. 63 “probability of which is approximately 0” -> replace with “probability of which is <0.0001” or similar.

Summary of findings related to terms of reference

The terms of reference for the CIE review specifically asks for discussion around the three topics below. Here I briefly respond.

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.

Yes, the report does consider all of the best available data and represents an appropriate approach. It focuses on projecting current conditions (roughly 2010-2019 for most assumptions) into the future, which is reasonable, especially when trying to predict the risk to North Atlantic right whales from anthropogenic causes.

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.

Yes, these scenarios are appropriate to forecast the risk of continuing the status quo in the 2010s in terms of the extinction risk to North Atlantic right whales. As I argue above, I would also include at least one scenario that models the system as being driven by different regimes during which prey availability, entanglement risk, ship strike risk, and calving rates are “good” for a period of years, followed by a period of “bad” years. This scenario would be better at reproducing the conditions that allowed the population to increase at around 3% per year in the past.

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.

Yes, the scientific conclusions are valid: a reduction in the entanglement risk would greatly reduce the quasi-extinction risk to the population, suggesting an immediate management response to turn around the population decline. Sources of uncertainty are well described and included in the model.

Conclusions and recommendations

Major recommendations arising from the review points above are as follows:

1. Conduct retrospective diagnostics to validate the model, to demonstrate that model fits are not biased at the end of the time series.
2. Include a regime shifts scenario that models good periods (like those in the early 1990s and 2000s) and bad periods (like the late 1990s and 2010s), and has some probability of the population returning to the population growth rates seen over 1990-2010.

3. Use a consistent period of time (2010-19) to model recent prey-influenced calving rates and injury and mortality rates, rather than using 2010-19 for the former and 2014-19 for the latter.
4. Include a scenario where calving rates are resampled from observed data in past periods of time, rather than being based on an uncertain relationship between prey availability and calving rates. Also provide diagnostics that compare predictions of calving rates from prey availability to past calving rates, to check how well this relationship matches the data.
5. Adjust the method used to initialize individuals in the first year of the simulations to better reflect the individual level at which uncertainty occurs (point 7 of the major points above).

Cited references

Caswell H, Fujiwara M, Brault S (1999) Declining survival probability threatens the North Atlantic right whale. Proceedings of the National Academy of Sciences 96:3308-3313

Fujiwara M, Caswell H (2001) Demography of the endangered North Atlantic right whale. Nature 414:537-541

Myers RA (1998) When do environment-recruitment correlations work? Reviews in Fish Biology and Fisheries 8:285-305

NOAA (2022) North Atlantic right whale (*Eubalaena glacialis*): Western Atlantic stock. Draft U.S. Atlantic and Gulf of Mexico Marine Mammal Stock Assessments 2021, pp 22-48. Available from https://media.fisheries.noaa.gov/2021-10/Draft%202021%20NE%26SE%20SARs.pdf

Pace III RM, Corkeron PJ, Kraus SD (2017) State–space mark–recapture estimates reveal a recent decline in abundance of North Atlantic right whales. Ecology and Evolution 7:8730-8741

Pace III RM, Williams R, Kraus SD, Knowlton AR, Pettis HM (2021) Cryptic mortality of North Atlantic right whales. Conservation Science and Practice 2021:e346

Runge MC, Linden DW, Hostetler JA, Borggaard DL, Garrison LP, Knowlton AR, Lesage V, Williams R, Pace III RM (2022) A management-focused population viability analysis for North Atlantic right whales, draft 08-2022.

Appendix 1: Bibliography of materials provided for review

Linden DW, Hostetler JA, Pace III RM, Garrison LP, Knowlton AR, Lesage V, Runge MC, Williams R (2022) Multistate capture-recapture models to estimate mortality and reproduction in North Atlantic right whales (1990-2019)

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

Pace III RM, Corkeron PJ, Kraus SD (2017) State–space mark–recapture estimates reveal a recent decline in abundance of North Atlantic right whales. Ecology and Evolution 7:8730-8741

Pace III RM, Williams R, Kraus SD, Knowlton AR, Pettis HM (2021) Cryptic mortality of North Atlantic right whales. Conservation Science and Practice 2021:e346

Runge MC, Linden DW, Hostetler JA, Borggaard DL, Garrison LP, Knowlton AR, Lesage V, Williams R, Pace III RM (2022) A management-focused population viability analysis for North Atlantic right whales, draft 08-2022.

Appendix 2: CIE performance work statement

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;

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:

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

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.

.

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.

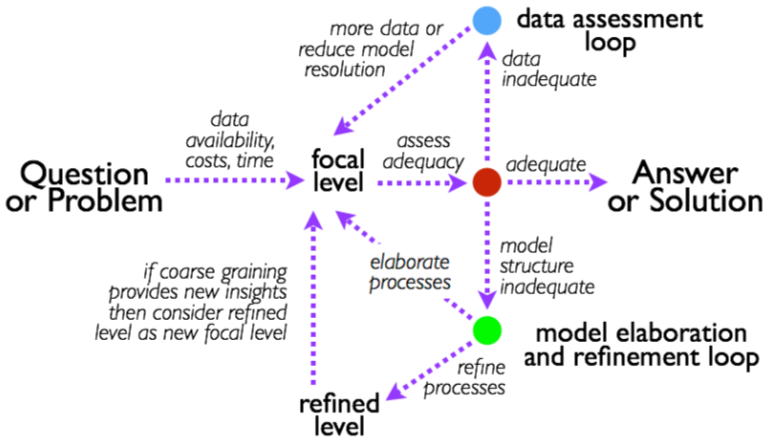

Figure from Getz et al, 2018. *Assessing model adequacy. The process of assessing model adequacy to provide an answer to a question or a solution to a problem begins at the focal level at which the model was initially formulated. Adequacy is assessed through evaluation of the relative benefits of new 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.*

Issue 4. Code verification and model validation and performance evaluation: 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.

   Pace et al. 2017: 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.

   Pace et al. 2020: 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.

   MMSA 2022: 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.*

   Linden et al. 2022: 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.

Independent Peer Review of the Population Viability Analysis for North Atlantic Right Whales

Center for Independent Experts (CIE)
External Independent Peer Review

by

Andrew J. Read

Gloucester, North Carolina

October 2022

1. Executive Summary

The North Atlantic Right Whale (NARW) is one of the world's most endangered marine mammals. Despite protection under relevant legislation in both the United States and Canada, and intensive conservation efforts in both countries, the population has been declining since 2010. In 2018 the National Marine Fisheries Service (NMFS) established the Population Evaluation Tool Subgroup to develop a population viability analysis (PVA) for the NARW. The goal of the Subgroup was to develop a predictive tool that would allow NMFS and its partners to characterize the future population trajectory and extinction risk for NARWs under a variety of scenarios. Development of the PVA is described in "*A Management-focused Population Viability Analysis for North Atlantic Right Whales.*"

The present report represents my independent review of the PVA developed by the Population Evaluation Tool Subgroup. It is my overall assessment that the PVA will, indeed, meet the demand for a predictive demographic tool for NMFS and its partners. And, as intended, the PVA will allow: prediction of the future status of NARWs under a variety of conditions; evaluation of the effects of individual anthropogenic threats on the demography of this population; and exploration of the consequences of various management interventions on their recovery. I believe the PVA will prove to be an extremely valuable addition to the conservation toolkit for this species.

It is also my assessment that the draft report considers the best available scientific information for NARWs, including the estimation of cryptic mortality by Pace et al. (2021). The report considers four major issues facing this population, including entanglement, vessel strikes, changes in prey resources, and anthropogenic noise. The two primary anthropogenic threats are entanglement in fixed fishing gear and vessel strikes, including both direct mortality and sub-lethal effects on mortality and reproduction. The PVA deals well with both of these anthropogenic threats and with the effects of a decrease in prey availability. The PVA deals less effectively with the potential effects of noise because we do not yet have direct linkages between exposure to noise and any demographic parameter for NARWs.

The goals laid out in the draft report are appropriate and reflect the clear need for a quantitative assessment tool for NARWs. The five specific objectives reflect typical goals for any Population Viability Analysis (PVA) of an endangered species.

Overall, the general structure of the core age and stage-structured model in PVA is reasonable and reflects the objectives of the approach, the desired model outputs needed to inform management, the current state of knowledge of NARWs, and the anthropogenic threats they face. Certain aspects of right whale demography are simplified in the model, including the assumption that there is no age-based variation in fecundity for mature females. In addition, the model incorporates an unrealistically simple form of density dependence. These simplifications are appropriate, given the current state of the NARW population (well below carrying capacity) and our knowledge of right whale reproduction, but more realistic simulations could be employed in future iterations. Likewise, the structure of the sub-models used to deal with the effects of mortality, reproduction, entanglement, vessel strike, and prey availability, is generally appropriate. As noted below, I have concerns regarding the sub-model used to estimate the effects of anthropogenic noise.

Parameter estimates for the PVA were developed using historical data from individual right whale sighting histories, mortality records, and an index of prey abundance. We are fortunate to have rich sighting histories of individual NARWs that span decades, derived from photo-identification records of whales contributed by members of the North Atlantic Right Whale Consortium, and curated by the New England Aquarium. In addition, we have extensive, but incomplete, records of right whale mortalities documented by stranding networks in eastern Canada and the U.S. NARWs are stenophagous predators, and primarily consume copepods of the genus *Calanus*, so it is possible to construct an index of prey abundance for the Gulf of Maine, Georges Bank, and Gulf of St. Lawrence. Of course, an index of prey abundance may not accurately reflect the true *availability* of prey to a predator; this may be particularly true for right whales, which are ram filter feeders and require dense aggregations of prey to feed profitably (van der Hoop et al. 2019).

The PVA employs sightings data and information on NARW carcass recoveries from 1990 – 2019 and employs a baseline abundance estimate from 2019. Importantly, the baseline scenario in the model employs a subset of data from the more recent past to parameterize reproductive rates (2010-2019) and injury and mortality rates (2014-2019). I believe this is an appropriate decision, given the changes in these parameters we have observed over the past decade, all of which have worsened the status of the population. Nevertheless, it is important to note that this choice assumes that the current negative conditions, especially with respect to lower prey availability, will persist into the future.

In general, I find the scientific conclusions of the draft report to be sound and interpreted appropriately from the model outputs. The scenarios considered in the report are reasonable, although not exhaustive. The authors have considered the major sources of uncertainty and have included appropriate caveats, where warranted. The *status quo* baseline scenario indicates that the NARW population is likely to continue to decline over the next century, leading to a median reduction of just over half of current abundance and a relatively high (0.684) likelihood of quasi-extinction (fewer than 50 reproductive females) at the end of that period. This rather dire conclusion is, sadly, consistent with our current knowledge of the demography of NARWs.

I have a few suggestions for improvement of the model and recommendations for future work, which are laid out below. However, I believe that the science described in the report is of a very high standard and that the PVA will be extremely useful to managers, scientists, and other stakeholders interested in the future of the NARW. It is clear from even a cursory reading of the report that development of the PVA required an enormous amount of work.

As noted in the draft report, we do not have any clear evidence of a demographic effect of anthropogenic noise on the vital rates of right whales. Instead, therefore, the authors chose to simulate the potential effects of noise through a control parameter in the prey submodel. On Page 32 they recognize that "…*this is a coarse way to represent a reduction in prey accessibility caused by environmental noise limiting the ability of whales to locate and acquire food*…" I agree, but do not believe that this is the most likely potential pathway for the effects of noise on the demography of NARWs. Instead, it is more likely that noise could mask the acoustic signals used by mothers and calves, causing a disruption of the bond between females and their dependent young. However, we have no way to parameterize such an effect. In addition, the approach taken in the draft report assumes that there are no effects of anthropogenic noise other than those experienced on the feeding grounds. Given these limitations in our present knowledge, therefore, I **recommend** that the Noise Submodel be removed from the PVA.

The management regime for NARWs is extremely fluid, reflecting the extremely complex and dynamic policy environments in both the U.S. and Canada. This management landscape is continually changing as new initiatives are developed to address the threats of entanglement and vessel strike, together with the added complications of litigation, at least in the U.S. The authors of the draft report were faced with a complicated decision, therefore, regarding what exactly constitutes a baseline, especially as a series of major mitigation measures are, or soon may be, launched in U.S waters, including the proposed ship strike rule and new measures to reduce the entanglement risk as part of the Atlantic Large Whale Take Reduction Rule. To address this uncertainty in exactly what constitutes 'current conditions,' the authors propose three different baseline scenarios, including the *status quo* (Baseline 1), and scenarios in which injury rates due to entanglement will be reduced by 25% (Baseline 2) or 50% (Baseline 3). Given the enormous uncertainty in what exactly will happen with these initiatives, at least in the very near future, I **recommend** that the PVA include only the *status quo* scenario as a Baseline.

One of the objectives of the work of the PET Subgroup was to *Facilitate communication, outreach, and education with stakeholders and the public.* As the results of the PVA will be of considerable interest to a very broad audience, including fishermen, managers, and other stakeholders, I **recommend** that NMS develop an extended, non-technical summary of the work described in the report, including an explanation of how PVAs are typically used in conservation planning.

I concur with the authors that the PVA should be viewed as a "living tool," which can be adapted, updated, and improved as more information becomes available. The draft report includes several sensible recommendations regarding future model development. I **recommend** that the highest priority of such future work should be to incorporate knowledge of the health status of individual right whales into the model. As the authors note, the core individual-based model is well-suited to the addition of such information, which would likely improve our ability to predict the probabilities of reproduction and mortality at the level of the individual whale.

As a second priority, I support the authors' suggestion that the PVA should be linked to the Decision Support Tool (DST) developed by NMFS as part of the 2021 Atlantic Large Whale Take

Reduction Rule. In the draft report the authors take care to note that "the purposes of the model do *not* include direct links between specific management actions and long-term population dynamics" (Section 2.3.3) and that links between such interventions and changes in demographic rates are the purview of other tools. But, perhaps unsurprisingly, they stray into exploring these linkages in the report, for example in Section 8.3. Therefore, I **recommend** that NMFS support future work to link the DST and PVA in a way that allows managers and other stakeholders to explore the potential demographic consequences of specific management interventions.

I also **recommend** that the authors explore the incorporation of explicit spatial structure into the model. Many PVAs account for spatial structure, especially when there are clear linkages between habitat quality and demographic processes, or in cases where the magnitude of anthropogenic threats varies significantly across the landscape (or seascape). The latter is certainly true for NARWs although, as the authors note, including such spatial structure would require more detailed information on the distribution of whales and the risk of entanglement and vessel strike than currently exists. In addition, the location of a large proportion of the population is unknown for much of the year. Nevertheless, this is an area that might be explored in future iterations of this 'living tool.' As noted in Section 9.3 of the draft report, with additional data it might be possible to separate the risk of ship strike and entanglement in the U.S. and Canada, which would have obvious benefits to the management process.

Finally, given the increasing amount of anthropogenic noise being introduced into the environment of the NARW, together with our uncertainty about the effects of this stressor, I **recommend** that future research address the potential linkages between the exposure to anthropogenic noise and the survival and reproduction of NARWs, so that this factor can be included in future iterations of the PVA.

2. Background

North Atlantic Right Whales (NARWs) are one of the most endangered species of marine mammals. Despite decades of protection under the U.S. Marine Mammal Protection Act, U.S. Endangered Species Act and the Canadian Species at Risk Act, the population remains very small and is declining. The population was estimated to consist of 483 individuals in 2010 (Pace et al. 2017) but declined to about 336 individuals in 2020 (Pettis et al. 2022).

The two primary anthropogenic threats to NARWs are entanglement in fixed fishing gear (primarily pot, trap, and gillnet fisheries) and vessel strikes. Most observed mortality in this population is attributable to these two factors and, apart from neonatal mortality, death from natural causes is rare. In addition, there have been recent changes in the availability of prey (*Calanus* spp.) in the feeding range of this species. The effects of entanglement and ship strike include both direct mortality and sub-lethal effects; this is particularly true for entanglement. Most of the population has been entangled in gear at some point in their lives and a quarter of the population becomes entangled each year (Knowlton et al. 2012). The sub-lethal effects of entanglement include adverse effects on health, a reduction in overall body size, and decreased reproductive output.

As described in the draft report, there have been several previous demographic analyses of NARWs for the purposes of projecting future population trends and for retrospective analyses of pre-exploitation abundance. In 2018 the National Marine Fisheries Service established the Population Evaluation Tool Subgroup to develop a population viability analysis (PVA) for NARWs. The goal of the Population Evaluation Tool Subgroup was to develop a quantitative, predictive tool that would allow NMFS and its partners to characterize the extinction risk for NARWs under a variety of scenarios. The need for such a PVA was specifically identified by NMFS in its most recent five-year review of the status of NARWs. Subsequent development of the PVA is described in the draft report entitled "*A Management-focused Population Viability Analysis for North Atlantic Right Whales.*"

The terms of reference for my review of the PVA and the associated report are as follows:

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.*

3. Description of My Role in the Review Activities

I am a conservation scientist with expertise on the ecology and demography of marine mammal species and my review should be viewed through this lens. The other two reviewers have considerably more expertise in population modeling and population viability analysis; I view my role here to ensure that the PVA is grounded in the best available science on NARWs and that the approach taken is of the greatest possible value to the many stakeholders with an interest in the conservation and recovery of this population.

I was first approached by the CIE to determine my interest in participating in this review in June 2022. I responded in the affirmative and, on August 1st, I was informed that I had been selected as one of the reviewers. I participated in a webinar on August 26th with the other two reviewers and some of the authors of the report. The webinar was extremely helpful and clarified several questions that we had regarding the report. We had one subsequent e-mail exchange to clarify an additional question; since that time, I have had no contact with the other reviewers, so the present report represents my independent desk review of the PVA.

4. Summary of Findings for Each TOR

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.

As noted by the authors of the draft report, PVAs can be extremely useful in evaluating extinction risk and comparing the potential efficacies of various management interventions. They note many prior instances in which PVAs have been used as an effective conservation tool, including several examples with other marine mammals (Regehr et al. 2015; Runge et al. 2017). It is my assessment that the PVA described in the draft report will meet the objectives of the PET Subgroup and allow NMFS and its partners to predict the future status of NARWs under a variety of conditions; evaluate the effects of individual anthropogenic threats on the demography of this population; and explore the consequences of various management interventions on their recovery. Overall, I believe the PVA will prove to be an extremely valuable addition to the conservation toolkit for this species. The goals of the work laid out in the draft report are appropriate and reflect the clear need for a quantitative assessment tool for NARWs. The five specific objectives reflect typical goals for a PVA of any endangered species.

The desired outputs of the model are reasonable and consistent with the recovery criteria laid out in the U.S. Recovery Plan of 2005 and the Canadian Species at Risk Act (SARA) Recovery Strategy of 2014. The desired demographic metrics for these two management strategies include future abundance, population growth rate, and the likelihood of quasi-extinction. As noted in the draft report, these metrics are all straightforward outputs from the PVA.

Overall, the structure of the model is reasonable and reflects the objectives of the approach, the desired model outputs required to inform management, and the current state of knowledge of NARWs and of the anthropogenic threats they face. The four threats incorporated into the model, entanglement, vessel strikes, changes in prey resources, and anthropogenic noise, are the major known issues facing this population. The two primary anthropogenic threats are entanglement in fixed fishing gear and vessel strikes, which include both effects on direct mortality and indirect effects on mortality and reproduction.

The ‘baseline and scenarios’ approach described in the draft report is appropriate and will allow managers and other stakeholders to predict the population trajectory of NARWs under baseline conditions and explore the potential effects of various management interventions.

The entanglement and vessel strike scenarios are relatively straightforward and involve varying the incidence and severity of such interactions. However, the approach taken with anthropogenic noise (the *Noise Submodel*) scenario strikes me as slightly odd. As noted in Section 3.5.5 of the draft report, modeling the effects of anthropogenic noise requires evidence that this stressor affects vital rates (survival and/or reproduction). Such an effect could be mediated, for example, if shipping noise interfered with feeding or increased the likelihood of separation of mothers and calves, especially given the very quiet calls made by these pairs (Parks et al. 2019). But, as noted in the report, we do not yet have clear evidence of any demographic effect of anthropogenic noise on right whale demography. Instead, therefore, the authors chose to simulate the potential effects of noise through a control parameter in the prey submodel. I do not see the advantage in retaining this submodel in the PVA given how little we understand about the effects of anthropogenic noise on the vital rates of NARWs and recommend that it be eliminated.

The combined age- and stage-structured approach of the base model makes sense. A fully age-structured model would be preferable, especially if there was evidence of age-related variation in reproductive rates in NARWs, such as a decline in fecundity with advancing age. As far as I am aware, there is no clear published evidence for such age-specific variation in reproductive output in female NARWs (although see Hamilton et al. 1998). We do not know the expected longevity of NARWs but based on the extensive sighting histories of a few animals, including a female that was observed in 1935 with a calf and later resighted in 1995, these are clearly very long-lived animals (Hamilton et al. 1998). Populations with such extended longevity are likely to experience some diminution in fecundity with age, so it would be useful to incorporate age-specific variation in the probability of a female weaning a calf when such information becomes available. Such variation could also influence the reproductive output of younger females. Are first-time mothers, for example, less likely to successfully wean an offspring than older, more experienced females?

The model structure must deal with an awkward mismatch between the reproductive seasonality of NARWs, in which most calves are born during winter, and a census date of July 1st. For example, neonatal mortality (one of the few apparent sources of natural mortality in this population) is incorporated into the early calf-loss rate (κ), and the first survival rate applied to calves is the survival from age 0.5 to age 1.5 ($s1$). It seems that “…calf survival does not depend on survival of its mother…” (P. 25), but what happens to a calf that loses its mother within the first six months of its life?

Given the very low current abundance of NARWs, relative to reasonable assumptions of original population size or carrying capacity, it seems very unlikely that density dependent factors will influence either reproduction or survival within the time frame explored by the PVA. The model considers the future trajectory of the NARW population over the period of a century, which spans only three or four generations for NARWs. Thus, although biologically implausible, the 'ceiling approach' to density dependence taken by the authors is appropriate in this instance.

It is my assessment that parameterization of the PVA employs the best available science for NARWs, including the important estimation of cryptic mortality by Pace et al. (2021). Good PVAs are data hungry models and this one is no exception. Fortunately, in comparison to most populations of marine mammals, NARWs are relatively well-studied, so we have a rich dataset on the histories of individual whales that serve as the foundation of the core model. Nevertheless, there are still limitations to our understanding of the survival and reproduction of NARWs and, particularly, to the sub-lethal effects of entanglement and vessel collisions. And, of course, there are still many important gaps in our knowledge of the distribution of fishing effort and vessel traffic that limit our ability to fully describe the effects of these threats. This limitation is more acute for the effects of prey limitation and even more so for anthropogenic noise, the effects of which are not captured effectively in the PVA.

It is important to note that the effects of prey availability are linked to an index of the abundance of *Calanus* copepods in the known feeding range of the species in the Gulf of Maine, Georges Bank, and Gulf of St. Lawrence. The feeding range of a significant proportion of the population is unknown, so it is unclear whether this index also reflects the abundance of Calanus populations in these unknown feeding areas. And, of course, an *index* of prey abundance may not accurately reflect the *availability* of prey to a predator; this may be particularly true for right whales, which are ram filter feeders and require dense aggregations of prey to feed profitably (van der Hoop et al. 2019). Nevertheless, the authors of the draft report have done a good job of capturing current knowledge of the prey base of NARWs for the purposes of the PVA.

Many PVAs account for explicit spatial structure, especially when there are clear linkages between habitat quality and demographic processes, or in cases where the magnitude of anthropogenic threats varies significantly across the landscape (or seascape). The latter is certainly true for NARWs although, as the authors note, including such spatial structure would require more detailed information on the distribution of whales and the risk of entanglement and vessel strike than currently exists. This is an area that might be explored in future iterations of this 'living tool.' As noted in Section 9.3, for example, with additional data it might be possible to separate the risk of ship strike and entanglement in the U.S. and Canada, which would have obvious benefits to the overall management process.

The authors take care to note that "the purposes of the model do *not* include direct links between specific management actions and long-term population dynamics" (Section 2.3.3) and that links between such interventions and changes in demographic rates are the purview of other tools. Nevertheless, I agree with their later conclusion (Section 9.2) that an integration of the PVA and these other approaches, such as the Decision Support Tool, is desirable and should be explored in the future. Such integration would allow managers and other stakeholders to explore the population-level effects of various management actions in a way that would allow direct evaluation of the potential benefits in terms of recovery criteria.

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.

The core model employs sightings data and information on NARW carcass recoveries obtained over a three-decade period from 1990 – 2019 and employs a baseline abundance estimate from 2019. Importantly, the baseline scenario in the model employs a subset of data from the recent past to parameterize reproductive rates (2010-2019) and injury and mortality rates (2014-2019). This is an important distinction, and one that deserves greater emphasis in the report. I believe the use of more recent data to estimation reproduction, injury, and mortality, is reasonable, given the changes in these parameters we have observed over the past decade, all of which have worsened the status of the population. If data from the entire period (1990-2019) were used, the baseline scenario would be overly optimistic. Nevertheless, it is important to note that this choice of baseline conditions assumes that the current rather negative conditions experienced by NARWs will persist into the future. It seems to me that this is likely to be true for prey availability and for injury and mortality rates from entanglement and vessel strikes.

I support use of the estimate of abundance for 2019 as a starting value for model projections, even though an estimate was available for 2020. As noted by Pace et al. (2017) the most recent estimate of abundance is likely to be negatively biased because some observations of living whales have not yet been reported.

It is important to note that our knowledge of the causes of mortality of NARWs, even during the 2010-2019 period, are limited. Our knowledge regarding cause of death in NARWs is based on painstaking examination of right whale carcasses by skilled veterinarians, pathologists, and anatomists. Remarkably, in 33 cases for which cause of death could be ascertained in adult or juvenile NARWs from 2003 to 2018, not a single case of natural mortality was identified (Sharp et al. 2019). Five cases of natural mortality were documented in perinatal animals, but it seems that, once a calf survives the rather perilous period around its birth, it is likely to experience very low subsequent rates of natural mortality. A similar pattern was documented in post-mortem examinations conducted between 1970 and 2002 (Moore et al. 2004), although necropsy procedures were not as fully developed during this period. However, as the authors of the draft report note, most (64%) deaths believed to have occurred between 1990 and 2017 were not documented, and cause of death could not be determined even for some recovered carcasses (Pace et al. 2021). Thus, our estimates of the relative importance of entanglement and vessel strike could be biased. Without further direct information on the cause of mortality on a larger proportion of deaths, we cannot assess the possibility of such bias.

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.

In general, I find the scientific conclusions of the draft report to be sound and interpreted appropriately from the model outputs. The scenarios considered in the report are reasonable, although not exhaustive. The authors have considered the major sources of uncertainty and, in general, have included appropriate caveats, where warranted. The primary baseline scenario indicates that the NARW population is likely to continue to decline steadily over the next century, leading to a median reduction of just over half of current abundance and a relatively high (0.684) likelihood of quasi-extinction (fewer than 50 reproductive females) at the end of that period. This rather dire conclusion is, sadly, consistent with our current knowledge of the demography of NARWs.

I do not support the inclusion of three baseline scenarios in the draft report. I recognize that the management regime for NARWs is extremely fluid, particularly at the present time, reflecting the extremely complex policy environment in both the U.S. and Canada. Thus, the authors were faced with a complicated decision regarding what constitutes a baseline, especially as a series of major mitigation measures may (or may not) be launched soon in U.S waters. To address this uncertainty, the authors proposed three different baseline scenarios, including the *status quo* (Baseline 1), and scenarios in which injury rates due to entanglement will be reduced by 25% (Baseline 2) or 50% (Baseline 3). I find this approach to be confusing and unnecessary.

I would also like to see more discussion of the assumptions regarding future conditions, particularly as it relates to patterns of fishing effort. For example, the authors take pains to acknowledge the difficulty of predicting future patterns of future vessel traffic in NARW habitat. On Page 45 they note "*The baseline scenario assumes that current vessel speed regulations will remain constant, as will the overall vessel traffic, thus, the vessel strike injury rate will remain constant over the period of projection.*" The same must certainly be true of future patterns of *effort* in fixed gear fisheries in the U.S. and Canada, especially under different scenarios of climate change. We know that patterns of fishing effort will change, but the PVA assumes that fishing effort will remain constant. And, as noted by the authors, further uncertainty exists regarding future industrial development in offshore waters, such as that associated with renewable marine energy installations. We do not yet understand what effects, if any, such development will have on this endangered population.

Finally, I would like to see a stronger explanation of the limitations of the PVA approach, including an expansion on some of the caveats laid out in Section 8.4. A clear exposition of these limitations would be especially important in any non-technical summary of this work produced for stakeholders outside the scientific community.

5. Conclusions

It is my assessment that the PVA described in "*A Management-focused Population Viability Analysis for North Atlantic Right Whales*" is an appropriate approach that meets the needs of NMFS and its partners responsible for recovery of the North Atlantic Right Whale. Furthermore, I believe that the report considers the best available scientific information, including data on the demography of right whales and the four major issues facing this population: entanglement; vessel strikes; changes in prey resources; and anthropogenic noise. The use of a subset of recent data to parameterize reproductive rates, injury and mortality rates is appropriate, given the changes in these parameters observed over the past decade, all of which have worsened the status of the population. Finally, the scientific conclusions contained in the report are sound and interpreted appropriately. The major sources of uncertainty have been considered and, in general, appropriate caveats have been described. The primary baseline scenario indicates that the NARW population is likely to continue to decline steadily over the next century, a conclusion that is consistent with our current knowledge of NARWs. Overall, I believe that the science described in the report is of a high standard and that the PVA will be extremely very useful to managers, scientists, and other stakeholders interested in the future of the NARW.

6. Recommendations

As noted in the draft report, we do not have any clear evidence of a demographic effect of anthropogenic noise on the vital rates of right whales. Instead, therefore, the authors chose to simulate the potential effects of noise through a control parameter in the prey submodel. On Page 32 they recognize that "…*this is a coarse way to represent a reduction in prey accessibility caused by environmental noise limiting the ability of whales to locate and acquire food*…" I agree, but do not believe that this is the most likely potential pathway for the effects of noise on the demography of NARWs. Instead, it is more likely that noise could mask the acoustic signals used by mothers and calves, causing a disruption of the bond between females and their dependent young. However, we have no way to parameterize such an effect. In addition, the approach taken in the draft report assumes that there are no effects of anthropogenic noise other than those experienced on the feeding grounds. Given these limitations in our present knowledge, therefore, I **recommend** that the Noise Submodel be removed from the PVA.

The management regime for NARWs is extremely fluid, reflecting the extremely complex policy environment in both the U.S. and Canada. This management landscape is continually changing as new initiatives are developed to address the threats of entanglement and vessel strike, together with the added complications of litigation, at least in the U.S. The authors of the draft report were faced with a complicated decision, therefore, regarding what exactly constitutes a baseline scenario, especially as a series of major mitigation measures are, or soon may be, launched in U.S waters, including the proposed ship strike rule and new measures to reduce the entanglement risk as part of the Atlantic Large Whale Take Reduction Rule. To address this uncertainty in exactly what constitutes 'current conditions,' the authors quite reasonably propose three different baseline scenarios, including the *status quo* (Baseline 1), and scenarios in which injury rates due to entanglement will be reduced by 25% (Baseline 2) or 50% (Baseline 3). Given the enormous uncertainty in what exactly will happen with these initiatives, at least in the very near future, I **recommend** that the PVA include only the status quo scenario as a Baseline.

One of the objectives of the work of the PET Subgroup was to *Facilitate communication, outreach, and education with stakeholders and the public.* As the results of the PVA will be of considerable interest to a very broad audience, including fishermen, managers, and other stakeholders, I **recommend** that NMFS develop an extended, non-technical summary of the work described in the report, including an explanation of how PVAs are used in conservation planning.

I concur with the authors that the PVA should be viewed as a "living tool," which can be adapted, updated, and improved as more information becomes available. The draft report includes several good recommendations regarding future model development. I **recommend** that the highest priority of future work should be to incorporate knowledge of the health status of individual right whales into the model. As the authors note, the core individual-based model is well-suited to the addition of such information, which would likely improve our ability to predict the probabilities of reproduction and mortality and the level of the individual.

As a second priority, I support the authors suggestion that the PVA should be linked to the Decision Support Tool (DST) developed by NMFS as part of the 2021 Atlantic Large Whale Take Reduction Rule. In the draft report the authors take care to note that "the purposes of the model do *not* include direct links between specific management actions and long-term population dynamics" (Section 2.3.3) and that links between such interventions and changes in demographic rates are the purview of other tools. But it seems to me that they stray into exploring these linkages in Section 8.3, for example. Therefore, I **recommend** that NMFS explore the potential to link the DST and PVA in a way that allows managers and other stakeholders to explore the potential demographic consequences of specific management interventions.

I also **recommend** that the authors explore the incorporation of explicit spatial structure into the model. Many PVAs account for spatial structure, especially when there are clear linkages between habitat quality and demographic processes, or in cases where the magnitude of anthropogenic threats varies significantly across the landscape (or seascape). The latter is certainly true for NARWs although, as the authors note, including such spatial structure would require more detailed information on the distribution of whales and the risk of entanglement and vessel strike than currently exists. In addition, the location of a large proportion of the population is unknown for much of the year. Nevertheless, this is an area that might be explored in future iterations of this 'living tool.' As noted in Section 9.3 of the draft report, with additional data it might be possible to separate the risk of ship strike and entanglement in the U.S. and Canada, which would have obvious benefits to the management process.

Finally, given the increasing amount of anthropogenic noise being introduced into the environment of the NARW, together with our uncertainty about the effects of this stressor, I **recommend** that future research address the potential linkages between the exposure to anthropogenic noise and the survival and reproduction of NARWs, so that this factor can be included in future iterations of the PVA.

7. References

Hamilton *et al.* 1998. Age structure and longevity in North Atlantic right whales *Eubalaena glacialis* and their relation to reproduction. *Marine Ecology Progress Series* 171: 285-292.

Knowlton *et al*. 2012. Monitoring North Atlantic right whale *Eubalaena glacialis* entanglement rates: a 30-year retrospective. *Marine Ecology Progress Series* 466: 293-302.

Moore *et al*. 2004. Morphometry, gross morphology and available histopathology in North Atlantic right whale (*Eubalaena glacialis*) mortalities (1970-2002). *Journal of Cetacean Research and Management* 6: 199-214.

Pace *et al*. 2017. State–space mark–recapture estimates reveal a recent decline in abundance of North Atlantic right whales. *Ecology and Evolution* 7: 8730-8741.

Pace *et al*. 2021. Cryptic mortality of North Atlantic right whales. *Conservation Science and Practice* 3:e346.

Parks *et al.* 2019. Acoustic crypsis in communication by North Atlantic right whale mother–calf pairs on the calving grounds. *Biology Letters* 15: 20190485.

Pettis *et al*. 2022. North Atlantic Right Whale Consortium 2021 Annual Report Card. *Report to the North Atlantic Right Whale Consortium*.

Regehr *et al*. 2015. Resilience and risk: a demographic model to inform conservation planning for polar bears. *U.S. Geological Survey Open-File Report* 2015-1029.

Runge *et al*. 2017. Status and threats analysis for the Florida manatee (*Trichechus manatus latirostri*s), 2016. *Scientific Investigations Report* 2017-5030.

Sharp *et al.* 2019. Gross and histopathologic diagnoses from North Atlantic right whale *Eubalaena glacialis* mortalities between 2003 and 2018. *Diseases of Aquatic Organisms* 135: 1-31.

van der Hoop *et al*. 2018. Foraging rates of ram-filtering North Atlantic right whales. Functional Ecology 33: 1290-1306.

Appendix 1

Bibliography of Materials Provide for Review

Moore *et al*. 2004. Morphometry, gross morphology and available histopathology in North Atlantic right whale (*Eubalaena glacialis*) mortalities (1970-2002). *Journal of Cetacean Research and Management* 6: 199-214.

Hayes *et al*. 2022. North Atlantic Right Whale (*Eubalaena glacialis*): Western Atlantic Stock. In: U.S. Atlantic and Gulf of Mexico Marine Mammal Stock Assessments 2021. *NOAA Technical Memorandum NMFS-NE 288*, Northeast Fisheries Science Center, Woods Hole, MA.

Pace *et al*. 2017. State–space mark–recapture estimates reveal a recent decline in abundance of North Atlantic right whales. *Ecology and Evolution* 7: 8730-8741.

Pace *et al*. 2021. Cryptic mortality of North Atlantic right whales. *Conservation Science and Practice* 3:e346.

Runge *et al*. 2022. A Management-focused Population Viability Analysis for North Atlantic Right Whales. Unpublished document.

Sharp *et al.* 2019. Gross and histopathologic diagnoses from North Atlantic right whale *Eubalaena glacialis* mortalities between 2003 and 2018. *Diseases of Aquatic Organisms* 135: 1-31.

Appendix 2

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

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:

Lesage, Fisheries and Oceans Canada; Dr. Daniel Linden, NOAA Fisheries, Greater Atlantic Regional Fisheries Office; Dr. Rob Williams, ORCA

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.