Summary Report of the

North Atlantic Right Whale Decision Support Tool Peer Review

January 30 – February 1, 2023

Report prepared by Panel Members:
Richard Merrick (Chair)
Dewayne Fox
Rebecca Lewison
Genny Nesslage
Erin Summers
Tim Tinker

Introduction

The Atlantic Large Whale Take Reduction Team (ALWTRT) asked NOAA in 2018-19 for better decision-making tools to help them understand how much entanglement risk to North Atlantic right whales (NARW) could be reduced under different management scenarios. In response, scientists at NOAA's Northeast Fisheries Science Center working with others developed the Decision Support Tool (DST). Phase 1 of the DST was peer reviewed in 2019[1] by the Center for Independent Experts (CIE) and subsequently used to evaluate proposed modifications to the U.S. commercial American lobster and Jonah crab trap/pot fisheries in New England. The DST has since been updated and modified in order to evaluate the ALWTRT's proposed Phase 2 and 3 modifications to the Atlantic Large Whale Take Reduction Plan (ALWTRP), including possible changes in the U.S. gillnet and other trap/pot commercial fisheries from Maine to Florida, as well as additional consideration of lobster and Jonah crab trap fisheries.

We report here on the peer review of the updated version of the DST. The primary changes to the DST compared to the Phase 1 version of the model reviewed by the Center for Independent Experts (CIE) in 2019 include:

- Expansion of the spatial scale of the model to include U.S coastal waters from Florida to the Scotian Shelf,
- Inclusion of all state and federal U.S. fixed gear fisheries under the jurisdiction of the Plan (including other trap/pot and all gillnet fisheries),
- Simulation of threats from this broader range of gear,
- Updates to the Duke whale habitat density model focusing on NARW sightings from 2010 through ~2020, as well as addition of data on fin, minke, and humpback whale

[1]https://www.fisheries.noaa.gov/feature-story/decision-support-tool-helpful-those-finding-ways-reduce-whale-entanglement-fishing

distributions,
- Consideration of vertical distribution of whales based on behavior (transiting, feeding, etc.),
- Modeling of the risk reduction resulting from additional potential fishery management actions (line caps, soak limits, etc.), and
- The addition of new methodologies for spatially allocating fishing using other data sources, such as depth.

The Peer Review Panel met via Google Meet on January30-February 1, 2023 (see agenda in Appendix 1). The review was organized cooperatively between the Atlantic Scientific Review Group (ASRG) and NOAA's Northeast Fisheries Science Center/Greater Atlantic Region Fisheries Office. The Panel was composed of six scientists including Richard Merrick (Chair, NOAA retired), Dewayne Fox (Delaware State University), Rebecca Lewison (San Diego State University), Genny Nesslage (University of Maryland Center for Environmental Science), Erin Summers (Maine Department of Marine Resources), and Martin Tim Tinker (University of California Santa Cruz, USGS retired).

The meeting opened at noon on Monday December 12, 2023 with welcoming remarks, a review of the five Terms of Reference (TOR, Appendix 2) for the review, and a summary of events leading up to this review by Michael Asaro (NOAA) and Panel chair Richard Merrick. The remainder of the day focused on presentations and discussion of the Fishery Data used within the DST. Day 2 was devoted to presentations and discussion on the Whale Habitat Density model and the core modeling within the DST. The morning of Day 3 concluded with a two-hour closed session among the Panelists to discuss their core strategic recommendations, followed by an hour open session with NOAA and the public to present these recommendations. The Panel Chair compiled and edited this Report with assistance (by correspondence) from the Panelists, before submission of the Report to NOAA. Documentation of the DST and a copy of the DST presentations can be downloaded from the Peer Review website:

https://www.fisheries.noaa.gov/event/peer-review-right-whale-decision-support-tool-0

The Panel considered whether the current version of the DST meets the Terms of Reference. While they agreed that the TORs are generally met, they believe explicit analysis and presentation of model sensitivity and uncertainty in the DST would provide a better tool for informing management of the risks of various approaches to reducing the Serious Injury and Mortality (SI/M) of large whales resulting from interactions with commercial fishing gear. The Panel's experience is that best practice standards suggest managers should first define the level of uncertainty they are willing to accept when attempting to meet the risk reduction goal (e.g., reducing takes to the stocks Potential Biological Removal level or PBR), and that an evaluation tool (like the DST) should then quantify whether an option (or combination of options) reaches the risk reduction goal with an acceptable level of uncertainty.

Evaluation of the Terms of Reference

TOR 1: Incorporates the U.S. lobster, gillnet, and other trap/pot[2] commercial fishery dependent (i.e., catch) data sets necessary to determine fishing effort and distribution for the Phase 2 and 3 DST

The Panel agreed that this term of reference had been met. NOAA staff should be commended for the work they have performed gathering and synthesizing the disparate fishery dependent data for those fixed gear fisheries interacting with large whales. However, the Panel was very concerned with how sensitive the model estimates of risk were to the spatial scale and resolution selected for the spatial allocation of gear from a fishery..

An example given (slide 107 of the DST presentation[3]) showed that co-occurrence and risk scores for one month in one fishery were halved by changing the allocation from very coarse Statistical/Subareas ("GeoAreas") to fine scale Vessel Monitoring System (VMS) coordinates. Given that four different approaches to spatial scaling are used for different fisheries, this could significantly bias scores in any comparison between fisheries and would make it more difficult to "roll up" risk reduction across fisheries.

Recommendations:

1. Evaluate DST model performance at different spatiotemporal scales to determine whether the bias resulting from data reporting resolution requires fisheries with low spatial resolution to adopt disproportionately large measures of risk reduction compared to fisheries with higher resolution reporting.
2. Based on the evaluation, consider, in consultation with NOAA fishery managers, whether a uniform approach to the spatial (and perhaps temporal) distribution of gear in the model would be appropriate.

TOR 2: Incorporates spatially and temporally relevant estimates of right whale distribution and densities.

The Panel was impressed by the significant amount of research and data that had gone into the whale habitat density modeling supporting the DST, and as such agreed that this TOR had generally been met. Still, the Panel discussed how using other approaches in the DST for estimating the likelihood of whale presence in an area might provide additional, useful information on the risk of a large whale interacting with fishing gear.

Recommendations:

1. Evaluate whale model performance through multi-model comparisons across varying levels of model complexity. We strongly encourage a comparison of the whale density

[2] Including at least the Mid-Atlantic mixed species trap/pot, Northeast/Mid-Atlantic American lobster trap/pot, Jonah crab trap/pot fisheries, and Mid-Atlantic/Northeast sink gillnet fisheries

[3] https://media.fisheries.noaa.gov/2023-01/DST-Model-Peer-Review-Documentation_Jan2023-nefsc.pdf

model with alternative models of whale distribution that do not focus on density (e.g., occurrence, dSDMs).
2. Further test the sensitivity of the risk calculations to the choice of a whale model. Less complex models may facilitate data type integration (either through an ensemble or INLA methods) across spatial areas.
3. As with the allocation of fishery effort/gear, consider whether the spatiotemporal scales used for whale presence are appropriate for estimating risk. Consider whether the fishery and whale datasets should be aligned using the same scales. Also, consider whether larger scales may more accurately capture the area a whale occupies given that a sighting location represents only a snapshot of a whale's location and does not consider that the whale is often moving from the location.

TOR 3: Uses appropriate entanglement risk coefficients by gear type.

The Panel agreed that TOR is generally met, though there are some additional analyses that could improve the quality of advice. This is an area where sensitivity analysis could provide important insight into the properties of the DST. While it is clear the model team has been conducting informal sensitivity analyses all along, the results of this will not be clear to users of the model. The panel's suggestion is thus to formalize the sensitivity analyses and standardize the presentation of results for users/decision makers.

<u>Recommendations</u>:

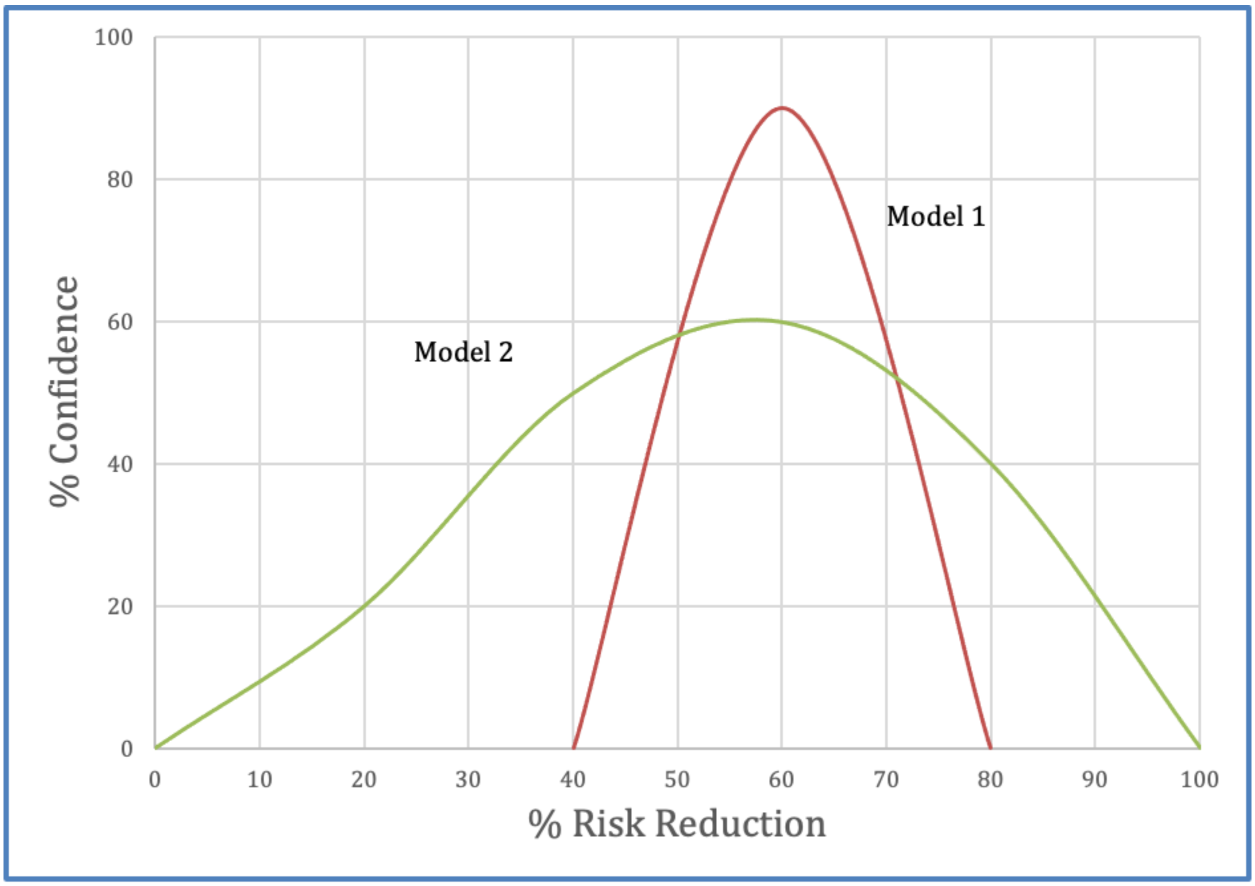
1. Consider the use of simulated data sets (i.e., simulate data where you know the true values/distributions, then add realistic levels of stochasticity) to measure model sensitivity to different parameters, assumptions, and data sources.
2. Consider focusing analysis of sensitivity of model results to the coefficients of threats. For example, much of the success of the ALWTRP rests on the assumption that rope of less than 1,700 lbs breaking strength is relatively safe for large whales. What would be the outcome of the analysis if this assumption was <u>incorrect</u>, and breaking strengths need to be reduced to much lower levels?
3. Consider the utility of simpler approaches to risk analysis for various gear types. For example, attempting to model whale shape to understand gill net interactions introduces additional uncertainty to the threat analysis, while a simpler approach (e.g., panel height versus fishing depth) may be just as useful and not prone to additional uncertainties.

TOR 4: Incorporates appropriate approaches to applying these coefficients to estimate relative risk (and risk uncertainty) of large whales encountering gear.

Again, the Panel agreed that the current version of the DST model does generally meet the TOR element of estimating relative risk, though there is insufficient attention to estimating risk uncertainty. A key benefit of the quantification of uncertainty is the ability to balance average expected risk reduction with the relative degree of confidence in that value. All estimates of risk reduction should be accompanied by an uncertainty measure, and managers should consider setting some boundaries on the acceptable level of uncertainty (e.g., X % of measures must have at least Y % probability of exceeding some confidence threshold).

Recommendations:

1. The DST needs to quantify scientific uncertainty in various data types and integrate all uncertainty to provide estimates of scientific uncertainty in risk reductions. Some possible approaches include:
    a. For whale Surface Density Maps (SDMs: run model with 100 different bootstrap maps and measure variation in risk, and
    b. For fisheries data: boot-strap samples across years for those portions of fisheries where there are sufficient data to estimate magnitude of variance in gear distributions.
2. We encourage the adoption of a graphical representation of the factors that strongly influence scientific uncertainty for communication to managers and stakeholders. This would allow the depiction in a simple case (as in the following figure), to show how managers decided to select a gear option (Model 1) based on an *a priori* decision to require at least 75% confidence in the model predictions of a 60% risk reduction.



TOR 5: Using the fishing effort and right whale density data reviewed under TORs 1-2, provides appropriate estimates of risk reduction of large whales encountering gear

The Panel understands the sequence of events that led to the current DST approach to risk estimation using arbitrary dimensionless units. However, the Panel strongly suggests it would be more appropriate to estimate risk in demographic units that are aligned with the MMPA's PBR approach to risk reductions. Moreover, it would also seem appropriate to introduce estimates of not only scientific risk/uncertainty but also management risk/uncertainty (though the latter may be more difficult to quantify).

<u>Recommendations:</u>

1. Consider reformulating risk units to instantaneous hazards, to create demographically meaningful measures of risk (and risk reduction) that can be more seamlessly integrated in PVA models and PBR evaluations. The current "risk units" metric is somewhat ad-hoc, difficult to interpret, and it is unclear how adding units across months/fisheries/areas affects the validity of the overall risk estimates (i.e. because they are not demographically explicit). Switching to a cumulative/competing hazards approach could be relatively straightforward and would make presentation and interpretation of results simpler. (See example in Appendix 3).
2. Consider whether it is appropriate to estimate risks at the level of 1 nm square. Although the Panel understands that it may at times be necessary to use a finer resolution to deal with coastline issues, the considerable uncertainty in both gear and whale locations suggests that this level of analysis introduces an unnecessary level of uncertainty into the risk calculations. If such fine scale analysis is required to accommodate coastline complexity in nearshore areas, consider a hierarchical approach (e.g., 1 nm squares in the near shore, and 10 nm squares beyond 3 nm).
3. Consider estimating management uncertainty and incorporating that uncertainty in estimates of risk. The Panel's understanding of the DST model analyses is that the risk estimates produced assume the fishery modifications will be fully implemented. Like all fishery actions, not all recommended actions will be adopted and implemented perfectly, thus producing what could be considerable management uncertainty. The DST could, if run as a tool for Management Strategy Evaluations, be used to evaluate how management uncertainty could impact effectiveness of the ALWTRPs plan to reduce takes to PBR.

Closing Comments

The Panel recommends strategic modifications to the DST to improve the quality of the DST and its utility for proactive management. In particular, the Panel urges NOAA to conduct sensitivity analyses of model parameters and quantify scientific uncertainties in the model results. These recommendations align with comments from the CIE reviewers in 2019 and the ASRG in 2020.

The Panel has an overarching final concern that while the DST does provide estimates of risk, the structure of the DST model is a set of independently developed submodels that do not (in many cases) align well spatiotemporally. This disjointed structure may be the source of unspecified levels of uncertainty with unknown sensitivity to model parameters. Given the need for a risk assessment tool to address an immediate management need, the current DST is a good first step. However, NOAA and the DST team should consider stepping back and reviewing whether the DST model structure is appropriate, and then consider redeveloping the modeling framework for long-term use in assessing NARW and possibly other large whale/fisheries interactions (e.g., humpback whales) . This could include creating a spatial/seasonal whale population model that integrated all the sighting and mark-recapture data to estimate mortality and natality in a way that linked them to fishing effort and food availability and other relevant physical/environmental covariates. If such a model can fit the available historical time series well enough, NOAA could then more reliably forecast future population response to changes in management, carrying the uncertainty in parameter estimates into those projections. Recommendations we have provided under the various TORs (e.g., exploring alternative whale modeling, and estimating risk units as instantaneous hazards) can be used in this new approach to provide potentially more robust advice to the ALWTRT and NOAA managers.

The Panel believes that these recommendations should be addressed immediately and strongly suggests that <u>before</u> NOAA makes further modifications to the current version of the DST, that they consider and address our concerns and recommendations. Recent Federal legislative actions have slowed implementation of further ALWTRP modifications for the American lobster and Jonah crab fisheries, so it is reasonable to expect that NOAA has sufficient time to fully consider and explore these recommendations.

The Panel and the ASRG are very grateful to the DST Team for the work that went into their analysis and their patience with the Panel in conducting the peer review. Assistance by Michael Asaro and Burton Shank made the meeting possible, and we thank them for this!

Appendix 1 - Agenda for An Independent Peer Review of the Phases 2-3 Expanded Decision Support Tool created for the Atlantic Large Whale Take Reduction Team

January 30: Data

10:00 Introduction of Panel, Terms of Reference, Agenda Review – *Richard Merrick*
10:15 ALWTRT Background and Overview of the Purpose of the Decision Support Tool (DST) – *Michael Asaro*
10:45 Clarifying questions and discussion – *Panel*
11:00 Model Overview – *Burton Shank*

- *General background*
- *Highlights of what has changed*
- *Linkages to Terms of Reference*

11:30 Clarifying Questions and discussion – *Panel*

12:00 Lunch

1:00 Presentation on Gillnet and OTP gear and fisheries – *Alicia Miller, Laura Solinger*
1:30 Presentation on Lobster/Jonah crab gear and fisheries – *Burton Shank*
2:15 Clarifying questions and discussion – *Panel*

2:45 Break

3:00 Clarifying questions and discussion about data – *Panel*
4:00 Public questions and comments
4:15 Discussion of findings/recommendations on DST supporting data - *Panel*
5:00 Adjourn

January 31: Models

10:00. Summary of Day 1 – *Richard Merrick*
10:15. Presentation on the Duke NARW "Habitat" (distribution and density) model – *Jason Roberts*

- What is new in recent model versions
- How this is integrated into the Decision Support Tool

11:15 Clarifying questions and discussion – *Panel*
11:45 Public questions and comments

12:00 Lunch

1:00 Presentation on the DST model – *Burton Shank*

- Co-occurrence model and scoring

- Risk/gear threat model and scoring
    - Threat weightings
    - Modeling the threats
    - Model outputs

2:00 Clarifying questions and discussion about models – *Panel*

2:30 Break

2:45 Continue discussion about models – *Panel*
4:00 Public questions and comments
4:15 Discussion of findings/recommendations on DST and supporting models
5:00 Adjourn

February 1 – Recommendations

9:00 Panel closed discussion – *Richard Merrick*
11:00 Recommendations on strategic improvements in the Decision Support Tool - *Panel*
12:00 End meeting

Appendix 2 - Terms of Reference of An Independent Peer Review of the Phase 2-3 expanded Decision Support Tool created for the Atlantic Large Whale Take Reduction Team

The Panel shall, based on information presented by NMFS in two publicly viewable Webinars, review whether the Decision Support Tool (DST):

1. Incorporates the U.S. lobster, gillnet, and other trap/pot[4] commercial fishery dependent (i.e., catch) data sets necessary to determine fishing effort and distribution for the Phase 2 and 3 DST,
2. Incorporates spatially and temporally relevant estimates of right whale distribution and densities,
3. Uses appropriate entanglement risk coefficients by gear type,
4. Incorporates appropriate approaches to applying these coefficients to estimate relative risk (and risk uncertainty) of large whales encountering gear[5], and
5. Using the fishing effort and right whale density data reviewed under TORs 1-2, provides appropriate estimates of risk reduction of large whales encountering gear.

Materials[6] supporting the review of these TORs shall be available to the Panel a minimum of two weeks before the meeting.

Based on this review the Panel will summarize and report their conclusions about the DST with recommendations as to:

1. Whether each of the TORs are met,
2. Short term improvements to the DST that can support the current Take Reduction Plan Phase 2 and 3 regulatory analyses, and
3. Long term modifications to data treatment or model methods that can improve the DST's support of Take Reduction Plan decision making and regulatory analysis.

A draft of this report shall be prepared by the Panel Chair, with Panelist's assistance and presented to NMFS within two weeks of the close of the last meeting.

---

[4] Including at least the Mid-Atlantic mixed species trap/pot, Northeast/Mid-Atlantic American lobster trap/pot, Jonah crab trap/pot fisheries, and Mid-Atlantic/Northeast sink gillnet fisheries

[5] At least two approaches to estimating panel encounter risk (and risk uncertainty) shall be provided by the DST developers - one that relies on information on the vertical distribution of whales and an alternative option such as not considering panel encounter risk.

[6] Including a detailed description of the model and examples risk reduction estimates based on different management decision

Appendix 3 Simple example of a Competing Hazards approach

$g_{j,i,t}$ = gear density for fishery *j* within cell *i* in time interval *t*

$D_{i,t}$ = estimated probability that a whale occurs within cell *i* in time interval *t*
(possibly re-scale density to per-capita probability of occurrence)

S = expected annual "survival" rate from entanglement (i.e., the probability of avoiding death or serious injury from entanglement)

R = expected annualized risk of mortality/serious injury from entanglement (calculated as 1 – *S*): this is the metric for which the target reduction from "baseline value" is 90%

H_j = instantaneous hazard for fishery type *j*, integrated across spatial grid and time intervals (but can also be calculated and compared for any desired sub-area or time interval)

γ_j = log hazard rate associated with fishery type *j*

X_k = fishery and gear covariates (e.g., X_1 = average line strength, X_2 = avg. line length, X_3 = avg. net length per line, etc.). Each should be scaled relative to some reference level.

β_k = log hazard ratios associated with each covariate – note that these effects are additive on a log scale. For example, if a unit increase in line strength (X_1) is associated with a 25% increase in mortality risk for that line, then β_1 = log(1.25) = 0.223.

ω_j = interpreted as "minimum possible log hazard from cause *j*"

α = tuning parameter, to achieve reasonably appropriate magnitude of overall cumulative hazards (could adjust this parameter so that realized S ≈ mean estimated survival rate based on PVA model)

$$\gamma_j = \alpha \cdot \sum_{t}^{T} \sum_{i}^{N} \left[g_{j,i,t} + \beta_1 \left(X_{1,j,i,t} \right) + \beta_2 \left(X_{2,j,i,t} \right) + \ldots \beta_k \cdot X_{k,j,i,t} \right] \cdot D_{i,t}$$
$$H_j = \exp(\omega_j + \gamma_j)$$
$$S = \exp\left(- \sum_{j}^{Fisheries} H_j \right), \quad R = 1 - S$$

<u>Notes:</u>

- Overall annual entanglement risk will vary on 0 – 1 scale, and would be demographically explicit (i.e., would correspond directly to reduction in expected mortalities/serious injuries)
- Effects of management actions can be presented and compared to the baseline scenario in the form of hazard ratios (unitless measures, comparable across fisheries or areas, not sensitive to absolute values)… e.g., scenario A might be associated with a 20% hazard reduction for fishery *j* within area W and season Q
- Adjusting instantaneous hazards for differing time intervals or sub-regions is very straightforward as they are additive: the hazard rate corresponding to a 3-month period is simply *H*÷4 (finite annual rates are not additive and so adjusting for differing temporal/special scales involves taking exponents)
- Fishery-specific hazards, and associated uncertainty, can be easily quantified and compared, while sources of uncertainty (in whale density, gear density, effects of gear variables or modifications) are all incorporated at the log hazard scale, where it is mathematically "safe" to assume gaussian distributions.
- This type of formulation should not change the existing model structure or data inputs, but would simply re-formulate the final computation of risk (and risk reduction)