Data Management: A Key to Effective Conservation

by

Cristiana Falvo

Dr. Douglas Nowacek, Adviser

April 29, 2021

Masters project submitted in partial fulfillment of the
requirements for the Master of Environmental Management degree in
the Nicholas School of the Environment of
Duke University

EXECUTIVE SUMMARY

My master's project leverages data management and visualization skills to support a five-year case study that ultimately will contribute to the conservation of the North Atlantic right whale, *Eubalaena glacialis*. The North Atlantic right whale is one of the most critically endangered species on the planet and, as such, is federally protected under various US environmental laws, one of which is the National Environmental Policy Act (NEPA). NEPA calls for federal agencies to consider the environmental impacts of their proposed actions, which under the law includes direct, indirect and cumulative impacts. The right whale experiences multiple sources of stress that have been shown to have negative impacts on its health. While the scientific understanding of impacts of individual stressors such as noise pollution, prey scarcity and fishing gear has grown, there is still a lack of understanding of the cumulative impacts of multiple stressors on marine mammals.

The Strategic Environmental Research and Development Program (SERDP) is supporting a five-year study that aims to advance the understanding of cumulative impacts of multiple stressors on marine mammals by using two case study examples to develop a modeling framework. The case study that this master's project supports examines how multiple stressors observed in Cape Cod Bay, MA have impacted the health of the North Atlantic right whale over time. By combining rates of exposure to stressors in Cape Cod Bay (underwater noise and prey availability), right whale visual health status (whale sighting data linked to photo ID catalog) and vital rate information (birth and mortality events), the case study team will aim to gain insight into how multiple stressors accumulate and potentially interact to impact the right whale on an individual and population level.

To develop the Population Consequences of Multiple Stressors (PCoMS) modeling framework, the right whale case study team will explore three historic long-term datasets that represent two sources of stress and a long-term health record for the right whale. Noise levels collected by the Cornell Lab of Ornithology's Bioacoustics Research Program and zooplankton distributions collected

by the Center for Coastal Studies represent sources of data to assess the stress to the right whale; these stressors are predicted to have more significant impacts on whale health when experienced in combination compared to on their own. The combination of the long-term sightings database owned by the North Atlantic Right Whale Consortium and the New England Aquarium's photo ID catalog provide long-term health records for identifiable individuals in the population. The existence of these three datasets provides an opportunity to test how fluctuations in noise levels and prey distributions may have interacted and impacted the health and vital rates of right whales over time.

In the context of understanding the cumulative impacts of multiple stressors on marine mammals, the data pieces described are only valuable when they overlap on meaningful spatial and temporal scales. In order to facilitate the exploration and identification of data overlaps, I designed a Python-based data exploration tool that allows the visualization and interactive exploration of the temporal extents of each of the three datasets. This identification of temporal overlaps will serve as the first step in the SERDP team's identification of viable modeling data; this step will narrow the scope of usable data to only those with temporal overlap, which will then set the stage for the team to determine of those data, which have spatial overlap that is meaningful for modeling.

The temporal data exploration tool was designed with two uses in mind: to serve as an interactive exploration tool for the case study team and to provide a usable output for the modelers. The interactive nature of the visualization tool is intended to allow active exploration of the datasets across time and facilitate discussion about what research questions are possible based on the data overlaps that exist. In addition to the visualization component, there is a reported list of 'triple overlap dates' (i.e., days on which all three data streams were collected) that is meant to help both the case study team and the modelers understand which days are viable for modeling. In addition to the reported list of triple overlap dates, the script outputs a CSV file that includes a full report of each survey date that is available for each dataset. This CSV report will allow the modelers to understand where there are overlaps between any combination of two datasets, which may become relevant at later stages of the

study. The Python framework for the tool was written with the awareness of the study's future needs, and thus has built-in flexibility to incorporate additional temporal and spatial data as they're acquired.

The data exploration tool I developed uses data management and visualization to organize disparate pieces of information into one meaningful interface. In doing this, the tool helps facilitate useful discussion and contributes to the development of informed and strategic research questions. In better understanding the overlaps of data in Cape Cod Bay, the right whale case study team can move closer to understanding the stressor dynamics within this and other habitats along the right whale's range. This effort will be a valuable step in advancing the understanding of the cumulative impacts of multiple stressors on the North Atlantic right whale, as well as other marine mammals, which will in turn contribute to more effective management of marine mammals into the future.

ACKNOWLEDGEMENTS

Many thanks to my master's advisor Doug Nowacek for ushering me into this project and working with me through a challenging year. Many thanks to Rob Schick, Enrico Pirotta, Theoni Photopoulou and Peter Tyack for their continued coordination and support through the project. Thanks to the right whale case study team, the PCoMS working group and especially the data contributors - Chris Clark and team from Cornell University's Bioacoustics Research Program, Stormy Mayo and team from the Center of Coastal Studies, and Heather Pettis, Amy Knowlton, Bob Kenny, Phil Hamilton and the numerous other folks who make up the North Atlantic Right Whale Consortium - for providing the incredible long-term data records that make this and future work possible.

A big thanks to Eric Monson at Duke's Center for Data and Visualization Sciences for the support and engagement with this project, his guidance makes me excited about carrying on my data skills to future endeavors. A big thanks to the many Duke professors who advised me through the planning stages of this project.

And tremendous thanks to the friends, family and community around me for giving me the support and energy that carried me through this master's experience.

TABLE OF CONTENTS

INTRODUCTION

The North Atlantic Right Whale

The North Atlantic right whale experiences multiple sources of stress that have been shown to have negative impacts on the health of the species (Kraus and Rolland, 2007). As a result of the right whale's challenging history and continued exposure to stressors, is one of the most critically endangered species of marine mammals on the planet (Cooke, 2021). As such, the right whale is federally protected under various US environmental laws, including the National Environmental Policy Act (NEPA) (Eagle et al., 2017). The goals of this master's project and the study it supports are ultimately aligned with the need to advance the scientific understanding of cumulative environmental impacts as required by NEPA.

The National Environmental Policy Act

The National Environmental Policy Act "forces [federal] agencies to consider the environmental impacts of their proposed actions" by requiring the documentation and approval of any "major action in the resource field" (Eagle et al., 2017). NEPA calls for an agency's considerations to include the "direct, indirect and cumulative effects", or impacts, of their actions on the surrounding environment (National Academies of Sciences [NAS], 2017). Cumulative impacts in this regulatory context are defined as "the incremental effect of a proposed human action when added to those of other human actions" (NAS, 2017). The consideration of cumulative impacts is crucial for the effective management of marine mammals given that the reality for these animals is experiencing the impacts of human activities, not in isolation of one another, but co-occurring constantly in their habitats (Tyack, 2020). NEPA relies on agencies utilizing the best available information to outline the impacts their activities are predicted to have on the environment (Eagle et al., 2017). Currently, there is a gap in existing scientific understanding that hampers the ability to predict the cumulative effects of multiple stressors on marine mammals (Tyack, 2020).

The Cumulative Impacts of Multiple Stressors

The National Academies of Sciences (NAS) has spearheaded past efforts to investigate cumulative impacts of multiple stressors in marine mammals. The Population Consequences of Multiple Stressors (PCoMS) working group, funded by the Office of Naval Research, will build on NAS's past efforts by supporting two marine mammal case studies. The case studies are supported by the Strategic Environmental Research and Development Program (SERDP) and are ultimately aimed at advancing the understanding of the cumulative effects of multiple stressors on marine animals. The data management tool I've developed for my master's project supports the North Atlantic right whale case study team who is working on developing a PCoMS modeling framework using historic data from Cape Cod Bay, Massachusetts.

The PCoMS modeling framework builds on previous efforts of modeling impacts of single stressors on marine mammal individuals and populations. There is a growing understanding of the impacts that single stressors have on marine mammals, including the impacts of noise on stress hormone levels (Rolland et al. 2012); impacts of increased drag (van der Hoop et al. 2017a) and reduced survival rates from entanglement (Robbins et al. 2015). A previous effort that arose from a National Academies of Sciences panel in 2005 developed a framework that translates the impacts that acoustic disturbance, or noise, has on an individual marine mammal to its population as a whole (Pirotta et al., 2018). A later framework known as the Population Consequences of Disturbance, or PCoD, model built on this effort by using the same approach to quantify impacts of stressors other than noise (Pirotta et al., 2018). These past efforts serve as the basis for the right whale case study team's efforts to model the population consequences of exposure to multiple stressors.

METHODOLOGY

The Right Whale Case Study

The North Atlantic right whale case study will use available data from Cape Cod Bay to explore the feasibility of connecting exposure to single and multiple stressors to right whale health, growth, reproduction and survival (NAS, 2017). Doing this will require quantifying and combining rates of exposure to stressors in Cape Cod Bay (underwater noise and prey availability), right whale visual health status (whale sighting data linked to photo ID catalog) and vital rates (birth and mortality events) of right whales observed in the bay within a common time period. The aim of modeling the available data in this way is ultimately to gain insight into the mechanistic pathways where stress impacts manifest and interact in individuals, and better understand how these interactions translate to population-level effects (NAS, 2017). Gaining these insights at both an individual and population level will help address current knowledge gaps in the field which should in turn help agencies better document the predicted impacts of their proposed activities under NEPA (Tyack, 2020).

The right whale case study team plans to approach their five-year study by meeting the following objectives: 1) determine the size of the current dataset of individual whale sightings and health assessments how they are distributed spatially, seasonally and interannually; 2) determine which stressors can be matched to individual whale sightings; 3) test the relationship between stressors and individual whale health status and vital rates through time and in different spatial regions; and 4) explore the relationship between stressors and their combined impact on different metrics of health status and vital rates, and work towards developing a model for the consequences of multi-stressors on North Atlantic right whales (NAS, 2017). In order to make progress in these areas, the case study must first answer the following questions: 1) where and when are individual whales seen; 2) what stressors are concurrent with individual whales in space and time; 3) is exposure to different stressors associated with a change in individual health status and vital rates; and 4) do these stressors interact, and if so, in what way? (NAS, 2017).

The Study Site

Cape Cod Bay, Massachusetts is a small but important area along the North Atlantic right whale's range; the well documented dynamics of this area make it an ideal study site to explore the research questions posed by the right whale case study (Watkins, 1982). Groups of right whales have been observed in Cape Cod Bay since the 1950s as they visit each spring season to feed on dense patches of zooplankton (Schevill et al. 1986, Winn et al. 1986, Hamilton and Mayo 1990, Payne et al. 1990). In recent decades, there have been shifts in the distributions of zooplankton in the bay which have in turn affected the availability of food for the whales (Mayo and Marx 1990). Meanwhile, low-frequency noise from vessel traffic in the bay continues to be a source of chronic stress for right whales (Clark et al. 2009, NAS 2017). The right whale case study is interested in examining how both noise and prey availability have affected the right whale over time, and will do so by exploring the long-term datasets that track these dynamics (NAS, 2017).

The Data Pieces

The right whale case study team is in its beginning stages of acquiring and exploring the data that are available for pursuing its research questions. The universe of existing data that provide information on either right whale health or sources of stress is extensive; available data include but are not limited to documentation of entanglement events, scarring from ship propellers, drone-captured whale blow samples with associated microbiome genetic sequencing, drone-captured photogrammetry measurements, regular and opportunistic zooplankton sampling locations and whale localization estimates derived from processed acoustic data (Tyack, 2020). There are plans to incorporate relevant available pieces going forward, however at this point in the study, the group has acquired data for the investigation of two primary stressors, prey availability and noise, as well as right whale sighting information that links to the long-term health record. These three pieces will serve as the starting point for the data exploration phase and have been assembled into the data exploration tool created for this master's project.

Acoustic Data

Underwater noise has been shown to cause injury, behavioral disturbance, acute and chronic stress in right whales (Rolland et al., 2012). The majority of the low-frequency noise that causes some of these problems comes from large vessels (Hatch et al., 2008), which pass through Cape Cod Bay on a regular basis (Friss, 2019). Underwater noise levels from vessel traffic in and around the bay have been recorded using 4-10 bottom mounted acoustic recorders continuously during the January-May season since 2001 by the Cornell Lab of Ornithology's Bioacoustics Research Program (cf. Clark et al., 2009). This long-term, continuous time series of noise levels in the bay will allow the right whale study team to examine the impacts of chronic noise on right whale health and vital rates, as well as investigate potential interactions between noise and other stressors (Tyack, 2020).

Zooplankton Data

Zooplankton serve as a critical food supply, or prey, for the North Atlantic right whale and are undergoing distribution shifts as a result of changing climatic conditions (Record, 2019). Changes in prey availability for right whales have the potential to cause stress in the individuals and population that have been coming to Cape Cod Bay regularly to feed (Tyack, 2020). The prey distributions in Cape Cod Bay, MA have been measured seasonally since 1984 by Stormy Mayo et al. at the Center for Coastal Studies (Mayo et al., 1990). Prey surveys have been conducted both in conjunction with whale presence (opportunistic sampling) and when whales are absent (regular sampling). The availability of both long-term prey and whale sightings data allow for the examination of how prey availability might be interacting with other forms of stress and ultimately impact the health and vital rates of right whales (Tyack, 2020).

Whale Data

The North Atlantic right whale is a species that has been well-studied for multiple decades (Knowlton et al., 2012). In 1986, a cooperative whale research program with partners from the University of Rhode Island, New England Aquarium, Center for Coastal Studies, Woods Hole Oceanographic Institution, and other organizations formed the North Atlantic Right Whale Consortium (NARWC)

which now holds a wealth of data on the species (North Atlantic Right Whale Consortium [NARWC], 2021). The combination of the NARWC's two primary databases, the sightings database (maintained by the University of Rhode Island) and the Photo-Identification database (maintained and curated by the New England Aquarium), provide a long-term health record for the species that can be traced to individual whales spotted in certain places and time periods (NARWC, 2021).

RESULTS

The Importance of Temporal Overlaps

In the context of understanding the cumulative impacts of multiple stressors on marine mammals, the data pieces described are only valuable when they overlap on meaningful spatial and temporal scales. Understanding the overlap between available acoustic, zooplankton and whale datasets will be a valuable first step in the realistic and strategic development of the multi stressor model. Because it is hypothesized that prey limitation and noise may be two stressors that amplify the negative health effects brought about by other sources of stress (such as entanglement), it is crucial to be able identify when and where the acoustic, zooplankton and whale datasets overlap (Tyack, 2020).

The Data Exploration Tool -- Design and Utility

In order to facilitate the exploration and identification of data overlaps, I developed a Python-based data exploration tool. At this stage in the study, the right whale case study team has acquired all survey dates for each dataset (between 2008 and 2018) but not yet all survey locations. As a result, the tool allows the visualization and interactive exploration of the temporal extents of each of the three datasets and reports 'triple overlap dates' (i.e., days on which all three data streams were collected). This identification of temporal overlap will serve as the first step in the right whale case study team's identification of viable modeling data; this step will narrow the scope of usable data to only those with temporal overlap, which will then set the stage for the team to determine of those data, which have spatial overlap that is meaningful for modeling.

The tool was designed with two uses in mind: to serve as an interactive exploration tool for the case study team and to provide a usable output for the modelers. **Figure 1** shows the two timeline views of the three datasets (distinguished by color); the full temporal extent is shown along the bottom bar and the detailed selection view is shown along the top bar (VanderPlas, 2018). The Altair Python package this tool was built in allows the user to select a time period of interest along the bottom bar that will display in more detail along the top bar. When a selection is made, dates that have occurrences of all

three datasets are reported in a list to the right of the timeline visual. The interactive nature of the visualization is intended to allow active exploration of the datasets across time and facilitate discussion about what research questions are possible based on the data overlaps that exist. The reported list of triple overlap dates is meant to help both the case study team and the modelers understand which days are viable for modeling. In addition to the reported list of triple overlap dates, the script outputs a CSV file that includes a full report of each survey date that is available for each dataset. This CSV report will allow the modelers to understand where there are overlaps between any combination of two datasets, which may become relevant at later stages of the study.

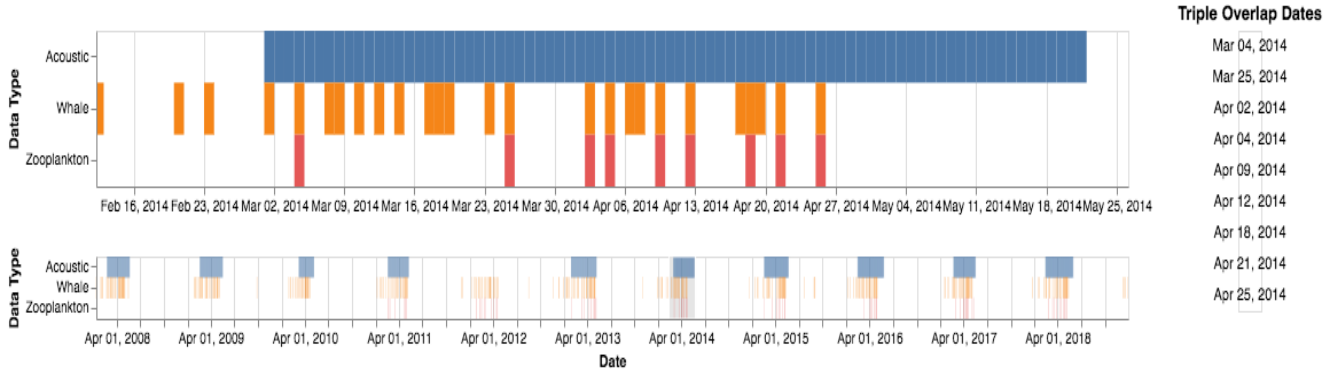

Figure 1 Temporal data exploration tool showing the acoustic (blue), whale (orange), and zooplankton (red) data availability within the selected time range of February - May 2014. The full temporal extent (bottom) and the detailed selection view (top) is shown on the left and the list of dates where all three datasets occur within the selected time period is shown on the right.

DISCUSSION

The Data Exploration Tool -- Flexibility for Future Use

The temporal data exploration tool supports the right whale case study team in their current exploratory phase of their study by facilitating the exploration and output of temporal overlaps between datasets. Creating this tool amidst the early stages of the case study prompted the design to be well documented and flexible in order to accommodate two anticipated needs of the study in the future: 1) the incorporation of additional temporal data and 2) the consideration of spatial overlap between temporally overlapping data.

At the time this master's project was written, the right whale case study team was still in the process of acquiring data for the modeling effort. While the tool's ability to identify temporal overlap between acoustic, whale and zooplankton data between 2008 and 2018 serves as a valuable first step in the exploration process, it will be helpful to understand the overlaps between all data acquired for the study. With the continued acquisition of data in mind, the Python script at the base of this tool provides guidance for formatting new data pieces in a way that allows for their incorporation into the visualization. The Python scripts were thoroughly documented and explained to the team in a closeout meeting to allow for the continued use of the tool.

The Importance of Spatial Overlaps

In order to investigate the cumulative impacts of multiple stressors on marine mammals, an understanding of both the temporal and spatial contexts of observed whales and sources of stress is necessary. The case study team will be acquiring all locations associated with whale sightings, zooplankton surveys and hydrophone deployments, and will eventually need to understand the spatial context associated with each temporal overlap. In anticipation of this need, I modified the existing temporal data exploration tool framework to incorporate a spatial view. The spatiotemporal data exploration tool shown in **figure 2** uses the Altair package to display the locations associated with each survey date selected (note that the zooplankton and whale locations shown are not real, but merely

placeholder data, created for conceptual visualization purposes). The Python framework for the spatiotemporal exploration tool will also be shared with the case study team to facilitate the future needs of their research.

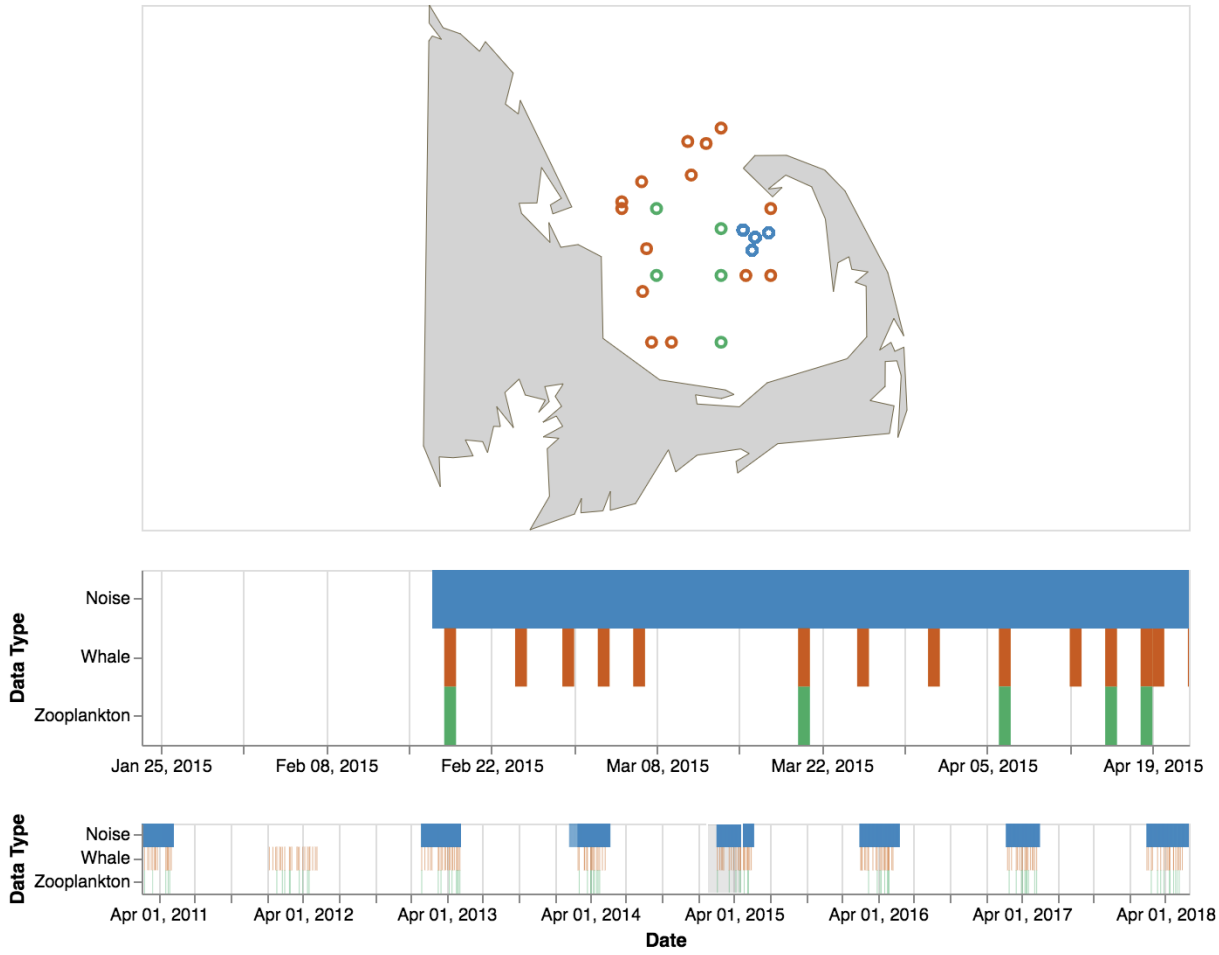

Figure 2 Spatiotemporal data exploration tool showing the acoustic, whale, and zooplankton data availability within the selected time range of February - April 2015. The temporal extent is shown along the two bottom bars (functionality described in **figure 1**) and the spatial context associated with the selected time frame is displayed in the map above (note that the zooplankton and whale locations shown are not real, but merely placeholder data, created for conceptual visualization purposes).

Conclusion

The data exploration tool I developed uses data management and visualization to organize disparate pieces of information into one meaningful interface. In doing this, the tool helps facilitate useful discussion and contributes to the development of informed and strategic research questions. In better understanding the overlaps of data in Cape Cod Bay, the right whale case study team can move closer to understanding the stressor dynamics within this and other habitats along the right whale's range. This effort will be a valuable step in advancing the understanding of the cumulative impacts of multiple stressors on the North Atlantic right whale, as well as other marine mammals, which will in turn contribute to more effective management of marine mammals into the future.

REFERENCES

Clark, C.W., W.T. Ellison, B.L. Southall, L. Hatch, S.M. Van Parijs, A. Frankel and D. Ponirakis. 2009. Acoustic masking in marine ecosystems: intuitions, analysis, and implication. Marine Ecology Progress Series 395:201–222.

Cooke, J.G. 2020. *Eubalaena glacialis* (errata version published in 2020). *The IUCN Red List of Threatened Species* 2020: e.T41712A178589687. https://dx.doi.org/10.2305/IUCN.UK.2020-2.RLTS.T41712A178589687.en. Downloaded on 25 April 2021.

Eagle, J., Salzman, J., & Thompson, B. H. (2017). *Natural resources law and policy*. St Paul, MN: Foundation Press.

Friss, Gwenn. Curious Cape Cod: Canal Rules of the Road. Cape Cod Times. February 2, 2019.

Hamilton, P. K., and C. A. Mayo. 1990. Population characteristics of right whales (*Eubalaena glacialis*) observed in Cape Cod and Massachusetts Bays, 1978–1986. Report of the International Whaling Commission (Special Issue 12):203–208.

Hatch, L., Clark, C., Merrick, R. *et al.* Characterizing the Relative Contributions of Large Vessels to Total Ocean Noise Fields: A Case Study Using the Gerry E. Studds Stellwagen Bank National Marine Sanctuary. *Environmental Management* 42, 735–752 (2008). https://doi.org/10.1007/s00267-008-9169-4

Knowlton AR, Hamilton PK, Marx MK, Pettis HM, Kraus SD (2012) Monitoring North Atlantic right whale *Eubalaena glacialis* entanglement rates: a 30 yr retrospective. Mar Ecol Prog Ser 466:293-302. https://doi.org/10.3354/meps09923.

Kraus, S.D., and R. Rolland. 2007. The Urban Whale: North Atlantic Right Whales at the Crossroads. Cambridge, MA: Harvard University Press.

Mayo, C. A., and M. K. Marx. 1990. Surface foraging behavior of the North Atlantic right whale, *Eubalaena glacialis*, and associated zooplankton characteristics. *Canadian Journal of Zoology* 68:2214–2220.

National Academies of Sciences (NAS), Engineering, and Medicine 2017. Approaches to Understanding the Cumulative Effects of Stressors on Marine Mammals. Washington, DC: The National Academies Press. https://doi.org/10.17226/23479.

North Atlantic Right Whale Consortium 2021 Database 02/20/2021 (Anderson Cabot Center for Ocean Life at the New England Aquarium, Boston, MA, U.S.A.).

Payne, R., V. Rowntree and J. S. Perkins. 1990. Population size, trends and reproductive parameters of right whales (*Eubalaena australis)* off Peninsula Valdes, Argentina. Report of the International Whaling Commission (Special Issue 12):271–278.

Pirotta, E, Booth, CG, Costa, DP, et al. Understanding the population consequences of disturbance.

Ecol Evol. 2018; 8: 9934– 9946. https://doi.org/10.1002/ece3.4458.

Record, Nicholas R., et al. "Rapid Climate-Driven Circulation Changes Threaten Conservation of Endangered North Atlantic Right Whales." *Oceanography*, vol. 32, no. 2, 2019, pp. 162–169. *JSTOR*, www.jstor.org/stable/26651192. Accessed 27 Apr. 2021.

Robbins, J., A.R. Knowlton and S. Landry. 2015. Apparent survival of north Atlantic right whales after entanglement in fishing gear. Biological Conservation 191:421-427.

Rolland, R.M., S.E. Parks, K.E. Hunt, M. Castellote, P.J. Corkeron, D.P. Nowacek, S.K. Wasser and S.D. Kraus. 2012. Evidence that ship noise increases stress in right whales. Proceedings of the Royal Society B: Biological Sciences 279: 2363-2368.

Schevill, W. E., W. A. Watkins and K. E. Moore. 1986. Status of *Eubalaena glacialis* off Cape Cod. Report of the International Whaling Commission (Special Issue 10):79–82.

Tyack, Peter. Towards an Understanding of the Cumulative Effects of Multiple Stressors on Marine Mammals - an Interdisciplinary Working Group with Case Studies. Strategic Environmental Research and Development Program. RC20-1097. June, 2020. https://serdp-estcp.org/Program-Areas/Resource-Conservation-and-Resiliency/Natural-Resources/Living-Marine-Resources-Ecology-and-Management/RC20-1097/RC20-1097.

van der Hoop, J., P. Corkeron, M. Moore. 2017a. Entanglement is a costly life-history stage in large whales. Ecology and Evolution 7(1):92-106.

VanderPlas, J., Granger, B., Heer, J., Moritz, D., Wongsuphasawat, K., Satyanarayan, A., ... Sievert, S. (2018). Altair: Interactive statistical visualizations for python. *Journal of Open Source Software*, *3*(32), 1057.

Watkins, W. A., and W. E. Schevill. 1982. Observations of right whales, *Eubalaena-glacialis*, in Cape Cod waters. *Fishery Bulletin* 80:875–880.

Winn, H. E., C. A. Price and P. W. Sorensen. 1986. The distributional biology of the right whale (*Eubalaena glacialis*) in the Western North Atlantic. Report of the International Whaling Commission (Special Issue 10):129–138.