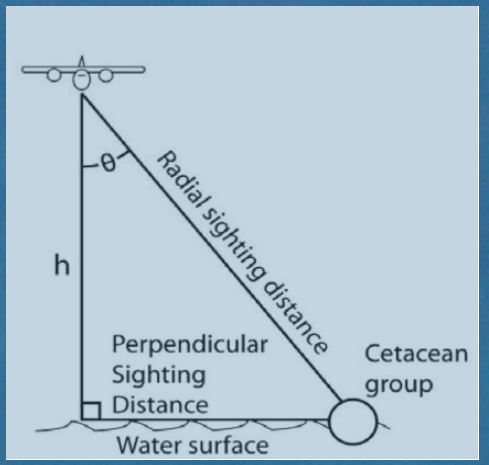
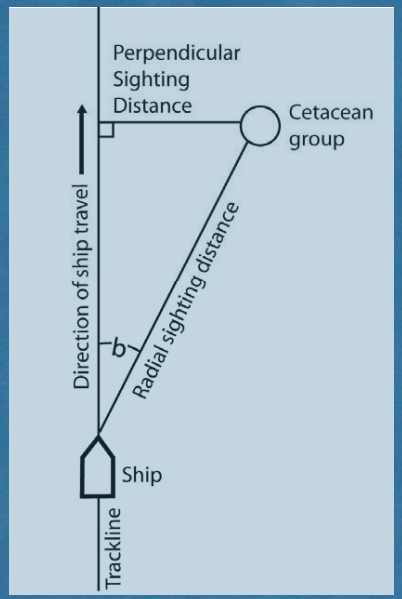
Density Surface Modeling

WMMC Distance Workshop
7 December 2019





Jason J. Roberts[1]
David L. Miller[2]

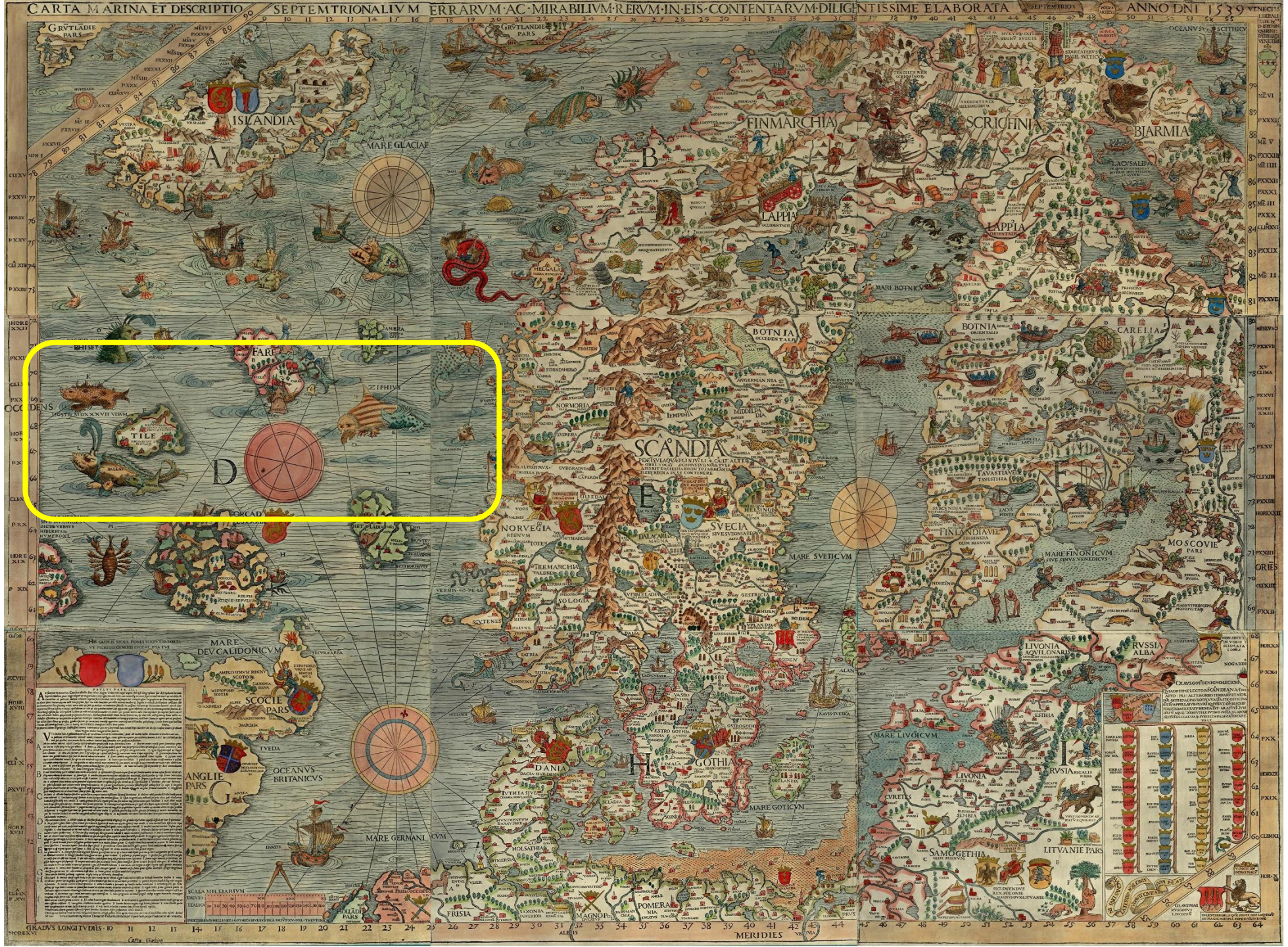
[1] Marine Geospatial Ecology Lab, Duke University

[2] Centre for Research into Ecological and Environmental Modelling, University of St. Andrews

Photo: NOAA/NEFSC, Permit #17355

Münster (1552)
Meerwunder und seltzame Thier

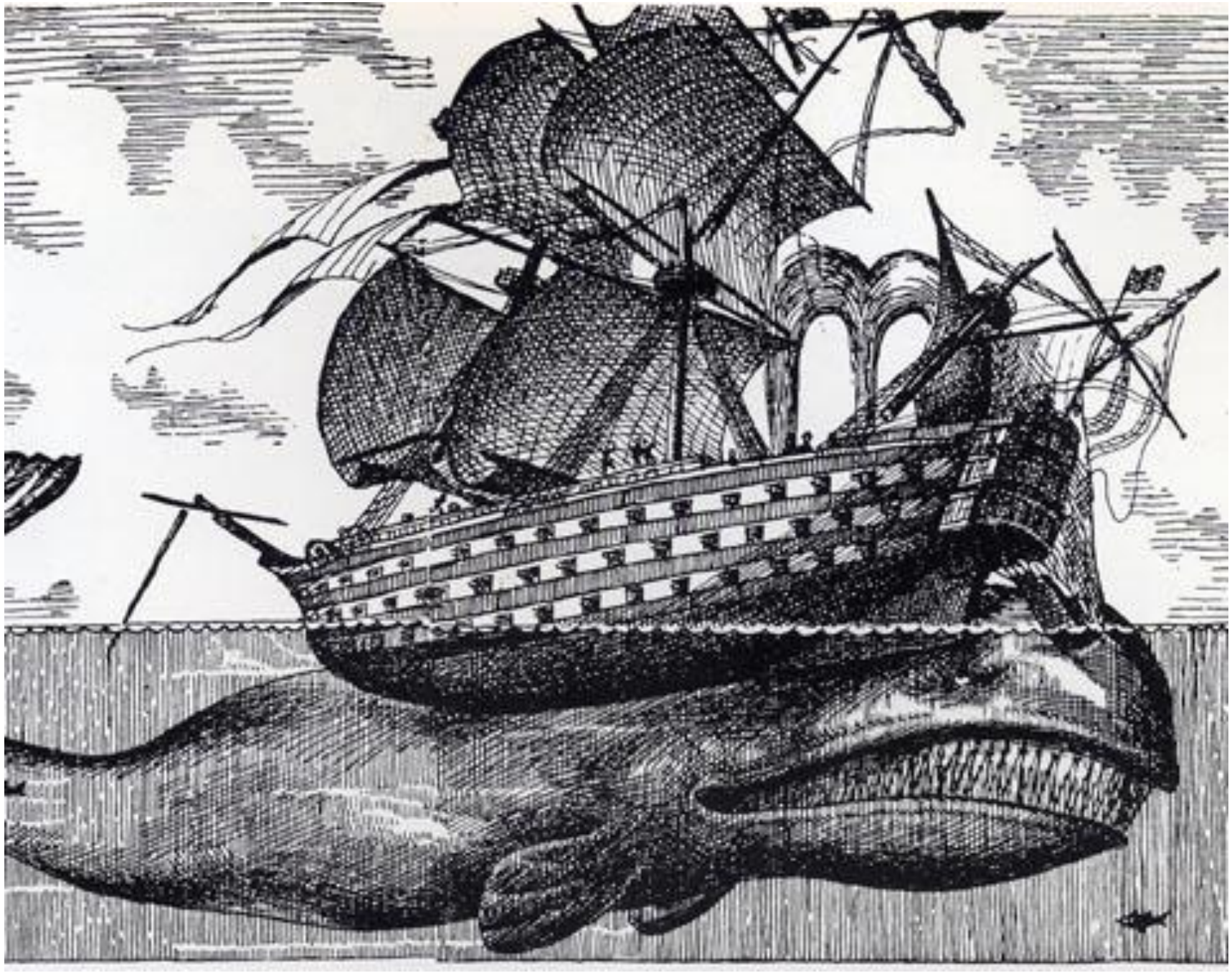
"Here be dragons!"



Olaus Magnus (1539) *Carta Marina*

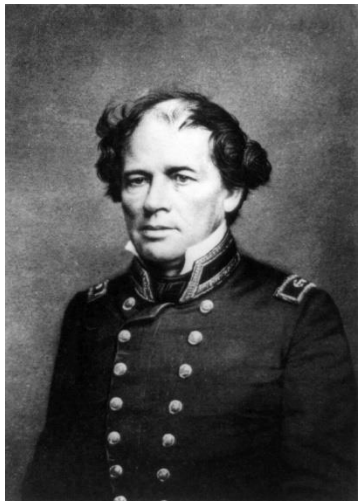
Olaus Magnus (1539) *Carta Marina*

Olaus Magnus (1539) *Carta Marina*

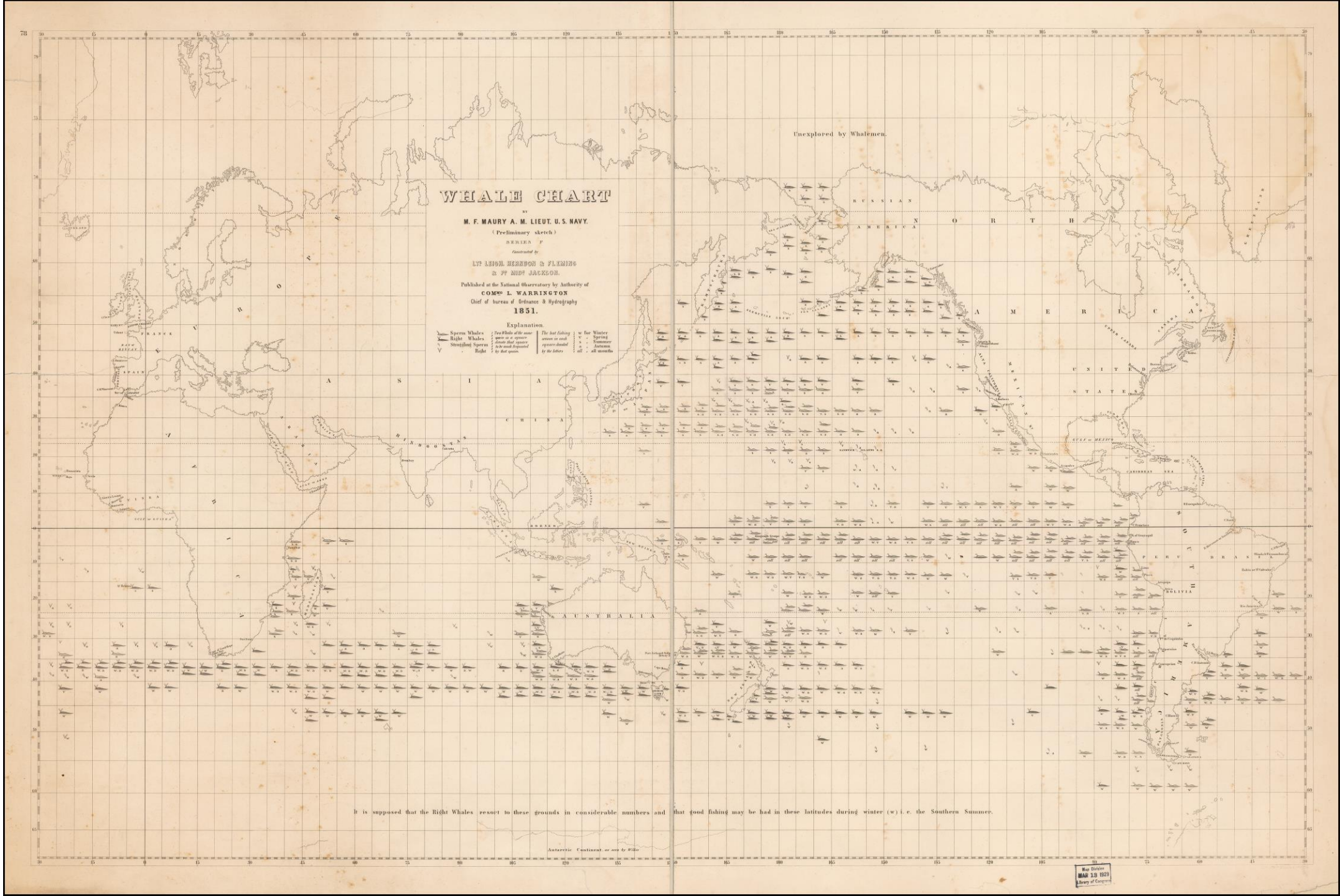


Whaleship
Essex, 1820

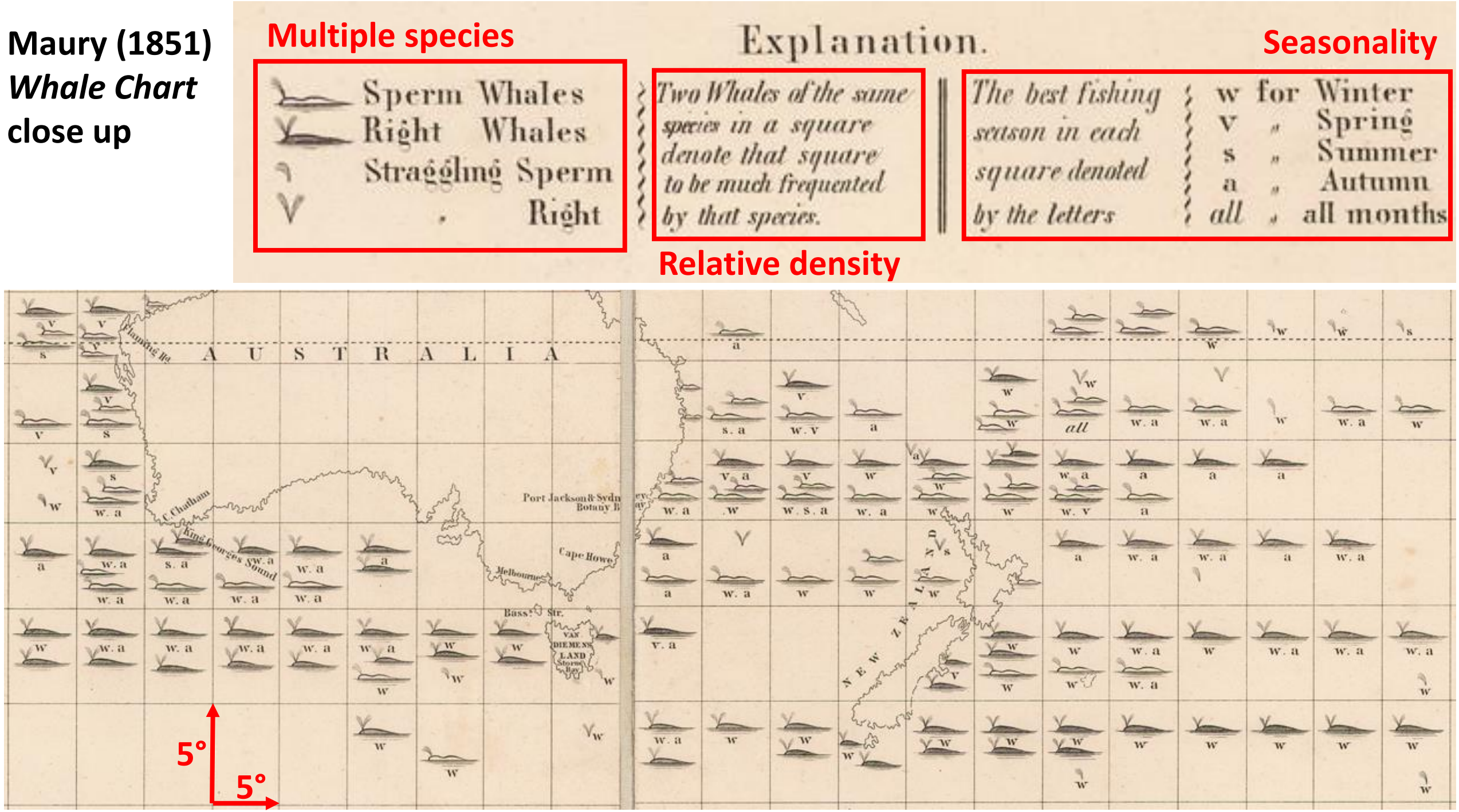
Maury (1851)
Whale Chart



Lieut. Matthew Fontaine Maury, U.S. Navy



Maury (1851)
Whale Chart
close up



New Bedford docks, Massachusetts, USA 1870

Photos: New Bedford Whaling Museum

Ready to shoot the Whale—Harpoon shown in position.


Photo: www.grayline.com

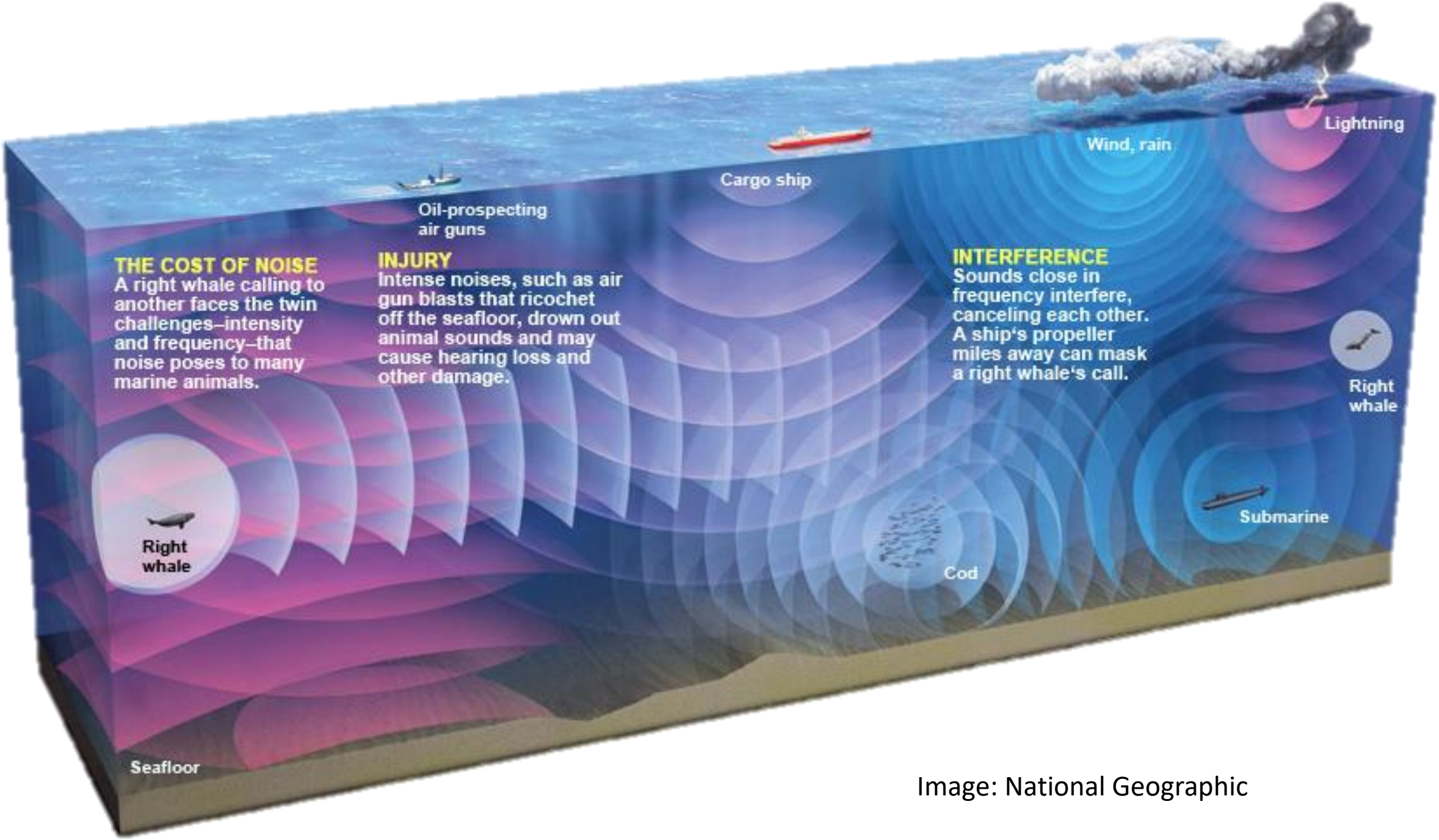


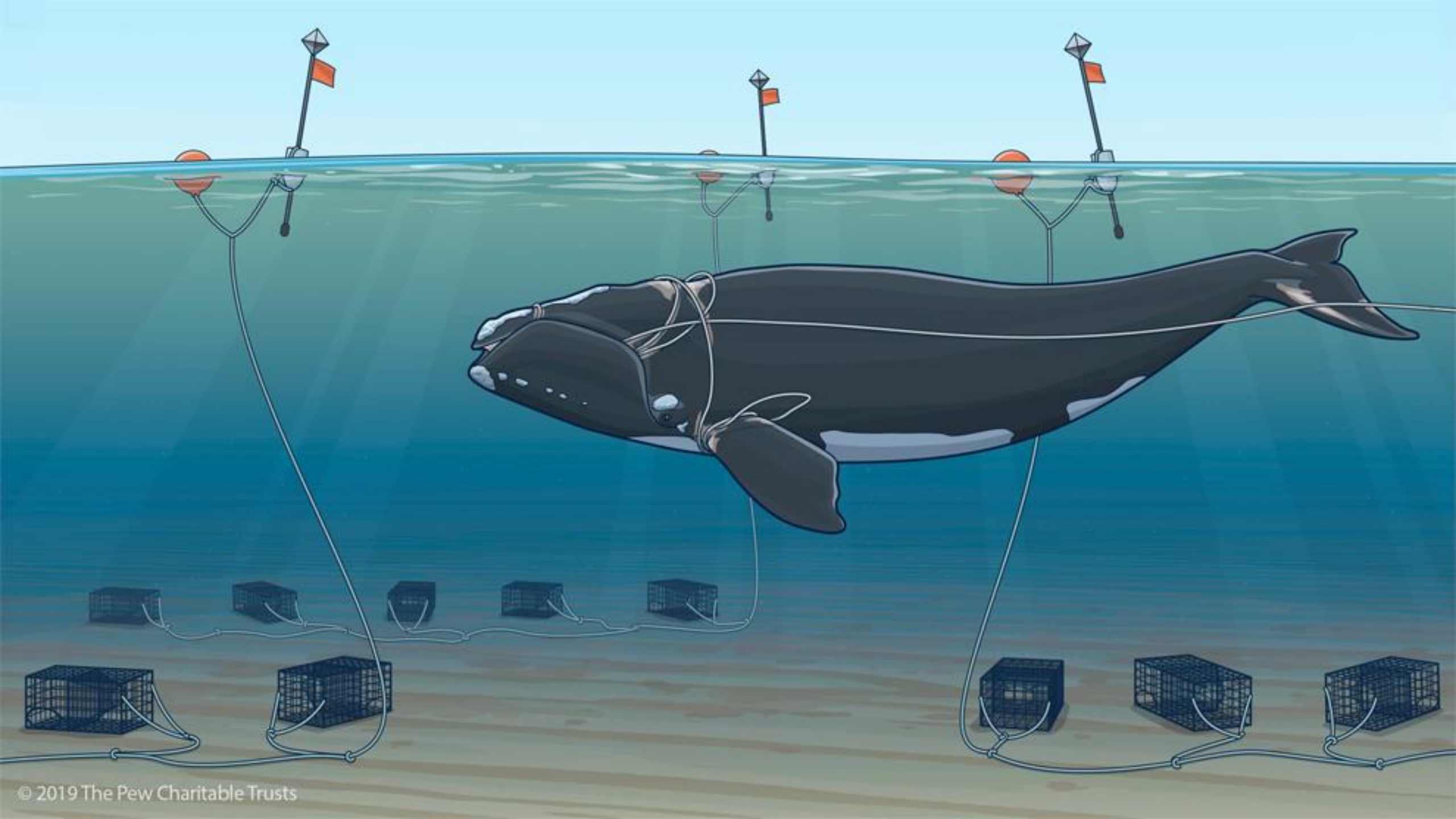
Image: National Geographic

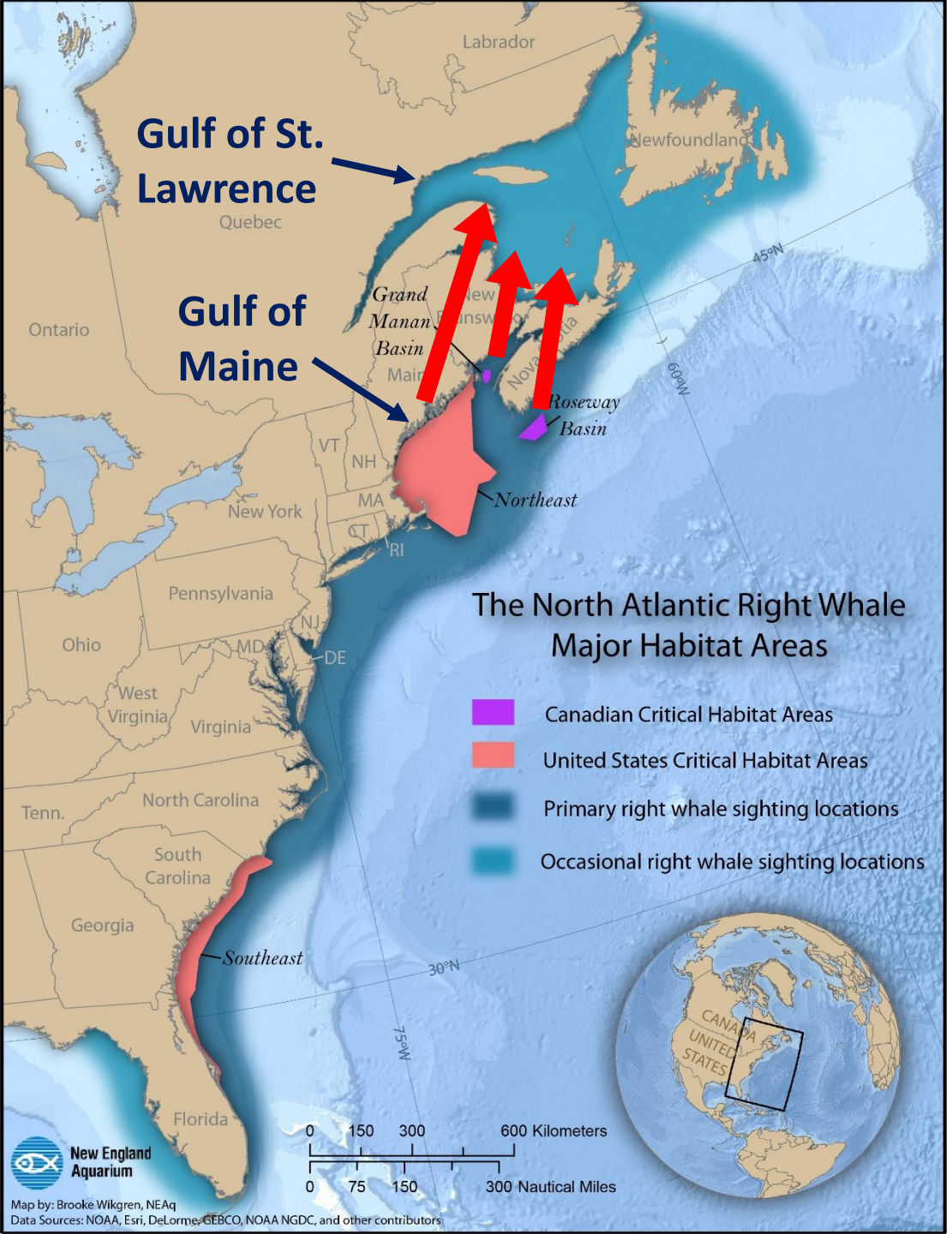
Photo: NOAA

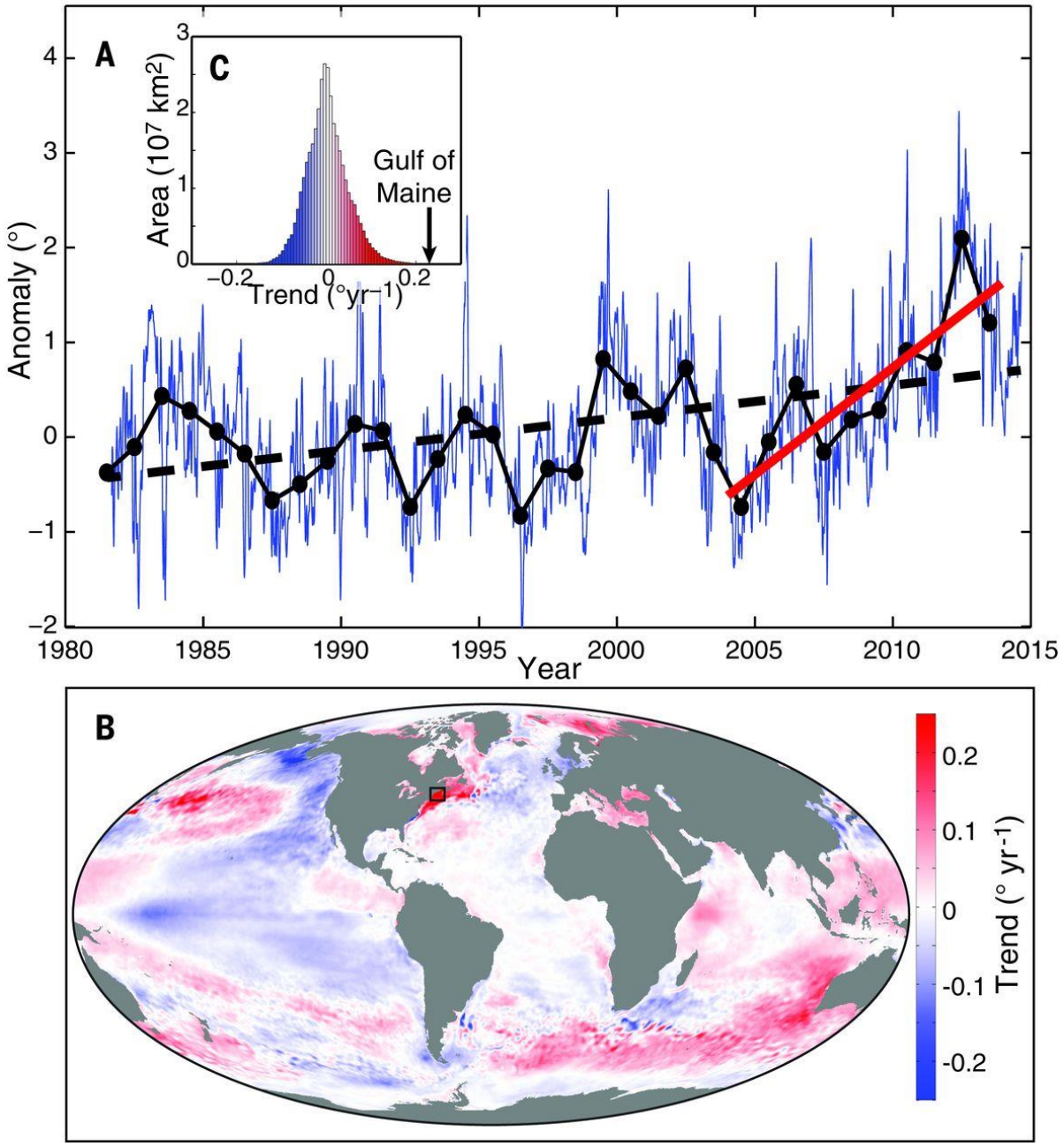

Photo: Kalli de Meyer, Bonaire Marine Park


Photo: Souffleurs d'Écume

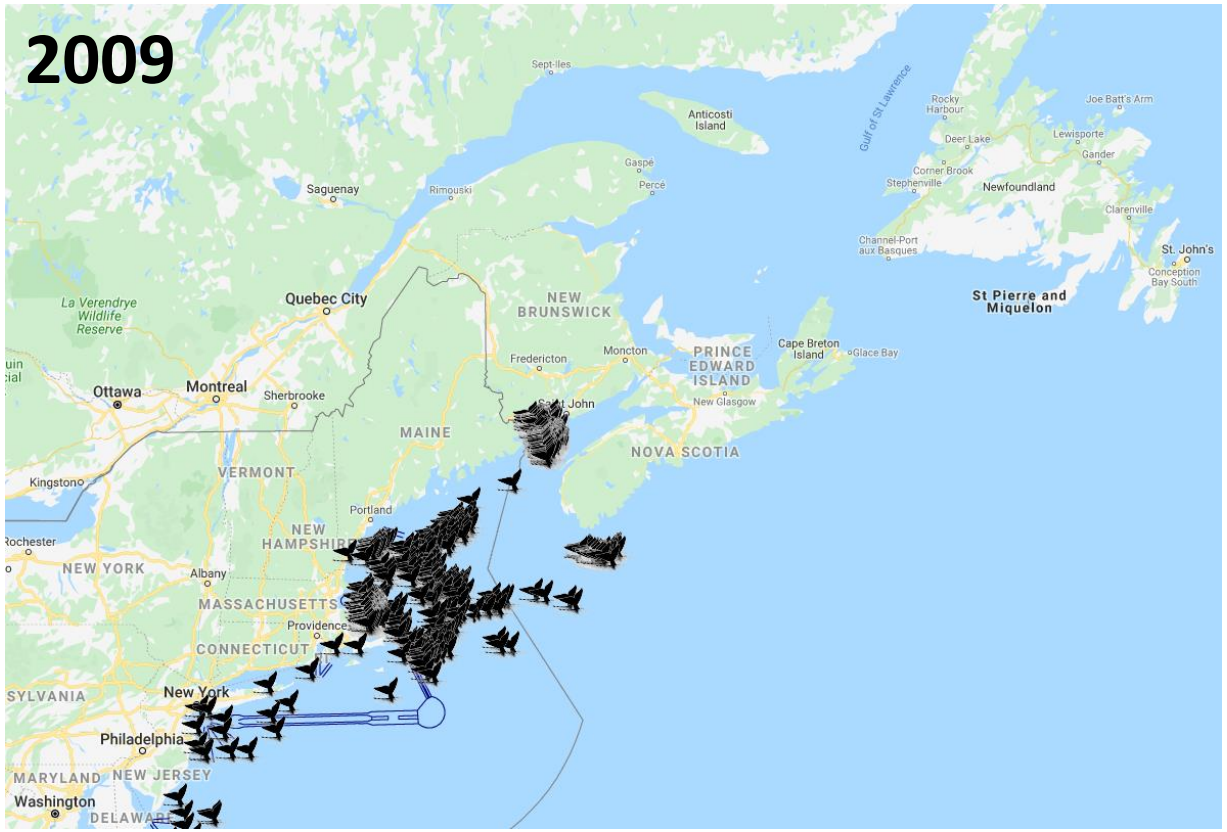

Photo: The Maritime Executive

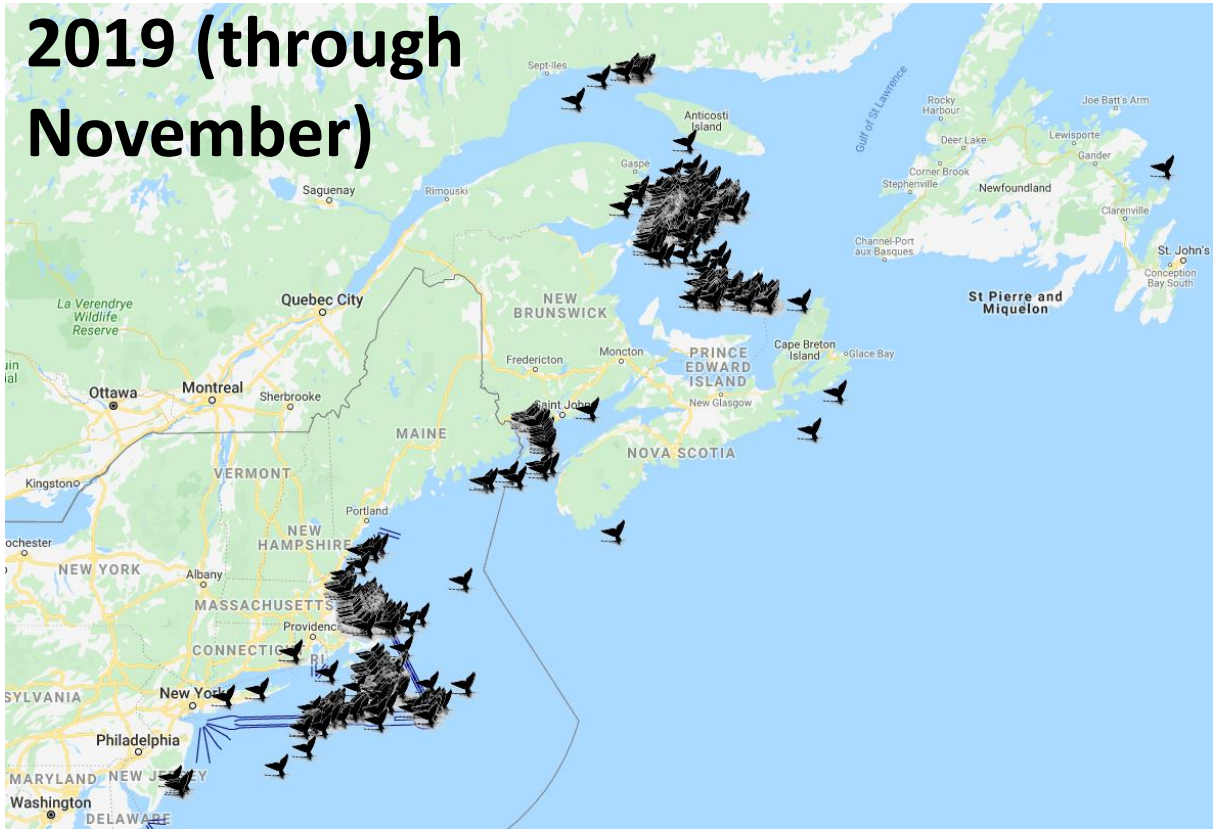






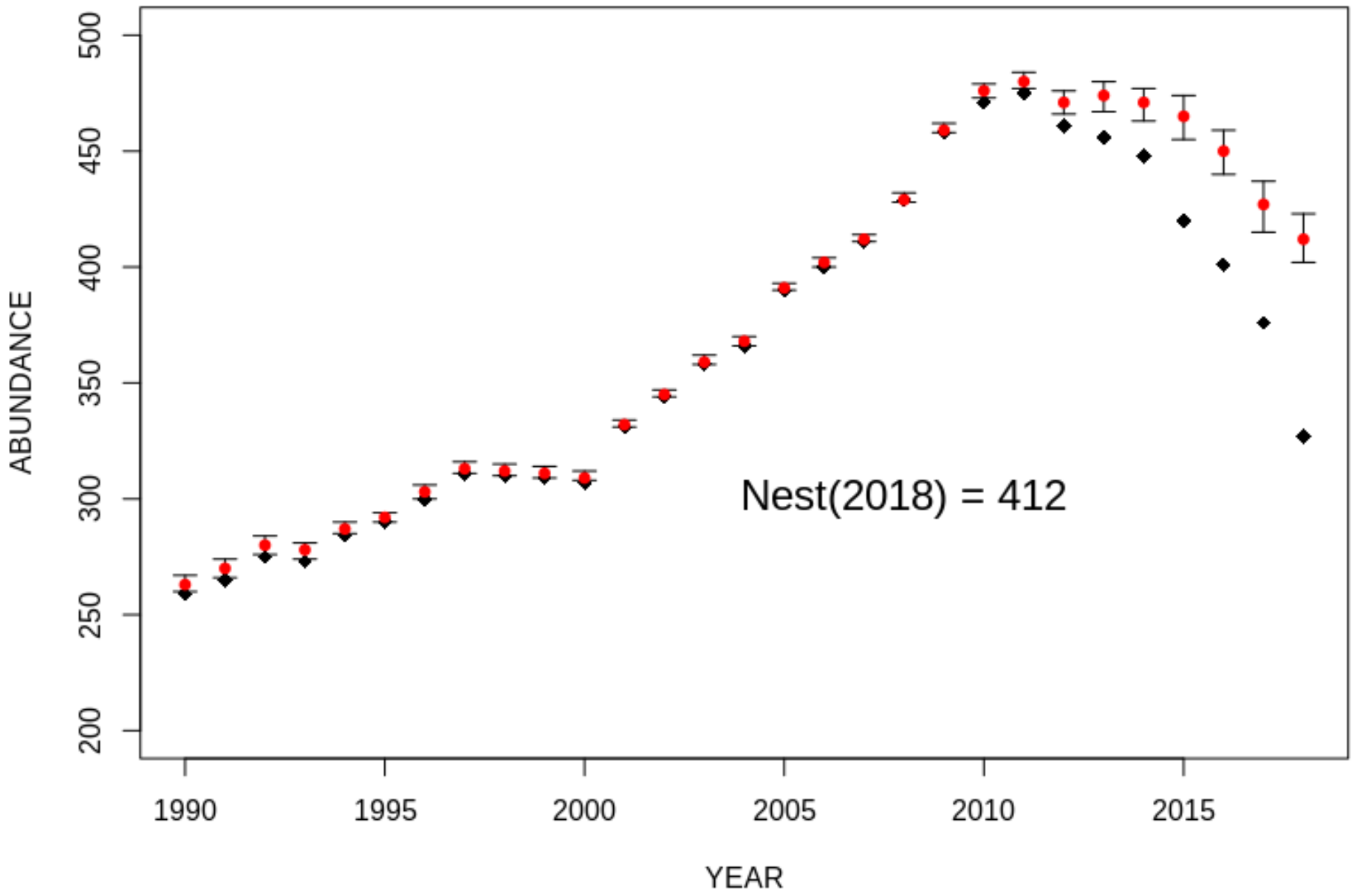
Pershing et al. (2015)





Data: NOAA NARWSS

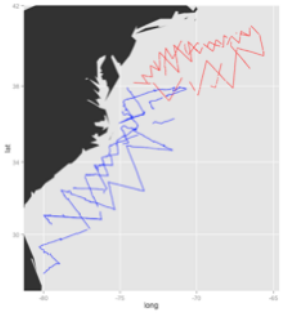
North Atlantic Right Whale Abundance (Pace, 2019)

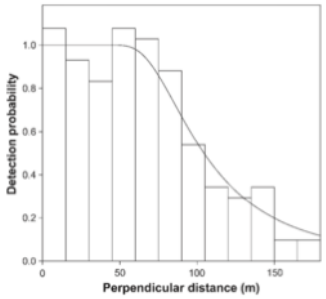


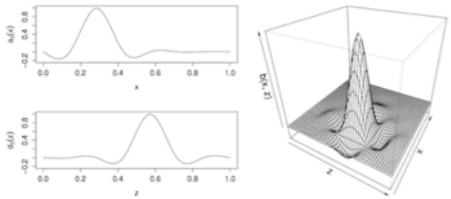
We need some maps:

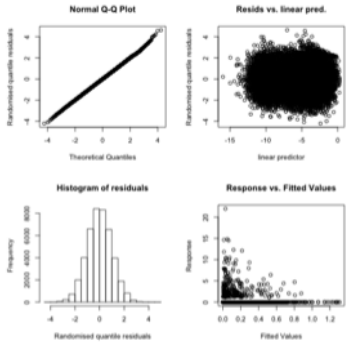
- Where are they?
- How many are there?
- How do they move seasonally?
- Where will they go in the future?

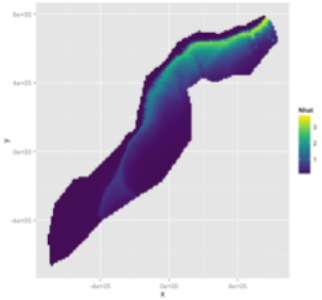
Density surface modeling











Line transects → Detectability → Spatial model (GAM) → Model checking/ criticism → Predictions

(or point transects)

Spatial covariates → Spatial model (GAM)

Spatial covariates → Predictions

Stage 1

Stage 2

Adapted from figure by David L. Miller

The seminal paper (2004)

Spatial Models for Line Transect Sampling

Sharon L. HEDLEY and Stephen T. BUCKLAND

This article develops methods for fitting spatial models to line transect data. These allow animal density to be related to topographical, environmental, habitat, and other spatial variables, helping wildlife managers to identify the factors that affect abundance. They also enable estimation of abundance for any subarea of interest within the surveyed region, and potentially yield estimates of abundance from sightings surveys for which the survey design could not be randomized, such as surveys conducted from platforms of opportunity. The methods are illustrated through analyses of data from a shipboard sightings survey of minke whales in the Antarctic.

Journal of Agricultural, Biological, and Environmental Statistics, Volume 9, Number 2, Pages 181–199
DOI: 10.1198/1085711043578

Update with example R code (2013)

Methods in Ecology and Evolution

British Ecological Society

Methods in Ecology and Evolution 2013, **4**, 1001–1010 doi: 10.1111/2041-210X.12105

Spatial models for distance sampling data: recent developments and future directions

David L. Miller[1]*, M. Louise Burt[2], Eric A. Rexstad[2] and Len Thomas[2]

[1]*Department of Natural Resources Science, University of Rhode Island, Kingston, RI 02881, USA; and* [2]*Centre for Research into Ecological and Environmental Modelling, The Observatory, University of St Andrews, St Andrews KY16 9LZ, UK*

Stage 2 of traditional DSM line transect analysis

1. Split transects into segments

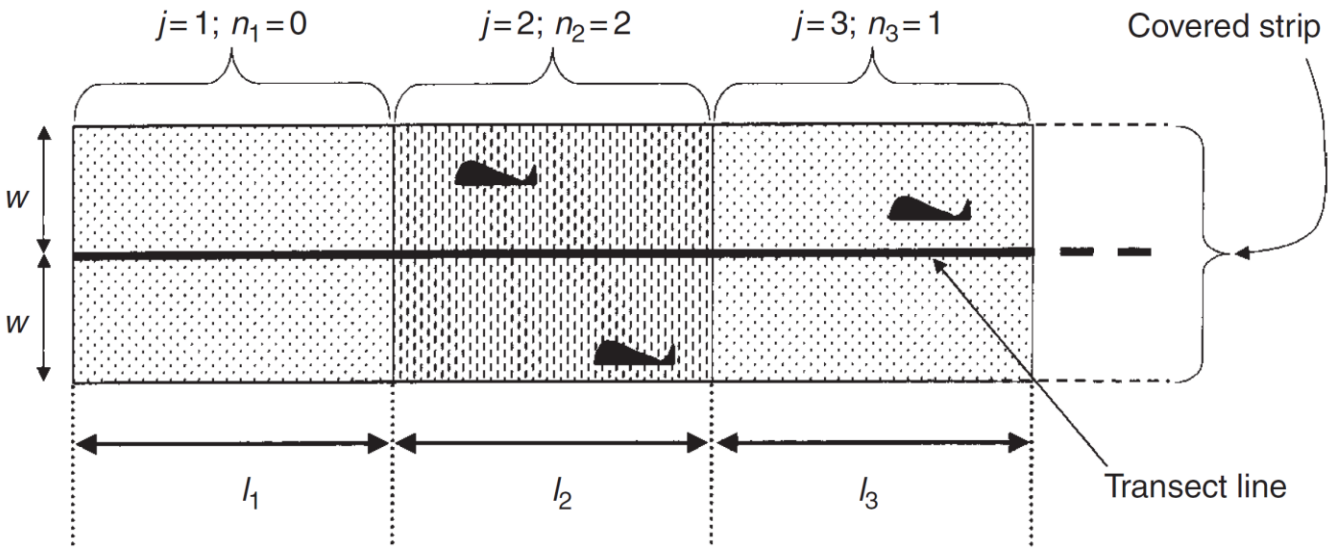


Fig. 4.4. Illustration of notation for the count model, using three hypothetical segments of lengths l_1, l_2, and l_3.

j : segment index
k : spatial model covariate index

2. Model count per segment with spatial covariates

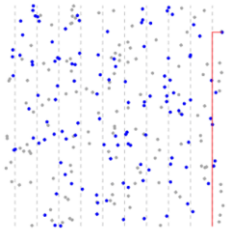
$$\mathbb{E}(n_j) = \hat{p}_j A_j \exp\left[\beta_0 + \sum_k f_k(z_{jk})\right]$$

$$A_j = 2wl_j$$

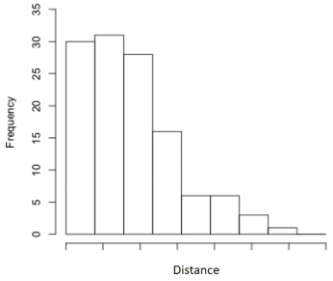
n_j : count of individuals for segment j
$\hat{p}_j$: probability of detection for segment j
A_j : area for segment j
f_k : smooth of covariate k for values z_{jk}

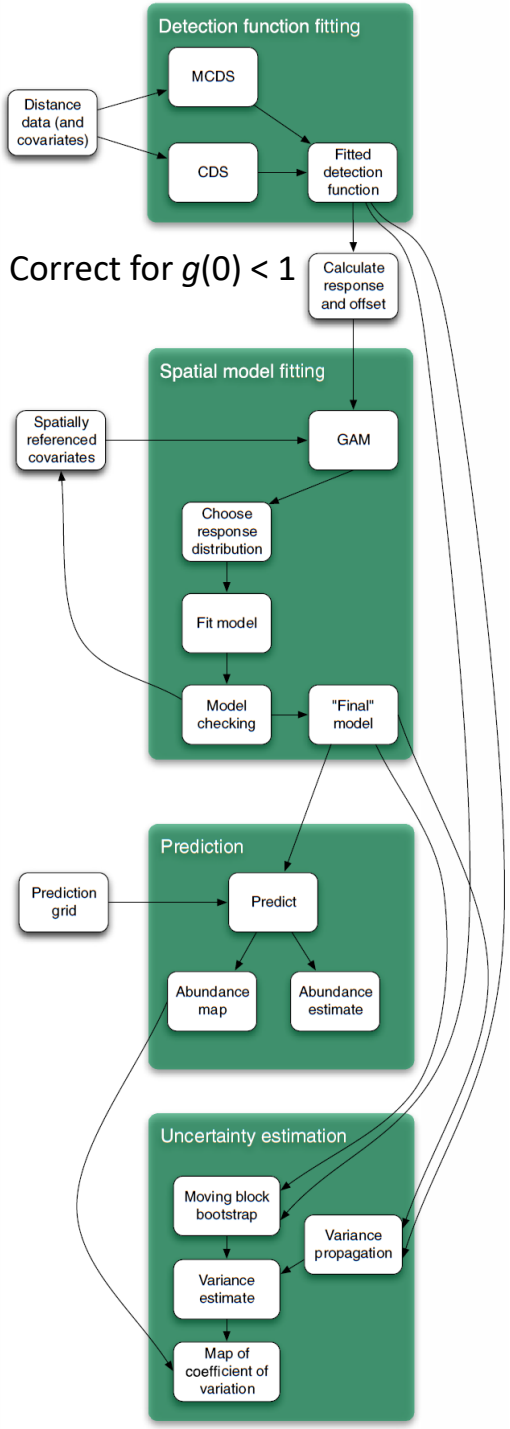
Line transect surveys

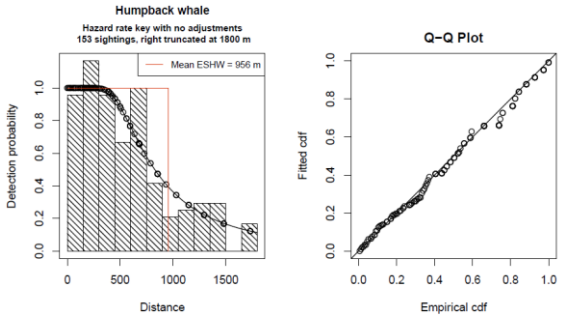
Survey area



Histogram of distances







Model the effects of:

- Distance to sighting
- Group size
- Sea state / glare / etc.

on probability of sighting

Availability bias due to diving

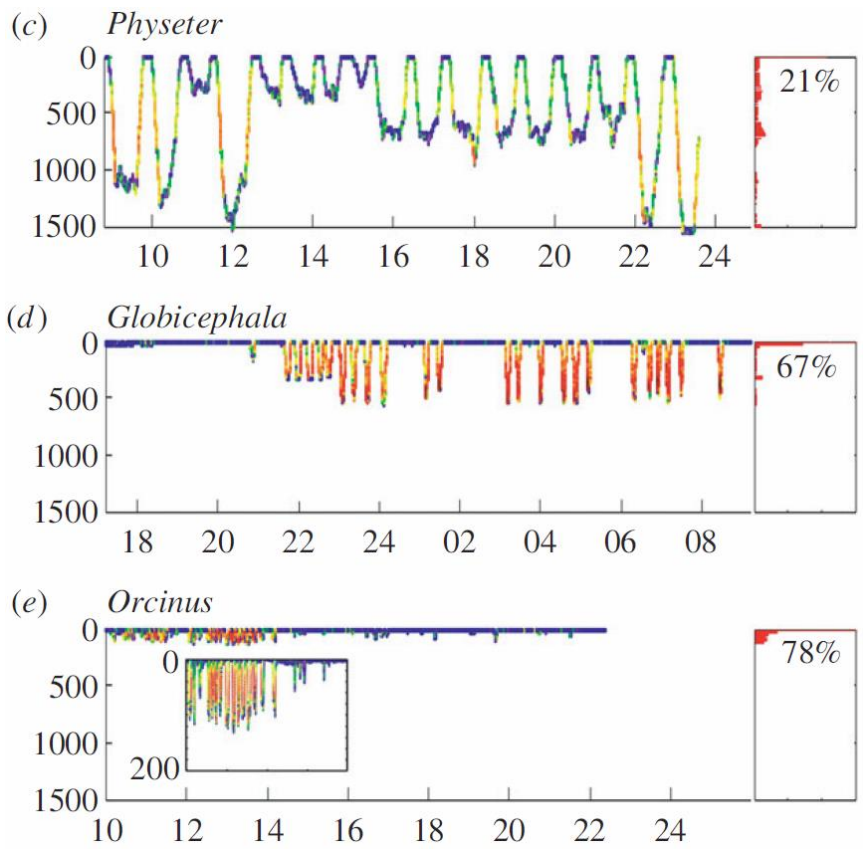


Figure: Hooker et al. (2011)

DSM workflow

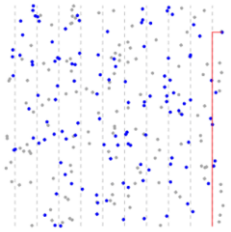
Adapted from Miller et al. (2013)

DSM workflow

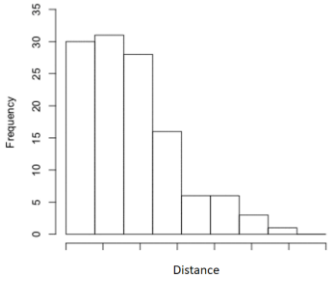
Adapted from Miller et al. (2013)

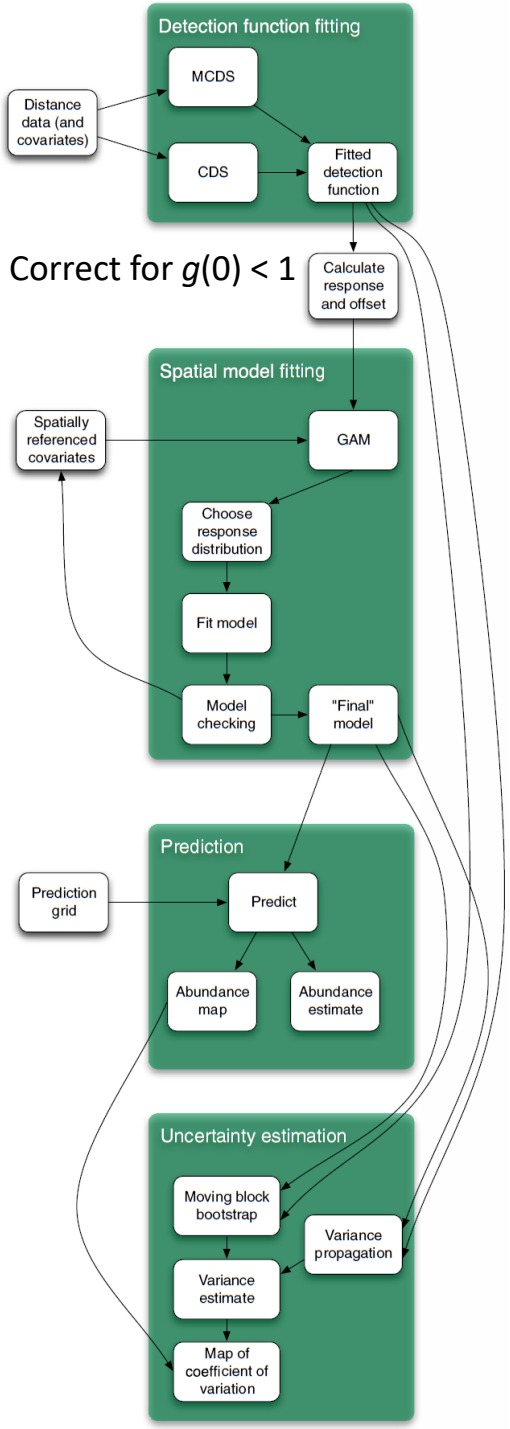
Line transect surveys

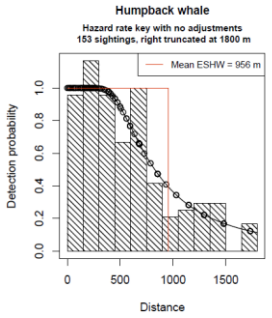
Survey area

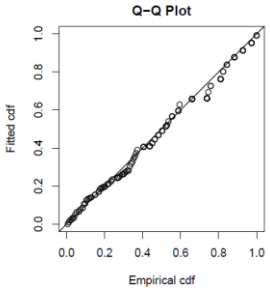


Histogram of distances









Model the effects of:

- Distance to sighting
- Group size
- Sea state / glare / etc.

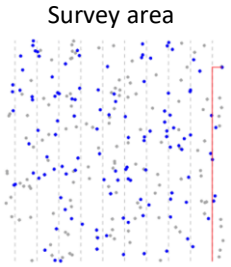
on probability of sighting

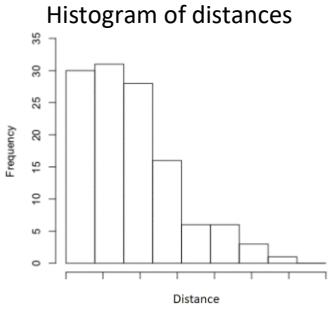
Perception bias

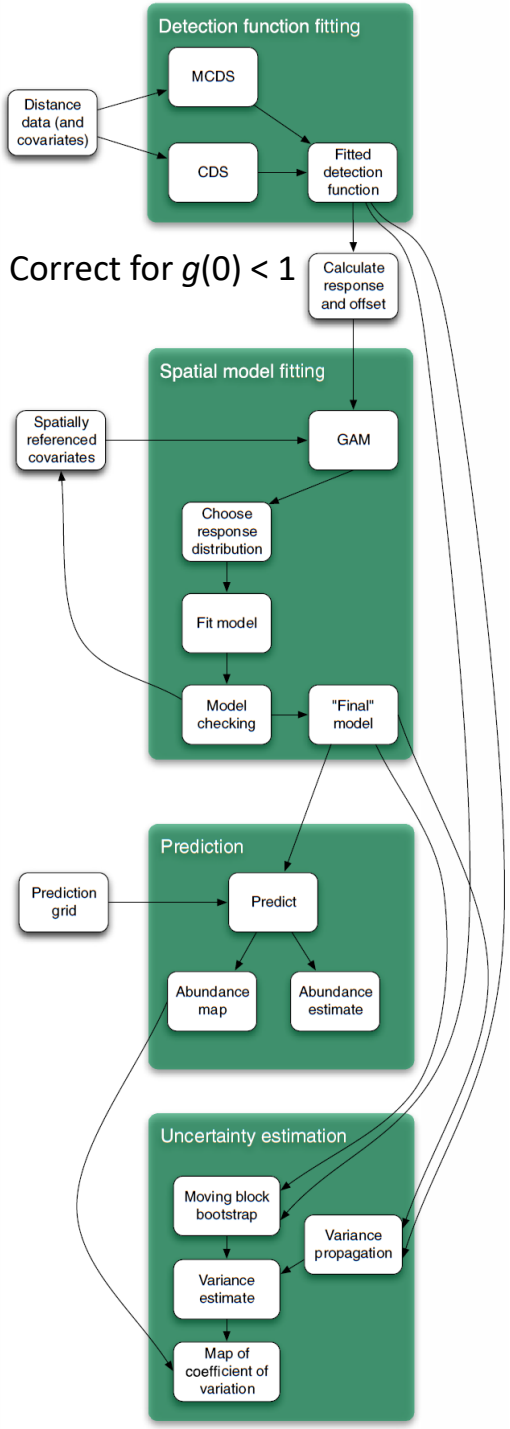
Photo: MARY_OREL

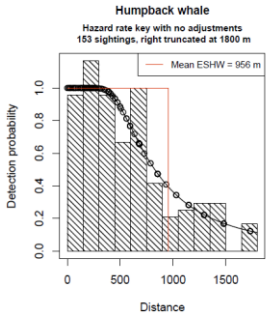
Photo: Minette Layne

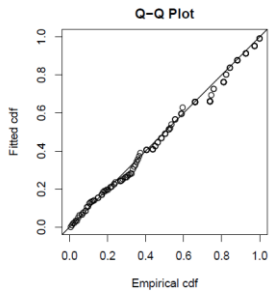
Line transect surveys











Model the effects of:

- Distance to sighting
- Group size
- Sea state / glare / etc.

on probability of sighting

Typical approach to bias correction:

1. Estimate availability bias correction factor from diving data
2. Estimate perception bias correction factor with an independent observer team, or similar methods
3. Multiply correction factors to get $g(0)$
4. Multiply $\hat{p}_j$ by $g(0)$ in second stage:

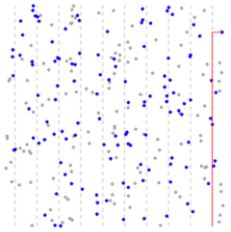
$$\mathbb{E}(n_j) = \hat{p}_j \, g(0) \, A_j \exp\left[\beta_0 + \sum_k f_k(z_{jk})\right]$$

DSM workflow

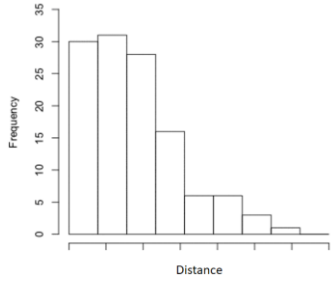
Adapted from Miller et al. (2013)

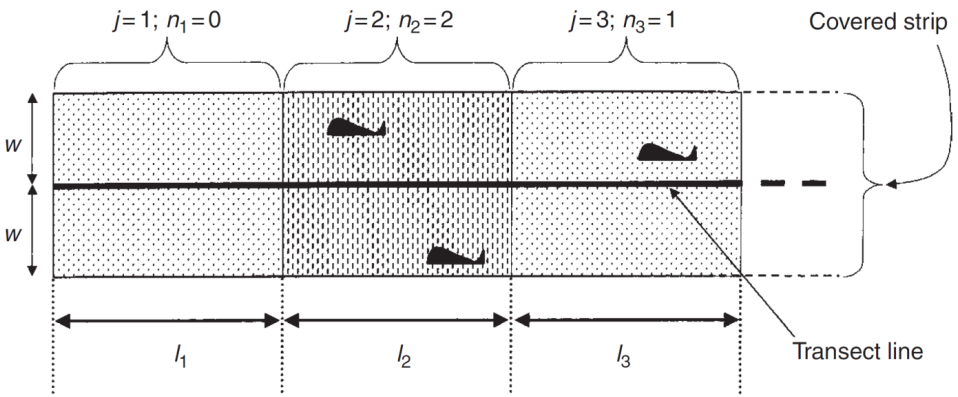
Line transect surveys

Survey area

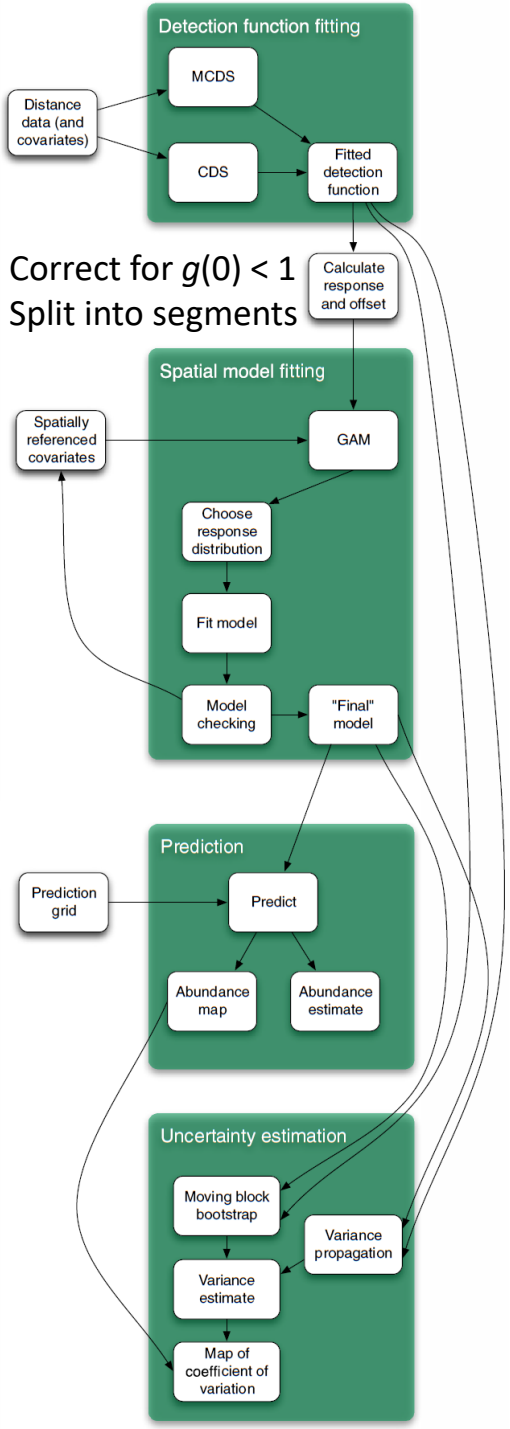


Histogram of distances

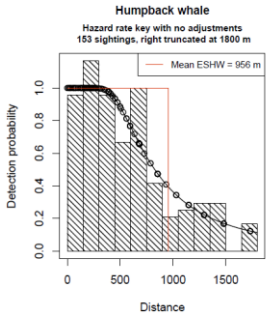


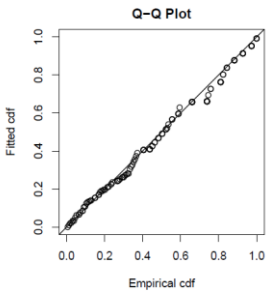


Can also split into grid cells



Correct for $g(0) < 1$
Split into segments





Model the effects of:

- Distance to sighting
- Group size
- Sea state / glare / etc.

on probability of sighting

DSM workflow

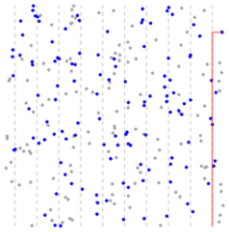
Adapted from Miller et al. (2013)

DSM workflow

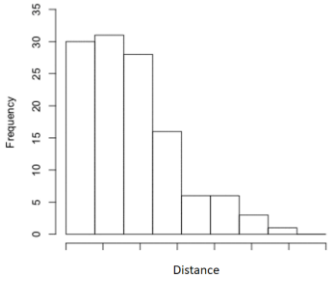
Adapted from Miller et al. (2013)

Line transect surveys

Survey area

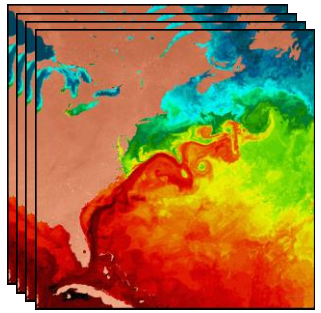


Histogram of distances



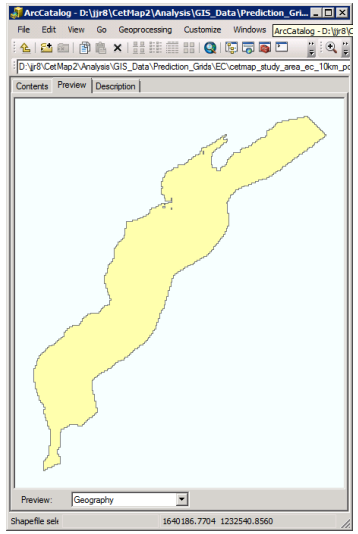
Sample environmental covariates at segments

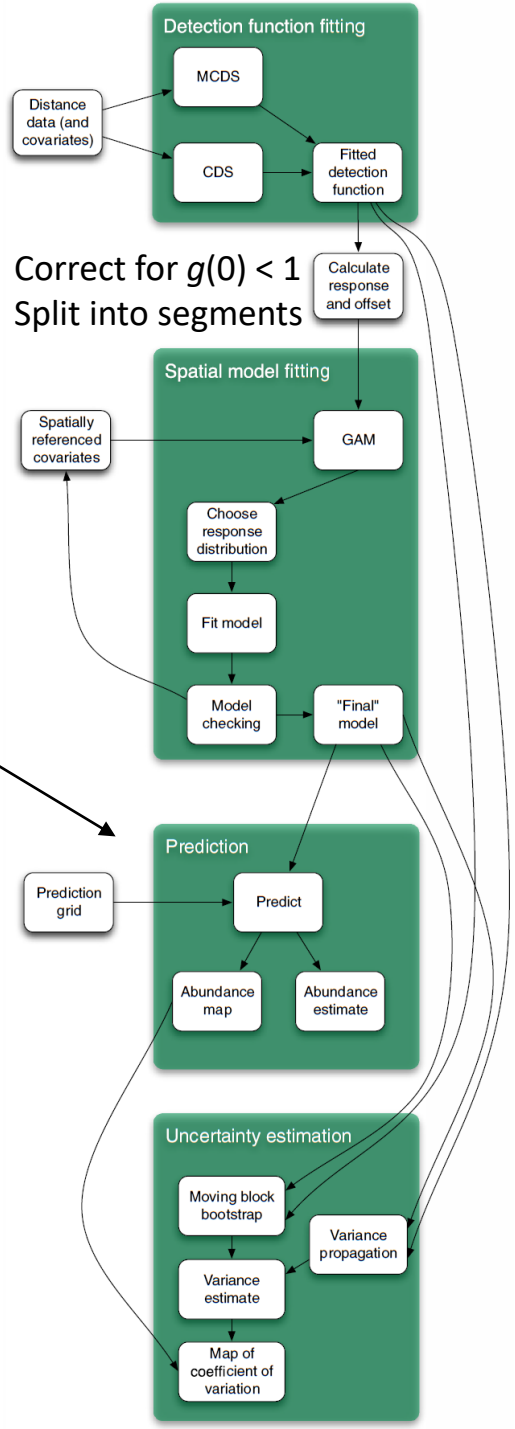
- Physiographic (e.g. depth)
- Physical (e.g. SST)
- Biological (e.g. prey)

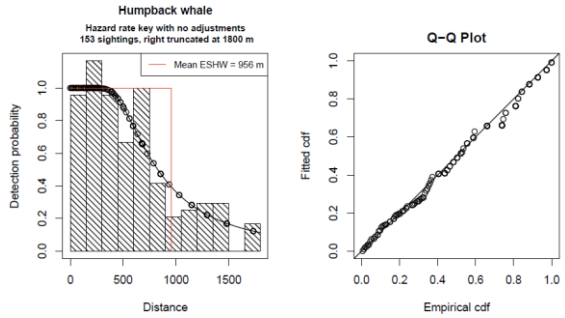


Grids

Study area extent



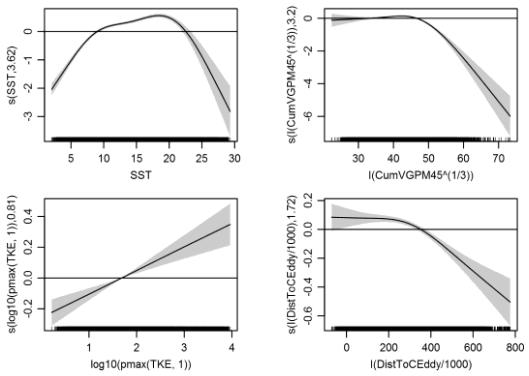




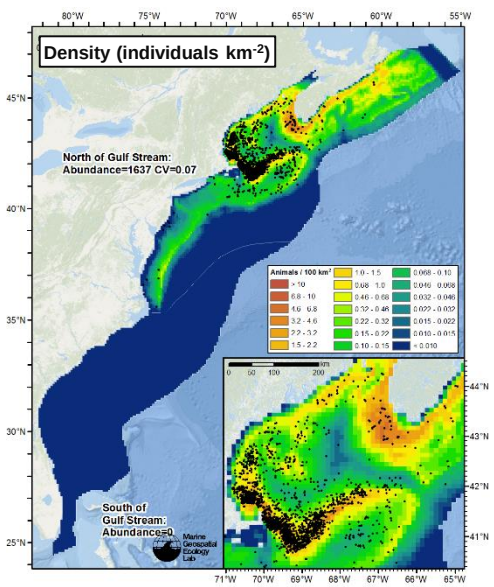
Model the effects of:

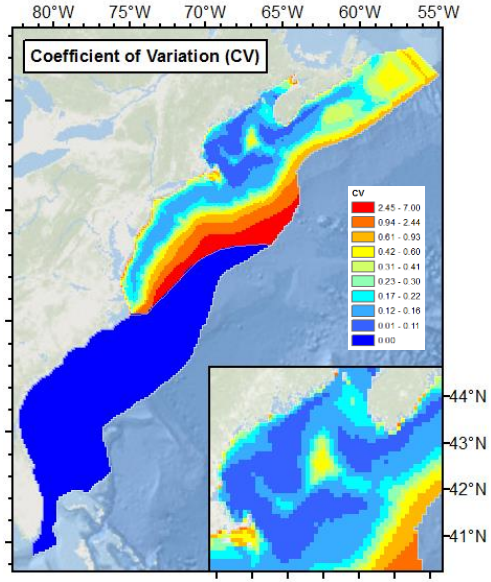
- Distance to sighting
- Group size
- Sea state / glare / etc.

on probability of sighting



Model density from spatial, temporal, and environmental covariates





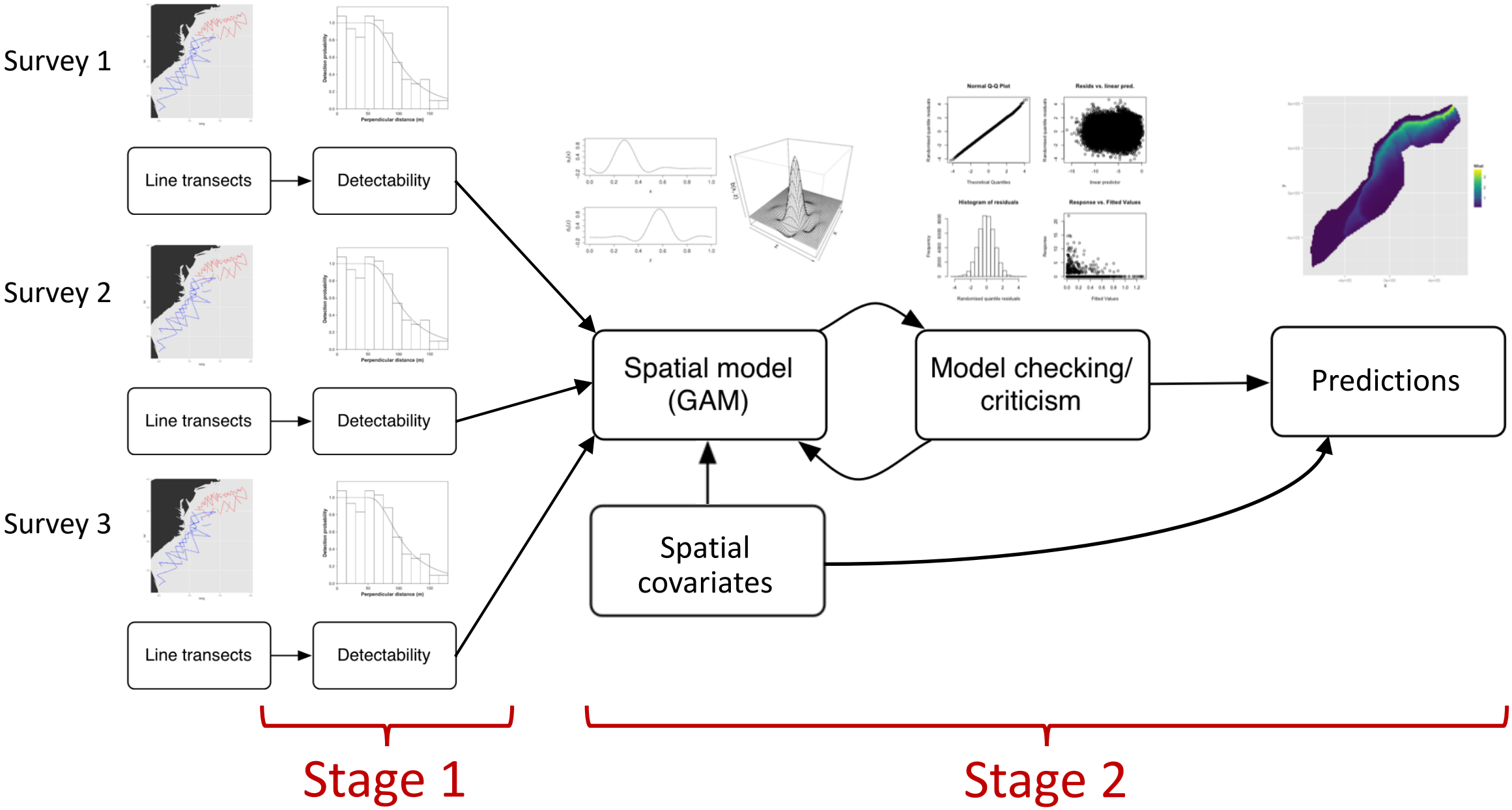
Recent research trends

- Integrating multiple surveys to facilitate models with greater spatial, temporal, and taxonomic coverage
- Extrapolating models across unsurveyed areas, times, and the future
- Incorporating multiple sources of uncertainty into density surface predictions

Recent research trends

- Integrating multiple surveys to facilitate models with greater spatial, temporal, and taxonomic coverage
- Extrapolating models across unsurveyed areas, times, and the future
- Incorporating multiple sources of uncertainty into density surface predictions

Common idea: standardize surveys via detection modeling



Roberts et al. (2016)

www.nature.com/scientificreports

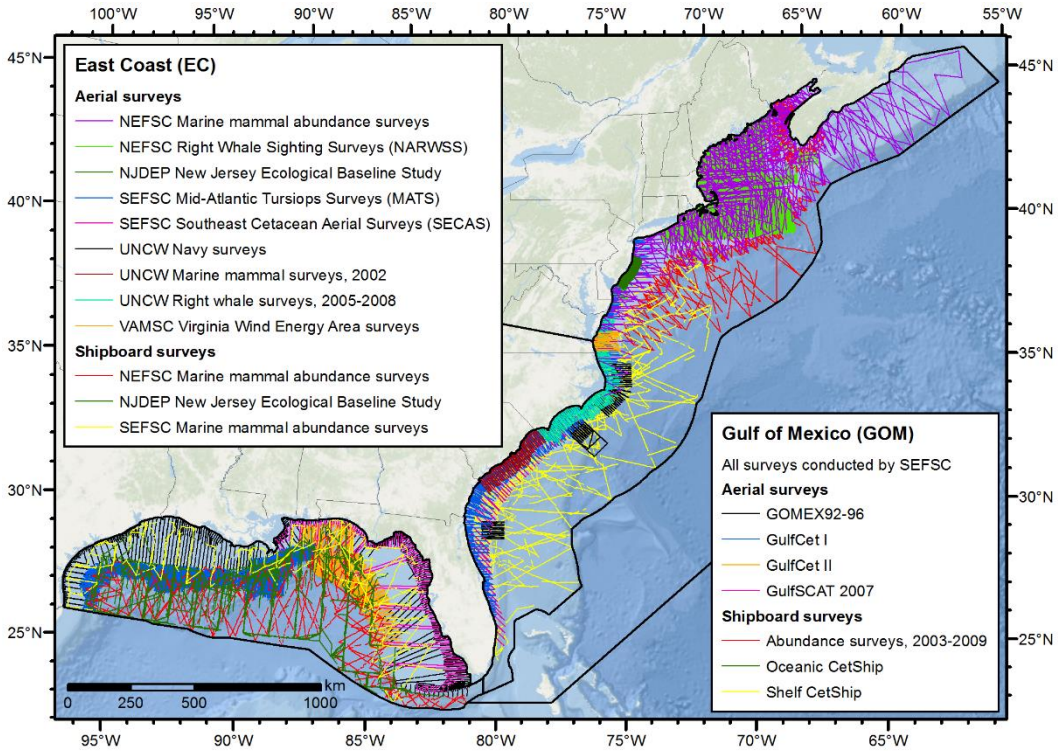
SCIENTIFIC REPORTS

OPEN

Habitat-based cetacean density models for the U.S. Atlantic and Gulf of Mexico

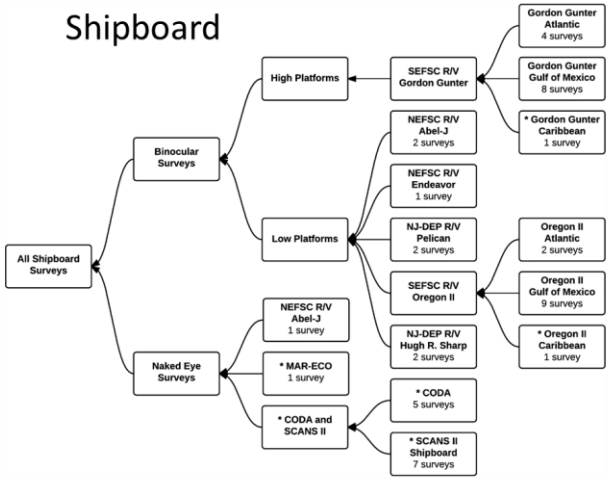
Received: 18 November 2015
Accepted: 17 February 2016
Published: 03 March 2016

Jason J. Roberts[1], Benjamin D. Best[1,2], Laura Mannocci[1], Ei Fujioka[1], Patrick N. Halpin[1], Debra L. Palka[3], Lance P. Garrison[4], Keith D. Mullin[5], Timothy V. N. Cole[3], Christin B. Khan[3], William A. McLellan[6], D. Ann Pabst[6] & Gwen G. Lockhart[7]

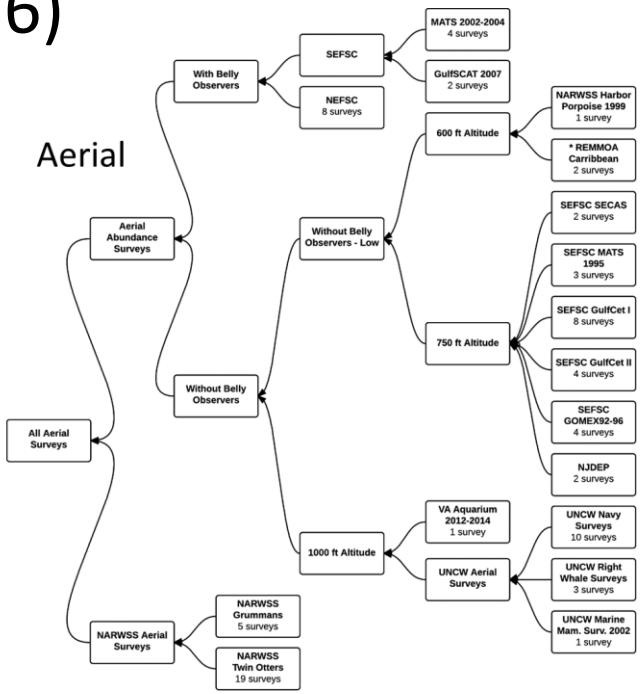


Detection hierarchies

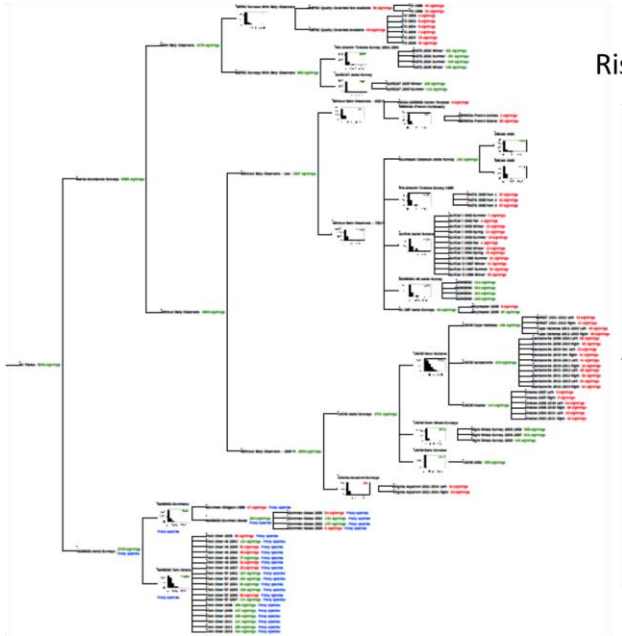
Shipboard



Aerial

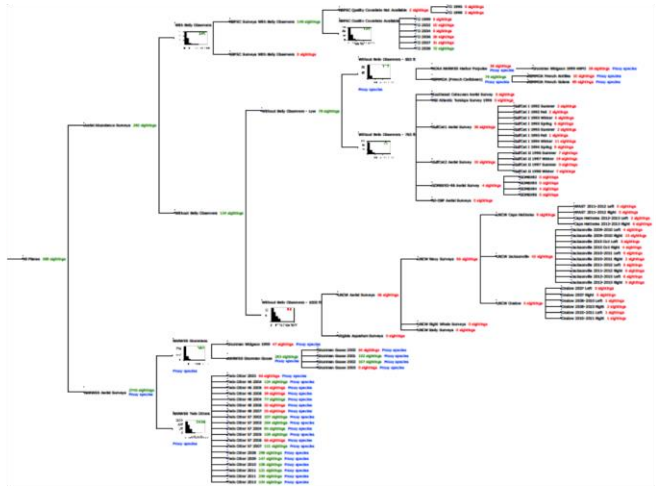


Bottlenose dolphin: 5630 sightings, 17 functions

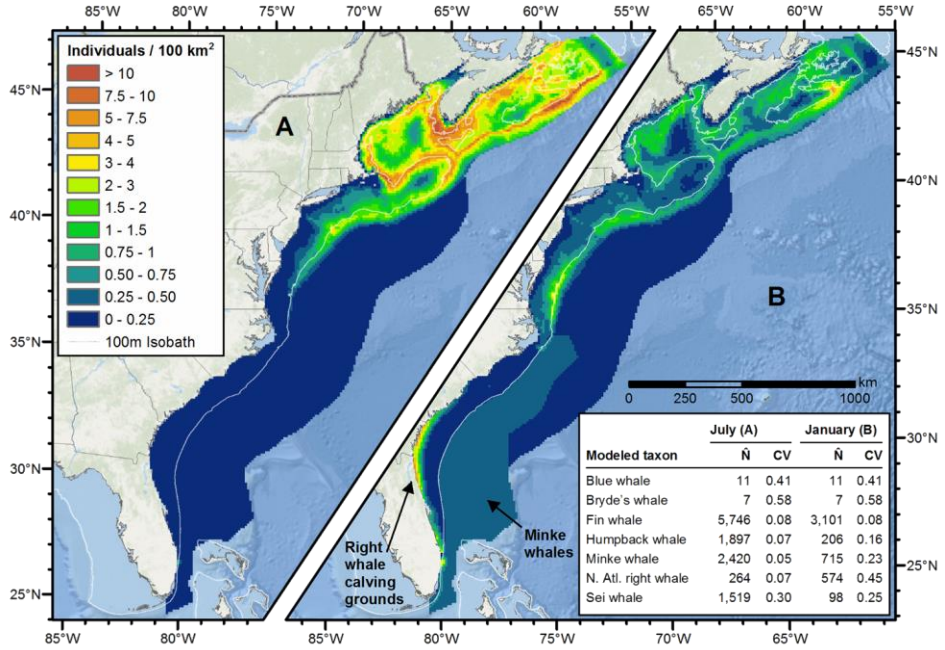


Examples: aerial surveys

Risso's dolphin: 368 sightings, 7 detection functions

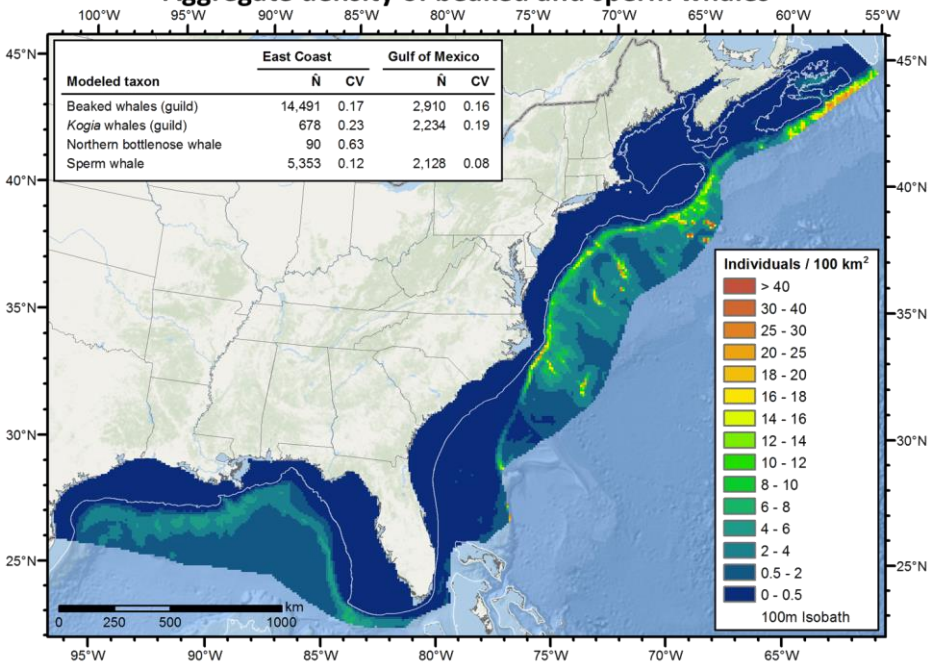


Aggregate density of baleen whales



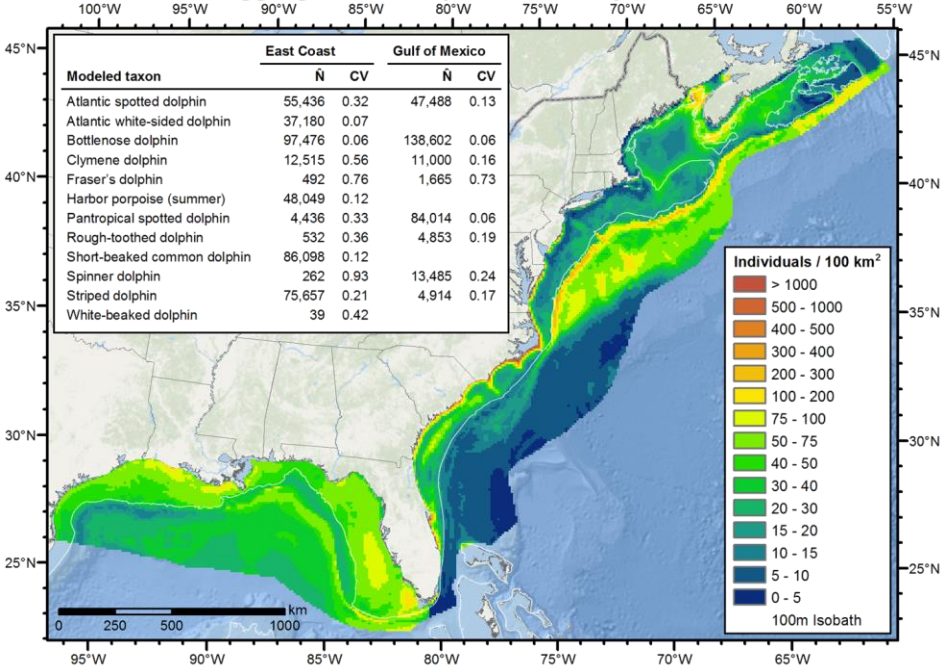
Modeled taxon	July (A) Ñ	July (A) CV	January (B) Ñ	January (B) CV
Blue whale	11	0.41	11	0.41
Bryde's whale	7	0.58	7	0.58
Fin whale	5,746	0.08	3,101	0.08
Humpback whale	1,897	0.07	206	0.16
Minke whale	2,420	0.05	715	0.23
N. Atl. right whale	264	0.07	574	0.45
Sei whale	1,519	0.30	98	0.25

Aggregate density of beaked and sperm whales



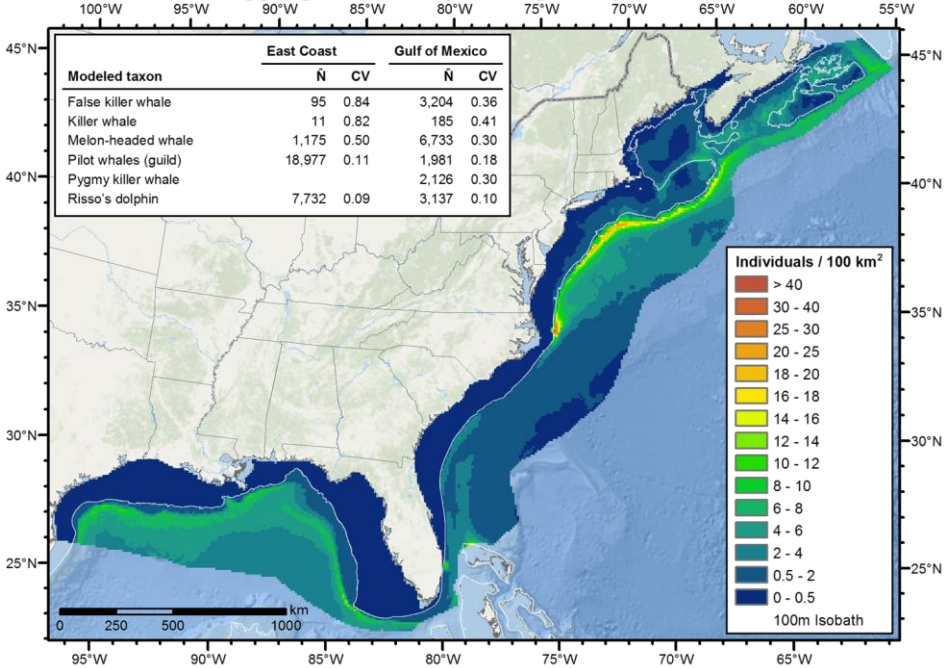
Modeled taxon	East Coast Ñ	East Coast CV	Gulf of Mexico Ñ	Gulf of Mexico CV
Beaked whales (guild)	14,491	0.17	2,910	0.16
Kogia whales (guild)	678	0.23	2,234	0.19
Northern bottlenose whale	90	0.63		
Sperm whale	5,353	0.12	2,128	0.08

Aggregate density of small delphinoids



Modeled taxon	East Coast Ñ	East Coast CV	Gulf of Mexico Ñ	Gulf of Mexico CV
Atlantic spotted dolphin	55,436	0.32	47,488	0.13
Atlantic white-sided dolphin	37,180	0.07		
Bottlenose dolphin	97,476	0.06	138,602	0.06
Clymene dolphin	12,515	0.56	11,000	0.16
Fraser's dolphin	492	0.76	1,665	0.73
Harbor porpoise (summer)	48,049	0.12		
Pantropical spotted dolphin	4,436	0.33	84,014	0.06
Rough-toothed dolphin	532	0.36	4,853	0.19
Short-beaked common dolphin	86,098	0.12		
Spinner dolphin	262	0.93	13,485	0.24
Striped dolphin	75,657	0.21	4,914	0.17
White-beaked dolphin	39	0.42		

Aggregate density of large delphinoids



Modeled taxon	East Coast Ñ	East Coast CV	Gulf of Mexico Ñ	Gulf of Mexico CV
False killer whale	95	0.84	3,204	0.36
Killer whale	11	0.82	185	0.41
Melon-headed whale	1,175	0.50	6,733	0.30
Pilot whales (guild)	18,977	0.11	1,981	0.18
Pygmy killer whale			2,126	0.30
Risso's dolphin	7,732	0.09	3,137	0.10

Taxonomic coverage

- 26 single-species models
- 3 guild models

Seasonal coverage

- 11 taxa predicted at monthly resolution
- 18 at year-round resolution

Virgili et al. (2019)

Received: 20 February 2018 | Revised: 27 August 2018 | Accepted: 19 September 2018

DOI: 10.1111/geb.12850

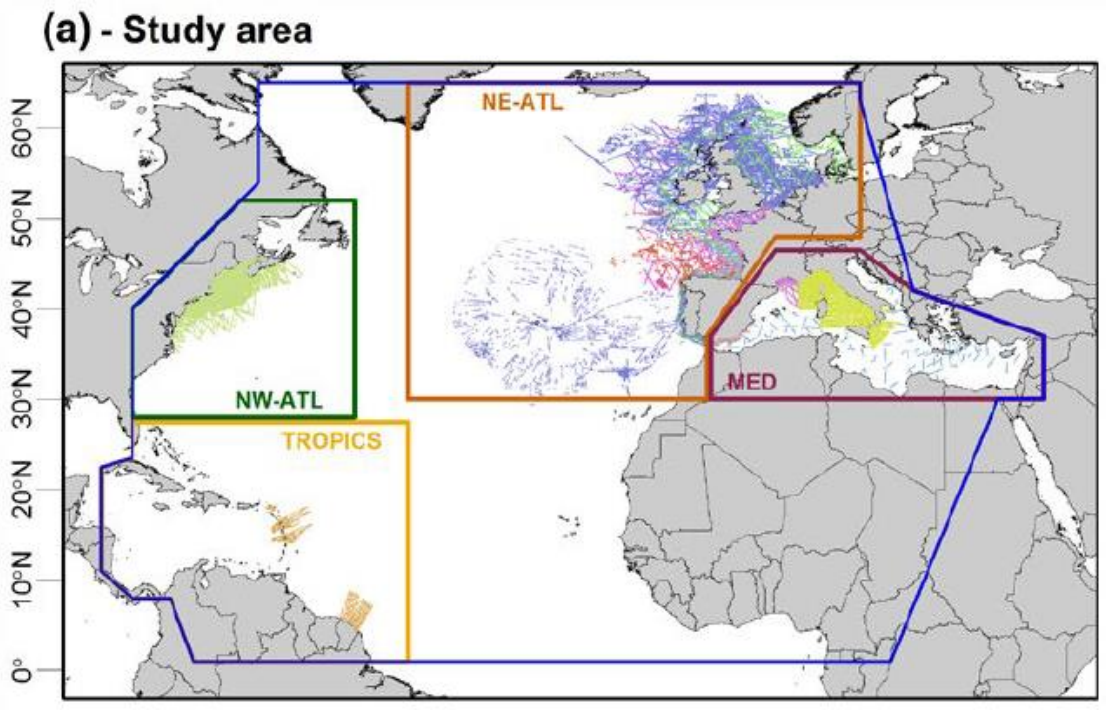
RESEARCH PAPER

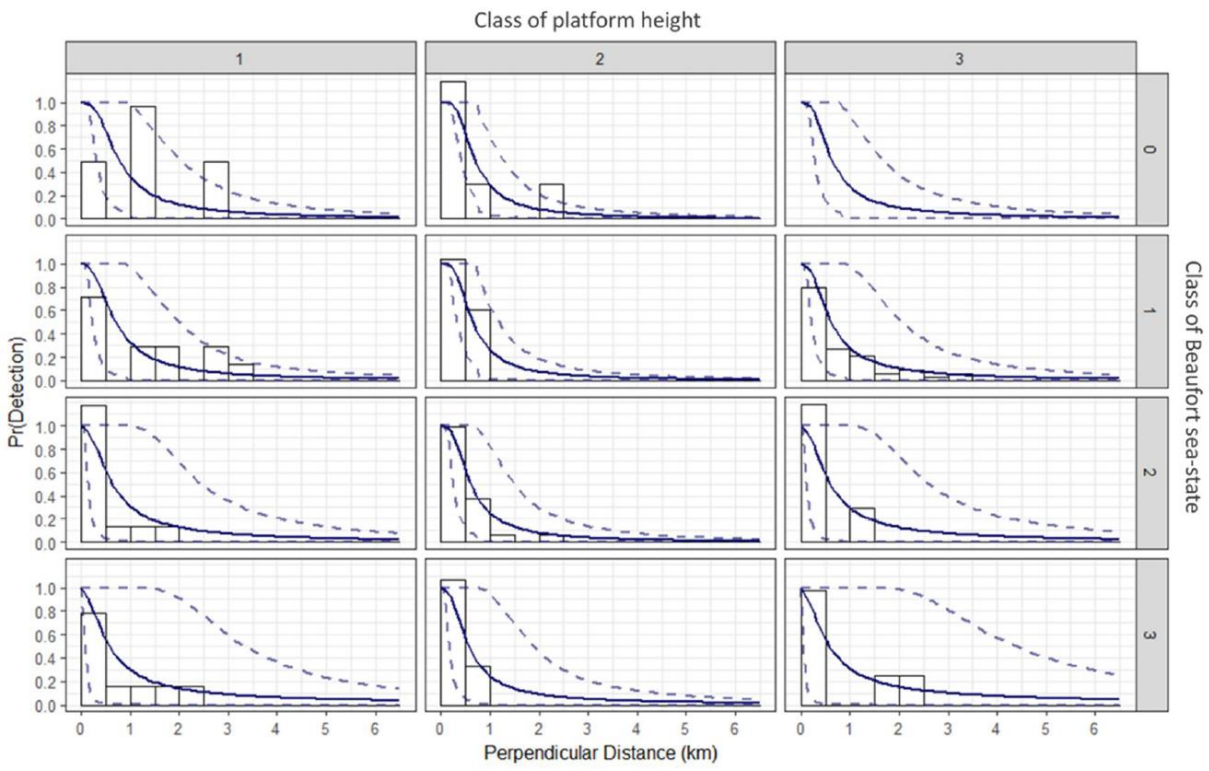
WILEY Global Ecology and Biogeography A Journal of Macroecology

Combining multiple visual surveys to model the habitat of deep-diving cetaceans at the basin scale

Large-scale modelling of deep-diving cetacean habitats

Auriane Virgili[1] | Matthieu Authier[2] | Oliver Boisseau[3,4] | Ana Cañadas[5] | Diane Claridge[6] | Tim Cole[7] | Peter Corkeron[7] | Ghislain Dorémus[2] | Léa David[8] | Nathalie Di-Méglio[8] | Charlotte Dunn[6] | Tim E. Dunn[9] | Isabel García-Barón[10] | Sophie Laran[2] | Giancarlo Lauriano[11] | Mark Lewis[9] | Maite Louzao[10] | Laura Mannocci[12] | José Martínez-Cedeira[13] | Debra Palka[7] | Simone Panigada[14] | Emeline Pettex[2] | Jason J. Roberts[12] | Leire Ruiz[15] | Camilo Saavedra[16] | M. Begoña Santos[16] | Olivier Van Canneyt[2] | José Antonio Vázquez Bonales[5] | Pascal Monestiez[1,17] | Vincent Ridoux[1,2]





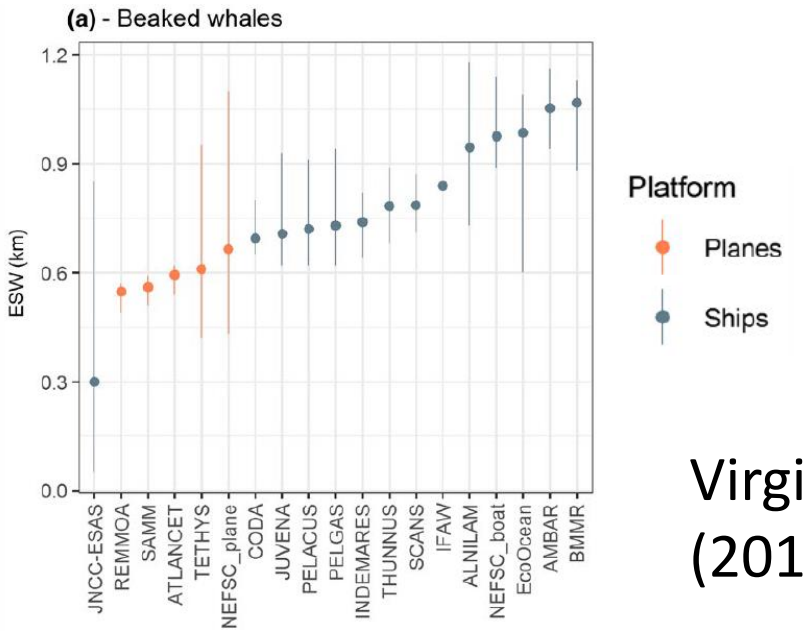
Hierarchical Bayesian model for hazard rate detection:

$$\mathrm{ESW}_{jks} = \int_0^w g_s(x)\,dx = \int_0^w \left[1 - \exp\left(-\left(\frac{x}{e^{\beta_{j0}+\beta_{j1}\times k+\alpha_s}}\right)^{-\nu_s}\right)\right] dx$$

s = survey
j = platform height
k = sea-state

Effects of platform height

Survey random effects



Virgili et al. (2019)

Recent research trends

- Integrating multiple surveys to facilitate models with greater spatial, temporal, and taxonomic coverage
- Extrapolating models across unsurveyed areas, times, and the future
- Incorporating multiple sources of uncertainty into density surface predictions

Recent research trends

- Integrating multiple surveys to facilitate models with greater spatial, temporal, and taxonomic coverage
- Extrapolating models across unsurveyed areas, times, and the future
- Incorporating multiple sources of uncertainty into density surface predictions

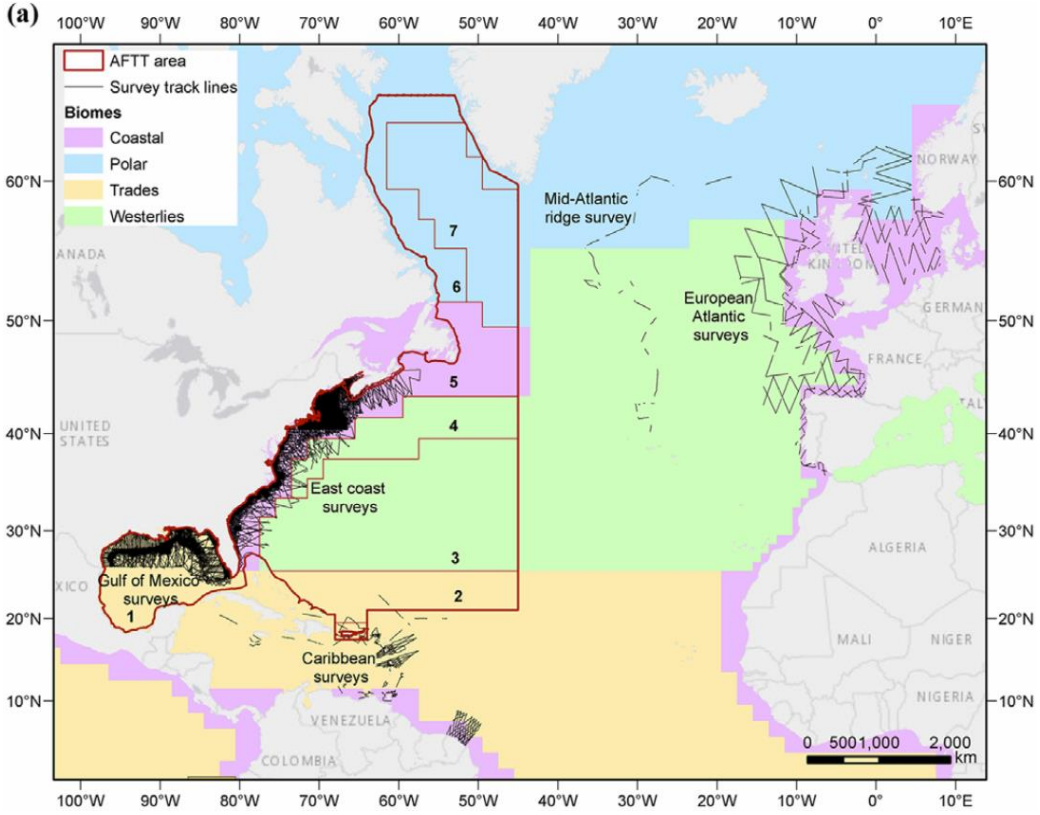
Conservation Biology

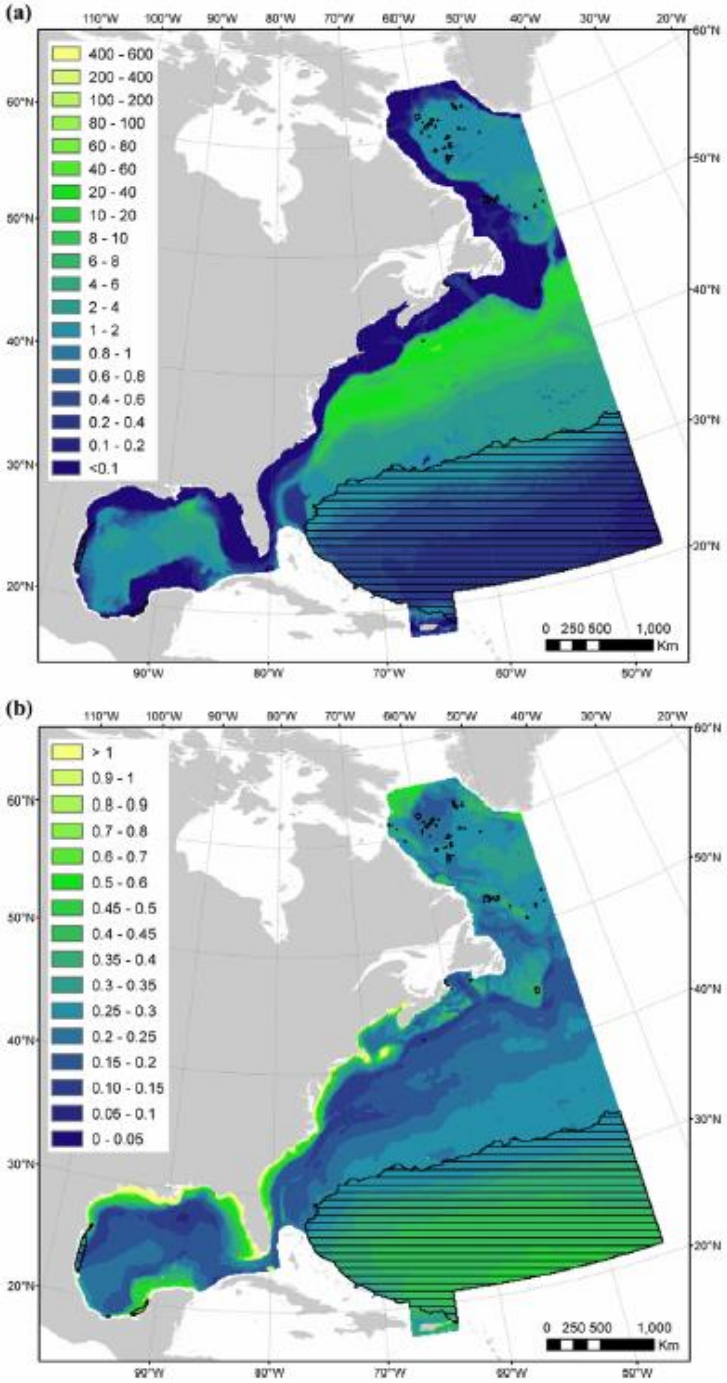
Contributed Paper

Extrapolating cetacean densities to quantitatively assess human impacts on populations in the high seas

Laura Mannocci,[1]* Jason J. Roberts,[1] David L. Miller,[2,3,4] and Patrick N. Halpin[1]

[1]Marine Geospatial Ecology Laboratory, Nicholas School of the Environment, Duke University, Durham, NC 27708, U.S.A.
[2]Integrated Statistics, 16 Sumner Street, Woods Hole, MA 02543, U.S.A.
[3]Centre for Research into Ecological and Environmental Modelling and School of Mathematics and Statistics, University of St Andrews, The Observatory, Buchanan Gardens, St Andrews, Fife KY16 9LZ, Scotland
[4]Woods Hole Oceanographic Institution, Woods Hole, MA 02543, U.S.A.





Mannocci et al. (2017)

Figure 4. Maps of (a) mean year-round predicted densities (individuals per 100 km^2) and (b) mean year-round coefficients of variation from the spatial part of the density model for striped dolphin (black horizontal lines, extrapolations beyond the predictor ranges [these predicted densities should be considered with extreme caution]). An Albers equal area projection is used for the study area.

Virgili et al. (2019)

Received: 20 February 2018 | Revised: 27 August 2018 | Accepted: 19 September 2018

DOI: 10.1111/geb.12850

RESEARCH PAPER

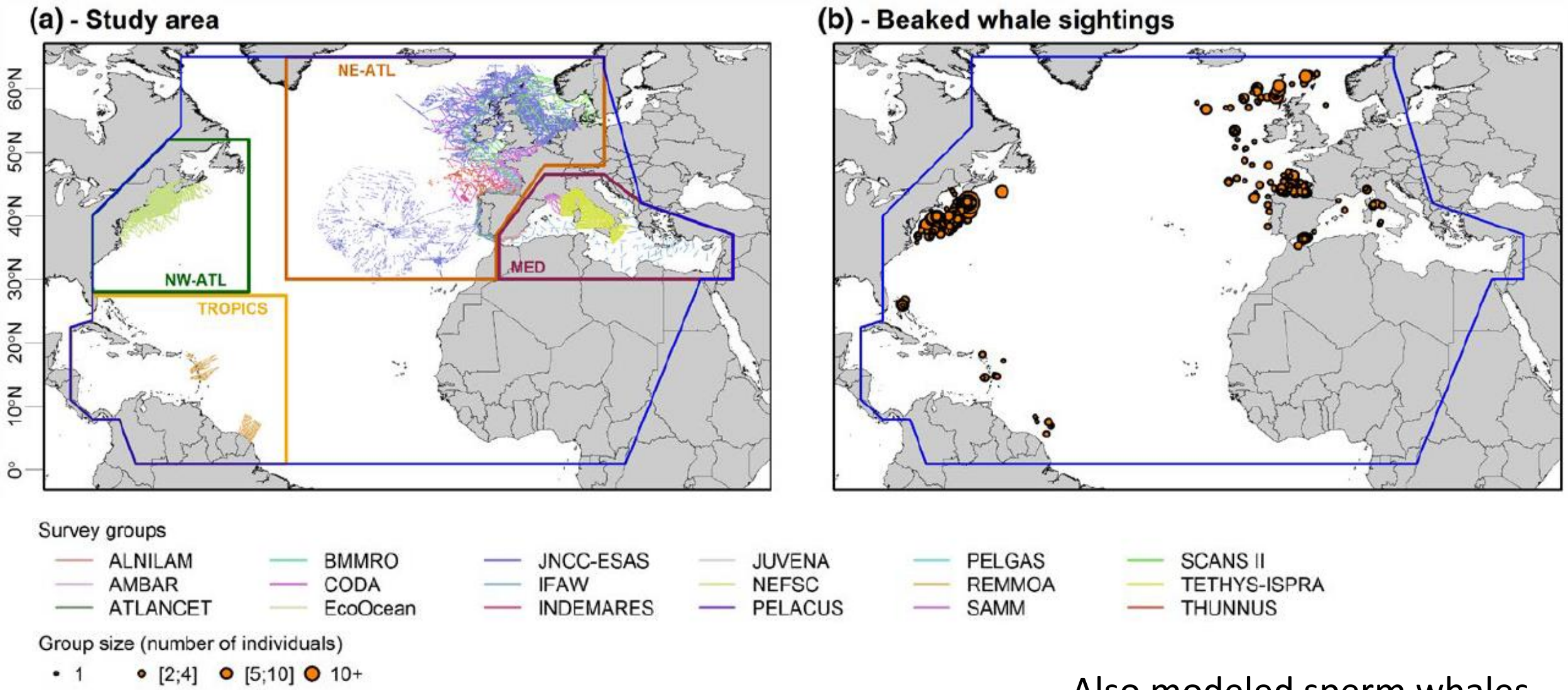
WILEY Global Ecology and Biogeography A Journal of Macroecology

Combining multiple visual surveys to model the habitat of deep-diving cetaceans at the basin scale

Large-scale modelling of deep-diving cetacean habitats

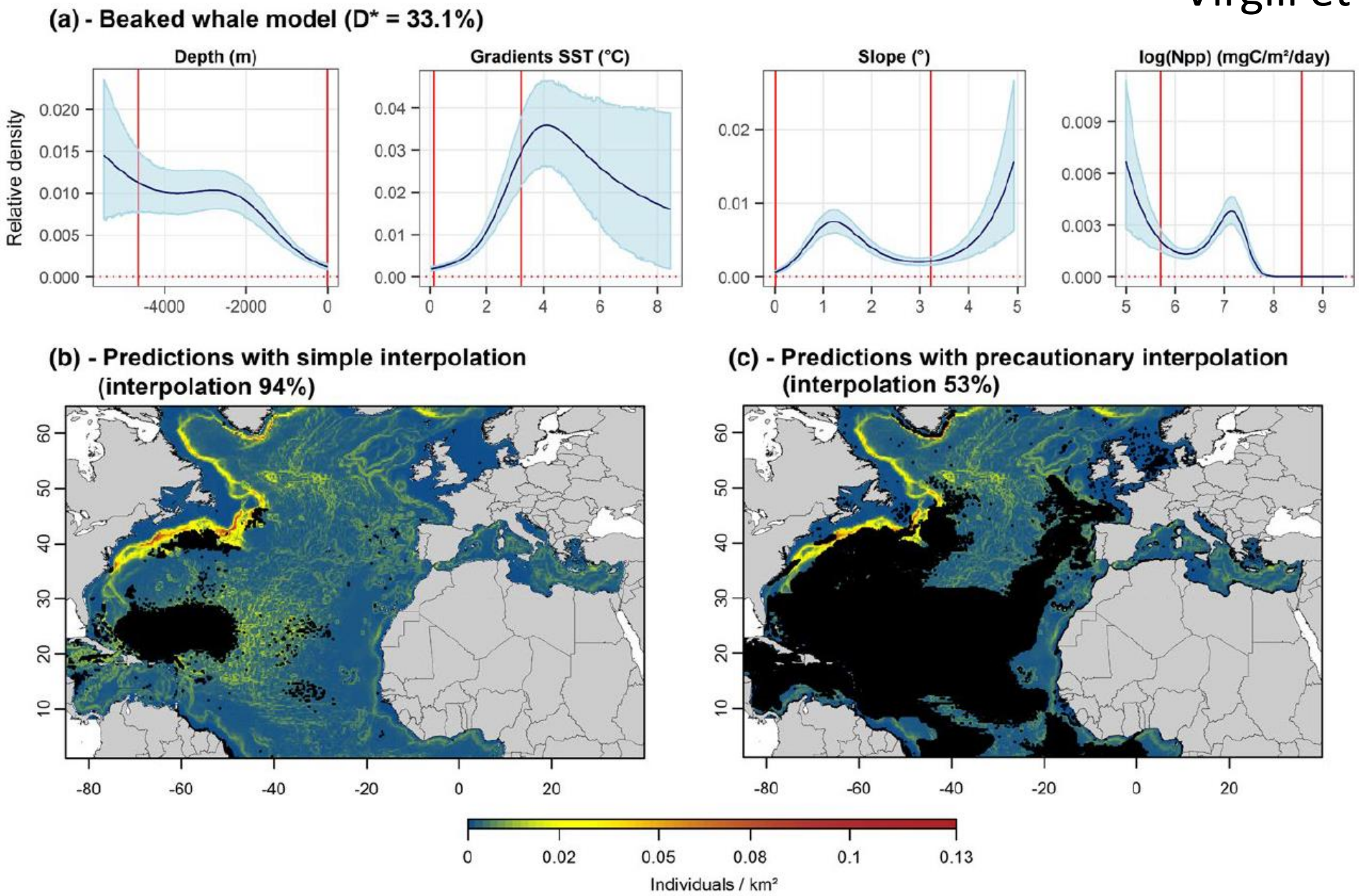
Auriane Virgili[1] | Matthieu Authier[2] | Oliver Boisseau[3,4] | Ana Cañadas[5] | Diane Claridge[6] | Tim Cole[7] | Peter Corkeron[7] | Ghislain Dorémus[2] | Léa David[8] | Nathalie Di-Méglio[8] | Charlotte Dunn[6] | Tim E. Dunn[9] | Isabel García-Barón[10] | Sophie Laran[2] | Giancarlo Lauriano[11] | Mark Lewis[9] | Maite Louzao[10] | Laura Mannocci[12] | José Martínez-Cedeira[13] | Debra Palka[7] | Simone Panigada[14] | Emeline Pettex[2] | Jason J. Roberts[12] | Leire Ruiz[15] | Camilo Saavedra[16] | M. Begoña Santos[16] | Olivier Van Canneyt[2] | José Antonio Vázquez Bonales[5] | Pascal Monestiez[1,17] | Vincent Ridoux[1,2]

Virgili et al. (2019)



Also modeled sperm whales and *Kogia* (results not shown)

Virgili et al. (2019)



Received: 5 May 2018 | Revised: 28 September 2018 | Accepted: 5 October 2018

DOI: 10.1111/ddi.12867

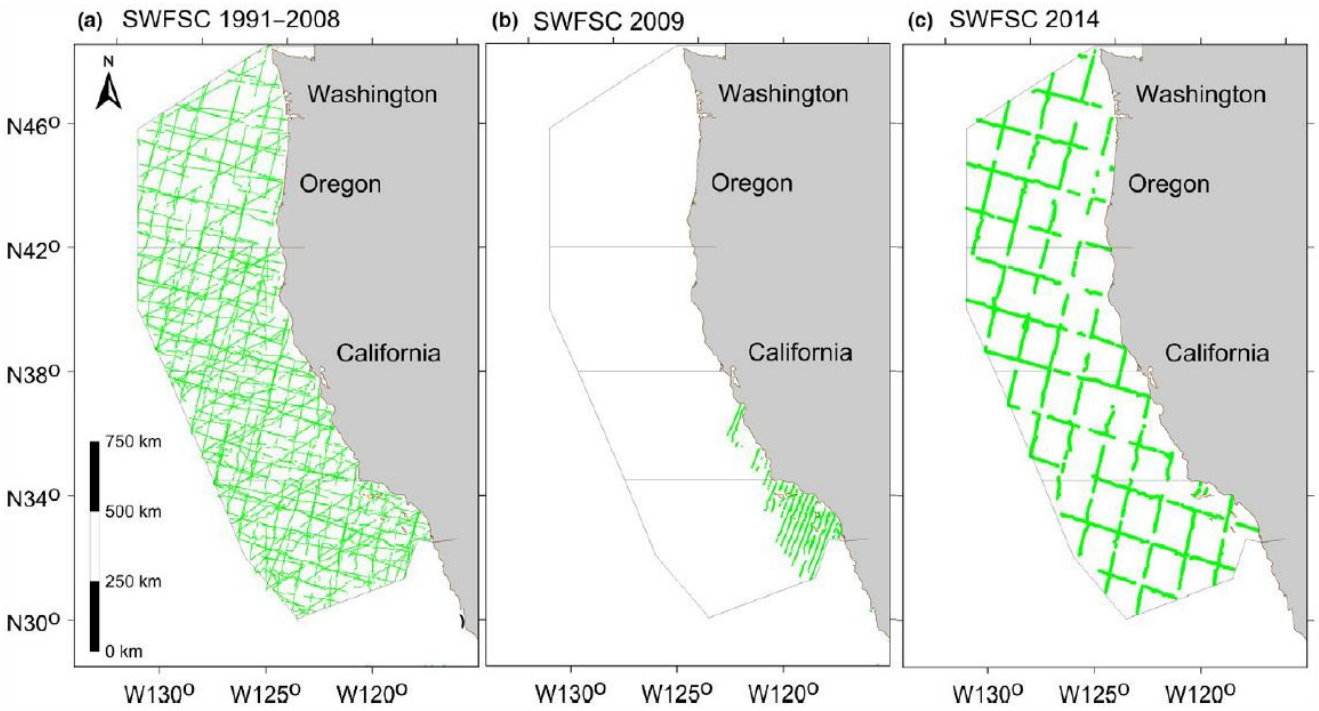
Becker et al. (2019)

BIODIVERSITY RESEARCH

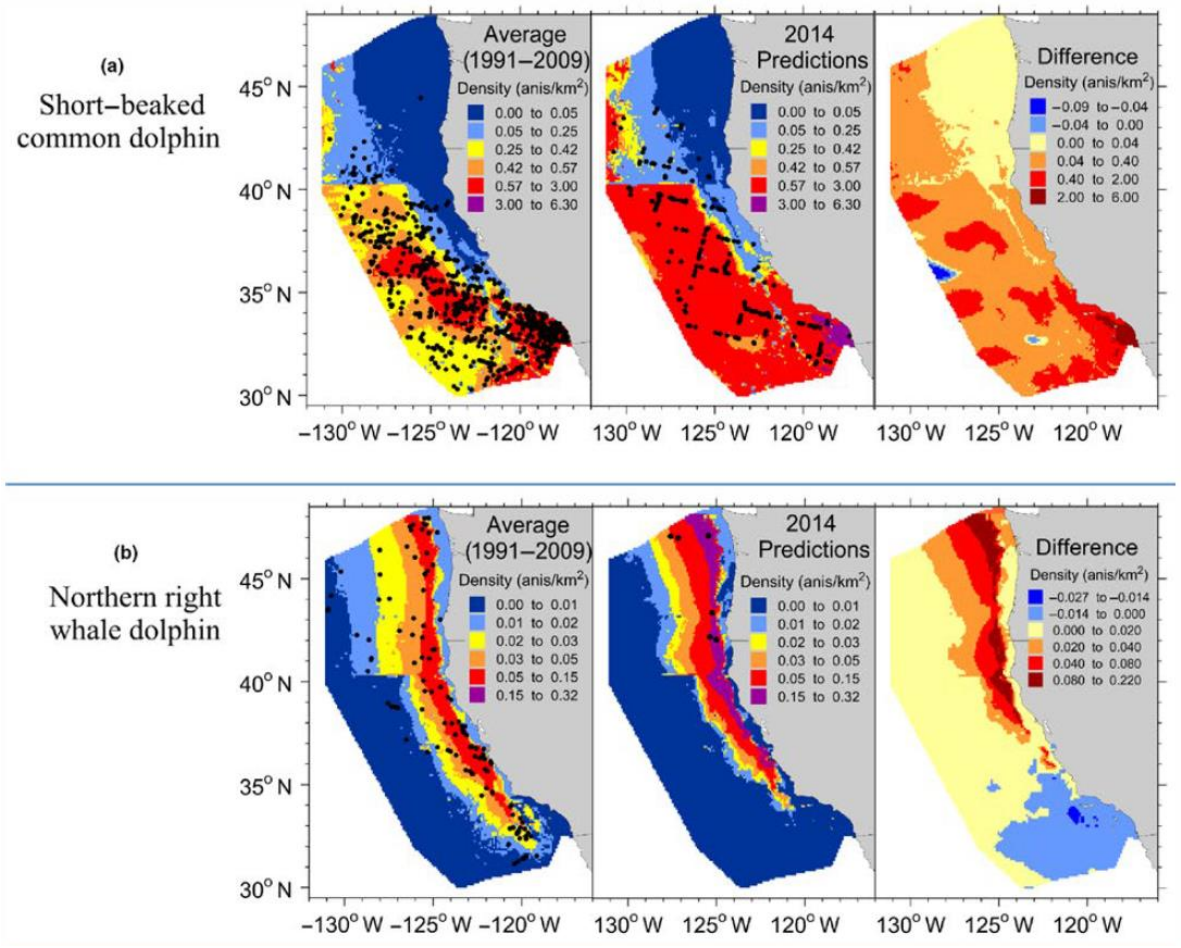
WILEY Diversity and Distributions

Predicting cetacean abundance and distribution in a changing climate

Elizabeth A. Becker[1,2] | Karin A. Forney[3,4] | Jessica V. Redfern[1] | Jay Barlow[1] | Michael G. Jacox[5,6] | Jason J. Roberts[7] | Daniel M. Palacios[8]



Example results for 2 species (8 modeled in total):



Extrapolation tools developed by Phil Bouchet et al.

dsmextra v 1.0.0

Extrapolation diagnosis in R

- New R package available freely on GitHub
- Functions for detecting, summarising and visualising extrapolation (ExDet + %N)
- Designed to easily integrate with outputs generated by the dsm package
- Step-by-step vignette available, with practical examples on sperm/beaked whale data
- More features to come …

dsmextra 1.0.0 | Reference | Articles

Extrapolation tools for density surface models

`dsmextra` provides a toolkit for quantifying and visualising extrapolation in density surface models (as implemented in package dsm) projected into novel environmental space. Currently, `dsmextra` defines extrapolation on the basis of two metrics: **(1) ExDet** (Mesgaran et al. 2014), and **(2) %N** (percentage of data nearby, Mannocci et al. 2018).

`dsmextra` offers a variety of numerical and graphical outputs, including summary plots and interactive maps created as ggplot2 and html objects, respectively. Additional functionality (e.g. assessment methods for dynamic covariates) will be added in future releases.

The idea behind `dsmextra` is to aid ecologists, practitioners, and model end-users in identifying conditions (e.g. areas) under which predicted density surfaces may be prone to errors. In so doing, `dsmextra` may support:

- Better-informed interpretations of density surface model outputs and their associated uncertainties.
- Improvements to model development and selection protocols.
- Cost-effective allocation of future survey effort towards priority, data-poor areas.

GPL-3 + file LICENSE

Phil J Bouchet
Author, maintainer

David L Miller
Author

Laura Mannocci
Author

Jason J Roberts
Contributor

Catriona M Harris
Contributor

Len Thomas
Contributor

lifecycle maturing

DOI 10.5281.zenodo.3529465

Getting started

If you are just getting started with `dsmextra`, we recommend reading the tutorial vignette, which provides a quick introduction to the package. You may also find the below technical report useful:

Bouchet et al. (2019). From here and now to there and then: Practical recommendations for extrapolating cetacean density surface models to novel conditions. CREEM technical report 2019-01, Centre for Research into Ecological & Environmental Modelling (CREEM), University of St Andrews, 59 p.

https://densitymodelling.github.io/dsmextra/

Recent research trends

- Integrating multiple surveys to facilitate models with greater spatial, temporal, and taxonomic coverage
- Extrapolating models across unsurveyed areas, times, and the future
- Incorporating multiple sources of uncertainty into density surface predictions

Recent research trends

- Integrating multiple surveys to facilitate models with greater spatial, temporal, and taxonomic coverage
- Extrapolating models across unsurveyed areas, times, and the future
- Incorporating multiple sources of uncertainty into density surface predictions

Variance in density surface models

DSM equation

$$n_i = \exp\left(\log_e A_i \hat{p}_i + \beta_0 + \underbrace{f_{\text{space}}(\texttt{lat}_i, \texttt{lon}_i)}_{\text{smooth of space}} + \underbrace{f_{\text{depth}}(\texttt{Depth}_i)}_{\text{smooth of depth}}\right)$$

log link

detectability in segment *i*

count in segment *i*

area of segment *i*

intercept

Multiple sources of variance

$$n_i = \exp\left(\log_e A_i \hat{p}_i + \underbrace{\beta_0 + f_{\text{space}}(\texttt{lat}_i, \texttt{lon}_i) + f_{\text{depth}}(\texttt{Depth}_i)}_{\text{covariance from GLM/GAM theory, } \boldsymbol{V}_{\hat{\boldsymbol{\beta}}}}\right)$$

covariance from distance sampling, $\mathbf{V}_{\boldsymbol{\theta}}$

How to combine variances?

1. Use the “delta method” (Buckland et al. 2001, p. 52)

$$\left\{\mathrm{CV}\left(\hat{N}\right)\right\}^2 = \left\{\mathrm{CV}\left(\hat{N}_{\mathrm{GAM}}\right)\right\}^2 + \left\{\mathrm{CV}\left(\hat{N}_{\mathrm{DS}}\right)\right\}^2$$

Drawbacks:

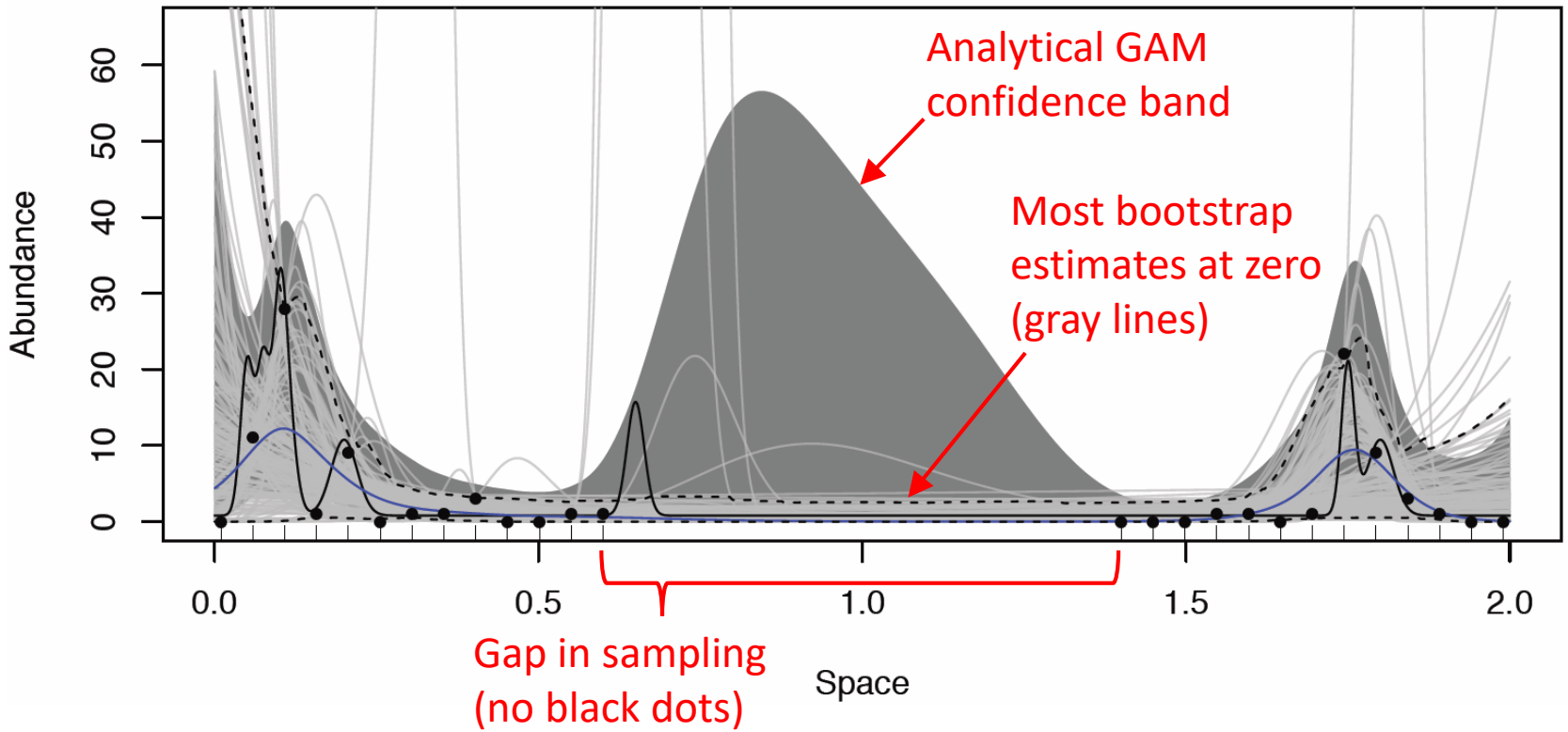
- Assumes independence: does not account for covariance between the GAM and the detection function
- There can be only 1 detection function

How to combine variances?

2. Use a moving block bootstrap

Drawbacks:

- Ignores prior information about the GAM smooths
- Assumes covariates have been sampled throughout their ranges, with no gaps



How to combine variances?

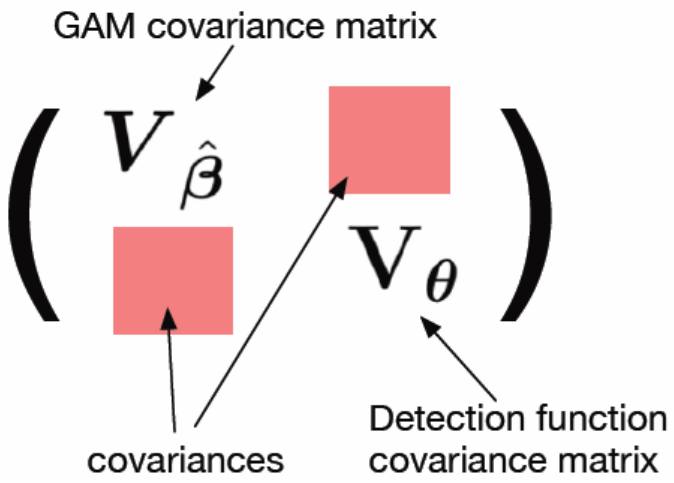
3. Propagate variances (Bravington, Miller, and Hedley; in review)

Going back to the DSM formula (on the log scale) and "adding" a random effect, then refitting the model

$$\log_e n_i = \log_e A_i \hat{p}_i + \beta_0 + f_{\text{space}}(\texttt{lat}_i, \texttt{lon}_i) + f_{\text{depth}}(\texttt{Depth}_i) + \kappa_i \boldsymbol{\delta}$$

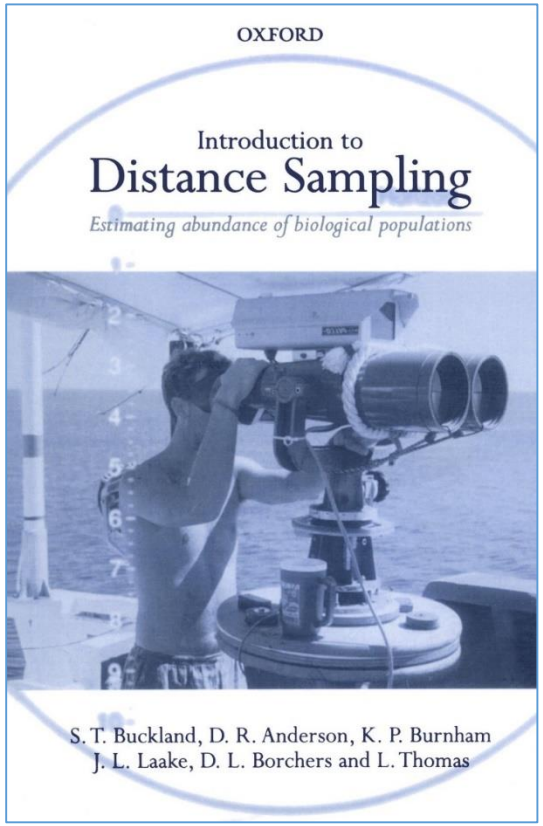
Random effect adds the missing covariances but does not add to the mean

In the refitted model, we have adjusted the detectability and have a combined variance-covariance matrix:

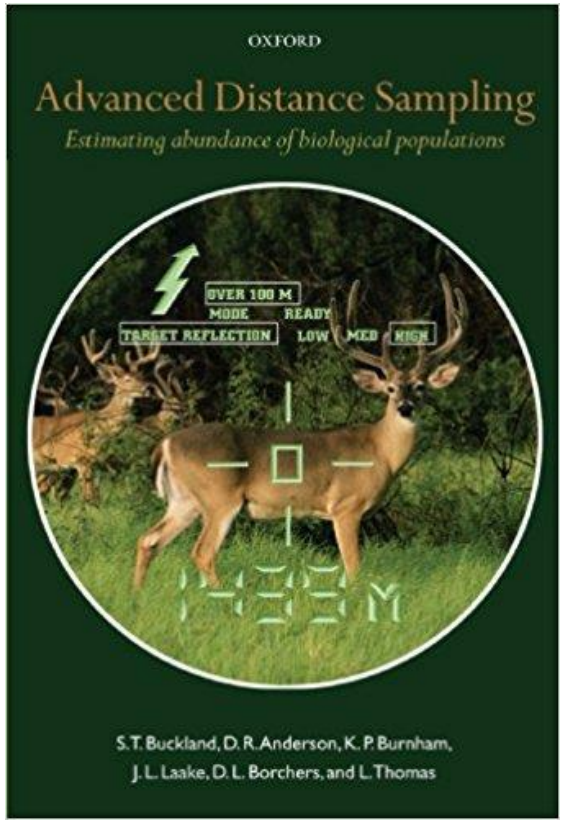


Density Surface Modeling Resources

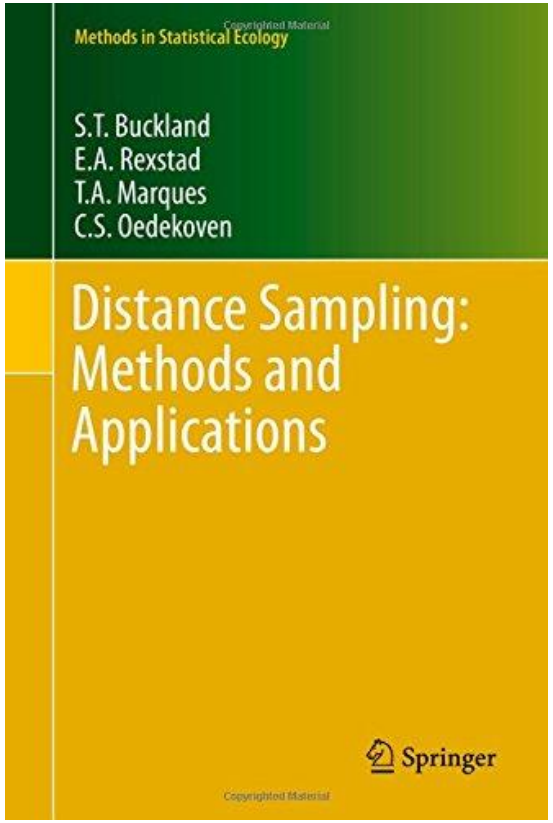
- Reference books



Buckland et al. (2001)



Buckland et al. (2004)



Buckland et al. (2015)

Density Surface Modeling Resources

- Reference books
- https://distancesampling.org/
  - Free software: program *Distance* and R packages
  - Upcoming and archived workshops
  - Free online course (26 lectures, 12 exercises)
  - Mailing list and searchable archive (Google Groups)
  - Bibliography (1500 references)
- FAQ for marine mammal density surface modelers: https://osf.io/5eza8/wiki/

- Project Wiki Pages
  - Home
- Component Wiki Pages
  - DSM Workflows
  - Surveys and Survey Data
  - Detection Models
  - Spatial Models

Questions Frequently Asked by Marine Mammal Density Surface Modelers

In 2018, the U.S. Navy Living Marine Resources (LMR) program initiated the "DenMod" collaboration between us--the Centre for Research into Ecological and Environmental Modelling (CREEM) at the University of St. Andrews, the Marine Geospatial Ecology Laboratory at Duke University, and the NOAA Fisheries Science Centers--to develop best practices for density surface modeling of marine species and provide practical guidance to the modeling community, particularly those involved in modeling marine mammals for the Navy's training and testing areas. Pursuant to this goal, we developed this collection of answers to technical modeling questions frequently asked by ourselves and external colleagues. Through this corpus, we offer ready solutions to common modeling problems, distilled from our experiences developing distance sampling and density surface modeling methodology, and applying it in the development of marine mammal density models for the Navy.

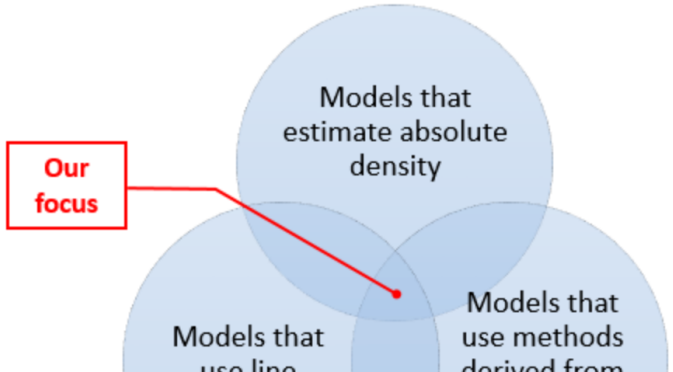
How the questions are organized

The questions are organized into wiki pages, with one or more related questions occurring in sequence on each page. The pages are grouped into several "components". You can drill into the navigation hierarchy on the left or click the component links below:

- DSM Workflows
- Surveys and Survey Data
- Detection Models
- Spatial Models

Scope of the questions

Most of the practitioners within our collaboration have focused on modeling the absolute abundance (number of individuals) and density (individuals per unit area) of marine mammals, usually over large expanses (100s-1000s km) of marine waters of the United States, typically by applying methods derived from distance sampling to visual line transect surveys conducted from ships and aircraft. The questions that appear here reflect this focus.



Home

- Home
- Covariates: steps and tools for acquiring
- Covariates: transformations and outliers
- Covariates: which should I use?
- mgcv: diagnostics
- mgcv: distributions
- mgcv: model selection
- mgcv: soap film smoother
- Predictions: summarizing
- Predictions: validating
- Splitting transects into segments
- Unidentified and ambiguous sightings

View

Wiki Version: (Current) Jason Roberts: 2019-12-06 06:50:53+00:00 UTC

Spatial Models

This wiki contains questions related to the second stage of traditional density surface modeling, which is the spatial model. Typically, this is a generalized additive model that relates abundance observed on segments of a line transect survey to covariates, which often include an explicit spatial term (a 2D smooth of x and y coodinates) and environmental variables believed to relate to the habitat of the modeled taxon. The kinds of questions that appear here reflect this approach.

The questions and their answers are grouped by topic. Select a topic below or from the alphabetical list in the left navigation bar.

- Splitting transects into segments
- Covariates: which should I use?
- Covariates: steps and tools for acquiring
- Covariates: transformations and outliers
- mgcv: distributions
- mgcv: model selection
- mgcv: diagnostics
- mgcv: soap film smoother
- Predictions: summarizing
- Predictions: validating
- Unidentified and ambiguous sightings

Thank you!

DSM FAQ:
https://osf.io/5eza8/wiki/

Ziphius cavirostris