Density Model for Common Minke Whale (*Balaenoptera acutorostrata*) for the U.S. East Coast: Supplementary Report

Model Version 10.1

Duke University Marine Geospatial Ecology Laboratory[*]

2023-05-27

Citation

When citing our methodology or results generally, please cite Roberts et al. (2016, 2023). The complete references appear at the end of this document. We are preparing a new article for a peer-reviewed journal that will eventually replace those. Until that is published, those are the best general citations.

When citing this model specifically, please use this reference:

Roberts JJ, Yack TM, Cañadas A, Fujioka E, Halpin PN, Barco SG, Boisseau O, Chavez-Rosales S, Cole TVN, Cotter MP, Cummings EW, Davis GE, DiGiovanni Jr. RA, Garrison LP, Gowan TA, Jackson KA, Kenney RD, Khan CB, Lockhart GG, Lomac-MacNair KS, McAlarney RJ, McLellan WA, Mullin KD, Nowacek DP, O'Brien O, Pabst DA, Palka DL, Quintana-Rizzo E, Redfern JV, Rickard ME, White M, Whitt AD, Zoidis AM (2022) Density Model for Common Minke Whale (*Balaenoptera acutorostrata*) for the U.S. East Coast, Version 10.1, 2023-05-27, and Supplementary Report. Marine Geospatial Ecology Laboratory, Duke University, Durham, North Carolina.

Copyright and License

Model Version History

Version	Date	Description
1	2013-05-07	Initial version.
2	2013-05-08	Text edited to correct minor errors.

[*]For questions or to offer feedback please contact Jason Roberts (jason.roberts@duke.edu) and Tina Yack (tina.yack@duke.edu)

Version	Date	Description
3	2014-03-01	Switched from four seasonal models to two. Reformulated density model using a Horvitz-Thompson estimator. Eliminated GAM for group size (consequence of above). Added group size as a candidate covariate in detection functions (benefit of above). Added survey ID as a candidate covariate in NOAA NARWSS detection functions. Took more care in selecting right-truncation distances. Fitted models with contemporaneous predictors, for comparison to climatological. Switched SST and SST fronts predictors from NOAA Pathfinder to GHRSST CMC0.2deg L4. Changed SST fronts algorithm to use Canny operator instead of Cayula-Cornillon. Switched winds predictors from SCOW to CCMP (SCOW only gives climatol. estimates.) Added DistToEddy predictors, based on Chelton et al. (2011) eddy database. Added cumulative VGPM predictors, summing productivity for 45, 90, and 180 days. Added North Atlantic Oscillation (NAO) predictor; included 3 and 6 month lags. Transformed predictors more carefully, to better minimize leverage of outliers. Implemented hybrid hierarchical-forward / exhaustive model selection procedure. Model selection procedure better avoids concurvity between predictors. Allowed GAMs to select between multiple formulations of dynamic predictors. Adjusted land mask to eliminate additional estuaries and hard-to-predict cells.
4	2014-05-14	Added discussion of acoustic monitoring studies to text. Density models unchanged.
5	2014-05-20	Fixed bug in temporal variability plots. Density models unchanged.
6	2014-10-18	Added surveys: NJ-DEP, Virginia Aquarium, NARWSS 2013, UNCW 2013. Extended study area up Scotian Shelf. Added SEAPODYM predictors. Switched to mgcv estimation of Tweedie p parameter (family=tw()). Added Palka (2006) survey-specific g(0) estimates. Removed distance to eddy predictors and wind speed predictor from all models; they were not ecologically justified. Fixed missing pixels in several climatological predictors, which led to not all segments being utilized. Adjusted subregion extents. Eliminated Cape Cod Bay subregion.
7	2014-11-13	Reconfigured detection hierarchy and adjusted NARWSS detection functions based on additional information from Tim Cole. Switched to uniform distribution of density for southeast slope and abyss in winter. Removed CumVGPM180 predictor. Updated documentation.
8	2014-12-03	Fixed bug that applied the wrong detection function to segments NE_narwss_1999_widgeon_hapo dataset. Refitted model. Updated documentation.
8.1	2015-02-02	Updated the documentation. No changes to the model.
8.2	2015-05-14	Updated calculation of CVs. Switched density rasters to logarithmic breaks. No changes to the model.
8.3	2015-09-26	Updated the documentation. No changes to the model.
8.4	2016-04-21	Switched calculation of monthly 5% and 95% confidence interval rasters to the method used to produce the year-round rasters. (We intended this to happen in version 8.2 but I did not implement it properly.) Updated the monthly CV rasters to have value 0 where we assumed the species was absent, consistent with the year-round CV raster. No changes to the other (non-zero) CV values, the mean abundance rasters, or the model itself. Model files released as supplementary information to Roberts et al. (2016).
9	2017-06-01	Began update to Roberts et al. (2015) model. Introduced new surveys from AMAPPS, NARWSS, UNCW, and VAMSC. Updated modeling methodology. Refitted detection functions and spatial models from scratch using new and reprocessed covariates. Model released as part of a scheduled update to the U.S. Navy Marine Species Density Database (NMSDD).

Version	Date	Description
10	2022-06-20	This model is a major update over the prior version, with substantial additional data, improved statistical methods, and an increased spatial resolution. It was released as part of the final delivery of the U.S. Navy Marine Species Density Database (NMSDD) for the Atlantic Fleet Testing and Training (AFTT) Phase IV Environmental Impact Statement. Several new collaborators joined and contributed survey data: New York State Department of Environmental Conservation, TetraTech, HDR, and Marine Conservation Research. We incorporated additional surveys from all continuing and new collaborators through the end of 2020. (Because some environmental covariates were only available through 2019, certain models only extend through 2019.) We increased the spatial resolution to 5 km and, at NOAA's request, we extended the model further inshore from New York through Maine. We reformulated and refitted all detection functions and spatial models. We updated all enviromental covariates to newer products, when available, and added several covariates to the set of candidates. For models that incorporated dynamic covariates, we estimated model uncertainty using a new method that accounts for both model parameter error and temporal variability.
10.1	2023-05-27	Completed the supplementary report documenting the details of this model. Corrected the 5 and 95 percent rasters so that they contain the value 0 where the taxon was asssumed absent, rather than NoData. Nothing else was changed.

1 Survey Data

We built this model from data collected between 1998-2019 (Table 1, Figure 1). We excluded surveys that did not target minke whales or were otherwise problematic for modeling them. In keeping with our primary strategy for the 2022 modeling cycle, we excluded data prior to 1998 in order to utilize biological covariates derived from satellite ocean color observations, which were only available for a few months before 1998. We excluded data after 2019 in order to utilize zooplankton and micronekton biomass estimates from SEAPODYM (Lehodey et al. 2008), which preliminary modeling indicated were were effective spatial covariates but were only available through 2019. For the winter model for the region covering the Blake Plateau and off-shelf waters south of the Gulf Stream, we expanded the temporal range back to the winter of 1991/92 in order to include the wintertime NOAA SEFSC OT-92 shipboard survey, but no minke whales were sighted during it.

For all surveys, we restricted the model to survey transects with sea states of Beaufort 4 or less, except for aerial surveys conducted by New England Aquarium, for which we used Beaufort 2 or less, owing to a lack of sightings at higher sea states. For all surveys from all collaborators, we also excluded transects with poor weather or visibility for surveys that reported those conditions.

Table 1: Survey effort and observations considered for this model. Effort is tallied as the cumulative length of on-effort transects. Observations are the number of groups and individuals encountered while on effort. Off effort observations and those lacking an estimate of group size or distance to the group were excluded.

Institution	Program	Period	Effort 1000s km	Groups	Individuals	Mean Group Size
Aerial Surveys						
HDR	Navy Norfolk Canyon	2018-2019	11	4	4	1.0
NEAq	CNM	2017-2019	1	0	0	
NEAq	MMS-WEA	2017-2019	21	90	121	1.3
NEAq	NLPSC	2011-2015	30	58	67	1.2
NEFSC	AMAPPS	2010-2019	83	124	145	1.2
NEFSC	NARWSS	2003-2019	429	1,169	1,316	1.1
NEFSC	Pre-AMAPPS	1999-2008	45	59	60	1.0
NYS-DEC/TT	NYBWM	2017-2019	58	26	30	1.2
SEFSC	AMAPPS	2010-2019	110	16	19	1.2
UNCW	MidA Bottlenose	2002-2002	17	0	0	
UNCW	Navy Cape Hatteras	2011-2017	34	3	6	2.0
UNCW	Navy Jacksonville	2009-2017	92	11	16	1.5
UNCW	Navy Norfolk Canyon	2015-2017	14	3	3	1.0
UNCW	Navy Onslow Bay	2007-2011	49	1	2	2.0
UNCW	SEUS NARW EWS	2005-2008	114	0	0	
VAMSC	MD DNR WEA	2013-2015	15	1	1	1.0
VAMSC	Navy VACAPES	2016-2017	19	0	0	
VAMSC	VA CZM WEA	2012-2015	21	2	3	1.5
	Total		**1,164**	**1,567**	**1,793**	**1.1**
Shipboard Surveys						
MCR	SOTW Visual	2012-2019	8	31	33	1.1
NEFSC	AMAPPS	2011-2016	12	25	28	1.1
NEFSC	Pre-AMAPPS	1998-2007	13	142	145	1.0
SEFSC	AMAPPS	2011-2016	15	1	1	1.0
SEFSC	Pre-AMAPPS	1992-2006	29	1	1	1.0
	Total		**76**	**200**	**208**	**1.0**
	Grand Total		**1,240**	**1,767**	**2,001**	**1.1**

Table 2: Institutions that contributed surveys used in this model.

Institution	Full Name
HDR	HDR, Inc.
MCR	Marine Conservation Research

Table 2: Institutions that contributed surveys used in this model. *(continued)*

Institution	Full Name
NEAq	New England Aquarium
NEFSC	NOAA Northeast Fisheries Science Center
NYS-DEC/TT	New York State Department of Environmental Conservation and Tetra Tech, Inc.
SEFSC	NOAA Southeast Fisheries Science Center
UNCW	University of North Carolina Wilmington
VAMSC	Virginia Aquarium & Marine Science Center

Table 3: Descriptions and references for survey programs used in this model.

| Program	Description	References
AMAPPS	Atlantic Marine Assessment Program for Protected Species	Palka et al. (2017), Palka et al. (2021)
CNM	Northeast Canyons Marine National Monument Aerial Surveys	Redfern et al. (2021)
MD DNR WEA	Aerial Surveys of the Maryland Wind Energy Area	Barco et al. (2015)
MidA Bottlenose	Mid-Atlantic Onshore/Offshore Bottlenose Dolphin Surveys	Torres et al. (2005)
MMS-WEA	Marine Mammal Surveys of the MA and RI Wind Energy Areas	Quintana-Rizzo et al. (2021), O'Brien et al. (2022)
NARWSS	North Atlantic Right Whale Sighting Surveys	Cole et al. (2007)
Navy Cape Hatteras	Aerial Surveys of the Navy's Cape Hatteras Study Area	McLellan et al. (2018)
Navy Jacksonville	Aerial Surveys of the Navy's Jacksonville Study Area	Foley et al. (2019)
Navy Norfolk Canyon	Aerial Surveys of the Navy's Norfolk Canyon Study Area	Cotter (2019), McAlarney et al. (2018)
Navy Onslow Bay	Aerial Surveys of the Navy's Onslow Bay Study Area	Read et al. (2014)
Navy VACAPES	Aerial Survey Baseline Monitoring in the Continental Shelf Region of the VACAPES OPAREA	Mallette et al. (2017)
NLPSC	Northeast Large Pelagic Survey Collaborative Aerial Surveys	Leiter et al. (2017), Stone et al. (2017)
NYBWM	New York Bight Whale Monitoring Surveys	Zoidis et al. (2021)
Pre-AMAPPS	Pre-AMAPPS Marine Mammal Abundance Surveys	Mullin and Fulling (2003), Garrison et al. (2010), Palka (2006)
SEUS NARW EWS	Southeast U.S. Right Whale Early Warning System Surveys	
SOTW Visual	R/V Song of the Whale Visual Surveys	Ryan et al. (2013)
VA CZM WEA	Virginia CZM Wind Energy Area Surveys	Mallette et al. (2014), Mallette et al. (2015)
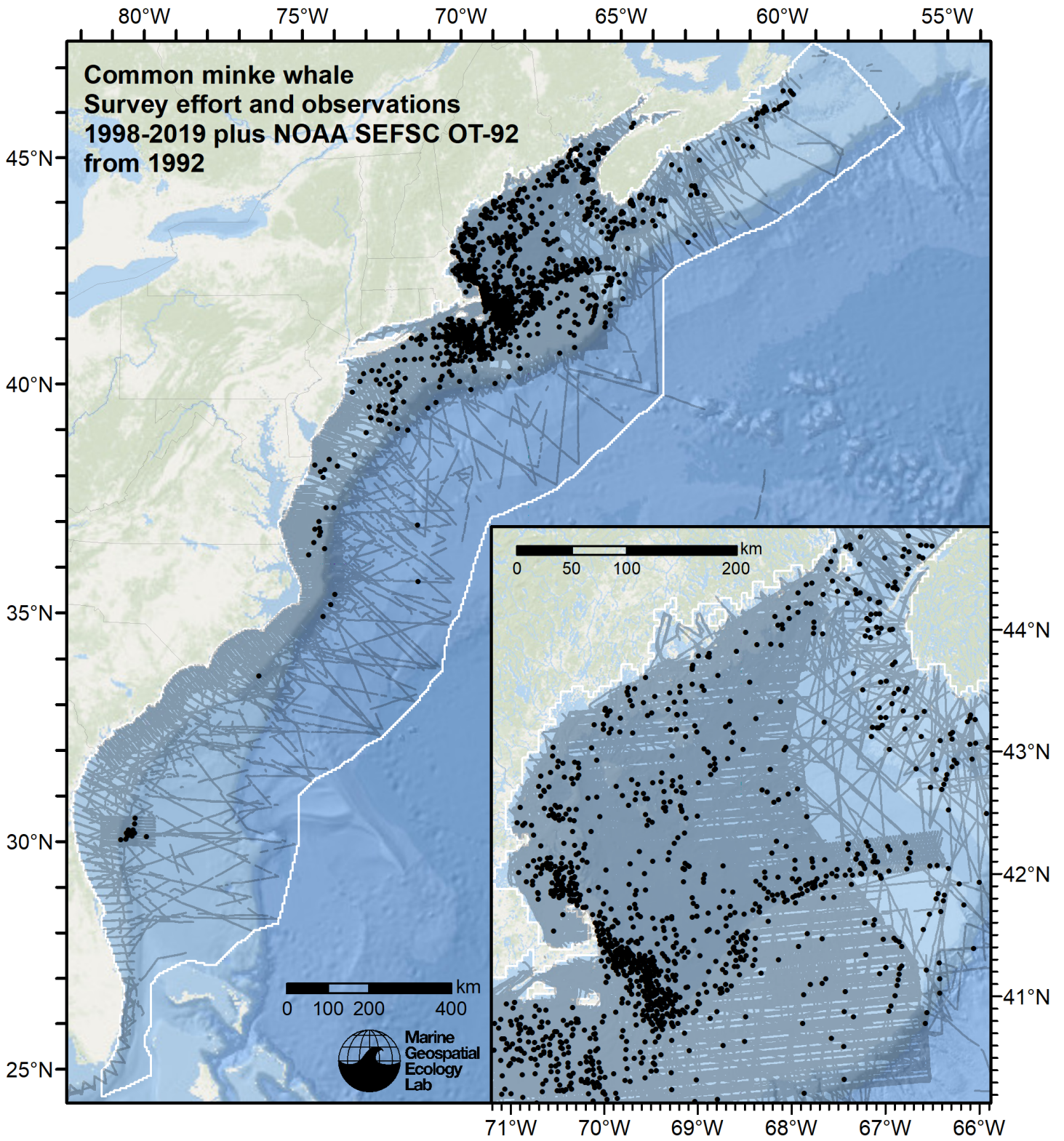
Figure 1: Survey effort and common minke whale observations available for density modeling, after detection functions were applied, and excluded segments and truncated observations were removed.

2 Detection Functions

2.1 Taxon Specific

We fitted the detection functions in this section to common minke whale observations exclusively, without pooling in other species. We usually adopted this approach when we had enough sightings of this taxon to fit a detection function without pooling and we judged that this taxon's detectability differed in important respects from others that pooling should be avoided if possible. We also occasionally used this approach for certain taxa that had similar detectability to others but for which we had so many sightings that pooling in other species provided little benefit.

2.1.1 Aaerial Surveys
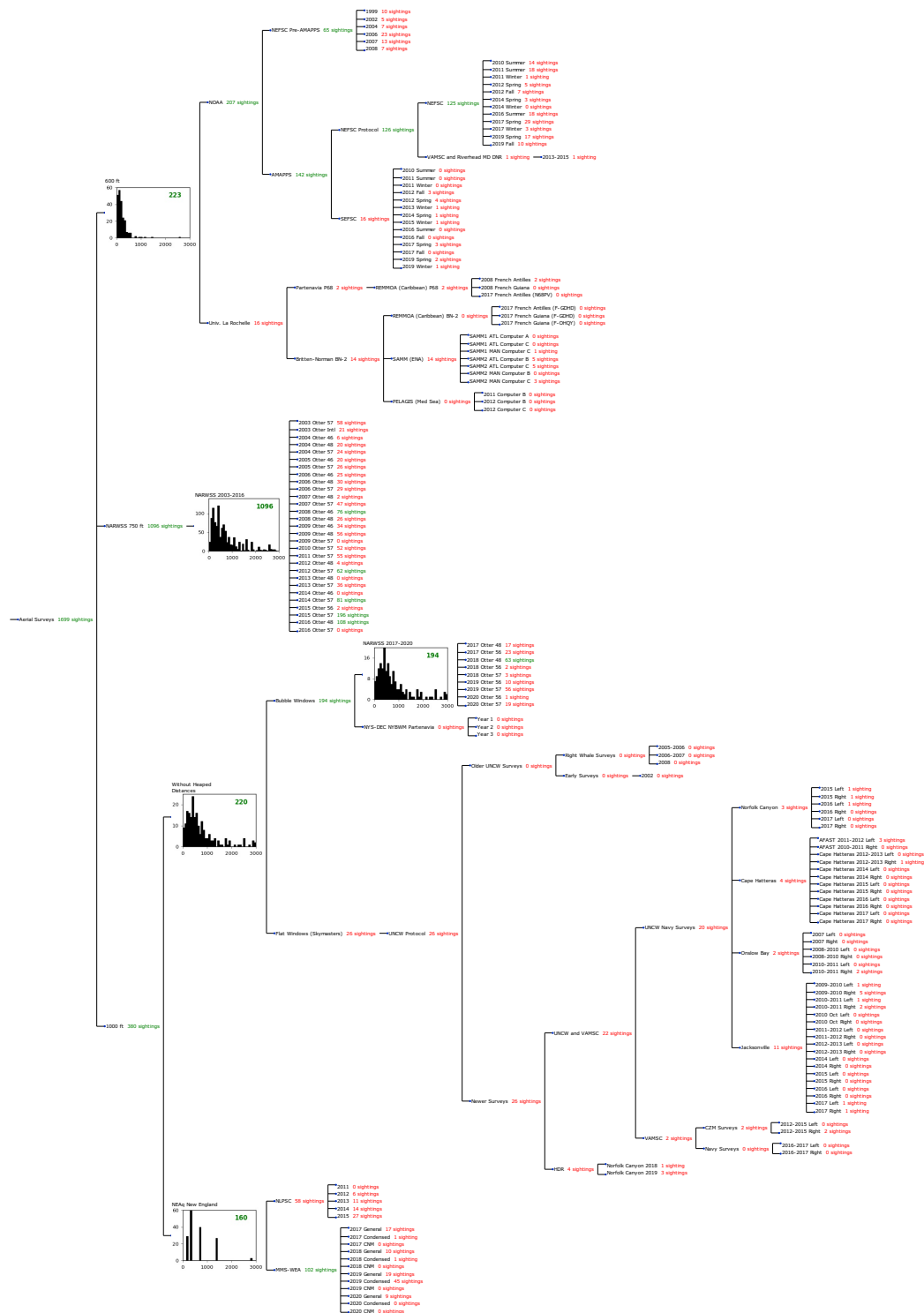
Figure 2: Detection hierarchy for aerial surveys, showing how they were pooled during detectability modeling, for taxon-specific detection functions. Each histogram represents a detection function and summarizes the perpendicular distances of observations that were pooled to fit it, prior to truncation. Observation counts, also prior to truncation, are shown in green when they met the recommendation of Buckland et al. (2001) that detection functions utilize at least 60 sightings, and red otherwise. For rare taxa, it was not always possible to meet this recommendation, yielding higher statistical uncertainty. During the spatial modeling stage of the analysis, effective strip widths were computed for each survey using the closest detection function above it in the hierarchy (i.e. moving from right to left in the figure). Surveys that do not have a detection function above them in this figure were either addressed by a detection function presented in a different section of this report, or were omitted from the analysis.

2.1.1.1 600 ft

After right-truncating observations greater than 600 m, we fitted the detection function to the 216 observations that remained. The selected detection function (Figure 3) used a hazard rate key function with Program (Figure 4) as a covariate.
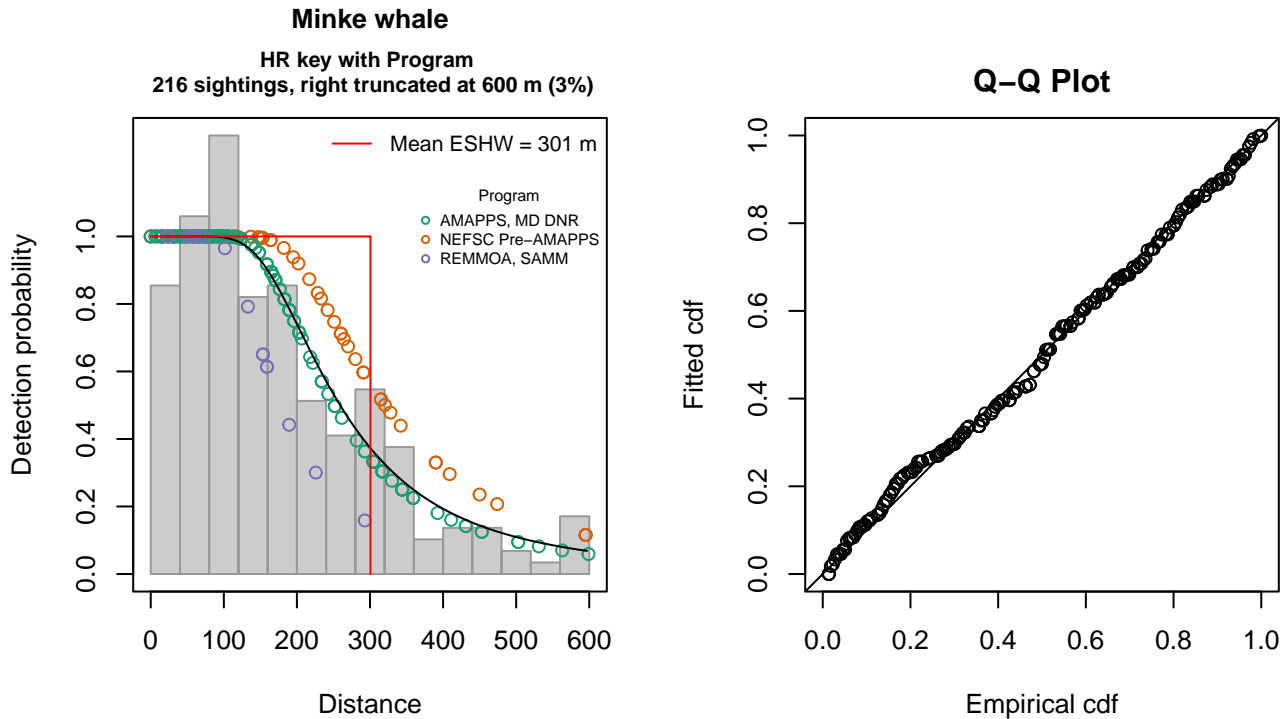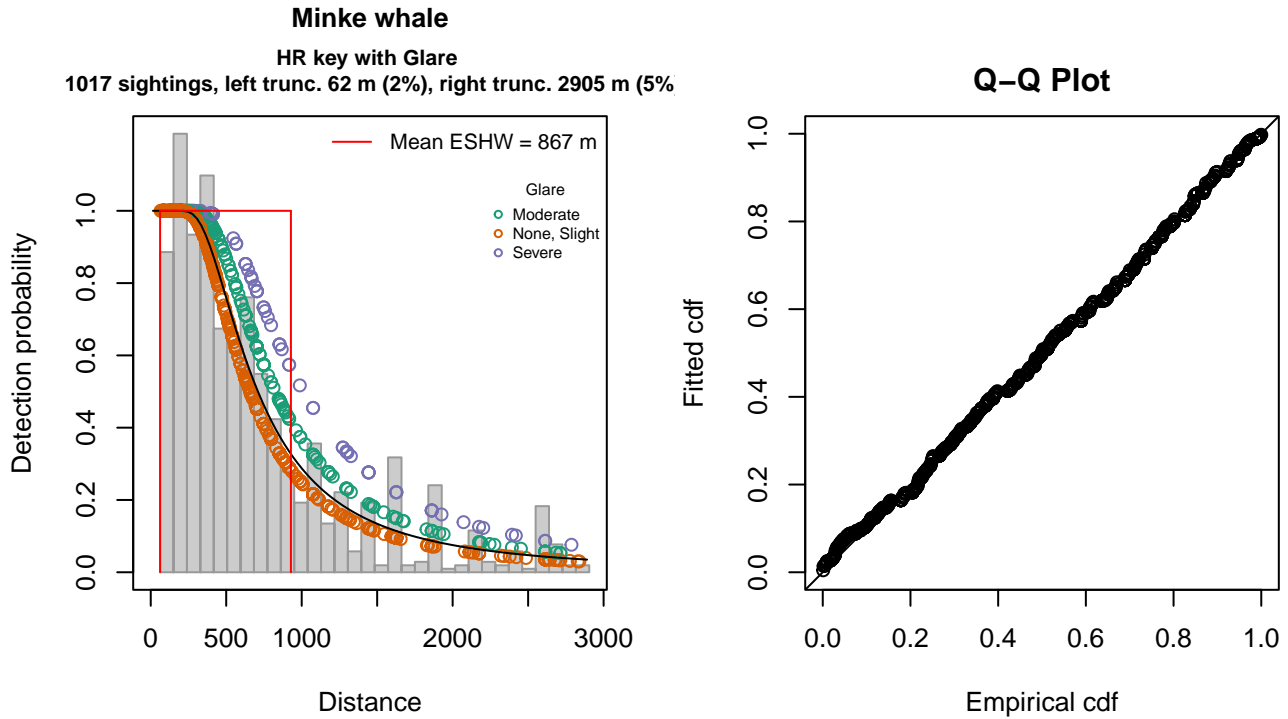

Figure 3: 600 ft detection function and Q-Q plot showing its goodness of fit.

Statistical output for this detection function:

```
Summary for ds object
Number of observations :  216
Distance range         :  0  -  600
AIC                    :  2642.407

Detection function:
 Hazard-rate key function

Detection function parameters
Scale coefficient(s):
                          estimate         se
(Intercept)              5.3954215  0.1186837
ProgramNEFSC Pre-AMAPPS  0.2437832  0.1628141
ProgramREMMOA, SAMM     -0.3436494  0.2798051

Shape coefficient(s):
            estimate         se
(Intercept) 1.028465  0.1708956


                      Estimate          SE         CV
Average p             0.492268  0.03333238 0.06771186
N in covered region 438.785415 36.64934025 0.08352452

Distance sampling Cramer-von Mises test (unweighted)
Test statistic = 0.057999  p = 0.826585
```
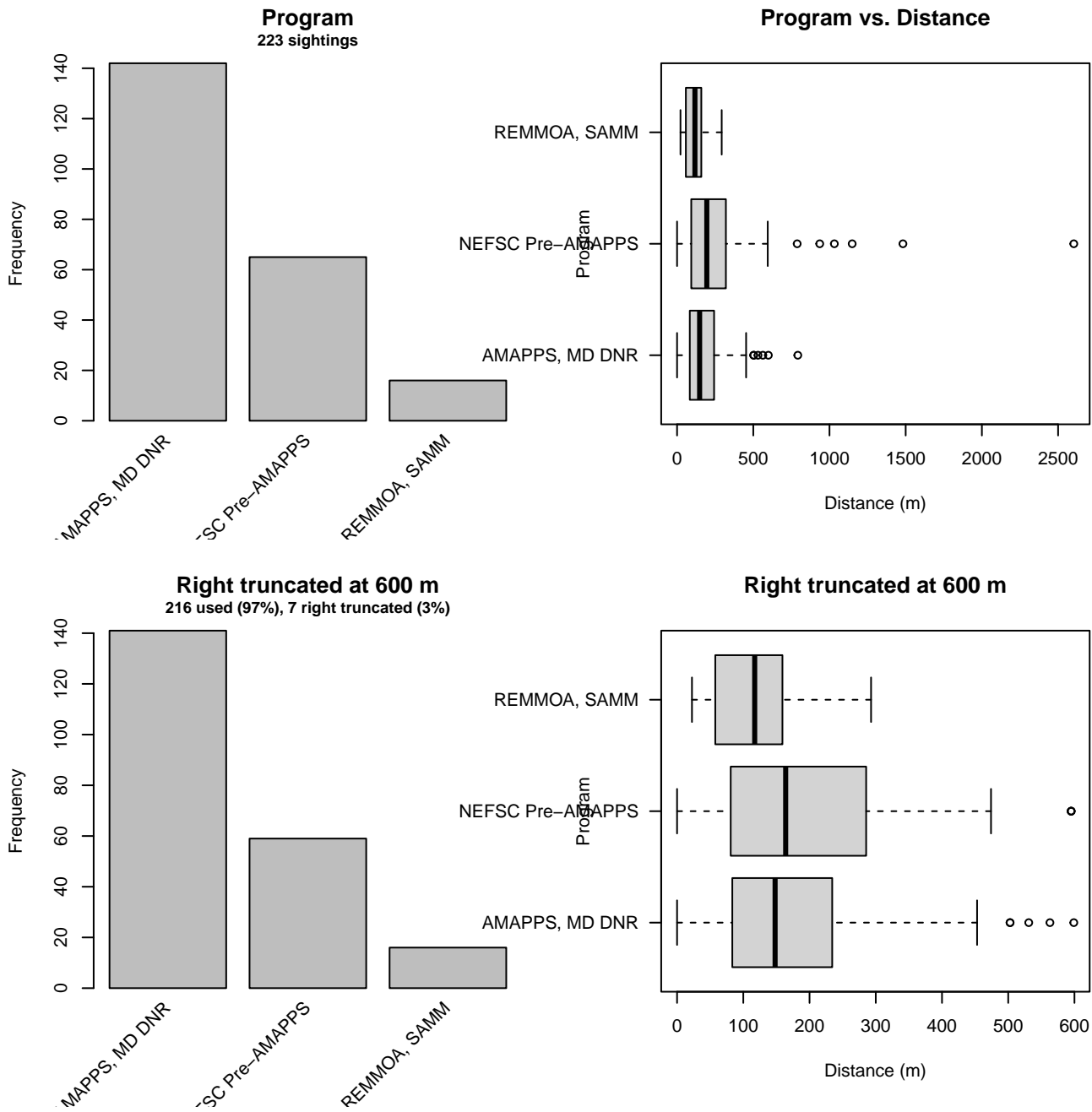
Figure 4: Distribution of the Program covariate before (top row) and after (bottom row) observations were truncated to fit the 600 ft detection function.

2.1.1.2 NARWSS 2003-2016

After right-truncating observations greater than 2905 m and left-truncating observations less than 62 m (Figure 6), we fitted the detection function to the 1017 observations that remained. The selected detection function (Figure 5) used a hazard rate key function with Glare (Figure 7) as a covariate.

Minke whale
HR key with Glare
1017 sightings, left trunc. 62 m (2%), right trunc. 2905 m (5%

Q–Q Plot

Figure 5: NARWSS 2003-2016 detection function and Q-Q plot showing its goodness of fit.

Statistical output for this detection function:

```
Summary for ds object
Number of observations :   1017
Distance range         :   62  -  2905
AIC                    :   15149.11

Detection function:
 Hazard-rate key function

Detection function parameters
Scale coefficient(s):
                     estimate          se
(Intercept)         6.5439116 0.09680953
GlareNone, Slight  -0.2252438 0.09935993
GlareSevere         0.2028217 0.15796343

Shape coefficient(s):
            estimate          se
(Intercept) 0.7603569 0.05405745


                      Estimate          SE          CV
Average p              0.29948   0.01315968 0.04394177
N in covered region 3395.88674 173.98472747 0.05123396

Distance sampling Cramer-von Mises test (unweighted)
Test statistic = 0.115253  p = 0.515150
```

Left trucated sightings (in red)

1017 used (93%), 22 left trunc. (2%), 59 right trunc. (5%)
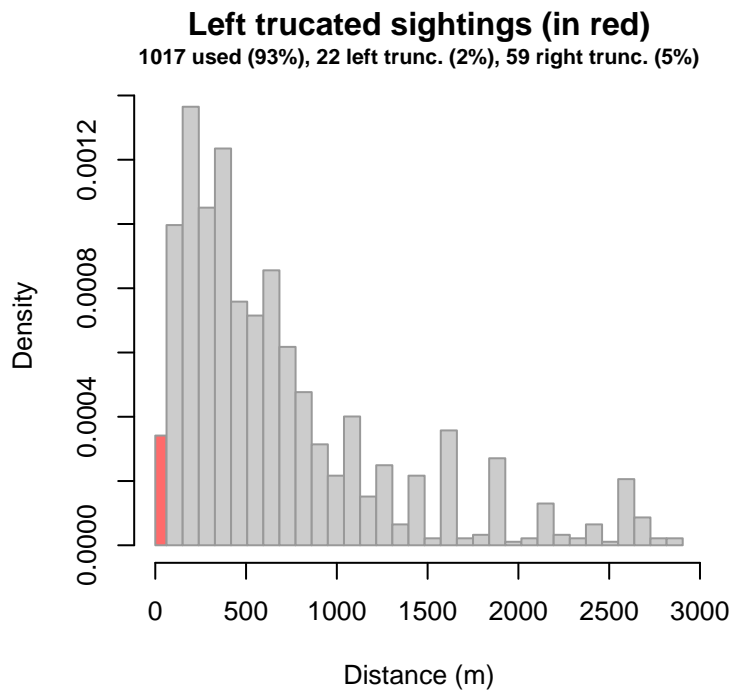
Figure 6: Density histogram of observations used to fit the NARWSS 2003-2016 detection function, with the left-most bar showing observations at distances less than 62 m, which were left-truncated and excluded from the analysis [Buckland et al. (2001)]. (This bar may be very short if there were very few left-truncated sightings, or very narrow if the left truncation distance was very small; in either case it may not appear red.)
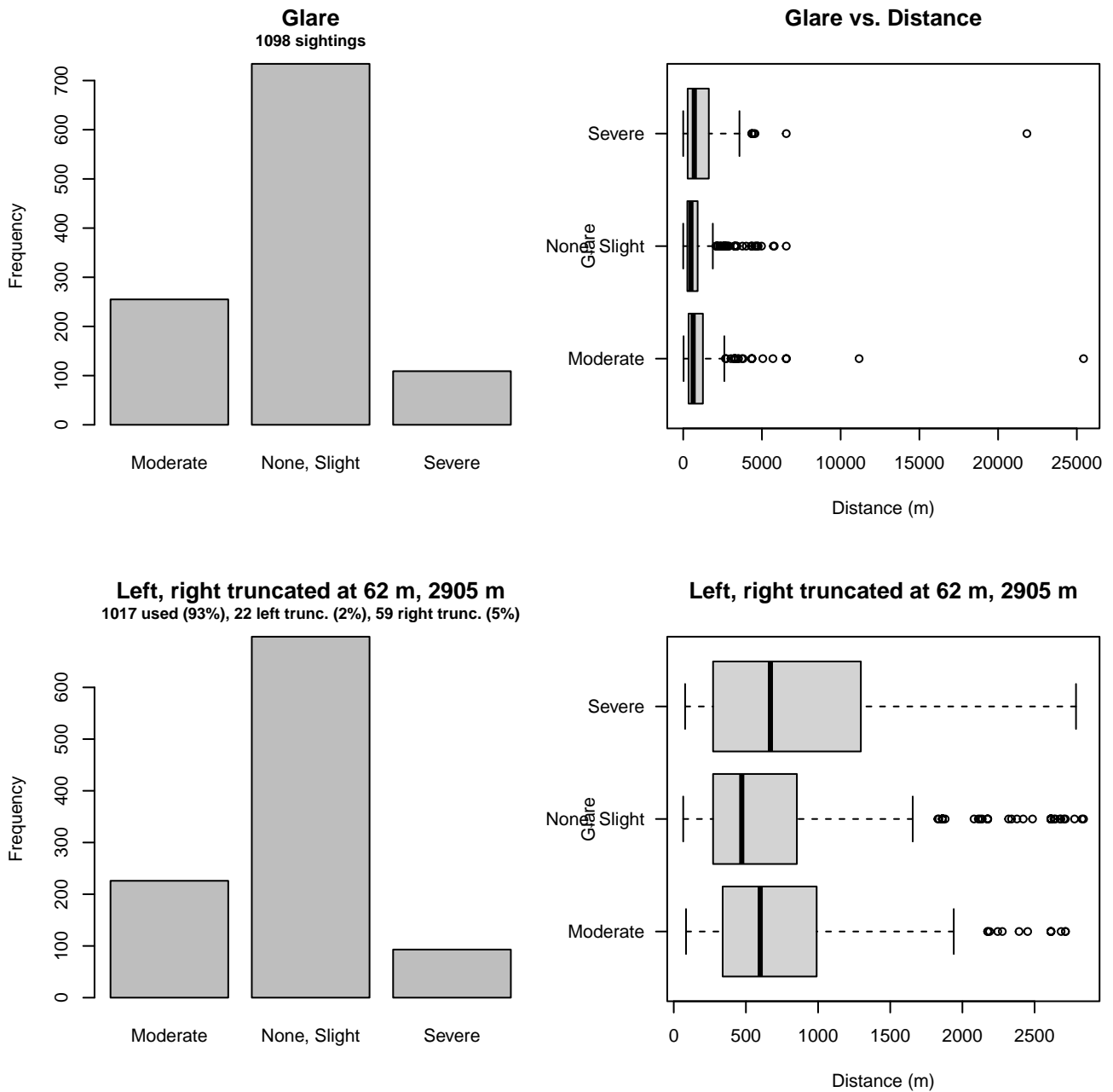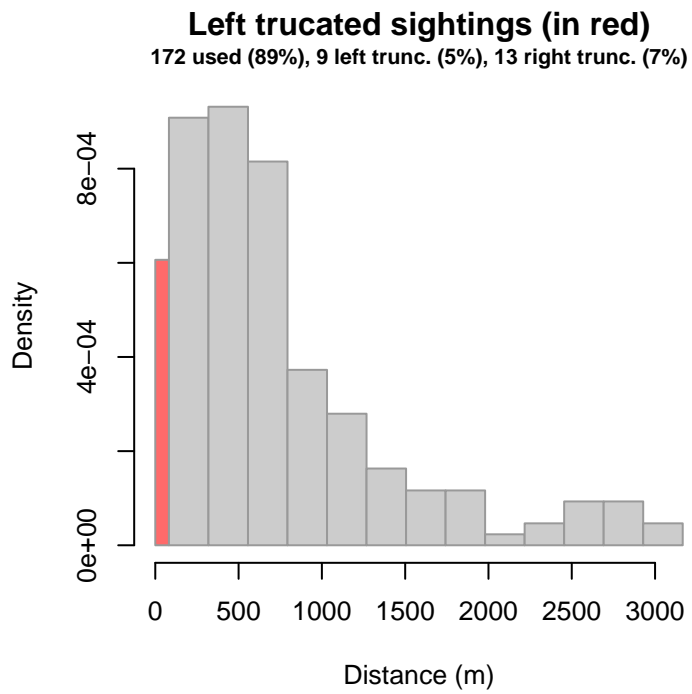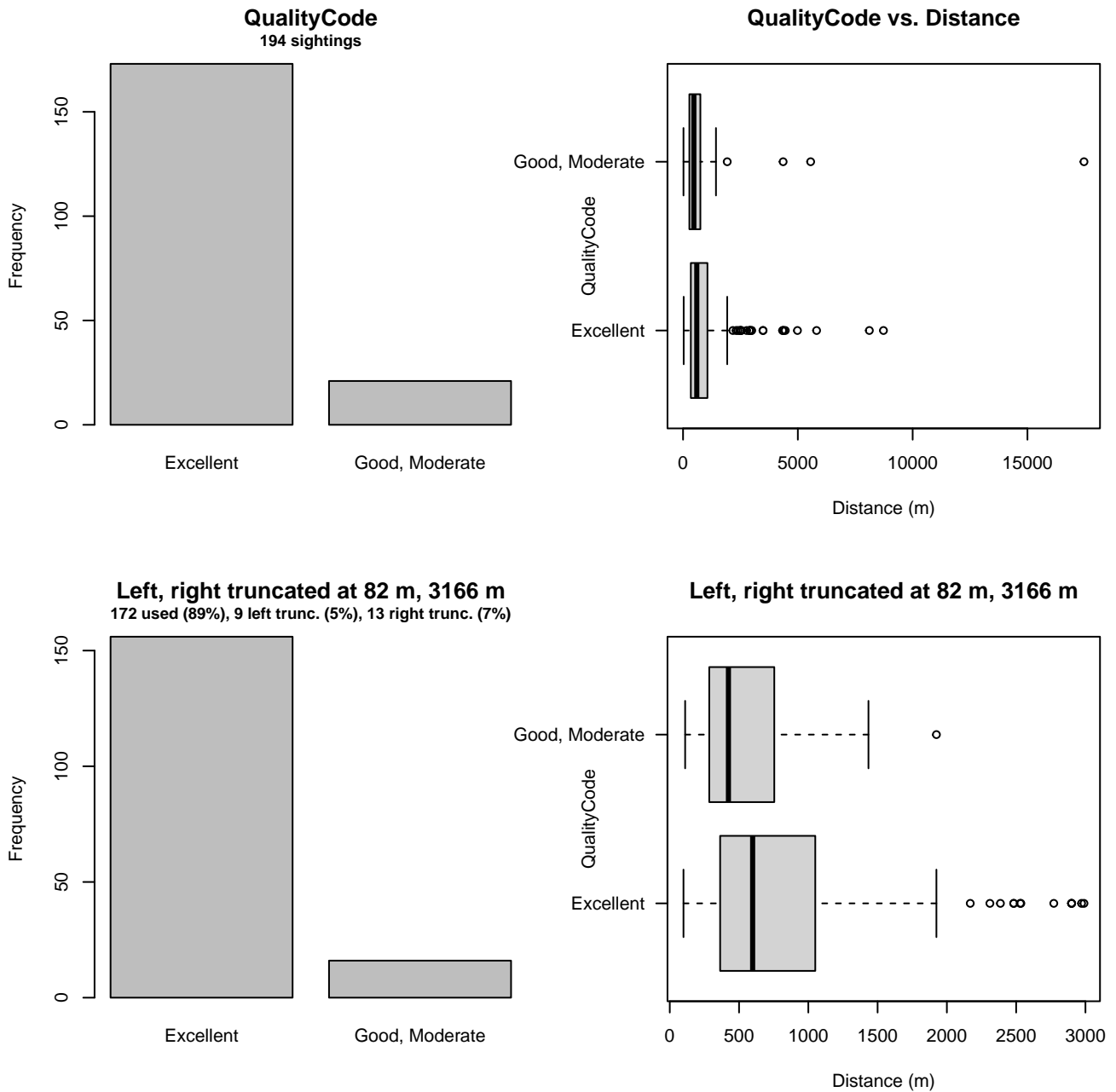
Figure 7: Distribution of the Glare covariate before (top row) and after (bottom row) observations were truncated to fit the NARWSS 2003-2016 detection function.

2.1.1.3 NARWSS 2017-2020

After right-truncating observations greater than 3166 m and left-truncating observations less than 82 m (Figure 9), we fitted the detection function to the 172 observations that remained. The selected detection function (Figure 8) used a hazard rate key function with QualityCode (Figure 10) as a covariate.
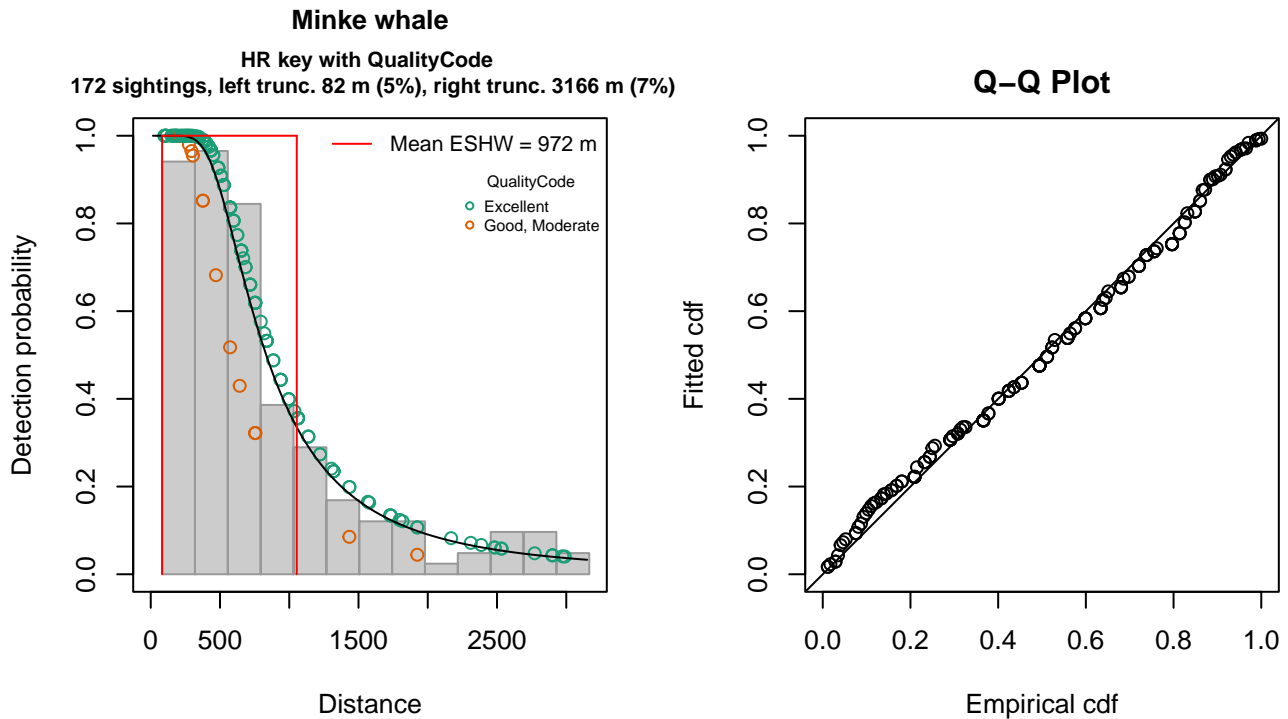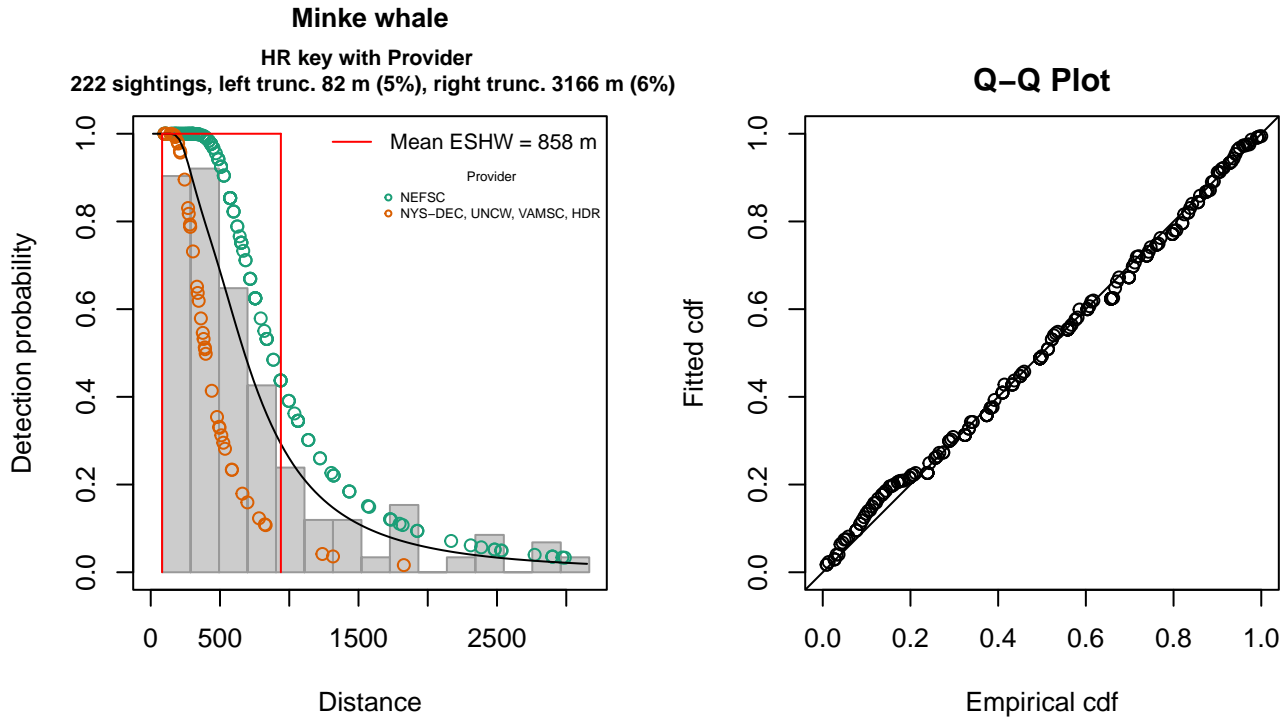
Figure 8: NARWSS 2017-2020 detection function and Q-Q plot showing its goodness of fit.

Statistical output for this detection function:

```
Summary for ds object
Number of observations :  172
Distance range         :  82  -  3166
AIC                    :  2595.786

Detection function:
 Hazard-rate key function

Detection function parameters
Scale coefficient(s):
                           estimate        se
(Intercept)               6.6103494 0.1425903
QualityCodeGood, Moderate -0.3971451 0.3196362

Shape coefficient(s):
            estimate        se
(Intercept) 0.8285285 0.1288261

                    Estimate        SE         CV
Average p          0.3109887  0.03230521 0.1038790
N in covered region 553.0746973 67.33379185 0.1217445

Distance sampling Cramer-von Mises test (unweighted)
Test statistic = 0.084119  p = 0.668383
```

Left trucated sightings (in red)

172 used (89%), 9 left trunc. (5%), 13 right trunc. (7%)

Figure 9: Density histogram of observations used to fit the NARWSS 2017-2020 detection function, with the left-most bar showing observations at distances less than 82 m, which were left-truncated and excluded from the analysis [Buckland et al. (2001)]. (This bar may be very short if there were very few left-truncated sightings, or very narrow if the left truncation distance was very small; in either case it may not appear red.)

Figure 10: Distribution of the QualityCode covariate before (top row) and after (bottom row) observations were truncated to fit the NARWSS 2017-2020 detection function.

2.1.1.4 Without Heaped Distances

After right-truncating observations greater than 3166 m and left-truncating observations less than 82 m (Figure 12), we fitted the detection function to the 222 observations that remained. The selected detection function (Figure 11) used a hazard rate key function with Provider (Figure 13) as a covariate.

Minke whale

Figure 11: Without Heaped Distances detection function and Q-Q plot showing its goodness of fit.

Statistical output for this detection function:

```
Summary for ds object
Number of observations :   222
Distance range         :   82  -  3166
AIC                    :   3278.884

Detection function:
 Hazard-rate key function

Detection function parameters
Scale coefficient(s):
                                   estimate          se
(Intercept)                       6.6177869 0.1185000
ProviderNYS-DEC, UNCW, VAMSC, HDR -0.7854538 0.1785224

Shape coefficient(s):
            estimate          se
(Intercept) 0.8948026 0.1065167


                    Estimate          SE          CV
Average p          0.2457718   0.02438505 0.09921823
N in covered region 903.2768302 104.51166763 0.11570281

Distance sampling Cramer-von Mises test (unweighted)
Test statistic = 0.069187  p = 0.756566
```

Left trucated sightings (in red)

222 used (90%), 12 left trunc. (5%), 14 right trunc. (6%)
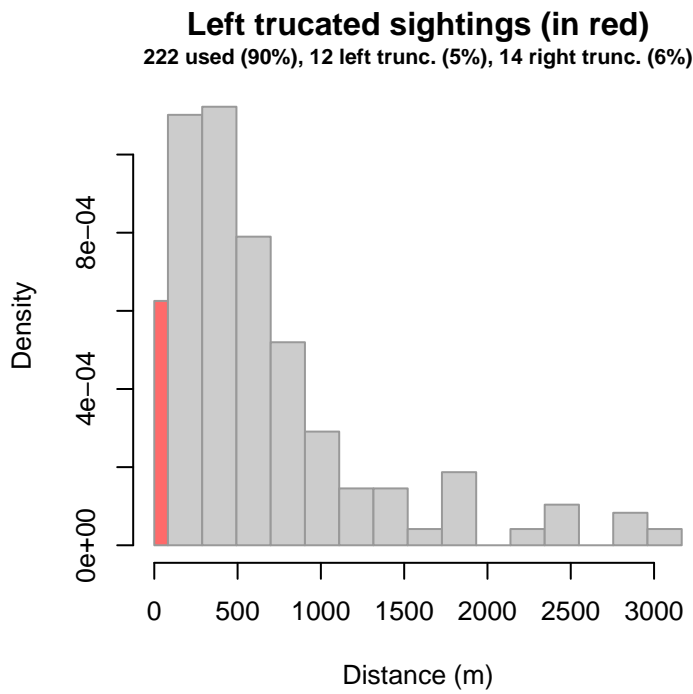
Figure 12: Density histogram of observations used to fit the Without Heaped Distances detection function, with the left-most bar showing observations at distances less than 82 m, which were left-truncated and excluded from the analysis [Buckland et al. (2001)]. (This bar may be very short if there were very few left-truncated sightings, or very narrow if the left truncation distance was very small; in either case it may not appear red.)
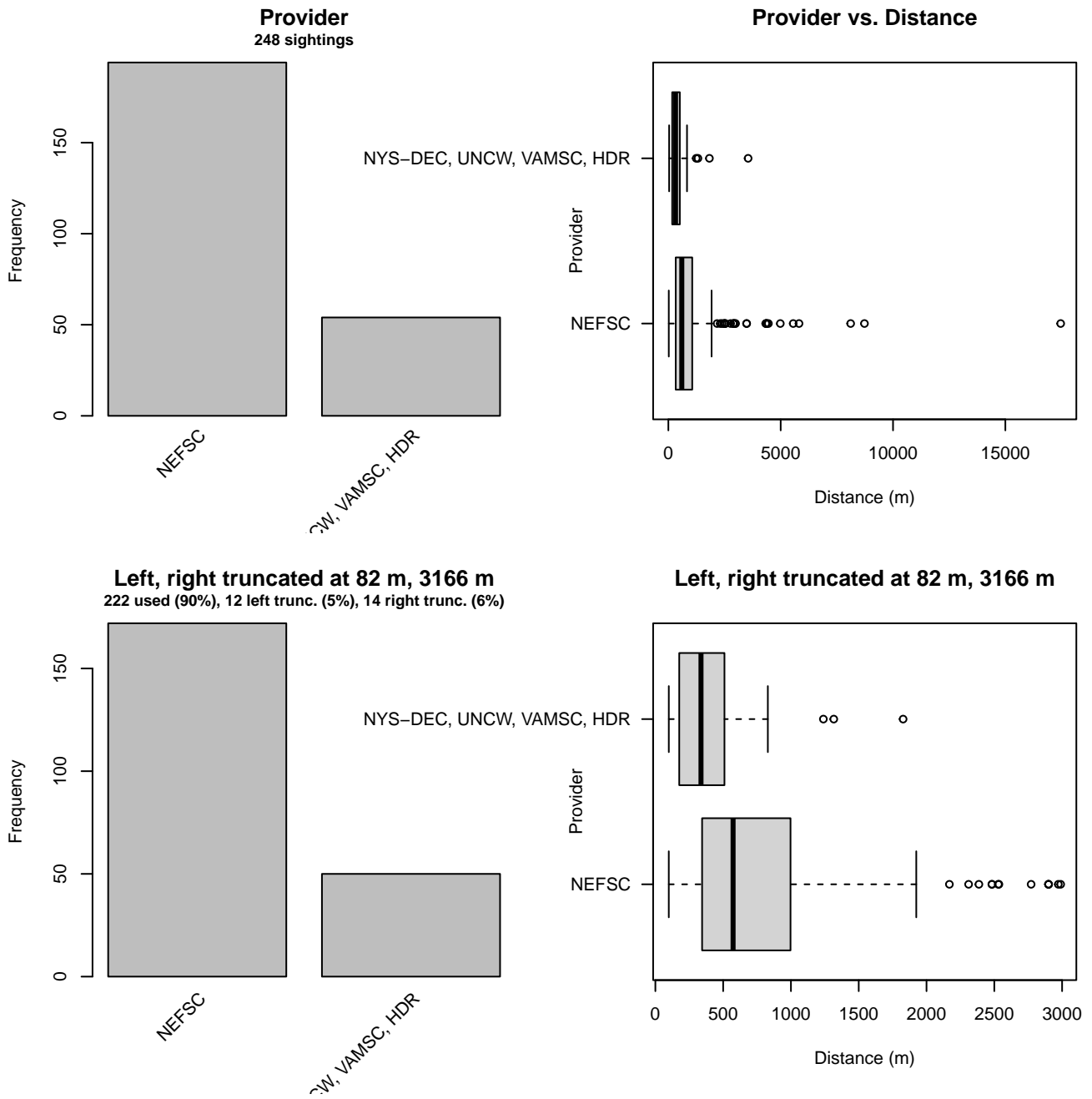
Figure 13: Distribution of the Provider covariate before (top row) and after (bottom row) observations were truncated to fit the Without Heaped Distances detection function.

2.1.1.5 NEAq New England

After right-truncating observations greater than 1852 m and left-truncating observations less than 71 m (Figure 15), we fitted the detection function to the 156 observations that remained. The selected detection function (Figure 14) used a half normal key function with Beaufort (Figure 16) and Glare (Figure 17) as covariates.
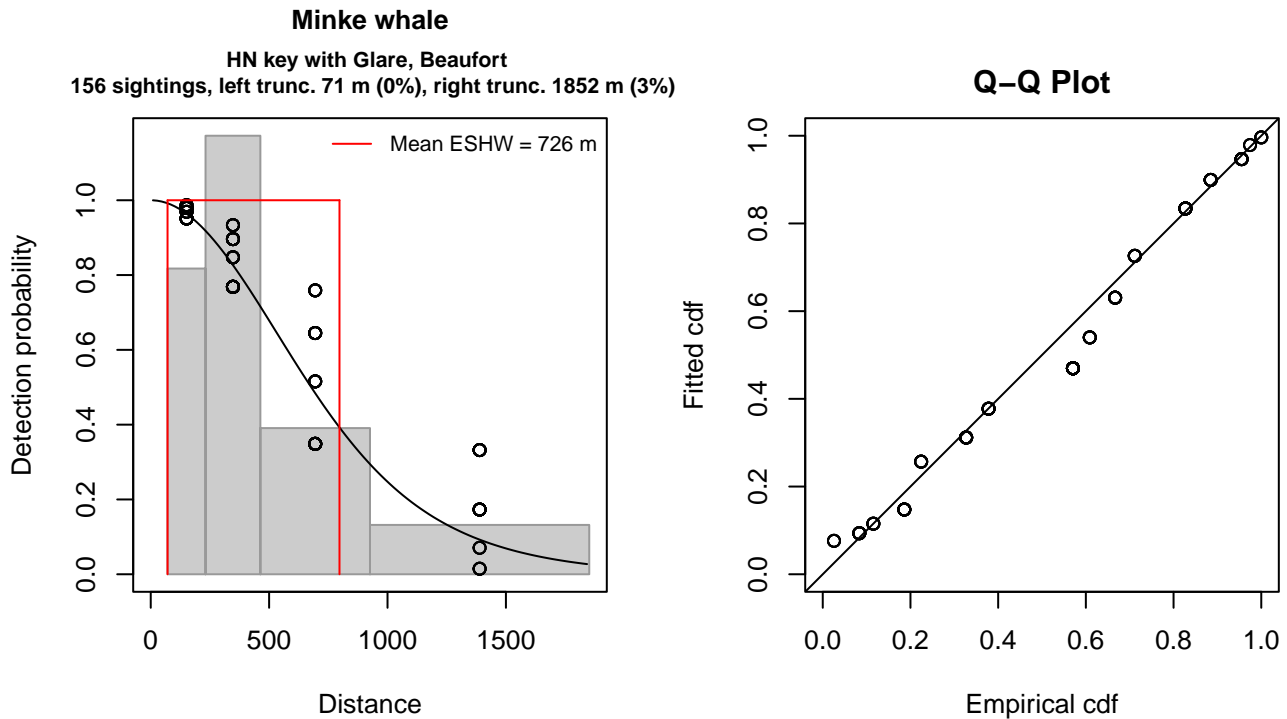
Figure 14: NEAq New England detection function and Q-Q plot showing its goodness of fit.

Statistical output for this detection function:

```
Summary for ds object
Number of observations :   156
Distance range         :   71  -  1852
AIC                    :   420.4569

Detection function:
 Half-normal key function

Detection function parameters
Scale coefficient(s):
                      estimate        se
(Intercept)          6.4026993 0.1487021
GlareNone, Slight   -0.2322762 0.1608882
Beaufort2            0.4382890 0.1443732


                      Estimate        SE          CV
Average p            0.3810585  0.0283158  0.07430827
N in covered region 409.3859868 40.2369358 0.09828606

Distance sampling Cramer-von Mises test (unweighted)
Test statistic = 0.348869  p = 0.099025
```

Left trucated sightings (in red)

156 used (98%), 0 left trunc. (0%), 4 right trunc. (2%)
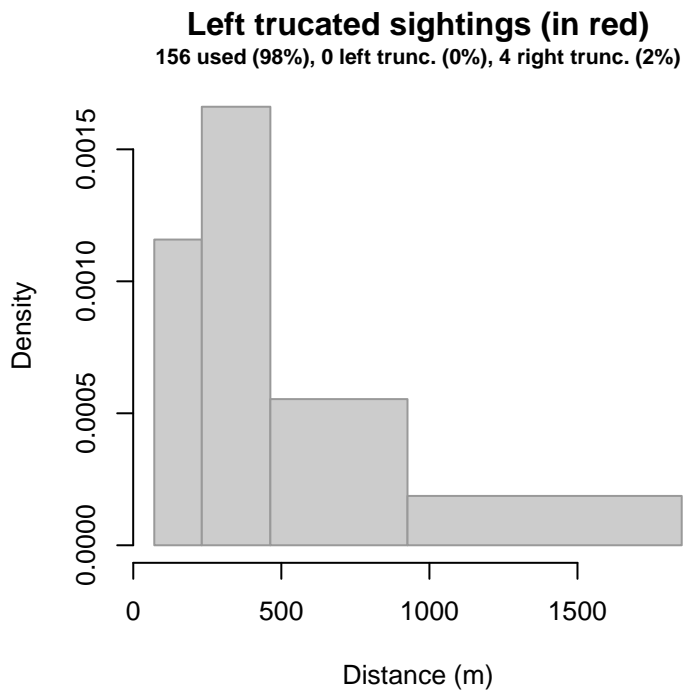
Figure 15: Density histogram of observations used to fit the NEAq New England detection function, with the left-most bar showing observations at distances less than 71 m, which were left-truncated and excluded from the analysis [Buckland et al. (2001)]. (This bar may be very short if there were very few left-truncated sightings, or very narrow if the left truncation distance was very small; in either case it may not appear red.)
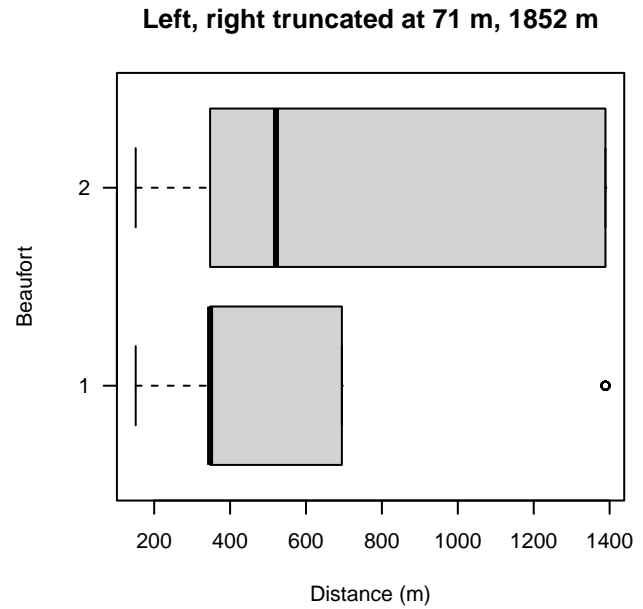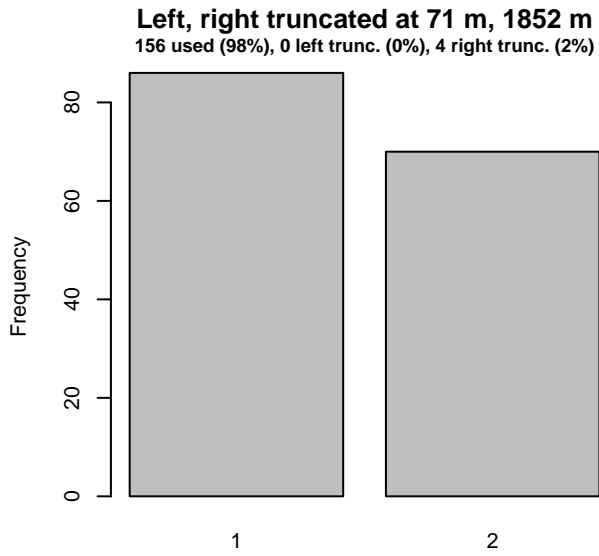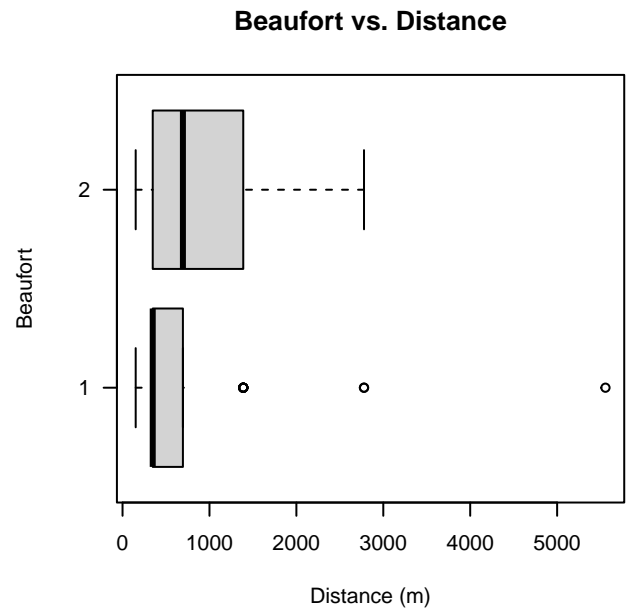
Figure 16: Distribution of the Beaufort covariate before (top row) and after (bottom row) observations were truncated to fit the NEAq New England detection function.
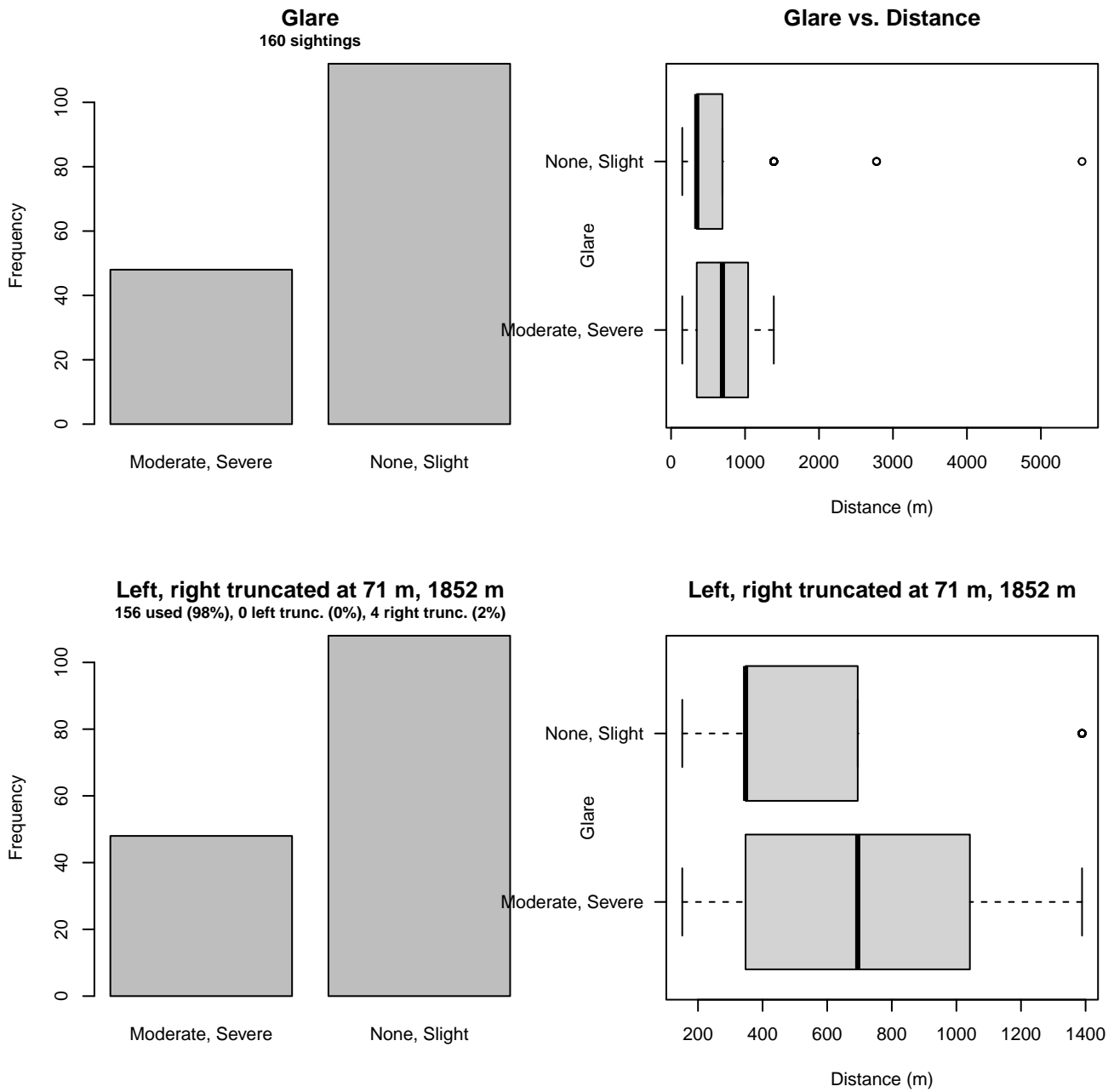
Figure 17: Distribution of the Glare covariate before (top row) and after (bottom row) observations were truncated to fit the NEAq New England detection function.
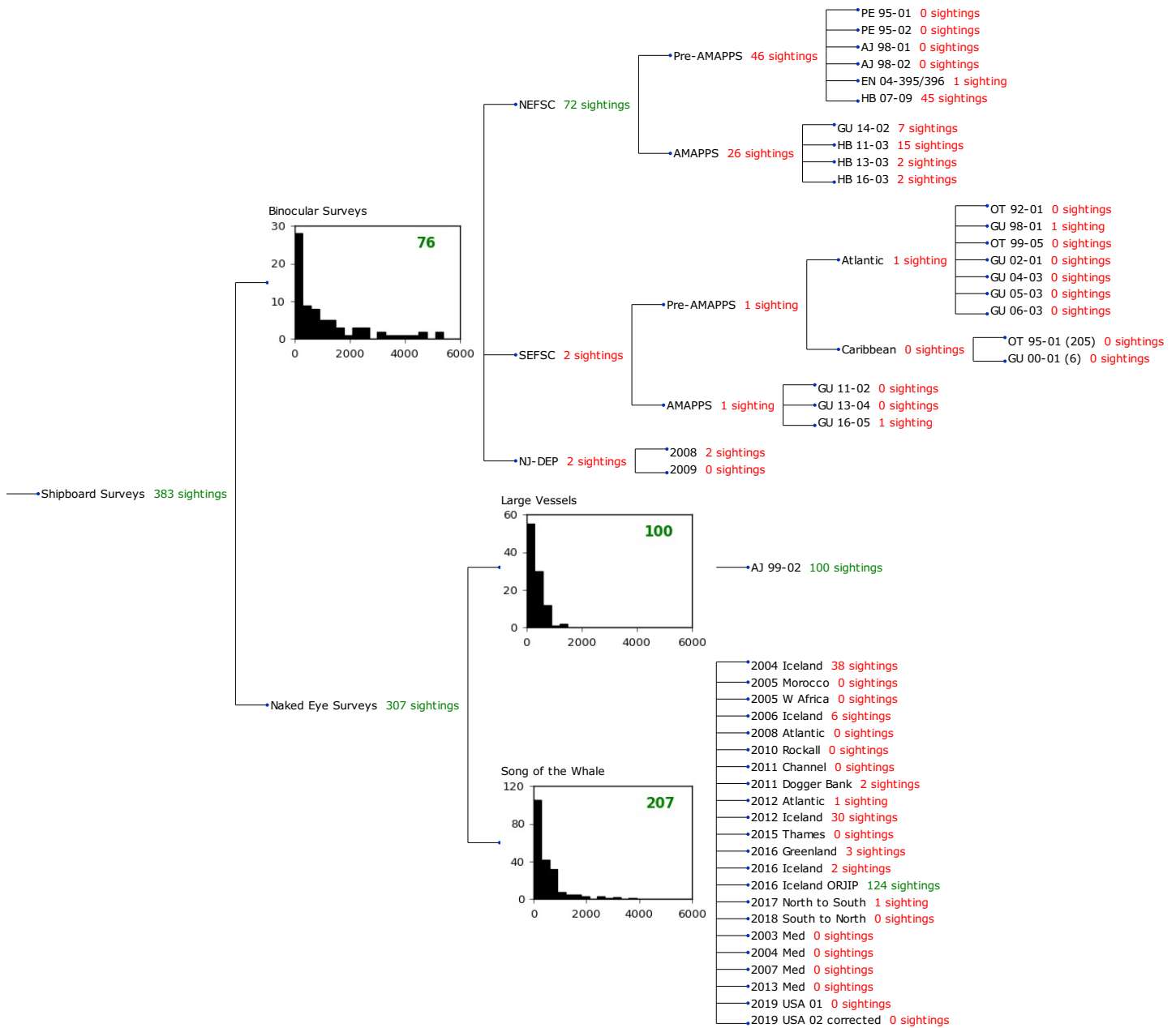
2.1.2 Shipboard Surveys

Figure 18: Detection hierarchy for shipboard surveys, showing how they were pooled during detectability modeling, for taxon-specific detection functions. Each histogram represents a detection function and summarizes the perpendicular distances of observations that were pooled to fit it, prior to truncation. Observation counts, also prior to truncation, are shown in green when they met the recommendation of Buckland et al. (2001) that detection functions utilize at least 60 sightings, and red otherwise. For rare taxa, it was not always possible to meet this recommendation, yielding higher statistical uncertainty. During the spatial modeling stage of the analysis, effective strip widths were computed for each survey using the closest detection function above it in the hierarchy (i.e. moving from right to left in the figure). Surveys that do not have a detection function above them in this figure were either addressed by a detection function presented in a different section of this report, or were omitted from the analysis.

2.1.2.1 Binocular Surveys

After right-truncating observations greater than 5000 m, we fitted the detection function to the 73 observations that remained. The selected detection function (Figure 19) used a hazard rate key function with no covariates.
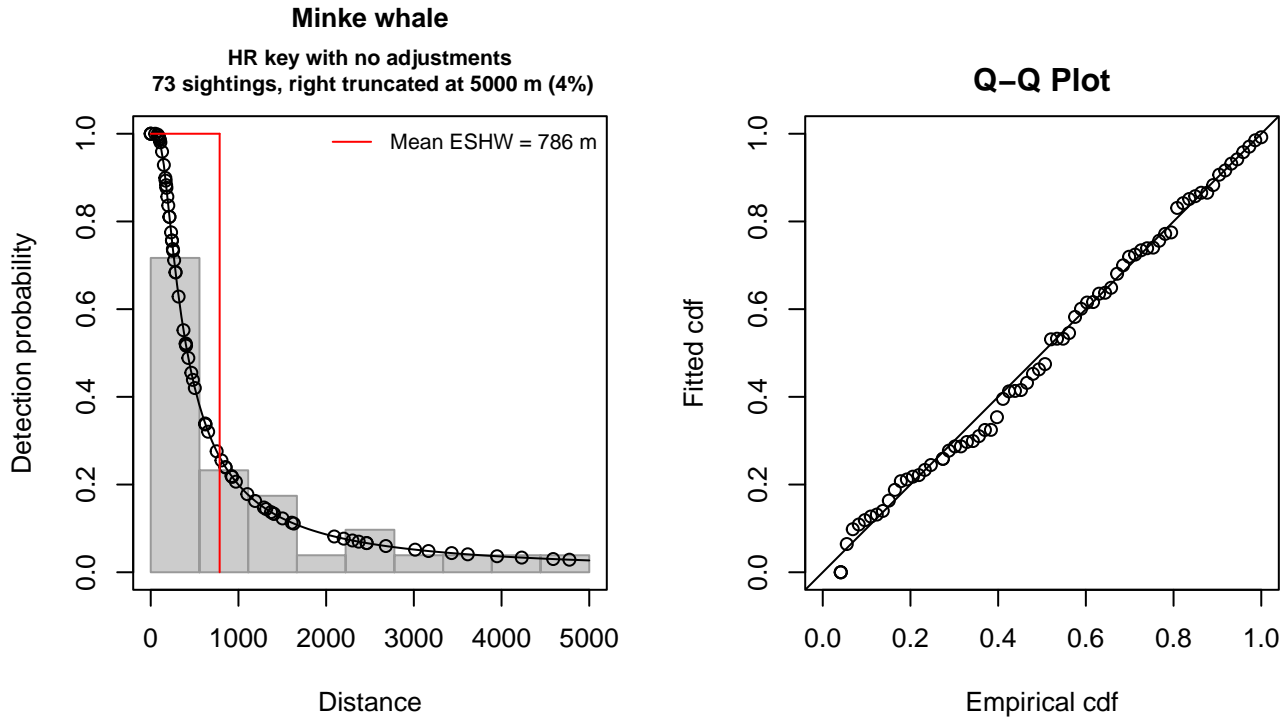
Minke whale

HR key with no adjustments
73 sightings, right truncated at 5000 m (4%)

Q–Q Plot

Figure 19: Binocular Surveys detection function and Q-Q plot showing its goodness of fit.

Statistical output for this detection function:

```
Summary for ds object
Number of observations :   73
Distance range         :   0  -  5000
AIC                    :   1155.783

Detection function:
 Hazard-rate key function

Detection function parameters
Scale coefficient(s):
            estimate         se
(Intercept) 5.753977 0.4508822

Shape coefficient(s):
            estimate         se
(Intercept) 0.2639556 0.1672159

                     Estimate          SE         CV
Average p           0.1571369  0.04103682 0.2611533
N in covered region 464.5631114 131.19045726 0.2823953

Distance sampling Cramer-von Mises test (unweighted)
Test statistic = 0.031039  p = 0.972631
```

2.1.2.2   Large Vessels

After right-truncating observations greater than 1000 m, we fitted the detection function to the 98 observations that remained. The selected detection function (Figure 20) used a half normal key function with no covariates.
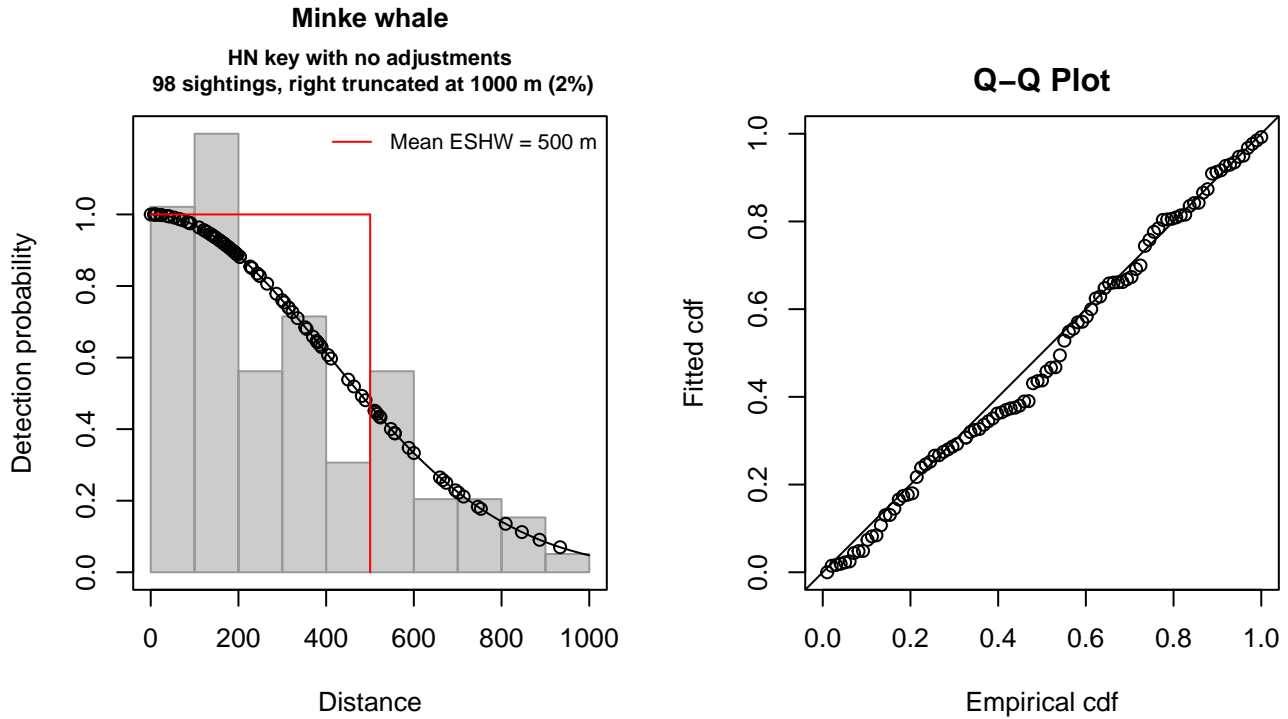
Minke whale

HN key with no adjustments
98 sightings, right truncated at 1000 m (2%)

Q–Q Plot

Figure 20: Large Vessels detection function and Q-Q plot showing its goodness of fit.

Statistical output for this detection function:

```
Summary for ds object
Number of observations :  98
Distance range        :  0  -  1000
AIC                   :  1308.981

Detection function:
 Half-normal key function

Detection function parameters
Scale coefficient(s):
          estimate          se
(Intercept) 6.003282 0.08605203


                    Estimate        SE          CV
Average p            0.5004431  0.03899715 0.07792524
N in covered region 195.8264459 20.69642140 0.10568757


Distance sampling Cramer-von Mises test (unweighted)
Test statistic = 0.064317  p = 0.786839
```

2.1.2.3   Song of the Whale

After right-truncating observations greater than 2000 m, we fitted the detection function to the 200 observations that remained. The selected detection function (Figure 21) used a hazard rate key function with Clouds (Figure 22) and WaveHeight (Figure 23) as covariates.
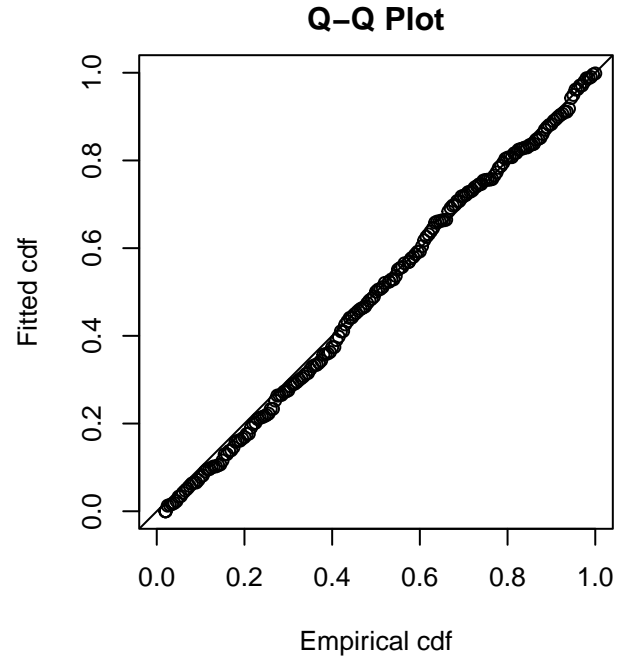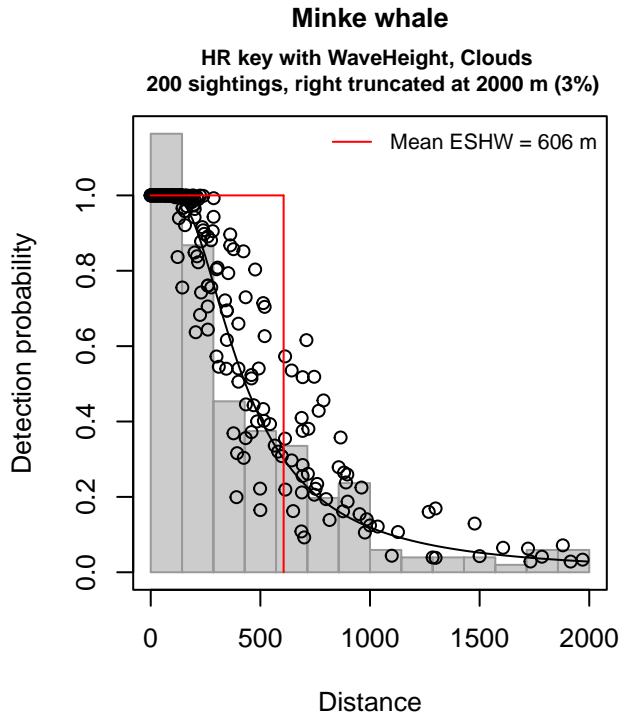
Figure 21: Song of the Whale detection function and Q-Q plot showing its goodness of fit.

Statistical output for this detection function:

```
Summary for ds object
Number of observations :   200
Distance range         :   0  -   2000
AIC                    :   2815.351

Detection function:
 Hazard-rate key function

Detection function parameters
Scale coefficient(s):
              estimate          se
(Intercept)  6.54523085  0.21123279
WaveHeight  -1.49547058  0.69178639
Clouds      -0.05194221  0.02552016

Shape coefficient(s):
            estimate         se
(Intercept) 0.7780387  0.1248031

                    Estimate          SE          CV
Average p          0.2817449  0.02616734  0.0928760
N in covered region 709.8618891 78.80285301 0.1110115

Distance sampling Cramer-von Mises test (unweighted)
Test statistic = 0.057586   p = 0.829184
```
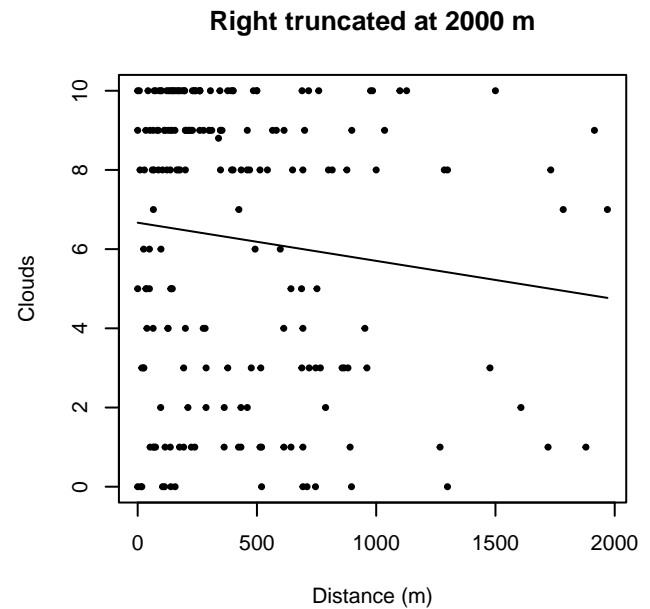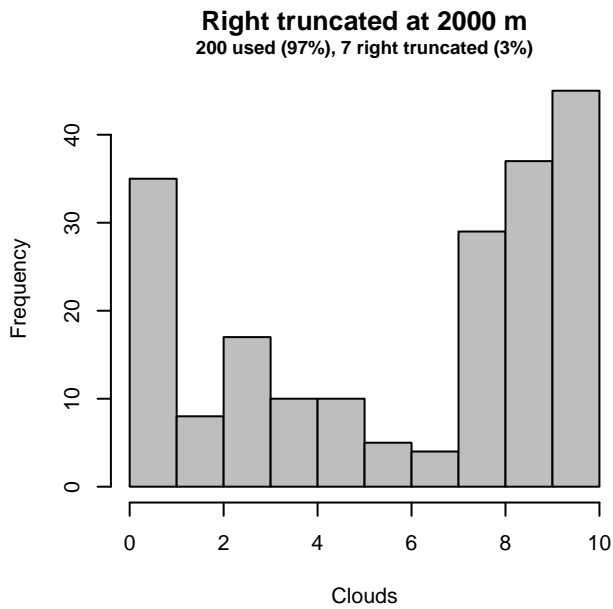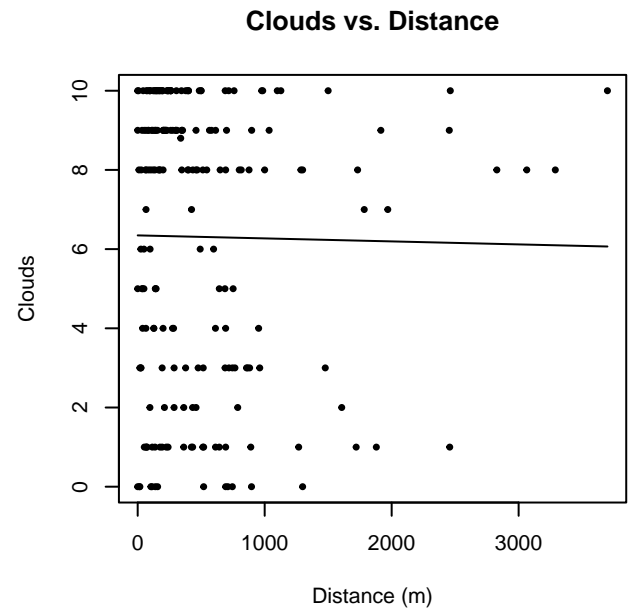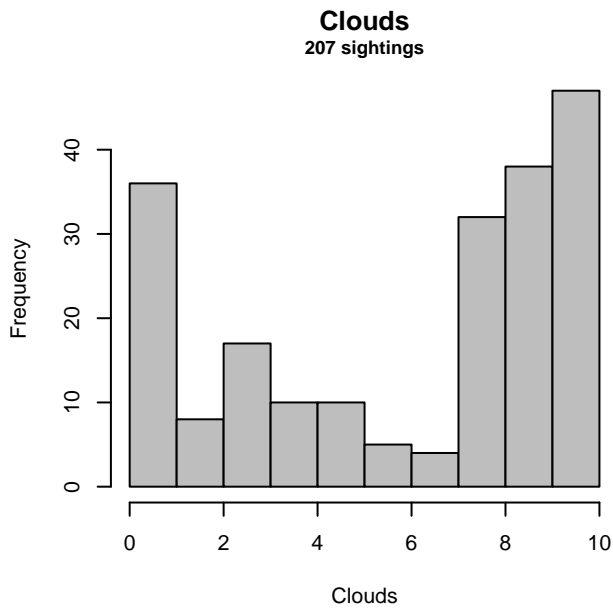
Figure 22: Distribution of the Clouds covariate before (top row) and after (bottom row) observations were truncated to fit the Song of the Whale detection function.
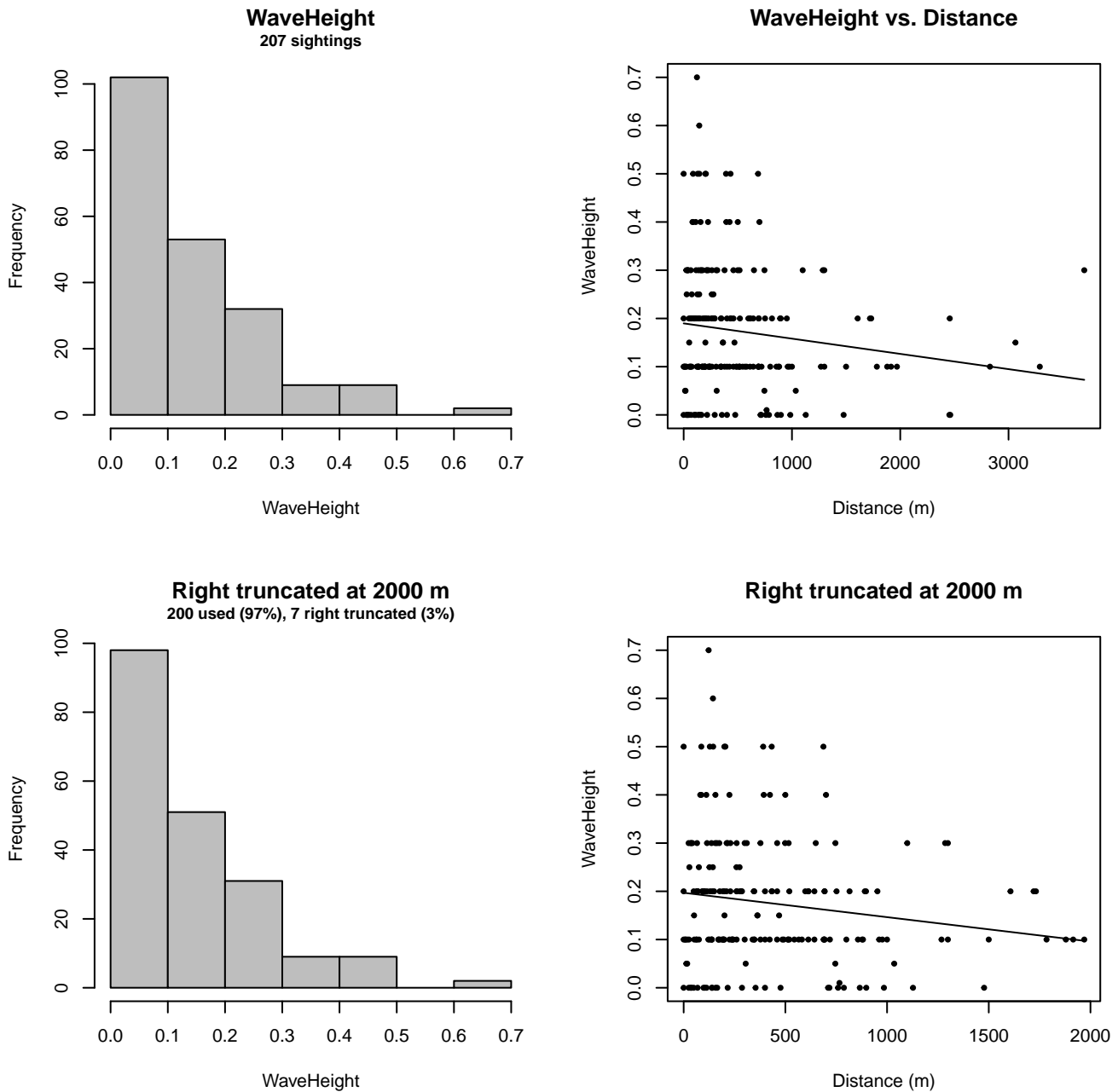
Figure 23: Distribution of the WaveHeight covariate before (top row) and after (bottom row) observations were truncated to fit the Song of the Whale detection function.

3 Bias Corrections

Density surface modeling methodology uses *distance sampling* (Buckland et al. 2001) to model the probability that an observer on a line transect survey will detect an animal given the perpendicular distance to it from the transect line. Distance sampling assumes that detection probability is 1 when perpendicular distance is 0. When this assumption is not met, detection probability is biased high, leading to an underestimation of density and abundance. This is known as the $g_0 < 1$ problem, where g_0 refers to the detection probability at distance 0. Modelers often try to address this problem by estimating g_0 empirically and dividing it into estimated density or abundance, thereby correcting those estimates to account for the animals that were presumed missed.

Two important sources of bias for visual surveys are known as *availability bias*, in which an animal was present on the transect line but impossible to detect, e.g. because it was under water, and *perception bias*, in which an animal was present and available but not noticed, e.g. because of its small size or cryptic coloration or behavior (Marsh and Sinclair 1989). Modelers often

estimate the influence of these two sources of bias on detection probability independently, yielding two estimates of g_0, hereafter referred to as g_{0A} and g_{0P}, and multiply them together to obtain a final, combined estimate: $g_0 = g_{0A}.g_{0P}$.

Our overall approach was to perform this correction on a per-observation basis, to have the flexibility to account for many factors such as platform type, surveyor institution, group size, group composition (e.g. singleton, mother-calf pair, or surface active group), and geographic location (e.g. feeding grounds vs. calving grounds). The level of complexity of the corrections varied by species according to the amount of information available, with North Atlantic right whale having the most elaborate corrections, derived from a substantial set of publications documenting its behavior, and various lesser known odontocetes having corrections based only on platform type (aerial or shipboard), derived from comparatively sparse information. Here we document the corrections used for common minke whale.

3.1 Aerial Surveys

Palka et al. (2021) developed perception bias corrections using two team, mark recapture distance sampling (MRDS) methodology (Burt et al. 2014) for aerial surveys conducted in 2010-2017 by NOAA NEFSC and SEFSC during the AMAPPS program. These were the only extant perception bias estimates developed from aerial surveys used in our analysis, aside from estimates developed earlier by Palka and colleagues (Palka 2006; Palka et al. 2017). Those earlier efforts utilized older methods and less data than their 2021 analysis, so we applied the Palka et al. (2021) estimates to all aerial survey programs (Table 4).

We applied Palka's estimate for NEFSC to all programs other than SEFSC on the basis that those programs employed a similar visual scanning protocol that allowed observers to scan from the trackline up to the horizon, while SEFSC's protocol generally limited scanning only up to 50° from the trackline, resulting in a smaller effective strip width. Also, the estimate for NEFSC was made from a pooling of species that was strongly dominated by minke whales, while the estimate for SEFSC was a more even mixture of many species, dominated by larger whales and minke whales only comprising 20% of the pool.

We caution that it is possible that perception bias was different on the other aerial programs, as they often used different aircraft, flew at different altitudes, and were staffed by different personnel. Of particular concern are that many programs flew Cessna 337 Skymasters, which had flat windows, while NOAA flew de Havilland Twin Otters, which had bubble windows, which likely afforded a better view of the transect line and therefore might have required less of a correction than the Skymasters. Correcting the other programs using NOAA's estimate as we have done is likely to yield less bias than leaving them uncorrected, but we urge all programs to undertake their own efforts to estimate perception bias, as resources allow.

We estimated availability bias corrections using the Laake et al. (1997) estimator and dive intervals reported by Palka et al. (2017) (Table 5). To estimate time in view, needed by the Laake estimator, we used results reported by Robertson et al. (2015), rescaled linearly for each survey program according to its target altitude and speed. We caution that Robertson's analysis was done for a de Havilland Twin Otter, which may have a different field of view than that of the other aircraft used here, which mainly comprised Cessna 337 Skymasters with flat windows but also a Partenavia P-68 with bubble windows (on the NYS-DEC/TT surveys). However, we note that McLellan et al. (2018) conducted a sensitivity analysis on the influence of the length of the "window of opportunity" to view beaked whales from a Cessna Skymaster on their final density estimates and found that they varied by only a few thousandths of an animal per kilometer when the window of opportunity more than doubled. Still, we urge additional program-specific research into estimation of availability bias.

To address the influence of group size on availability bias, we applied the group availability estimator of McLellan et al. (2018) on a per-observation basis. Following Palka et al. (2021), who also used that method, we assumed that individuals in the group dived asynchronously. The resulting g_{0A} corrections ranged from about 0.37 to 1 (Figure 24), with the majority less than 0.5. We caution that the assumption of asynchronous diving can lead to an underestimation of density and abundance if diving is actually synchronous; see McLellan et al. (2018) for an exploration of this effect. However, if future research finds that this species conducts synchronous dives and characterizes the degree of synchronicity, the model can be updated to account for this knowledge.

Table 4: Perception bias corrections for common minke whale applied to aerial surveys.

Surveys	Group Size	g_{0P}	g_{0P} Source
SEFSC	Any	0.86	Palka et al. (2021): SEFSC
All others	Any	0.62	Palka et al. (2021): NEFSC

Table 5: Surface and dive intervals for common minke whale used to estimate availability bias corrections.

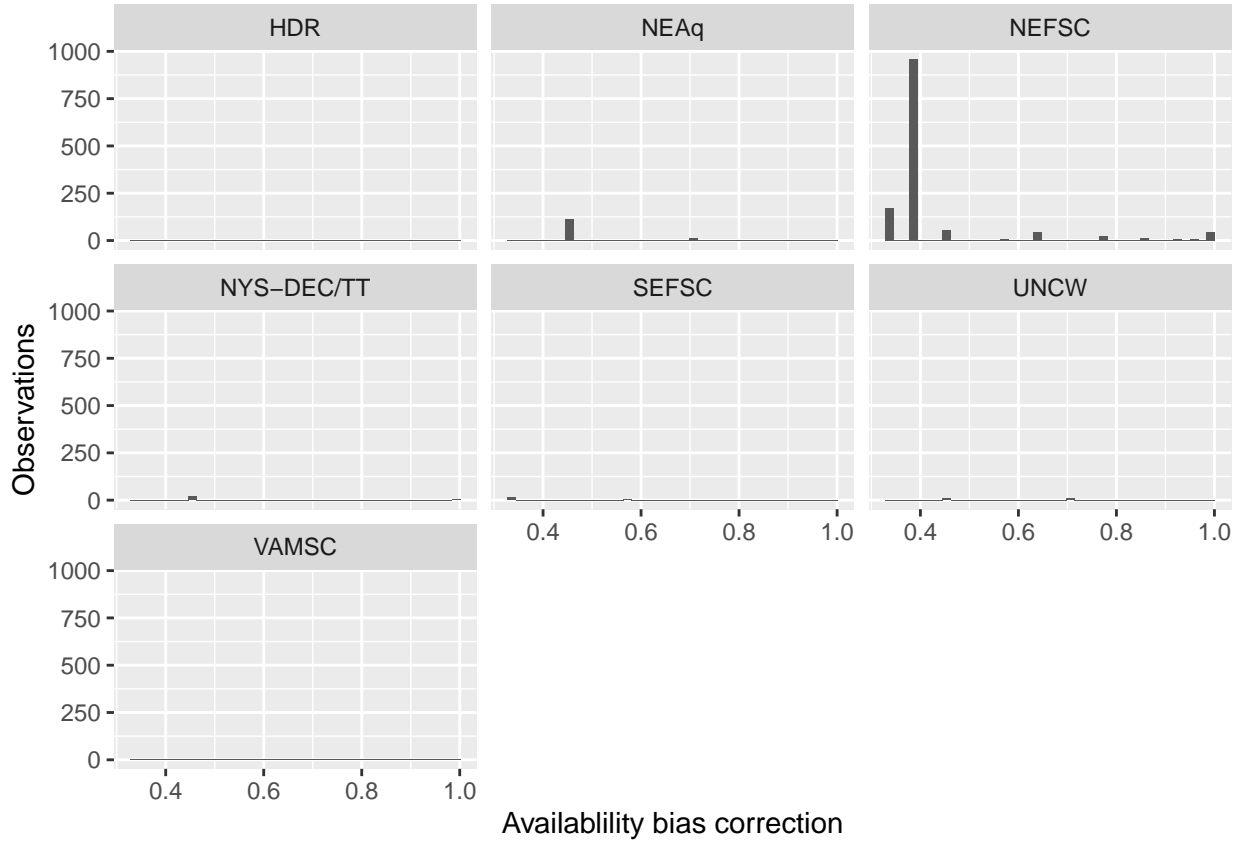| Surface Interval (s)	Dive Interval (s)	Source
26.8	116.4	Palka et al. (2017)

Figure 24: Availability bias corrections for common minke whale for aerial surveys, by institution.

3.2 Shipboard Surveys

Most of the shipboard surveys in our analysis used high-power (25x150), pedestal-mounted binoculars. Similar to aerial surveys, Palka et al. (2021) developed perception bias corrections using two team, MRDS methodology (Burt et al. 2014) for high-power binocular surveys conducted in 2010-2017 by NOAA NEFSC and SEFSC during the AMAPPS program. These were the only extant perception bias estimates developed from high-power binocular surveys used in our analysis, aside from estimates developed earlier by Palka and colleagues (Palka 2006; Palka et al. 2017). Those earlier efforts utilized older methods and less data than their 2021 analysis, so we applied the Palka et al. (2021) estimates to all shipboard surveys that searched with high-power binoculars (Table 6).

Table 6: Perception and availability bias corrections for common minke whale applied to shipboard surveys.

Surveys	Searching Method	Group Size	g_{0P}	g_{0P} Source	g_{0A}	g_{0A} Source
NEFSC	Binoculars	Any	0.48	Palka et al. (2021): NEFSC	1	Assumed
SEFSC	Binoculars	Any	0.57	Palka et al. (2021): SEFSC	1	Assumed
NEFSC, MCR	Naked eye	Any	0.69	Palka et al. (2006)	1	Assumed

Palka (2006) also developed a correction for an NEFSC shipboard survey (AJ 99-02) in which the primary team observers searched by naked eye. We applied this estimate to that survey as well as to the MCR Song of the Whale surveys, which also searched by naked eye but did not have a program-specific estimate. We caution that the platform height for the MCR surveys was substantially (~9m) lower than the NEFSC survey, and the target survey speed was slower (6 knots for MCR

vs. 10 knots for NEFSC)

Given that the dive interval of this species (Table 5) was short relative to the amount of time a given patch of water remained in view to shipboard observers, we assumed that no availability bias correction was needed ($g_{0A} = 1$).

4   Density Model

Minke whales are distributed widely across the North Atlantic, from the tropics to high latitudes. They are believed to follow a typical baleen whale migration pattern, moving to high latitudes in summer to feed and low latitudes in winter to breed or calve (Risch et al. 2014). While a number of summer feeding grounds have been identified across the productive shelf waters of temperate and sub-polar regions of the North Atlantic, the winter breeding grounds are relatively unknown and may lie far offshore in the tropics or subtropics. Víkingsson and Heide-Jørgensen (2015) tracked three minke whales by satellite departing Icelandic feeding grounds in late summer of 2002 and 2004. The whales moved south along the Mid-Atlantic Ridge; the final transmission from the longest-transmitting tag was received on 5 December 2004 from about latitude 25 °N, 1000 km northwest of the Cape Verde Islands. Nieukirk et al. (2004) deployed an array of six hydrophones along either side of the Mid-Atlantic ridge between 15-35 °N and 33-50 °W from February 1999 to March 2001 and reported minkle whale pulse trains from October to June, with the bulk of detections occurring in November through March, suggesting that the central North Atlantic could be a wintering ground for minke whales.

Risch et al. (2014) characterized minke whale seasonality along the east coast of North America as observed through passive acoustic monitoring. Minkes were detected in high numbers south of Cape Hatteras from December-March, off New York in April and May, and at Stellwagen Bank, and at Emerald and Roseway Basins from April through November, with a peak in September and October. Aerial surveys of the U.S. Navy's Undersea Warfare Training Range off of Jacksonville, Florida recorded sightings on the deep side of the shelf break in the months of December through February, including mother-calf pairs (Foley et al. 2019). More extensive acoustic monitoring from 2017-2020 indicated that minke whales are present across the Blake Plateau from December through April with the peak concentration at southern locations from December through March (Kiehbadroudinezhad et al. 2021; Kowarski et al. 2022).

Given the evidence that minke whales exhibit strong migratory behavior that is potentially indicative of a seasonal switch in habitat relationships—e.g., a preference for cold, temperate waters during feeding season vs. warm, subtropical waters during calving season—we split the year into Winter (November-March) and Summer (April-October) seasonal models.

4.1   Winter

For the Winter (November-March) season, we initially attempted to model sei whale with a single density surface model that spanned the entire study area, but this model proved problematic because of a lack of survey effort beyond the shelf break. Instead, we split the study area at Cape Hatteras, an ecoregional boundary between distinct cetacean communities (Schick et al. 2011) where the Gulf Stream separates from the continental shelf and heads into the Atlantic, and we modeled the regions north and south of here separately.

In the north (Figure 25, region A), we fitted a full density surface model (Section 4.1.1). South of Cape Hatteras, along the upper shelf (depths <125 m), which was intensely surveyed by the North Atlantic Right Whale Early Warning System (NARW EWS) surveys since 2003 with no minke whale sightings (Figure 25, region D), we assumed the species was absent. (The NARW EWS surveys are not shown on this map.) In the region spanning the Blake Plateau and offshore waters for which some surveying was available and sightings were reported (Figure 25, region C), as well as acoustic detections (Kiehbadroudinezhad et al. 2021; Kowarski et al. 2022), we made a pseudo-stratified density estimate (Section 4.1.2). In the distant offshore region for which no surveying was available during the Winter months (Figure 25, region B), we did not estimate density, owing to a lack of data and our resulting low confidence in our ability to correctly guess whether minke whales are present in this area during these months.
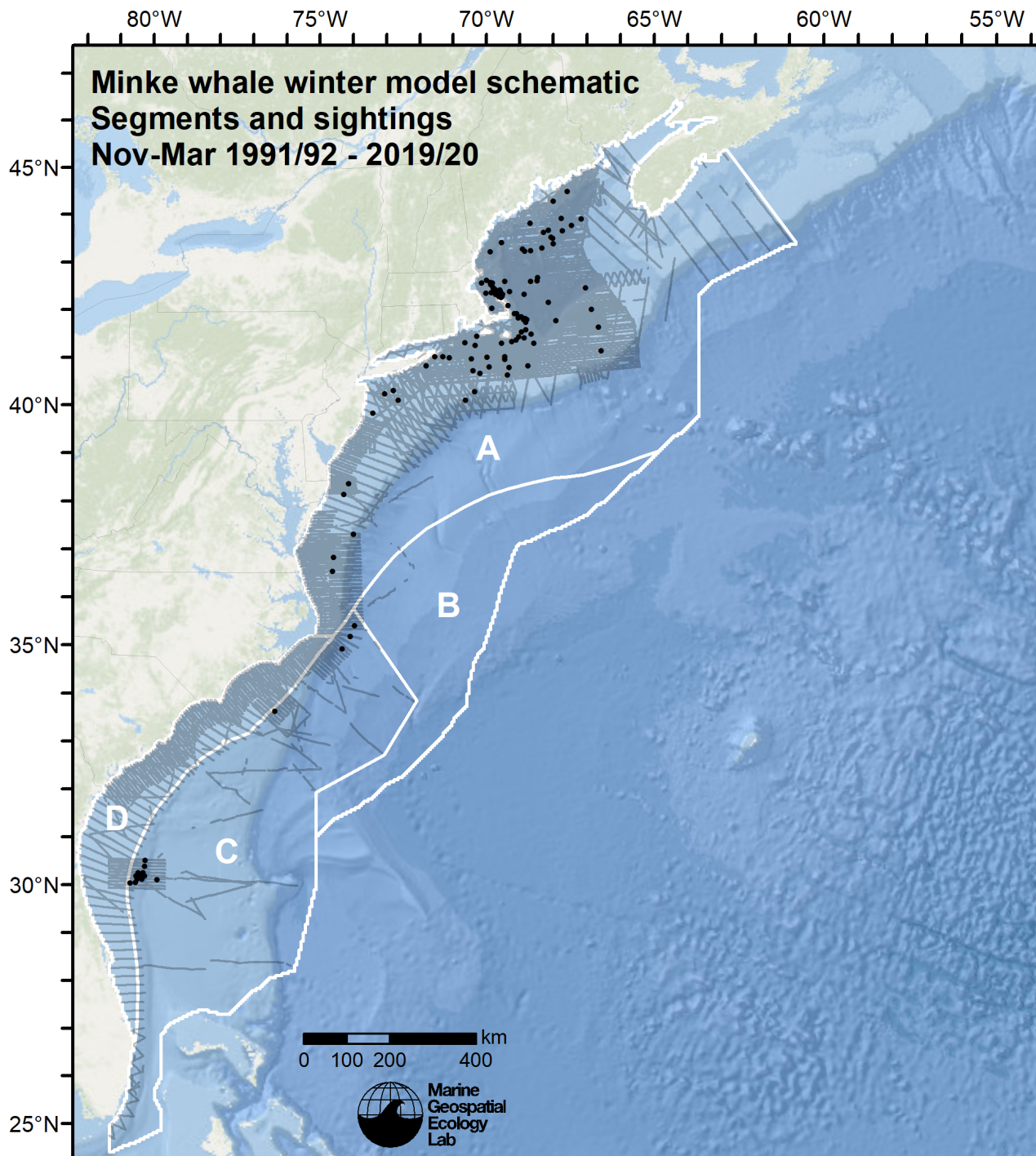
Figure 25: Schematic showing how we split the study area in Winter into regional models as described in the text above.

4.1.1 North of the Gulf Stream

The surveys incorporated into the model for waters north of the Gulf Stream, spanning November 1998 through March 2020, reported 112 sightings (Figure 26), scattered across the continental shelf from North Carolina to Maine. The highest concentration occurred north of Cape Cod Bay, but effort was also high in this region, so the concentration of sightings did not necessarily imply a concentration of density.

The model selection procedure was straightforward. When ranked by REML score (Wood 2011), models with climatological covariates outranked those with contemporaneous covariates. We selected as our final choice the climatological-covariate model with the best AIC score, which ranked a close second place by REML; this model explained 1.5% more deviance than the highest ranked contemporaneous-covariate model.

The model retained three static covariates and two dynamic covariates (Table 7). Fitted functional relationships were relatively simple for all covariates (Figure 29), likely owing to the relatively small number of sightings available for this model. The functional relationship for depth was negative, indicating decreasing density with increasing depth, while the relationship for distance to the 125 m isobath was positive, indicating increasing density moving down the slope to the isobath and beyond it. Finally, the relationship with fetch, a measure of how enclosed an area is, was hump-shaped, indicating higher density in the vicinity of land but lower density in both enclosed areas and wide-open areas, although confidence limits were quite wide for the lowest values, owing to sparse surveying of such enclosed areas.

Taken together, we find these three static relationships difficult to interpret, as they appear to contradict each other. Minkes whale are found in deep offshore waters, as exemplified by the sightings in the central Gulf of Maine, but they are also found in inshore, at least in summer, in areas such as the lower Piscataqua River (Zoidis et al. 2022), the entrance to Passamaquoddy Bay (Ingram et al. 2007), and within the Mingan Islands (Doniol-Valcroze et al. 2007). We suspect these relationships were fitted to reflect the high densities in the vicinity of Cape Cod as well as the low densities along the continental slope (Figure 27).

The two dynamic covariates were similarly simple. The relationship for sea surface temperature was negative, indicating decreasing density with increasing temperature, indicating a preference for the colder, on-shelf waters in the northern half of the modeled region. The relationship with $\log_{10}$-transformed primary productivity as estimated by the Carbon-based Productivity Model (CbPM) (Behrenfeld et al. 2005; Westberry et al. 2008) was hump-shaped, suggesting an avoidance of areas with the lowest productivity during these months, which included the western central Gulf of Maine and Great South Channel as well as waters beyond the shelf break, and those with the highest, which included the mid-Atlantic shelf.
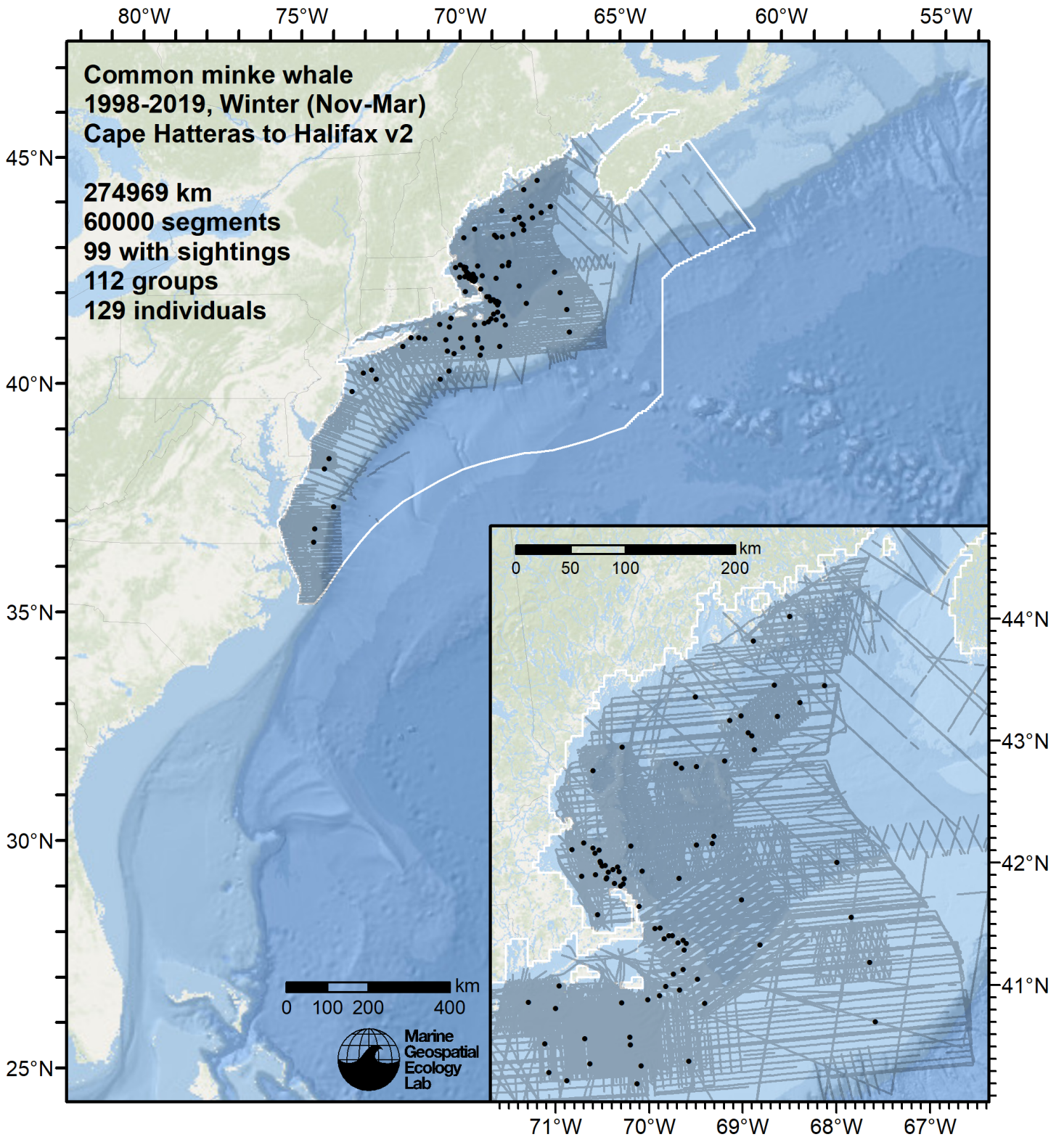
Figure 26: Survey segments used to fit the model for the region Cape Hatteras to Halifax v2 for Winter. Black points indicate segments with observations.
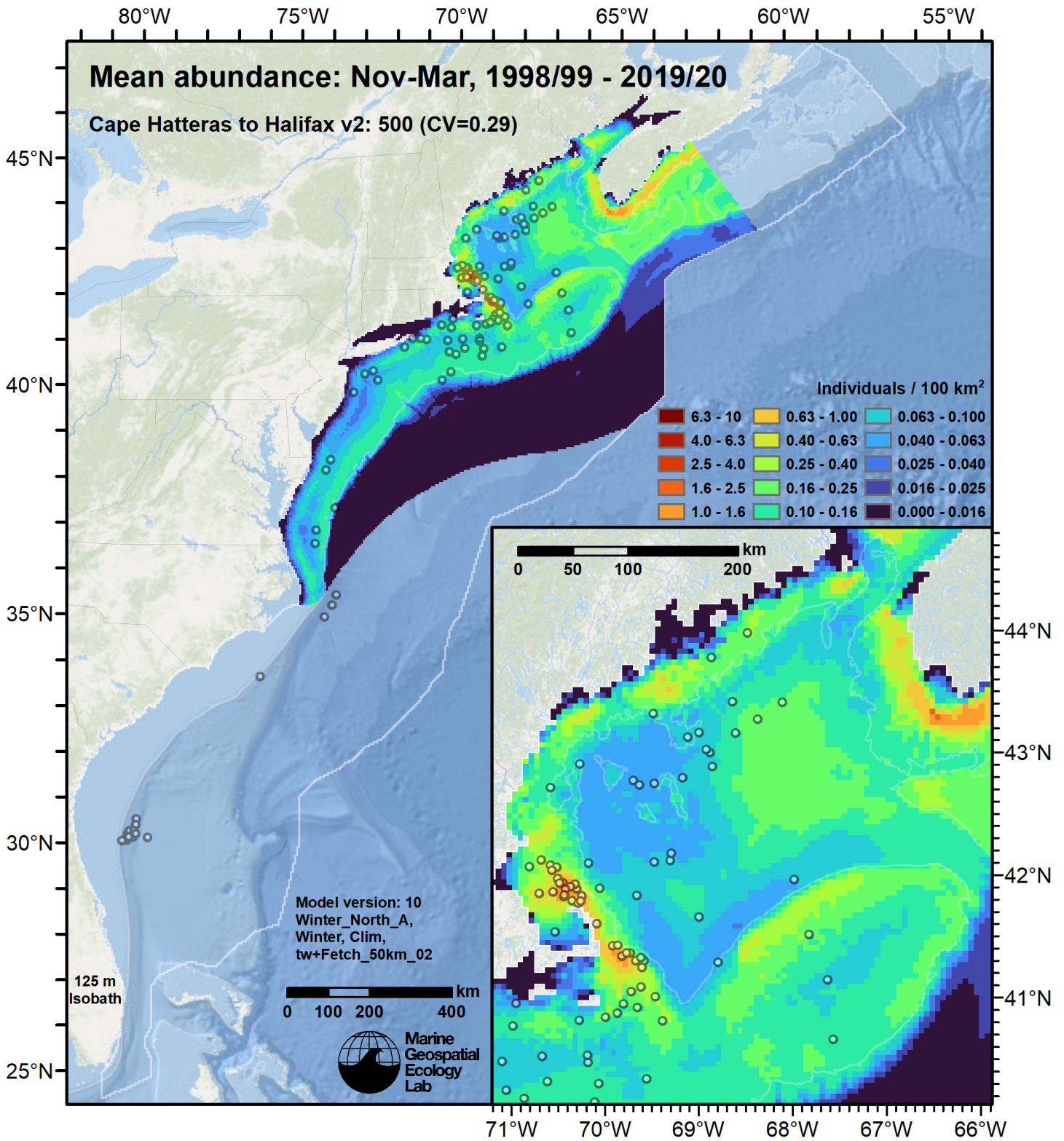
Figure 27: Common minke whale mean density for the indicated period, as predicted by the model for the region Cape Hatteras to Halifax v2 for Winter. Open circles indicate segments with observations. Mean total abundance and its coefficient of variation (CV) are given in the subtitle. Variance was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for seasonal variability in dynamic covariates but not interannual variability in them, as these covariates were monthly climatological averages.
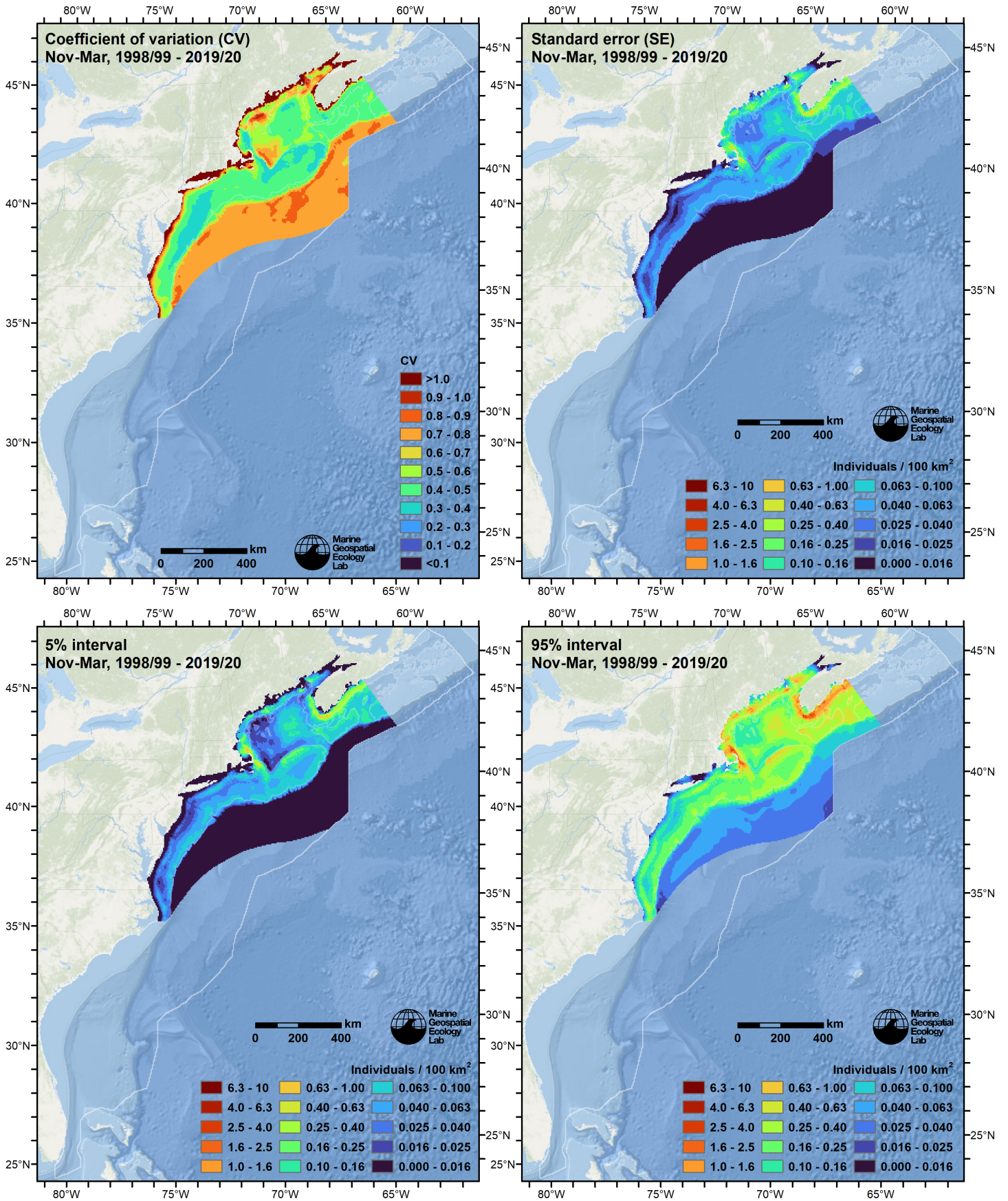
Figure 28: Uncertainty statistics for the common minke whale mean density surface (Figure 27) predicted by the model for the region Cape Hatteras to Halifax v2 for Winter. Variance was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for seasonal variability in dynamic covariates but not interannual variability in them, as these covariates were monthly climatological averages.

Statistical output for this model:

```
Family: Tweedie(p=1.06)
Link function: log

Formula:
IndividualsCorrected ~ offset(log(SegmentArea)) + s(log10(pmax(5,
    pmin(Depth, 2000)))), bs = "ts") + s(pmax(-120, pmin(I(DistTo125m/1000),
    60)), bs = "ts") + s(Fetch_50km, bs = "ts") + s(pmax(3, pmin(ClimSST_CMC,
    19)), bs = "ts") + s(log10(pmax(35, pmin(ClimPP_CbPM, 3000)))),
    bs = "ts")

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -21.0959     0.1494  -141.2   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:
                                             edf Ref.df     F  p-value
s(log10(pmax(5, pmin(Depth, 2000))))       1.024      9 1.570 5.84e-05 ***
s(pmax(-120, pmin(I(DistTo125m/1000), 60))) 1.224     9 2.860  < 2e-16 ***
s(Fetch_50km)                              3.450      9 2.644 1.13e-05 ***
s(pmax(3, pmin(ClimSST_CMC, 19)))          0.955      9 1.171 0.000663 ***
s(log10(pmax(35, pmin(ClimPP_CbPM, 3000)))) 3.286     9 4.159  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) =  0.00703   Deviance explained = 9.61%
-REML = 861.58  Scale est. = 5.9601     n = 60000

Method: REML   Optimizer: outer newton
full convergence after 12 iterations.
Gradient range [-8.120995e-05,6.336521e-05]
(score 861.578 & scale 5.960094).
Hessian positive definite, eigenvalue range [0.2922932,1655.749].
Model rank =  46 / 46

Basis dimension (k) checking results. Low p-value (k-index<1) may
indicate that k is too low, especially if edf is close to k'.

                                              k'    edf k-index p-value
s(log10(pmax(5, pmin(Depth, 2000))))       9.000 1.024    0.93    0.46
s(pmax(-120, pmin(I(DistTo125m/1000), 60))) 9.000 1.224   0.94    0.99
s(Fetch_50km)                              9.000 3.450    0.93    0.85
s(pmax(3, pmin(ClimSST_CMC, 19)))          9.000 0.955    0.92    0.05 *
s(log10(pmax(35, pmin(ClimPP_CbPM, 3000)))) 9.000 3.286   0.92    0.10 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(a) Climatological net primary productivity (mg C m^{-2} day^{-1}) (CbPM)

(b) Climatological sea surface temperature (°C)

(c) Seafloor depth (m)

(d) Distance to 125m isobath (km)
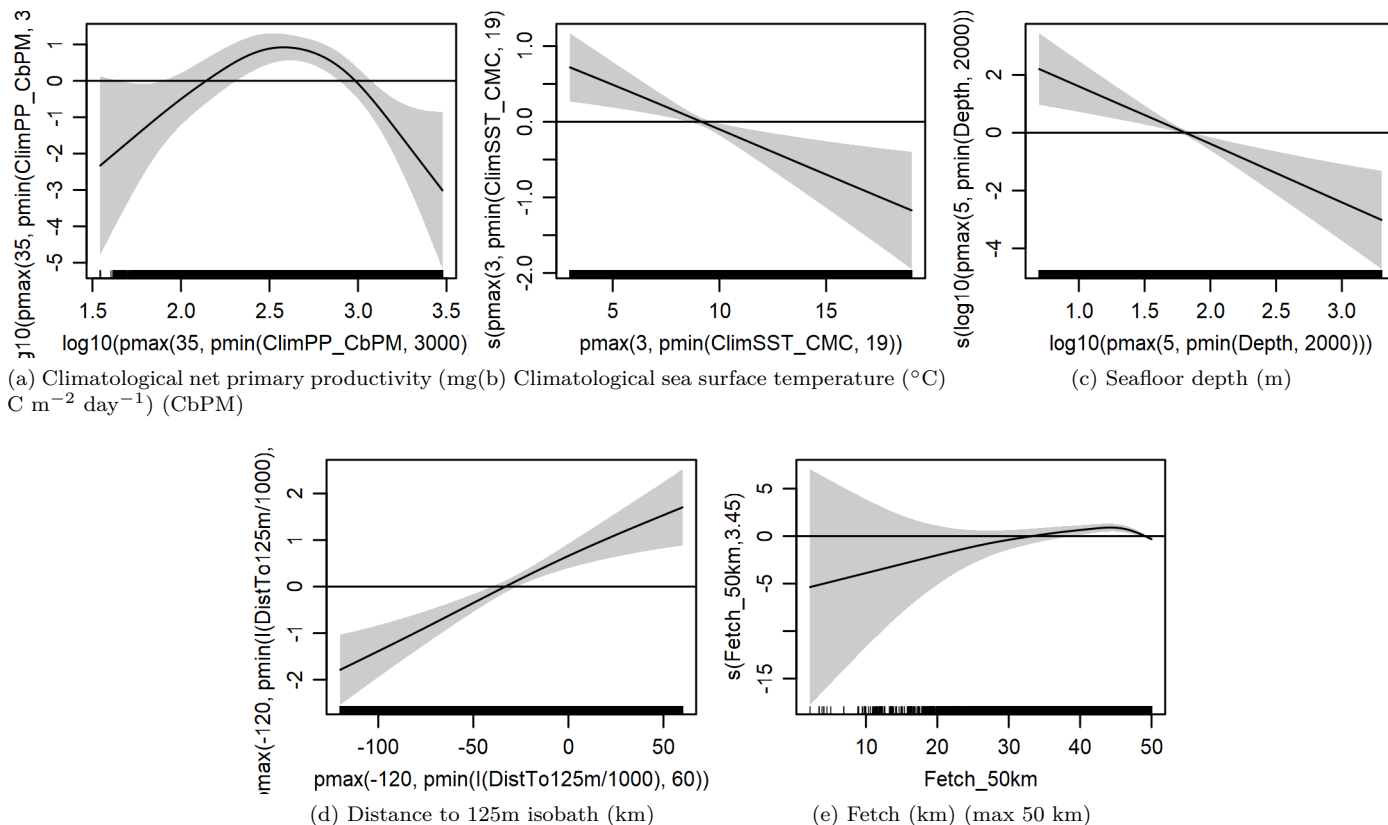
(e) Fetch (km) (max 50 km)

Figure 29: Functional plots for the final model for the region Cape Hatteras to Halifax v2 for Winter. Transforms and other treatments are indicated in axis labels. *log10* indicates the covariate was $\log_{10}$ transformed. *pmax* and *pmin* indicate the covariate's minimum and maximum values, respectively, were Winsorized to the values shown. Winsorization was used to prevent runaway extrapolations during prediction when covariates exceeded sampled ranges, or for ecological reasons, depending on the covariate. */1000* indicates meters were transformed to kilometers for interpretation convenience.

Table 7: Covariates used in the final model for the region Cape Hatteras to Halifax v2 for Winter.

| Covariate	Description
ClimPP_CbPM	Climatological monthly mean net primary productivity (mg C m^{-2} day^{-1}) from the Carbon-based Production Model (CbPM) (Behrenfeld et al. (2005); Westberry et al. (2008))
ClimSST_CMC	Climatological monthly mean sea surface temperature (°C) from GHRSST Level 4 CMC0.2deg and CMC0.1deg (Brasnett (2008); Canada Meteorological Center (2012); Meissner et al. (2016); Canada Meteorological Center (2016))
Depth	Depth (m) of the seafloor, from SRTM30_PLUS (Becker et al. (2009))
DistTo125m	Distance (km) to the 125m isobath, derived from SRTM30_PLUS (Becker et al. (2009))
Fetch_50km	Fetch (km): mean distance to shore averaged over 16 radial directions, limited to a maximum of 50 km

4.1.1.2   Diagnostic Plots
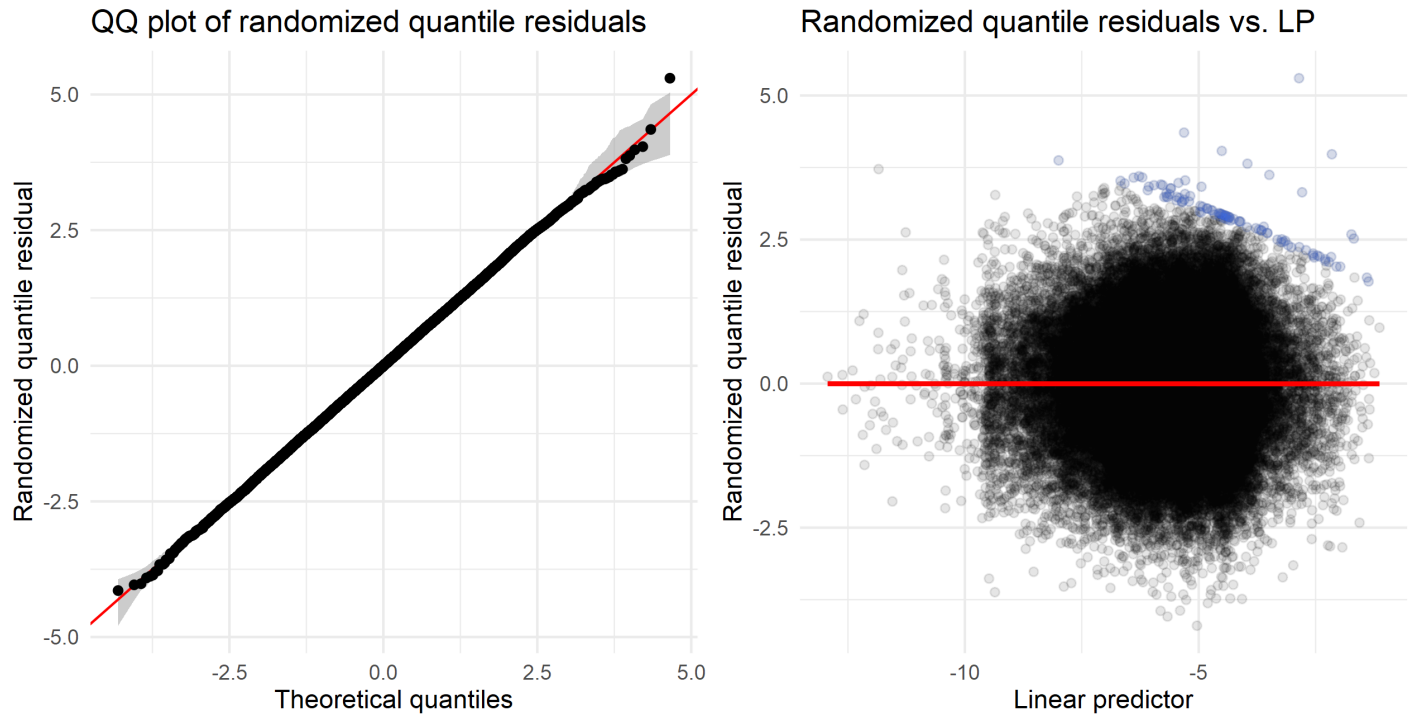
Figure 30: Residual plots for the final model for the region Cape Hatteras to Halifax v2 for Winter.
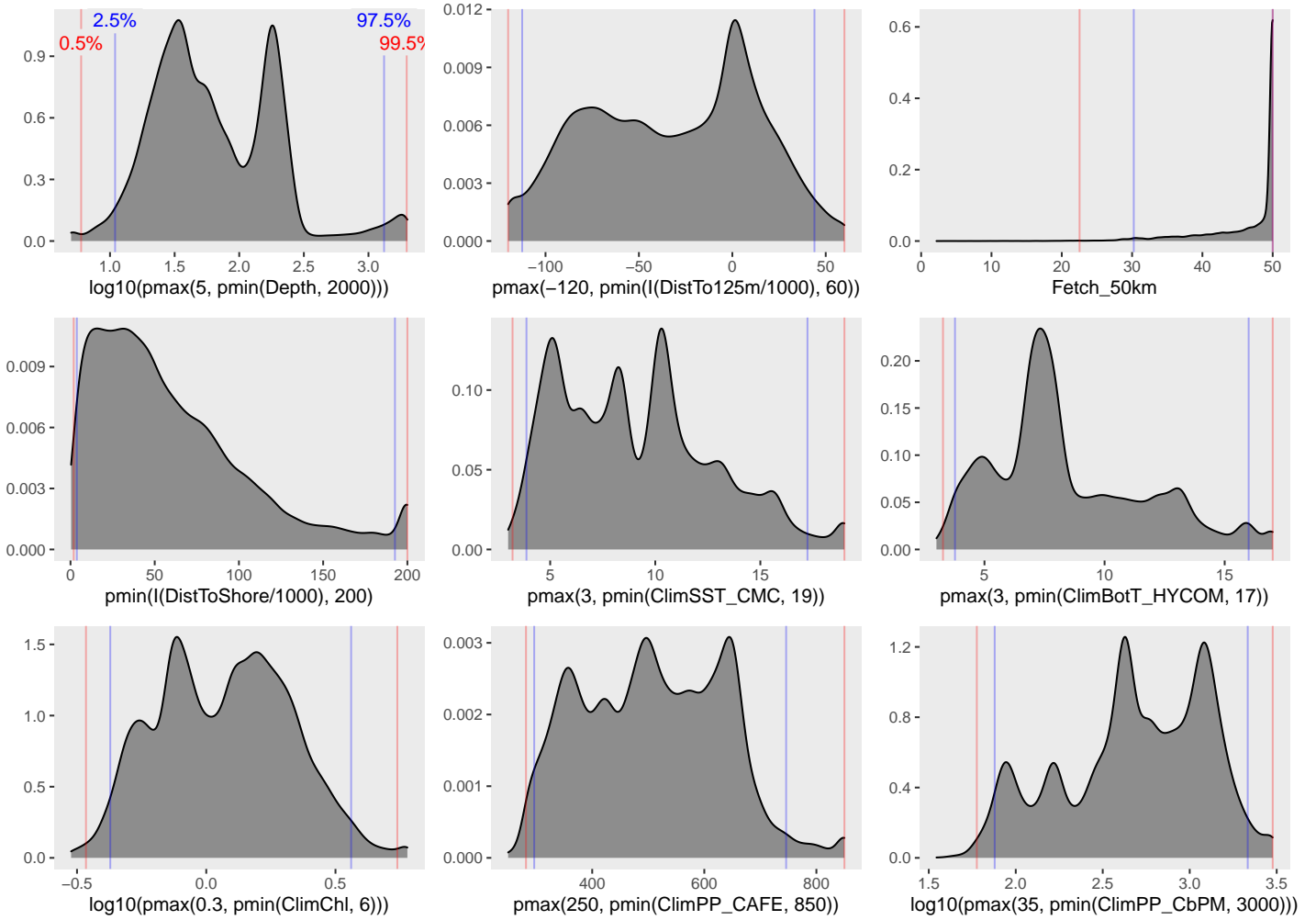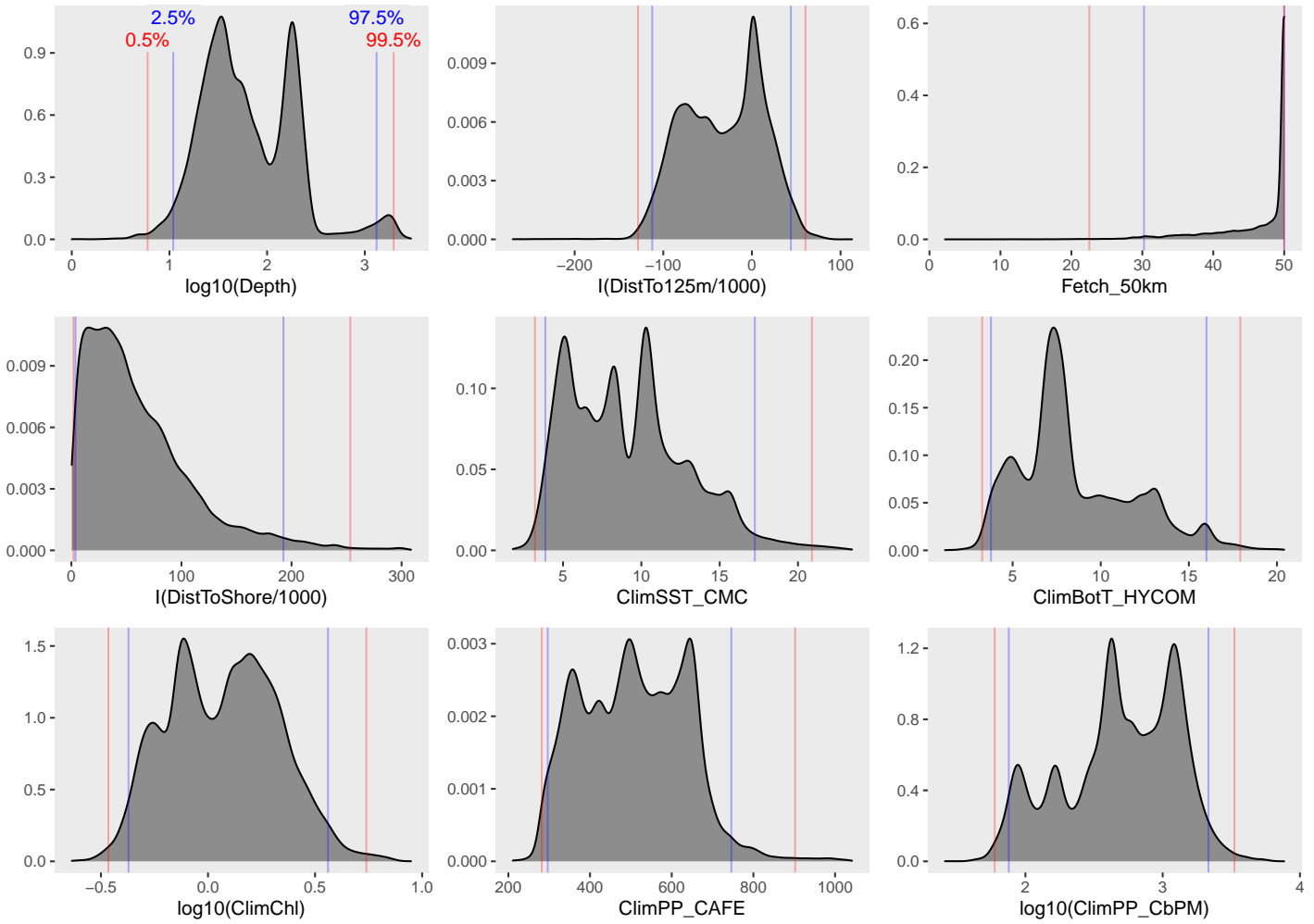
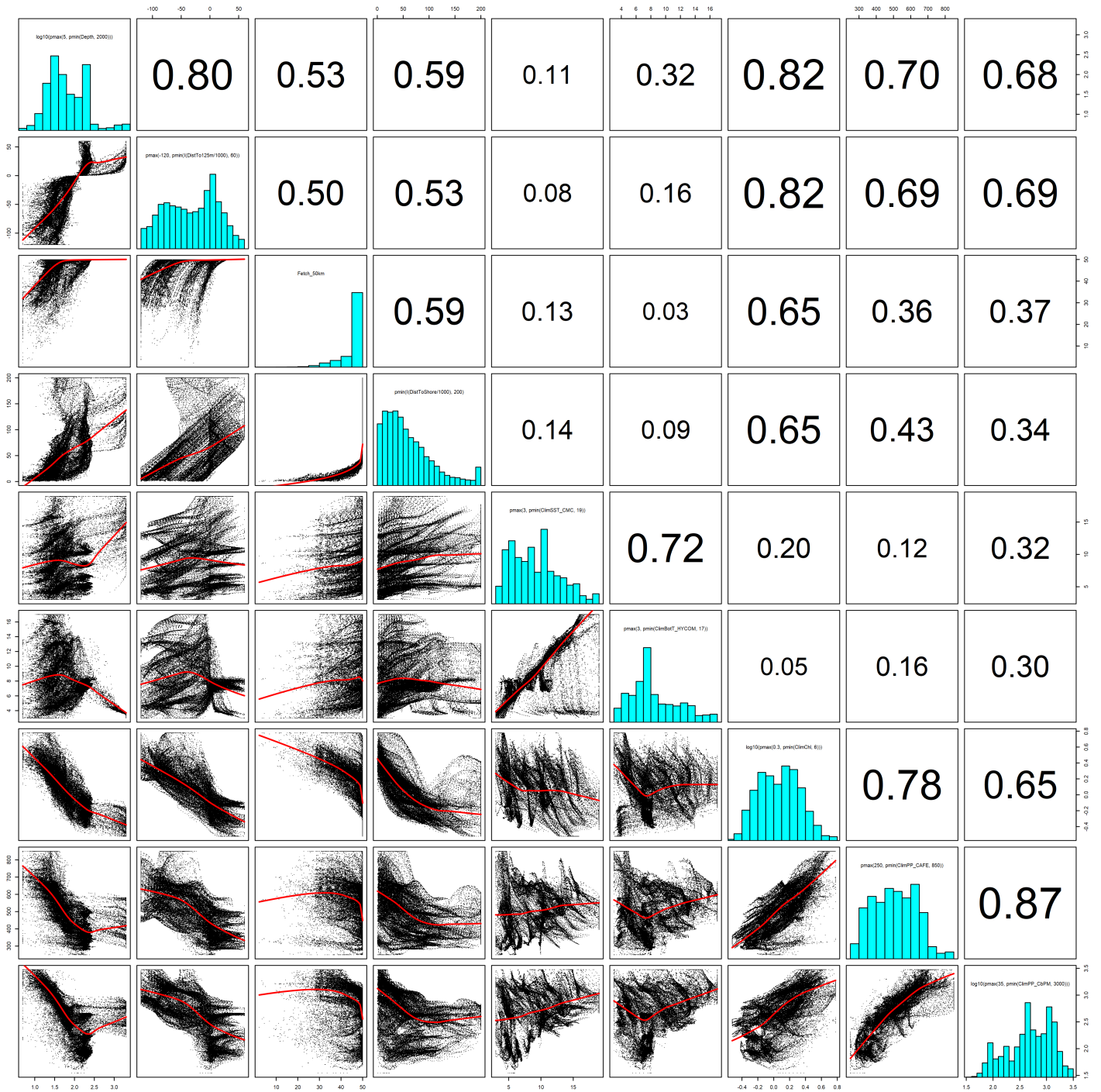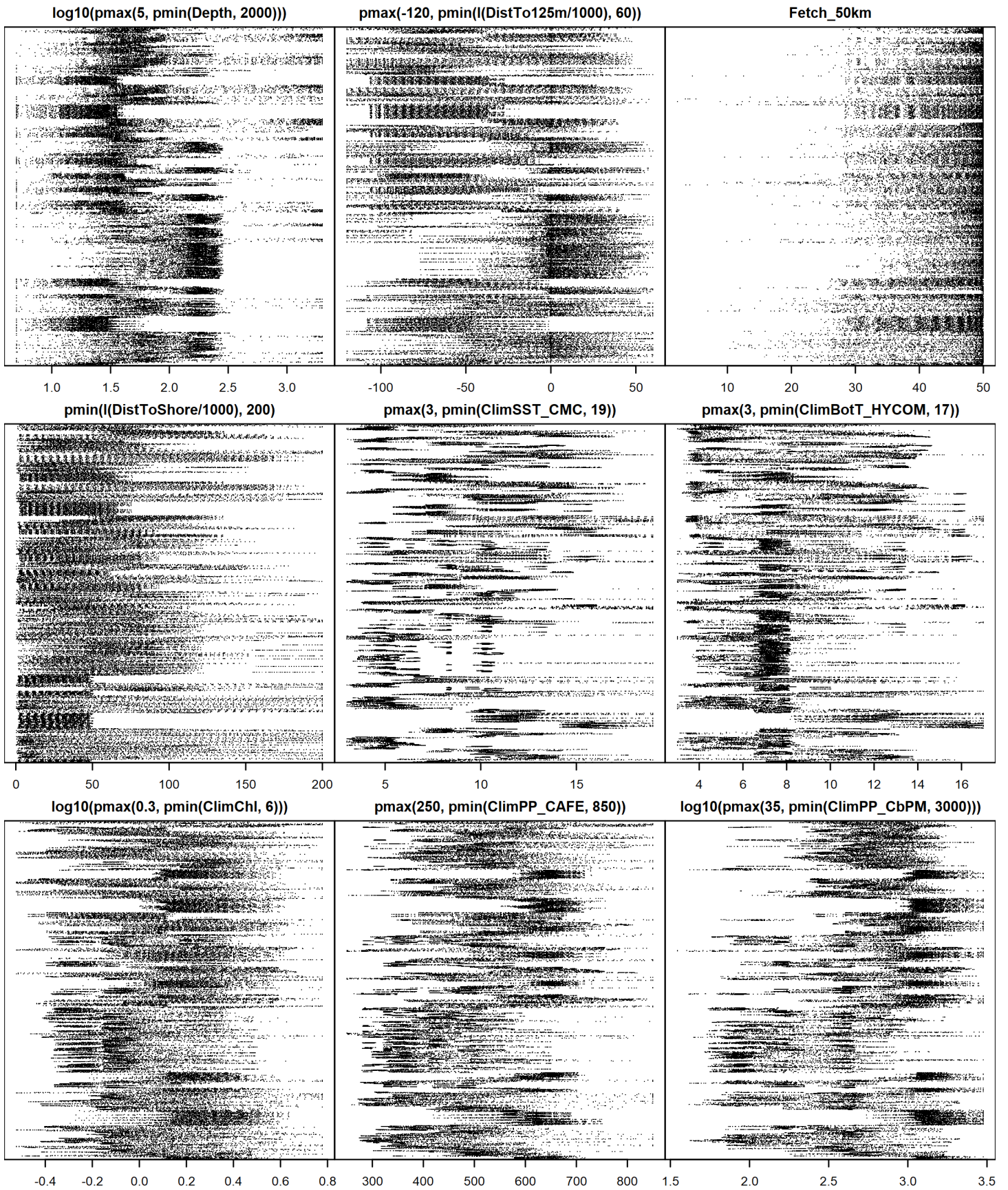Figure 31: Density histograms showing the distributions of the covariates considered during the final model selection step. The final model may have included only a subset of the covariates shown here (see Figure 29), and additional covariates may have been considered in preceding selection steps. Red and blue lines enclose 99% and 95% of the distributions, respectively. Transforms and other treatments are indicated in axis labels. *log10* indicates the covariate was $\log_{10}$ transformed. *pmax* and *pmin* indicate the covariate's minimum and maximum values, respectively, were Winsorized to the values shown. Winsorization was used to prevent runaway extrapolations during prediction when covariates exceeded sampled ranges, or for ecological reasons, depending on the covariate. */1000* indicates meters were transformed to kilometers for interpretation convenience.

Figure 32: Density histograms shown in Figure 31 replotted without Winsorization, to show the full range of sampling represented by survey segments.

Figure 33: Scatterplot matrix of the covariates considered during the final model selection step. The final model may have included only a subset of the covariates shown here (see Figure 29), and additional covariates may have been considered in preceding selection steps. Covariates are transformed and Winsorized as shown in Figure 31. This plot is used to check simple correlations between covariates (via pairwise Pearson coefficients above the diagonal) and visually inspect for concurvity (via scatterplots and red lowess curves below the diagonal).

Figure 34: Dotplot of the covariates considered during the final model selection step. The final model may have included only a subset of the covariates shown here (see Figure 29), and additional covariates may have been considered in preceding selection steps. Covariates are transformed and Winsorized as shown in Figure 31. This plot is used to check for suspicious patterns and outliers in the data. Points are ordered vertically by segment ID, sequentially in time.

4.1.1.3 Extrapolation Diagnostics

4.1.1.3.1 Univariate Extrapolation

(a) Depth covariate
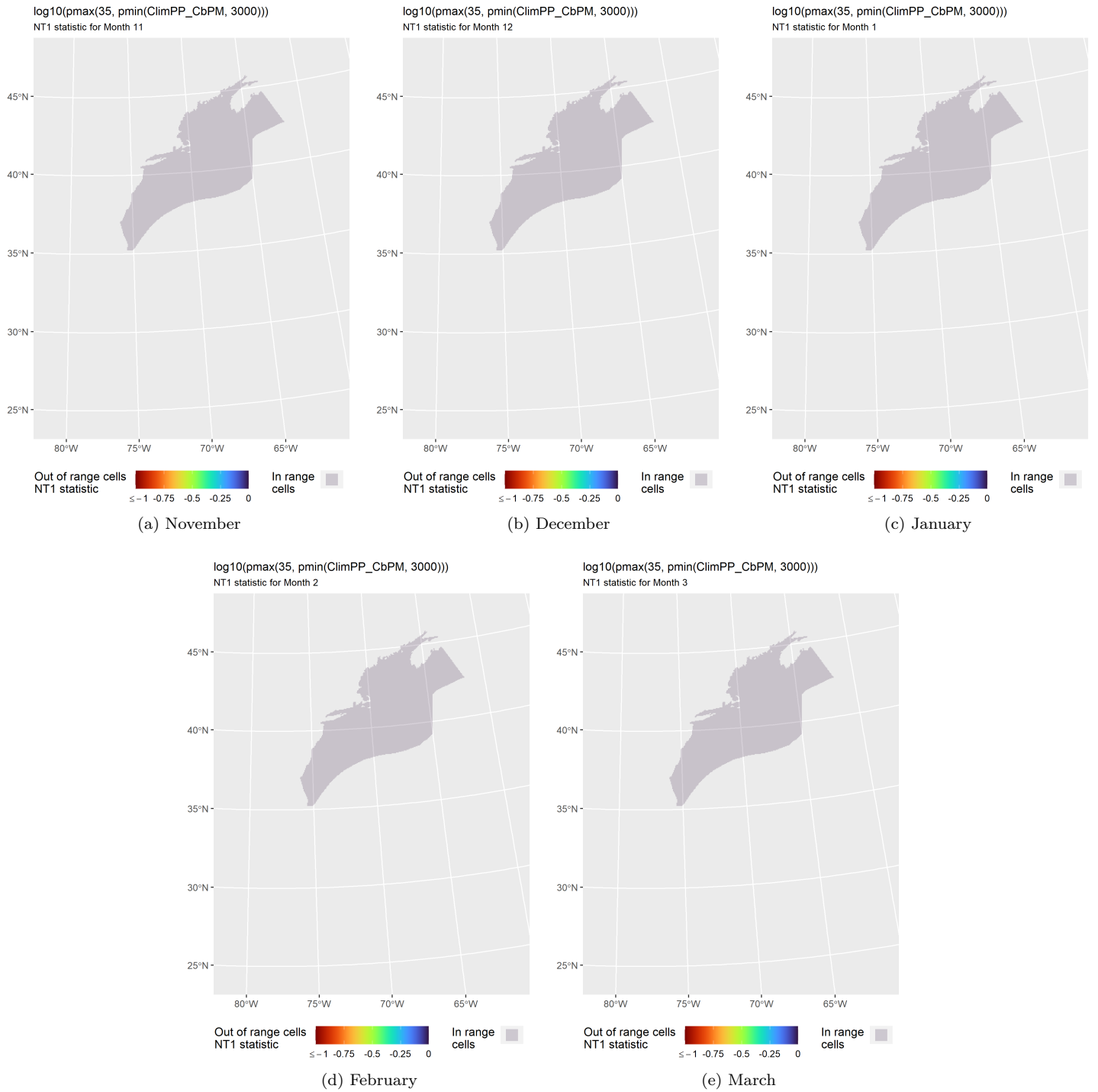

(b) DistTo125m covariate

(c) Fetch_50km covariate
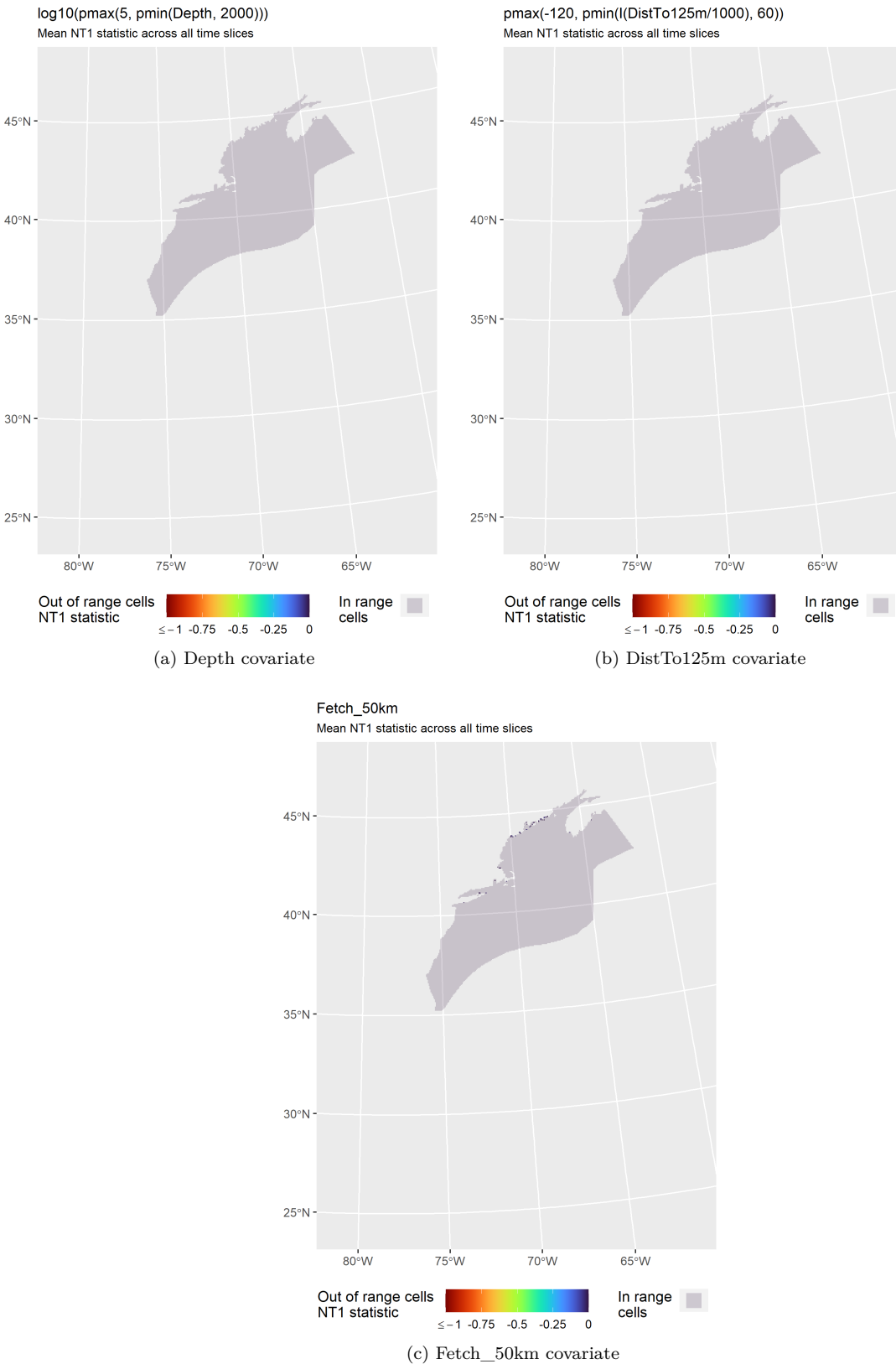
Figure 35: NT1 statistic (Mesgaran et al. (2014)) for static covariates used in the model for the region Cape Hatteras to Halifax v2 for Winter. Areas outside the sampled range of a covariate appear in color, indicating univariate extrapolation of that covariate occurred there. Areas within the sampled range appear in gray, indicating it did not occur.

(a) November

(b) December

(c) January

(d) February

(e) March

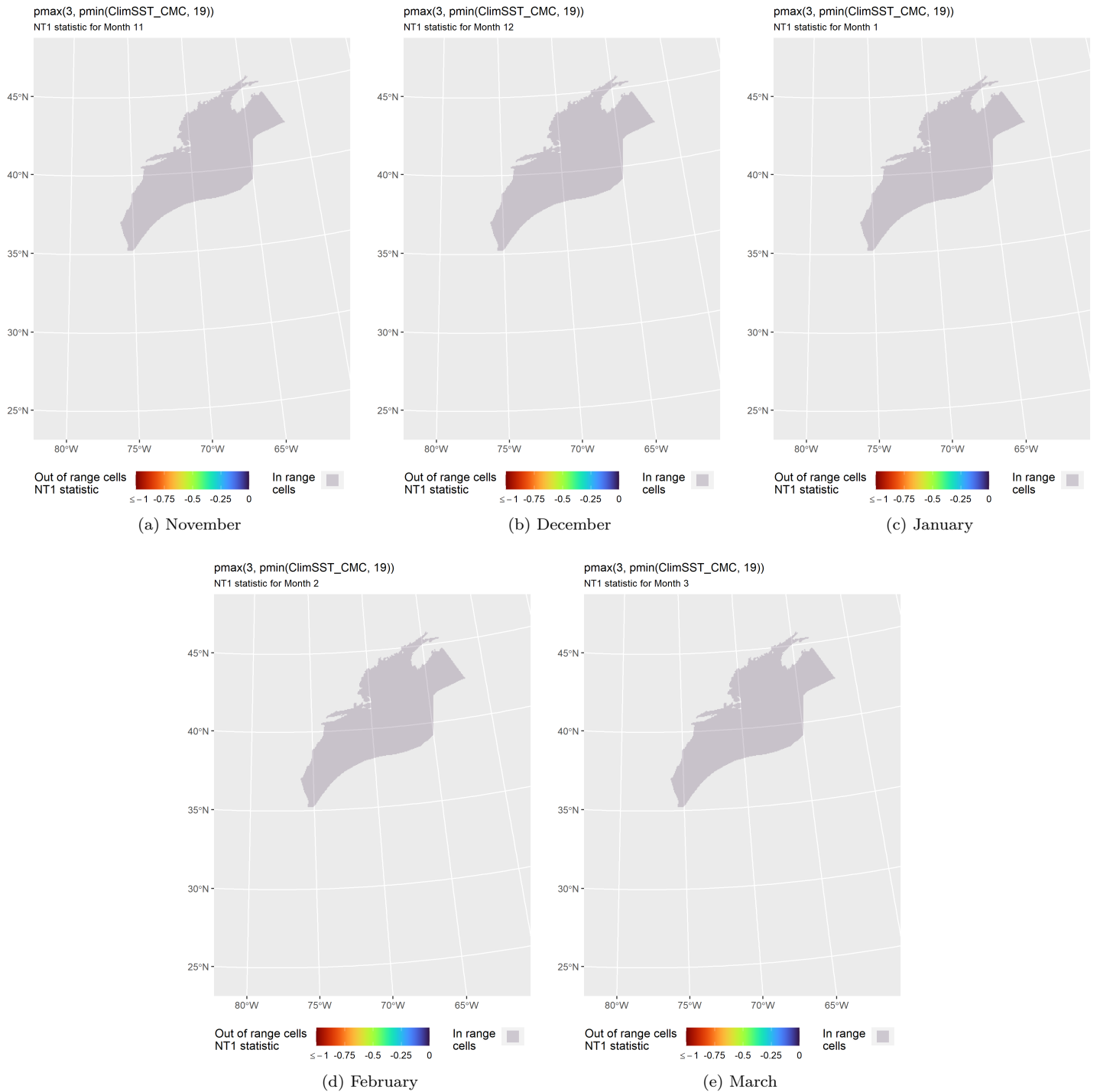Figure 36: NT1 statistic (Mesgaran et al. (2014)) for the ClimPP_CbPM covariate in the model for the region Cape Hatteras to Halifax v2 for Winter. Areas outside the sampled range of a covariate appear in color, indicating univariate extrapolation of that covariate occurred there during the month. Areas within the sampled range appear in gray, indicating it did not occur.

Figure 37: NT1 statistic (Mesgaran et al. (2014)) for the ClimSST_CMC covariate in the model for the region Cape Hatteras to Halifax v2 for Winter. Areas outside the sampled range of a covariate appear in color, indicating univariate extrapolation of that covariate occurred there during the month. Areas within the sampled range appear in gray, indicating it did not occur.

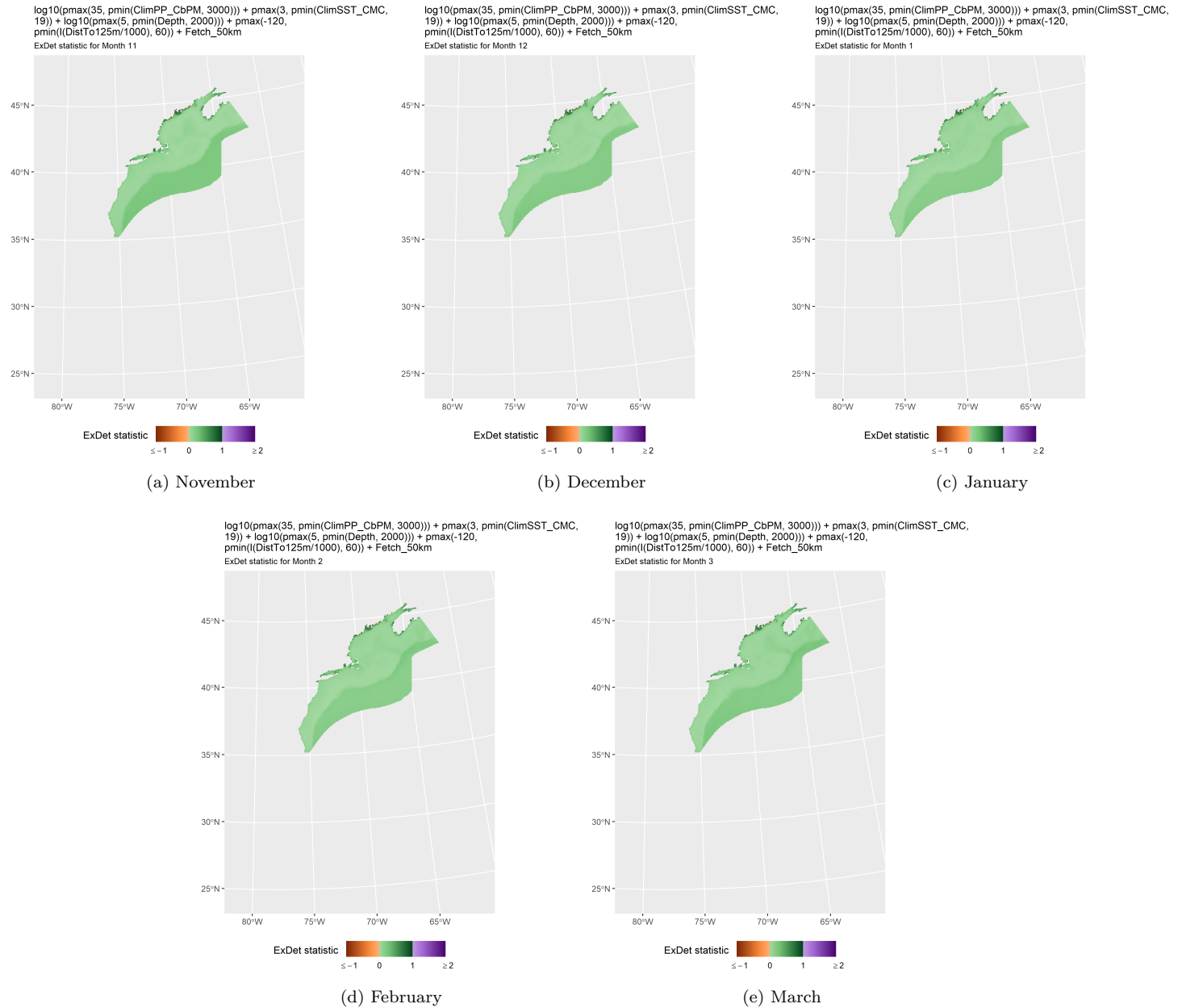#### 4.1.1.3.2   Multivariate Extrapolation

(a) November

(b) December

(c) January

(d) February

(e) March

Figure 38: ExDet statistic (Mesgaran et al. (2014)) for all of the covariates used in the model for the region Cape Hatteras to Halifax v2 for Winter. Areas in orange (ExDet < 0) required univariate extrapolation of one or more covariates (see previous section). Areas in purple (ExDet > 1), did not require univariate extrapolation but did require multivariate extrapolation, by virtue of having novel combinations of covariates not represented in the survey data, according to the NT2 statistic (Mesgaran et al. (2014)). Areas in green (0 ≥ ExDet ≤ 1) did not require either type of extrapolation.

4.1.2   South of the Gulf Stream

With so few sightings, it was not possible to fit a traditional density surface model that related density observed on survey segments to environmental covariates. Nor was it possible to make a proper design-based abundance estimate using traditional distance sampling (Buckland et al. 2001), because the aggregate surveys provided heterogeneous coverage that did not together constitute a proper systematic survey design.

To provide interested parties with at least a rough estimate of density for the region, we fitted a model with no covariates, under the assumption that density was distributed uniformly within it. This assumption, if true, would mean we could obtain similar density estimates under any sampling design, and therefore it would not matter if there was some heterogeneity in sampling. However, we strongly caution that this assumption of uniform density may not be correct, based on recent passive

acoustic monitoring (Kiehbadroudinezhad et al. 2021). We discuss this further in Section 6. Additional visual surveying is required to properly model how minke whale density varies spatially. Until such data have been collected, we offer this simplified approach as a rough-and-ready substitute for a full density surface model.

In this section, we present a map and tallies of the survey effort and sightings (Figure 39). Section 5.1 gives the resulting map of density and the total abundance and coefficient of variation.
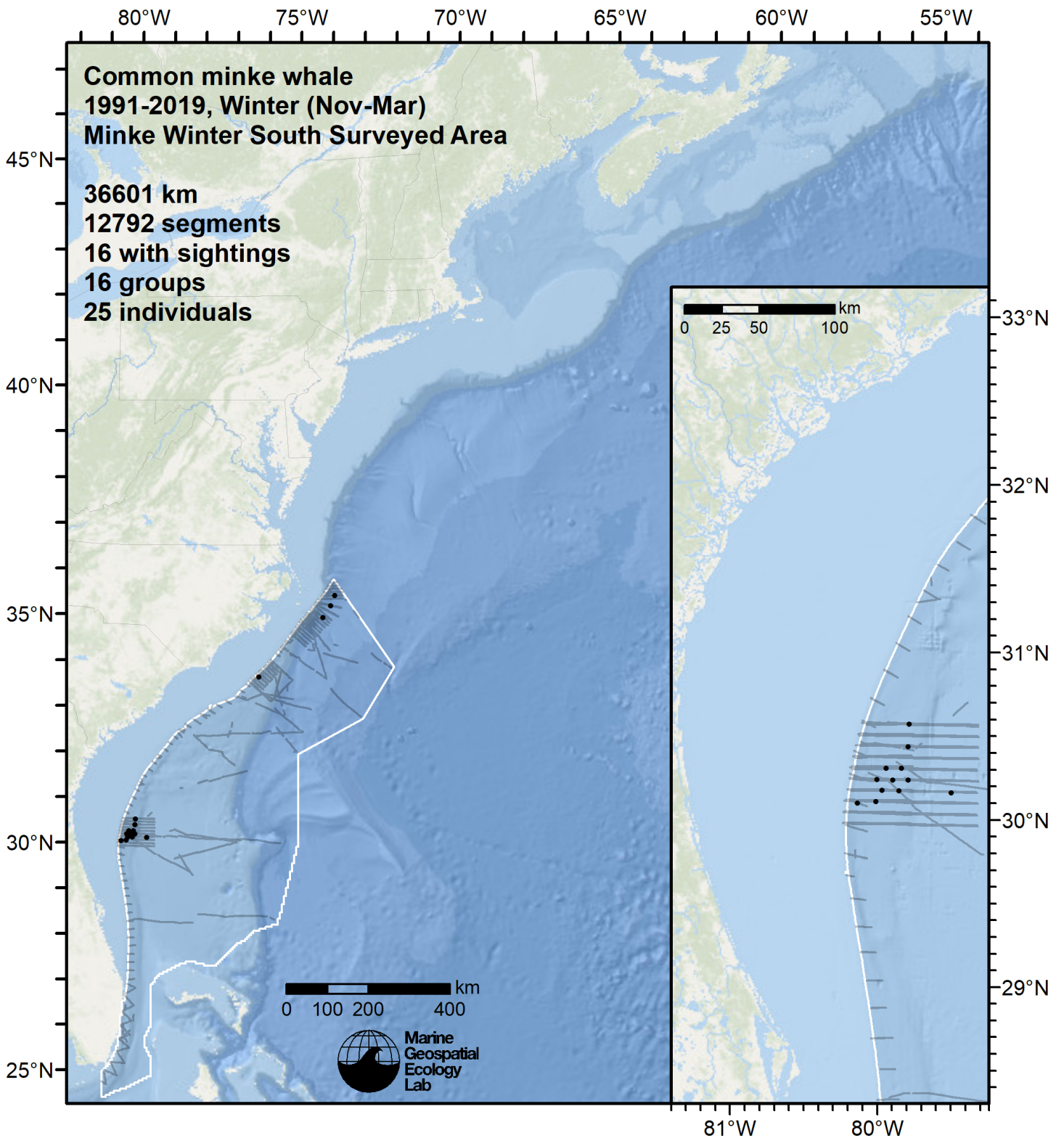
Figure 39: Survey segments and sightings used to estimate common minke whale density for the off-shelf region south of Gulf Stream. Black points indicate segments with observations.

4.2 Summer

For Summer (April-October), we fitted a single model spanning the entire study area except the upper shelf (depths <125 m) south of Cape Hatteras. On the upper shelf, we assumed density was zero, on the basis of no sightings having occurred there throughout the study period (even opportunistically, even in winter), and it being poor summer feeding habitat for minke whales.

The surveys incorporated into the model reported over 1600 sightings (Figure 40). Most were scattered across the continental shelf, from North Carolina to the northernmost extent of surveying, including the upper Bay of Fundy and the northeastern Scotian Shelf. But a number of sightings also occurred over the continental slope, and two occurred far beyond it, one in the Gulf Stream in July 1998 in waters 3500 m deep and another south of the Gulf Stream in July 2016 in waters 4160 m deep, both with a surface temperature of 26 °C. Given these sightings, and that infrequent acoustic detections had been reported over the Blake Plateau in April 2018-2019 (Kiehbadroudinezhad et al. 2021), and at Mid-Atlantic Ridge sites as late as June in 2000-2001 (Nieukirk et al. 2004; Risch et al. 2014), we deemed it appropriate to include off-shelf habitat south of the Gulf Stream in our model's spatial extent, rather than assume minke whales were entirely absent south of the Gulf Stream.

The model retained two static covariates and four dynamic covariates (Table 8). The relationship fitted for depth indicated a positive influence on density for a relatively shallow range of depths, about 13-126 m (or 1.1-2.1 in $\log_{10}$ scale) (Figure 43), while for fetch a positive relationship was shown for locations for which distance to land in all directions averaged about 32-48 km. Together these relationships suggest a preference for shallow and nearshore waters, which may be seen in the mean density prediction (Figure 41) over Georges Bank, around the Gulf of Maine, and along the Scotian Shelf. Predictions for this species extended further into bays, sounds, and other enclosed areas than our models for other baleen whale species, consistent with previous reports of minke whales foraging in nearshore areas (Doniol-Valcroze et al. 2007; Ingram et al. 2007; Zoidis et al. 2022).

The relationship fitted for sea surface temperature (SST) showed a positive influence on density between 6-21 °C, a relatively wide range, while the relationship for bottom temperature was negative, indicating an avoidance of locations with warm bottom temperatures, as occurs in the nearshore mid-Atlantic during summer and early fall, where no minke sightings were reported. The relationship fitted for distance to SST fronts was negative, indicating higher density closer to fronts, consistent with the finding of Ingram et al. (2007) who reported an association between minkes and tidal fronts near Campobello and Grand Manan Islands in the lower Bay of Fundy. (However, we note that Doniol-Valcroze et al. (2007) failed to detect a statistically-significant effect of the distribution of SST fronts on the distribution of minkes at the Mingan Islands, once autocorrelation was accounted for.) Finally, the relationship for biomass of epipelagic micronekton, representing likely prey for minke whales, showed a hump at medium to high values and a negative influence at the lowest values, generally consistent with what would be expected for this covariate.
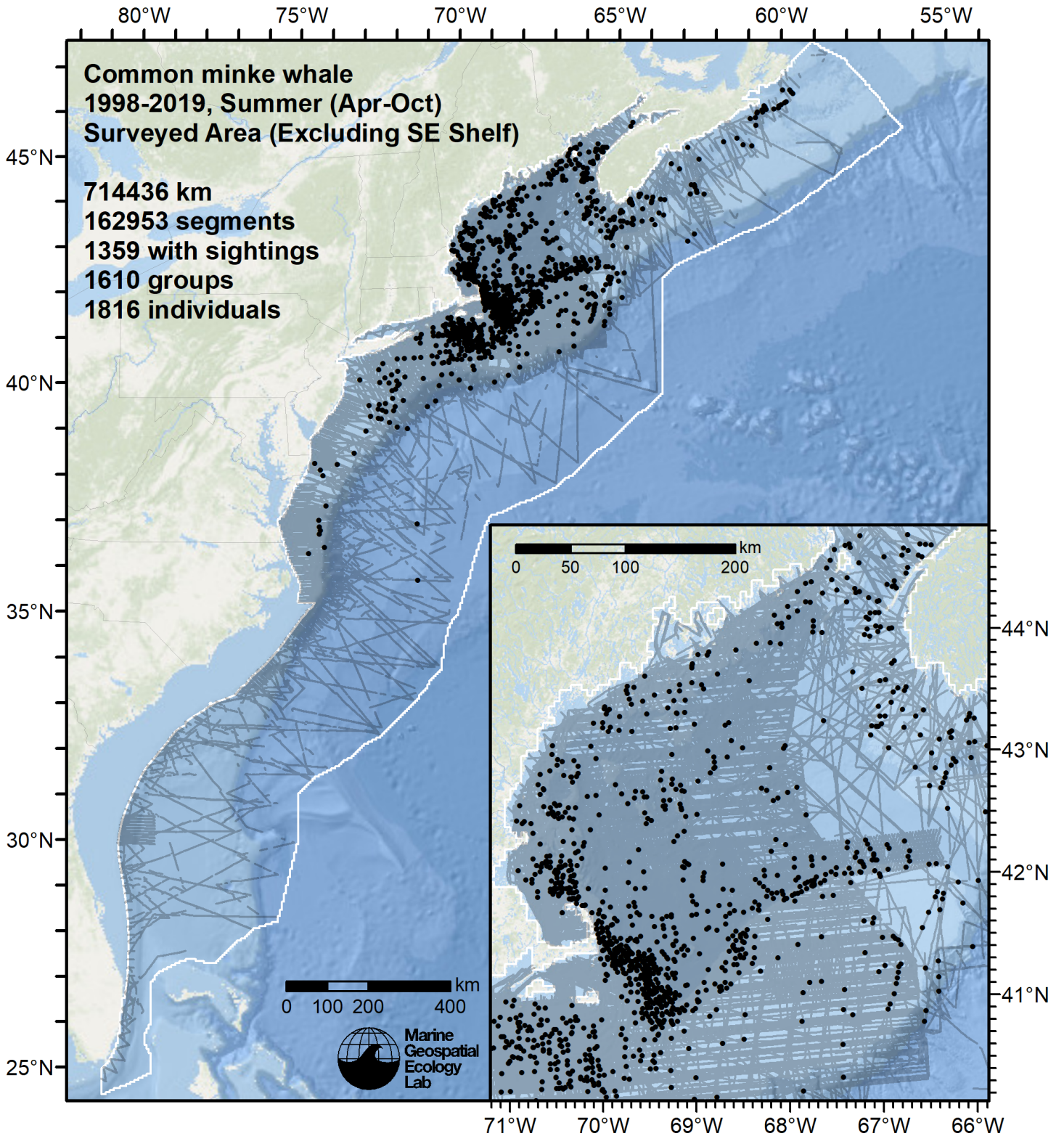
Figure 40: Survey segments used to fit the model for the region Surveyed Area (Excluding SE Shelf) for Summer. Black points indicate segments with observations.
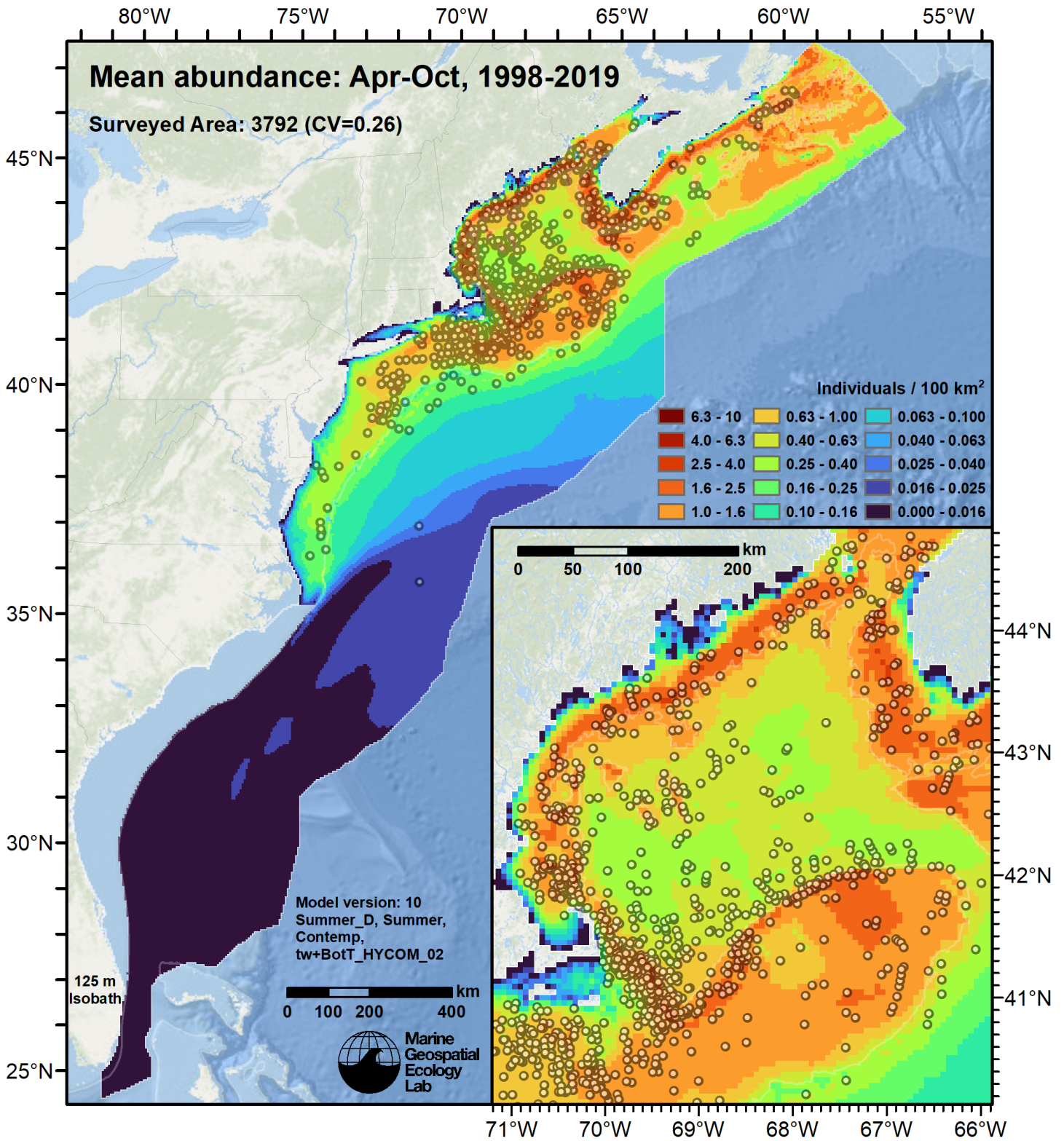
Figure 41: Common minke whale mean density for the indicated period, as predicted by the model for the region Surveyed Area (Excluding SE Shelf) for Summer. Open circles indicate segments with observations. Mean total abundance and its coefficient of variation (CV) are given in the subtitle. Variance was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for seasonal and interannual variability in dynamic covariates.
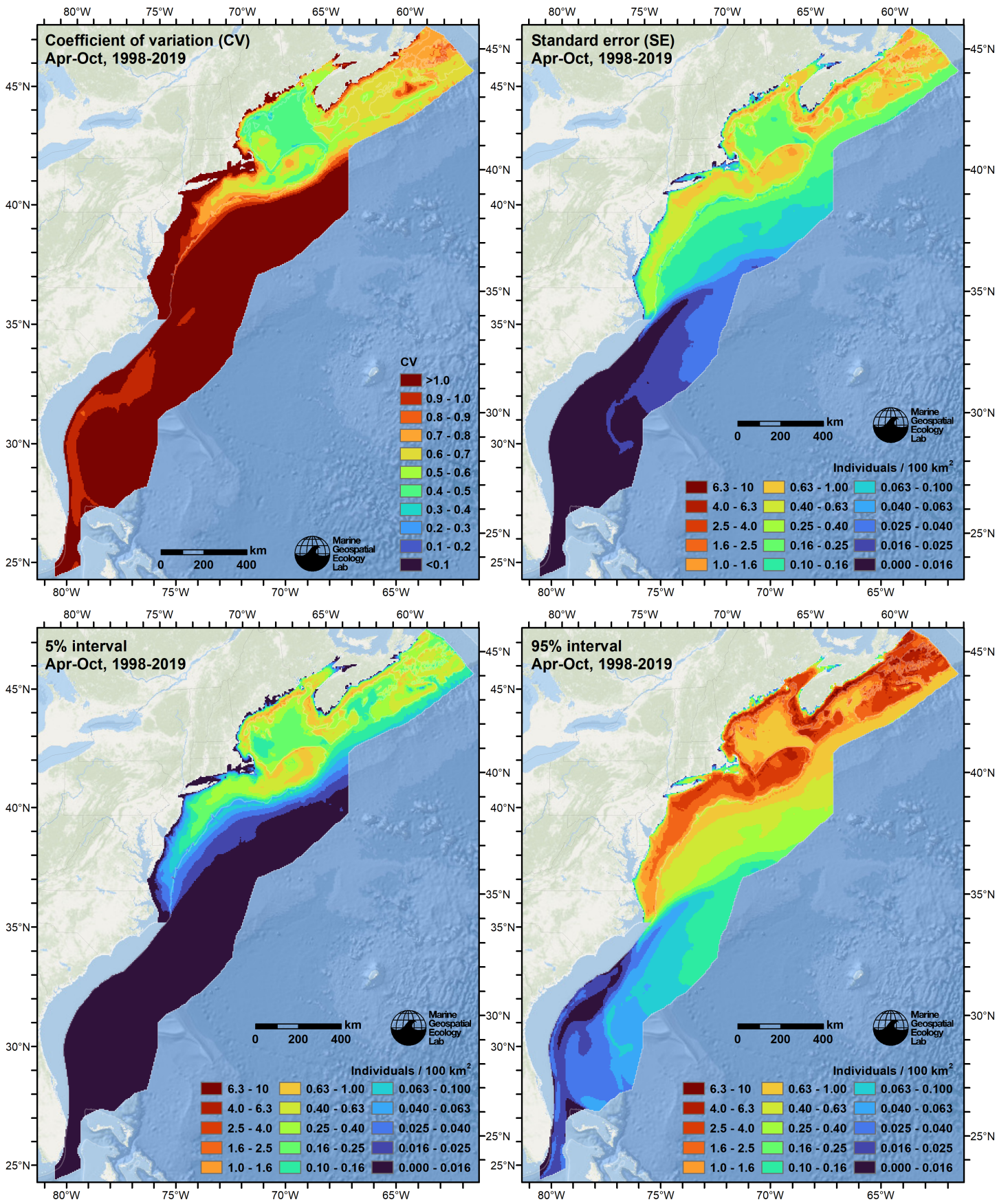
Figure 42: Uncertainty statistics for the common minke whale mean density surface (Figure 41) predicted by the model for the region Surveyed Area (Excluding SE Shelf) for Summer. Variance was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for seasonal and interannual variability in dynamic covariates.

Statistical output for this model:

```
Family: Tweedie(p=1.126)
Link function: log

Formula:
IndividualsCorrected ~ offset(log(SegmentArea)) + s(Fetch_50km,
    bs = "ts") + s(log10(pmax(3, pmin(Depth, 4000))), bs = "ts") +
    s(pmax(3, pmin(SST_CMC, 30)), bs = "ts") + s(pmax(3, pmin(BotT_HYCOM,
    27)), bs = "ts") + s(pmin(I(DistToFront063/1000), 75), bs = "ts") +
    s(pmax(0.5, pmin(MnkEpi, 26)), bs = "ts")

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -20.3615     0.1355  -150.3   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:
                                       edf Ref.df      F  p-value
s(Fetch_50km)                        2.782      9  1.959 0.000144 ***
s(log10(pmax(3, pmin(Depth, 4000)))) 5.510      9 18.474  < 2e-16 ***
s(pmax(3, pmin(SST_CMC, 30)))        6.586      9 24.661  < 2e-16 ***
s(pmax(3, pmin(BotT_HYCOM, 27)))     2.688      9 14.488  < 2e-16 ***
s(pmin(I(DistToFront063/1000), 75))  1.661      9  6.285  < 2e-16 ***
s(pmax(0.5, pmin(MnkEpi, 26)))       5.609      9  6.790  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) =  0.0107   Deviance explained = 13.3%
-REML = 9718.3  Scale est. = 6.9674    n = 162953

Method: REML   Optimizer: outer newton
full convergence after 11 iterations.
Gradient range [-0.002087588,0.0005670777]
(score 9718.283 & scale 6.967387).
Hessian positive definite, eigenvalue range [0.1694261,11901.02].
Model rank =  55 / 55

Basis dimension (k) checking results. Low p-value (k-index<1) may
indicate that k is too low, especially if edf is close to k'.

                                        k'  edf k-index p-value
s(Fetch_50km)                         9.00 2.78    0.88   0.705
s(log10(pmax(3, pmin(Depth, 4000)))) 9.00 5.51    0.84   0.015 *
s(pmax(3, pmin(SST_CMC, 30)))        9.00 6.59    0.84   0.030 *
s(pmax(3, pmin(BotT_HYCOM, 27)))     9.00 2.69    0.84   0.005 **
s(pmin(I(DistToFront063/1000), 75))  9.00 1.66    0.89   0.900
s(pmax(0.5, pmin(MnkEpi, 26)))       9.00 5.61    0.85   0.055 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
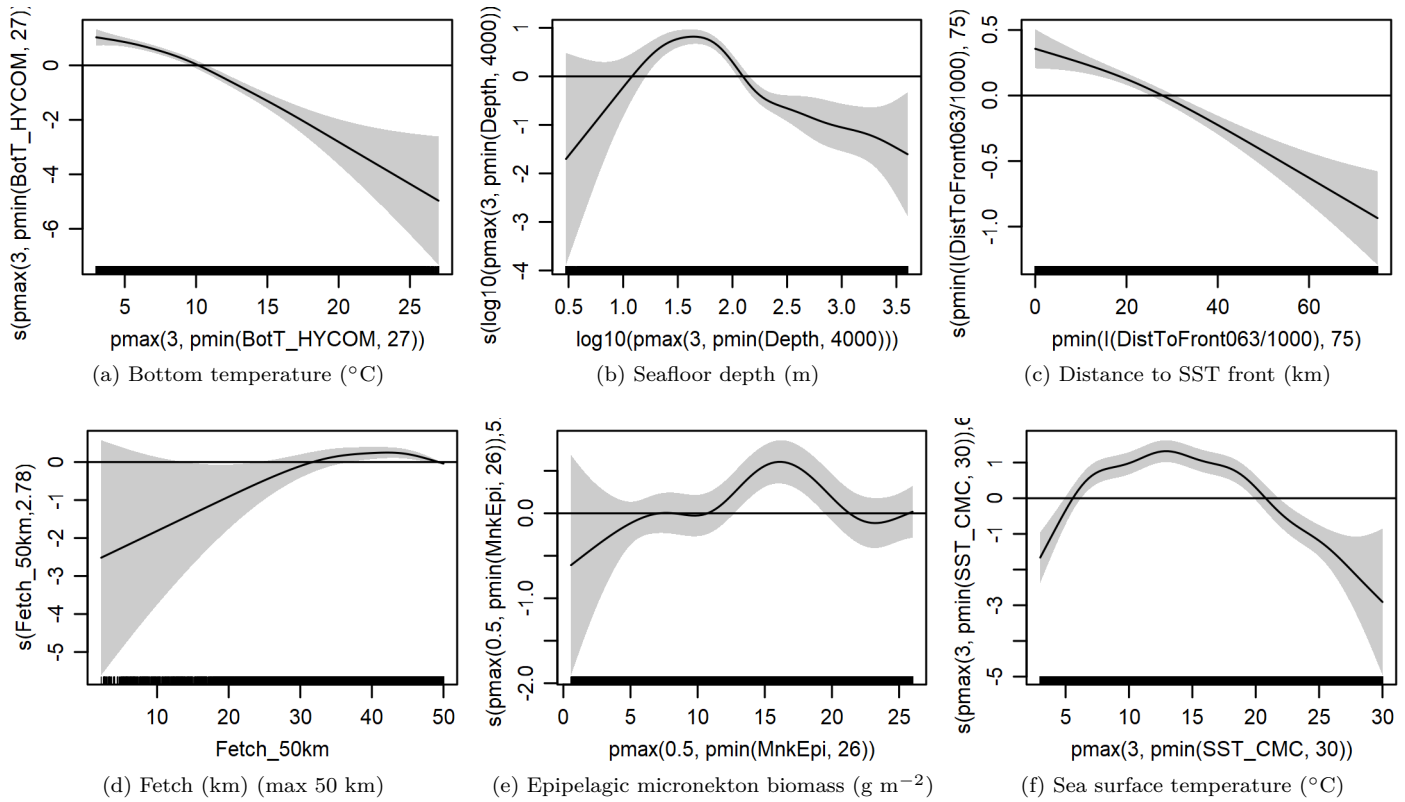
Figure 43: Functional plots for the final model for the region Surveyed Area (Excluding SE Shelf) for Summer. Transforms and other treatments are indicated in axis labels. *log10* indicates the covariate was $\log_{10}$ transformed. *pmax* and *pmin* indicate the covariate's minimum and maximum values, respectively, were Winsorized to the values shown. Winsorization was used to prevent runaway extrapolations during prediction when covariates exceeded sampled ranges, or for ecological reasons, depending on the covariate. */1000* indicates meters were transformed to kilometers for interpretation convenience.

Table 8: Covariates used in the final model for the region Surveyed Area (Excluding SE Shelf) for Summer.

| Covariate	Description
BotT_HYCOM	Monthly mean bottom temperature (°C) from the HYCOM GOFS 3.1 1/12° ocean model (Chassignet et al. (2009))
Depth	Depth (m) of the seafloor, from SRTM30_PLUS (Becker et al. (2009))
DistToFront063	Monthly mean distance (km) to the closest sea surface temperature front detected in daily GHRSST Level 4 CMC0.2deg and CMC0.1deg images (Brasnett (2008); Canada Meteorological Center (2012); Meissner et al. (2016); Canada Meteorological Center (2016)) with MGET's implementation of the Canny edge detector (Roberts et al. (2010); Canny (1986))
Fetch_50km	Fetch (km): mean distance to shore averaged over 16 radial directions, limited to a maximum of 50 km
MnkEpi	Monthly mean micronekton biomass available in the epipelagic zone, expressed as wet weight (g m^{-2}), from SEAPODYM (Lehodey et al. (2008); Lehodey et al. (2015)), provided by E.U. Copernicus Marine Service. doi: 10.48670/moi-00020. Computed as the sum of the SEAPODYM mnkc_epi, mnkc_mumeso, and mnkc_hmlmeso variables.
SST_CMC	Monthly mean sea surface temperature (°C) from GHRSST Level 4 CMC0.2deg and CMC0.1deg (Brasnett (2008); Canada Meteorological Center (2012); Meissner et al. (2016); Canada Meteorological Center (2016))

4.2.2 Diagnostic Plots
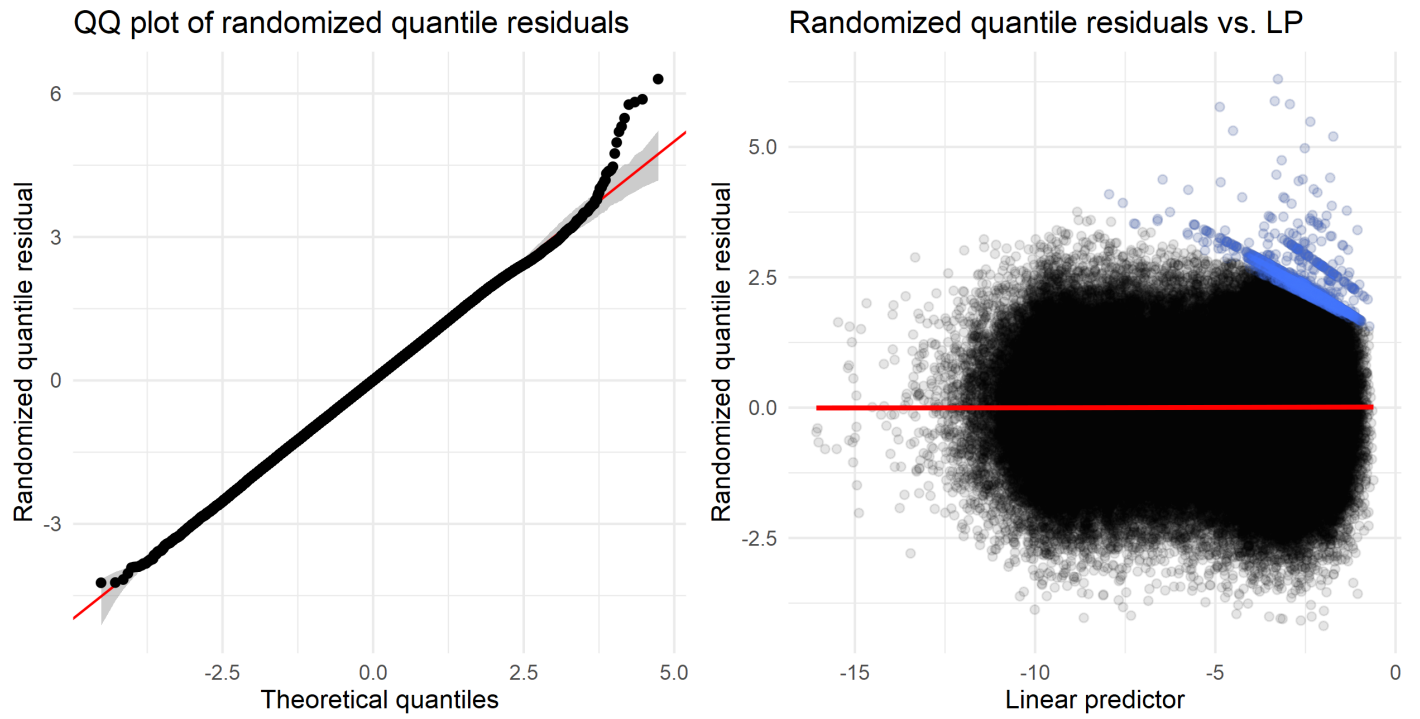
Figure 44: Residual plots for the final model for the region Surveyed Area (Excluding SE Shelf) for Summer.
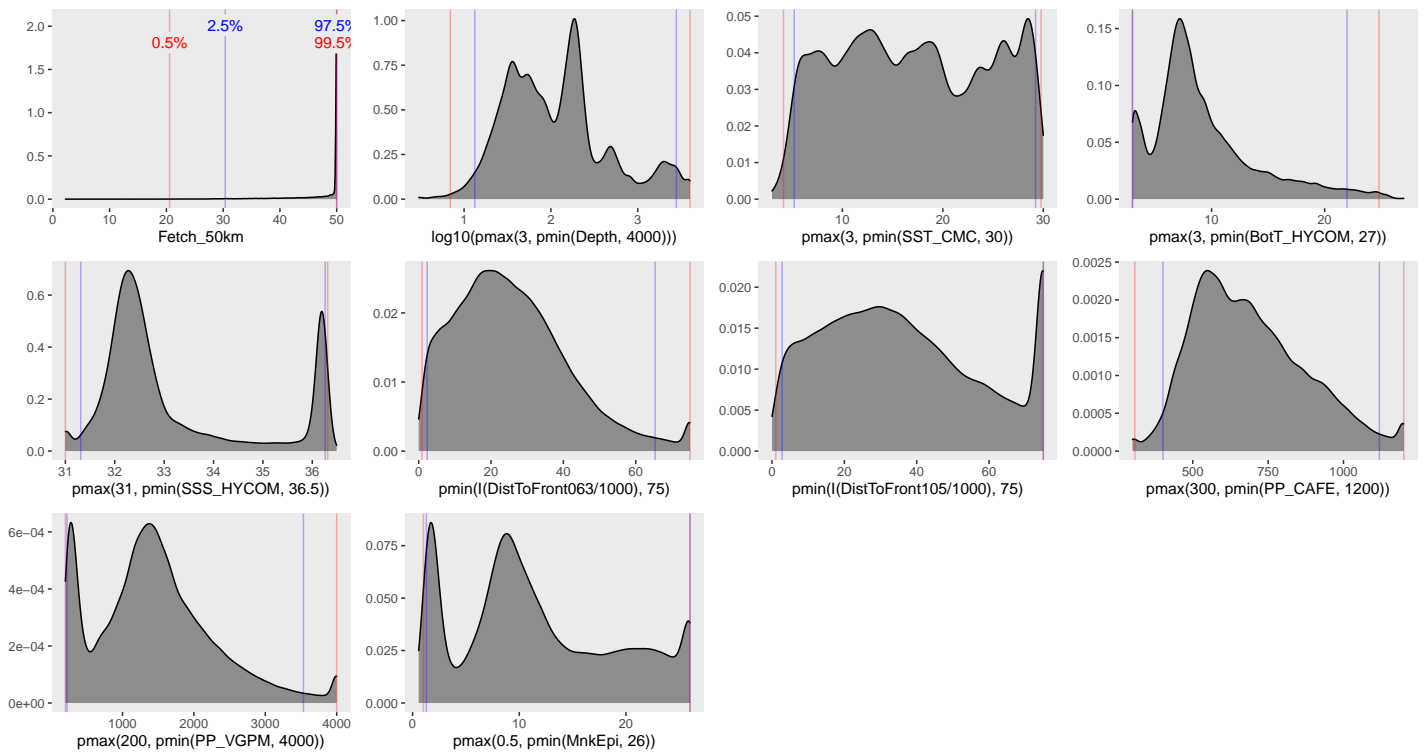
Figure 45: Density histograms showing the distributions of the covariates considered during the final model selection step. The final model may have included only a subset of the covariates shown here (see Figure 43), and additional covariates may have been considered in preceding selection steps. Red and blue lines enclose 99% and 95% of the distributions, respectively. Transforms and other treatments are indicated in axis labels. *log10* indicates the covariate was $\log_{10}$ transformed. *pmax* and *pmin* indicate the covariate's minimum and maximum values, respectively, were Winsorized to the values shown. Winsorization was used to prevent runaway extrapolations during prediction when covariates exceeded sampled ranges, or for ecological reasons, depending on the covariate. */1000* indicates meters were transformed to kilometers for interpretation convenience.
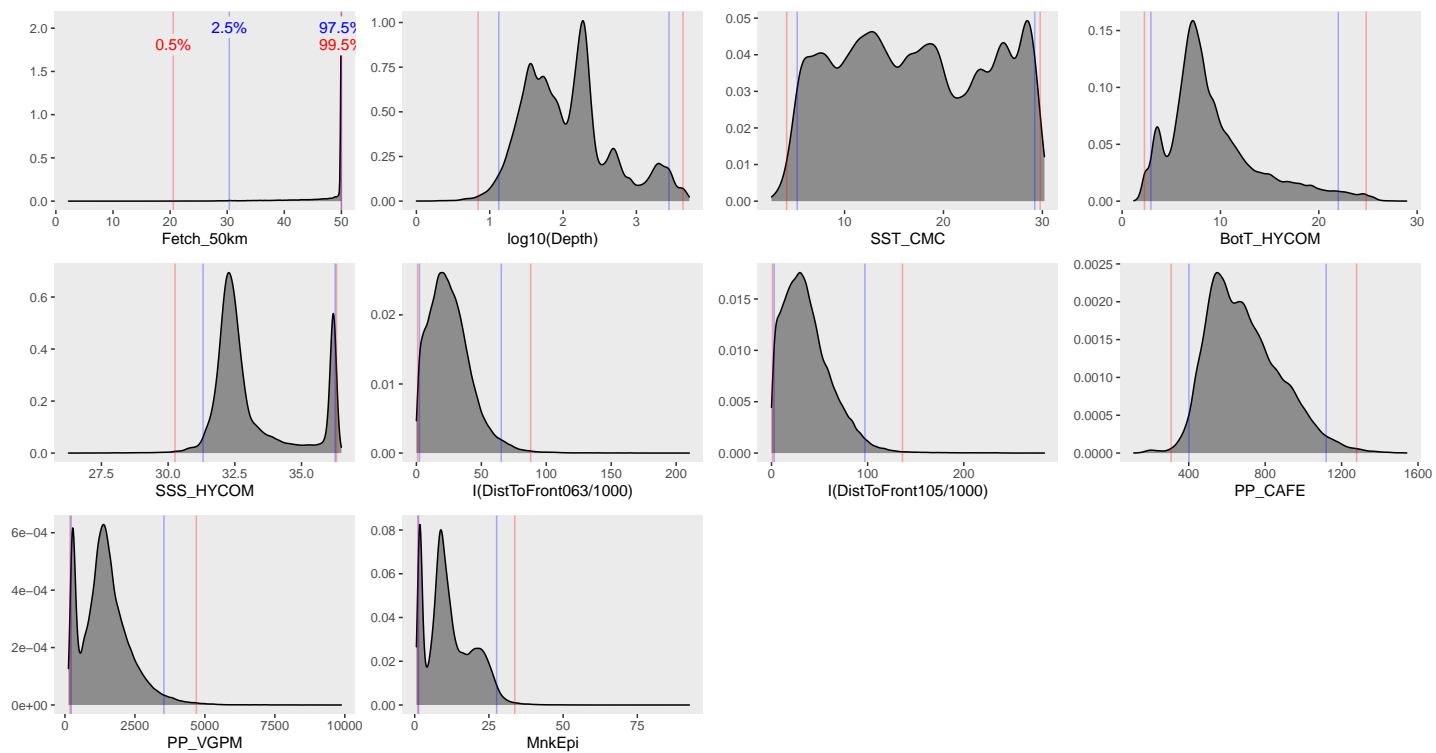
Figure 46: Density histograms shown in Figure 45 replotted without Winsorization, to show the full range of sampling represented by survey segments.
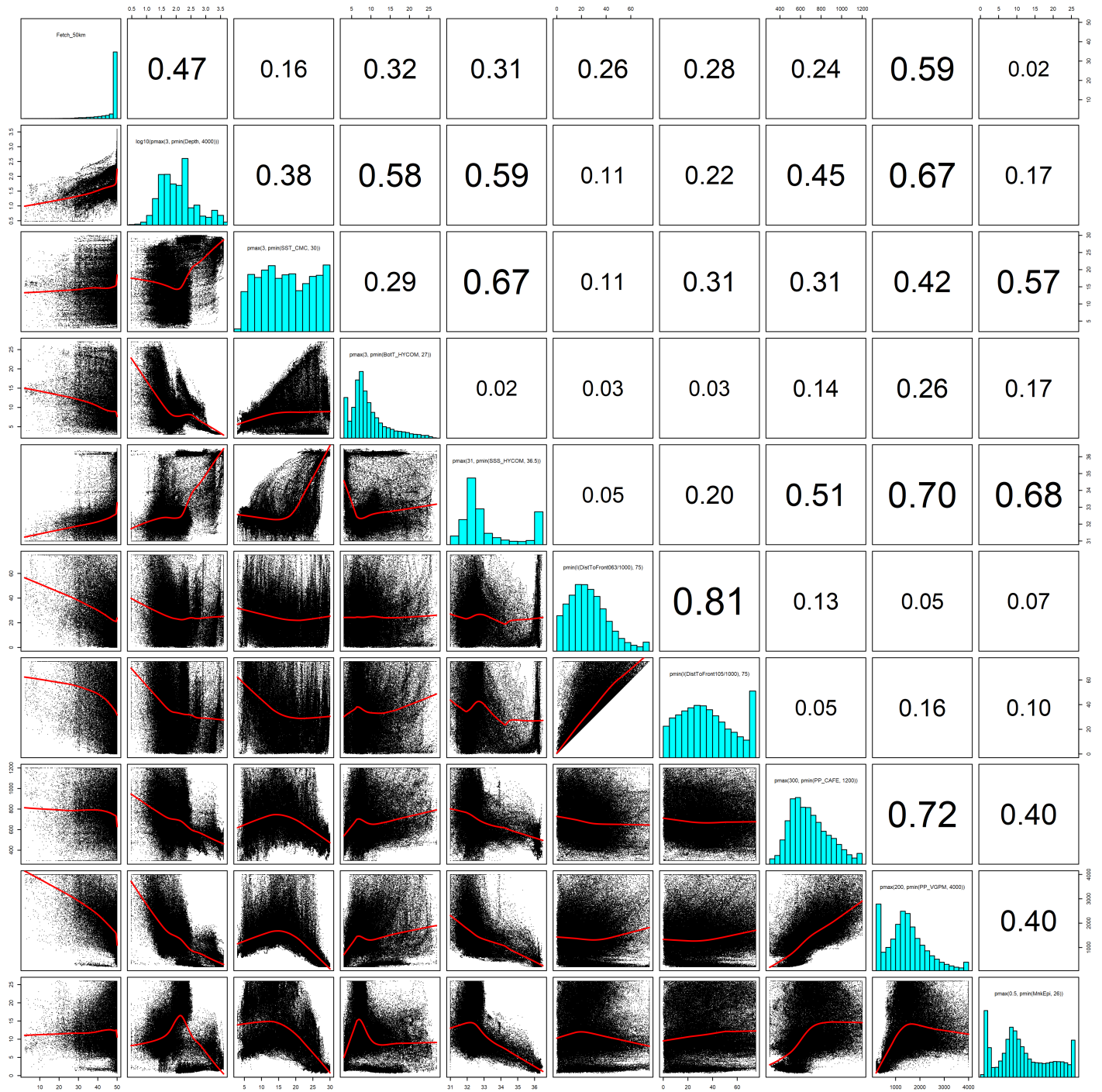
Figure 47: Scatterplot matrix of the covariates considered during the final model selection step. The final model may have included only a subset of the covariates shown here (see Figure 43), and additional covariates may have been considered in preceding selection steps. Covariates are transformed and Winsorized as shown in Figure 45. This plot is used to check simple correlations between covariates (via pairwise Pearson coefficients above the diagonal) and visually inspect for concurvity (via scatterplots and red lowess curves below the diagonal).
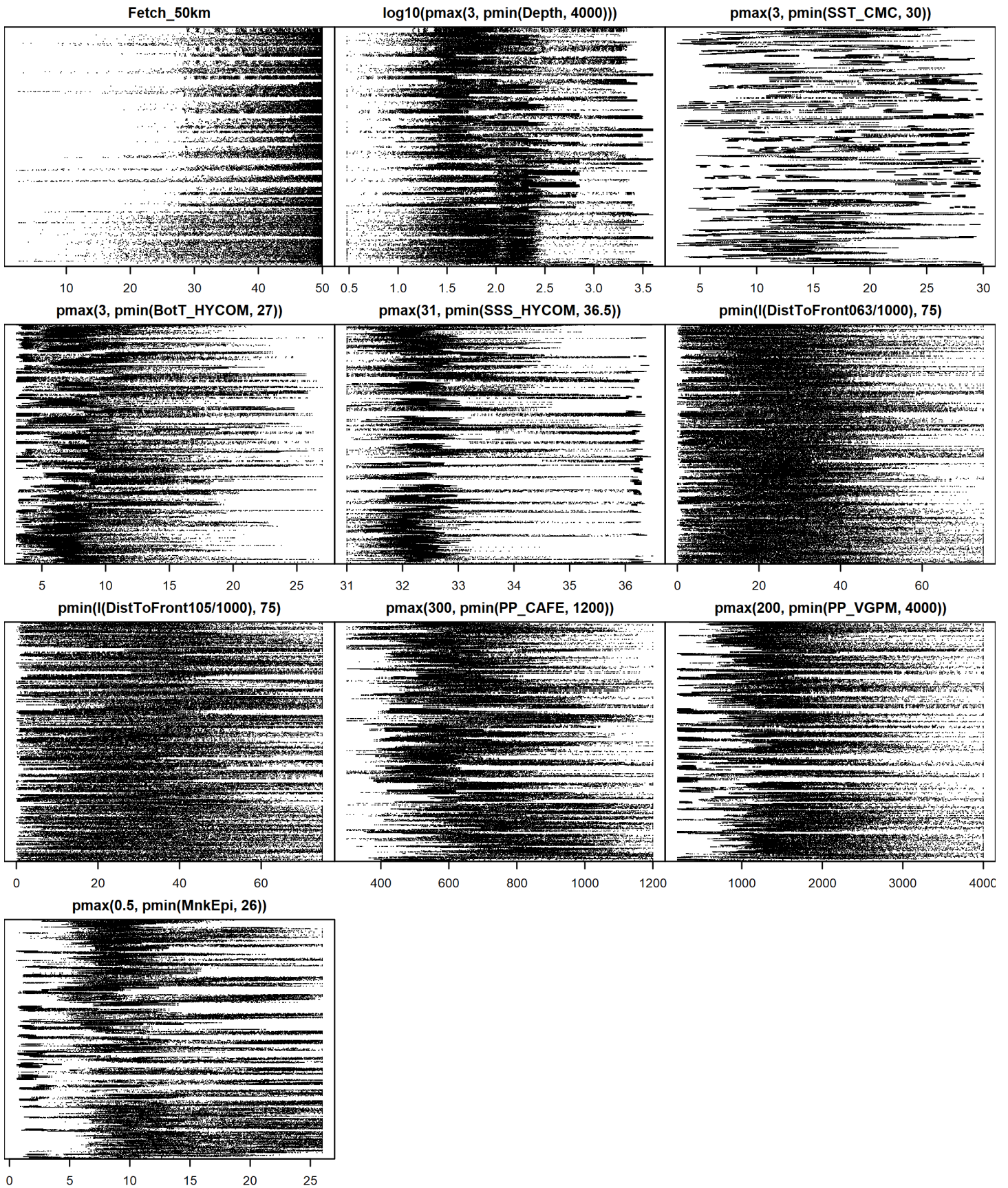
Figure 48: Dotplot of the covariates considered during the final model selection step. The final model may have included only a subset of the covariates shown here (see Figure 43), and additional covariates may have been considered in preceding selection steps. Covariates are transformed and Winsorized as shown in Figure 45. This plot is used to check for suspicious patterns and outliers in the data. Points are ordered vertically by segment ID, sequentially in time.

4.2.3 Extrapolation Diagnostics

4.2.3.1 Univariate Extrapolation

(a) Depth covariate
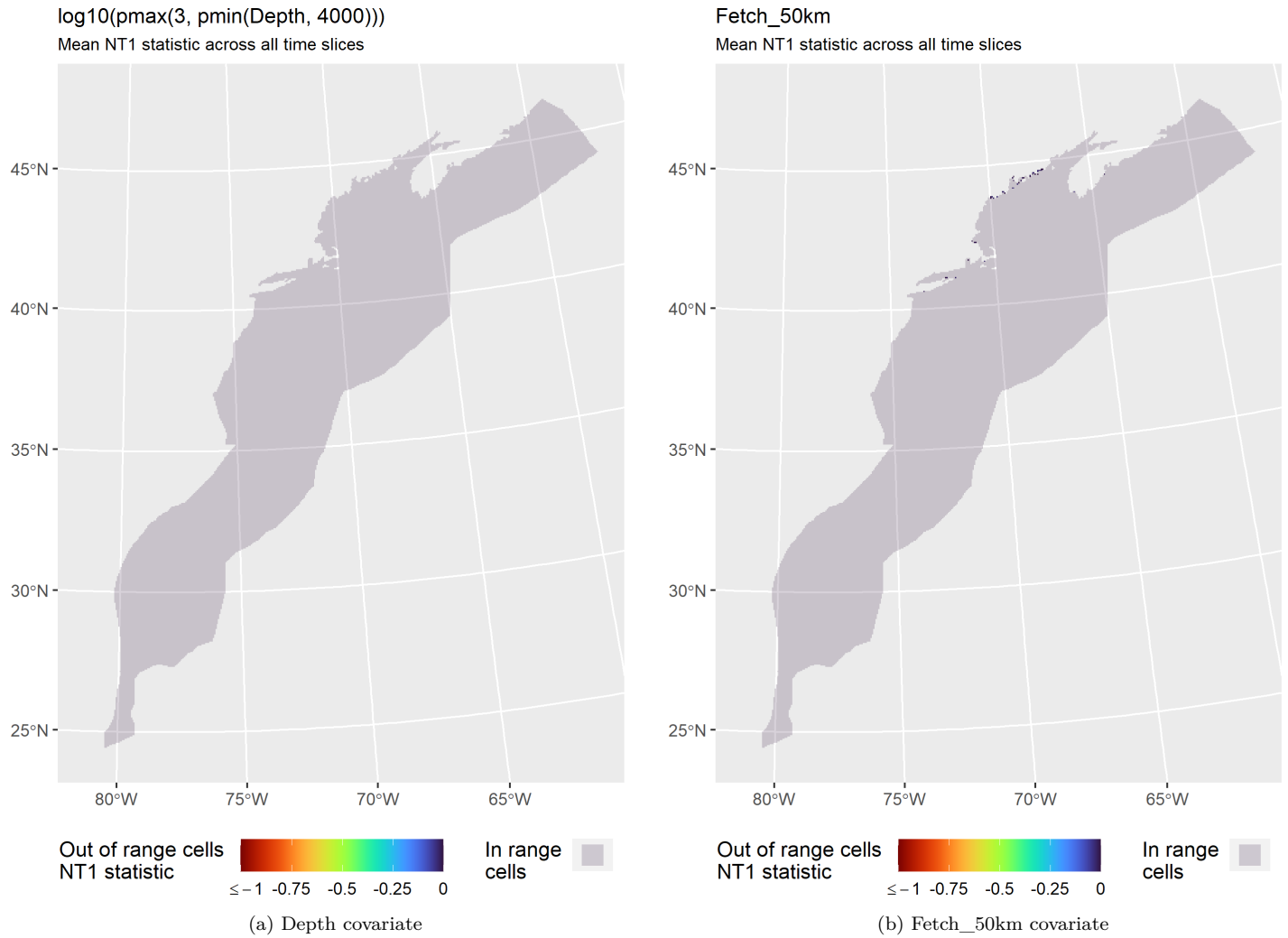
(b) Fetch_50km covariate

Figure 49: NT1 statistic (Mesgaran et al. (2014)) for static covariates used in the model for the region Surveyed Area (Excluding SE Shelf) for Summer. Areas outside the sampled range of a covariate appear in color, indicating univariate extrapolation of that covariate occurred there. Areas within the sampled range appear in gray, indicating it did not occur.
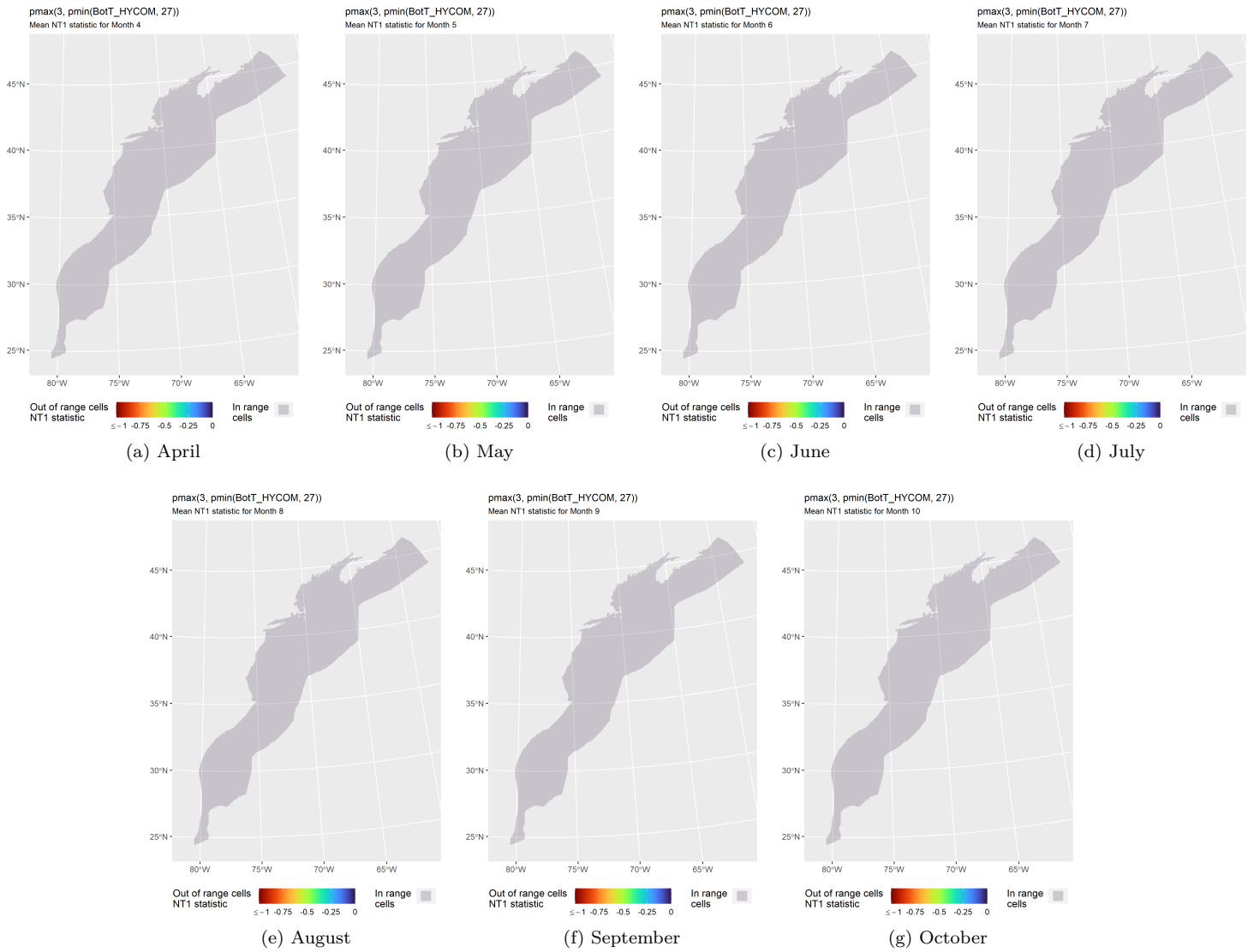
(a) April   (b) May   (c) June   (d) July

(e) August   (f) September   (g) October

Figure 50: NT1 statistic (Mesgaran et al. (2014)) for the BotT_HYCOM covariate in the model for the region Surveyed Area (Excluding SE Shelf) for Summer. Areas outside the sampled range of a covariate appear in color, indicating univariate extrapolation of that covariate occurred there during the month. Areas within the sampled range appear in gray, indicating it did not occur.

(a) April (b) May (c) June (d) July
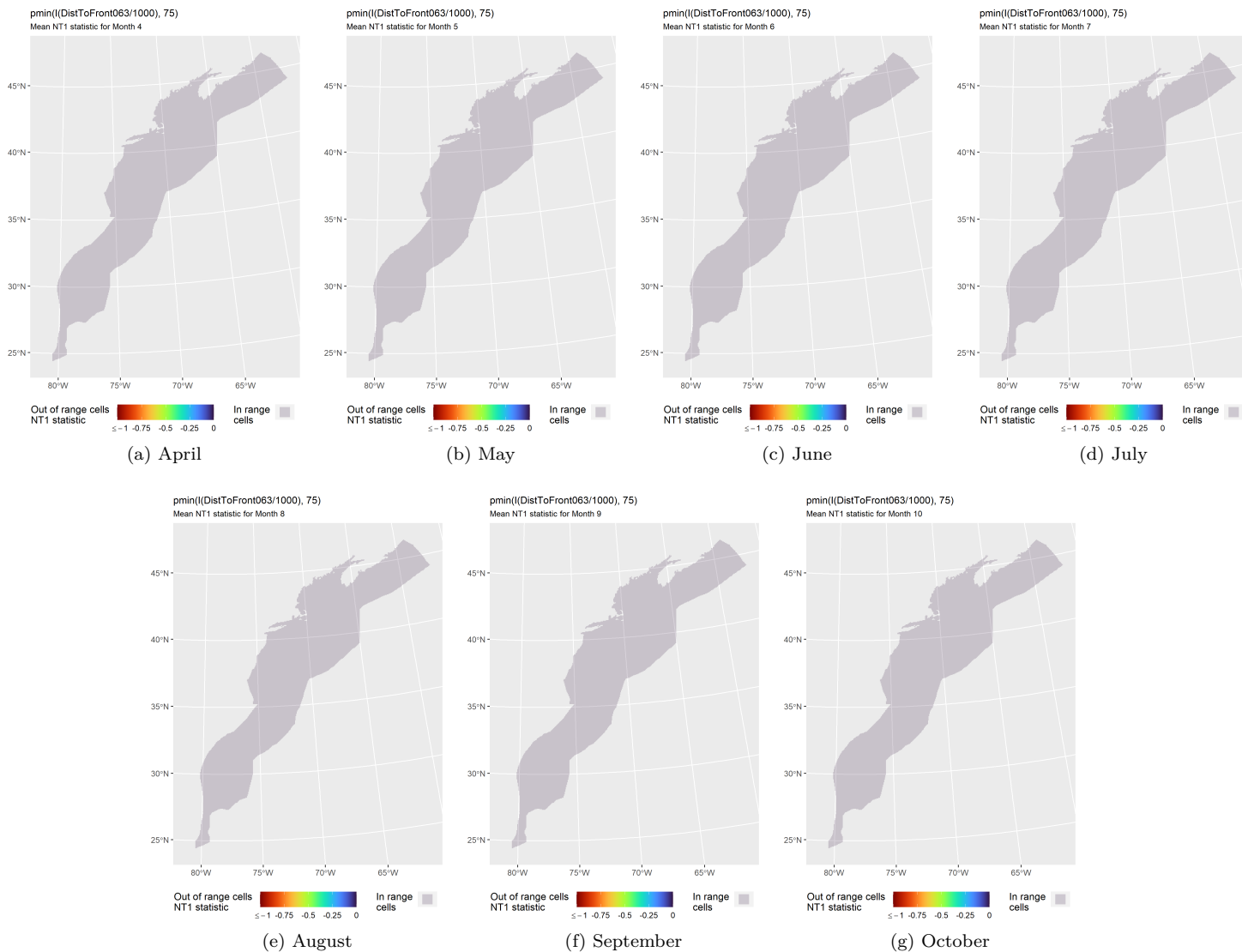
(e) August (f) September (g) October

Figure 51: NT1 statistic (Mesgaran et al. (2014)) for the DistToFront063 covariate in the model for the region Surveyed Area (Excluding SE Shelf) for Summer. Areas outside the sampled range of a covariate appear in color, indicating univariate extrapolation of that covariate occurred there during the month. Areas within the sampled range appear in gray, indicating it did not occur.
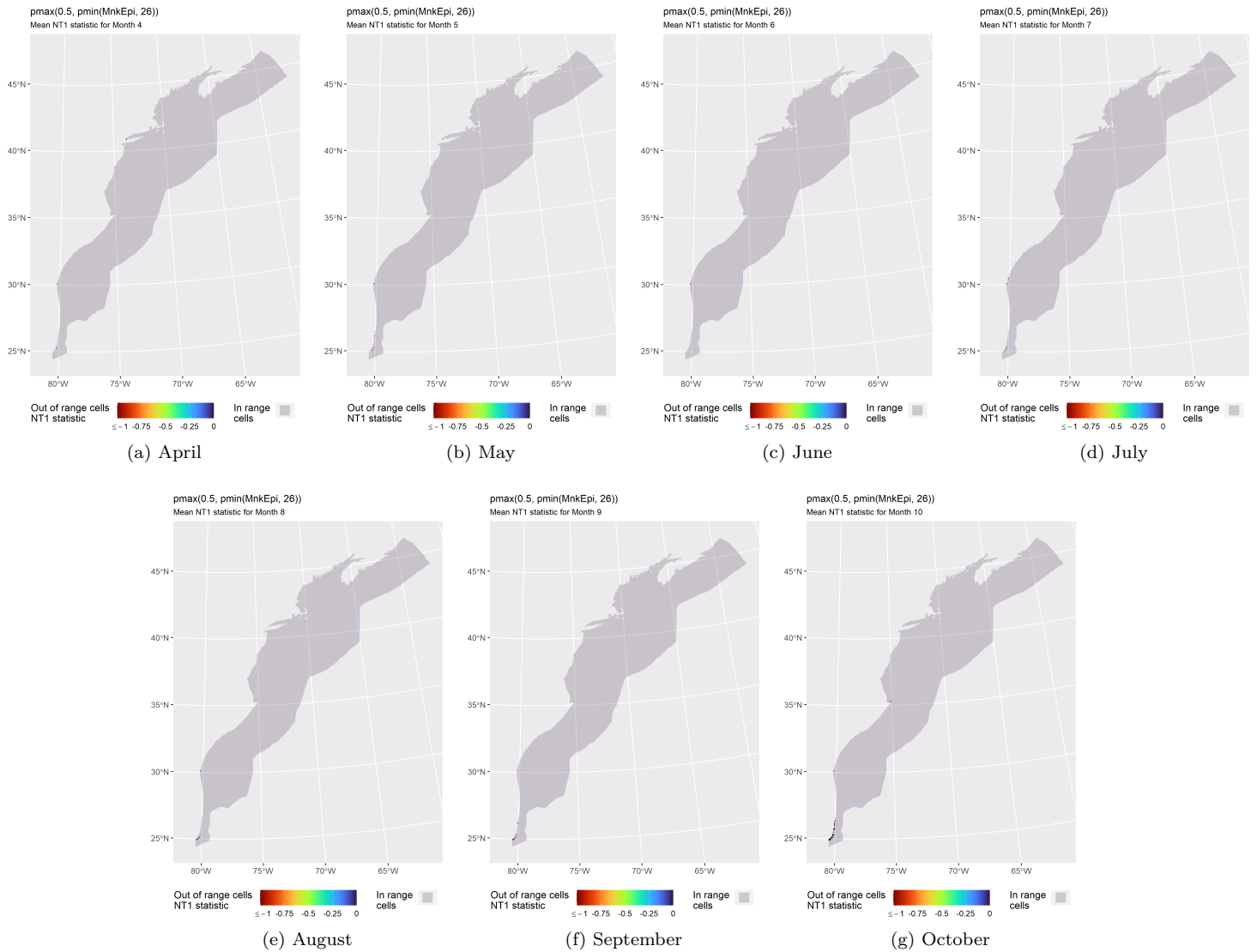
Figure 52: NT1 statistic (Mesgaran et al. (2014)) for the MnkEpi covariate in the model for the region Surveyed Area (Excluding SE Shelf) for Summer. Areas outside the sampled range of a covariate appear in color, indicating univariate extrapolation of that covariate occurred there during the month. Areas within the sampled range appear in gray, indicating it did not occur.
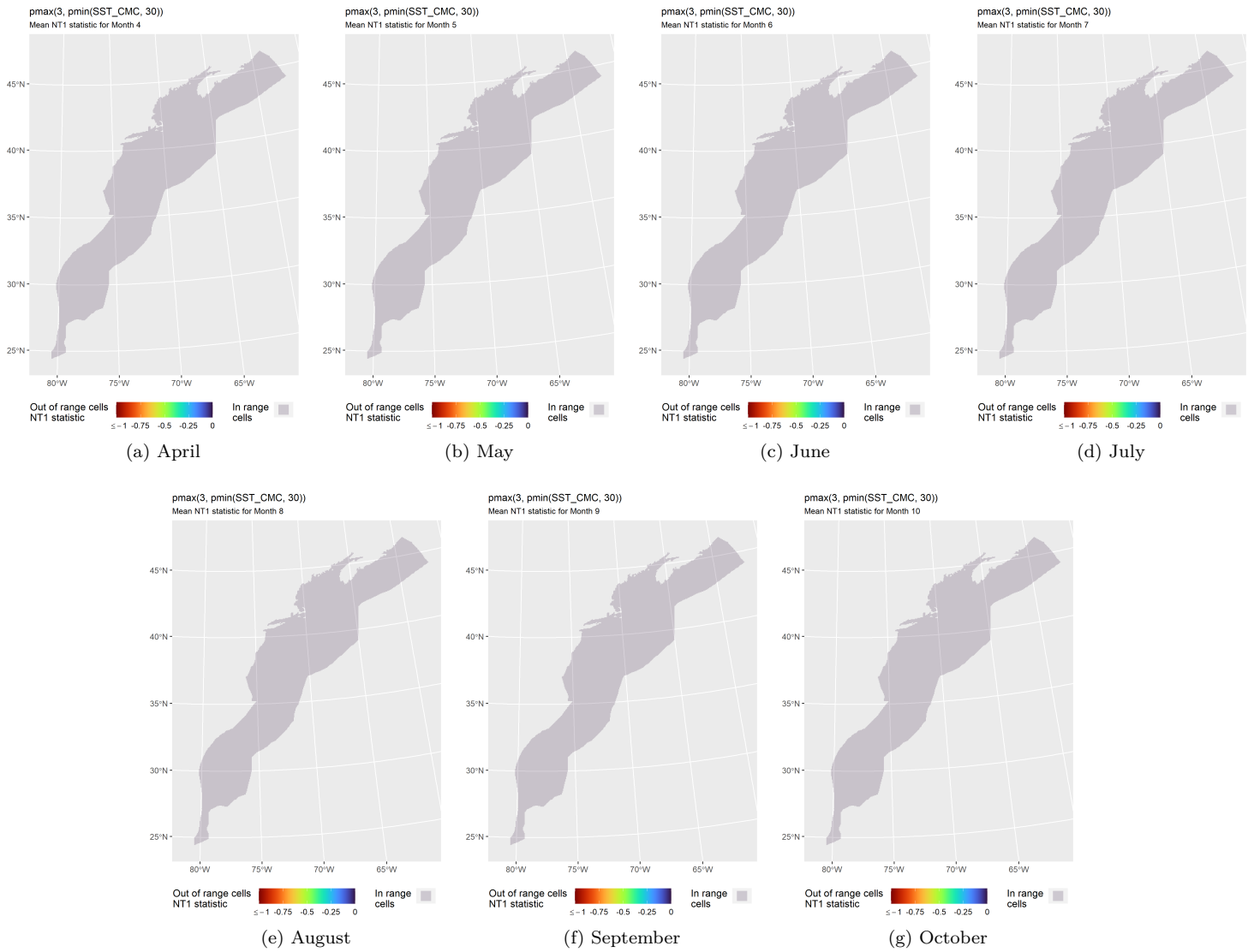
(a) April   (b) May   (c) June   (d) July

(e) August   (f) September   (g) October

Figure 53: NT1 statistic (Mesgaran et al. (2014)) for the SST_CMC covariate in the model for the region Surveyed Area (Excluding SE Shelf) for Summer. Areas outside the sampled range of a covariate appear in color, indicating univariate extrapolation of that covariate occurred there during the month. Areas within the sampled range appear in gray, indicating it did not occur.

4.2.3.2 Multivariate Extrapolation

(a) April   (b) May   (c) June   (d) July
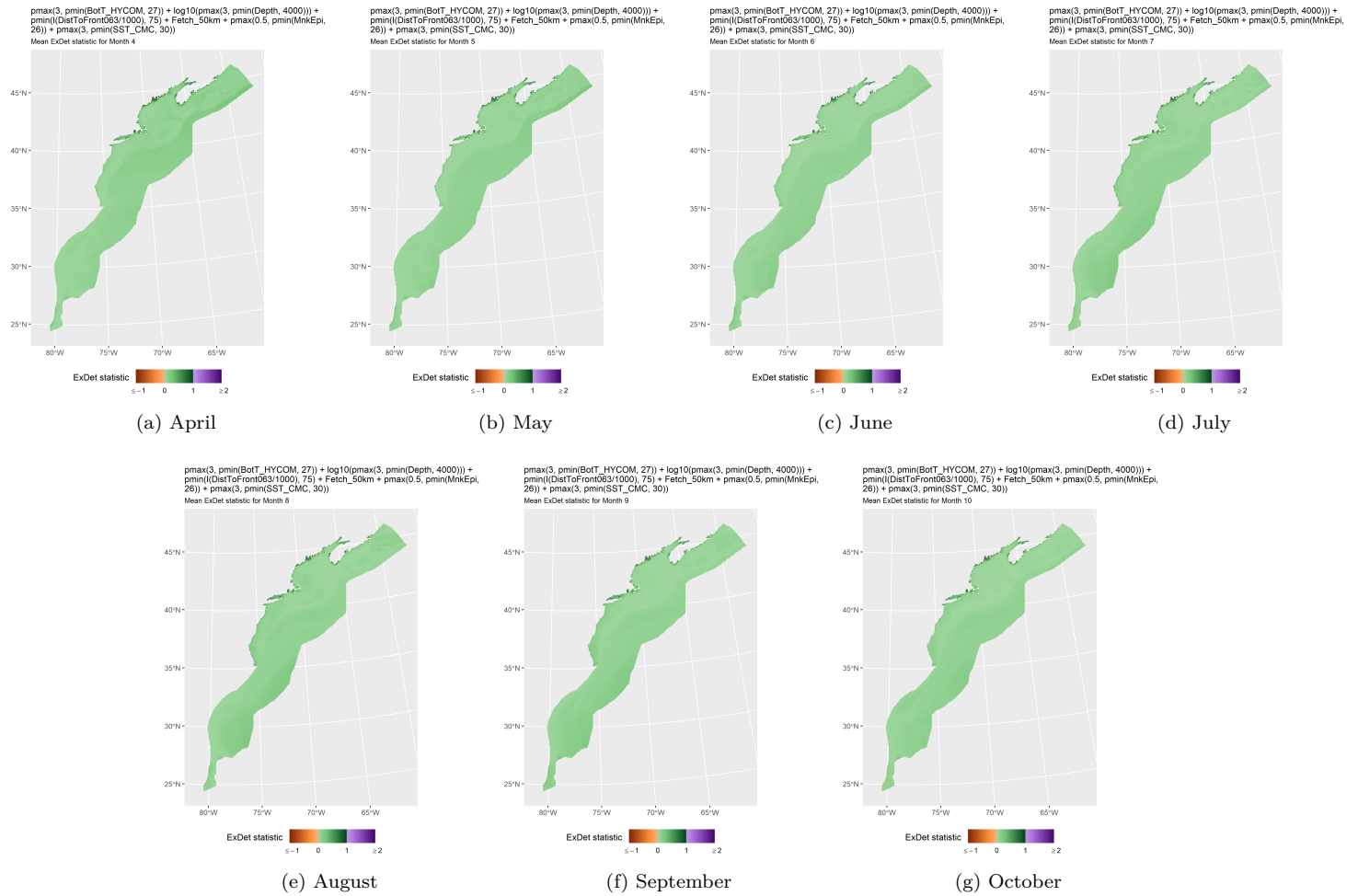
(e) August   (f) September   (g) October

Figure 54: ExDet statistic (Mesgaran et al. (2014)) for all of the covariates used in the model for the region Surveyed Area (Excluding SE Shelf) for Summer. Areas in orange (ExDet < 0) required univariate extrapolation of one or more covariates (see previous section). Areas in purple (ExDet > 1), did not require univariate extrapolation but did require multivariate extrapolation, by virtue of having novel combinations of covariates not represented in the survey data, according to the NT2 statistic (Mesgaran et al. (2014)). Areas in green ($0 \geq \text{ExDet} \leq 1$) did not require either type of extrapolation.

5 Predictions

Based on our evaluation of this model in the context of what is known of this species (see Section 6), we summarized its predictions into monthly climatological density and uncertainty surfaces, shown in the maps below.
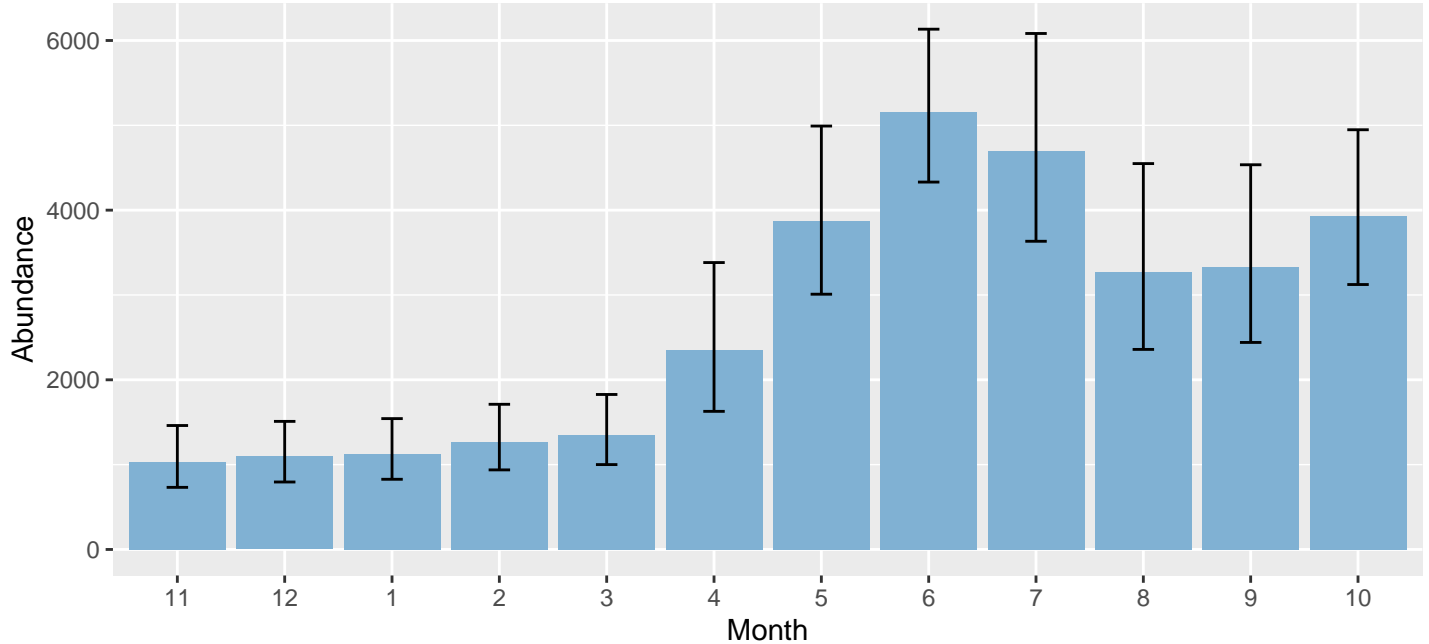
5.1 Summarized Predictions

Figure 55: Mean monthly abundance for the prediction area for November 1998 - October 2019. Note that the prediction area was not the same for all months (see Table 9 below and maps following). Error bars are a 95% interval, made with a log-normal approximation using the prediction's CV. The CV was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for temporal variability in dynamic covariates.

Table 9: Mean monthly abundance and density for the prediction area for November 1998 - October 2019. CV and intervals estimated as described for the previous figure.

| Month	Abundance	CV	95% Interval	Area (km²)	Density (individuals / 100 km²)
11	1,034	0.178	732 - 1,461	984,200	0.105
12	1,096	0.165	795 - 1,510	984,200	0.111
1	1,130	0.160	828 - 1,543	984,200	0.115
2	1,267	0.154	938 - 1,711	984,200	0.129
3	1,352	0.154	1,001 - 1,827	984,200	0.137
4	2,347	0.188	1,628 - 3,382	1,273,075	0.184
5	3,876	0.130	3,009 - 4,992	1,273,075	0.304
6	5,154	0.089	4,331 - 6,133	1,273,075	0.405
7	4,701	0.132	3,633 - 6,083	1,273,075	0.369
8	3,275	0.169	2,359 - 4,549	1,273,075	0.257
9	3,327	0.159	2,441 - 4,535	1,273,075	0.261
10	3,931	0.118	3,123 - 4,948	1,273,075	0.309
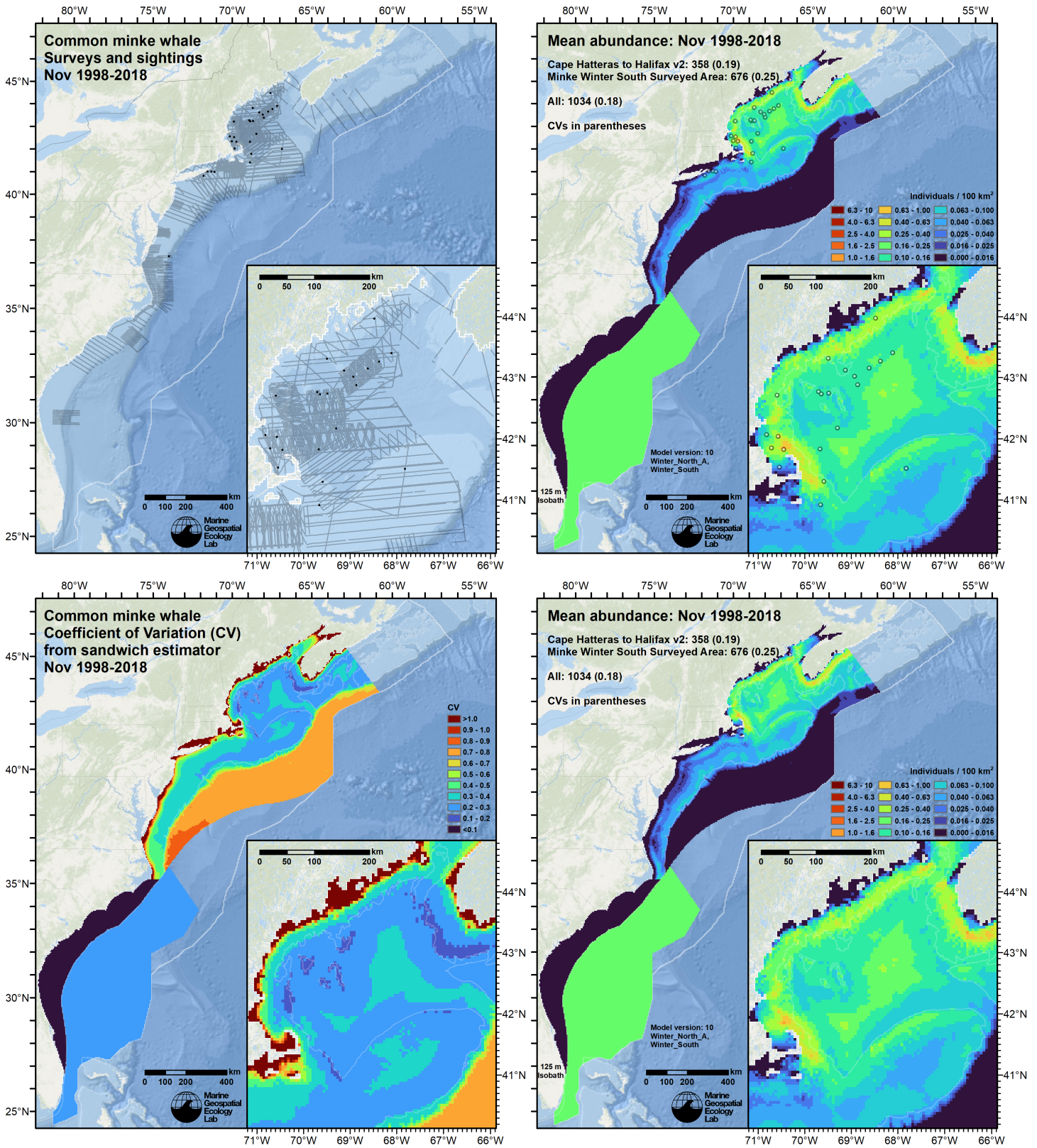
Figure 56: Survey effort and observations (top left), predicted density with observations (top right), predicted density without observations (bottom right), and coefficient of variation of predicted density (bottom left), for the month of November for the given era. Variance was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for temporal variability in dynamic covariates.
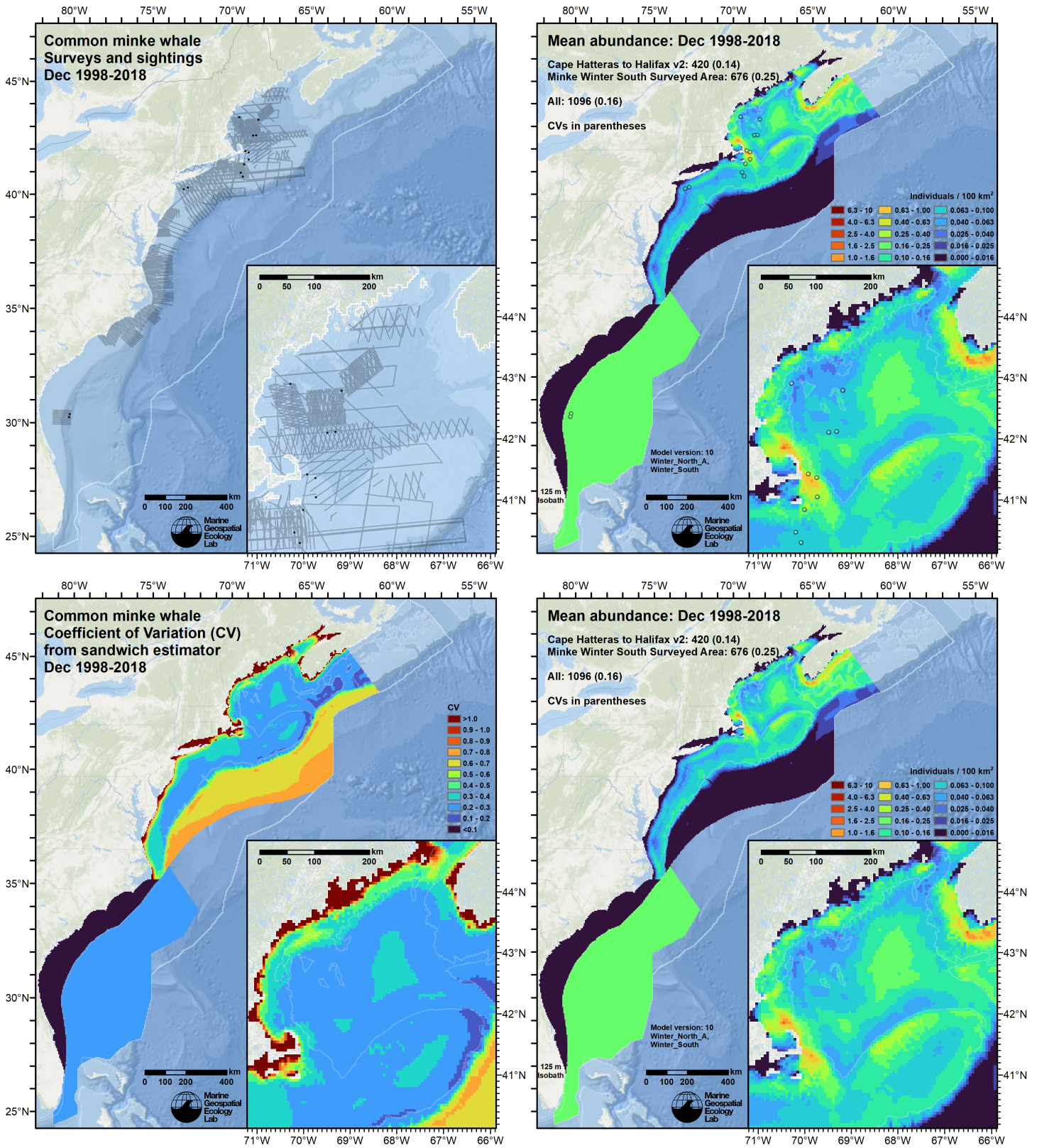
Figure 57: Survey effort and observations (top left), predicted density with observations (top right), predicted density without observations (bottom right), and coefficient of variation of predicted density (bottom left), for the month of December for the given era. Variance was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for temporal variability in dynamic covariates.
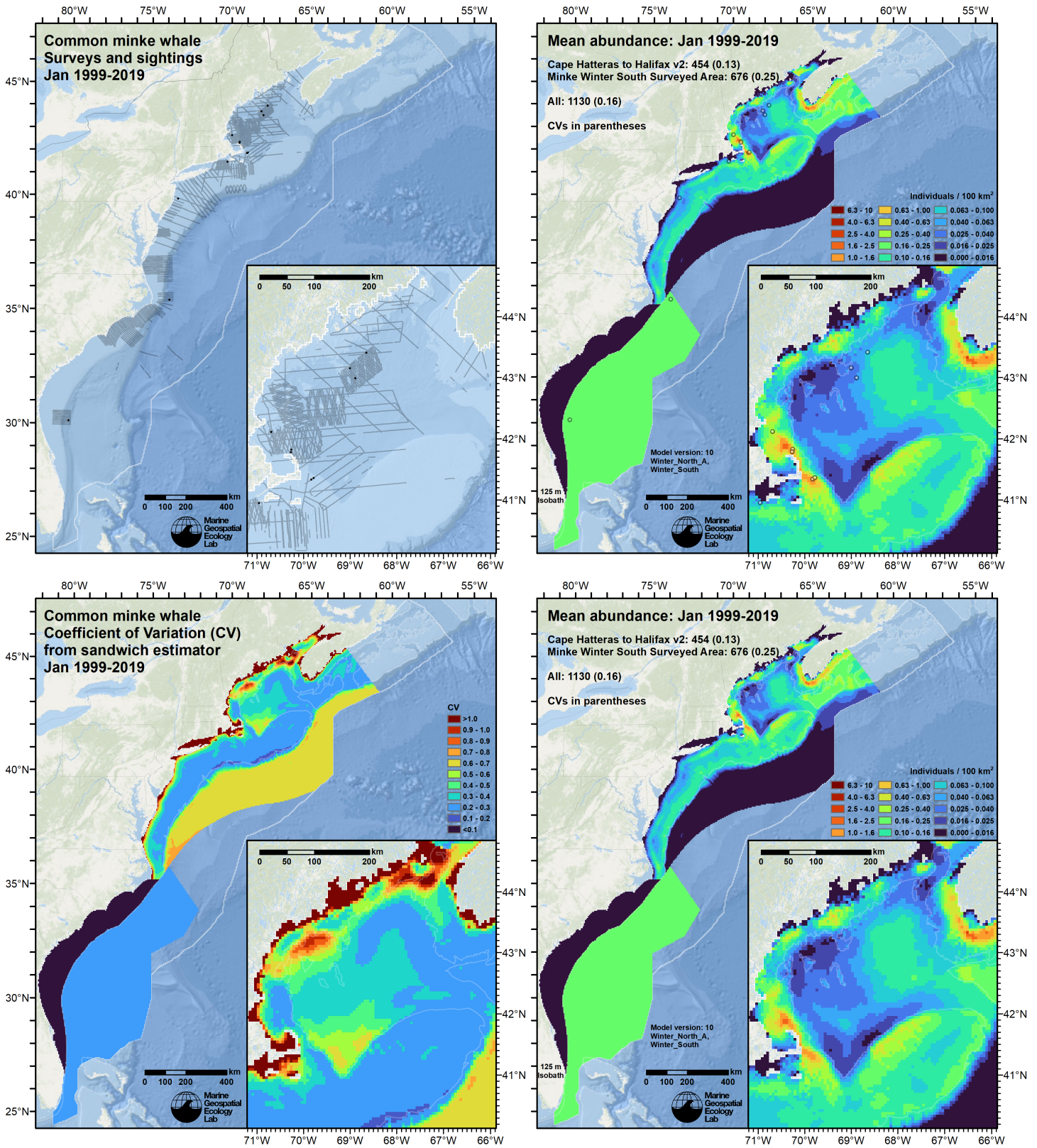
Figure 58: Survey effort and observations (top left), predicted density with observations (top right), predicted density without observations (bottom right), and coefficient of variation of predicted density (bottom left), for the month of January for the given era. Variance was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for temporal variability in dynamic covariates.
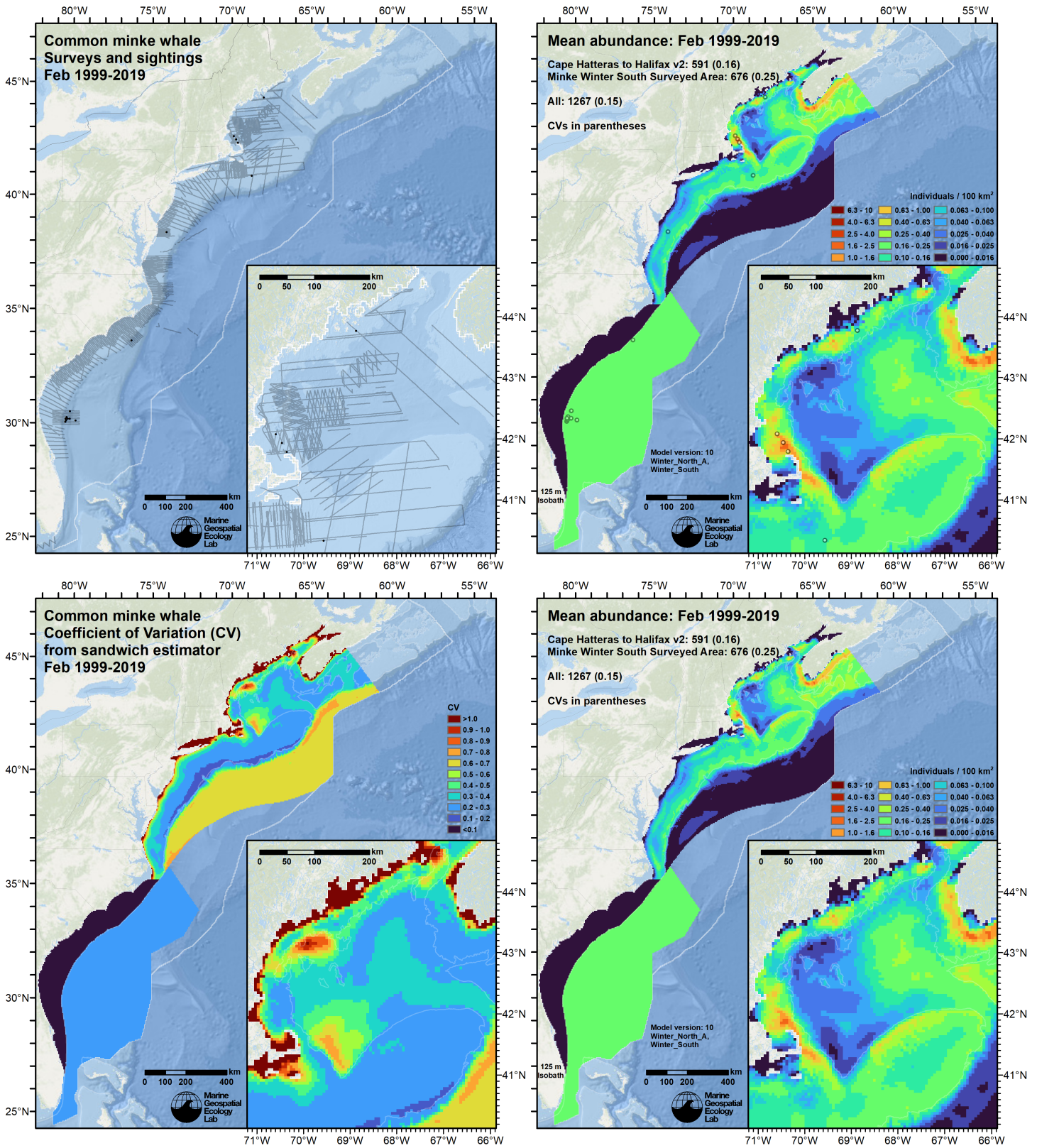
Figure 59: Survey effort and observations (top left), predicted density with observations (top right), predicted density without observations (bottom right), and coefficient of variation of predicted density (bottom left), for the month of February for the given era. Variance was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for temporal variability in dynamic covariates.
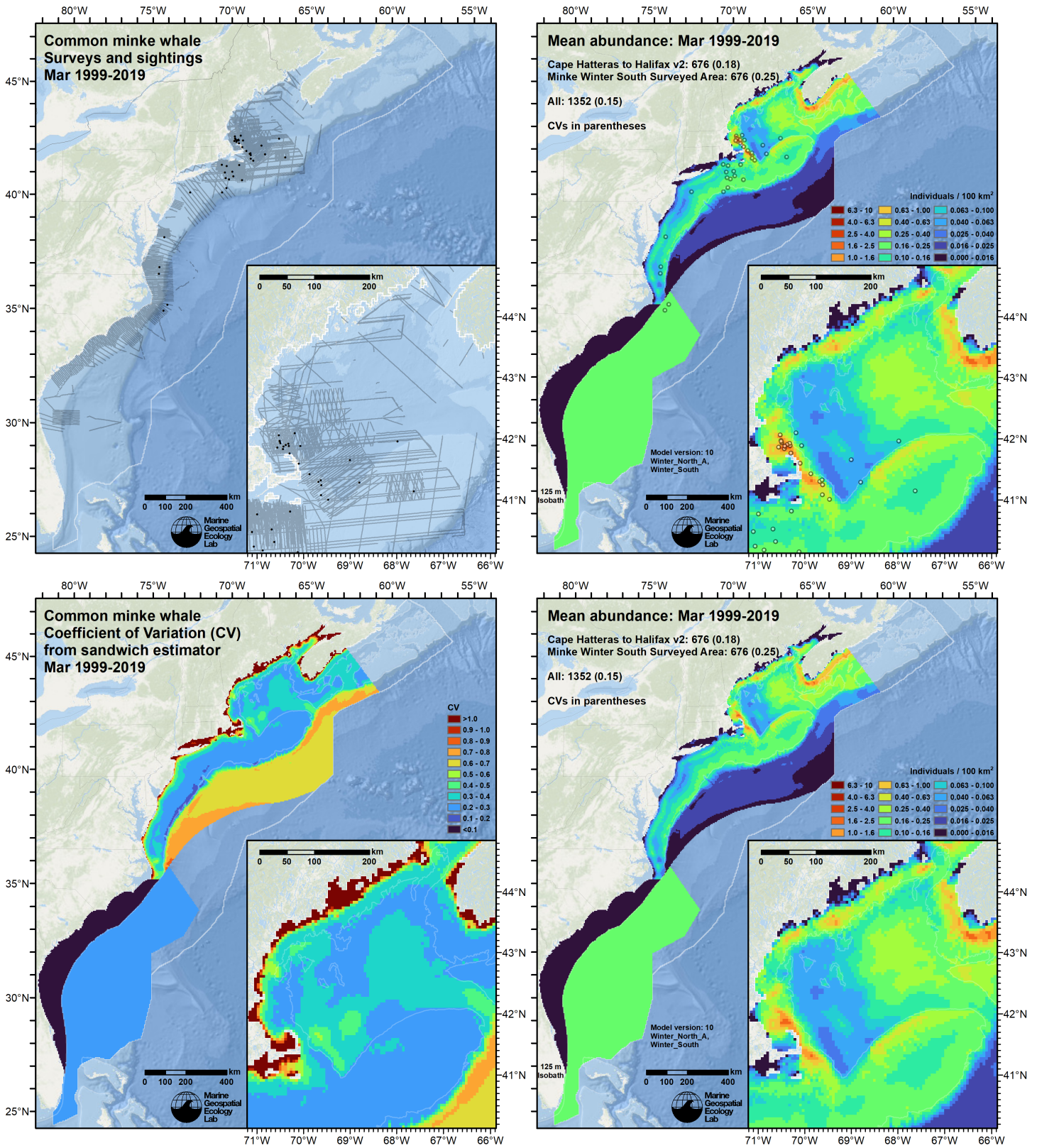
Figure 60: Survey effort and observations (top left), predicted density with observations (top right), predicted density without observations (bottom right), and coefficient of variation of predicted density (bottom left), for the month of March for the given era. Variance was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for temporal variability in dynamic covariates.

Figure 61: Survey effort and observations (top left), predicted density with observations (top right), predicted density without observations (bottom right), and coefficient of variation of predicted density (bottom left), for the month of April for the given era. Variance was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for temporal variability in dynamic covariates.
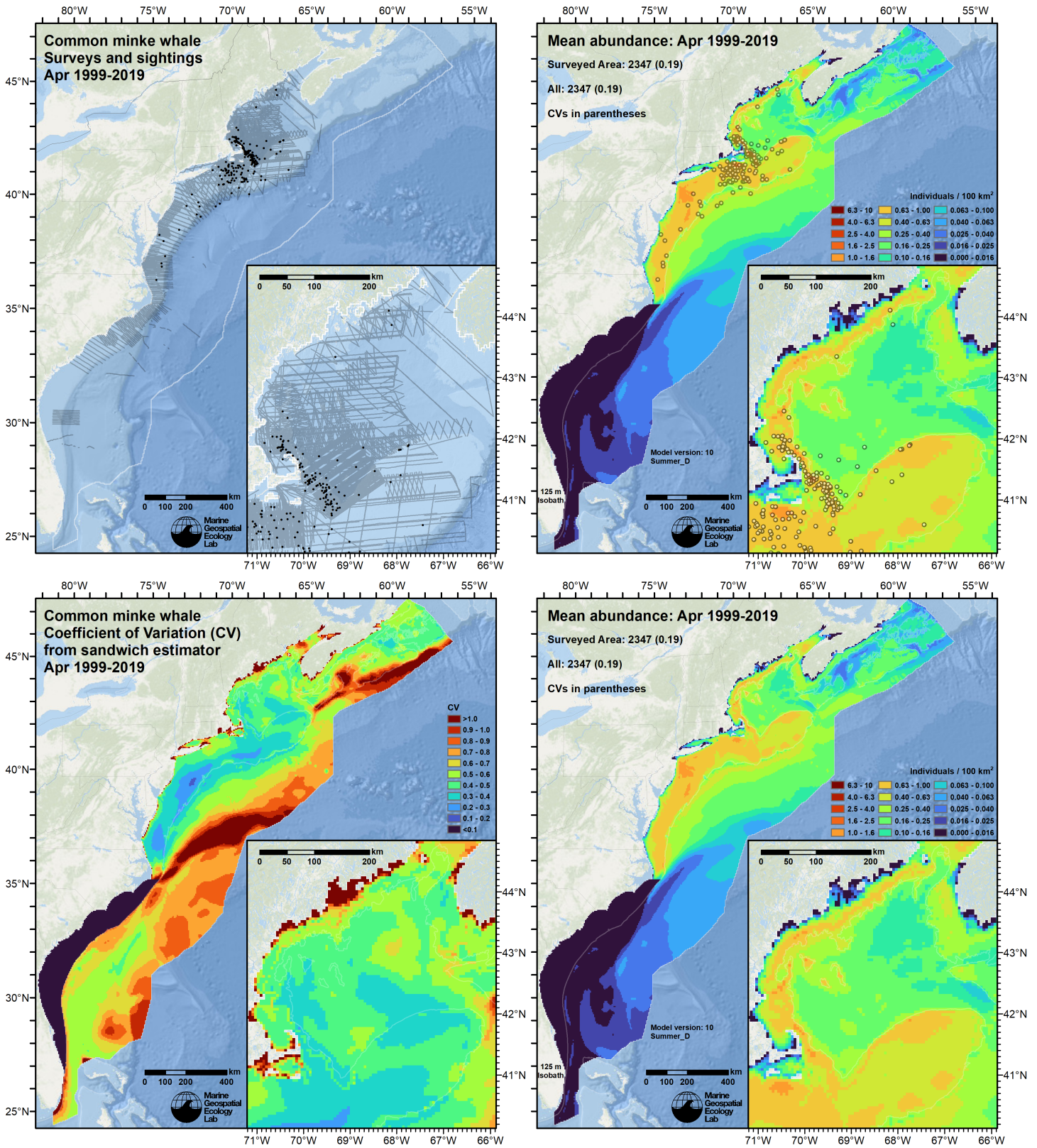
Figure 62: Survey effort and observations (top left), predicted density with observations (top right), predicted density without observations (bottom right), and coefficient of variation of predicted density (bottom left), for the month of May for the given era. Variance was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for temporal variability in dynamic covariates.
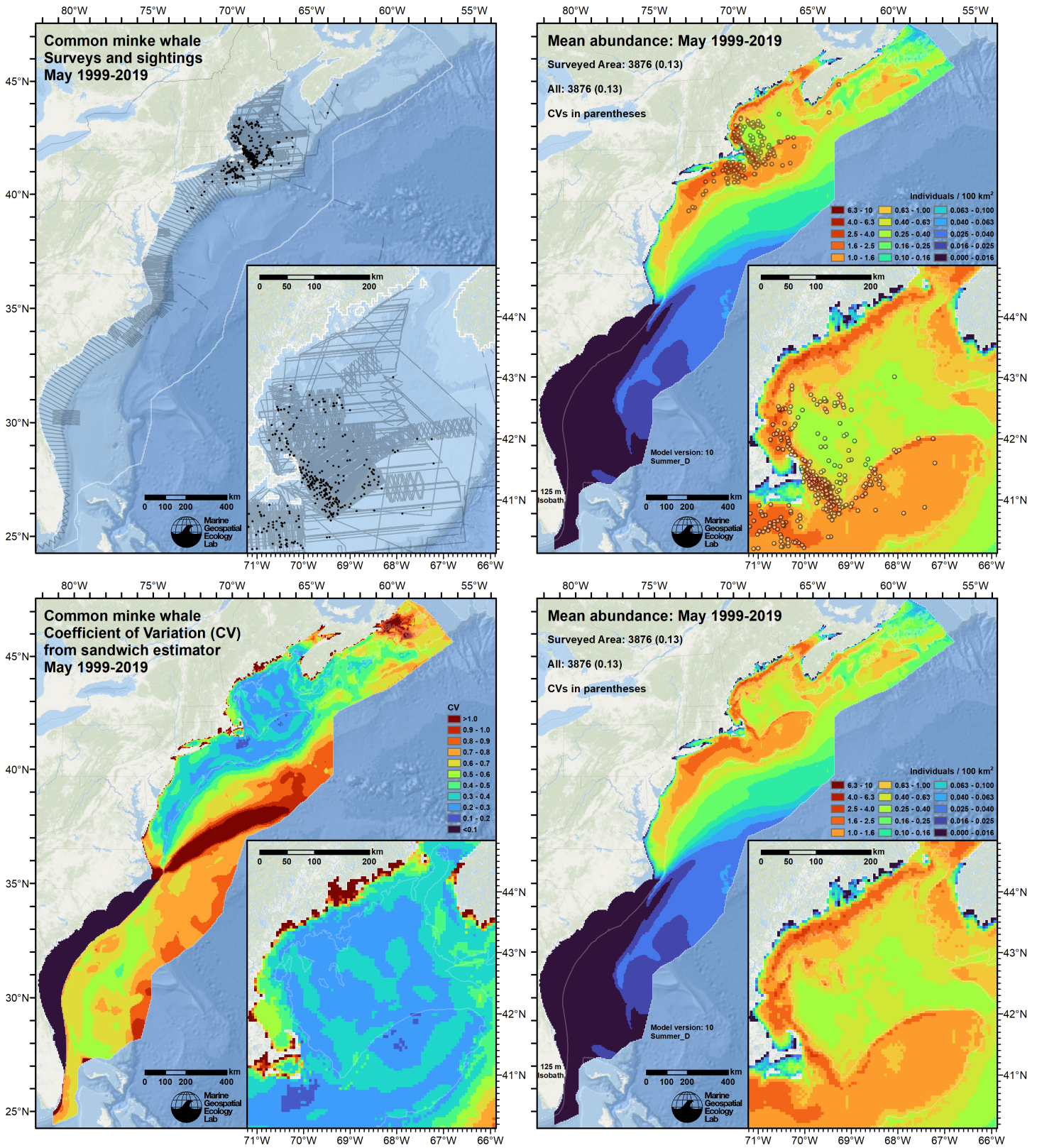
Figure 63: Survey effort and observations (top left), predicted density with observations (top right), predicted density without observations (bottom right), and coefficient of variation of predicted density (bottom left), for the month of June for the given era. Variance was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for temporal variability in dynamic covariates.

Figure 64: Survey effort and observations (top left), predicted density with observations (top right), predicted density without observations (bottom right), and coefficient of variation of predicted density (bottom left), for the month of July for the given era. Variance was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for temporal variability in dynamic covariates.
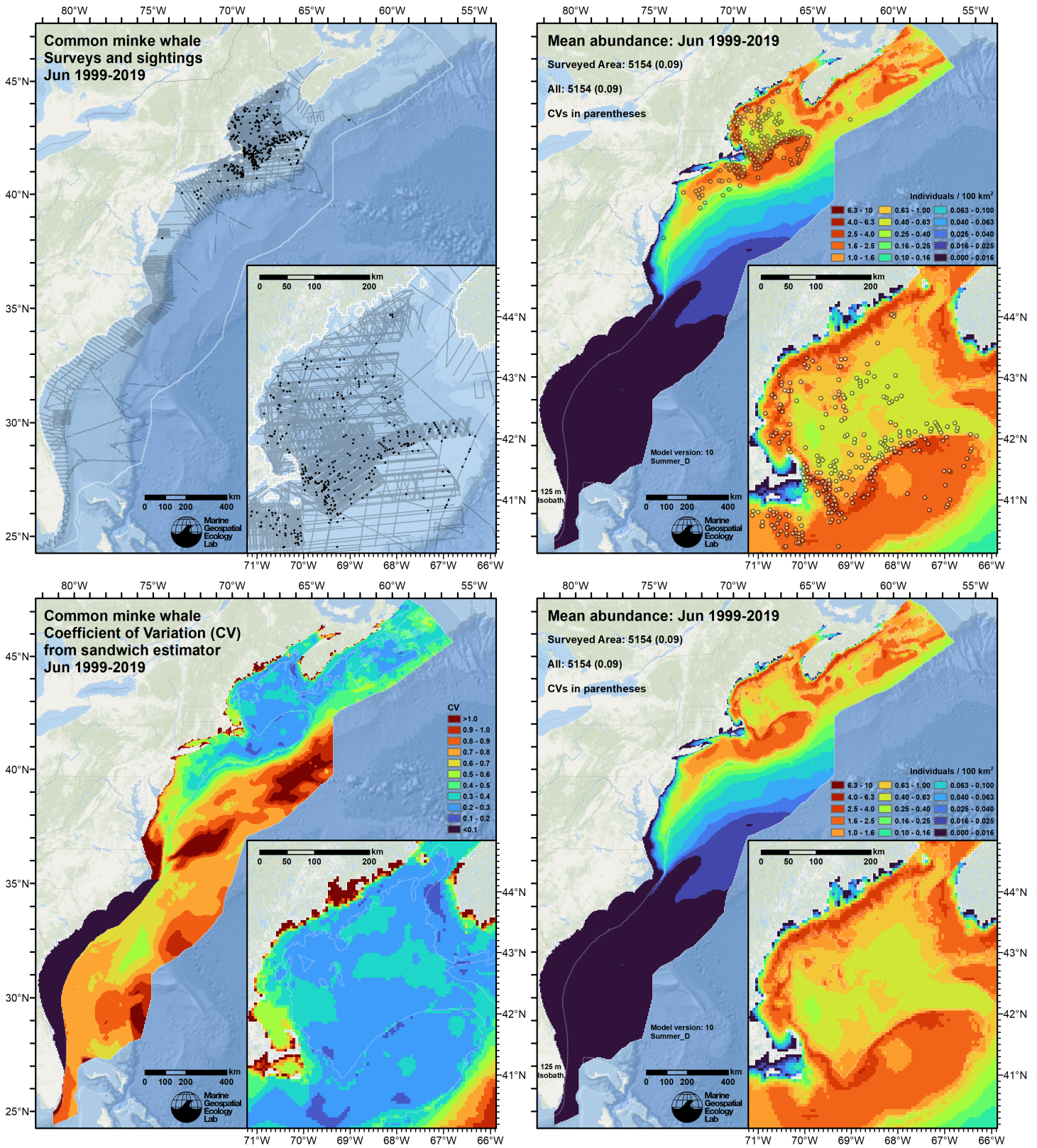
Figure 65: Survey effort and observations (top left), predicted density with observations (top right), predicted density without observations (bottom right), and coefficient of variation of predicted density (bottom left), for the month of August for the given era. Variance was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for temporal variability in dynamic covariates.
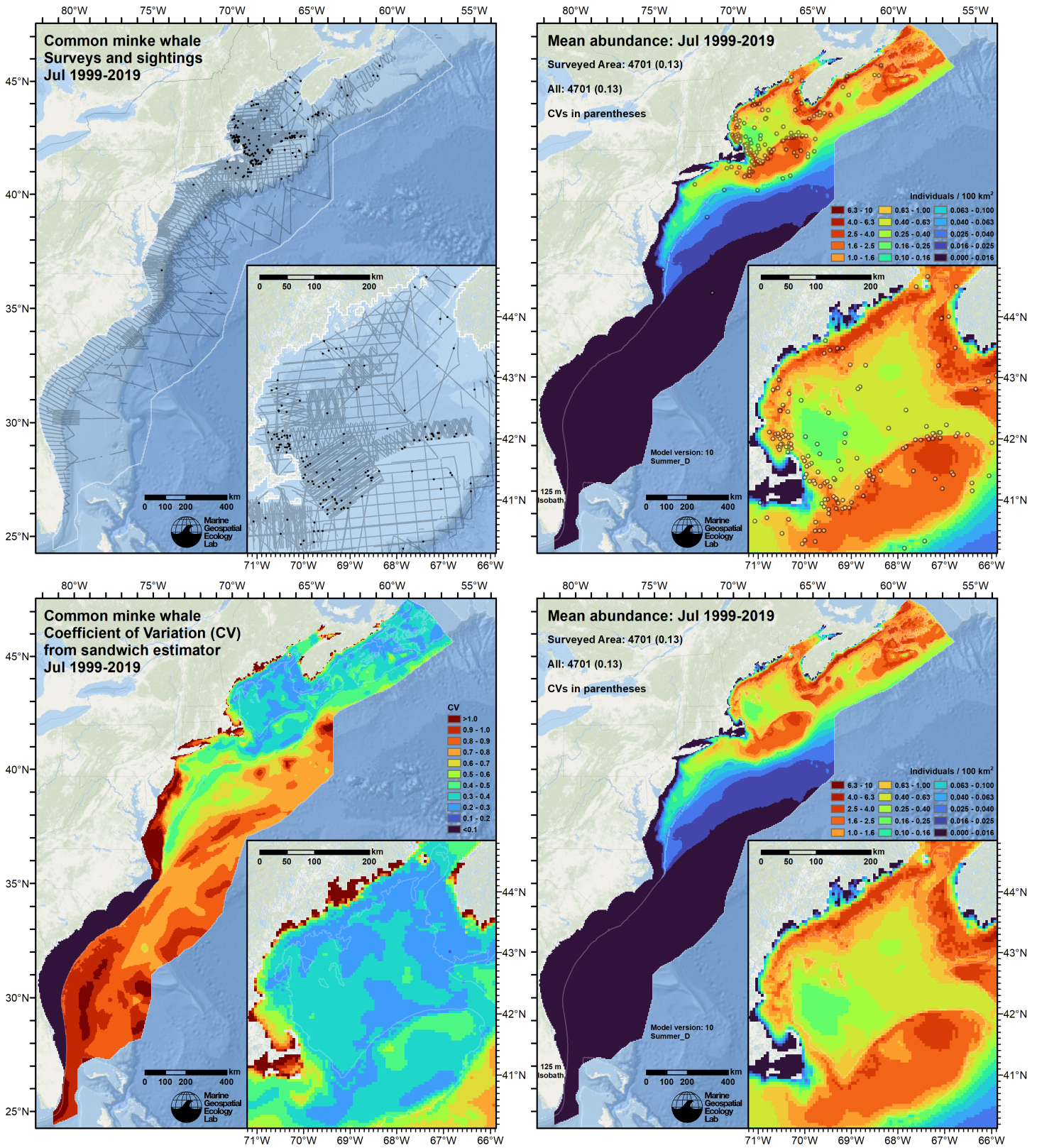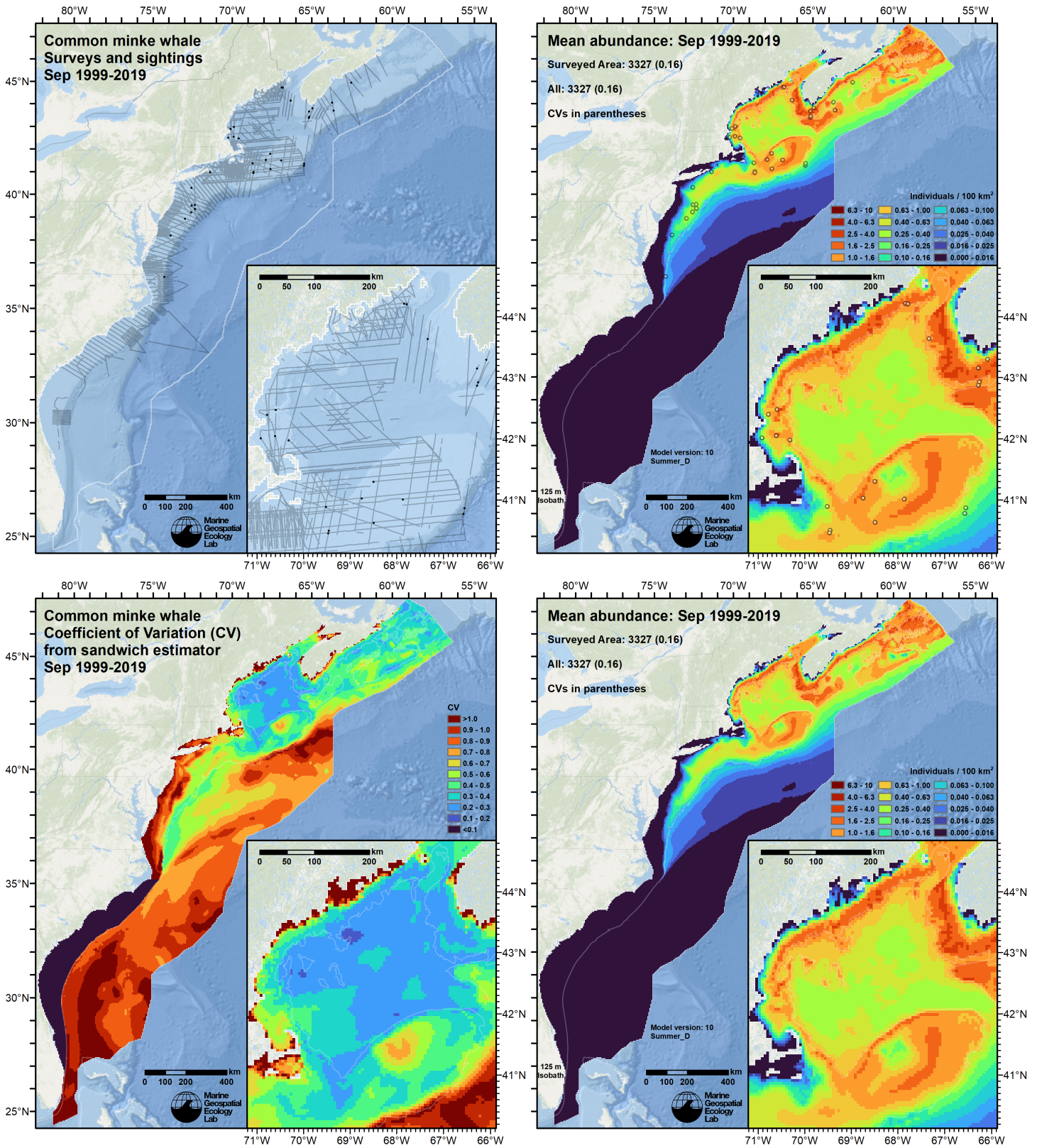
Figure 66: Survey effort and observations (top left), predicted density with observations (top right), predicted density without observations (bottom right), and coefficient of variation of predicted density (bottom left), for the month of September for the given era. Variance was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for temporal variability in dynamic covariates.
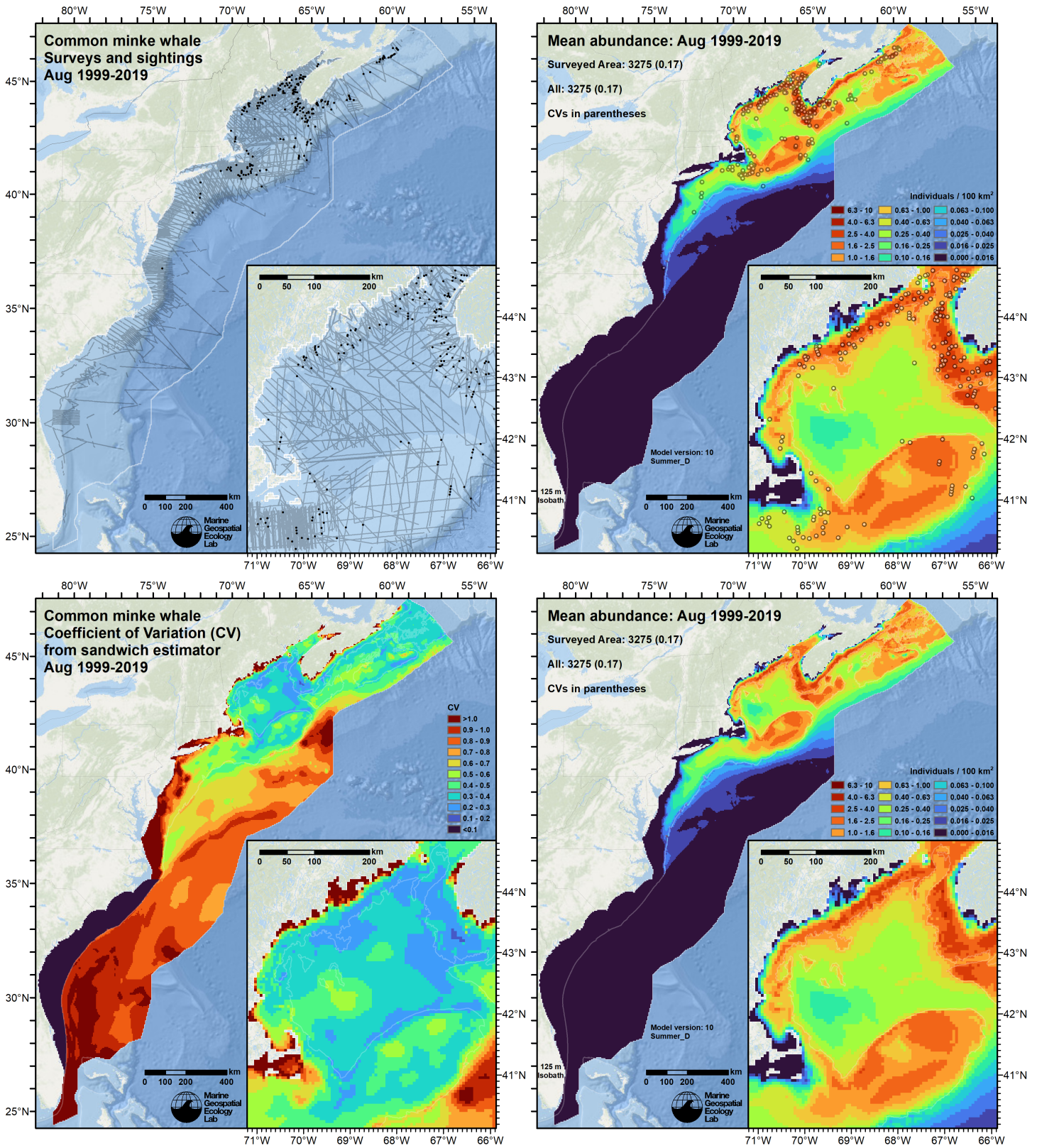
Figure 67: Survey effort and observations (top left), predicted density with observations (top right), predicted density without observations (bottom right), and coefficient of variation of predicted density (bottom left), for the month of October for the given era. Variance was estimated with the analytic approach given by Miller et al. (2022), Appendix S1, and accounts both for uncertainty in model parameter estimates and for temporal variability in dynamic covariates.
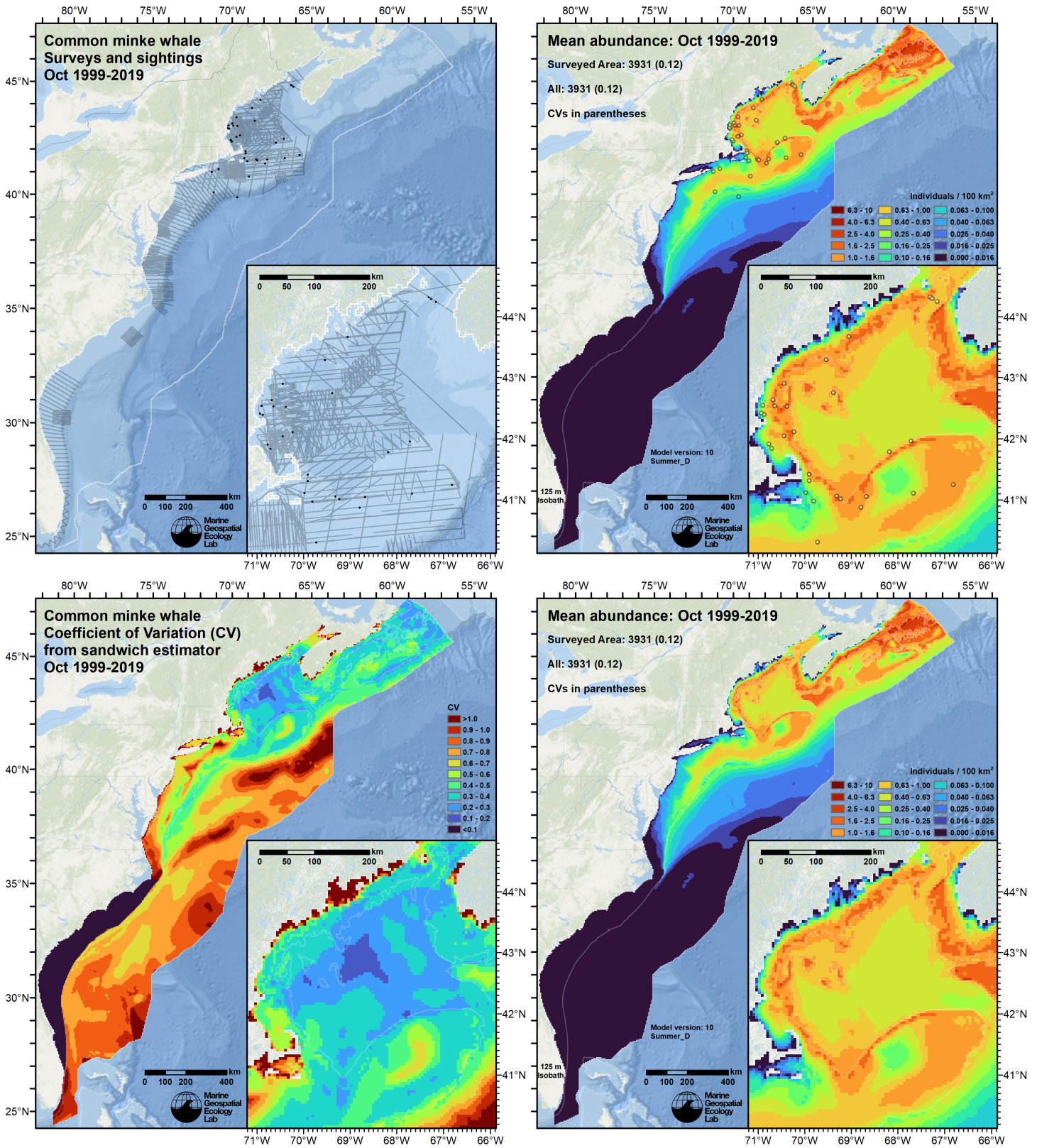
5.2 Abundance Comparisons

5.2.1 NOAA Stock Assessment Reports

Table 10: Comparison of regional abundance estimates from the 2021 NOAA Stock Assessment Report (SAR) (Hayes et al. (2022)) to estimates from this density model extracted from roughly comparable zones (Figure 68 below). The SAR estimates were based on a single year of surveying, while the model estimates were taken from the multi-year mean density surfaces we provide to model users (Section 5.1).

2021 Stock Assessment Report			Density Model		
Month/Year	Area	N_{est}	Period	Zone	Abundance
Jun-Sep 2016	Central Virginia to lower Bay of Fundy[a]	2,802	Jun-Sep 1998-2019	NEFSC	1,547
Jun-Aug 2016	Florida to Central Virginia[b]	17	Jun-Aug 1998-2019	SEFSC	82
Aug-Sep 2016	GSL/Bay of Fundy/Scotian Shelf[c]	6,158	Jun-Sep 1998-2019	Canada[d]	2,136
Jun-Sep 2016	Total	8,977	Jun-Sep 1998-2019	Total	3,765

[a] Estimate originally from Palka (2020).

[b] Estimate originally from Garrison (2020). The SAR excluded waters south of Central Virginia from its table of abundance estimates, presumably because abundance was negligible compared to the estimates for other regions.

[c] Estimate originally from Lawson and Gosselin (2018).

[d] Our Canada zone is roughly comparable to the SAR's "Bay of Fundy/Scotian Shelf" area but does not include the Gulf of St. Lawrence (GSL), where Lawson and Gosselin (2018) reported over half of their sightings (see discussion in Section 6).
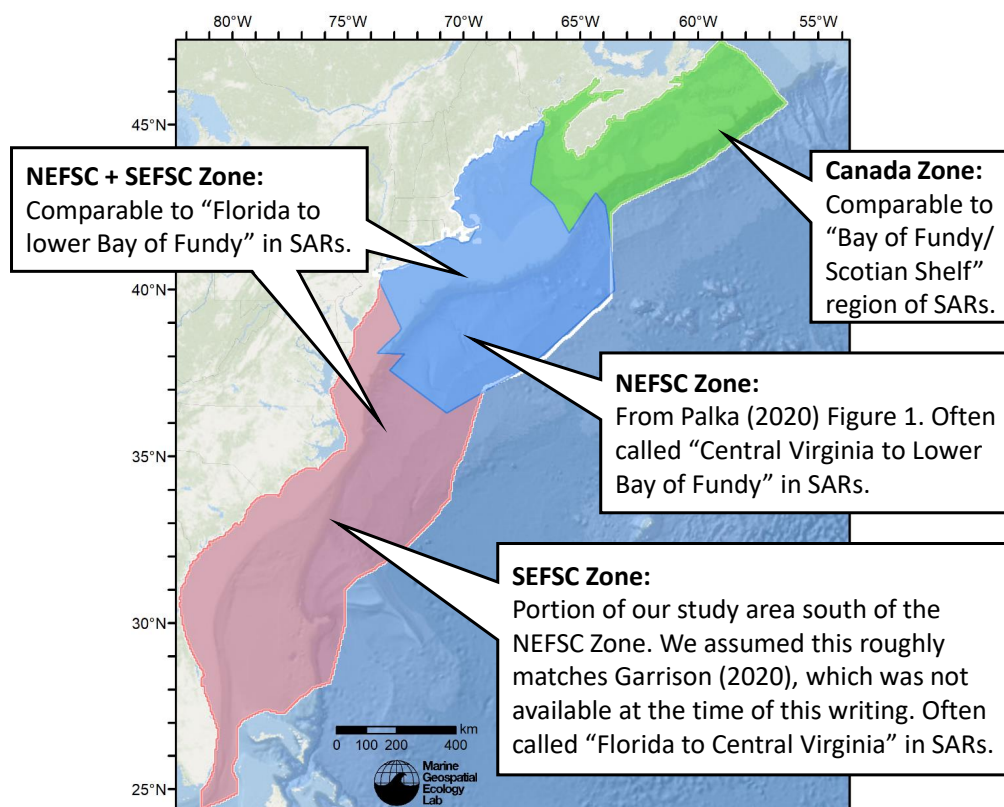
Figure 68: Zones for which we extracted abundance estimates from the density model for comparison to estimates from the NOAA Stock Assessment Report.

5.2.2 Previous Density Model
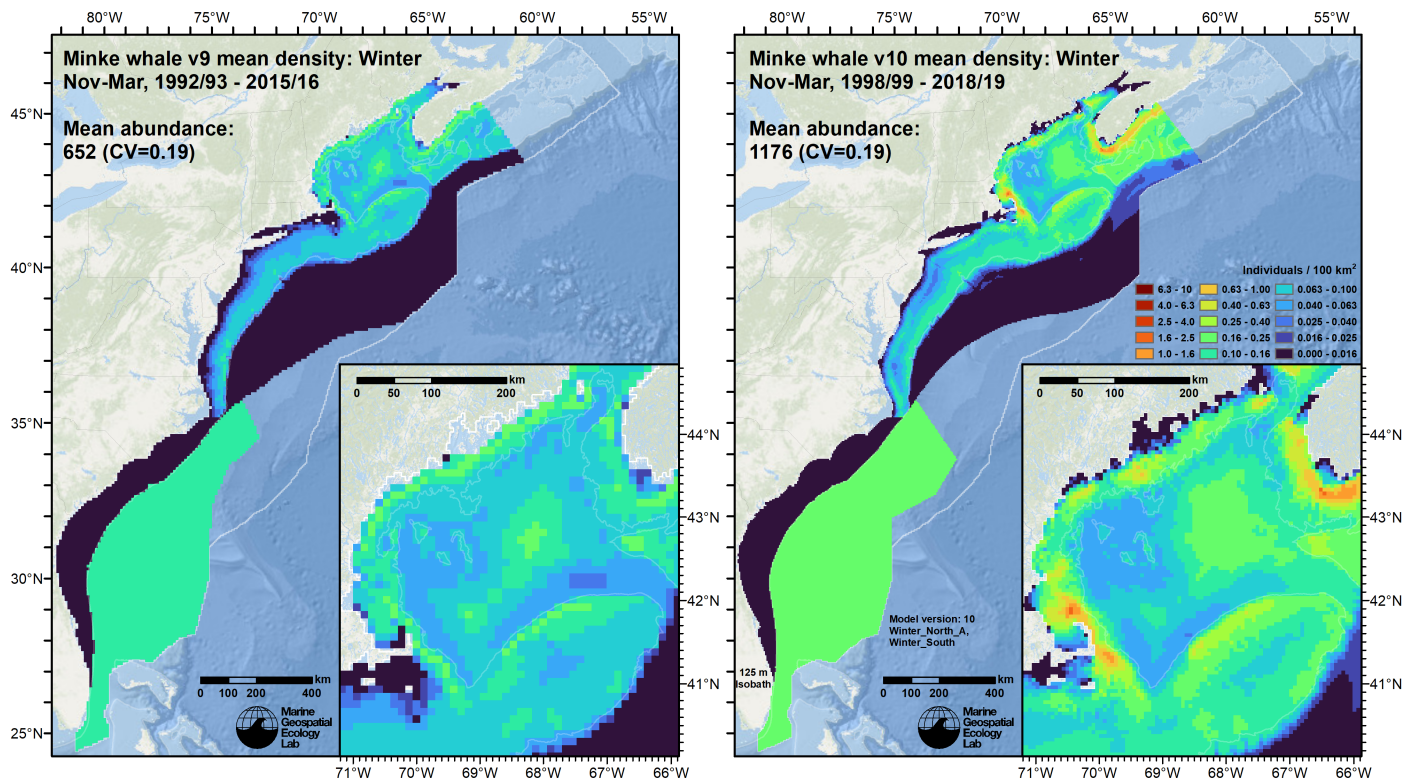
Figure 69: Comparison of the mean density predictions from the previous model (left) released by Roberts et al. (2017) to those from this model (right) for the Winter season.
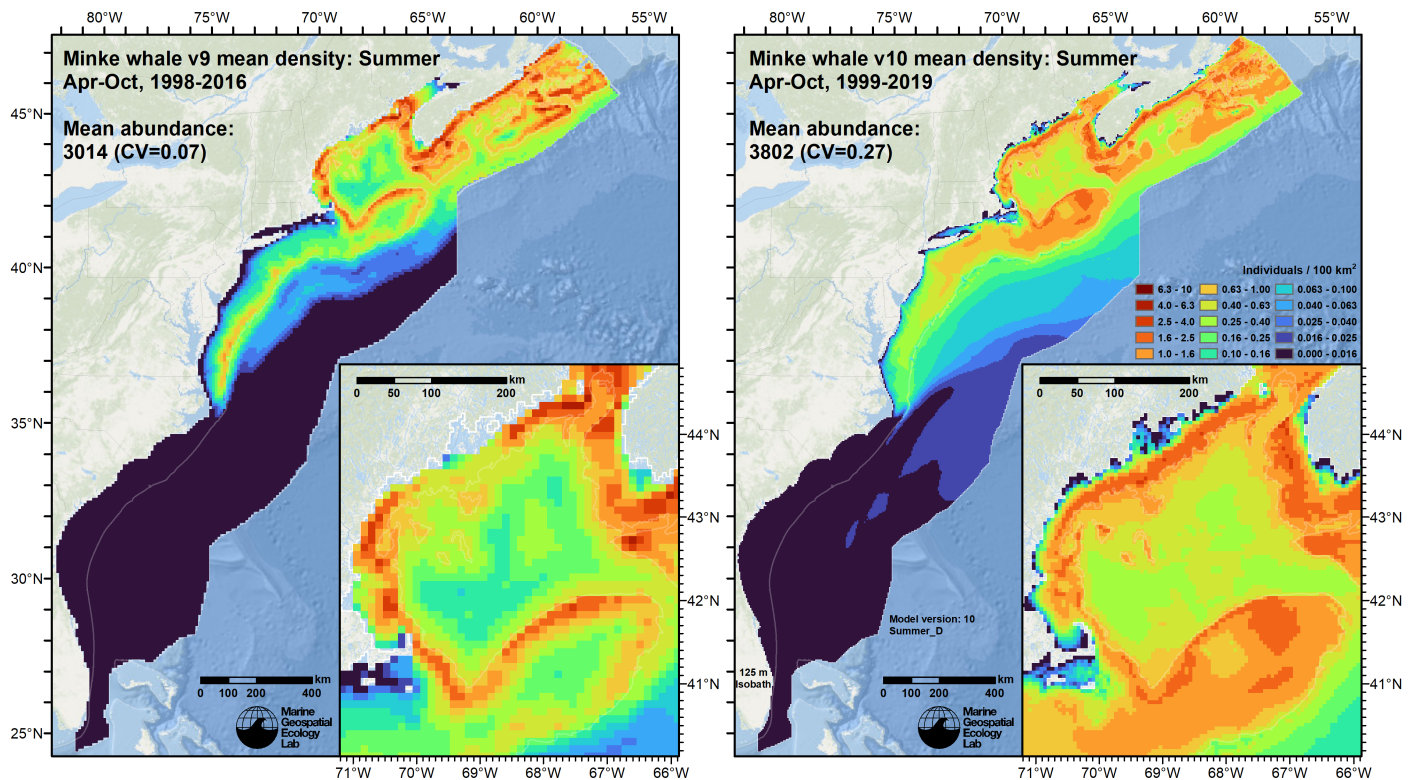
Figure 70: Comparison of the mean density predictions from the previous model (left) released by Roberts et al. (2017) to those from this model (right) for the Summer season.

6 Discussion

Minke whales are believed to follow a typical baleen whale migration pattern, moving to high latitudes in summer to feed and low latitudes in winter to breed or calve (Risch et al. 2014). Over the past two decades, passive acoustic monitoring has begun to tease out the details of these distributional dynamics, but several challenges have complicated this effort. Minke whales have a large vocal repertoire, but the behavioral function of many of these vocalizations remains unknown, owing to a lack of contemporaneous visual observations (Risch 2022). There appear to be substantial geographic variations in production of certain vocalizations, which may be due simply to seasonal variability in those vocalizations, or to those vocalizations being produced mainly by one segment of the population (e.g. mature males) that has a different geographical distribution than other segments (e.g. mature females) during certain seasons.

The lack of knowledge regarding these details, as well as the difficulty in incorporating the full vocal repertoire into distributional studies, has limited the research community's ability to fully characterize minke whale distributions in the North Atlantic with passive acoustic monitoring. For example, Risch et al. (2014) elucidated minke whale migration patterns and seasonal distribution in the North Atlantic via detection of specific pulse trains at hydrophones deployed from the tropical North Atlantic to the Davis Strait and southwest Iceland. This study tallied a substantial number of vocalizations in fall, winter, and spring, but almost none in summer, despite regular presence of minke whales observed visually at traditional summer feeding habitats. Similarly, Delarue et al. (2022) monitored waters of eastern Canada for baleen whales with 25 acoustic recorders deployed across the Scotian Shelf, Newfoundland, and Labrador and examined recordings for pulse trains similar to those used by Risch et al. (2014). Delarue et al. reported minke whale detections along the Scotian Shelf but none in Newfoundland or Labrador, where minke whales have been observed in abundance in summer (Lawson and Gosselin 2011, 2018).

A further complication to characterizing minke whale distributions, particularly by means of visual observation, is that, similar to fin whales, some fraction of the population appears to remain at high latitudes throughout the winter, when visual effort is typically reduced or suspended. The presence of minkes at high latitudes year-round has been confirmed for Antarctic minke whales both by acoustics and by satellite tagging (Risch (2022), and references therein). In our study area, the surveys contributed for our analysis recorded minke whales in the Gulf of Maine in every month of the year. Additionally, researchers at Halifax, Nova Scotia reported opportunistic sightings of minke whales year round (Simard et al. 2006; Bartha et al. 2011). Similarly, the OBIS-SEAMAP system (Halpin et al. 2009) cataloged several winter sightings off Westport, Nova Scotia reported by Canada DFO. (Our methodology does not allow opportunistic data to be incorporated into our analysis.)

When we summarized our model across the modeled period (1998-2019), mean monthly density maps (Figures 56-67) broadly agreed with the overall distribution and seasonal pattern described in the literature so far. In winter months, defined in our model as November through March, mean total abundance ranged from 1,034-1,352, with 676 whales predicted offshore of the southeast U.S. and the remainder predicted to be mainly over the continental shelf from Cape Hatteras to Halifax, the northernmost extent of our winter model. We caution that our model for the southeast was based on a pseudo-stratified approach that was prone to bias if density was, in fact, not distributed uniformly as the model assumed. Kiehbadroudinezhad et al. (2021) developed an index of relative abundance from acoustic detections at several locations across the Blake Plateau from November 2017 to November 2019 and reported that the index showed a gradient with fewer whales in the west and north, which suggests that our model's assumption did not hold. Our model was based on sightings obtained mainly at the western edge of the region, which raises the prospect that density could be higher to the east.

As in our prior model (Roberts et al. 2017), we studied alternative model designs for winter for the southern offshore area. For example, we tested habitat-based density surface models with various environmental covariates, both with and without sightings from the north in addition to those in the south, but all of these models strongly concentrated minke whale density into the Gulf Stream, which flows along the western side of the Blake Plateau. Risch et al. (2014) advanced a hypothesis that minke whales might use the Gulf Stream during migrations north in the spring. Our alternative models, showing density concentrated in the Gulf Stream, might be consistent with that hypothesis but would be contrary to the findings of Kiehbadroudinezhad et al. (2021), who inferred a larger number of whales to the east side of the Blake Plateau, away from the Gulf Stream. We believe our uniform density model is a reasonable treatment of the data currently available, but strongly urge caution and that additional surveying be performed over the Blake Plateau and other offshore waters in winter.

North of Cape Hatteras in winter, densities were highest near Cape Cod, where sightings occurred every month, and the south and east of Nova Scotia, where no sightings were reported but survey effort was very sparse. Although the opportunistic sightings reported around Nova Scotia confirm minke whales were present, we note that the model's prediction is largely a geographic extrapolation of species-environment relationships fitted mainly to data collected further south, and therefore urge caution with this prediction.

In summer, our model predicted a sweep of density up the coast from April to June, peaking at over 5000 whales in June, followed by slight decline through September, with a slight rise again in October (Figure 55; Table 9). Predictions for April to June included low to moderate density predicted beyond the continental shelf, including south of the Gulf Stream. This prediction is speculative, given the lack of off-shelf surveying during these months, but is consistent with acoustic

detections of minke whales in deep waters of the temperate North Atlantic in spring (Nieukirk et al. 2004; Risch et al. 2014; Kiehbadroudinezhad et al. 2021; Kowarski et al. 2022) and with the report of on-effort visual sightings far beyond the shelf break in July 1998 and July 2016, used in this model (see Section 4.2). Density was predicted closer to shore than for some of our other baleen whale models; these predictions are supported by reports of minke whales in near shore throughout the study area, such as near Montauk, New York (Chou et al. 2022), Long Island Sound (Collins 2015), the lower Piscataqua River (Zoidis et al. 2022), the entrance to Passamaquoddy Bay (Ingram et al. 2007), as well as other sightings archived in the OBIS-SEAMAP system (Halpin et al. 2009).

Given the general match between the model's predictions and what has been reported in the literature, the knowledge gaps discussed above notwithstanding, we elected to offer density predictions for this species at monthly temporal resolution.

Comparison of our model's abundance estimates to those of the most recent NOAA Stock Assessment Report (SAR) is complicated by differences in the regions given by the SAR and the regions over which we modeled density. The most directly comparable region, "Central Virginia to lower Bay of Fundy", was estimated by Palka (2020) to have an abundance of 2,802 (CV=0.81; 95% CI 696-11,273). Our estimate of 1,547 was 45% lower but within the wide confidence limits Palka's estimate. The SAR did not list an estimate for the "Florida to Central Virginia" region, but Garrison (2020), whose estimates were used for most of the other species in the 2021 SARs, did make an estimate of 17 minke whales, based on the single offshore sighting made in July 2016 (see Section 4.2). We presume this estimate was omitted from the SAR because it was negligible compared to the estimates for other regions. In any case, our estimate of 82 was nearly five times higher, but still a small fraction of those of other regions. Finally, for Canada, the SAR listed the estimate of Lawson and Gosselin (2018) for the "Gulf of St. Lawrence/Bay of Fundy/Scotian Shelf" region of 6,158. Our study area only covered the Bay of Fundy and the Scotian Shelf but Lawson and Gosselin (2018) did not give estimates for those individual strata. Prorating their estimate by the number of sightings in those two strata (given in their Table 9, not counting sightings missing distances or cluster sizes), to exclude those in the Gulf of St. Lawrence, gives 2,998 whales. Prorating their estimate by stratum area (given in their Table 4) to the just those two strata gives 3,389 whales. By comparison, our model estimated 2,136 for this region, 29 or 37% lower, depending on which proration method is used.

Tallying up the SAR's estimate for U.S. waters plus the prorated estimates for Canada, the total abundance is 5,817 or 6,208; our estimate for the same months is 35 or 39% lower. Several reasons could explain this difference, including: the SAR estimate could have been made during an anomalously high year (2016) relative to our long term average (1998-2019); the seasonal peak in abundance estimated by our model, 5,154 minkes in June, could be occurring too early in summer; or the minke population could have grown over our modeled period. Diagnosing the exact reasons for the difference is difficult and beyond the scope of our analysis, but in any case, we caution that our model might be underestimating minke whale density and abundance.

Compared to our prior model, this model's seasonal abundance was 80% higher in winter (1,176 vs. 652) (Figure 69) and 26% higher in summer (3,802 vs. 3,014) (Figure 70). In winter, higher density was predicted offshore in the southeast and widely across the continental shelf north of Cape Hatteras, with hotspots around Cape Cod and southeast Nova Scotia. In summer, higher density was predicted beyond the continental shelf, mainly occurring in early months (i.e., spring); we believe this density is plausible based on acoustics, as discussed above, but urge additional surveying of off-shelf waters in spring. Across the continental shelf north of Cape Hatteras, more diffuse but generally higher density was predicted, rather than being as concentrated along the 125 m isobath. Density was markedly lower along the mid-Atlantic shelf break, but notably higher at the mid-shelf north of Hudson Canyon through Georges Bank. Finally, we note that the prior model's CV of mean abundance (0.07) was substantially lower than that (0.27) of the new model. The CV of the prior model was unrealistically low because that model only accounted for uncertainty in model parameter estimates, while the new model also accounts for seasonal and interannual variability in predictions.

References

Barco SG, Burt L, DePerte A, Digiovanni R Jr. (2015) Marine Mammal and Sea Turtle Sightings in the Vicinity of the Maryland Wind Energy Area July 2013-June 2015, VAQF Scientific Report #2015-06. Virginia Aquarium & Marine Science Center Foundation, Virginia Beach, VA

Bartha G, Gowans S, Simard P, Tetley MJ, Keith EO (2011) Population Size and Site Fidelity of North Atlantic Minke Whales (Balaenoptera acutorostrata acutorostrata) off the Atlantic Coast of Nova Scotia, Canada. Aquatic Mammals 37:454–463. doi: 10.1578/AM.37.4.2011.454

Becker JJ, Sandwell DT, Smith WHF, Braud J, Binder B, Depner J, Fabre D, Factor J, Ingalls S, Kim S-H, Ladner R, Marks K, Nelson S, Pharaoh A, Trimmer R, Von Rosenberg J, Wallace G, Weatherall P (2009) Global Bathymetry and Elevation Data at 30 Arc Seconds Resolution: SRTM30_PLUS. Marine Geodesy 32:355–371. doi: 10.1080/01490410903297766

Behrenfeld MJ, Boss E, Siegel DA, Shea DM (2005) Carbon-based ocean productivity and phytoplankton physiology from space. Global Biogeochem Cycles 19:1–14. doi: 10.1029/2004GB002299

Brasnett B (2008) The impact of satellite retrievals in a global sea-surface-temperature analysis. Quarterly Journal of the Royal Meteorological Society 134:1745–1760. doi: 10.1002/qj.319

Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L (2001) Introduction to Distance Sampling: Estimating Abundance of Biological Populations. Oxford University Press, Oxford, UK

Burt ML, Borchers DL, Jenkins KJ, Marques TA (2014) Using mark-recapture distance sampling methods on line transect surveys. Methods in Ecology and Evolution 5:1180–1191. doi: 10.1111/2041-210X.12294

Canada Meteorological Center (2012) GHRSST Level 4 CMC0.2deg Global Foundation Sea Surface Temperature Analysis Version 2.0. PODAAC, CA, USA. doi: 10.5067/GHCMC-4FM02

Canada Meteorological Center (2016) GHRSST Level 4 CMC0.1deg Global Foundation Sea Surface Temperature Analysis Version 3.0. PODAAC, CA, USA. doi: 10.5067/GHCMC-4FM03

Canny JF (1986) A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 8:679–698. doi: 10.1016/B978-0-08-051581-6.50024-6

Chassignet E, Hurlburt H, Metzger EJ, Smedstad O, Cummings J, Halliwell G, Bleck R, Baraille R, Wallcraft A, Lozano C, Tolman H, Srinivasan A, Hankin S, Cornillon P, Weisberg R, Barth A, He R, Werner F, Wilkin J (2009) US GODAE: Global Ocean Prediction with the HYbrid Coordinate Ocean Model (HYCOM). Oceanog 22:64–75. doi: 10.5670/oceanog.2009.39

Chou E, Rekdahl ML, Kopelman AH, Brown DM, Sieswerda PL, DiGiovanni RA, Rosenbaum HC (2022) Occurrence of baleen whales in the New York Bight, 1998–2017: Insights from opportunistic data. J Mar Biol Ass 102:438–444. doi: 10.1017/S0025315422000716

Cole T, Gerrior P, Merrick RL (2007) Methodologies of the NOAA National Marine Fisheries Service Aerial Survey Program for Right Whales (Eubalaena glacialis) in the Northeast U.S., 1998-2006. U.S. Department of Commerce, Woods Hole, MA

Collins D (2015) Whales welcomed after a long hiatus. Associated Press

Cotter MP (2019) Aerial Surveys for Protected Marine Species in the Norfolk Canyon Region: 2018–2019 Final Report. HDR, Inc., Virginia Beach, VA

Delarue JJ-Y, Moors-Murphy H, Kowarski KA, Davis GE, Urazghildiiev IR, Martin SB (2022) Acoustic occurrence of baleen whales, particularly blue, fin, and humpback whales, off eastern Canada, 2015-2017. Endang Species Res 47:265–289. doi: 10.3354/esr01176

Doniol-Valcroze T, Berteaux D, Larouche P, Sears R (2007) Influence of thermal fronts on habitat selection by four rorqual whale species in the Gulf of St. Lawrence. Marine Ecology Progress Series 335:207–216.

Foley HJ, Paxton CGM, McAlarney RJ, Pabst DA, Read AJ (2019) Occurrence, Distribution, and Density of Protected Species in the Jacksonville, Florida, Atlantic Fleet Training and Testing (AFTT) Study Area. Duke University Marine Lab, Beaufort, NC

Garrison LP (2020) Abundance of cetaceans along the southeast U.S. East coast from a summer 2016 vessel survey. PRD Contribution # PRD-2020-04. NOAA National Marine Fisheries Service, Southeast Fisheries Science Center, Miami, FL

Garrison LP, Martinez A, Maze-Foley K (2010) Habitat and abundance of cetaceans in Atlantic Ocean continental slope waters off the eastern USA. Journal of Cetacean Research and Management 11:267–277.

Halpin P, Read A, Fujioka E, Best B, Donnelly B, Hazen L, Kot C, Urian K, LaBrecque E, Dimatteo A, Cleary J, Good C, Crowder L, Hyrenbach KD (2009) OBIS-SEAMAP: The World Data Center for Marine Mammal, Sea Bird, and Sea Turtle Distributions. Oceanography 22:104–115. doi: 10.5670/oceanog.2009.42

Hayes SA, Josephson E, Maze-Foley K, Rosel PE, Wallace J, Brossard A, Chavez-Rosales S, Cole TVN, Garrison LP, Hatch J, Henry A, Horstman SC, Litz J, Lyssikatos MC, Mullin KD, Murray K, Orphanides C, Ortega-Ortiz J, Pace RM, Palka DL, Powell J, Rappucci G, Soldevilla M, Wenzel FW (2022) US Atlantic and Gulf of Mexico Marine Mammal Stock Assessments 2021. NOAA National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA

Ingram SN, Walshe L, Johnston D, Rogan E (2007) Habitat partitioning and the influence of benthic topography and oceanography on the distribution of fin and minke whales in the Bay of Fundy, Canada. Journal of the Marine Biological Association of the United Kingdom 87:149–156. doi: 10.1017/S0025315407054884

Kiehbadroudinezhad S, Bruce Martin S, Mills Flemming J (2021) Estimating minke whale relative abundance in the North Atlantic using passive acoustic sensors. The Journal of the Acoustical Society of America 150:3569–3580. doi: 10.1121/10.0007063

Kowarski KA, Martin SB, Maxner EE, Lawrence CB, Delarue JJ-Y, Miksis-Olds JL (2022) Cetacean acoustic occurrence on the US Atlantic Outer Continental Shelf from 2017 to 2020. Marine Mammal Science mms.12962. doi: 10.1111/mms.12962

Laake JL, Calambokidis J, Osmek SD, Rugh DJ (1997) Probability of Detecting Harbor Porpoise From Aerial Surveys: Estimating g(0). Journal of Wildlife Management 61:63–75. doi: 10.2307/3802415

Lawson JW, Gosselin J-F (2011) Fully-corrected cetacean abundance estimates from the Canadian TNASS survey. National Marine Mammal Peer Review Meeting, Ottawa, Canada,

Lawson JW, Gosselin J-F (2018) Estimates of cetacean abundance from the 2016 NAISS aerial surveys of eastern Canadian waters, with a comparison to estimates from the 2007 TNASS. NAMMCO SC/25/AE/09. In: Proceedings of the NAMMCO 25th Scientific Committee (SC). North Atlantic Marine Mammal Commission, Bergen-Tromsø, Norway,

Lehodey P, Senina I, Murtugudde R (2008) A spatial ecosystem and populations dynamics model (SEAPODYM)–Modeling of tuna and tuna-like populations. Progress in Oceanography 78:304–318. doi: 10.1016/j.pocean.2008.06.004

Lehodey P, Conchon A, Senina I, Domokos R, Calmettes B, Jouanno J, Hernandez O, Kloser R (2015) Optimization of a micronekton model with acoustic data. ICES Journal of Marine Science 72:1399–1412. doi: 10.1093/icesjms/fsu233

Leiter S, Stone K, Thompson J, Accardo C, Wikgren B, Zani M, Cole T, Kenney R, Mayo C, Kraus S (2017) North Atlantic right whale Eubalaena glacialis occurrence in offshore wind energy areas near Massachusetts and Rhode Island, USA. Endang Species Res 34:45–59. doi: 10.3354/esr00827

Mallette SD, Lockhart GG, McAlarney RJ, Cummings EW, McLellan WA, Pabst DA, Barco SG (2014) Documenting Whale Migration off Virginia's Coast for Use in Marine Spatial Planning: Aerial and Vessel Surveys in the Proximity of the Virginia Wind Energy Area (VA WEA), VAQF Scientific Report 2014-08. Virginia Aquarium & Marine Science Center Foundation, Virginia Beach, VA

Mallette SD, Lockhart GG, McAlarney RJ, Cummings EW, McLellan WA, Pabst DA, Barco SG (2015) Documenting Whale Migration off Virginia's Coast for Use in Marine Spatial Planning: Aerial Surveys in the Proximity of the Virginia Wind Energy Area (VA WEA) Survey/Reporting Period: May 2014 - December 2014, VAQF Scientific Report 2015-02. Virginia Aquarium & Marine Science Center Foundation, Virginia Beach, VA

Mallette SD, McAlarney RJ, Lockhart GG, Cummings EW, Pabst DA, McLellan WA, Barco SG (2017) Aerial Survey Baseline Monitoring in the Continental Shelf Region of the VACAPES OPAREA: 2016 Annual Progress Report. Virginia Aquarium & Marine Science Center Foundation, Virginia Beach, VA

Marsh H, Sinclair DF (1989) Correcting for Visibility Bias in Strip Transect Aerial Surveys of Aquatic Fauna. The Journal of Wildlife Management 53:1017. doi: 10.2307/3809604

McAlarney R, Cummings E, McLellan W, Pabst A (2018) Aerial Surveys for Protected Marine Species in the Norfolk Canyon Region: 2017 Annual Progress Report. University of North Carolina Wilmington, Wilmington, NC

McLellan WA, McAlarney RJ, Cummings EW, Read AJ, Paxton CGM, Bell JT, Pabst DA (2018) Distribution and abundance of beaked whales (Family Ziphiidae) Off Cape Hatteras, North Carolina, U.S.A. Marine Mammal Science. doi: 10.1111/mms.12500

Meissner T, Wentz FJ, Scott J, Vazquez-Cuervo J (2016) Sensitivity of Ocean Surface Salinity Measurements From Spaceborne L-Band Radiometers to Ancillary Sea Surface Temperature. IEEE Trans Geosci Remote Sensing 54:7105–7111. doi: 10.1109/TGRS.2016.2596100

Mesgaran MB, Cousens RD, Webber BL (2014) Here be dragons: A tool for quantifying novelty due to covariate range and correlation change when projecting species distribution models. Diversity Distrib 20:1147–1159. doi: 10.1111/ddi.12209

Miller DL, Becker EA, Forney KA, Roberts JJ, Cañadas A, Schick RS (2022) Estimating uncertainty in density surface models. PeerJ 10:e13950. doi: 10.7717/peerj.13950

Mullin KD, Fulling GL (2003) Abundance of cetaceans in the southern U.S. North Atlantic Ocean during summer 1998. Fishery Bulletin 101:603–613.

Nieukirk SL, Stafford KM, Mellinger DK, Dziak RP, Fox CG (2004) Low-frequency whale and seismic airgun sounds recorded in the mid-Atlantic Ocean. The Journal of the Acoustical Society of America 115:1832–1843.

O'Brien O, Pendleton DE, Ganley LC, McKenna KR, Kenney RD, Quintana-Rizzo E, Mayo CA, Kraus SD, Redfern JV (2022) Repatriation of a historical North Atlantic right whale habitat during an era of rapid climate change. Sci Rep 12:12407. doi: 10.1038/s41598-022-16200-8

Palka D (2020) Cetacean Abundance in the US Northwestern Atlantic Ocean Summer 2016. Northeast Fish Sci Cent Ref Doc. 20-05. NOAA National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA

Palka D, Aichinger Dias L, Broughton E, Chavez-Rosales S, Cholewiak D, Davis G, DeAngelis A, Garrison L, Haas H, Hatch J, Hyde K, Jech M, Josephson E, Mueller-Brennan L, Orphanides C, Pegg N, Sasso C, Sigourney D, Soldevilla M, Walsh H (2021) Atlantic Marine Assessment Program for Protected Species: FY15 – FY19 (OCS Study BOEM 2021-051). U.S. Deptartment of the Interior, Bureau of Ocean Energy Management, Washington, DC

Palka DL (2006) Summer abundance estimates of cetaceans in US North Atlantic navy operating areas (NEFSC Reference Document 06-03). U.S. Department of Commerce, Northeast Fisheries Science Center, Woods Hole, MA

Palka DL, Chavez-Rosales S, Josephson E, Cholewiak D, Haas HL, Garrison L, Jones M, Sigourney D, Waring G, Jech M, Broughton E, Soldevilla M, Davis G, DeAngelis A, Sasso CR, Winton MV, Smolowitz RJ, Fay G, LaBrecque E, Leiness JB, Dettloff K, Warden M, Murray K, Orphanides C (2017) Atlantic Marine Assessment Program for Protected Species: 2010-2014 (OCS Study BOEM 2017-071). U.S. Deptartment of the Interior, Bureau of Ocean Energy Management, Washington, DC

Quintana-Rizzo E, Leiter S, Cole T, Hagbloom M, Knowlton A, Nagelkirk P, O'Brien O, Khan C, Henry A, Duley P, Crowe L, Mayo C, Kraus S (2021) Residency, demographics, and movement patterns of North Atlantic right whales Eubalaena glacialis in an offshore wind energy development area in southern New England, USA. Endang Species Res 45:251–268. doi: 10.3354/esr01137

Read AJ, Barco S, Bell J, Borchers DL, Burt ML, Cummings EW, Dunn J, Fougeres EM, Hazen L, Hodge LEW, Laura A-M, McAlarney RJ, Peter N, Pabst DA, Paxton CGM, Schneider SZ, Urian KW, Waples DM, McLellan WA (2014) Occurrence, distribution and abundance of cetaceans in Onslow Bay, North Carolina, USA. Journal of Cetacean Research and Management 14:23–35.

Redfern JV, Kryc KA, Weiss L, Hodge BC, O'Brien O, Kraus SD, Quintana-Rizzo E, Auster PJ (2021) Opening a Marine Monument to Commercial Fishing Compromises Species Protections. Front Mar Sci 8:645314. doi: 10.3389/fmars.2021.645314

Risch D (2022) Mysterious Minke Whales: Acoustic Diversity and Variability. In: Clark CW, Garland EC (eds) Ethology and Behavioral Ecology of Mysticetes. Springer International Publishing, Cham, pp 329–348

Risch D, Castellote M, Clark CW, Davis GE, Dugan PJ, Hodge LE, Kumar A, Lucke K, Mellinger DK, Nieukirk SL, others (2014) Seasonal migrations of North Atlantic minke whales: Novel insights from large-scale passive acoustic monitoring networks. Movement Ecology 2:24.

Roberts JJ, Best BD, Dunn DC, Treml EA, Halpin PN (2010) Marine Geospatial Ecology Tools: An integrated framework for ecological geoprocessing with ArcGIS, Python, R, MATLAB, and C++. Environmental Modelling & Software 25:1197–1207. doi: 10.1016/j.envsoft.2010.03.029

Roberts JJ, Best BD, Mannocci L, Fujioka E, Halpin PN, Palka DL, Garrison LP, Mullin KD, Cole TVN, Khan CB, McLellan WA, Pabst DA, Lockhart GG (2016) Habitat-based cetacean density models for the U.S. Atlantic and Gulf of Mexico. Scientific Reports 6:22615. doi: 10.1038/srep22615

Roberts JJ, Mannocci L, Halpin PN (2017) Final Project Report: Marine Species Density Data Gap Assessments and Update for the AFTT Study Area, 2016-2017 (Opt. Year 1), Document Version 1.4. Duke University Marine Geospatial Ecology Lab, Durham, NC

Roberts JJ, Yack TM, Halpin PN (2023) Marine mammal density models for the U.S. Navy Atlantic Fleet Training and Testing (AFTT) study area for the Phase IV Navy Marine Species Density Database (NMSDD), Document Version 1.3. Duke University Marine Geospatial Ecology Lab, Durham, NC

Robertson FC, Koski WR, Brandon JR, Thomas TA, Trites AW (2015) Correction factors account for the availability of bowhead whales exposed to seismic operations in the Beaufort Sea. Journal of Cetacean Research and Management 15:35–44.

Ryan C, Boisseau O, Cucknell A, Romagosa M, Moscrop A, McLanaghan R (2013) Final report for trans-Atlantic research passages between the UK and USA via the Azores and Iceland, conducted from R/V Song of the Whale 26 March to 28 September 2012. Marine Conservation Research International, Essex, UK

Schick R, Halpin P, Read A, Urban D, Best B, Good C, Roberts J, LaBrecque E, Dunn C, Garrison L, Hyrenbach K, McLellan W, Pabst D, Palka D, Stevick P (2011) Community structure in pelagic marine mammals at large spatial scales. Marine Ecology Progress Series 434:165–181. doi: 10.3354/meps09183

Simard P, Lawlor JL, Gowans S (2006) Temporal Variability of Cetaceans near Halifax, Nova Scotia. Can Field Nat 120:93. doi: 10.22621/cfn.v120i1.252

Stone KM, Leiter SM, Kenney RD, Wikgren BC, Thompson JL, Taylor JKD, Kraus SD (2017) Distribution and abundance of cetaceans in a wind energy development area offshore of Massachusetts and Rhode Island. J Coast Conserv 21:527–543. doi: 10.1007/s11852-017-0526-4

Torres LG, Mclellan WA, Meagher E, Pabst DA (2005) Seasonal distribution and relative abundance of bottlenose dolphins, Tursiops truncatus, along the US mid-Atlantic coast. Journal of Cetacean Research and Management 7:153.

Víkingsson GA, Heide-Jørgensen MP (2015) First indications of autumn migration routes and destination of common minke whales tracked by satellite in the North Atlantic during 2001–2011. Mar Mam Sci 31:376–385. doi: 10.1111/mms.12144

Westberry T, Behrenfeld MJ, Siegel DA, Boss E (2008) Carbon-based primary productivity modeling with vertically resolved photoacclimation. Global Biogeochemical Cycles 22:1–18. doi: 10.1029/2007GB003078

Wood SN (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 73:3–36. doi: 10.1111/j.1467-9868.2010.00749.x

Zoidis AM, Lomac-MacNair KS, Ireland DS, Rickard ME, McKown KA, Schlesinger MD (2021) Distribution and density of six large whale species in the New York Bight from monthly aerial surveys 2017 to 2020. Continental Shelf Research 230:104572. doi: 10.1016/j.csr.2021.104572

Zoidis AM, Olson PA, Jefferson TA, Johnson NC, Soucier CP, Bassi JH (2022) Distribution and Abundance of Marine Mammals in the Estuarine Waters of the Piscataqua River, Maine, USA. Aquat Mamm 48:25–35. doi: 10.1578/AM.48.1.2022.25