Density Model for Short-Finned Pilot Whale (*Globicephala macrorhynchus*) for the U.S. Gulf of Mexico: Supplementary Report

Duke University Marine Geospatial Ecology Lab*

Model Version 3.3 - 2015-09-30

Citation

When referencing our methodology or results generally, please cite our open-access article:

Roberts JJ, Best BD, Mannocci L, Fujioka E, Halpin PN, Palka DL, Garrison LP, Mullin KD, Cole TVN, Khan CB, McLellan WM, 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

To reference this specific model or Supplementary Report, please cite:

Roberts JJ, Best BD, Mannocci L, Fujioka E, Halpin PN, Palka DL, Garrison LP, Mullin KD, Cole TVN, Khan CB, McLellan WM, Pabst DA, Lockhart GG (2015) Density Model for Short-Finned Pilot Whale (*Globicephala macrorhynchus*) for the U.S. Gulf of Mexico Version 3.3, 2015-09-30, and Supplementary Report. Marine Geospatial Ecology Lab, Duke University, Durham, North Carolina.

Copyright and License

Revision History

Version	Date	Description of changes
1	2014-10-23	Initial version.
2	2014-11-20	Reconfigured detection hierarchy and adjusted detection functions. Removed CumVGPM180 predictor. Updated documentation.
3	2015-01-11	Added a missing sighting and refitted models.
3.1	2015-02-02	Updated the documentation. No changes to the model.
3.2	2015-05-14	Updated calculation of CVs. Switched density rasters to logarithmic breaks. No changes to the model.
3.3	2015-09-30	Updated the documentation. No changes to the model.

*For questions, or to offer feedback about this model or report, please contact Jason Roberts (jason.roberts@duke.edu)

Survey Data

Survey	Period	Length (1000 km)	Hours	Sightings
SEFSC GOMEX92-96 Aerial Surveys	1992-1996	27	152	0
SEFSC Gulf of Mexico Shipboard Surveys, 2003-2009	2003-2009	19	1156	17
SEFSC GulfCet I Aerial Surveys	1992-1994	50	257	11
SEFSC GulfCet II Aerial Surveys	1996-1998	22	124	1
SEFSC GulfSCAT 2007 Aerial Surveys	2007-2007	18	95	0
SEFSC Oceanic CetShip Surveys	1992-2001	49	3102	19
SEFSC Shelf CetShip Surveys	1994-2001	10	707	2
Total		195	5593	50

Table 2: Survey effort and sightings used in this model. Effort is tallied as the cumulative length of on-effort transects and hours the survey team was on effort. Sightings are the number of on-effort encounters of the modeled species for which a perpendicular sighting distance (PSD) was available. Off effort sightings and those without PSDs were omitted from the analysis.

Period	Length (1000 km)	Hours	Sightings
1992-2009	195	5592	50
1998-2009	62	2679	26
% Lost	68	52	48

Table 3: Survey effort and on-effort sightings having perpendicular sighting distances. % Lost shows the percentage of effort or sightings lost by restricting the analysis to surveys performed in 1998 and later, the era in which remotely-sensed chlorophyll and derived productivity estimates are available. See Figure 1 for more information.

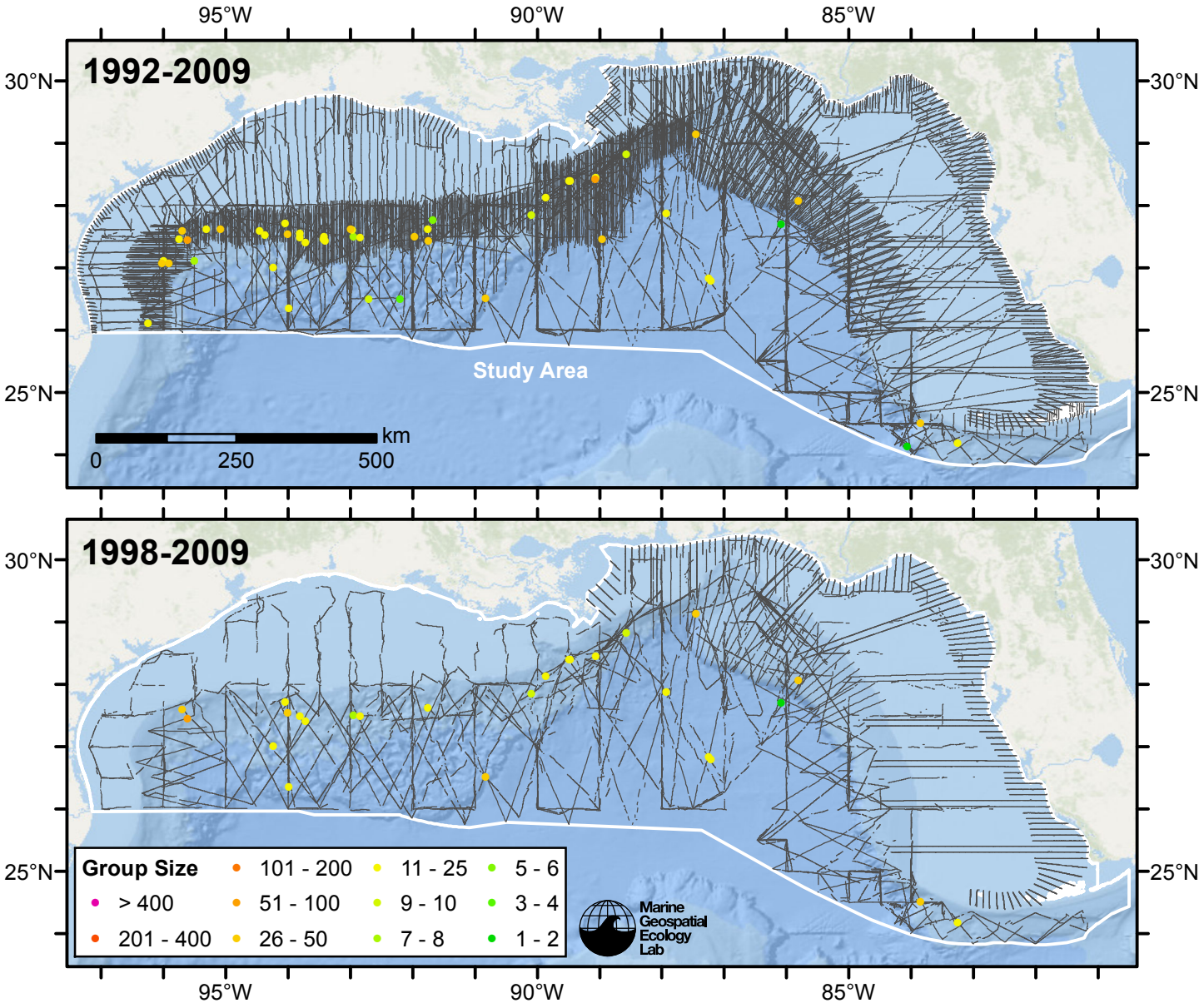


Figure 1: Pilot whales sightings and survey tracklines. The top map shows all surveys. The bottom map shows surveys performed in 1998 or later. the era in which remotely-sensed chlorophyll and derived productivity estimates are available. Models fitted to contemporaneous (day-of-sighting) estimates of those predictors only utilize these surveys. These maps illustrate the survey data lost in order to utilize those predictors. Models fitted to climatogical estimates of those predictors do not suffer this data loss.

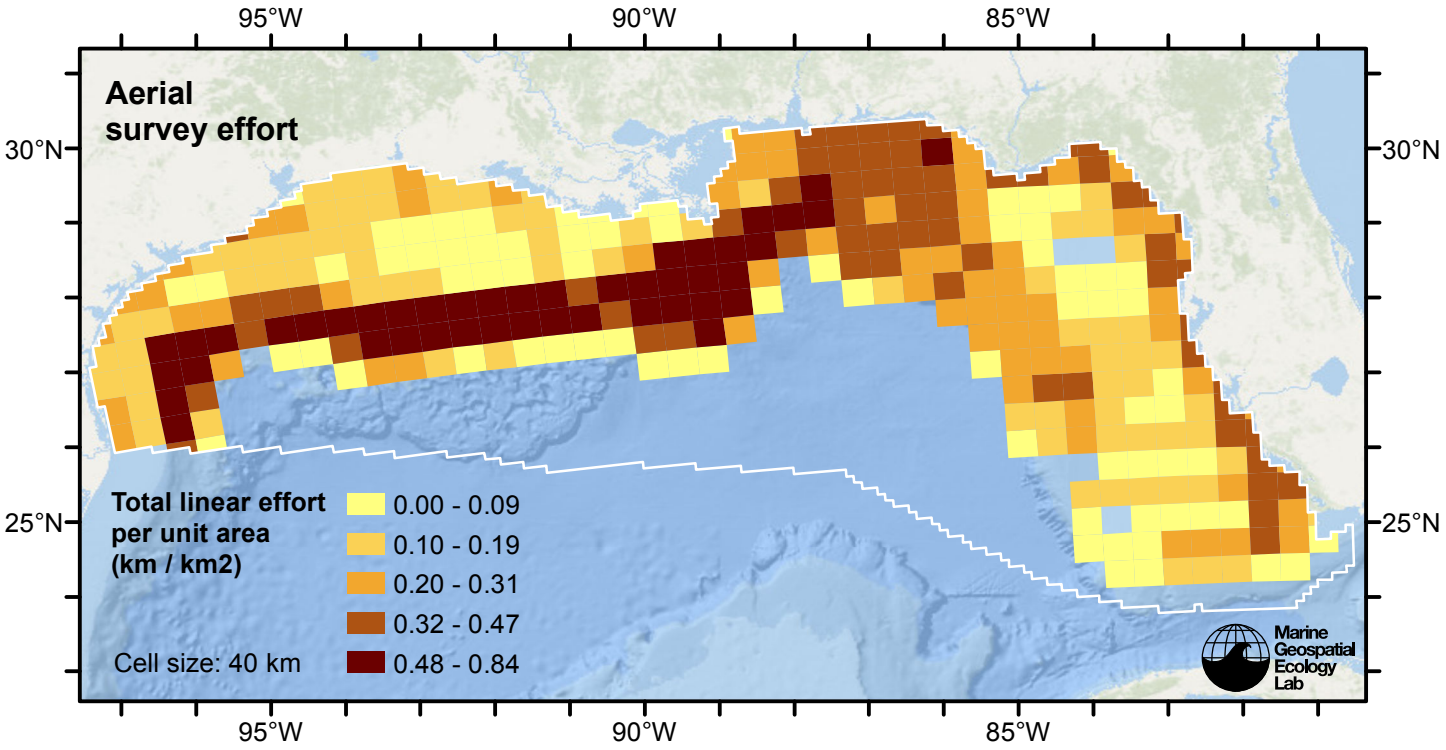


Figure 2: Aerial linear survey effort per unit area.

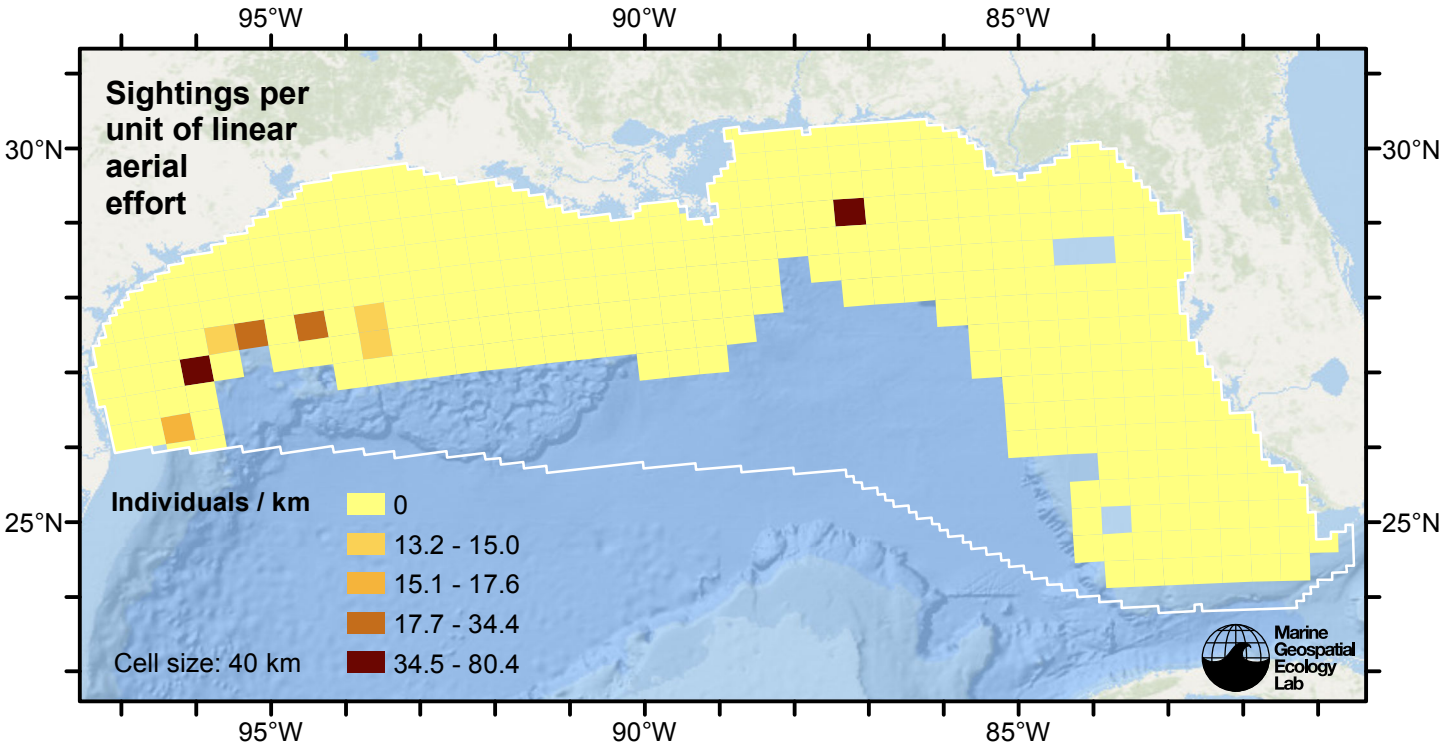


Figure 3: Pilot whales sightings per unit aerial linear survey effort.

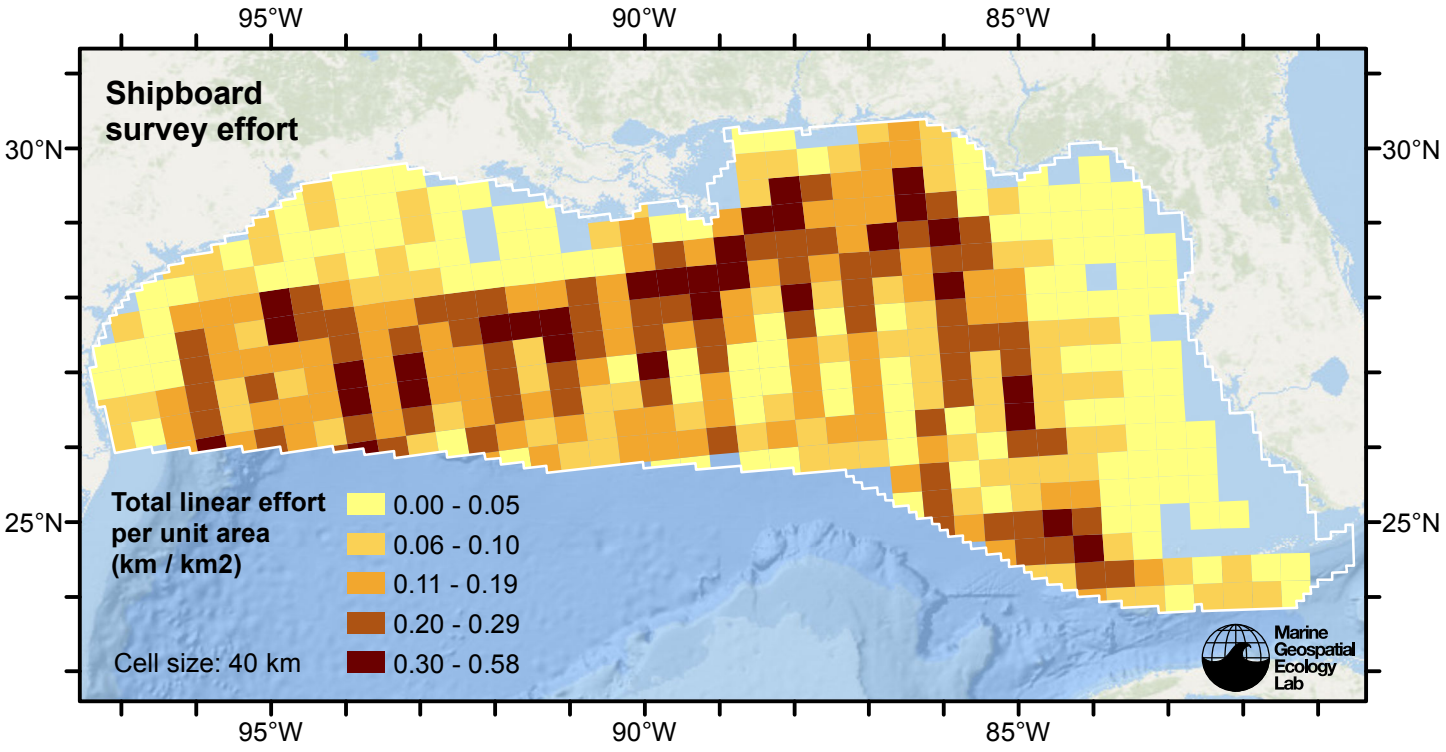


Figure 4: Shipboard linear survey effort per unit area.

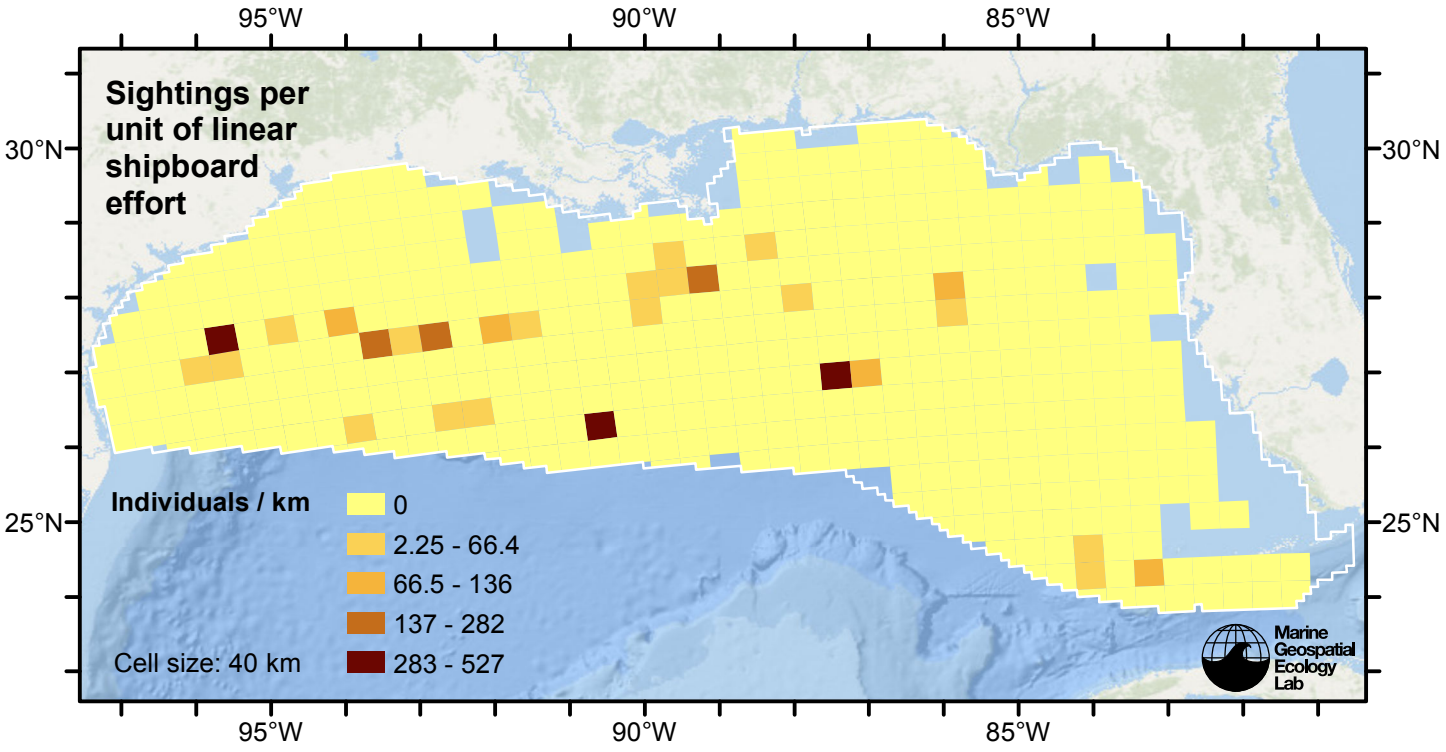


Figure 5: Pilot whales sightings per unit shipboard linear survey effort.

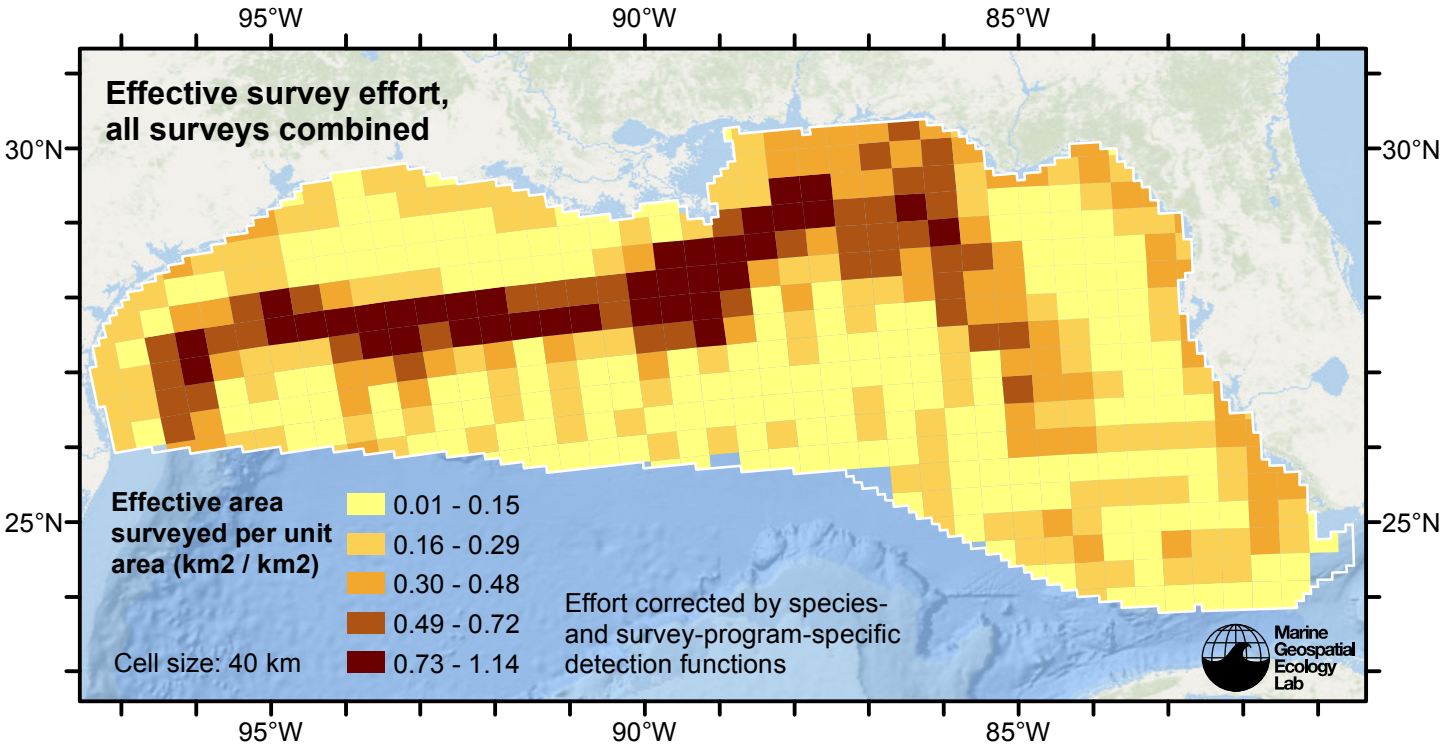


Figure 6: Effective survey effort per unit area, for all surveys combined. Here, effort is corrected by the species- and survey-program-specific detection functions used in fitting the density models.

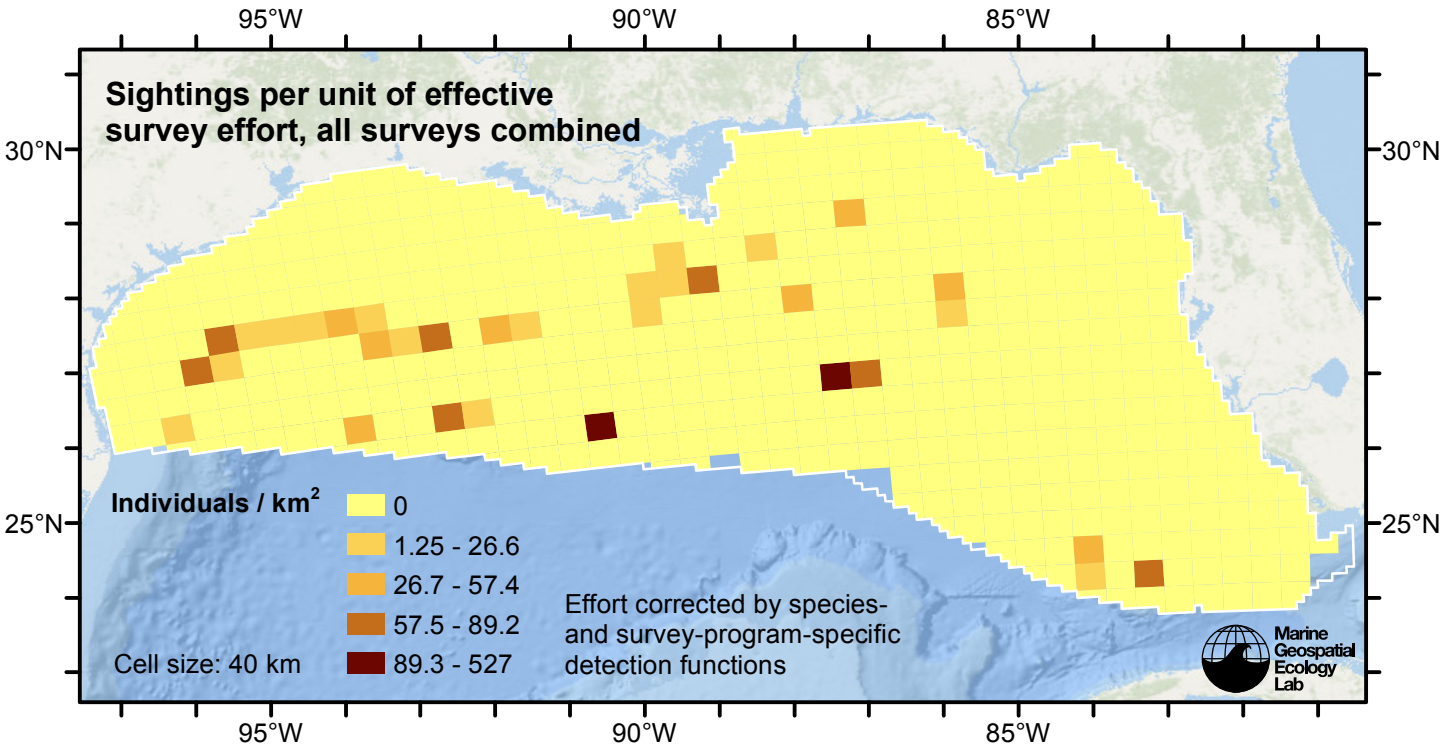


Figure 7: Pilot whales sightings per unit of effective survey effort, for all surveys combined. Here, effort is corrected by the species- and survey-program-specific detection functions used in fitting the density models.

Detection Functions

The detection hierarchy figures below show how sightings from multiple surveys were pooled to try to achieve Buckland et. al's (2001) recommendation that at least 60-80 sightings be used to fit a detection function. Leaf nodes, on the right, usually represent individual surveys, while the hierarchy to the left shows how they have been grouped according to how similar we believed the surveys were to each other in their detection performance.

At each node, the red or green number indicates the total number of sightings below that node in the hierarchy, and is colored green if 70 or more sightings were available, and red otherwise. If a grouping node has zero sightings–i.e. all of the surveys within it had zero sightings–it may be collapsed and shown as a leaf to save space.

Each histogram in the figure indicates a node where a detection function was fitted. The actual detection functions do not appear in this figure; they are presented in subsequent sections. The histogram shows the frequency of sightings by perpendicular sighting distance for all surveys contained by that node. Each survey (leaf node) recieves the detection function that is closest to it up the hierarchy. Thus, for common species, sufficient sightings may be available to fit detection functions deep in the hierarchy, with each function applying to only a few surveys, thereby allowing variability in detection performance between surveys to be addressed relatively finely. For rare species, so few sightings may be available that we have to pool many surveys together to try to meet Buckland's recommendation, and fit only a few coarse detection functions high in the hierarchy.

A blue Proxy Species tag indicates that so few sightings were available that, rather than ascend higher in the hierarchy to a point that we would pool grossly-incompatible surveys together, (e.g. shipboard surveys that used big-eye binoculars with those that used only naked eyes) we pooled sightings of similar species together instead. The list of species pooled is given in following sections.

Shipboard Surveys

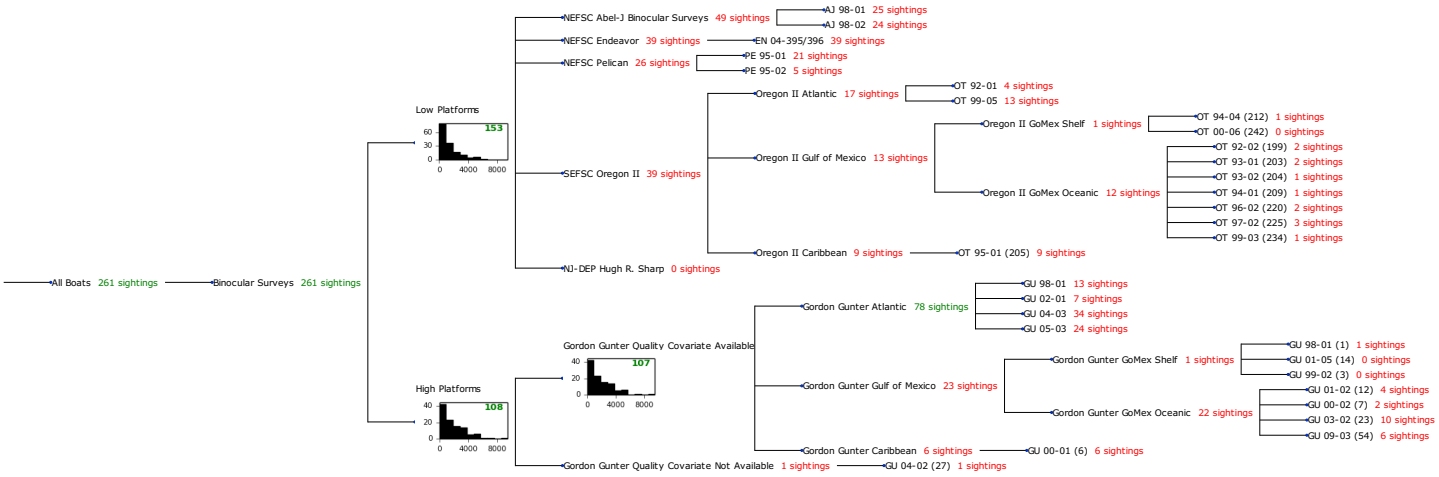


Figure 8: Detection hierarchy for shipboard surveys

Low Platforms

The sightings were right truncated at 7000m.

Covariate	Description
beaufort	Beaufort sea state.
size	Estimated size (number of individuals) of the sighted group.

Table 4: Covariates tested in candidate "multi-covariate distance sampling" (MCDS) detection functions.

Key	Adjustment	Order	Covariates	Succeeded	Δ AIC	Mean ESHW (m)
hr	poly	2		Yes	0.00	1685
hr	poly	4		Yes	0.12	1739
hr			beaufort	Yes	0.32	1804
hn	cos	2		Yes	0.86	1979
hr				Yes	1.07	1815
hr			beaufort, size	Yes	2.30	1801
hn	cos	3		Yes	2.87	1824
hn			beaufort	Yes	12.76	2408
hn				Yes	12.89	2415
hn	herm	4		Yes	14.72	2412
hr			size	No		
hn			size	No		
hn			beaufort, size	No		

Table 5: Candidate detection functions for Low Platforms. The first one listed was selected for the density model.

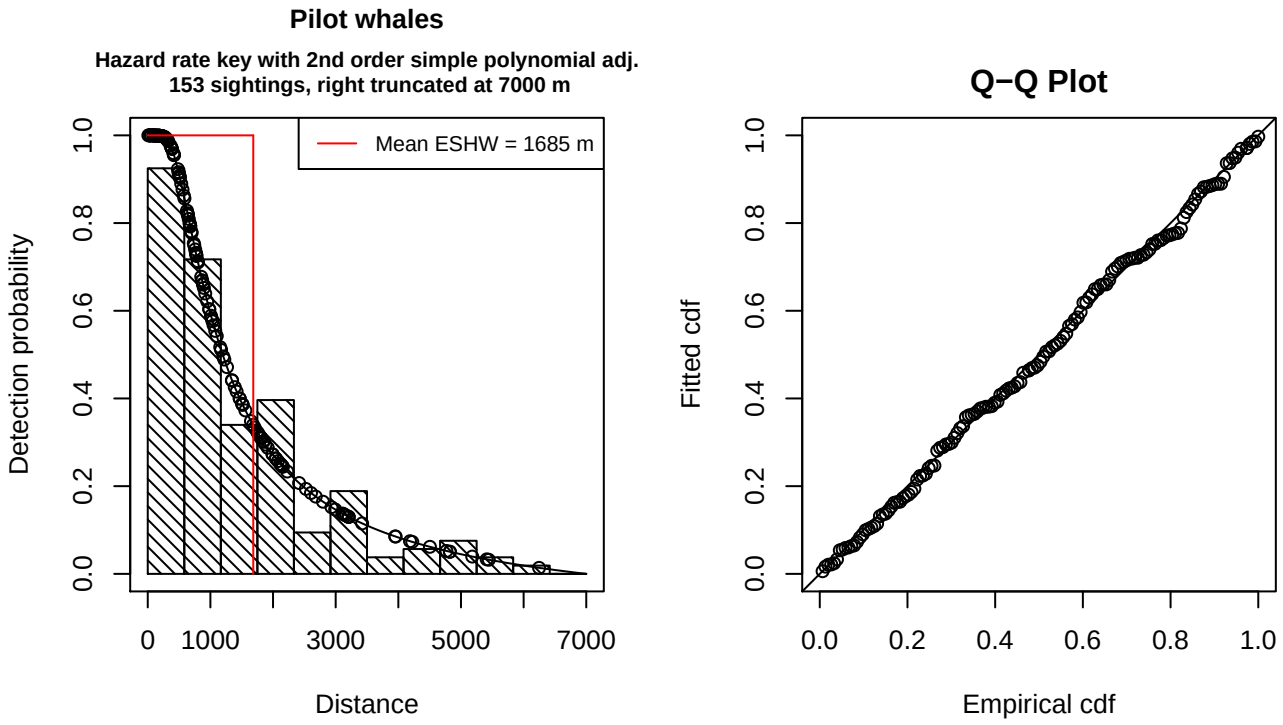


Figure 9: Detection function for Low Platforms that was selected for the density model

Statistical output for this detection function:

```
Summary for ds object
```

```
Number of observations :  153
Distance range         :  0  -  7000
AIC                    :  2525.03

Detection function:
 Hazard-rate key function with simple polynomial adjustment term of order 2

Detection function parameters
Scale Coefficients:
            estimate        se
(Intercept) 6.857564 0.2694695

Shape parameters:
             estimate        se
(Intercept) 0.3448006 0.2271376

Adjustment term parameter(s):
                estimate        se
poly, order 2 -0.9999998 0.2895039


Monotonicity constraints were enforced.
                       Estimate          SE        CV
Average p             0.2406467  0.03362653 0.1397340
N in covered region 635.7868377 99.49339784 0.1564886

Monotonicity constraints were enforced.
```

Additional diagnostic plots:

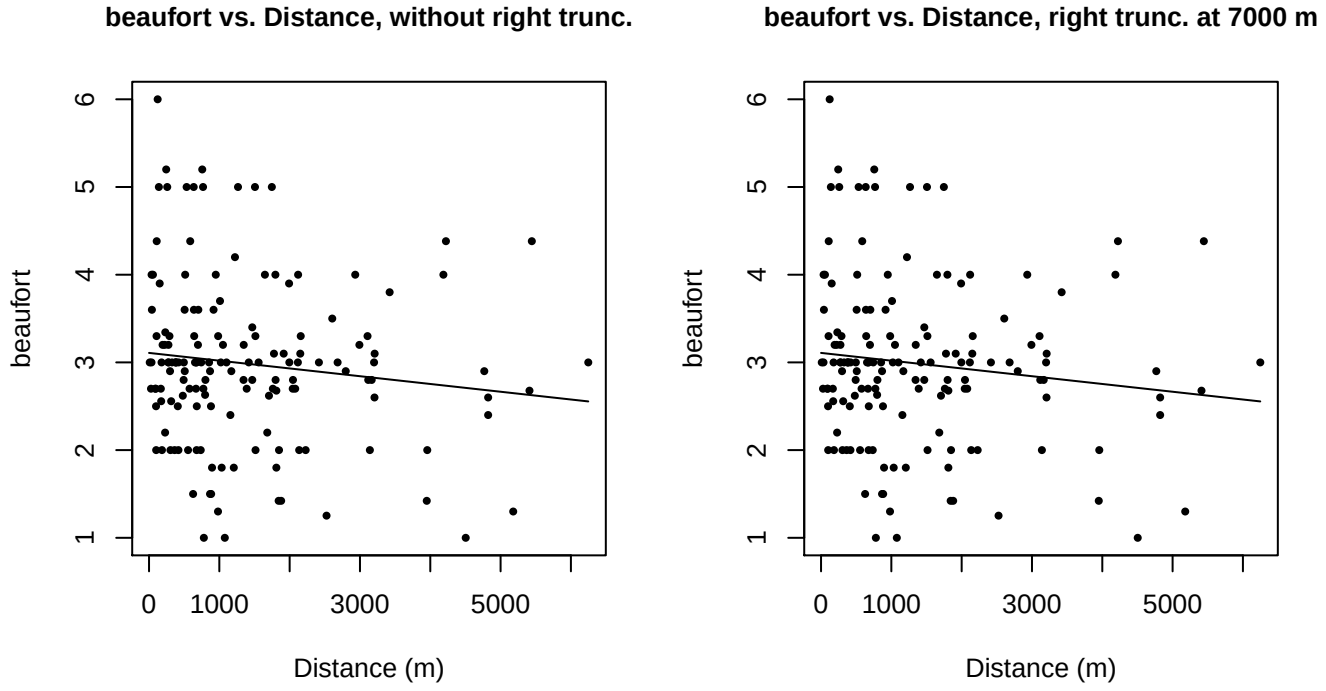


Figure 10: Scatterplots showing the relationship between Beaufort sea state and perpendicular sighting distance, for all sightings (left) and only those not right truncated (right). The line is a simple linear regression.

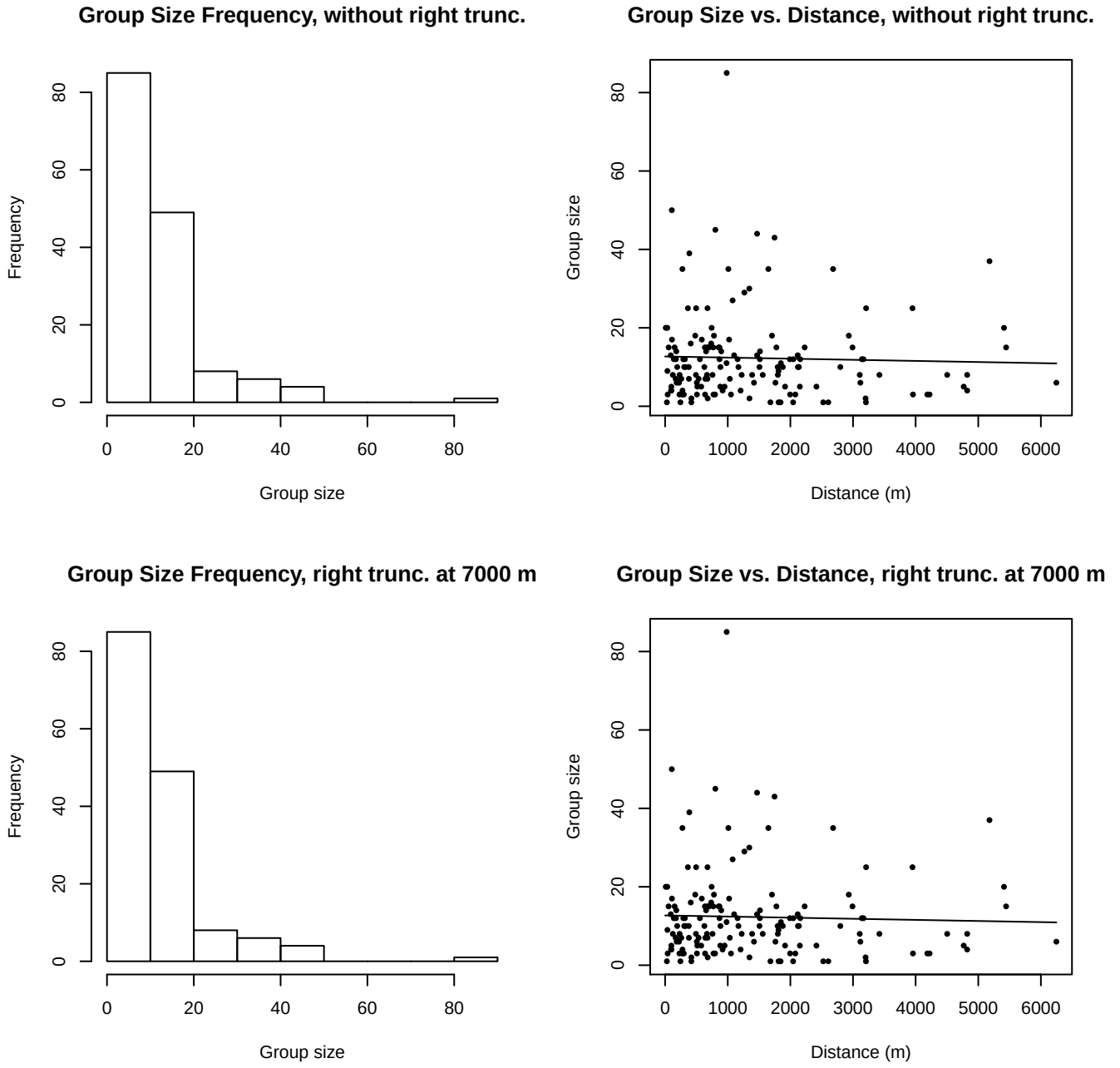


Figure 11: Histograms showing group size frequency and scatterplots showing the relationship between group size and perpendicular sighting distance, for all sightings (top row) and only those not right truncated (bottom row). In the scatterplot, the line is a simple linear regression.

High Platforms

The sightings were right truncated at 8000m.

Covariate	Description
beaufort	Beaufort sea state.
size	Estimated size (number of individuals) of the sighted group.

Table 6: Covariates tested in candidate "multi-covariate distance sampling" (MCDS) detection functions.

Key	Adjustment	Order	Covariates	Succeeded	Δ AIC	Mean ESHW (m)
hn			beaufort, size	Yes	0.00	2966
hn			beaufort	Yes	2.07	3008
hr			beaufort, size	Yes	9.84	3298
hr			beaufort	Yes	10.82	3450
hn			size	Yes	11.22	3031
hn	cos	3		Yes	16.30	2439
hr	poly	2		Yes	16.47	1910
hr			size	Yes	17.32	2828
hn	cos	2		Yes	17.90	2681
hr	poly	4		Yes	17.95	1994
hn				Yes	18.60	3042
hn	herm	4		Yes	20.54	3037
hr				Yes	21.05	2508

Table 7: Candidate detection functions for High Platforms. The first one listed was selected for the density model.

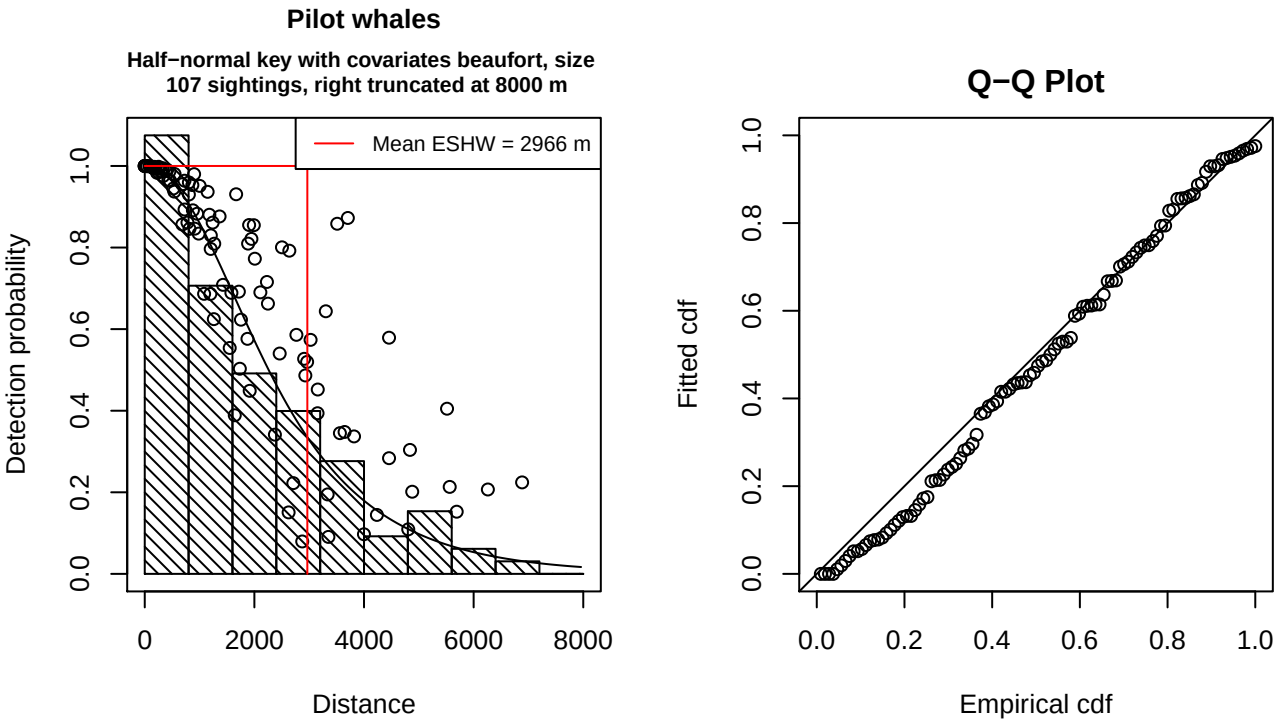


Figure 12: Detection function for High Platforms that was selected for the density model

Statistical output for this detection function:

```
Summary for ds object
```

```
Number of observations :  107
Distance range         :  0  -  8000
AIC                    :  1805.5

Detection function:
 Half-normal key function

Detection function parameters
Scale Coefficients:
              estimate         se
(Intercept)  8.0763226 0.27050263
beaufort    -0.2024908 0.06367844
size         0.1829904 0.10562413

                       Estimate          SE         CV
Average p             0.3286587  0.02660151 0.08093961
N in covered region 325.5657017 37.43191402 0.11497499
```

Additional diagnostic plots:

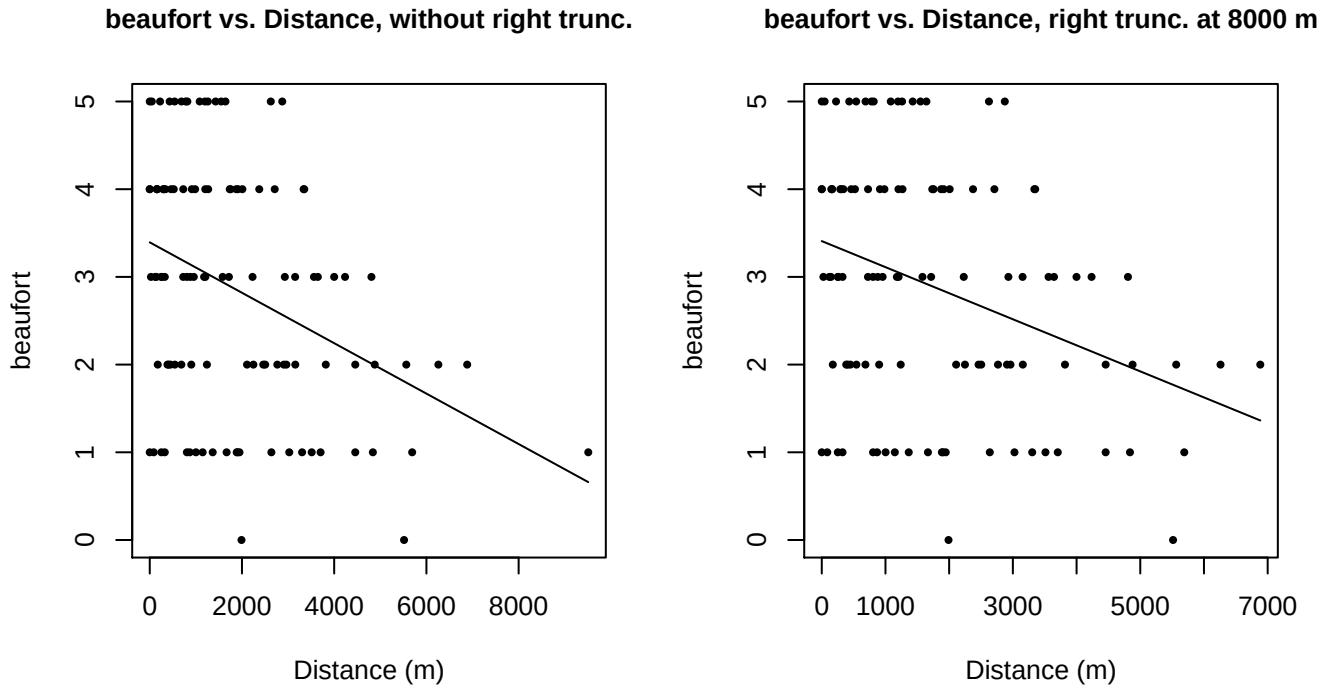


Figure 13: Scatterplots showing the relationship between Beaufort sea state and perpendicular sighting distance, for all sightings (left) and only those not right truncated (right). The line is a simple linear regression.

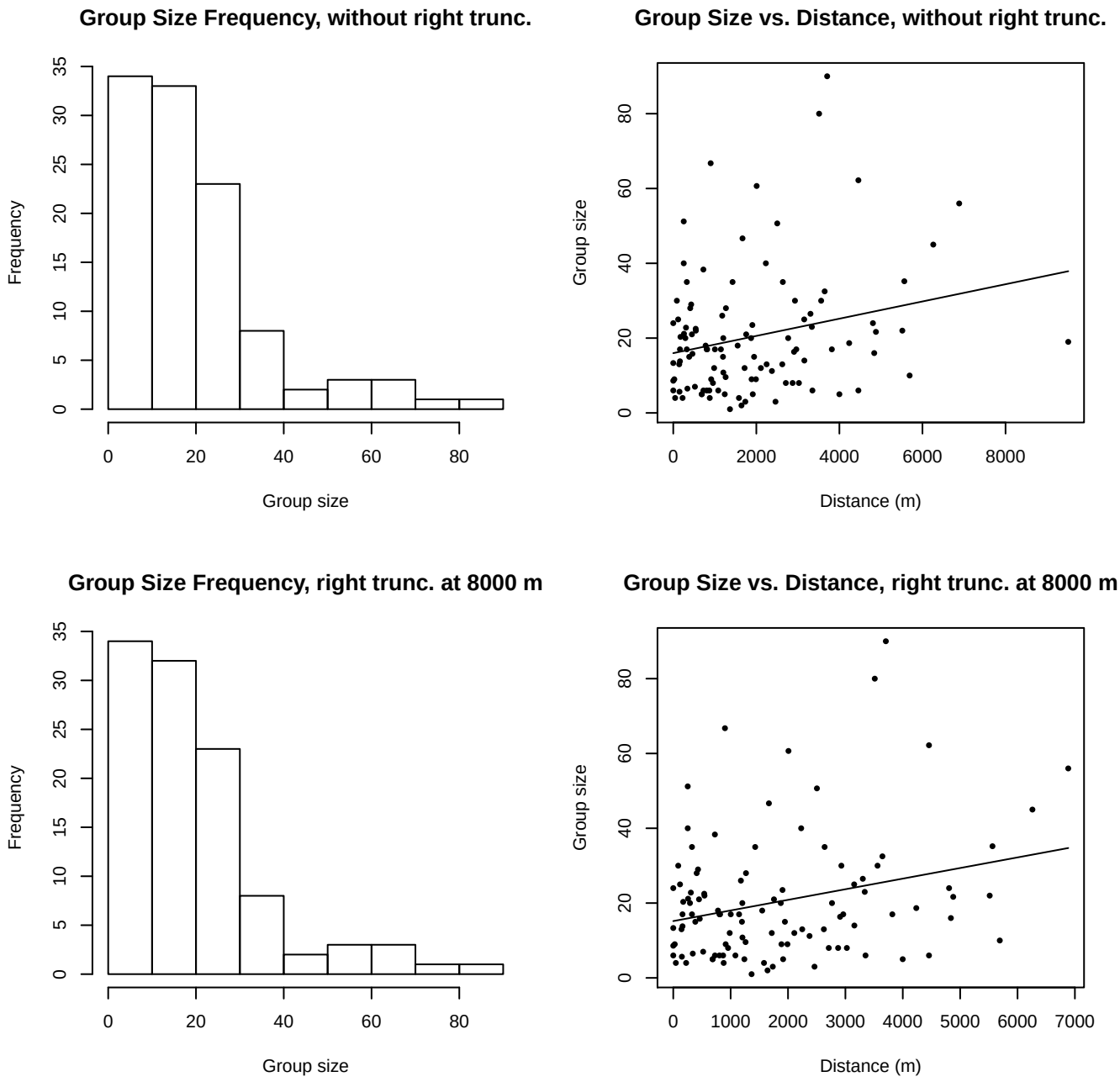


Figure 14: Histograms showing group size frequency and scatterplots showing the relationship between group size and perpendicular sighting distance, for all sightings (top row) and only those not right truncated (bottom row). In the scatterplot, the line is a simple linear regression.

Gordon Gunter Quality Covariate Available

The sightings were right truncated at 8000m.

Covariate	Description
beaufort	Beaufort sea state.
quality	Survey-specific index of the quality of observation conditions, utilizing relevant factors other than Beaufort sea state (see methods).
size	Estimated size (number of individuals) of the sighted group.

Table 8: Covariates tested in candidate "multi-covariate distance sampling" (MCDS) detection functions.

Key	Adjustment	Order	Covariates	Succeeded	Δ AIC	Mean ESHW (m)
hn			beaufort, quality	Yes	0.00	2858
hn			beaufort, size	Yes	0.44	2885
hn			beaufort, quality, size	Yes	0.56	2830
hn			beaufort	Yes	1.24	2911
hn			quality, size	Yes	3.88	2842
hn			quality	Yes	5.01	2873
hr			beaufort, quality, size	Yes	6.84	3482
hr			beaufort, quality	Yes	7.12	3556
hr			quality, size	Yes	9.41	3505
hr			beaufort, size	Yes	9.90	3269
hr			beaufort	Yes	9.97	3427
hn			size	Yes	11.48	2942
hr			quality	Yes	12.65	3382
hn	cos	3		Yes	14.41	2400
hr	poly	2		Yes	15.83	1867
hn				Yes	16.28	2955
hn	cos	2		Yes	16.42	2664
hr			size	Yes	17.39	2798
hr	poly	4		Yes	17.56	2062
hn	herm	4		Yes	18.23	2951
hr				Yes	20.48	2538

Table 9: Candidate detection functions for Gordon Gunter Quality Covariate Available. The first one listed was selected for the density model.

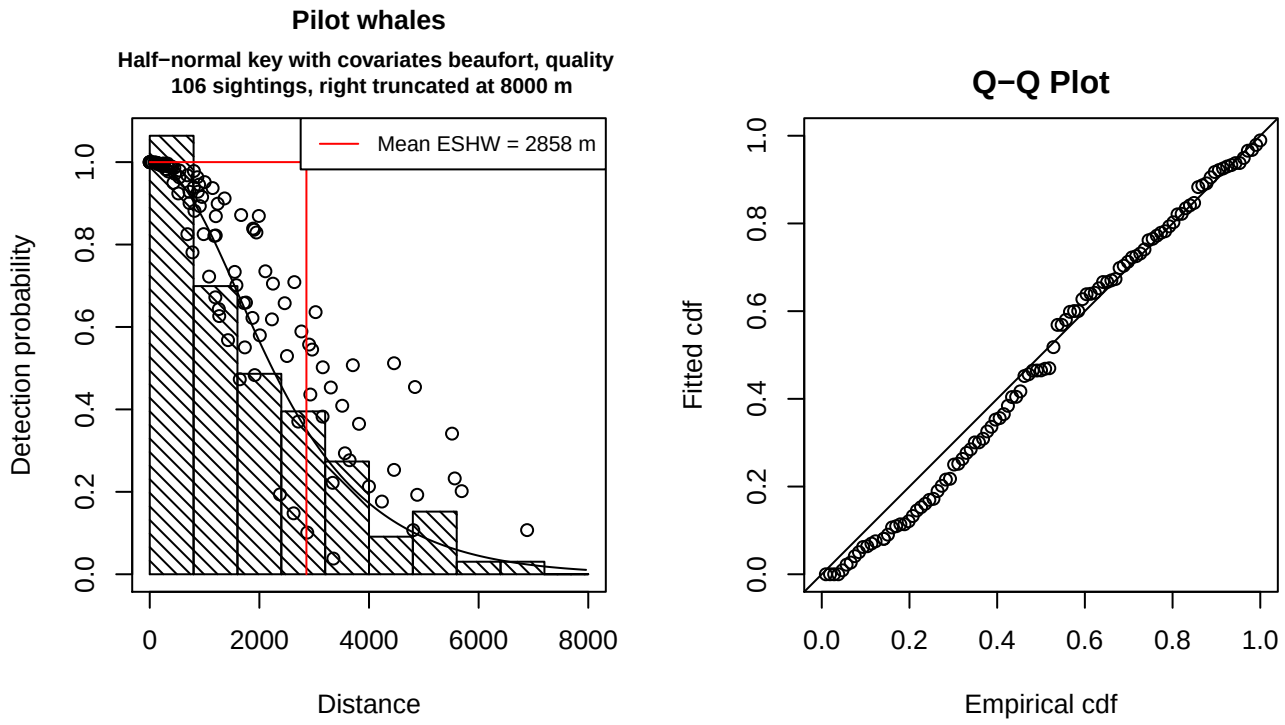


Figure 15: Detection function for Gordon Gunter Quality Covariate Available that was selected for the density model

Statistical output for this detection function:

```
Summary for ds object
Number of observations :  106
Distance range         :  0  -  8000
AIC                    :  1784.977

Detection function:
 Half-normal key function

Detection function parameters
Scale Coefficients:
             estimate         se
(Intercept)  8.4240449 0.21281649
beaufort    -0.1676876 0.07913117
quality     -0.1149824 0.06080005

                      Estimate          SE         CV
Average p             0.3222397  0.02598744 0.08064632
N in covered region 328.9476699 37.83382097 0.11501471
```

Additional diagnostic plots:

Figure 16: Scatterplots showing the relationship between Beaufort sea state and perpendicular sighting distance, for all sightings (left) and only those not right truncated (right). The line is a simple linear regression.

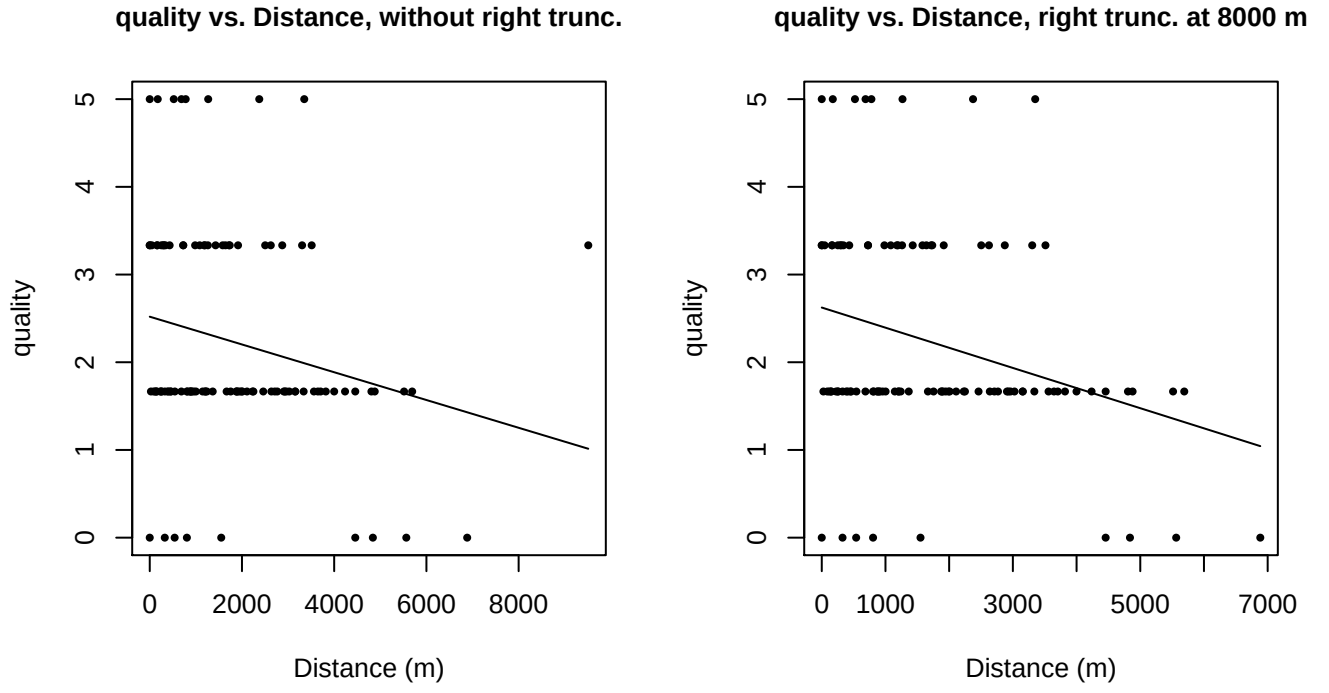


Figure 17: Scatterplots showing the relationship between the survey-specific index of the quality of observation conditions and perpendicular sighting distance, for all sightings (left) and only those not right truncated (right). Low values of the quality index correspond to better observation conditions. The line is a simple linear regression.

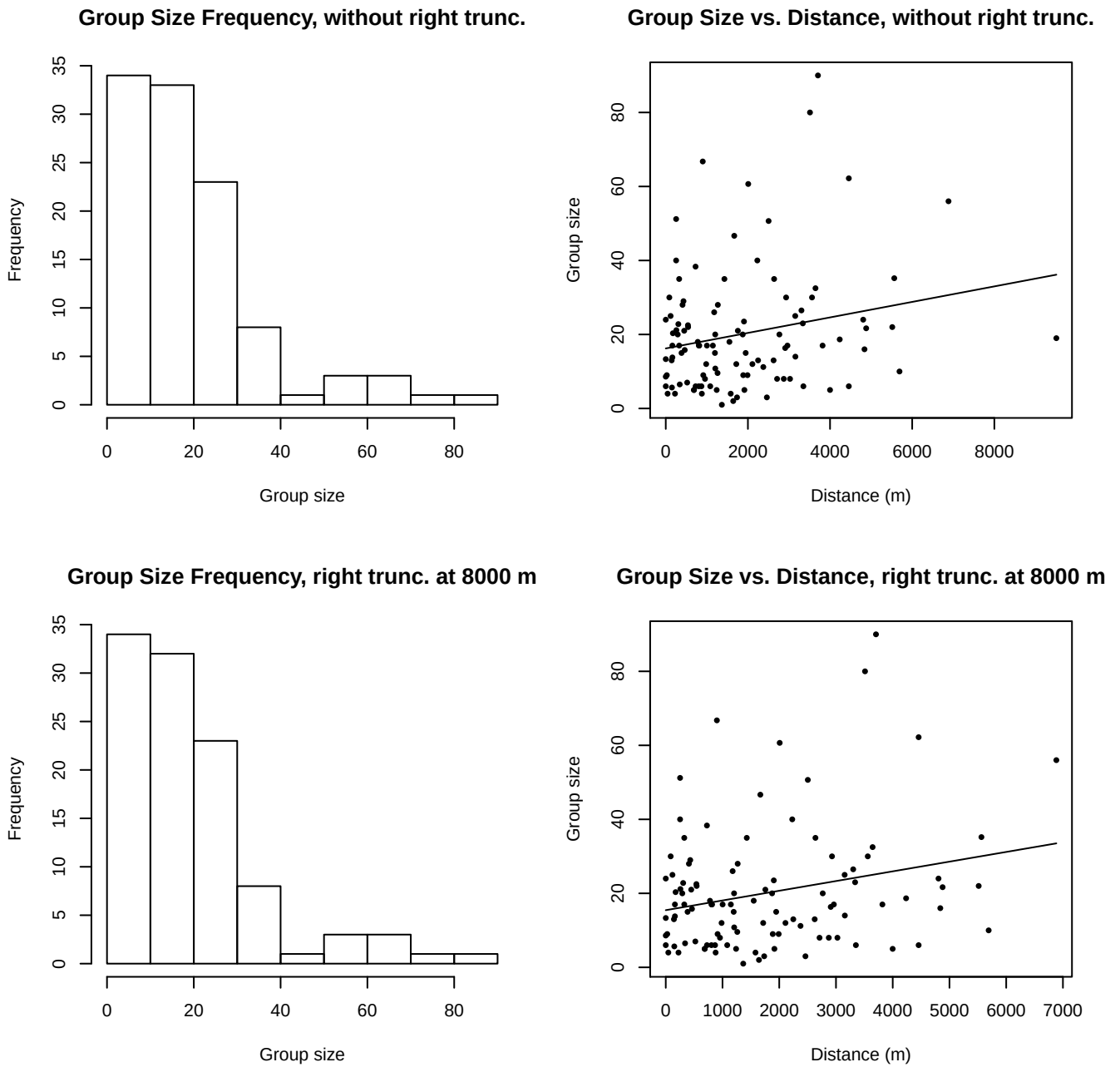


Figure 18: Histograms showing group size frequency and scatterplots showing the relationship between group size and perpendicular sighting distance, for all sightings (top row) and only those not right truncated (bottom row). In the scatterplot, the line is a simple linear regression.

Aerial Surveys

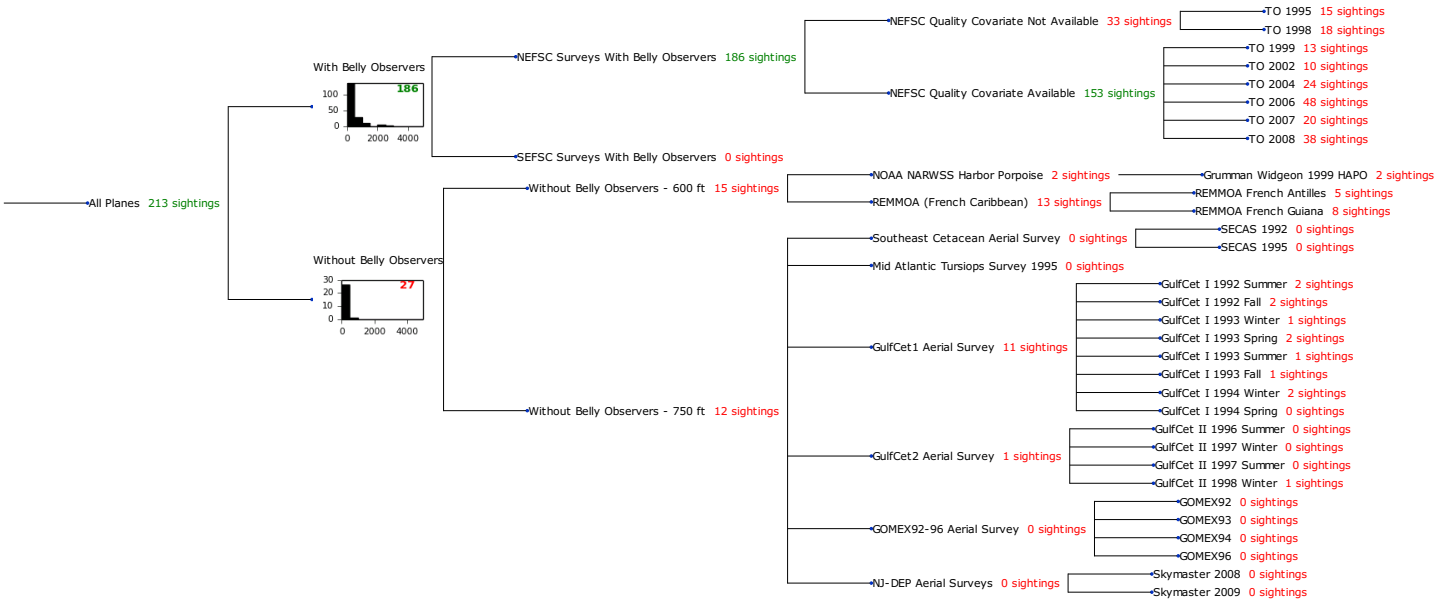


Figure 19: Detection hierarchy for aerial surveys

With Belly Observers

The sightings were right truncated at 1500m.

Covariate	Description
beaufort	Beaufort sea state.
size	Estimated size (number of individuals) of the sighted group.

Table 10: Covariates tested in candidate “multi-covariate distance sampling” (MCDS) detection functions.

Key	Adjustment	Order	Covariates	Succeeded	Δ AIC	Mean ESHW (m)
hr			size	Yes	0.00	394
hr			beaufort, size	Yes	1.78	398
hr				Yes	5.86	396
hr			beaufort	Yes	7.58	399
hr	poly	2		Yes	7.86	396
hr	poly	4		Yes	7.86	396
hn	cos	2		Yes	11.98	434
hn	cos	3		Yes	22.90	411
hn			size	Yes	24.93	596
hn				Yes	34.60	553
hn	herm	4		Yes	36.11	552
hn			beaufort	No		
hn			beaufort, size	No		

Table 11: Candidate detection functions for With Belly Observers. The first one listed was selected for the density model.

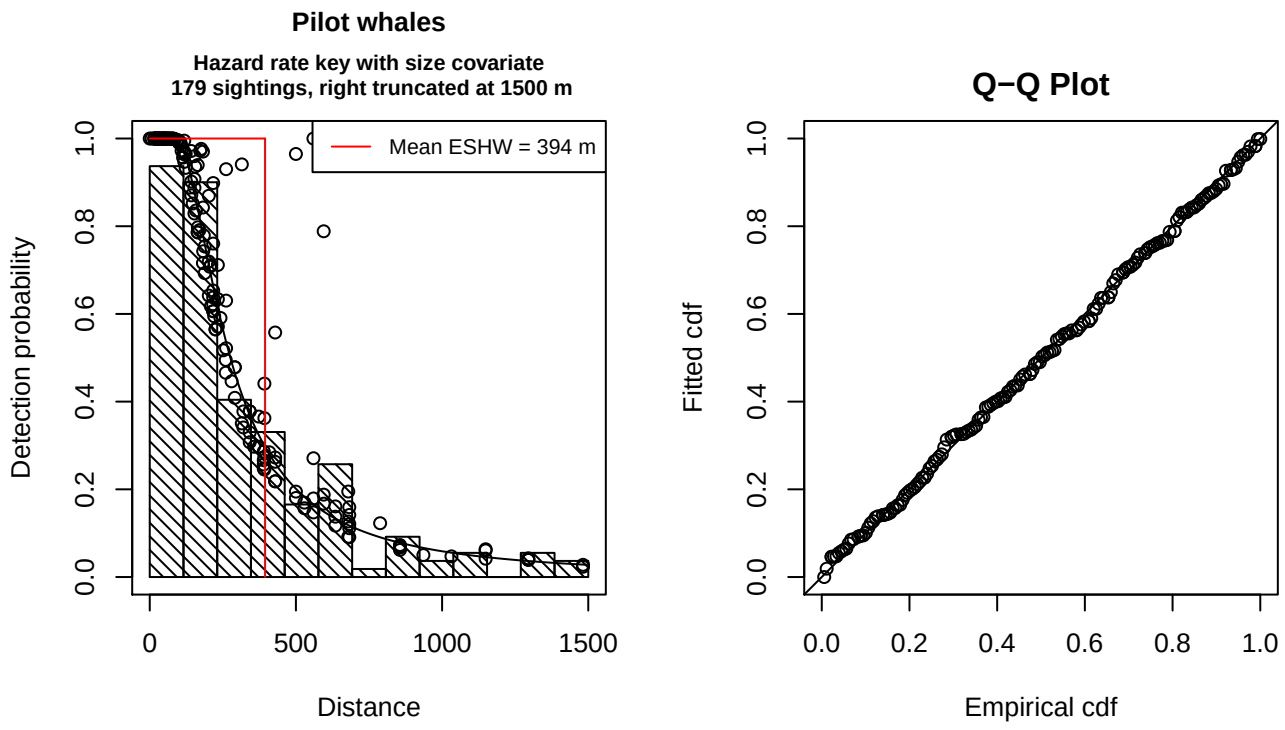


Figure 20: Detection function for With Belly Observers that was selected for the density model

Statistical output for this detection function:

```
Summary for ds object
Number of observations :  179
Distance range         :  0  -  1500
AIC                    :  2405.616

Detection function:
 Hazard-rate key function

Detection function parameters
Scale Coefficients:
             estimate        se
(Intercept) 5.2999331 0.1854678
size        0.2491643 0.1862108

Shape parameters:
             estimate        se
(Intercept) 0.6689847 0.1239738

                       Estimate         SE        CV
Average p              0.2529798  0.0258198 0.1020627
N in covered region 707.5664222 85.6297480 0.1210201
```

Additional diagnostic plots:

Figure 21: Scatterplots showing the relationship between Beaufort sea state and perpendicular sighting distance, for all sightings (left) and only those not right truncated (right). The line is a simple linear regression.

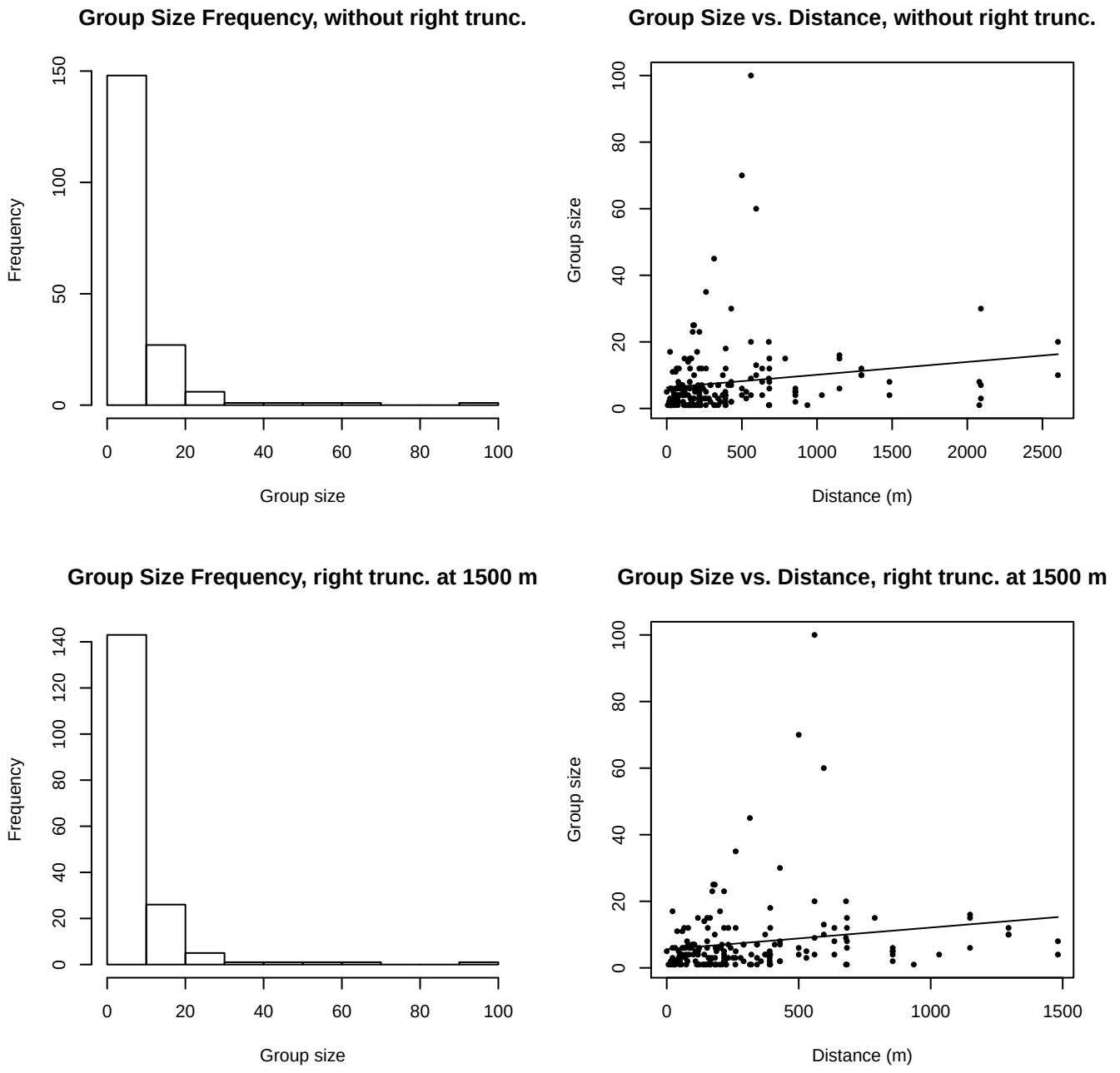


Figure 22: Histograms showing group size frequency and scatterplots showing the relationship between group size and perpendicular sighting distance, for all sightings (top row) and only those not right truncated (bottom row). In the scatterplot, the line is a simple linear regression.

Without Belly Observers

The sightings were right truncated at 800m. Due to a reduced frequency of sightings close to the trackline that plausibly resulted from the behavior of the observers and/or the configuration of the survey platform, the sightings were left truncted as well. Sightings closer than 83 m to the trackline were omitted from the analysis, and it was assumed that the the area closer to the trackline than this was not surveyed. This distance was estimated by inspecting histograms of perpendicular sighting distances.

Key	Adjustment	Order	Covariates	Succeeded	Δ AIC	Mean ESHW (m)

hn			Yes	0.00	235
hr			Yes	0.13	301
hn	cos	3	Yes	1.25	269
hn	cos	2	Yes	1.91	265
hn	herm	4	Yes	1.99	234
hr	poly	4	Yes	2.13	301
hr	poly	2	Yes	2.13	301

Table 12: Candidate detection functions for Without Belly Observers. The first one listed was selected for the density model.

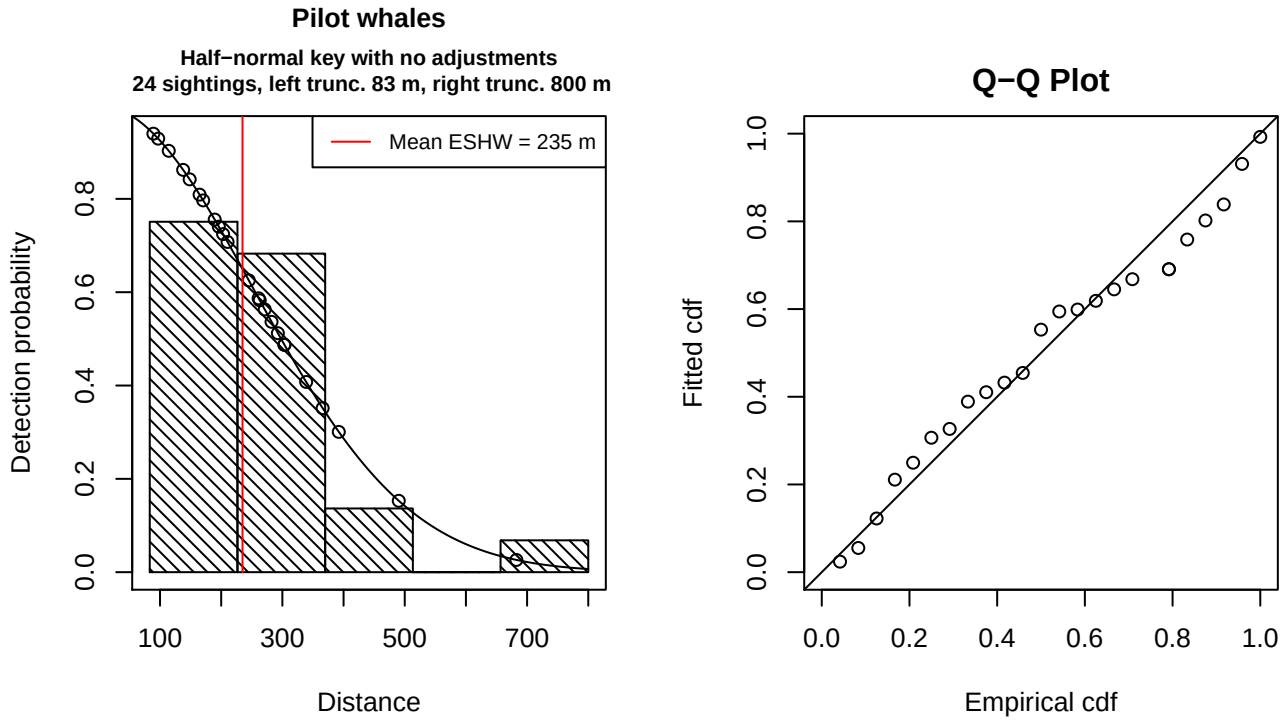


Figure 23: Detection function for Without Belly Observers that was selected for the density model

Statistical output for this detection function:

```
Summary for ds object
Number of observations :  24
Distance range         :  83.2036  -  800
AIC                    :  295.5364

Detection function:
 Half-normal key function

Detection function parameters
Scale Coefficients:
            estimate        se
```

```
(Intercept) 5.533402 0.1369899

                       Estimate          SE        CV
Average p             0.2935847  0.05279273 0.1798211
N in covered region 81.7481219 20.31726236 0.2485349
```

Additional diagnostic plots:

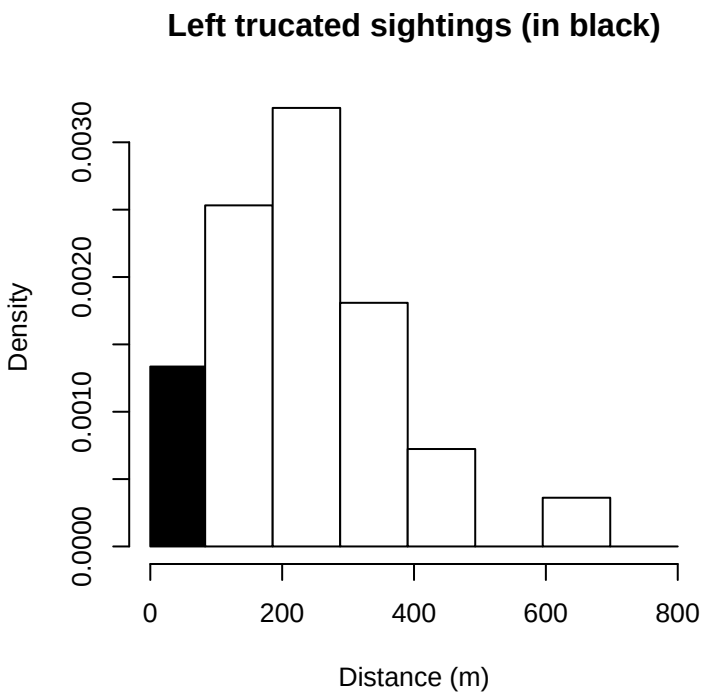


Figure 24: Density of sightings by perpendicular distance for Without Belly Observers. Black bars on the left show sightings that were left truncated.

g(0) Estimates

Platform	Surveys	Group Size	*g(0)*	Biases Addressed	Source
Shipboard	All	Any	0.585	Perception	Palka (2006)
Aerial	All	Any	0.607	Availability	Various

Table 13: Estimates of *g(0)* used in this density model.

Palka (2006) provided survey-specific g(0) estimates for two NOAA NEFSC shipboard surveys that used bigeye binoculars: the 1998 Abel-J survey (g(0)=0.50) and the 2004 Endeavor survey (g(0)=0.67). We used the estimates for the lower team, which was the primary team and the one for which we had sightings. These surveys occurred in the northwest Atlantic. All other binocular surveys, including all of those that occurred in the Gulf of Mexico, did not estimate g(0); for these we used the simple mean (g(0)=0.585) of Palka's two estimates. This estimate accounted for perception bias but not availability bias (Palka 2005b), so our estimate is likely to be biased low (but see discussion of pilot whale availability below).

We also considered Barlow (2006), which estimated g(0)=0.76 for groups of 1-20 animals and g(0)=1.00 for groups larger than 20 animals. Although Palka's estimate did not consider group size, and pilot whales can occur in large groups, we preferred Palka's estimate because it incorporated more sightings and it is specific to pilot whales (Barlow's estimate pooled pilot whales with many delphinds and estimated g(0) for the group). Like Palka's, Barlow's estimate only accounted for perception bias.

We did not find in the literature a g(0) estimate for pilot whales observed by aircraft. Pilot whales exhibit variable diving behavior that often includes long periods near the surface punctuated by occasional deep dives. Rather than base g(0) on availability bias estimated by from surface and dive intervals (following Carretta, 2000) as we did with other whales, we based it on percent- time-at-surface data reported from monitoring studies. Six short-finned pilot whales tracked with time-depth recorders near Madeira Island averaged 76.3% of the monitored period between 0-10m depth (Alves et al. 2013). Two pilot whales tracked with satellite tags near Florida spent 31% of the monitored period between 0-2m depth (Wells et al. 2013). One pilot whale tracked with a DTAG (species and location not given) spent 67% of the monitored period between 0-10m depth (Hooker et al. 2012). 14 long-finned pilot whales tracked with DTAGs in the Alboran Sea spent 57.49% of the monitored period at the surface (Canadas, 2011). Three long-finned pilot whales tracked with time-depth recorders near the Faroe Islands averaged 60.0% of the monitored period between 0-7m depth (Heide- Jorgensen et al. 2002). One large group of short-finned pilot whales tracked visually near the Gulf of California spent 66.6% of the monitored period visible at the surface (Barlow et al. 1997). We used the mean percent-time-at-surface (60.7%) for these 27 groups as the availability bias component of g(0). We did not incorporate an estimate of perception bias, thus our g(0) estimate is likely to be biased high.

Density Models

Short-finned pilot whales occur worldwide in tropical, sub-tropical, and warm temperate waters (Olson 2008). In the northern Gulf of Mexico, most sightings have been reported west of 89 W, primarily along the continental slope (Waring et al. 2013). All of the sightings in our database occurred off the continental shelf; many occurred along the slope but several occurred in the deep waters of the central Gulf. Our literature review did not yield any descriptions of seasonal movements for short-finned pilot whales in the Gulf of Mexico. Given this off-shelf distribution with no knowledge of seasonal patterns, we fitted a year-round model to off-shelf waters, defined here as those deeper than the 100m isobath.

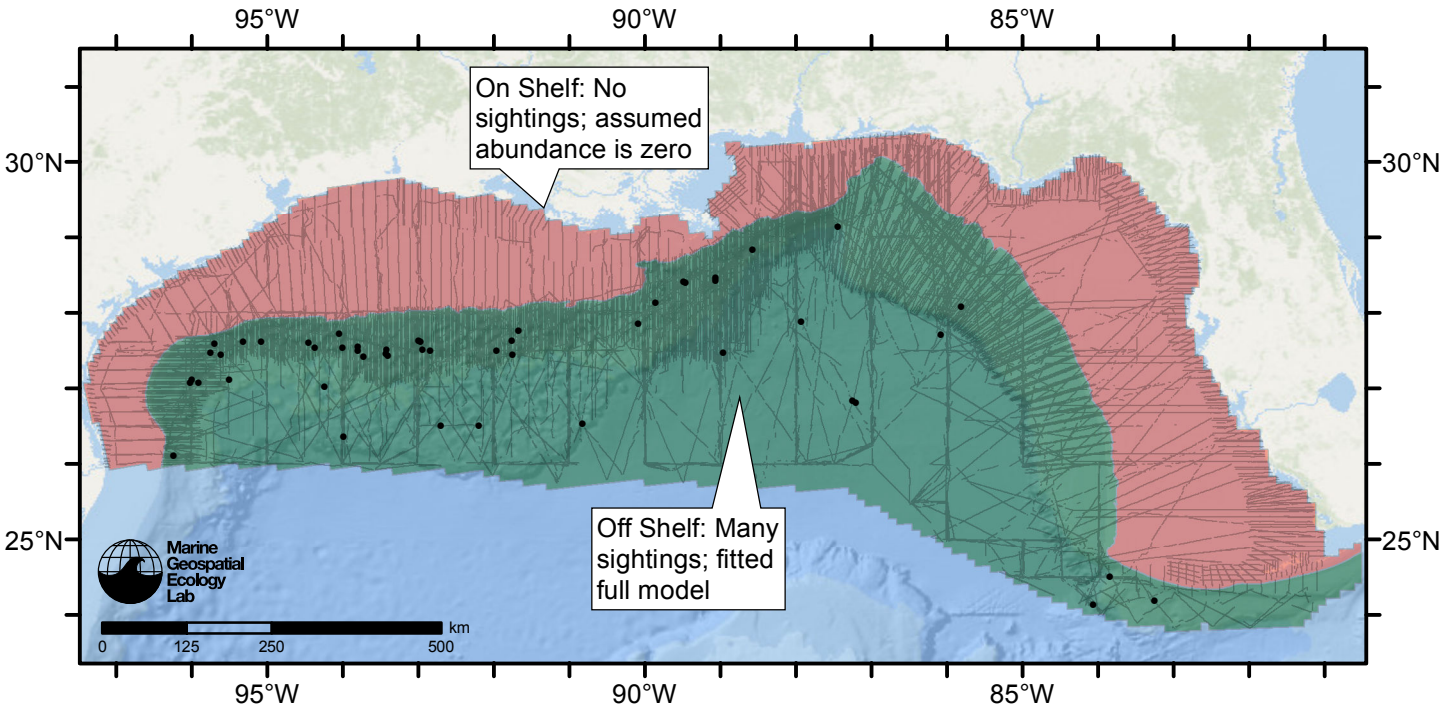


Figure 25: Pilot whales density model schematic. All on-effort sightings are shown, including those that were truncated when detection functions were fitted.

Climatological Model

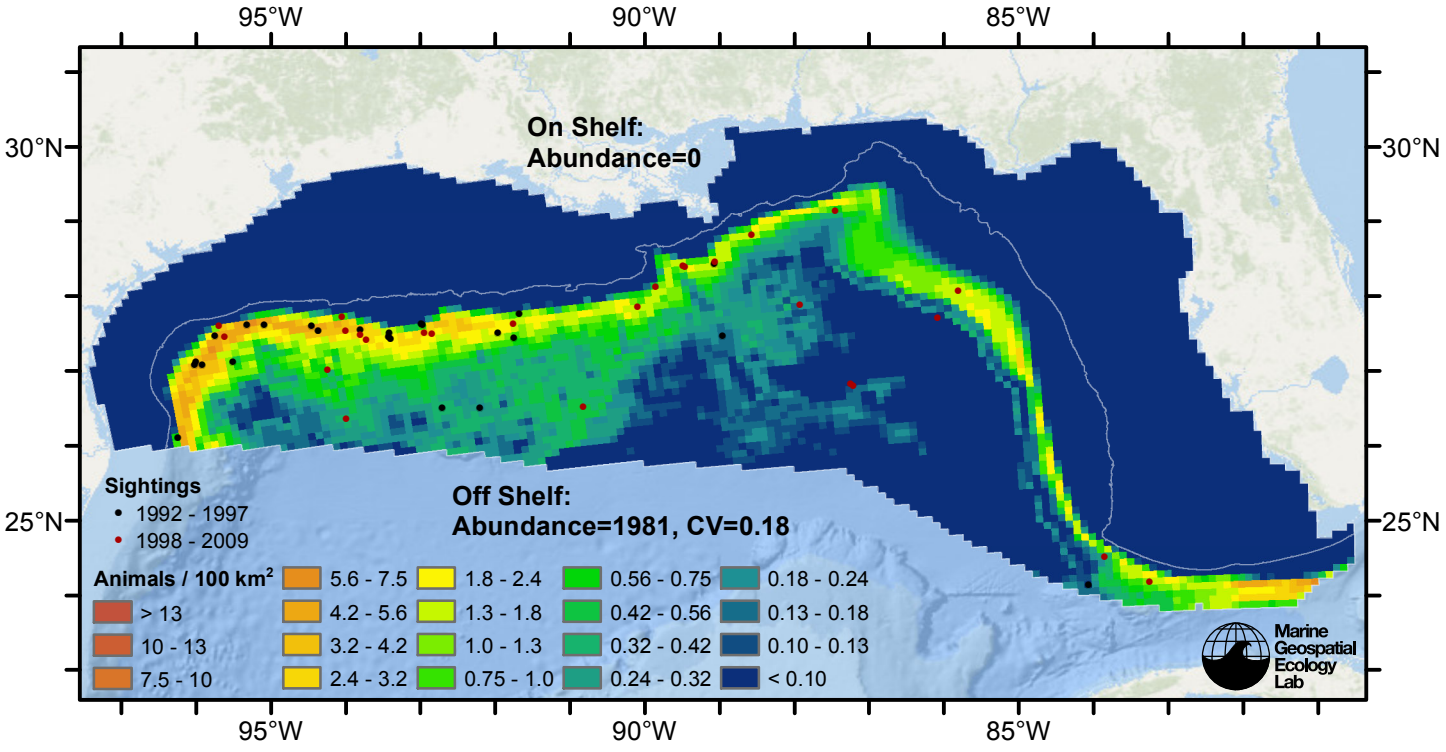


Figure 26: Pilot whales density predicted by the climatological model that explained the most deviance. Pixels are 10x10 km. The legend gives the estimated individuals per pixel; breaks are logarithmic. Abundance for each region was computed by summing the density cells occuring in that region.

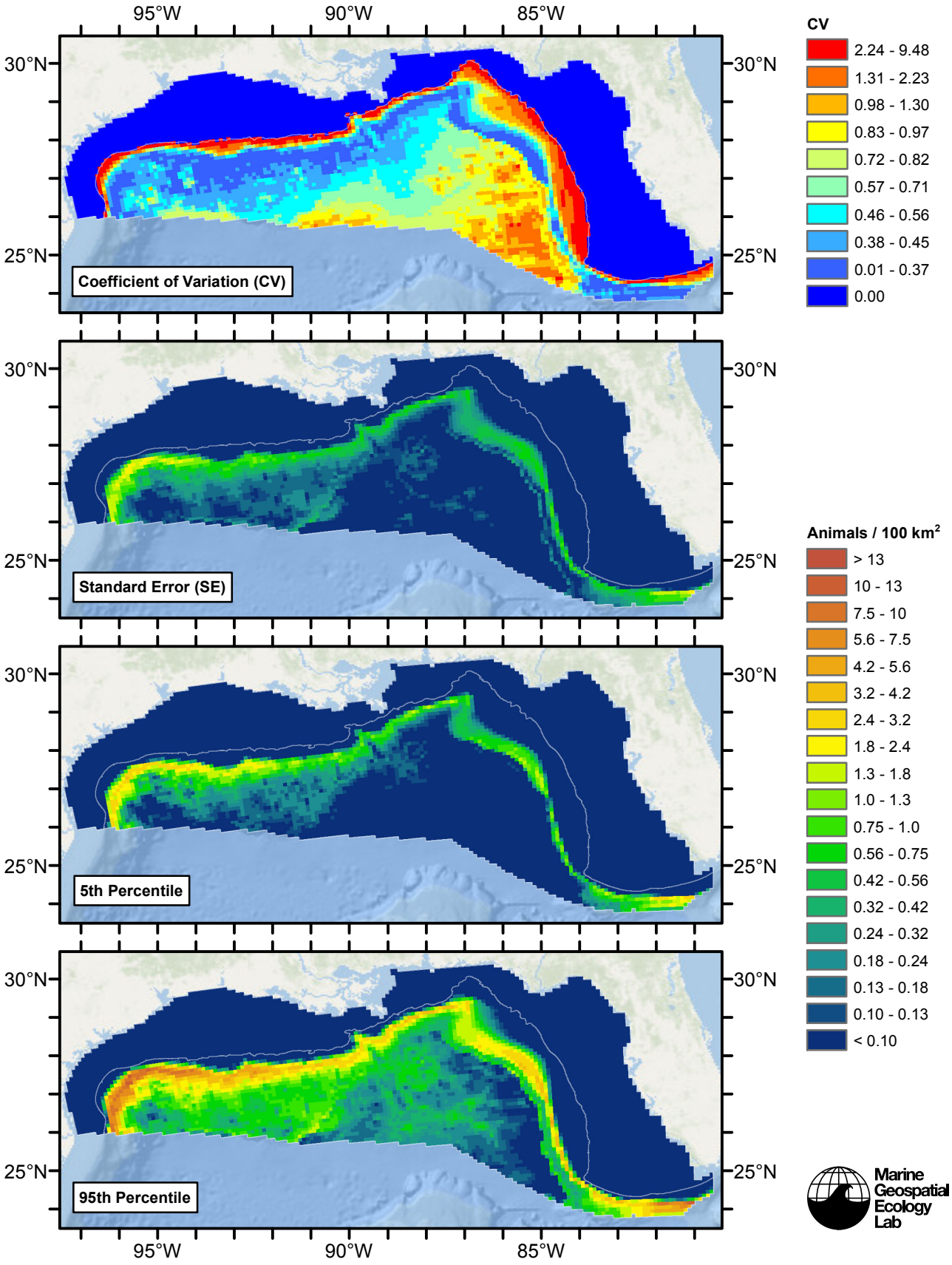


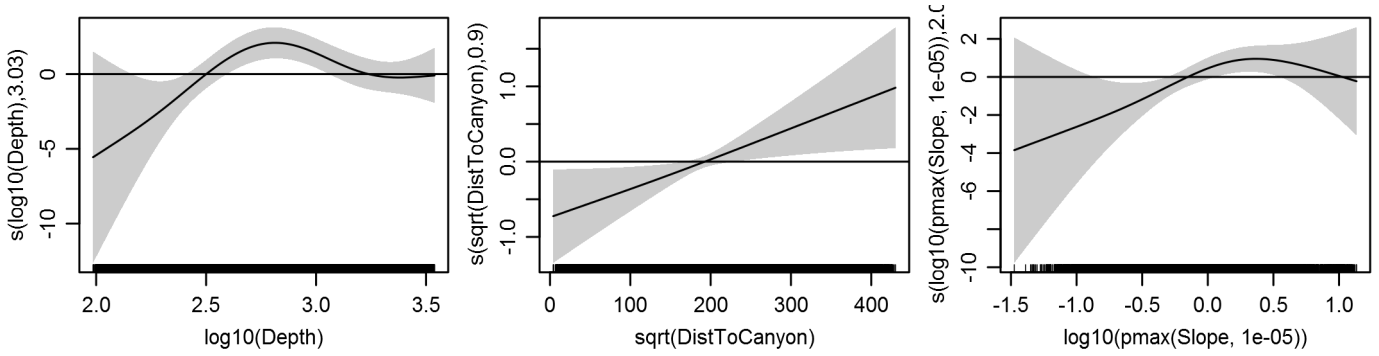
Figure 27: Estimated uncertainty for the climatological model that explained the most deviance. These estimates only incorporate the statistical uncertainty estimated for the spatial model (by the R mgcv package). They do not incorporate uncertainty in the detection functions, g(0) estimates, predictor variables, and so on.

Off Shelf

Statistical output

```
Rscript.exe: This is mgcv 1.8-3. For overview type 'help("mgcv-package")'.


Family: Tweedie(p=1.31)
Link function: log

Formula:
abundance ~ offset(log(area_km2)) + s(log10(Depth), bs = "ts",
    k = 5) + s(log10(pmax(Slope, 1e-05)), bs = "ts", k = 5) +
    s(sqrt(DistToCanyon), bs = "ts", k = 5)

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -6.4625     0.4674  -13.83   <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(Depth))              3.0290      4 5.971 7.11e-06 ***
s(log10(pmax(Slope, 1e-05))) 2.0499      4 2.091  0.00867 **
s(sqrt(DistToCanyon))        0.9034      4 1.549  0.00755 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) =  0.00116   Deviance explained = 18.2%
-REML = 576.11  Scale est. = 208.08    n = 14455

All predictors were significant. This is the final model.
Creating term plots.
Diagnostic output from gam.check():

Method: REML   Optimizer: outer newton
full convergence after 11 iterations.
Gradient range [-6.049325e-06,4.133098e-07]
(score 576.1082 & scale 208.0793).
Hessian positive definite, eigenvalue range [0.237974,182.0222].
Model rank =  13 / 13

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(Depth))              4.000 3.029   0.775    0.00
s(log10(pmax(Slope, 1e-05))) 4.000 2.050   0.818    0.04
s(sqrt(DistToCanyon))        4.000 0.903   0.824    0.12

Predictors retained during the model selection procedure: Depth, Slope, DistToCanyon

Predictors dropped during the model selection procedure:
```

Model term plots



Diagnostic plots

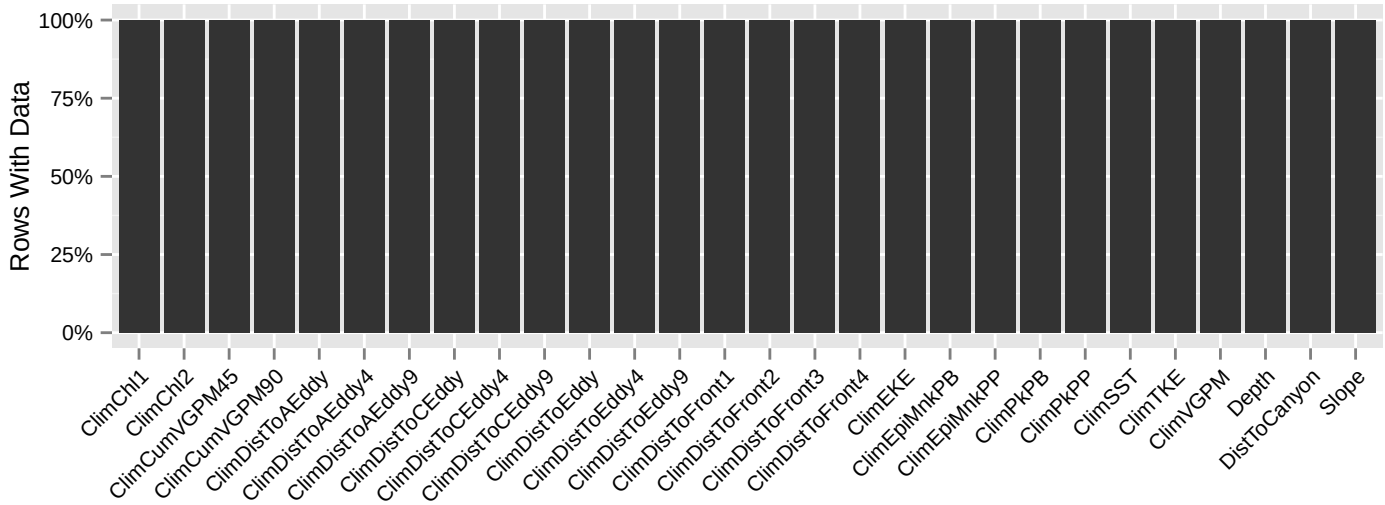


Figure 28: Segments with predictor values for the Pilot whales Climatological model, Off Shelf. This plot is used to assess how many segments would be lost by including a given predictor in a model.

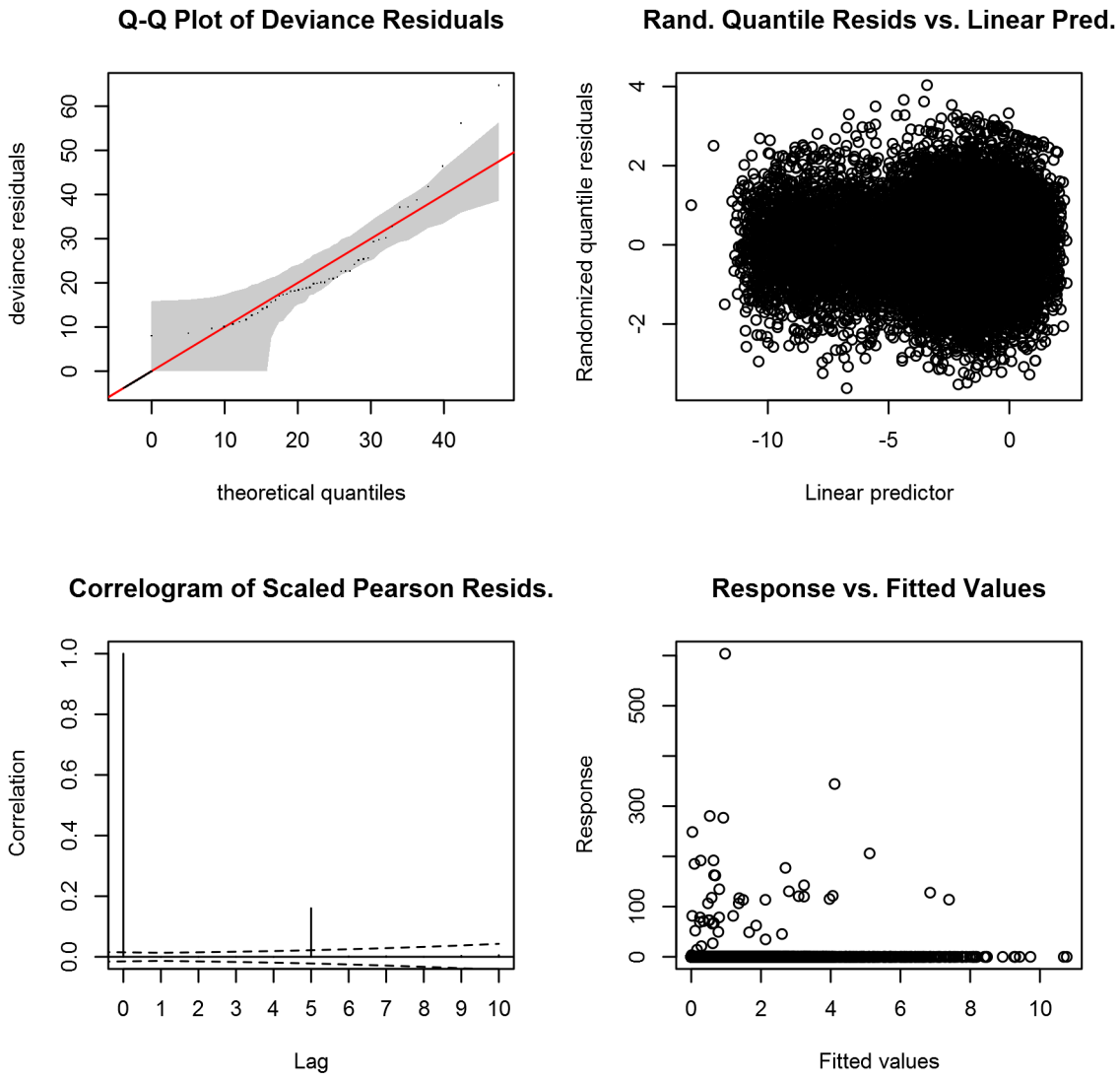


Figure 29: Statistical diagnostic plots for the Pilot whales Climatological model, Off Shelf.

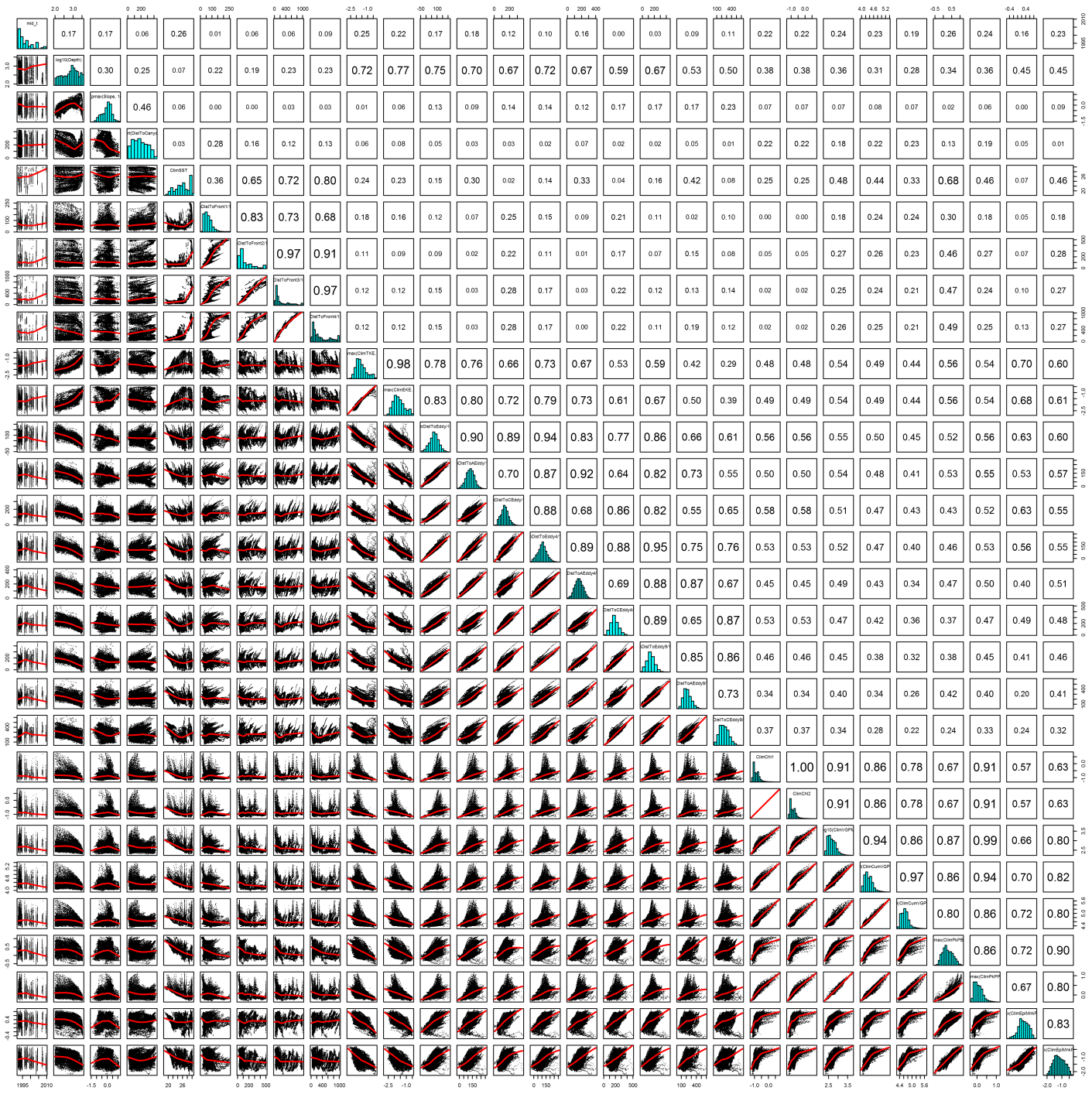


Figure 30: Scatterplot matrix for the Pilot whales Climatological model, Off Shelf. This plot is used to inspect the distribution of predictors (via histograms along the diagonal), simple correlation between predictors (via pairwise Pearson coefficients above the diagonal), and linearity of predictor correlations (via scatterplots below the diagonal). This plot is best viewed at high magnification.

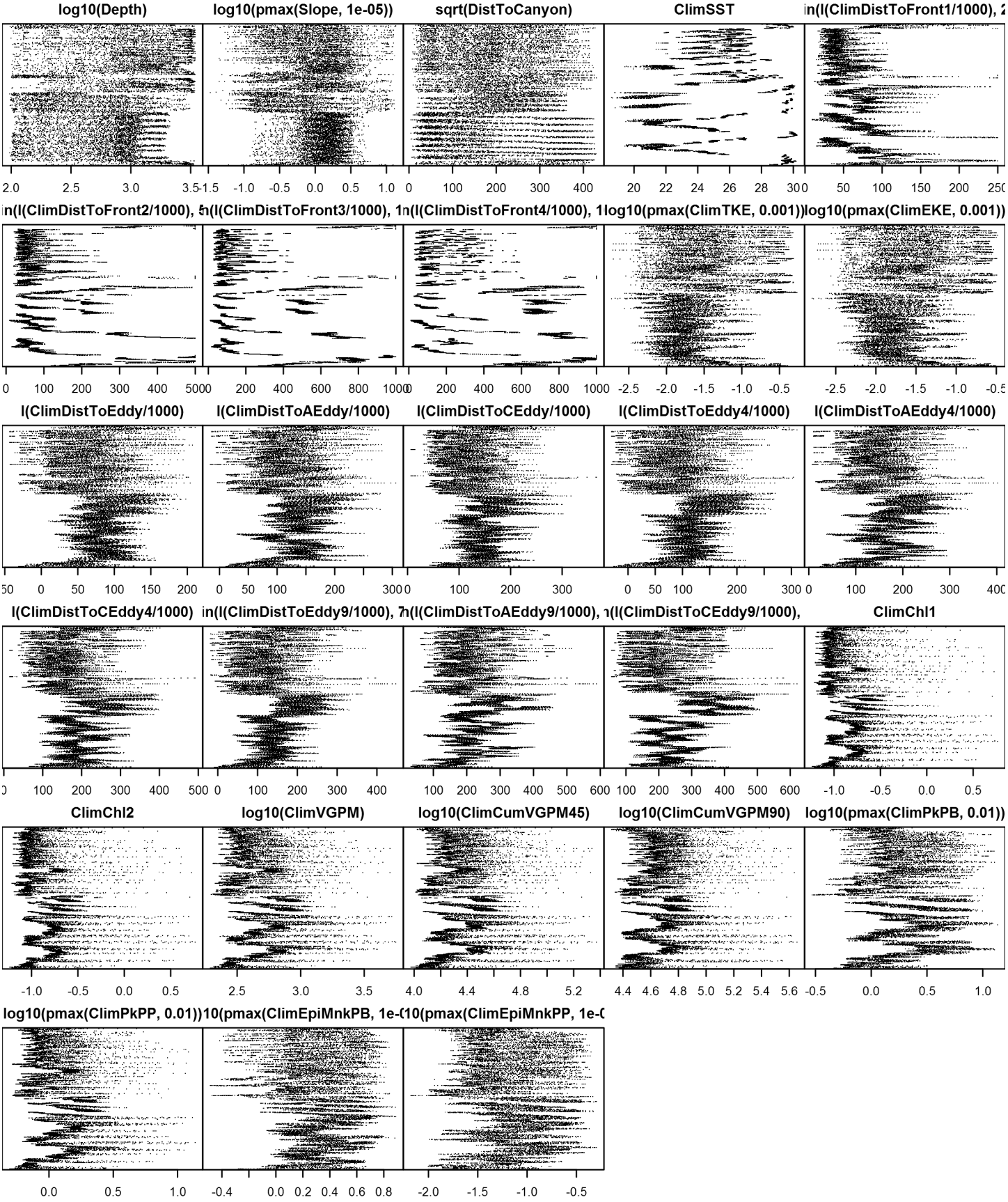


Figure 31: Dotplot for the Pilot whales Climatological model, Off Shelf. This plot is used to check for suspicious patterns and outliers in the data. Points are ordered vertically by transect ID, sequentially in time.

On Shelf

Density assumed to be 0 in this region.

Contemporaneous Model

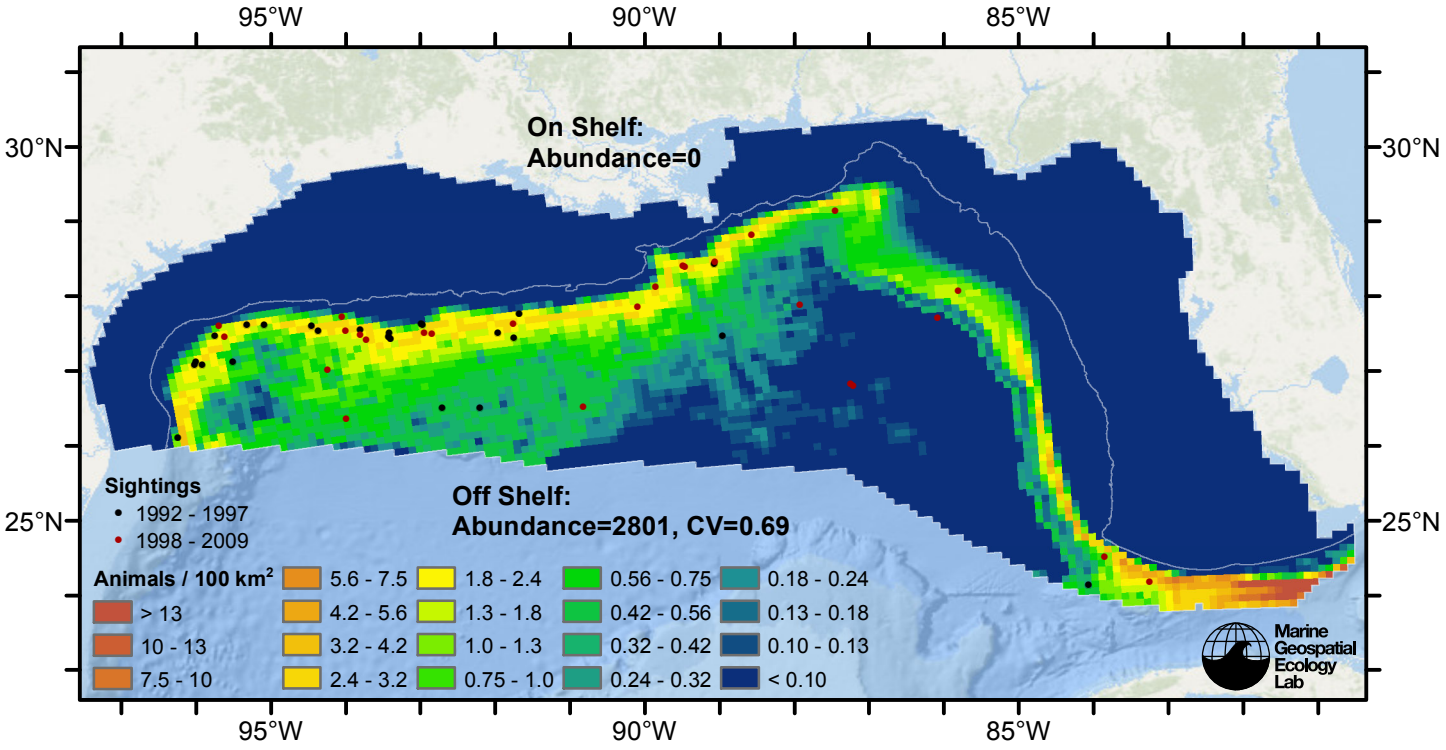


Figure 32: Pilot whales density predicted by the contemporaneous model that explained the most deviance. Pixels are 10x10 km. The legend gives the estimated individuals per pixel; breaks are logarithmic. Abundance for each region was computed by summing the density cells occuring in that region.

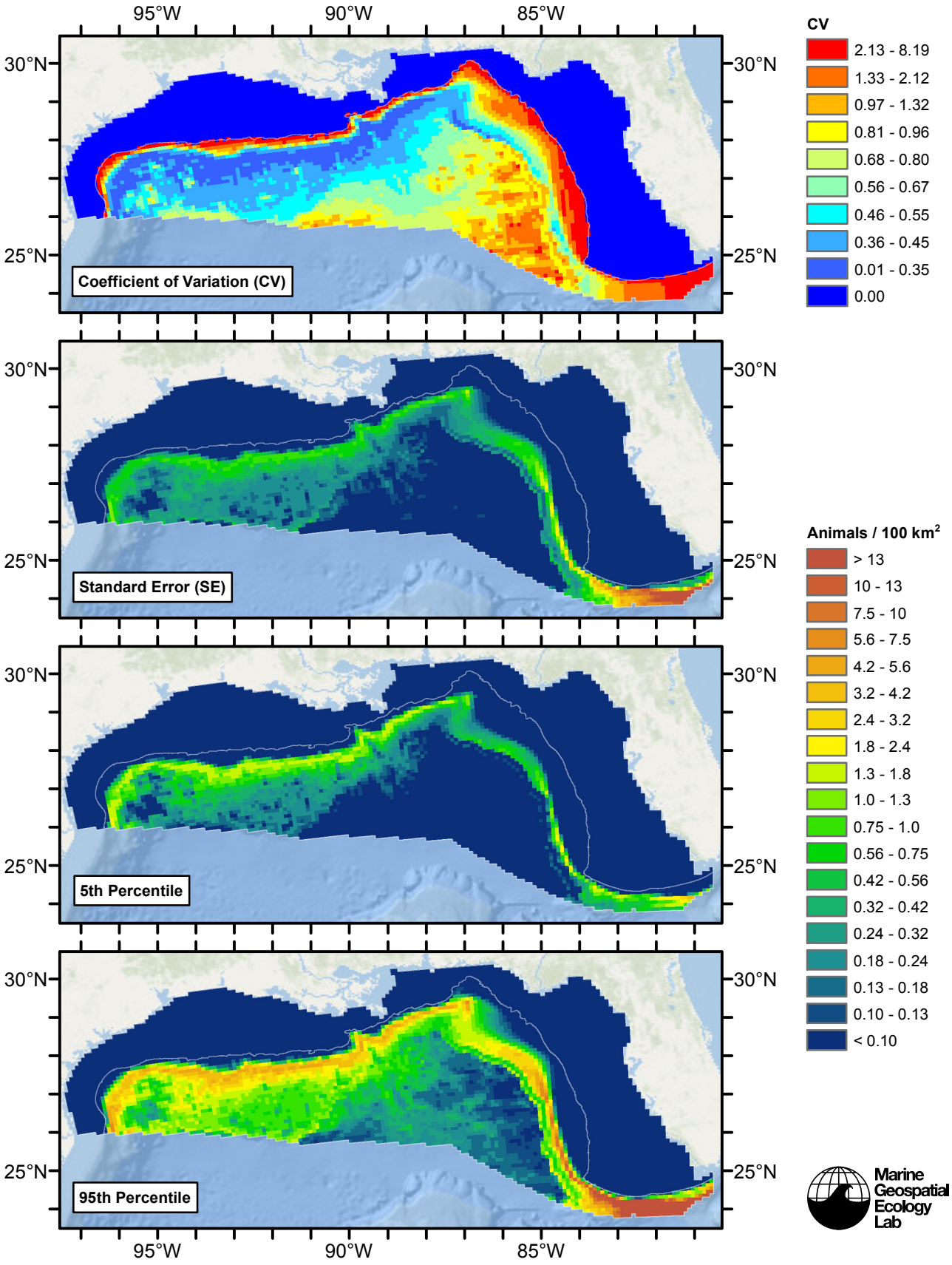


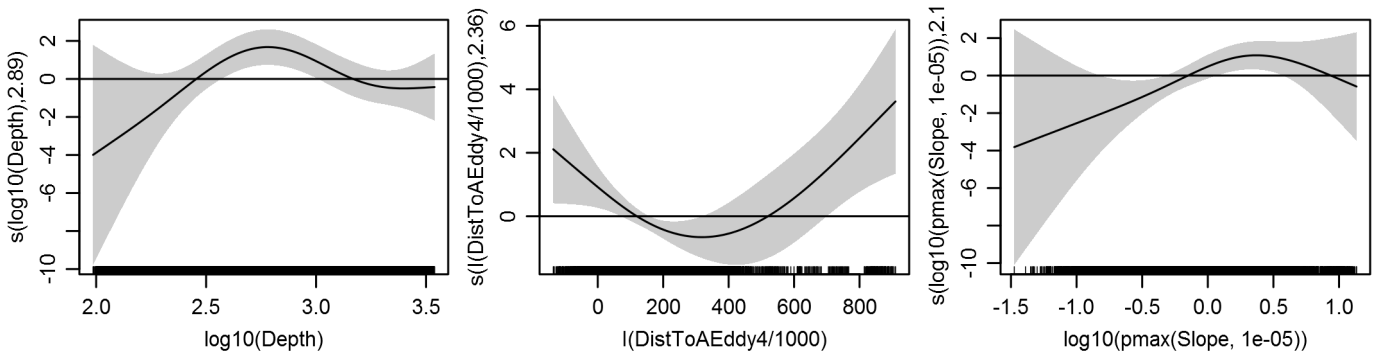
Figure 33: Estimated uncertainty for the contemporaneous model that explained the most deviance. These estimates only incorporate the statistical uncertainty estimated for the spatial model (by the R mgcv package). They do not incorporate uncertainty in the detection functions, g(0) estimates, predictor variables, and so on.

Off Shelf

Statistical output

```
Rscript.exe: This is mgcv 1.8-3. For overview type 'help("mgcv-package")'.


Family: Tweedie(p=1.312)
Link function: log

Formula:
abundance ~ offset(log(area_km2)) + s(log10(Depth), bs = "ts",
    k = 5) + s(log10(pmax(Slope, 1e-05)), bs = "ts", k = 5) +
    s(I(DistToAEddy4/1000), bs = "ts", k = 5)

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -6.3639     0.4186   -15.2   <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(Depth))              2.889      4 4.254 0.000232 ***
s(log10(pmax(Slope, 1e-05))) 2.191      4 2.237 0.007439 **
s(I(DistToAEddy4/1000))      2.359      4 4.559 7.16e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) =  0.0066   Deviance explained =   21%
-REML = 506.51  Scale est. = 204.72    n = 12354

All predictors were significant. This is the final model.
Creating term plots.
Diagnostic output from gam.check():

Method: REML   Optimizer: outer newton
full convergence after 11 iterations.
Gradient range [-2.310417e-05,5.80502e-07]
(score 506.5104 & scale 204.7213).
Hessian positive definite, eigenvalue range [0.2752962,158.8749].
Model rank =  13 / 13

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(Depth))              4.000 2.889   0.744    0.02
s(log10(pmax(Slope, 1e-05))) 4.000 2.191   0.743    0.03
s(I(DistToAEddy4/1000))      4.000 2.359   0.764    0.07

Predictors retained during the model selection procedure: Depth, Slope, DistToAEddy4

Predictors dropped during the model selection procedure: DistToCanyon, SST, DistToFront1, TKE,
DistToCEddy4
```

Model term plots



Diagnostic plots

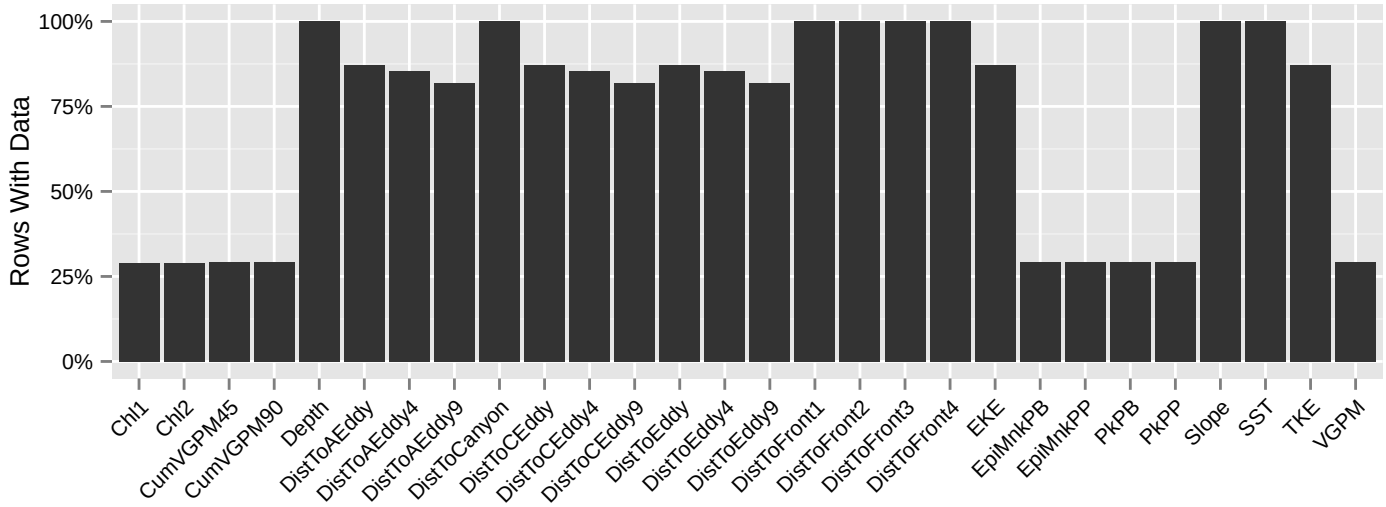


Figure 34: Segments with predictor values for the Pilot whales Contemporaneous model, Off Shelf. This plot is used to assess how many segments would be lost by including a given predictor in a model.

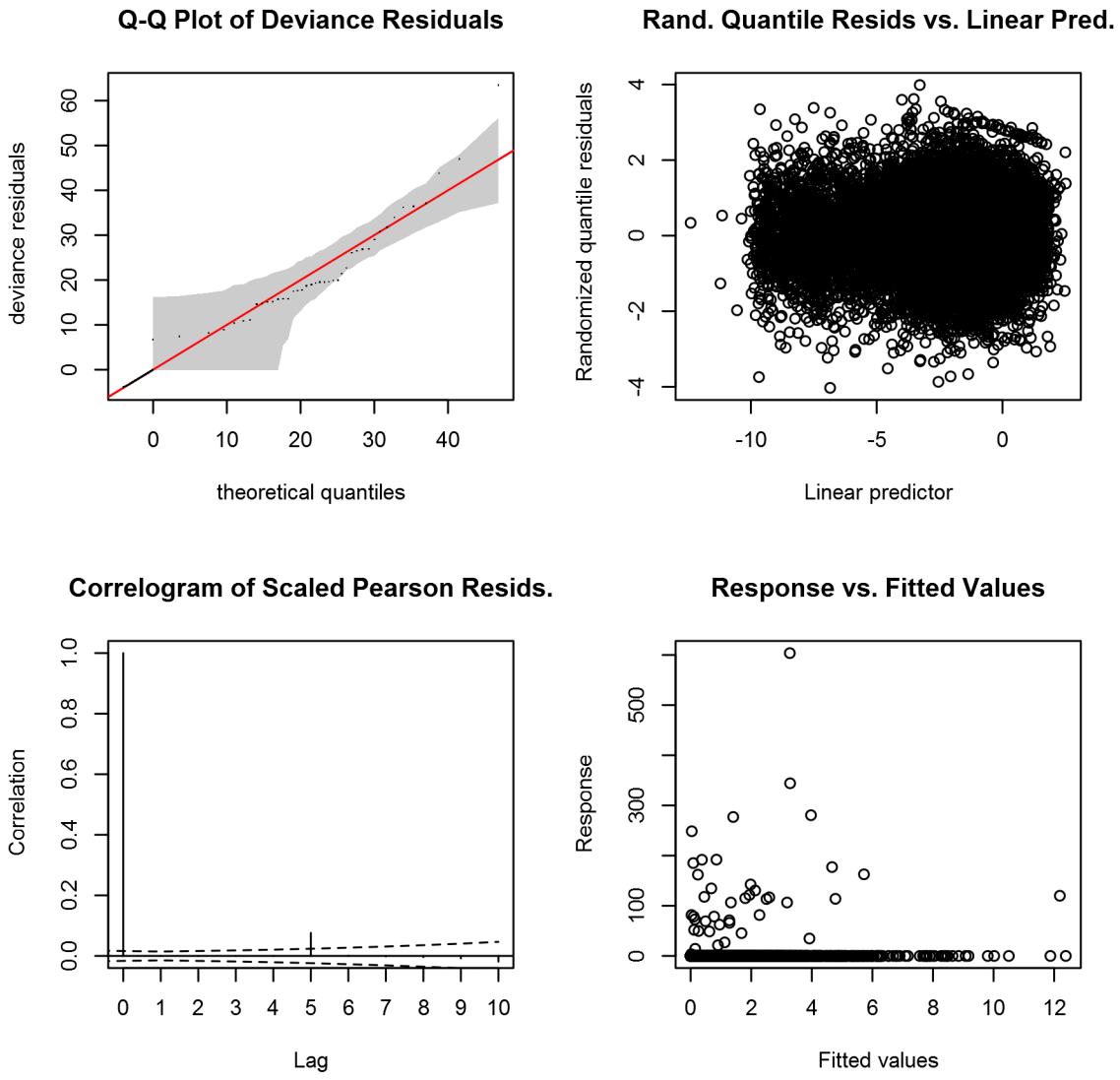


Figure 35: Statistical diagnostic plots for the Pilot whales Contemporaneous model, Off Shelf.

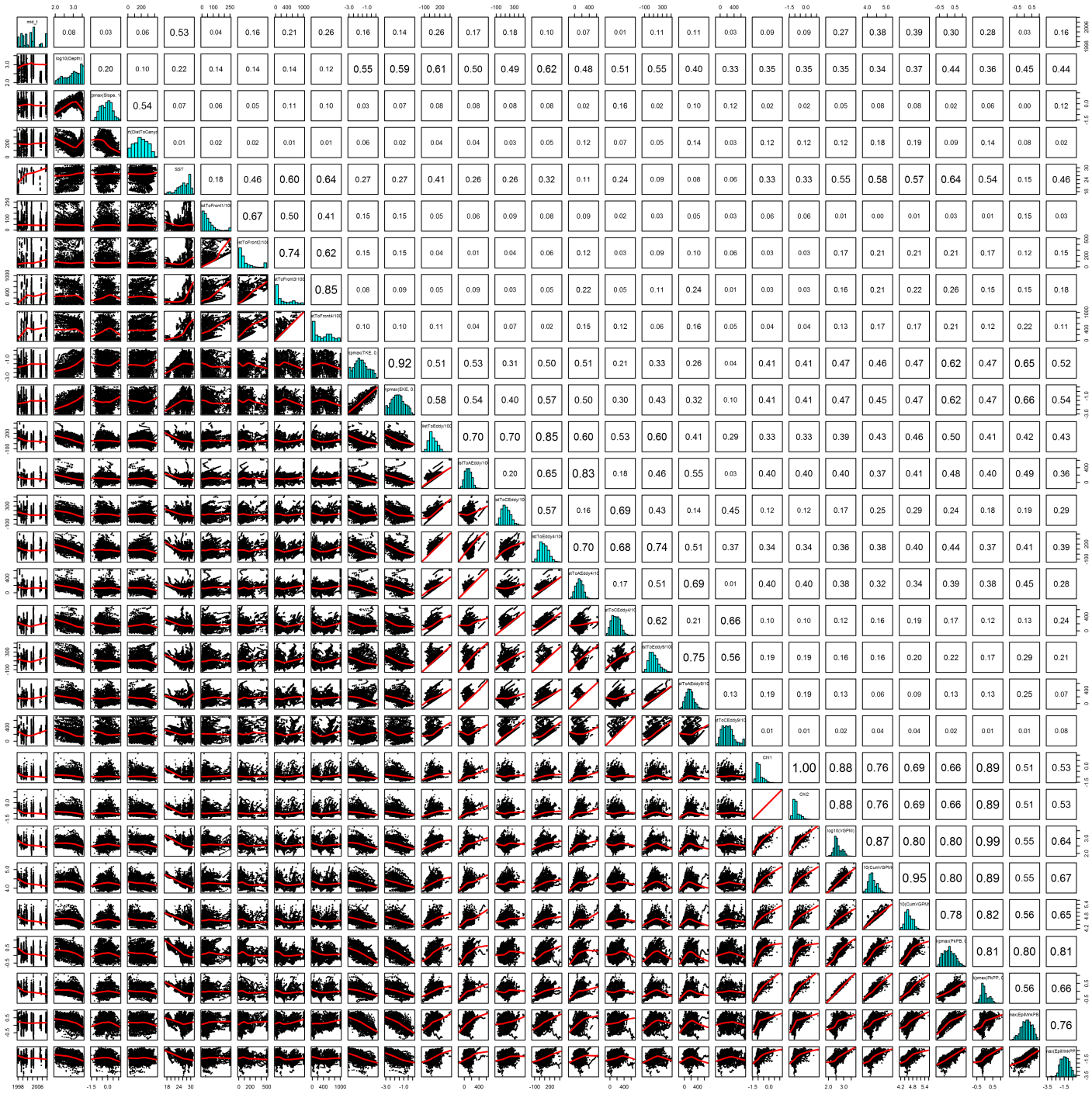


Figure 36: Scatterplot matrix for the Pilot whales Contemporaneous model, Off Shelf. This plot is used to inspect the distribution of predictors (via histograms along the diagonal), simple correlation between predictors (via pairwise Pearson coefficients above the diagonal), and linearity of predictor correlations (via scatterplots below the diagonal). This plot is best viewed at high magnification.

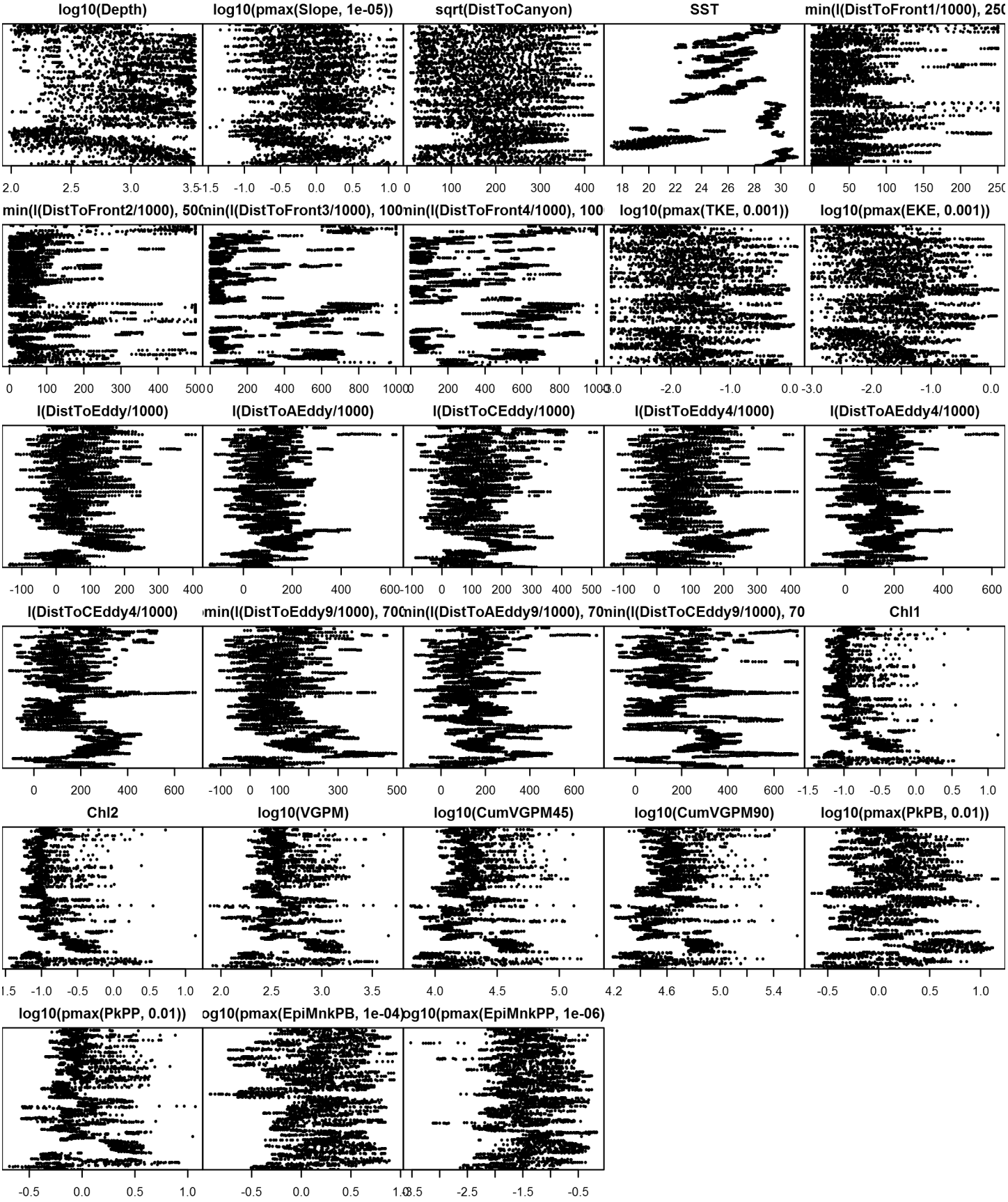


Figure 37: Dotplot for the Pilot whales Contemporaneous model, Off Shelf. This plot is used to check for suspicious patterns and outliers in the data. Points are ordered vertically by transect ID, sequentially in time.

On Shelf

Density assumed to be 0 in this region.

Climatological Same Segments Model

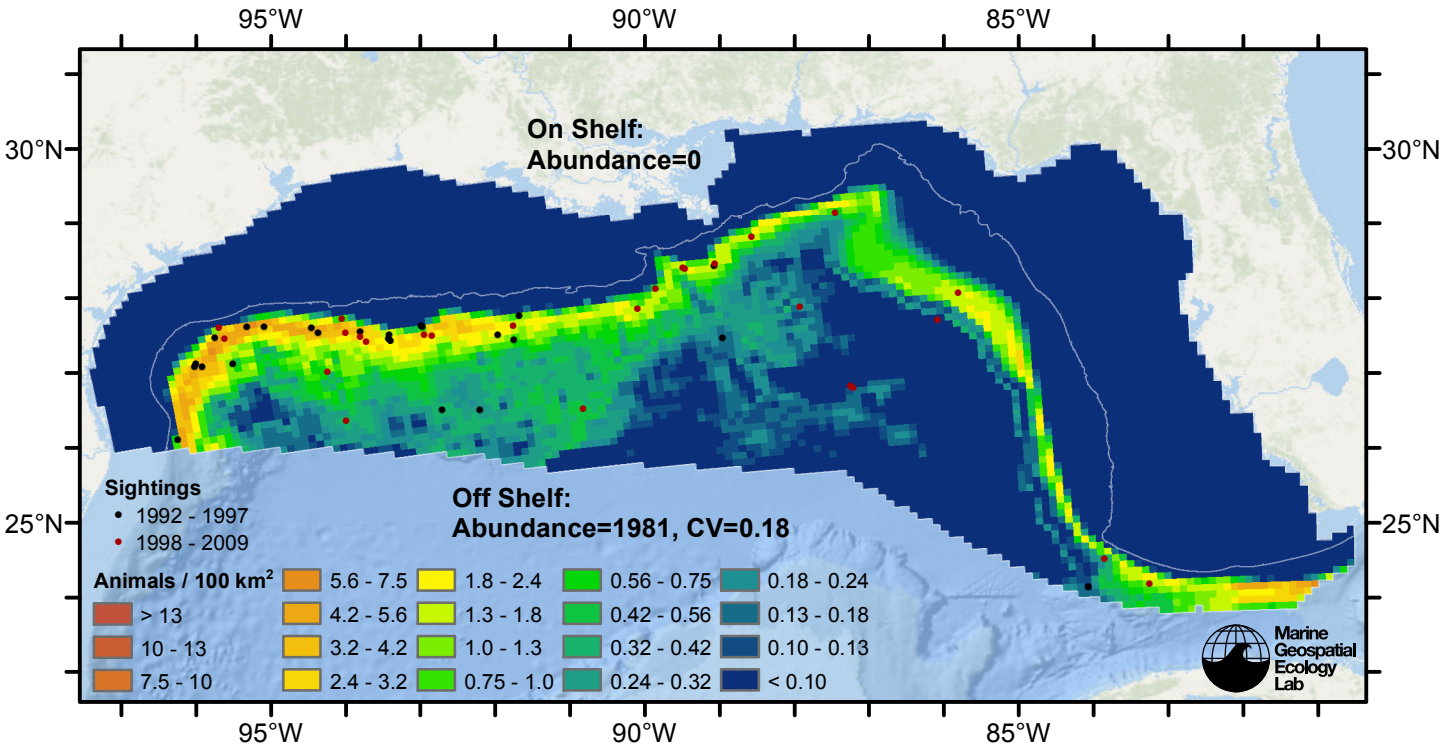


Figure 38: Pilot whales density predicted by the climatological same segments model that explained the most deviance. Pixels are 10x10 km. The legend gives the estimated individuals per pixel; breaks are logarithmic. Abundance for each region was computed by summing the density cells occuring in that region.

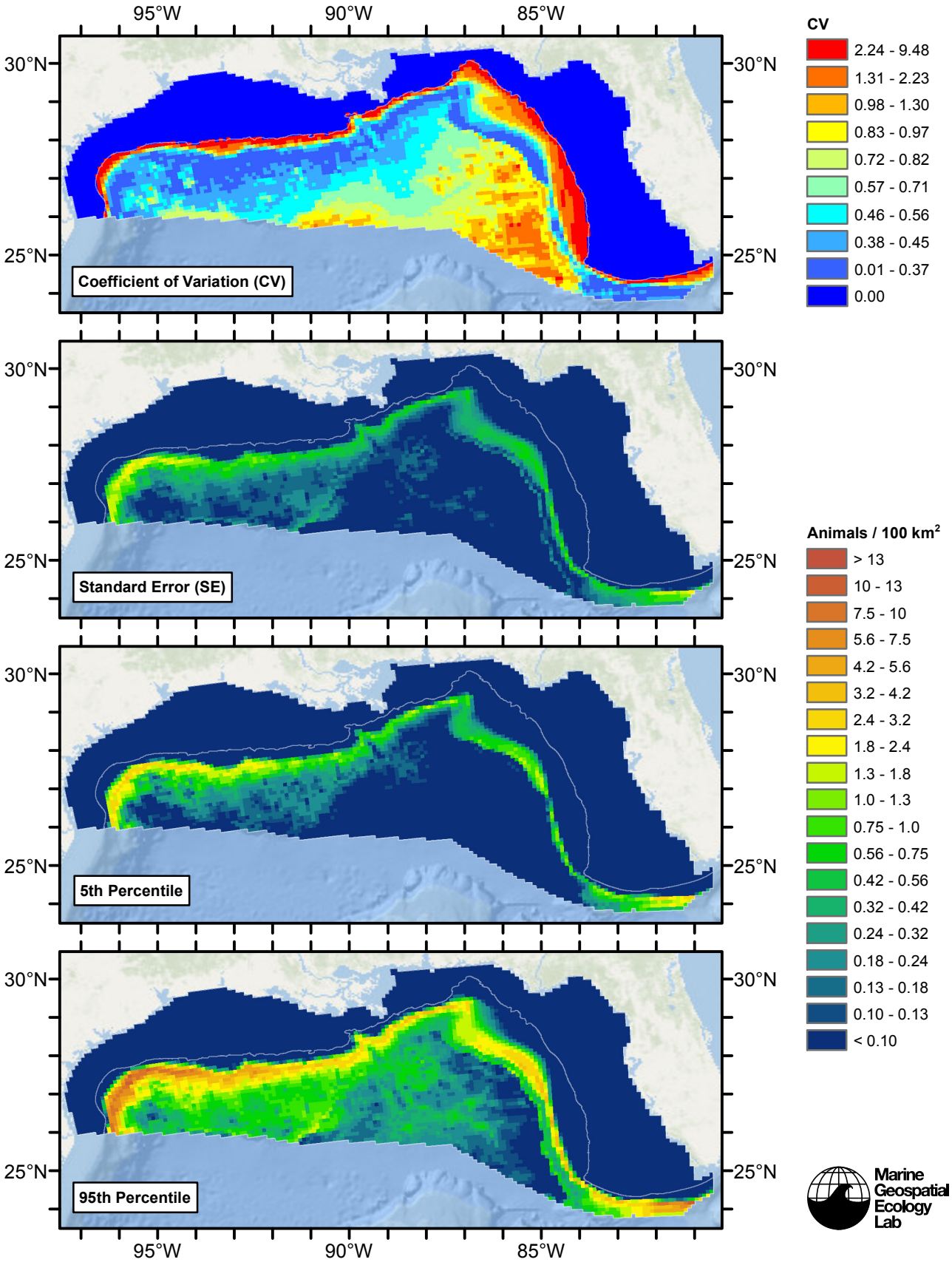


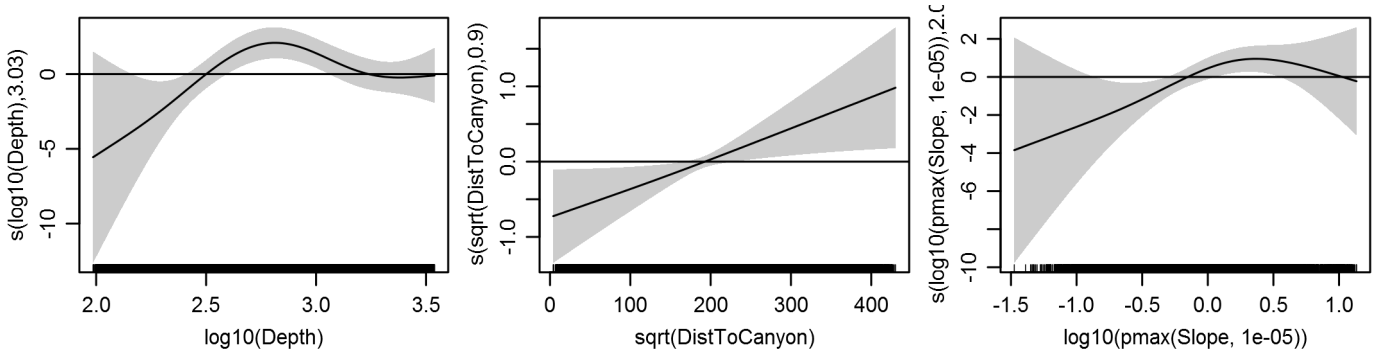
Figure 39: Estimated uncertainty for the climatological same segments model that explained the most deviance. These estimates only incorporate the statistical uncertainty estimated for the spatial model (by the R mgcv package). They do not incorporate uncertainty in the detection functions, g(0) estimates, predictor variables, and so on.

Off Shelf

Statistical output

```
Rscript.exe: This is mgcv 1.8-3. For overview type 'help("mgcv-package")'.


Family: Tweedie(p=1.31)
Link function: log

Formula:
abundance ~ offset(log(area_km2)) + s(log10(Depth), bs = "ts",
    k = 5) + s(log10(pmax(Slope, 1e-05)), bs = "ts", k = 5) +
    s(sqrt(DistToCanyon), bs = "ts", k = 5)

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -6.4625     0.4674  -13.83   <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(Depth))              3.0290      4 5.971 7.11e-06 ***
s(log10(pmax(Slope, 1e-05))) 2.0499      4 2.091  0.00867 **
s(sqrt(DistToCanyon))        0.9034      4 1.549  0.00755 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) =  0.00116   Deviance explained = 18.2%
-REML = 576.11  Scale est. = 208.08    n = 14455

All predictors were significant. This is the final model.
Creating term plots.
Diagnostic output from gam.check():

Method: REML   Optimizer: outer newton
full convergence after 11 iterations.
Gradient range [-6.049325e-06,4.133098e-07]
(score 576.1082 & scale 208.0793).
Hessian positive definite, eigenvalue range [0.237974,182.0222].
Model rank =  13 / 13

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(Depth))              4.000 3.029   0.517    0.00
s(log10(pmax(Slope, 1e-05))) 4.000 2.050   0.538    0.00
s(sqrt(DistToCanyon))        4.000 0.903   0.592    0.08

Predictors retained during the model selection procedure: Depth, Slope, DistToCanyon

Predictors dropped during the model selection procedure: ClimSST, ClimDistToFront1
```

Model term plots



Diagnostic plots

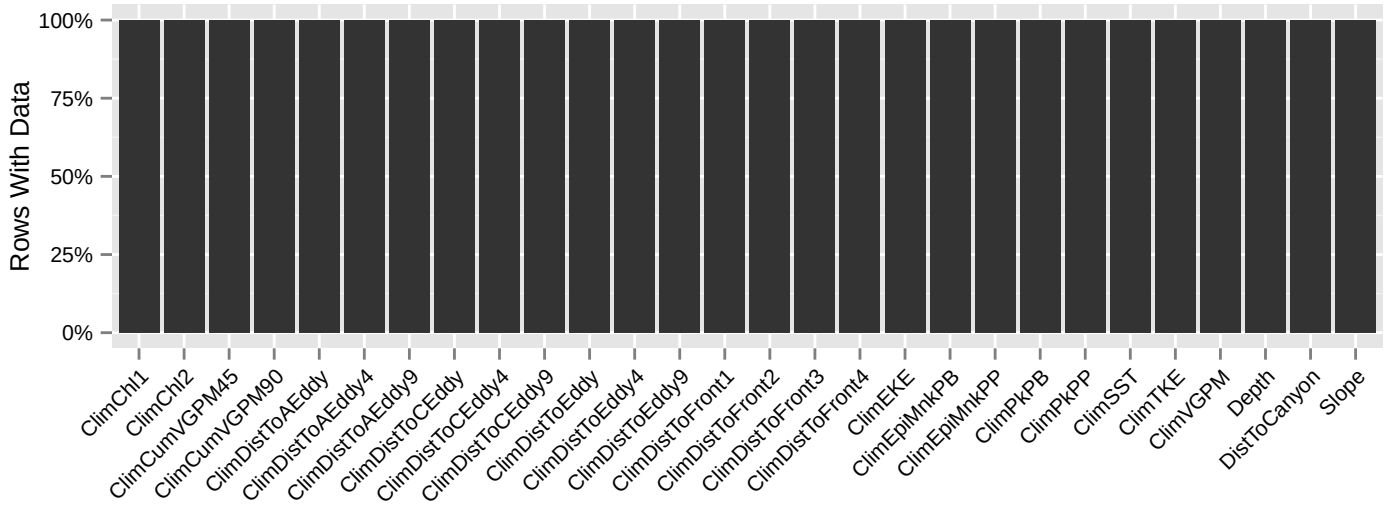


Figure 40: Segments with predictor values for the Pilot whales Climatological model, Off Shelf. This plot is used to assess how many segments would be lost by including a given predictor in a model.

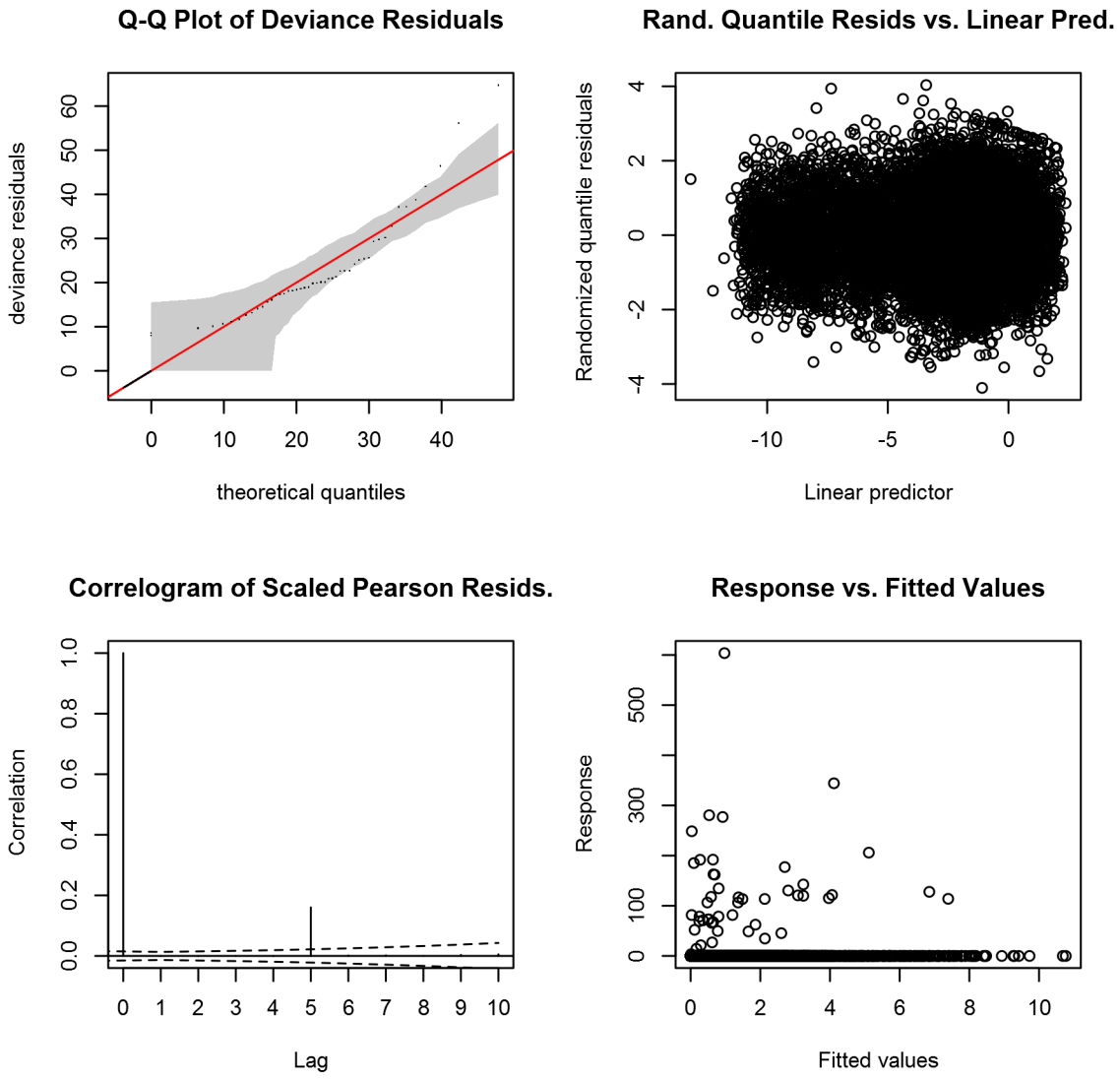


Figure 41: Statistical diagnostic plots for the Pilot whales Climatological model, Off Shelf.

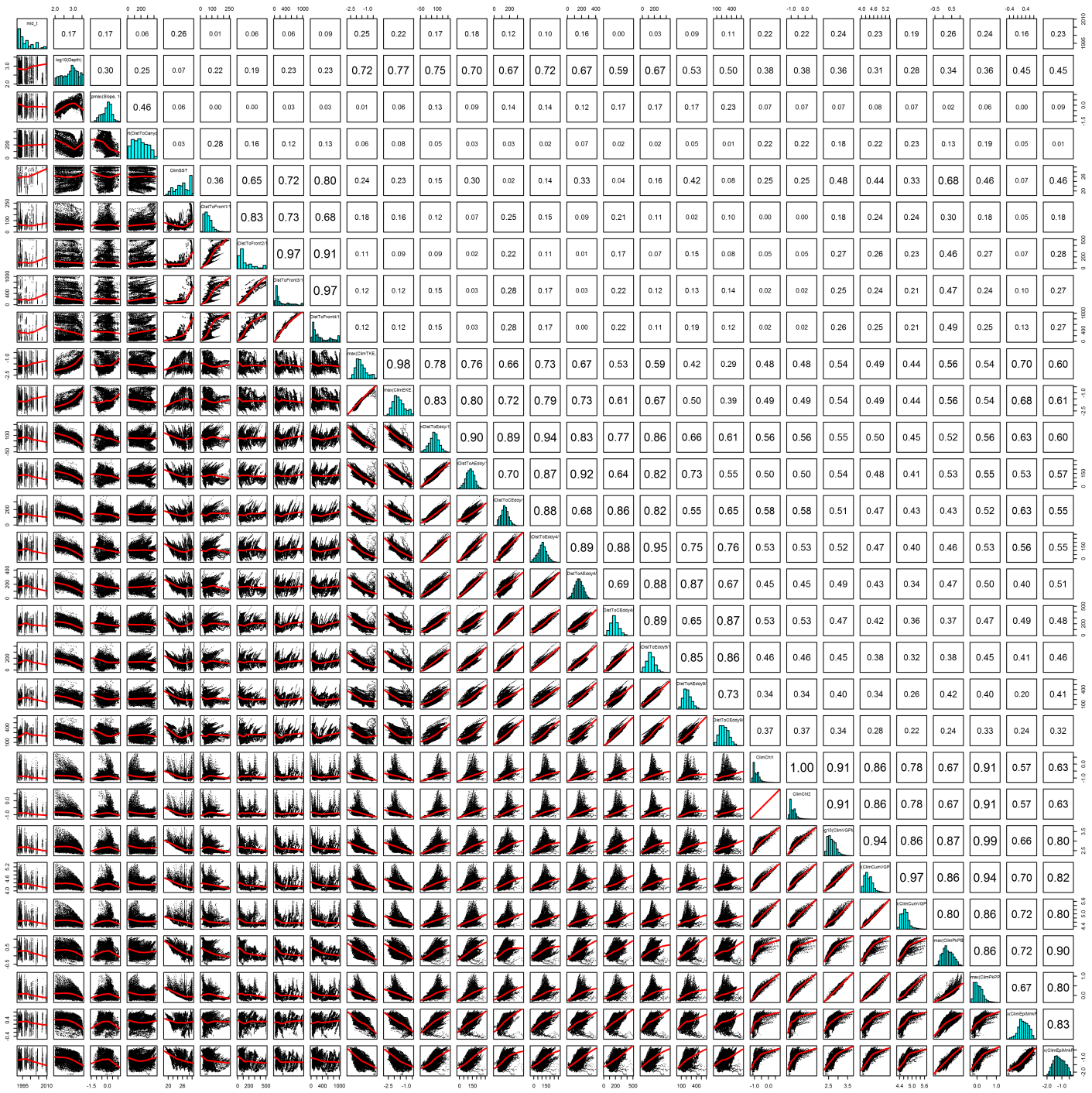


Figure 42: Scatterplot matrix for the Pilot whales Climatological model, Off Shelf. This plot is used to inspect the distribution of predictors (via histograms along the diagonal), simple correlation between predictors (via pairwise Pearson coefficients above the diagonal), and linearity of predictor correlations (via scatterplots below the diagonal). This plot is best viewed at high magnification.

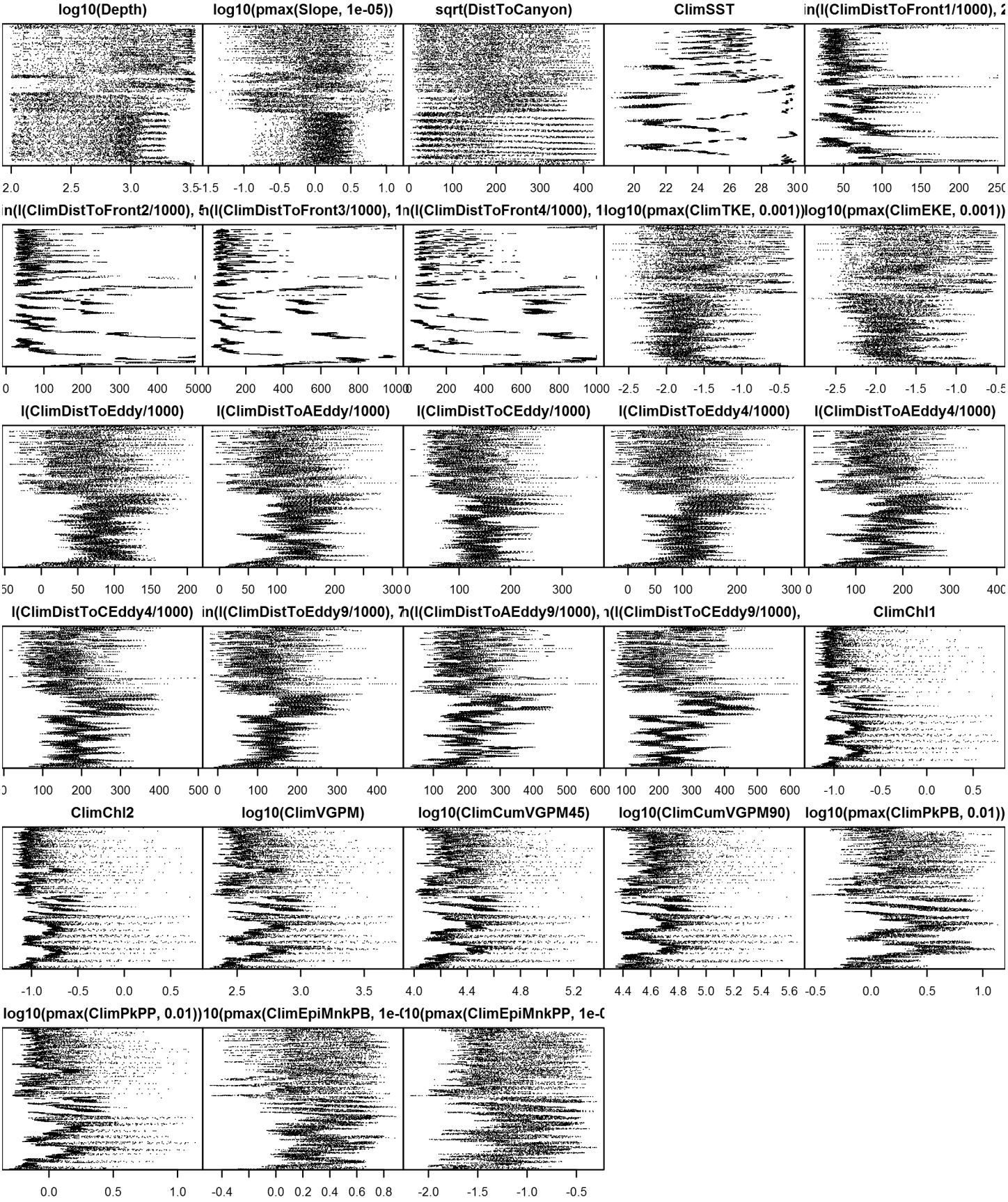


Figure 43: Dotplot for the Pilot whales Climatological model, Off Shelf. This plot is used to check for suspicious patterns and outliers in the data. Points are ordered vertically by transect ID, sequentially in time.

On Shelf

Density assumed to be 0 in this region.

Model Comparison

Spatial Model Performance

The table below summarizes the performance of the candidate spatial models that were tested. The first model contained only physiographic predictors. Subsequent models added additional suites of predictors of based on when they became available via remote sensing.

For each model, three versions were fitted; the % Dev Expl columns give the % deviance explained by each one. The "climatological" models were fitted to 8-day climatologies of the environmental predictors. Because the environmental predictors were always available, no segments were lost, allowing these models to consider the maximal amount of survey data. The "contemporaneous" models were fitted to day-of-sighting images of the environmental predictors; these were smoothed to reduce data loss due to clouds, but some segments still failed to retrieve environmental values and were lost. Finally, the "climatological same segments" models fitted climatological predictors to the segments retained by the contemporaneous model, so that the explantory power of the two types of predictors could be directly compared. For each of the three models, predictors were selected independently via shrinkage smoothers; thus the three models did not necessarily utilize the same predictors.

Predictors derived from ocean currents first became available in January 1993 after the launch of the TOPEX/Poseidon satellite; productivity predictors first became available in September 1997 after the launch of the SeaWiFS sensor. Contemporaneous and climatological same segments models considering these predictors usually suffered data loss. Date Range shows the years spanned by the retained segments. The Segments column gives the number of segments retained; % Lost gives the percentage lost.

Predictors	Climatol % Dev Expl	Contemp % Dev Expl	Climatol Same Segs % Dev Expl	Segments	% Lost	Date Range
Phys	18.2			14455		1992-2009
Phys+SST	18.2	18.2	18.2	14455	0.0	1992-2009
Phys+SST+Curr	18.2	21.0	15.9	12354	14.5	1993-2009
Phys+SST+Curr+Prod	18.2	13.0	18.2	14455	0.0	1992-2009

Table 14: Deviance explained by the candidate density models.

Abundance Estimates

The table below shows the estimated mean abundance (number of animals) within the study area, for the models that explained the most deviance for each model type. Mean abundance was calculated by first predicting density maps for a series of time steps, then computing the abundance for each map, and then averaging the abundances. For the climatological models, we used 8-day climatologies, resulting in 46 abundance maps. For the contemporaneous models, we used daily images, resulting in 365 predicted abundance maps per year that the prediction spanned. The Dates column gives the dates to which the estimates apply. For our models, these are the years for which both survey data and remote sensing data were available.

The Assumed g(0)=1 column specifies whether the abundance estimate assumed that detection was certain along the survey trackline. Studies that assumed this did not correct for availability or perception bias, and therefore underestimated abundance. The In our models column specifies whether the survey data from the study was also used in our models. If not, the study provides a completely independent estimate of abundance.

Dates	Model or study	Estimated abundance	CV	Assumed g(0)=1	In our models

1992-2009	Climatological model*	1981	0.18	No	
1993-2009	Contemporaneous model	2801	0.69	No	
1992-2009	Climatological same segments model	1981	0.18	No	
2009	Oceanic waters, Jun-Aug (Waring et al. 2013)	2415	0.66	Yes	Yes
2003-2004	Oceanic waters, Jun-Aug (Mullin 2007)	716	0.34	Yes	Yes
1996-2001	Oceanic waters, Apr-Jun (Mullin and Fulling 2004)	2388	0.48	Yes	Yes
1991-1994	Oceanic waters, Apr-Jun (Hansen et al. 1995)	353	0.89	Yes	Yes

Table 15: Estimated mean abundance within the study area. We selected the model marked with * as our best estimate of the abundance and distribution of this taxon. For comparison, independent abundance estimates from NOAA technical reports and/or the scientific literature are shown. Please see the Discussion section below for our evaluation of our models compared to the other estimates. Note that our abundance estimates are averaged over the whole year, while the other studies may have estimated abundance for specific months or seasons. Our coefficients of variation (CVs) underestimate the true uncertainty in our estimates, as they only incorporated the uncertainty of the GAM stage of our models. Other sources of uncertainty include the detection functions and g(0) estimates. It was not possible to incorporate these into our CVs without undertaking a computationally-prohibitive bootstrap; we hope to attempt that in a future version of our models.

Density Maps

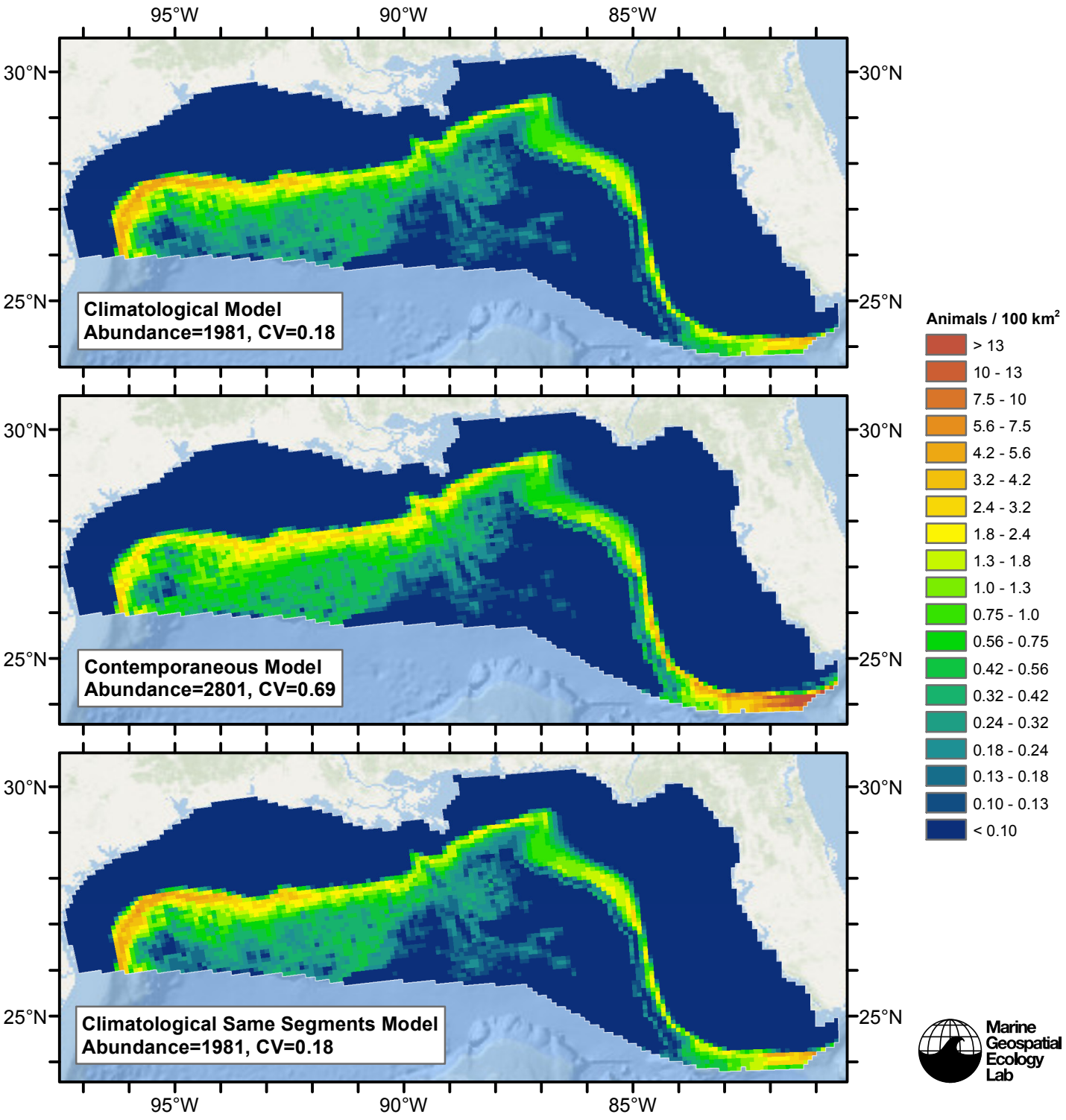


Figure 44: Pilot whales density and abundance predicted by the models that explained the most deviance. Regions inside the study area (white line) where the background map is visible are areas we did not model (see text).

Temporal Variability

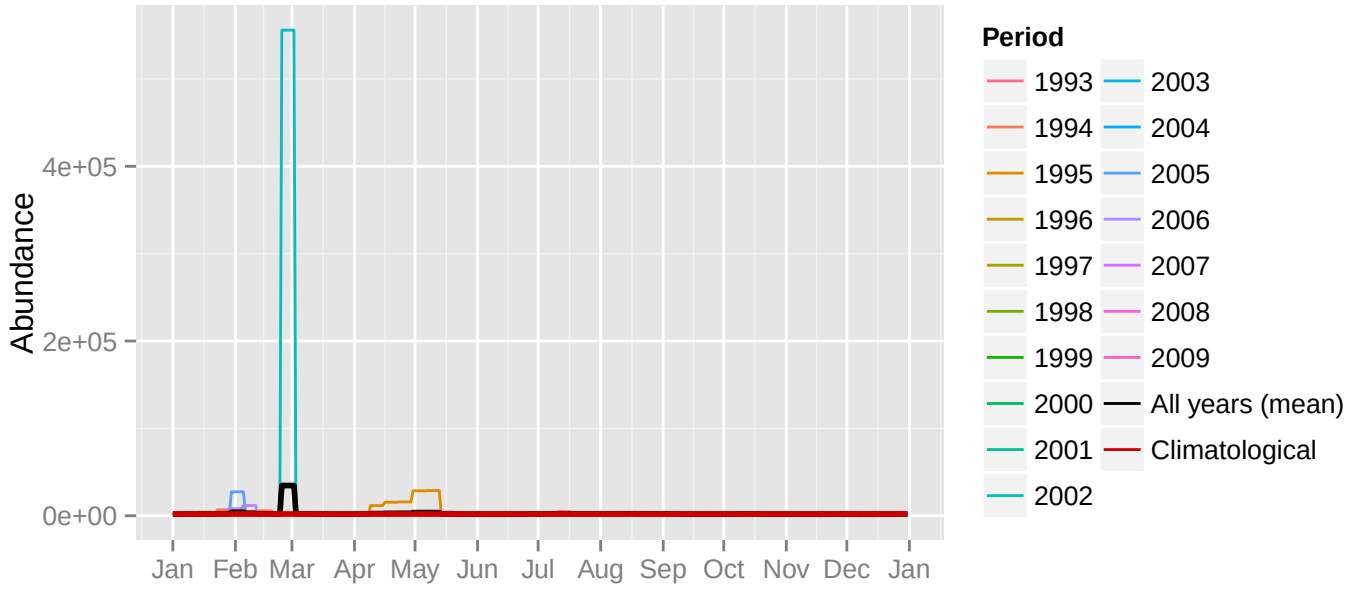


Figure 45: Comparison of Pilot whales abundance predicted at a daily time step for different time periods. Individual years were predicted using contemporaneous models. "All years (mean)" averages the individual years, giving the mean annual abundance of the contemporaneous model. "Climatological" was predicted using the climatological model. The results for the climatological same segments model are not shown.

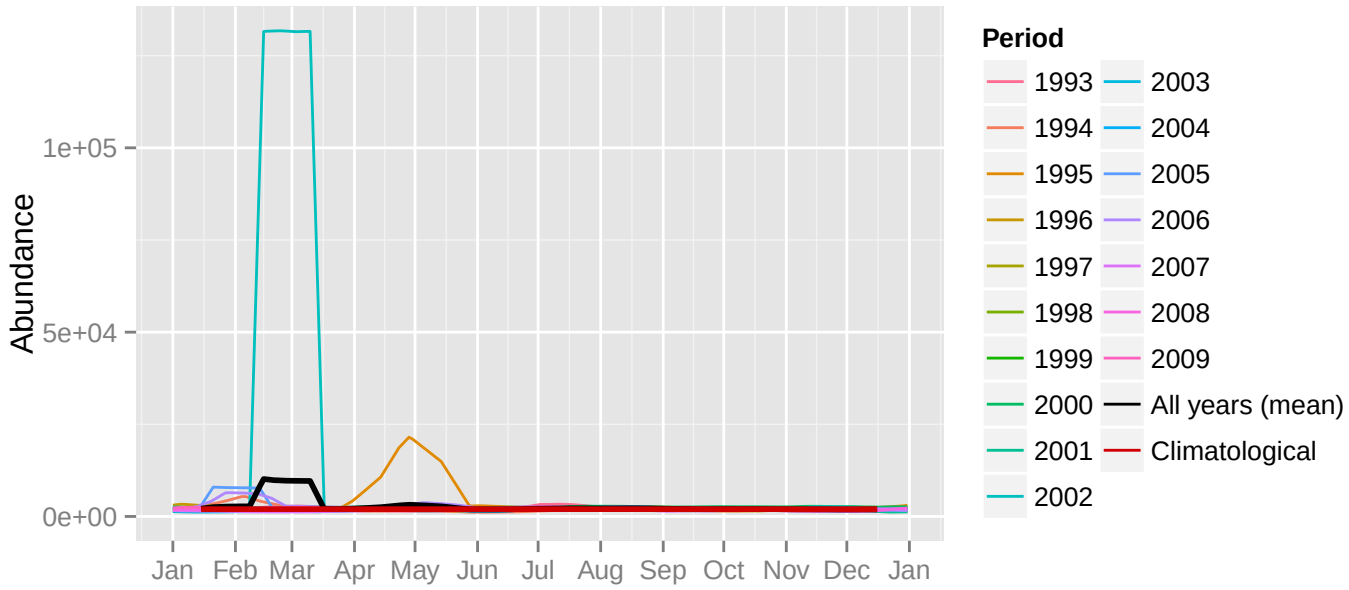
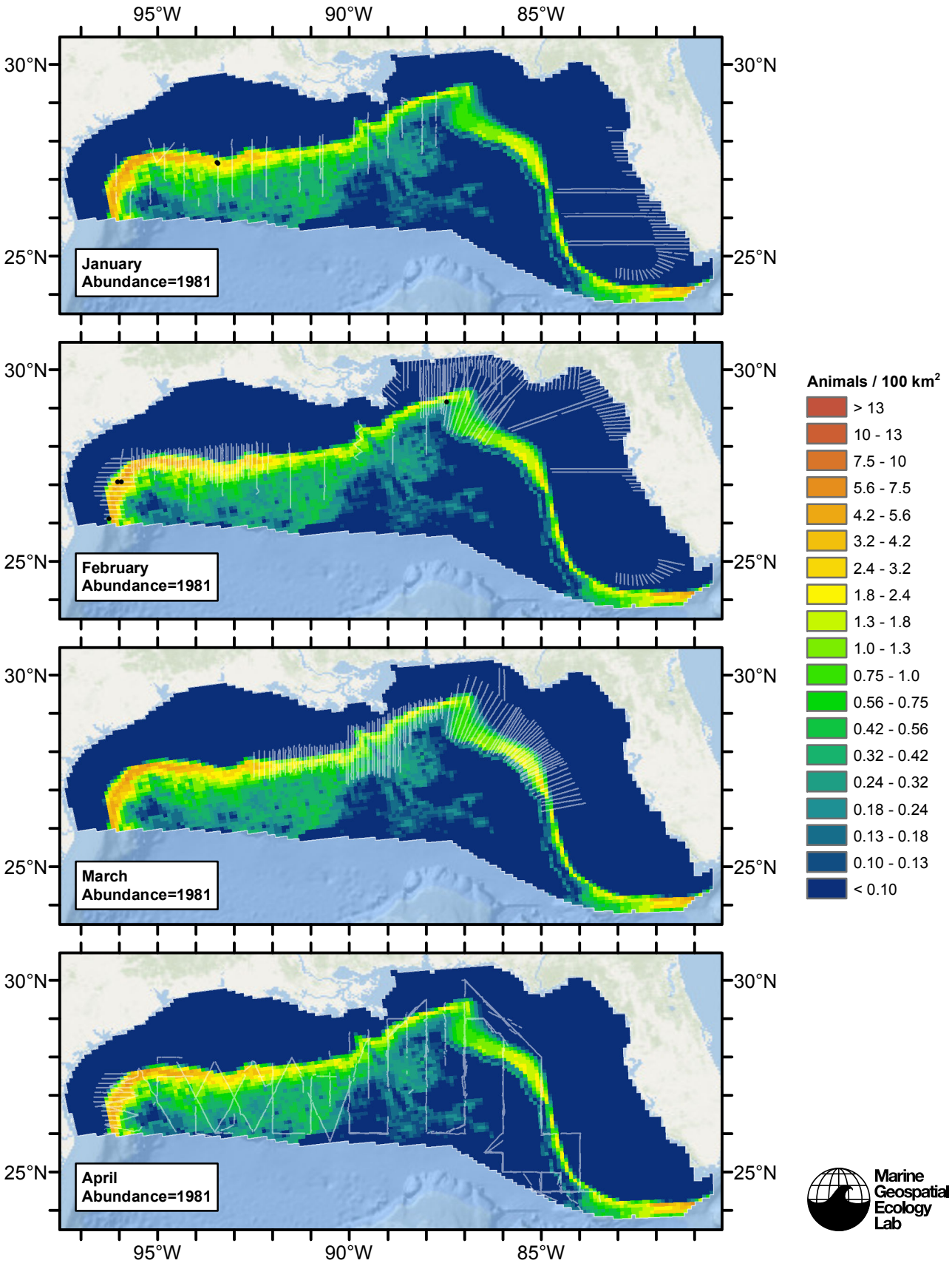
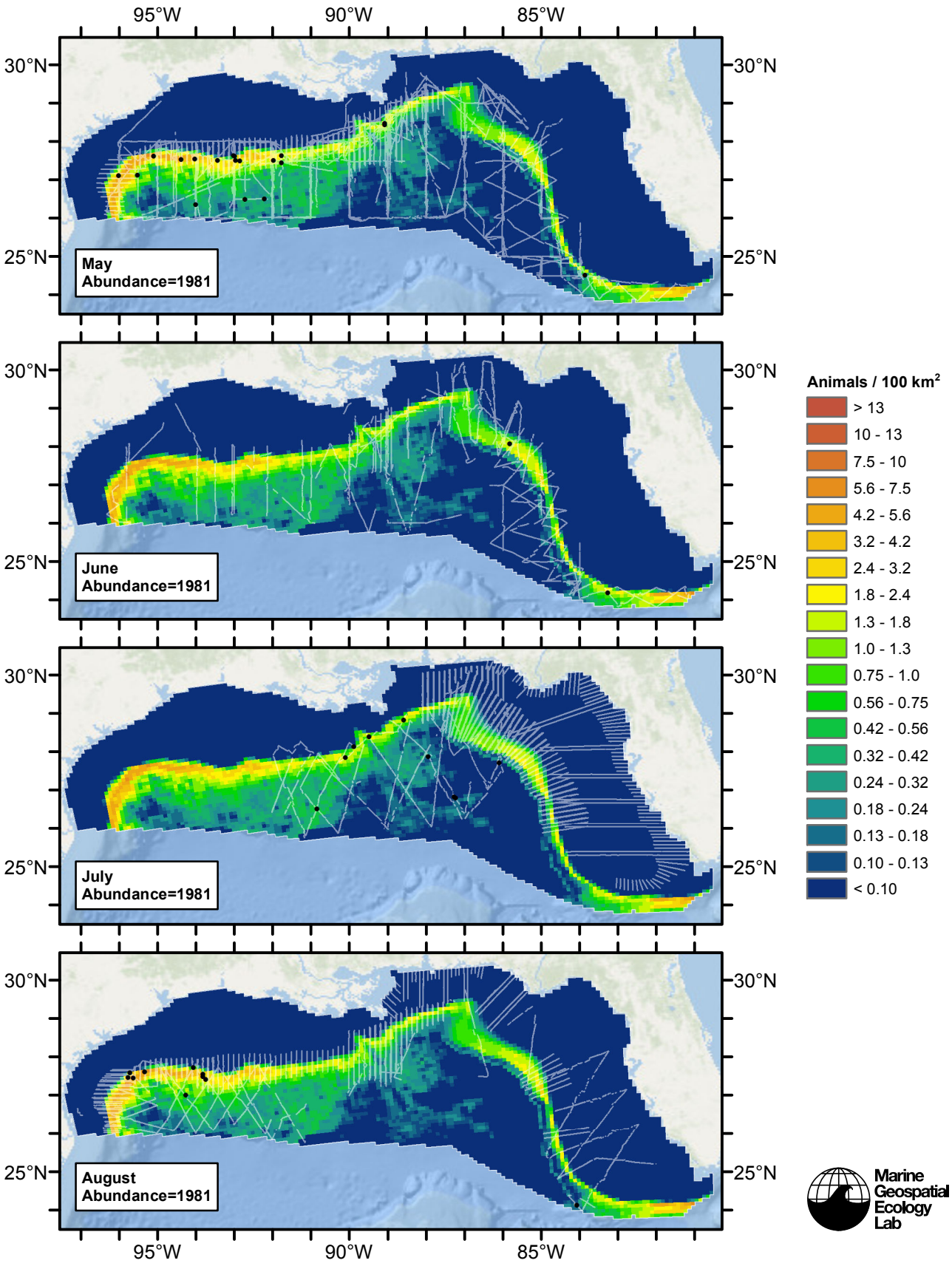
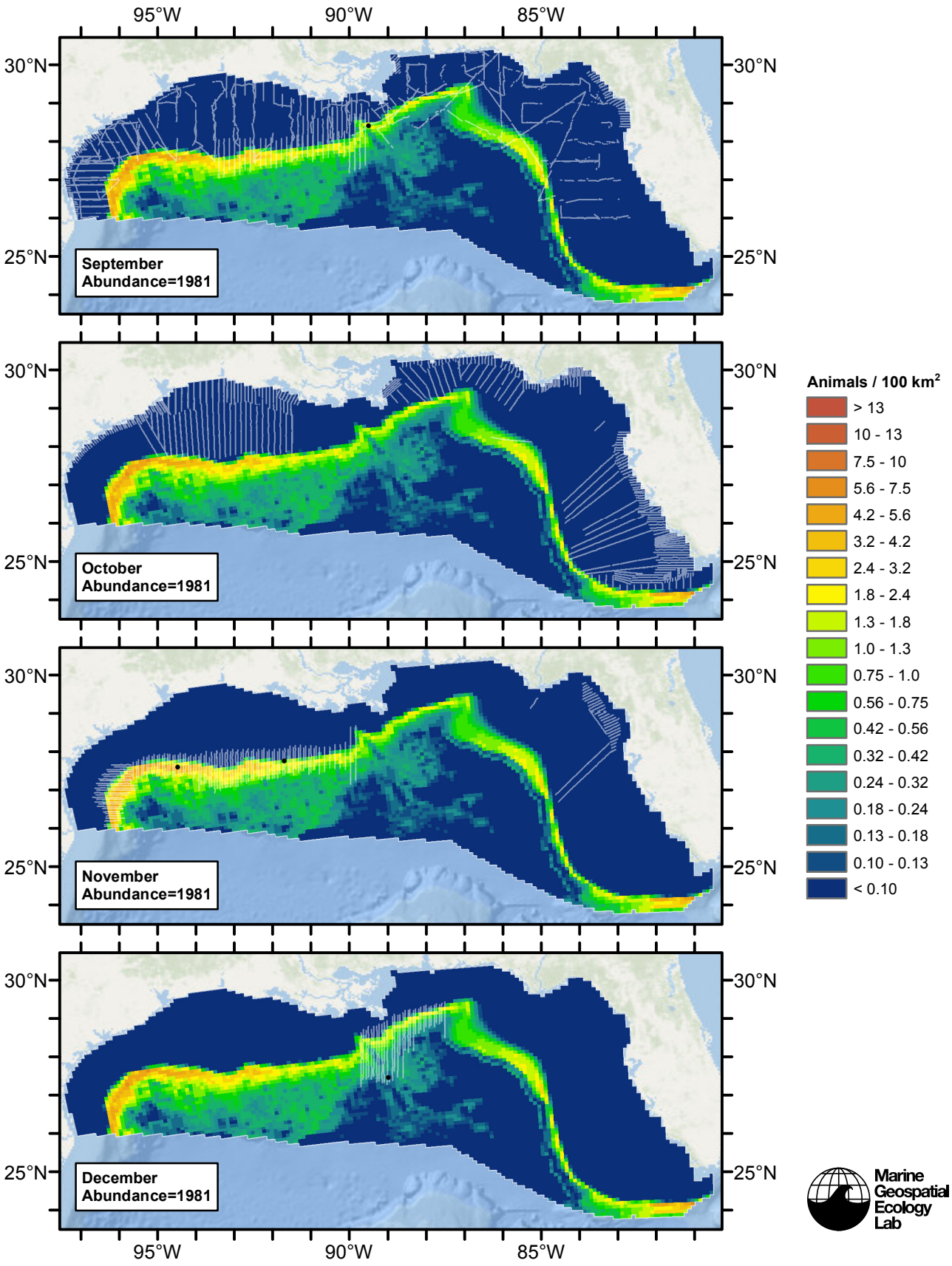
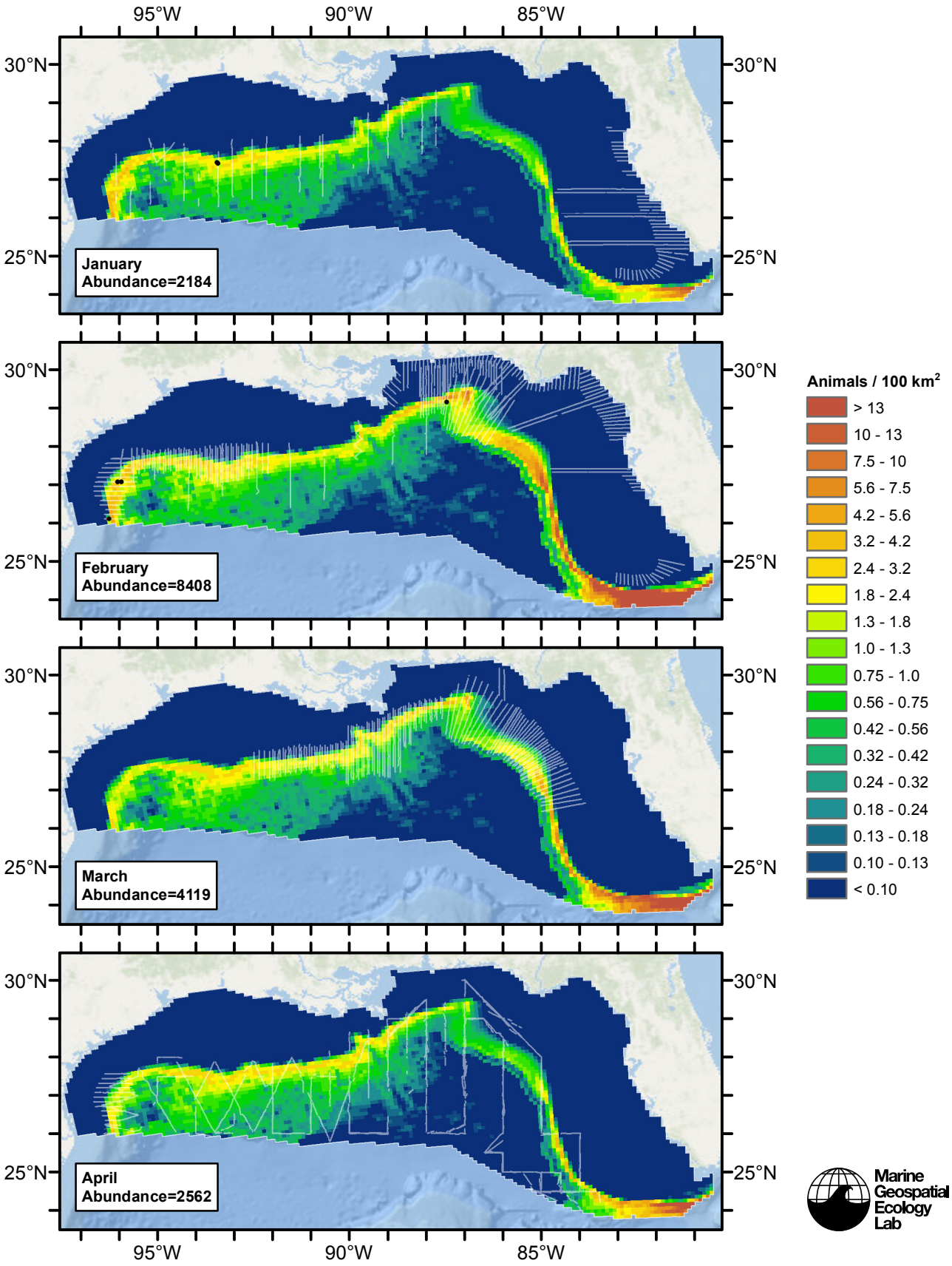


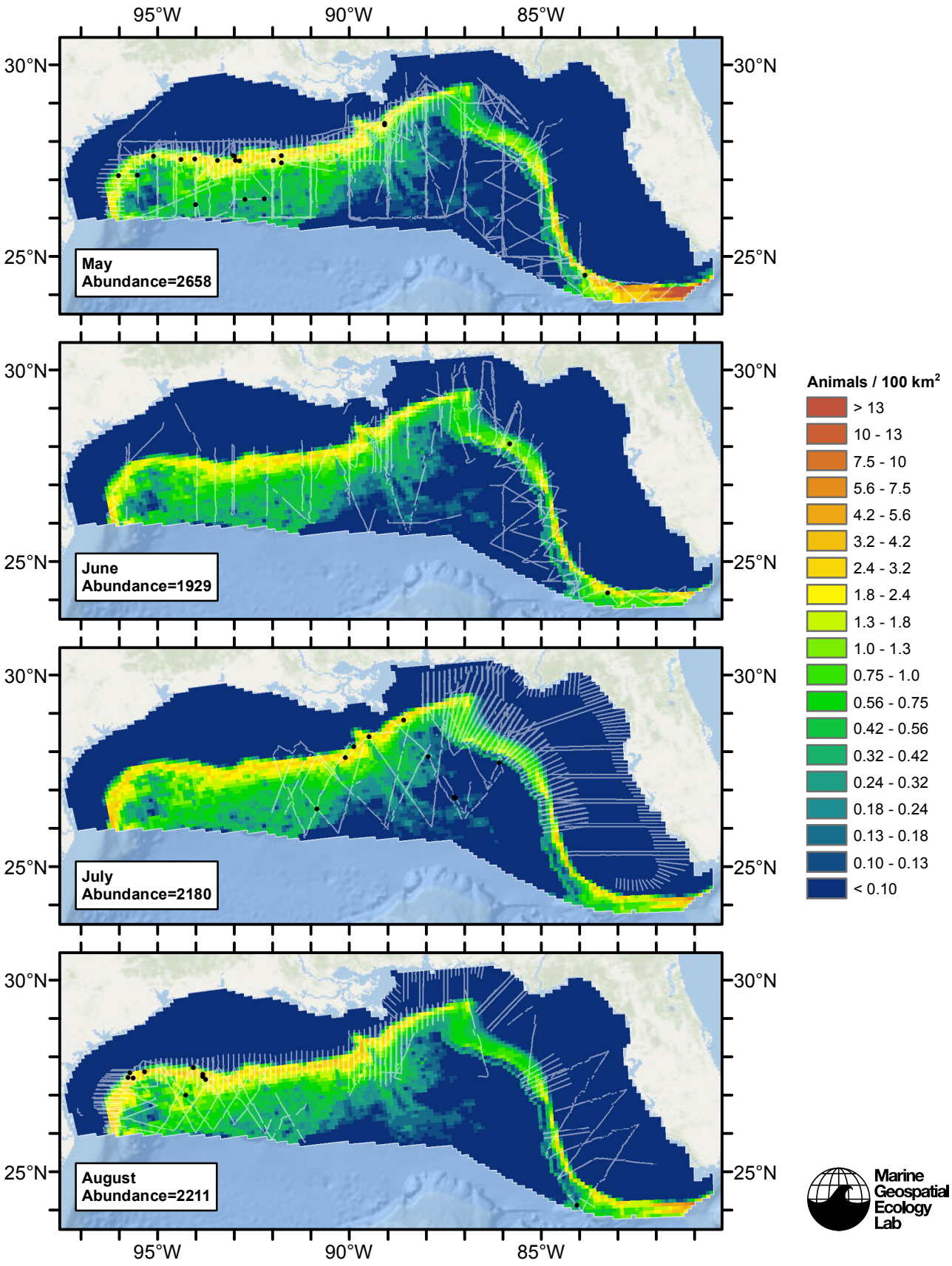
Figure 46: The same data as the preceding figure, but with a 30-day moving average applied.

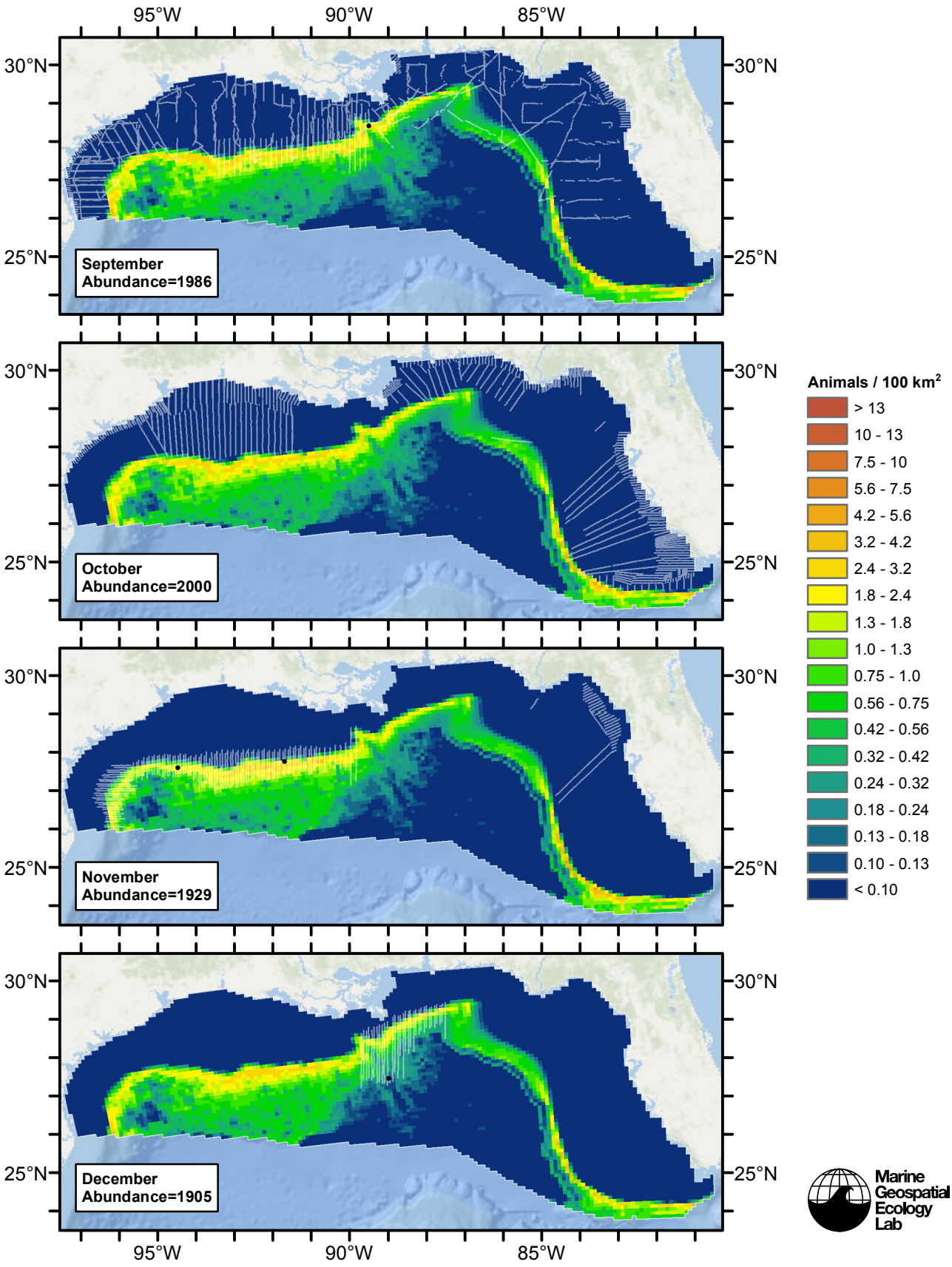

95°W
90°W
85°W
30°N
25°N
January
Abundance=1981
February
Abundance=1981
March
Abundance=1981
April
Abundance=1981
Animals / 100 km²
> 13
10 - 13
7.5 - 10
5.6 - 7.5
4.2 - 5.6
3.2 - 4.2
2.4 - 3.2
1.8 - 2.4
1.3 - 1.8
1.0 - 1.3
0.75 - 1.0
0.56 - 0.75
0.42 - 0.56
0.32 - 0.42
0.24 - 0.32
0.18 - 0.24
0.13 - 0.18
0.10 - 0.13
< 0.10
Marine
Geospatial
Ecology
Lab

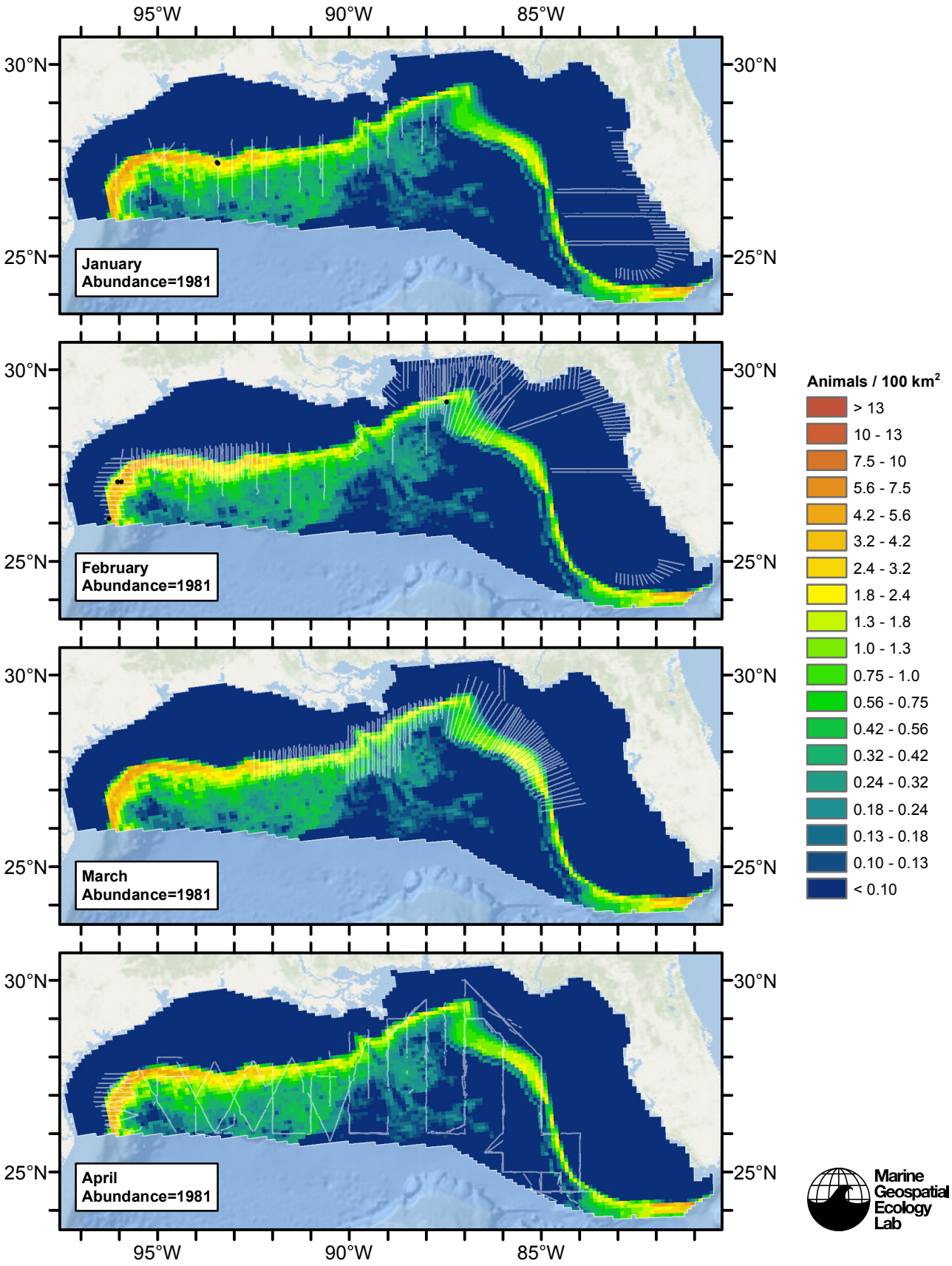

95°W
90°W
85°W
30°N
25°N
May
Abundance=1981
June
Abundance=1981
July
Abundance=1981
August
Abundance=1981
Animals / 100 km²
> 13
10 - 13
7.5 - 10
5.6 - 7.5
4.2 - 5.6
3.2 - 4.2
2.4 - 3.2
1.8 - 2.4
1.3 - 1.8
1.0 - 1.3
0.75 - 1.0
0.56 - 0.75
0.42 - 0.56
0.32 - 0.42
0.24 - 0.32
0.18 - 0.24
0.13 - 0.18
0.10 - 0.13
< 0.10
Marine
Geospatial
Ecology
Lab

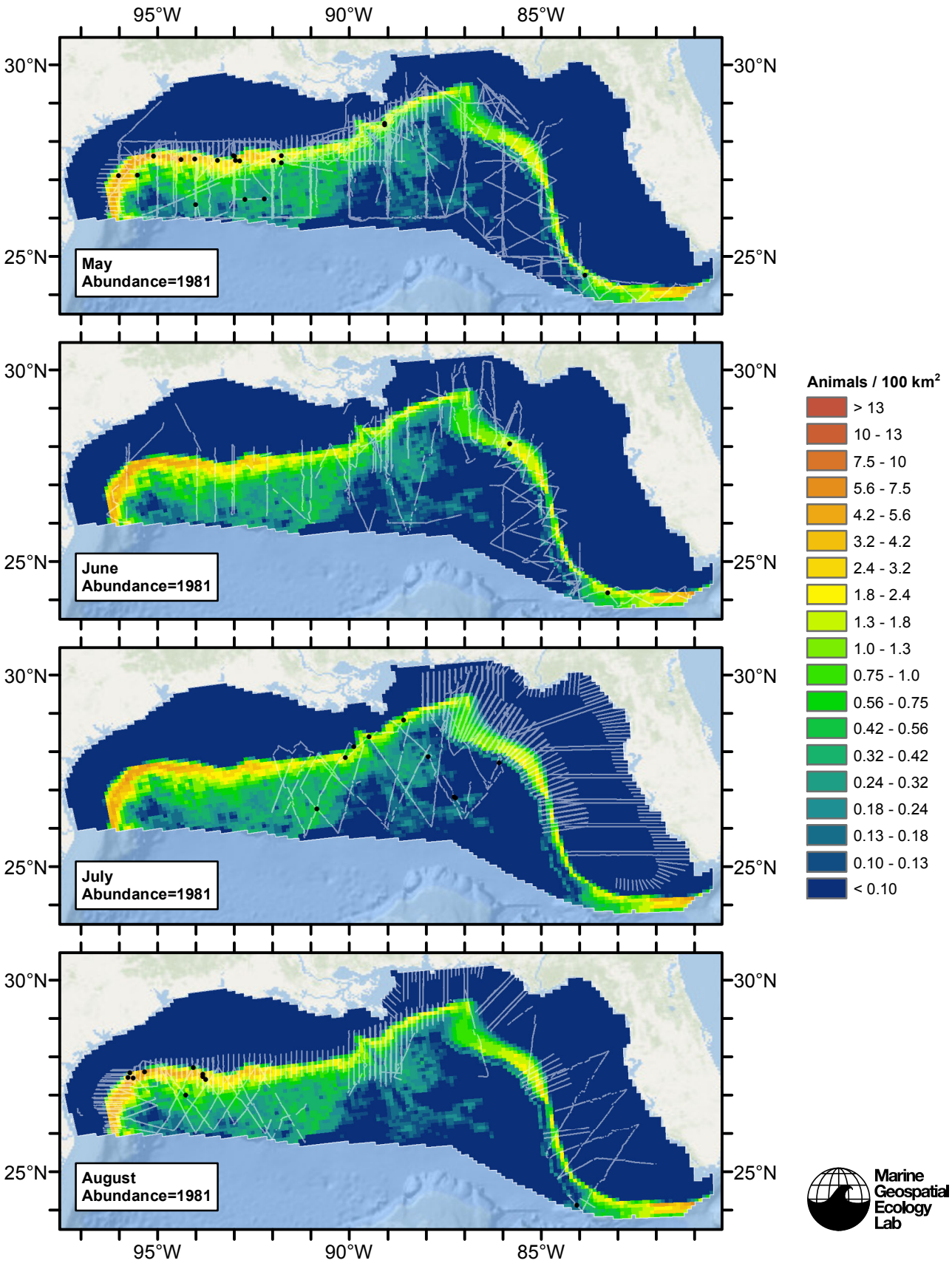

95°W
90°W
85°W
30°N
25°N
September
Abundance=1981
October
Abundance=1981
November
Abundance=1981
December
Abundance=1981
Animals / 100 km²
> 13
10 - 13
7.5 - 10
5.6 - 7.5
4.2 - 5.6
3.2 - 4.2
2.4 - 3.2
1.8 - 2.4
1.3 - 1.8
1.0 - 1.3
0.75 - 1.0
0.56 - 0.75
0.42 - 0.56
0.32 - 0.42
0.24 - 0.32
0.18 - 0.24
0.13 - 0.18
0.10 - 0.13
< 0.10
Marine
Geospatial
Ecology
Lab

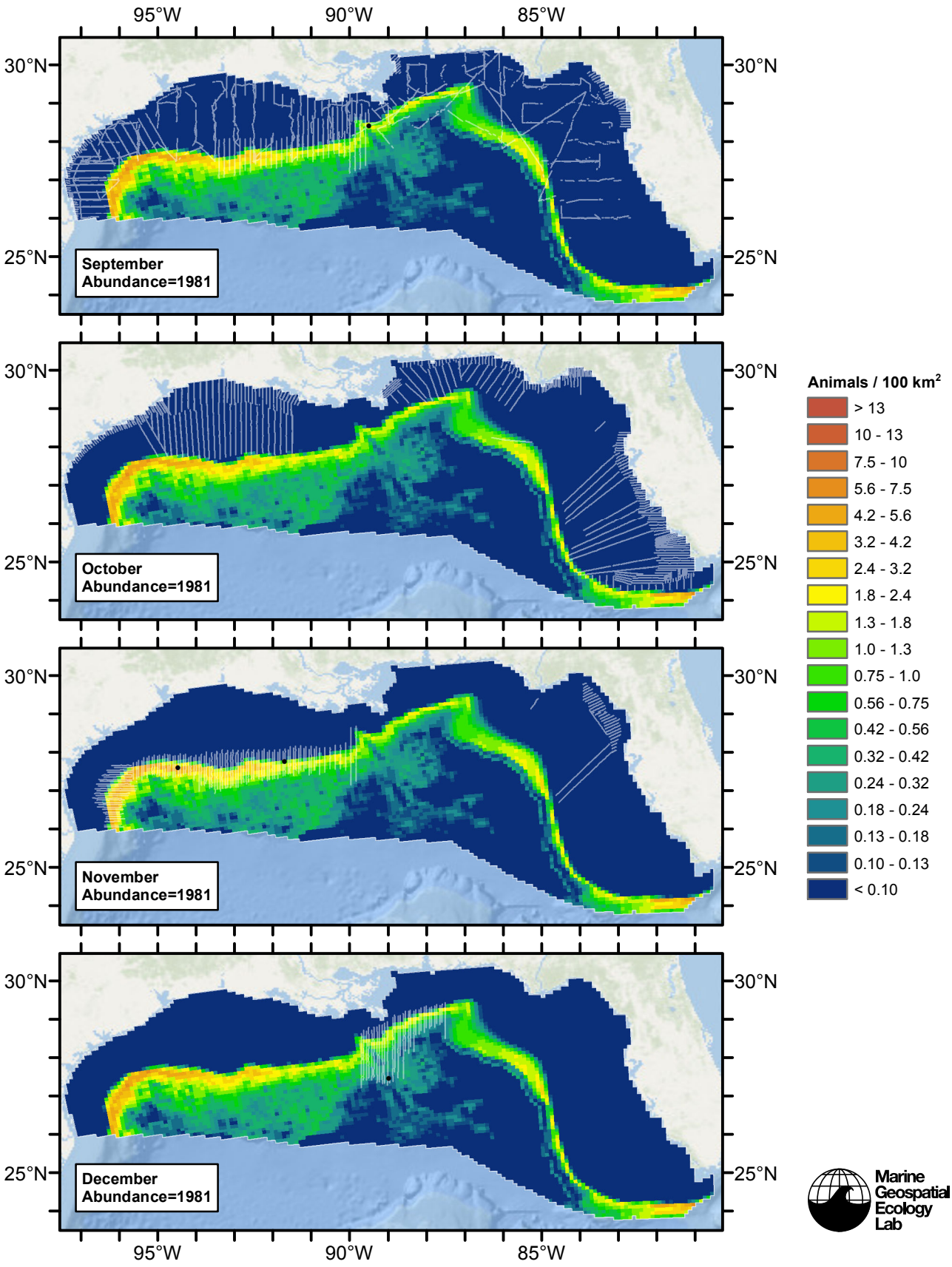

95°W
90°W
85°W
30°N
25°N
January
Abundance=2184
February
Abundance=8408
March
Abundance=4119
April
Abundance=2562
Animals / 100 km²
> 13
10 - 13
7.5 - 10
5.6 - 7.5
4.2 - 5.6
3.2 - 4.2
2.4 - 3.2
1.8 - 2.4
1.3 - 1.8
1.0 - 1.3
0.75 - 1.0
0.56 - 0.75
0.42 - 0.56
0.32 - 0.42
0.24 - 0.32
0.18 - 0.24
0.13 - 0.18
0.10 - 0.13
< 0.10
Marine
Geospatial
Ecology
Lab


95°W
90°W
85°W
30°N
25°N
May
Abundance=2658
June
Abundance=1929
July
Abundance=2180
August
Abundance=2211
Animals / 100 km²
> 13
10 - 13
7.5 - 10
5.6 - 7.5
4.2 - 5.6
3.2 - 4.2
2.4 - 3.2
1.8 - 2.4
1.3 - 1.8
1.0 - 1.3
0.75 - 1.0
0.56 - 0.75
0.42 - 0.56
0.32 - 0.42
0.24 - 0.32
0.18 - 0.24
0.13 - 0.18
0.10 - 0.13
< 0.10
Marine
Geospatial
Ecology
Lab


95°W
90°W
85°W
30°N
25°N
September
Abundance=1986
October
Abundance=2000
November
Abundance=1929
December
Abundance=1905
Animals / 100 km²
> 13
10 - 13
7.5 - 10
5.6 - 7.5
4.2 - 5.6
3.2 - 4.2
2.4 - 3.2
1.8 - 2.4
1.3 - 1.8
1.0 - 1.3
0.75 - 1.0
0.56 - 0.75
0.42 - 0.56
0.32 - 0.42
0.24 - 0.32
0.18 - 0.24
0.13 - 0.18
0.10 - 0.13
< 0.10
Marine
Geospatial
Ecology
Lab


95°W
90°W
85°W
30°N
25°N
January
Abundance=1981
February
Abundance=1981
March
Abundance=1981
April
Abundance=1981
Animals / 100 km²
> 13
10 - 13
7.5 - 10
5.6 - 7.5
4.2 - 5.6
3.2 - 4.2
2.4 - 3.2
1.8 - 2.4
1.3 - 1.8
1.0 - 1.3
0.75 - 1.0
0.56 - 0.75
0.42 - 0.56
0.32 - 0.42
0.24 - 0.32
0.18 - 0.24
0.13 - 0.18
0.10 - 0.13
< 0.10
Marine
Geospatial
Ecology
Lab


95°W
90°W
85°W
30°N
25°N
May
Abundance=1981
June
Abundance=1981
July
Abundance=1981
August
Abundance=1981
Animals / 100 km²
> 13
10 - 13
7.5 - 10
5.6 - 7.5
4.2 - 5.6
3.2 - 4.2
2.4 - 3.2
1.8 - 2.4
1.3 - 1.8
1.0 - 1.3
0.75 - 1.0
0.56 - 0.75
0.42 - 0.56
0.32 - 0.42
0.24 - 0.32
0.18 - 0.24
0.13 - 0.18
0.10 - 0.13
< 0.10
Marine
Geospatial
Ecology
Lab


95°W
90°W
85°W
30°N
25°N
September
Abundance=1981
October
Abundance=1981
November
Abundance=1981
December
Abundance=1981
Animals / 100 km²
> 13
10 - 13
7.5 - 10
5.6 - 7.5
4.2 - 5.6
3.2 - 4.2
2.4 - 3.2
1.8 - 2.4
1.3 - 1.8
1.0 - 1.3
0.75 - 1.0
0.56 - 0.75
0.42 - 0.56
0.32 - 0.42
0.24 - 0.32
0.18 - 0.24
0.13 - 0.18
0.10 - 0.13
< 0.10
Marine
Geospatial
Ecology
Lab

Discussion

The model built with contemporaneous physical oceanographic predictors related to sea surface temperature and currents explained the most deviance of all models but predicted very high density at the southeastern most edge of the study area in February and March, contributing to a year-round abundance estimate that was 41% higher than the best models built with climatological predictors, with a high CV (0.69). This area was sparsely surveyed, and only in the months of May and June. The extreme prediction occurred over a week in late February 2003, in which total abundance was predicted to be over 500,000. Given this obviously spurious prediction, we discarded this model and selected the climatological model fitted to all segments as our best estimate of short-finned pilot whale distribution and abundance.

The selected model retained only three predictors, which were all physiographic–depth, slope, and distance to canyon–and predicted density was highest along the continental slope away from the canyons that occur in the northeast Gulf of Mexico. This result concerning canyons was somewhat unexpected given the importance of this predictor in our east coast model of pilot whales, but it is consistent with the pattern of sightings reported in the Gulf.

Because the model does not contain any time-varying predictors, there is no point in considering whether monthly or year-round average predictions be used. All predictions are the same, regardless of time of year, so the year-round prediction should be used.

NOAA's abundance estimates varied widely over the years (Table 15). Our estimate of 1981 was well within the confidence limits of NOAA's most recent estimate of 2415, made in 2009.

References

Alves F, Dinis A, Ribeiro C, Nicolau C, Kaufmann M, Fortuna CM, et al. (2013) Daytime dive characteristics from six short-finned pilot whales Globicephala macrorhynchus off Madeira Island. Arquipelago - Life and Marine Sciences 31: 1-8.

Barlow J (2006) Cetacean abundance in Hawaiian waters estimated from a summer/fall survey in 2002. Marine Mammal Science 22: 446-464.

Barlow J, Forney KA, Von Saunder A, Urban-Ramirez J (1997) A report of cetacean acoustic detection and dive interval studies (CADDIS) conducted in the southern Gulf of California. NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-250. 48 p.

Hansen LJ, Mullin KD, Roden CL (1995) Estimates of cetacean abundance in the northern Gulf of Mexico from vessel surveys. Southeast Fisheries Science Center, Miami Laboratory, Contribution No. MIA-94/95-25, 9 pp.

Heide-Jorgensen MP, Bloch D, Stefansson E, Mikkelsen B, Ofstad LH, Dietz R (2002) Diving behaviour of long-finned pilot whales Globicephala melas around the Faroe Islands. Wildlife Biology 8: 307-313.

Hooker SK, Fahlman A, Moore MJ, Soto NA de, Quiros YB de, Brubakk AO, et al. (2012) Deadly diving? Physiological and behavioural management of decompression stress in diving mammals. Proc R Soc B 279: 1041-1050.

Mullin KD (2007) Abundance of cetaceans in the oceanic Gulf of Mexico based on 2003-2004 ship surveys. 26 pp.

Mullin KD, Fulling GL (2004) Abundance of cetaceans in the oceanic northern Gulf of Mexico. Mar. Mamm. Sci. 20(4): 787-807.

Olson PA (2008) Pilot whales: Globicephala melas and G. muerorhynchus. In: Encyclopedia of Marine Mammals, 2nd ed. (Perrin WF, Wursig B, Thewissen JGM, eds.) Academic Press, San Diego, California. pp. 847-852.

Palka DL (2005b) Shipboard surveys in the northwest Atlantic: estimation of g(0). In: Proceedings of a Workshop on Estimation of g(0) in Line-Transect Surveys of Cetaceans (Thomsen F, Ugarte F, Evans PGH, eds.). European Cetacean Society's 18th Annual Conference; Kolmarden, Sweden; Mar. 28, 2004. pp. 32-37.

Palka DL (2006) Summer Abundance Estimates of Cetaceans in US North Atlantic Navy Operating Areas. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 06-03: 41 p.

Waring GT, Josephson E, Maze-Foley K, Rosel PE, eds. (2013) U.S. Atlantic and Gulf of Mexico Marine Mammal Stock Assessments – 2012. NOAA Tech Memo NMFS NE 223; 419 p.

Wells RS, Fougeres EM, Cooper AG, Stevens RO, Brodsky M, Lingenfelser R, et al. (2013) Movements and Dive Patterns of Short-Finned Pilot Whales (Globicephala macrorhynchus) Released from a Mass Stranding in the Florida Keys. Aquatic Mammals 39: 61-72.