Density Model for Pantropical Spotted Dolphin (*Stenella attenuata*) for the U.S. Gulf of Mexico: Supplementary Report

Duke University Marine Geospatial Ecology Lab*

Model Version 3.3 - 2015-10-06

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 Pantropical Spotted Dolphin (*Stenella attenuata*) for the U.S. Gulf of Mexico Version 3.3, 2015-10-06, 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	2014-11-13	Adjusted detection hierarchy and detection functions for aerial surveys. Updated documentation.
3	2015-01-09	Added three missing sightings 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	2015-10-06	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	178
SEFSC GulfCet I Aerial Surveys	1992-1994	50	257	47
SEFSC GulfCet II Aerial Surveys	1996-1998	22	124	47
SEFSC GulfSCAT 2007 Aerial Surveys	2007-2007	18	95	0
SEFSC Oceanic CetShip Surveys	1992-2001	49	3102	415
SEFSC Shelf CetShip Surveys	1994-2001	10	707	32
Total		195	5593	719

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	719
1998-2009	62	2679	339
% Lost	68	52	53

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.

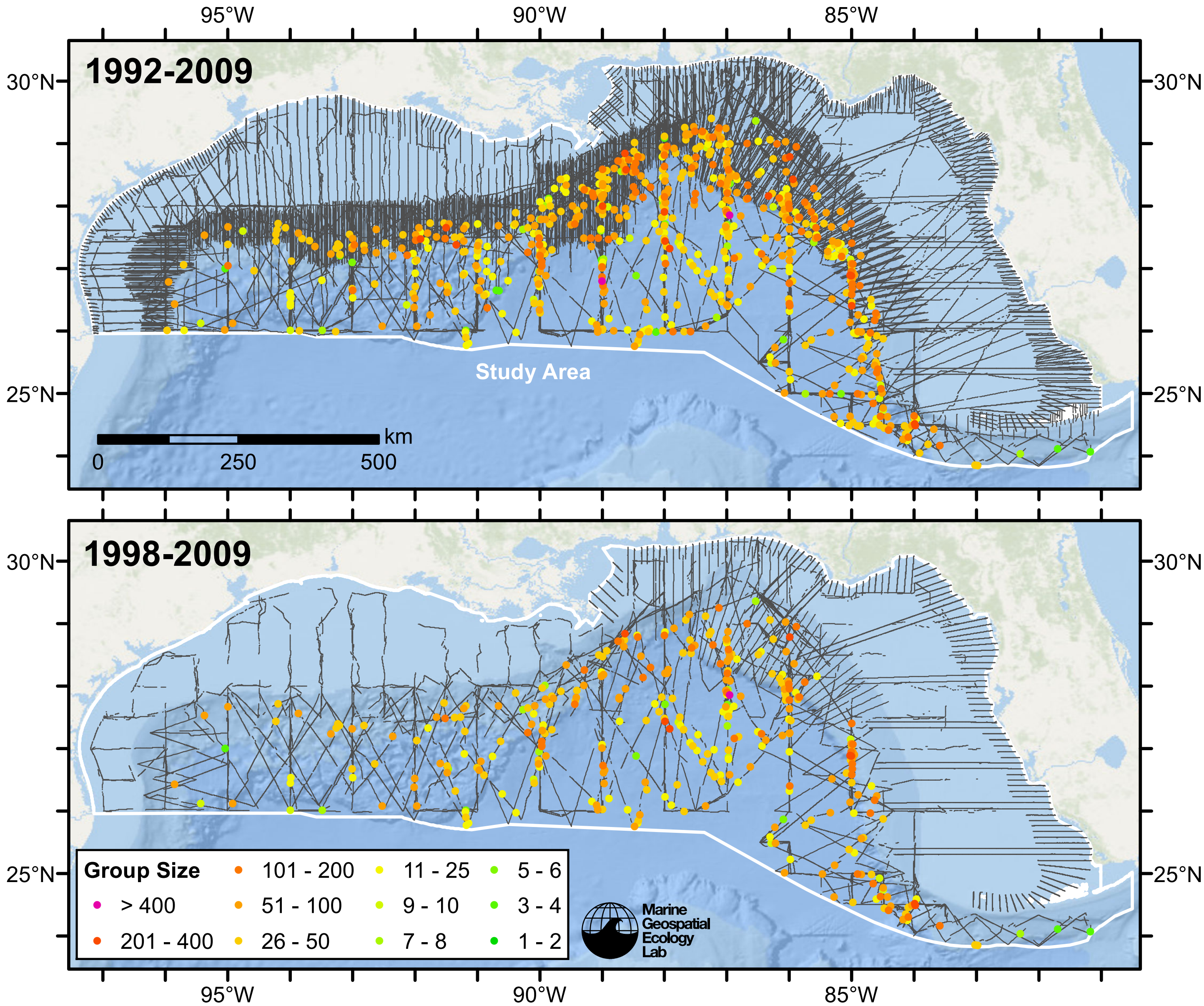


Figure 1: Pantropical spotted dolphin sightings and survey tracklines. The top map shows all surveys. The bottom map shows surveys performed in 1998 or later. the era in which remotely-sensed chlorophyll and derived productivity estimates are available. Models fitted to contemporaneous (day-of-sighting) 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.

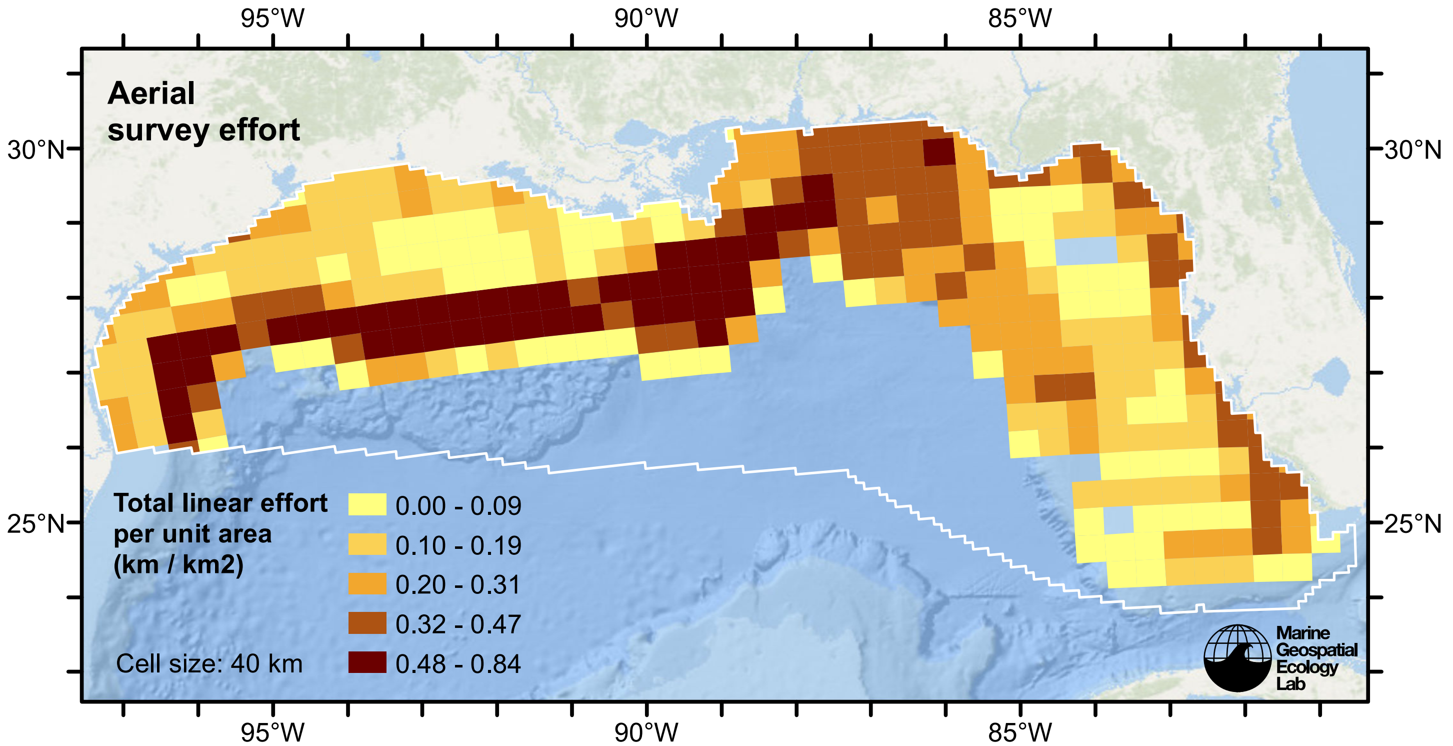


Figure 2: Aerial linear survey effort per unit area.

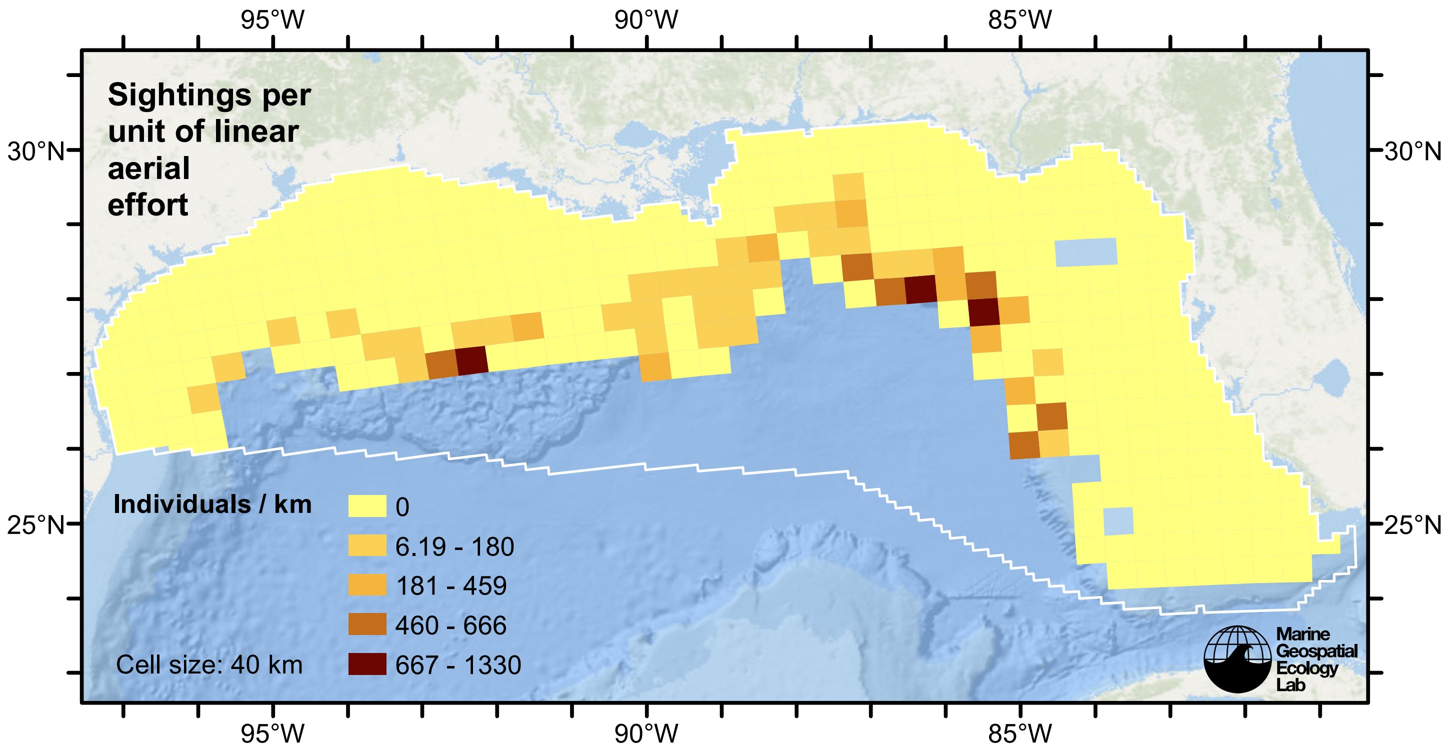


Figure 3: Pantropical spotted dolphin sightings per unit aerial linear survey effort.

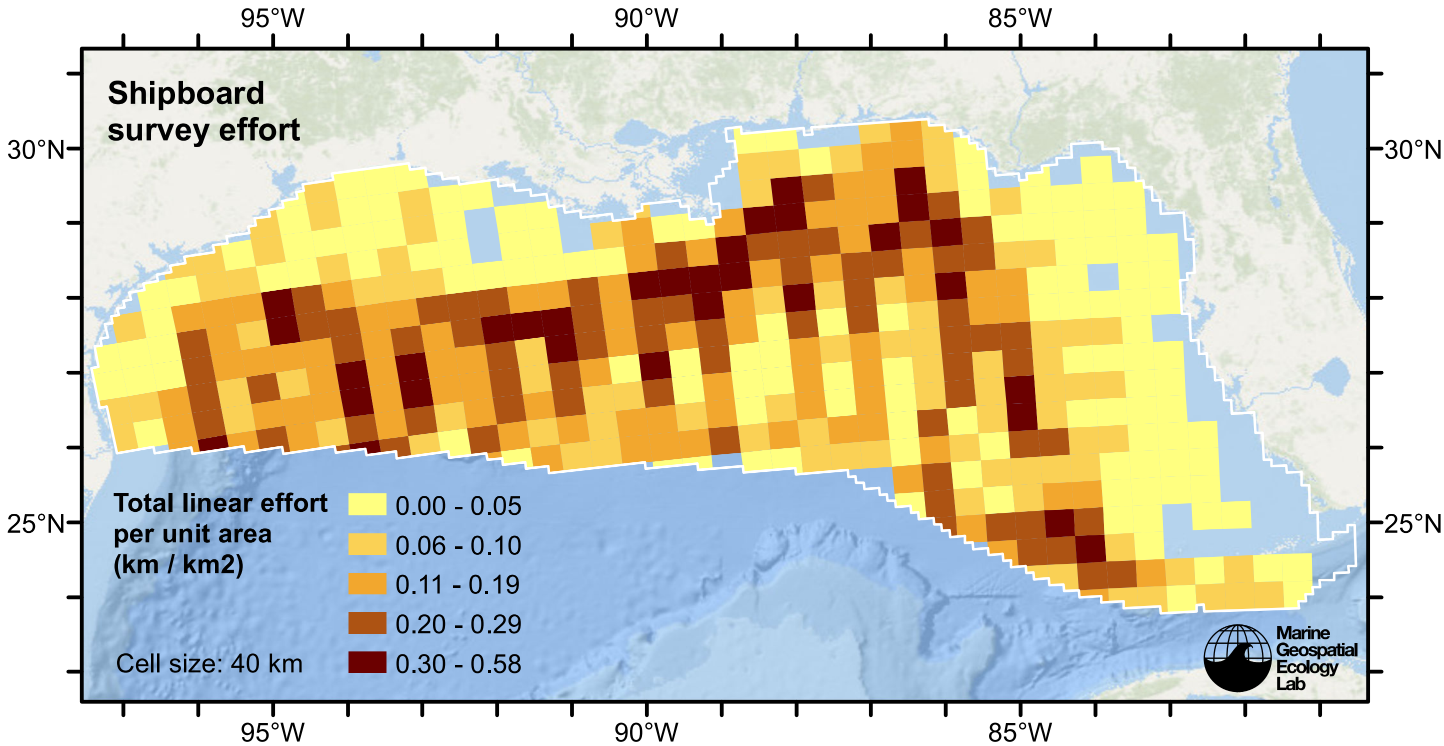


Figure 4: Shipboard linear survey effort per unit area.

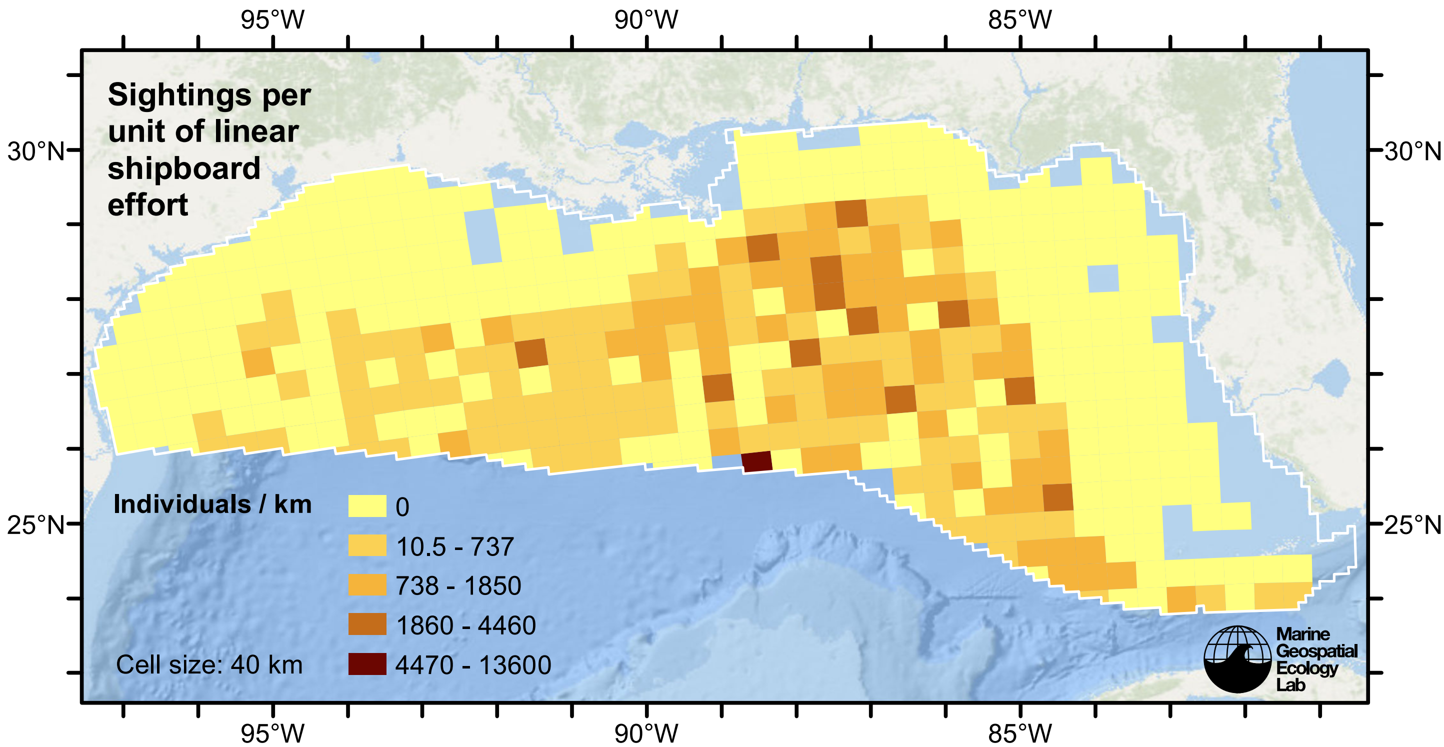


Figure 5: Pantropical spotted dolphin sightings per unit shipboard linear survey effort.

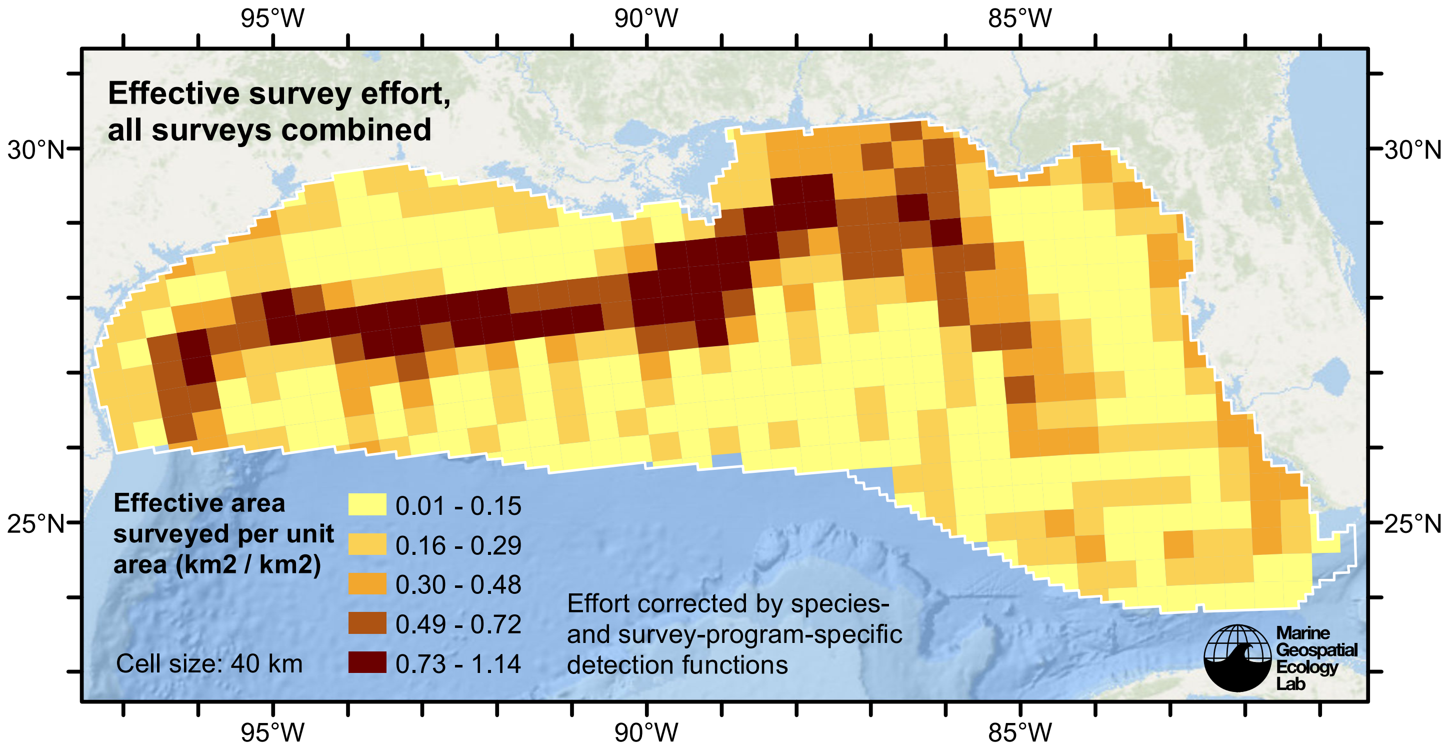


Figure 6: Effective survey effort per unit area, for all surveys combined. Here, effort is corrected by the species- and survey-program-specific detection functions used in fitting the density models.

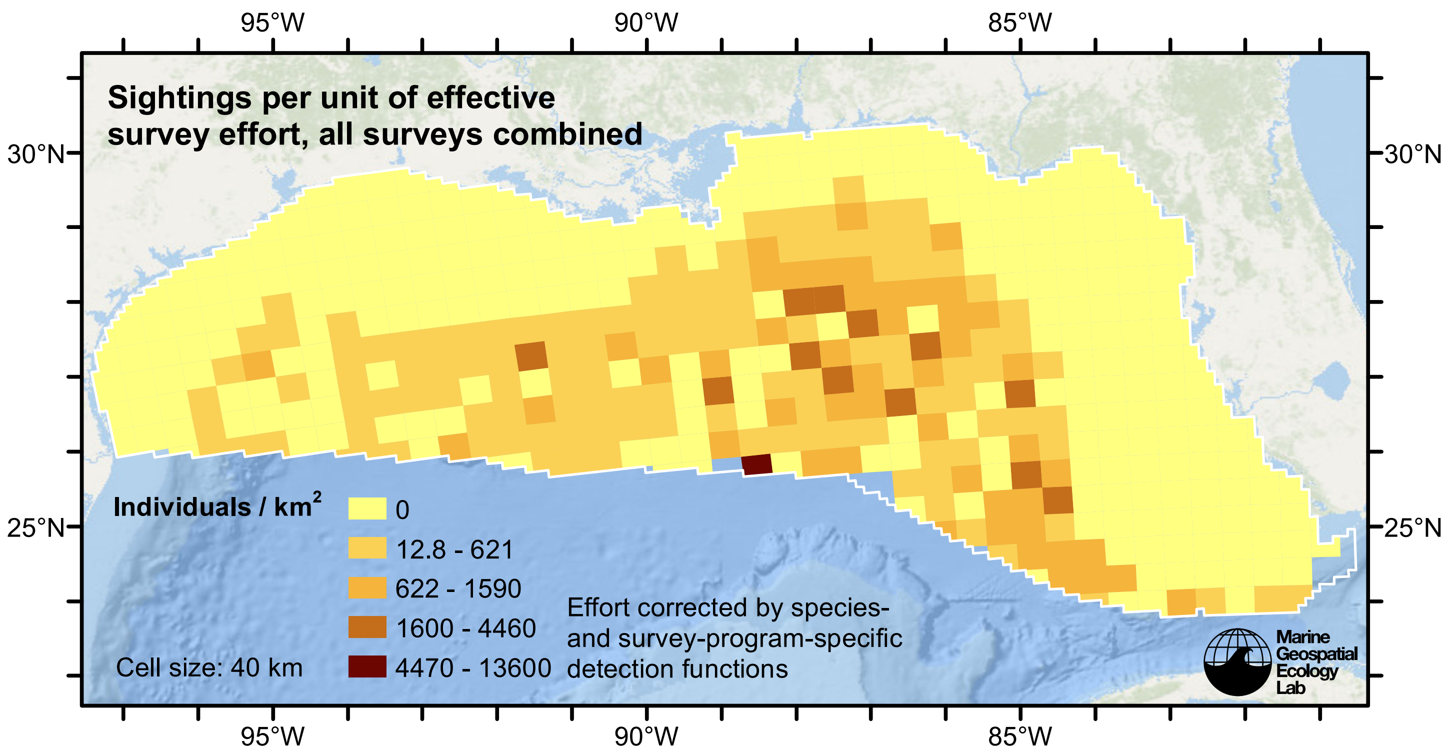


Figure 7: Pantropical spotted dolphin sightings per unit of effective survey effort, for all surveys combined. Here, effort is corrected by the species- and survey-program-specific detection functions used in fitting the density models.

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

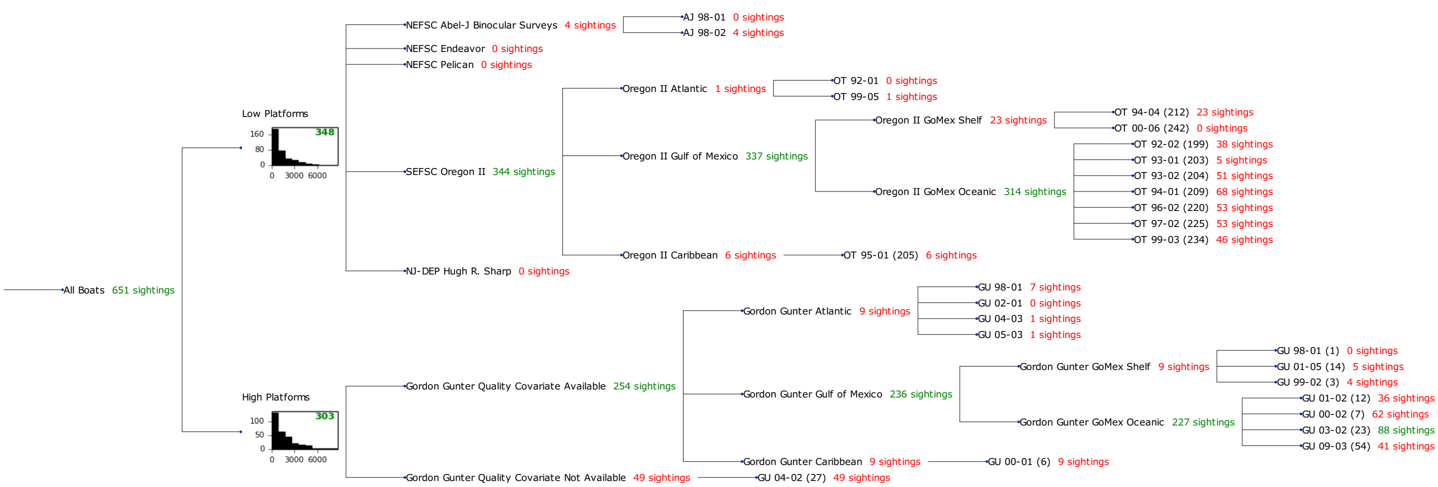


Figure 8: Detection hierarchy for shipboard surveys

Low Platforms

The sightings were right truncated at 5000m.

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			beaufort, size	Yes	0.00	1817
hr			size	Yes	0.58	1713
hr				Yes	53.23	728
hr			beaufort	Yes	53.57	742
hn	cos	2		Yes	78.00	1593
hn			beaufort, size	Yes	83.97	2182
hn			size	Yes	84.86	2174
hn	cos	3		Yes	87.93	1485
hn				Yes	123.93	2109
hn			beaufort	Yes	124.03	2108
hn	herm	4		Yes	125.00	2104
hr	poly	2		No		
hr	poly	4		No		

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

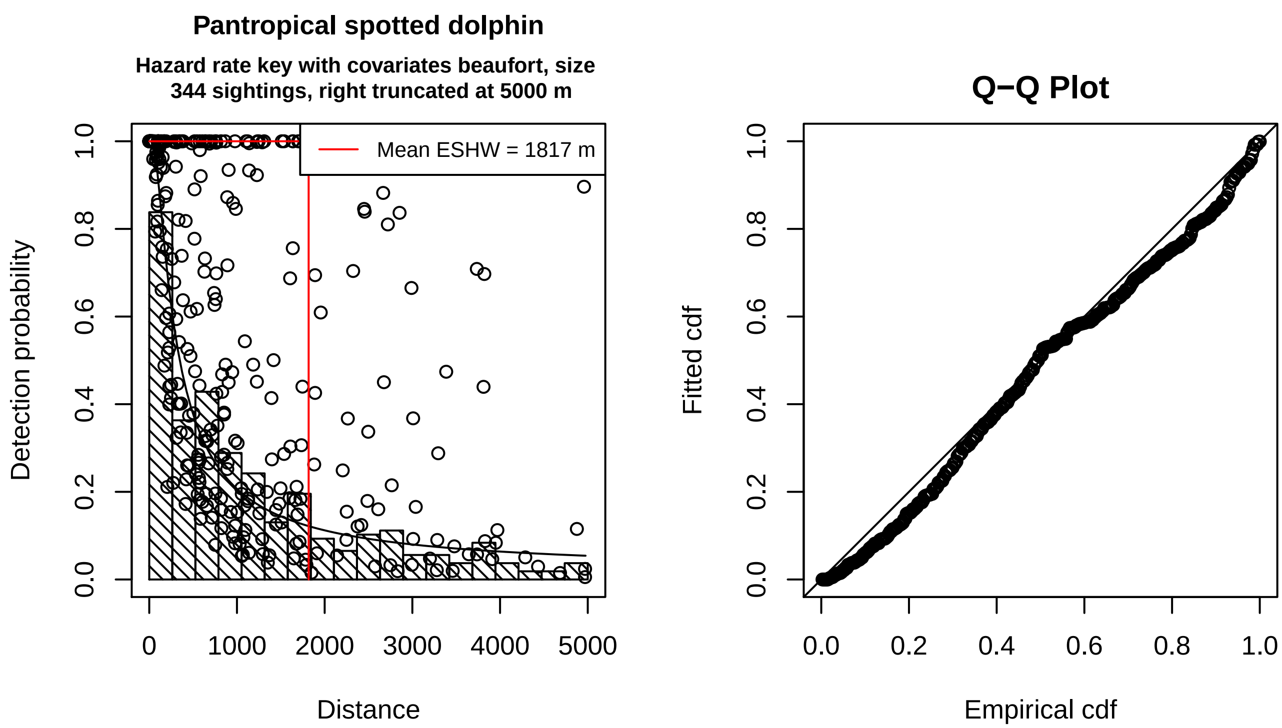


Figure 9: Detection function for Low Platforms that was selected for the density model

Statistical output for this detection function:

```
Summary for ds object
```

```
Number of observations :  344
Distance range         :  0  -  5000
AIC                    :  5477.79

Detection function:
 Hazard-rate key function

Detection function parameters
Scale Coefficients:
              estimate         se
(Intercept)  5.1390917 0.32755835
beaufort    -0.1542072 0.07613888
size         2.3802017 0.34732806

Shape parameters:
             estimate         se
(Intercept) 0.3576001 0.08845056

                        Estimate           SE        CV
Average p              0.1686068   0.02024788 0.1200893
N in covered region 2040.2498708 269.16384121 0.1319269
```

Additional diagnostic plots:

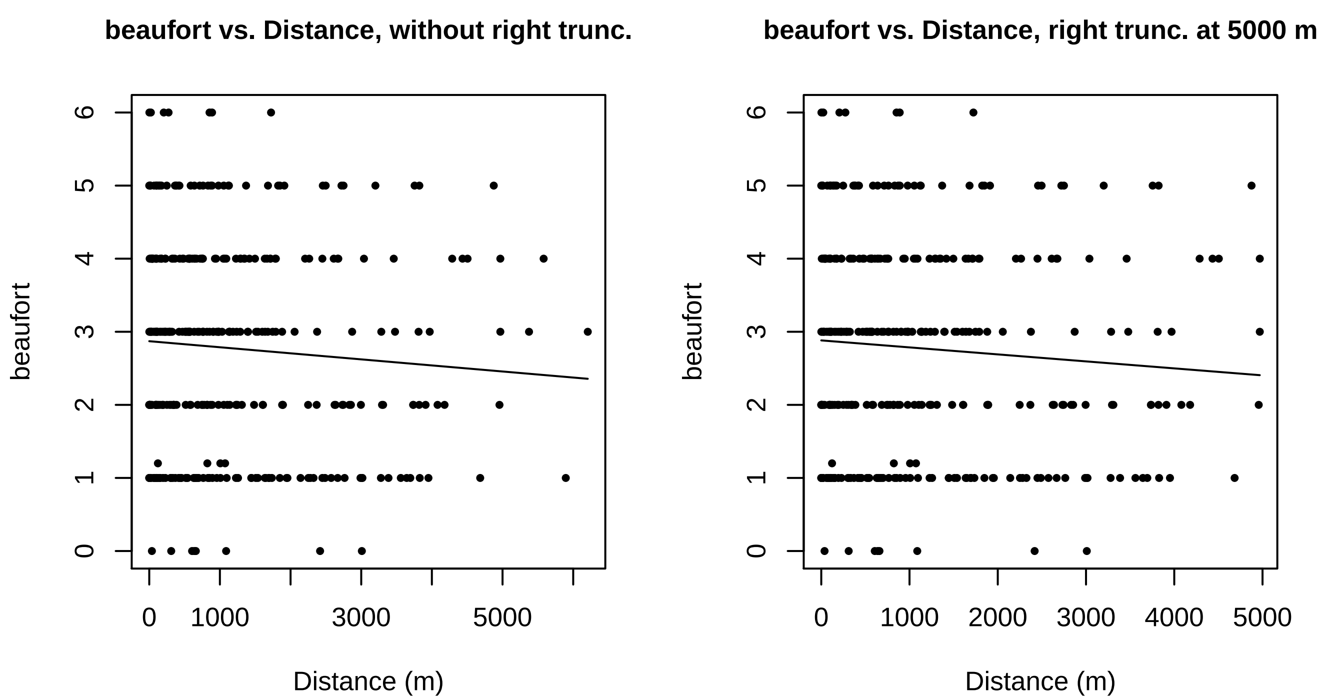


Figure 10: Scatterplots showing the relationship between Beaufort sea state and perpendicular sighting distance, for all sightings (left) and only those not right truncated (right). The line is a simple linear regression.

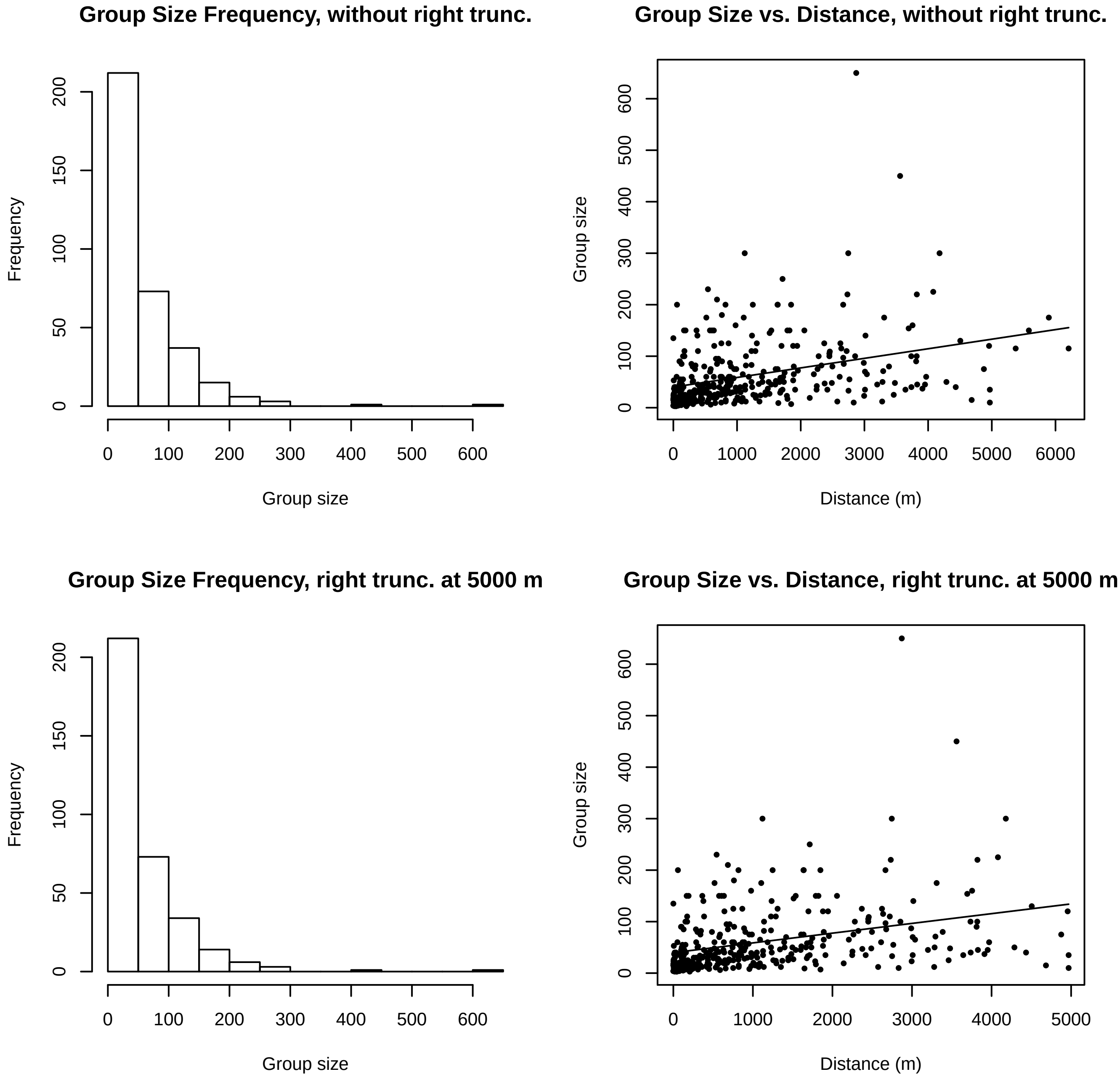


Figure 11: Histograms showing group size frequency and scatterplots showing the relationship between group size and perpendicular sighting distance, for all sightings (top row) and only those not right truncated (bottom row). In the scatterplot, the line is a simple linear regression.

High Platforms

The sightings were right truncated at 6000m.

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)
hr			beaufort, size	Yes	0.00	1404
hr			size	Yes	21.82	924
hr			beaufort	Yes	24.65	1013
hr	poly	4		Yes	42.14	827
hn			beaufort, size	Yes	42.29	2609
hr				Yes	46.56	726
hn			beaufort	Yes	69.58	2498
hn	cos	2		Yes	74.34	1977
hn	cos	3		Yes	80.32	1865
hn			size	Yes	88.70	2521
hn				Yes	99.38	2502
hn	herm	4		Yes	100.85	2496
hr	poly	2		No		

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

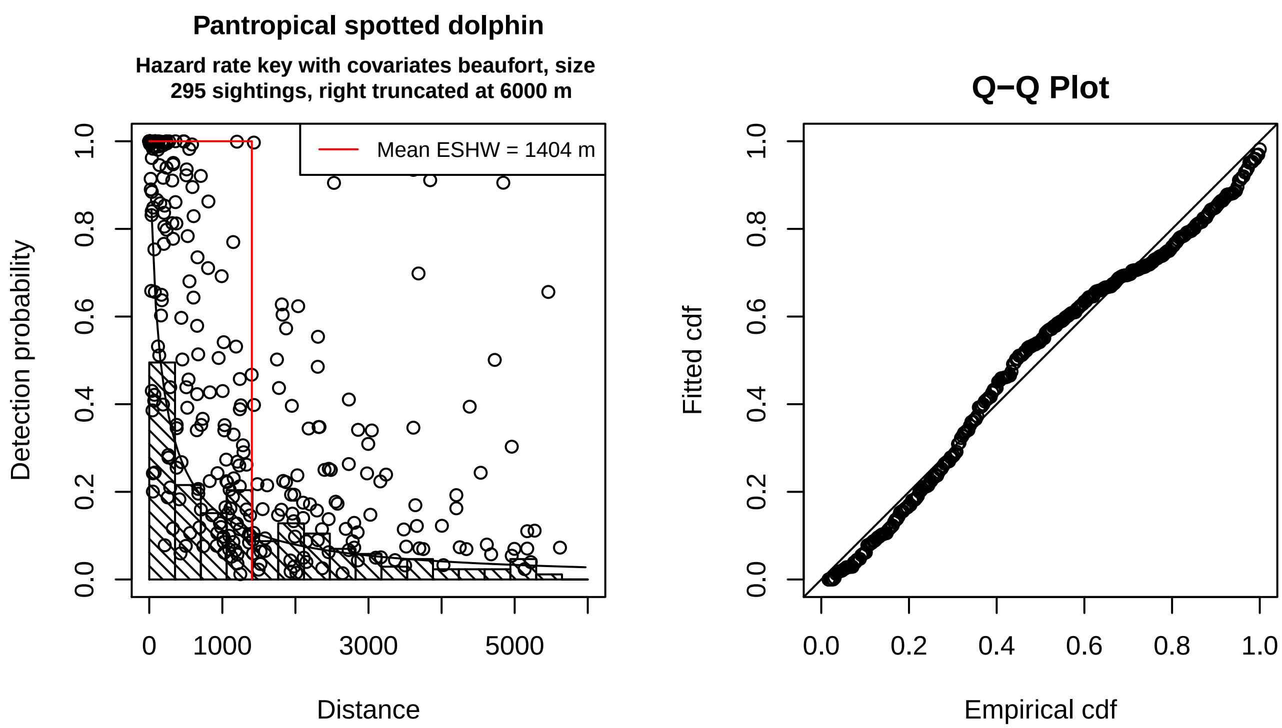


Figure 12: Detection function for High Platforms that was selected for the density model

Statistical output for this detection function:

```
Summary for ds object
```

```
Number of observations :  295
Distance range         :  0  -  6000
AIC                    :  4806.309

Detection function:
 Hazard-rate key function

Detection function parameters
Scale Coefficients:
              estimate        se
(Intercept)  6.3719618 0.4029306
beaufort    -0.6896037 0.1256594
size         0.9673811 0.1774448

Shape parameters:
             estimate         se
(Intercept) 0.1127456 0.08968933

                        Estimate           SE        CV
Average p              0.1010828   0.01972899 0.1951765
N in covered region 2918.3989143 595.17837580 0.2039400
```

Additional diagnostic plots:

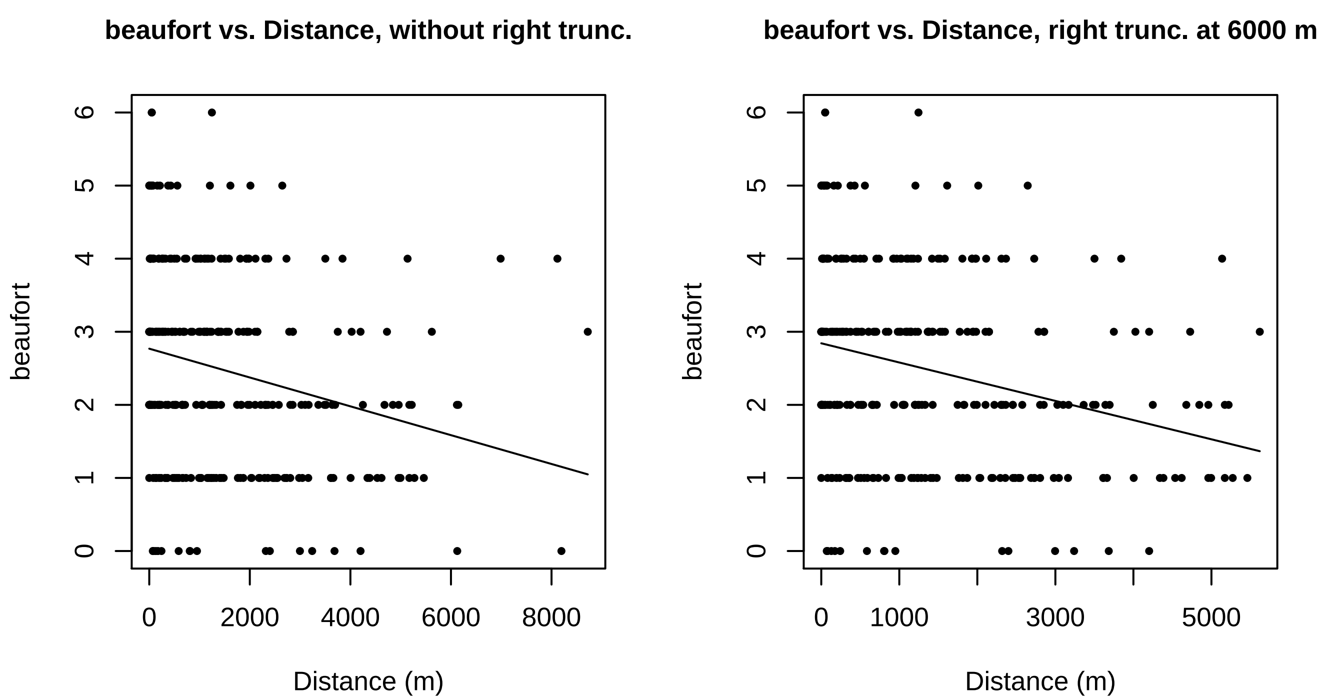


Figure 13: Scatterplots showing the relationship between Beaufort sea state and perpendicular sighting distance, for all sightings (left) and only those not right truncated (right). The line is a simple linear regression.

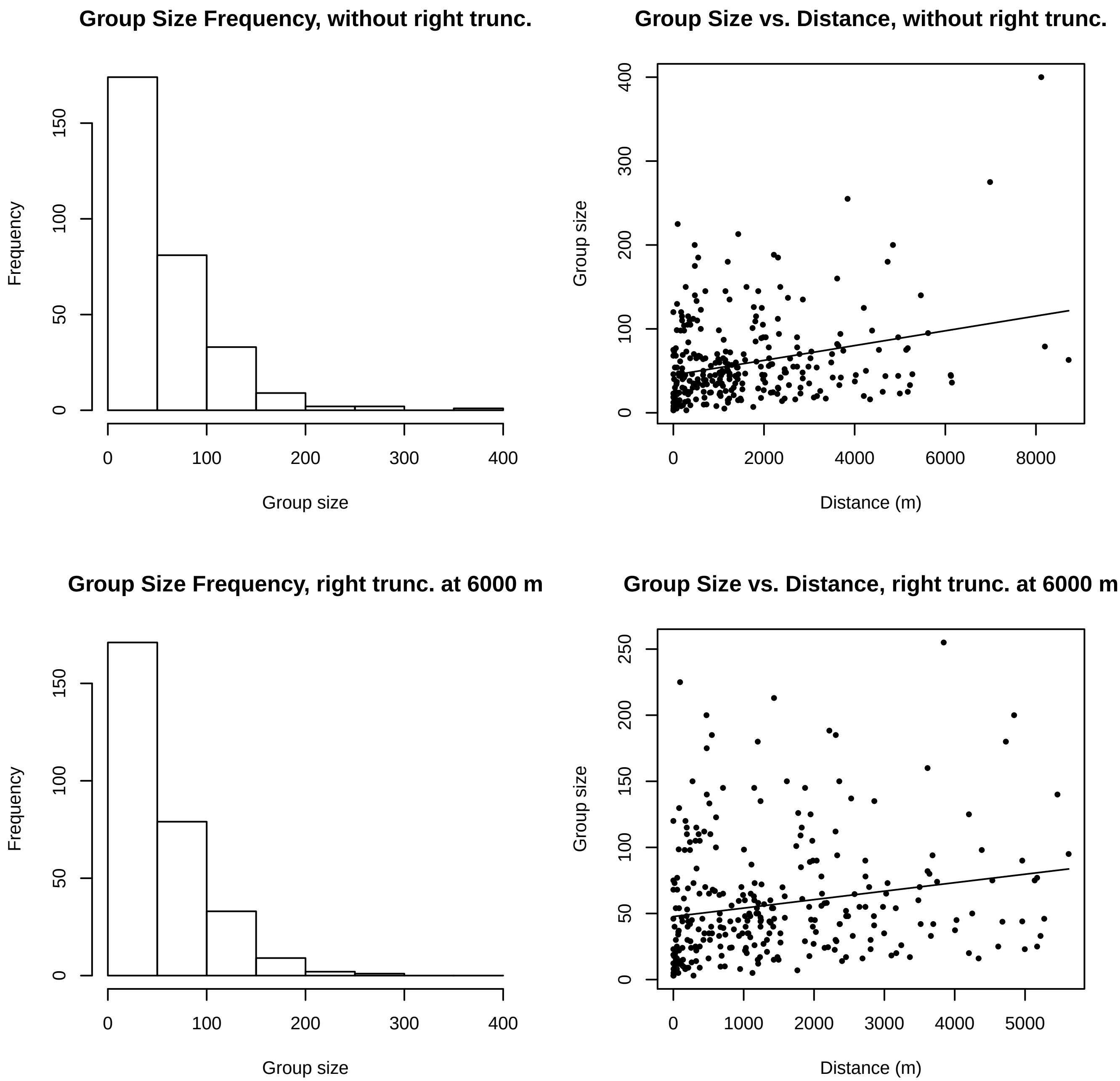


Figure 14: Histograms showing group size frequency and scatterplots showing the relationship between group size and perpendicular sighting distance, for all sightings (top row) and only those not right truncated (bottom row). In the scatterplot, the line is a simple linear regression.

Aerial Surveys

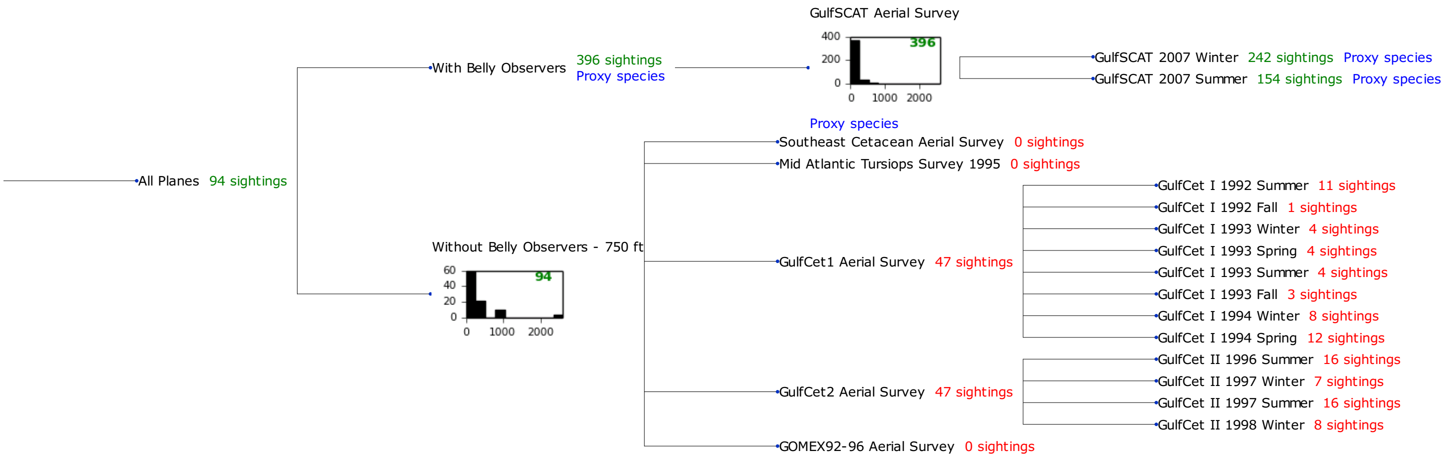


Figure 15: Detection hierarchy for aerial surveys

GulfSCAT Aerial Survey

Because this taxon was sighted too infrequently to fit a detection function to its sightings alone, we fit a detection function to the pooled sightings of several other species that we believed would exhibit similar detectability. These “proxy species” are listed below.

Reported By Observer	Common Name	n
Delphinus capensis	Long-beaked common dolphin	0
Delphinus delphis	Short-beaked common dolphin	0
Delphinus delphis/Lagenorhynchus acutus	Short-beaked common or Atlantic white-sided dolphin	0
Delphinus delphis/Stenella	Short-beaked common dolphin or Stenella spp.	0
Delphinus delphis/Stenella coeruleoalba	Short-beaked common or striped dolphin	0
Grampus griseus	Risso’s dolphin	0
Grampus griseus/Tursiops truncatus	Risso’s or Bottlenose dolphin	0
Lagenodelphis hosei	Fraser’s dolphin	0
Lagenorhynchus acutus	Atlantic white-sided dolphin	0
Lagenorhynchus albirostris	White-beaked dolphin	0
Lagenorhynchus albirostris/Lagenorhynchus acutus	White-beaked or white-sided dolphin	0
Stenella	Unidentified Stenella	0
Stenella attenuata	Pantropical spotted dolphin	0
Stenella attenuata/frontalis	Pantropical or Atlantic spotted dolphin	0
Stenella clymene	Clymene dolphin	0
Stenella coeruleoalba	Striped dolphin	0
Stenella frontalis	Atlantic spotted dolphin	15
Stenella frontalis/Tursiops truncatus	Atlantic spotted or Bottlenose dolphin	0
Stenella longirostris	Spinner dolphin	0
Steno bredanensis	Rough-toothed dolphin	0

Steno bredanensis/Tursiops truncatus	Bottlenose or rough-toothed dolphin	0
Tursiops truncatus	Bottlenose dolphin	381
Total		396

Table 8: Proxy species used to fit detection functions for GulfSCAT Aerial Survey. The number of sightings, n, is before truncation.

The sightings were right truncated at 400m.

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

Key	Adjustment	Order	Covariates	Succeeded	Δ AIC	Mean ESHW (m)
hn	herm	4		Yes	0.00	218
hn	cos	2		Yes	0.09	221
hn				Yes	0.90	199
hn			size	Yes	2.21	199
hn	cos	3		Yes	2.37	209
hr	poly	2		Yes	2.39	218
hr	poly	4		Yes	2.47	223
hr				Yes	4.46	230
hr			size	Yes	5.04	232
hn			beaufort	No		
hr			beaufort	No		
hn			quality	No		
hr			quality	No		
hn			beaufort, quality	No		
hr			beaufort, quality	No		
hn			beaufort, size	No		
hr			beaufort, size	No		
hn			quality, size	No		
hr			quality, size	No		
hn			beaufort, quality, size	No		
hr			beaufort, quality, size	No		

Table 10: Candidate detection functions for GulfSCAT Aerial Survey. The first one listed was selected for the density model.

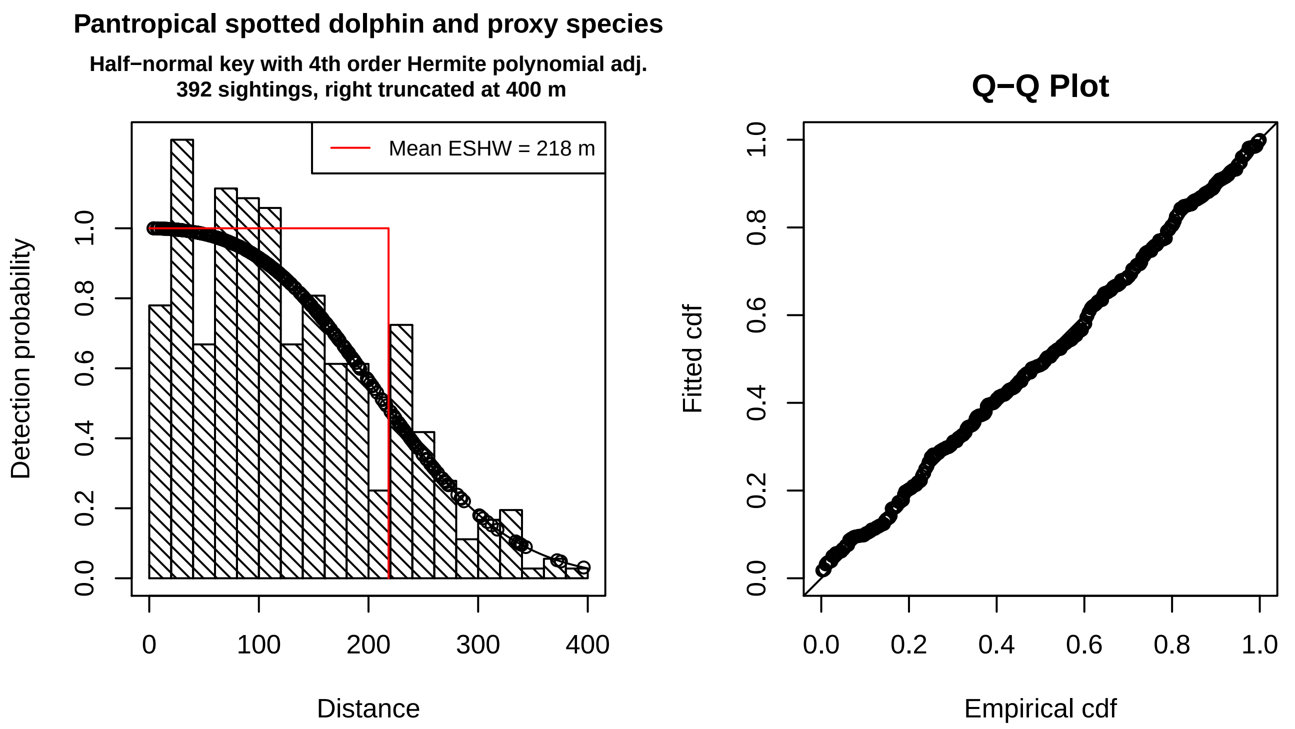


Figure 16: Detection function for GulfSCAT Aerial Survey that was selected for the density model

Statistical output for this detection function:

```
Summary for ds object
Number of observations :  392
Distance range         :  0  -  400
AIC                    :  4505.917

Detection function:
 Half-normal key function with Hermite polynomial adjustment term of order 4

Detection function parameters
Scale Coefficients:
            estimate        se
(Intercept) 4.855658 0.0741652

Adjustment term parameter(s):
                 estimate         se
herm, order 4 -0.04125642 0.01270664


Monotonicity constraints were enforced.
                       Estimate          SE         CV
Average p             0.5457537  0.04201324 0.07698205
N in covered region 718.2727866 60.45889329 0.08417261

Monotonicity constraints were enforced.
```

Additional diagnostic plots:

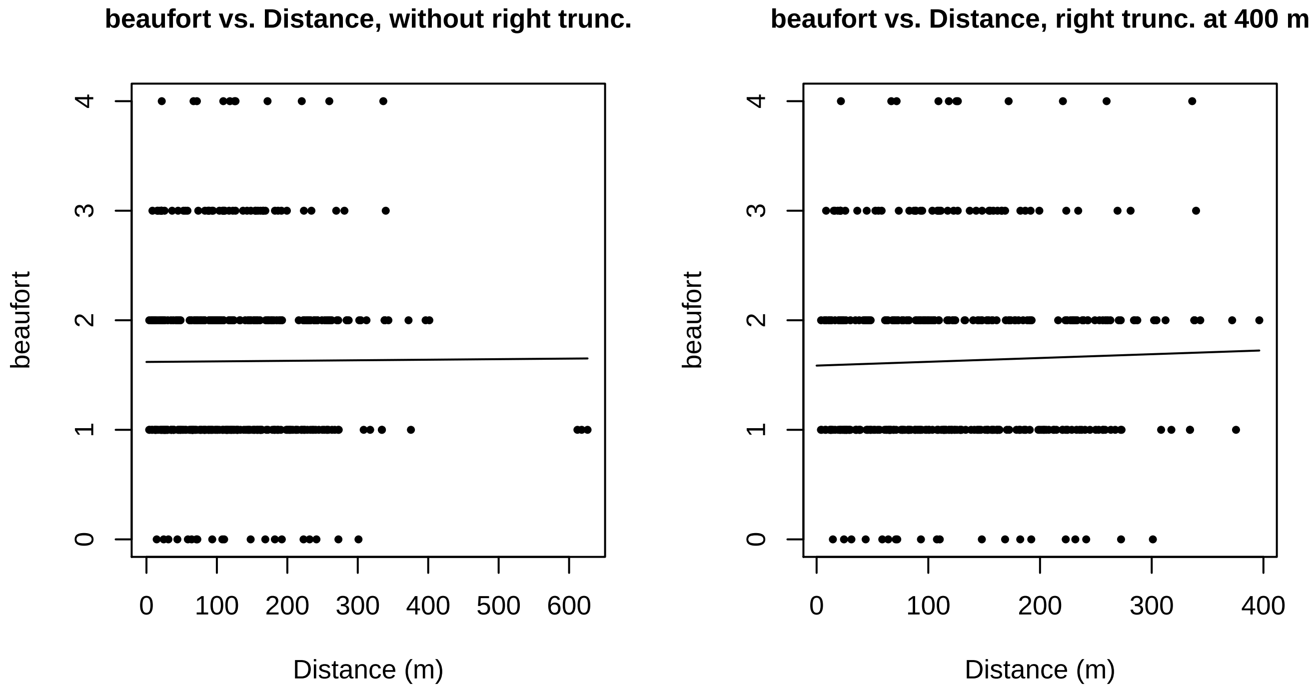


Figure 17: Scatterplots showing the relationship between Beaufort sea state and perpendicular sighting distance, for all sightings (left) and only those not right truncated (right). The line is a simple linear regression.

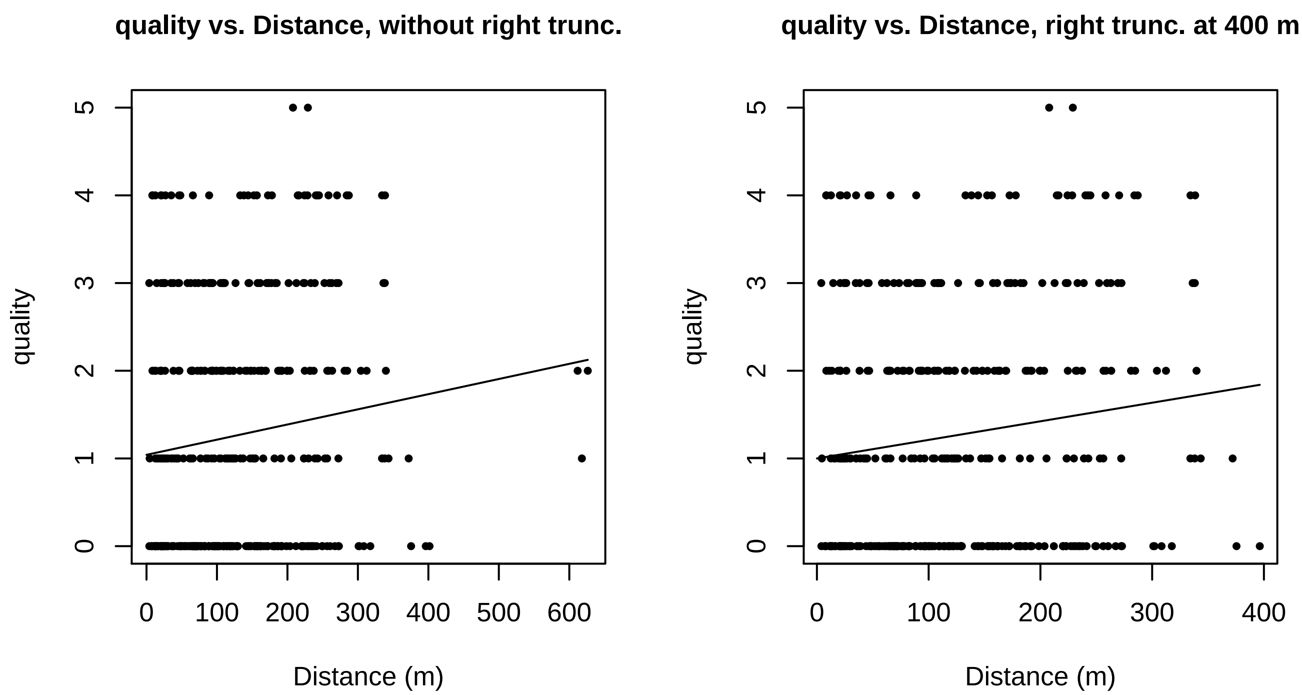


Figure 18: Scatterplots showing the relationship between the survey-specific index of the quality of observation conditions and perpendicular sighting distance, for all sightings (left) and only those not right truncated (right). Low values of the quality index correspond to better observation conditions. The line is a simple linear regression.

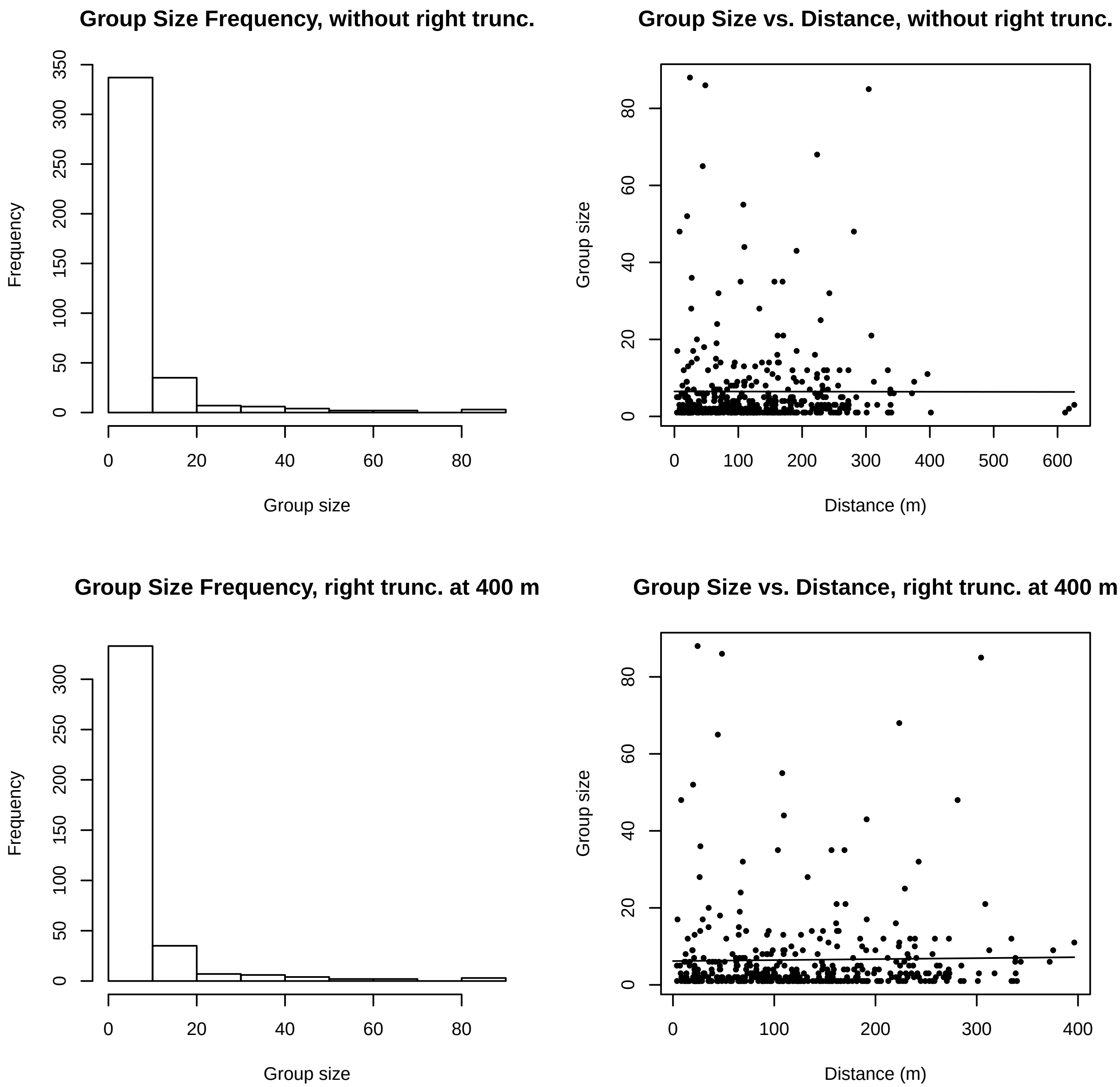


Figure 19: Histograms showing group size frequency and scatterplots showing the relationship between group size and perpendicular sighting distance, for all sightings (top row) and only those not right truncated (bottom row). In the scatterplot, the line is a simple linear regression.

Without Belly Observers - 750 ft

The sightings were right truncated at 900m.

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

Key	Adjustment	Order	Covariates	Succeeded	Δ AIC	Mean ESHW (m)
hr				Yes	0.00	385
hr			size	Yes	0.66	398
hn	cos	2		Yes	1.33	340
hr	poly	2		Yes	2.00	385
hr	poly	4		Yes	2.62	358
hn			size	Yes	10.29	486
hn			beaufort, size	Yes	11.10	501
hn				Yes	11.72	481
hn			quality, size	Yes	12.19	487
hn	cos	3		Yes	12.60	506
hn			beaufort, quality, size	Yes	12.84	502
hn	herm	4		Yes	13.03	479
hn			beaufort	Yes	13.51	481
hr			beaufort	No		
hn			quality	No		
hr			quality	No		
hn			beaufort, quality	No		
hr			beaufort, quality	No		
hr			beaufort, size	No		
hr			quality, size	No		
hr			beaufort, quality, size	No		

Table 12: Candidate detection functions for Without Belly Observers - 750 ft. The first one listed was selected for the density model.

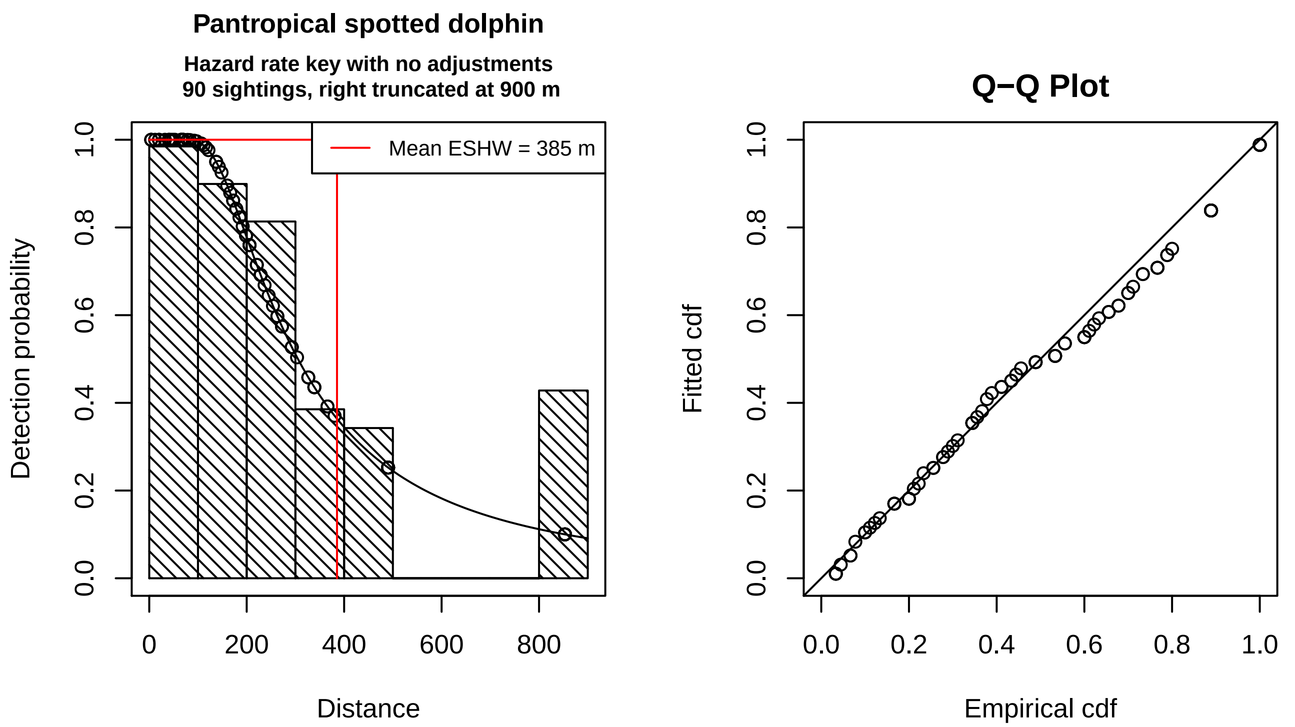


Figure 20: Detection function for Without Belly Observers - 750 ft that was selected for the density model

Statistical output for this detection function:

```
Summary for ds object
Number of observations :  90
Distance range         :  0  -  900
AIC                    :  1179.815

Detection function:
 Hazard-rate key function

Detection function parameters
Scale Coefficients:
            estimate        se
(Intercept) 5.520789 0.2463792

Shape parameters:
            estimate        se
(Intercept) 0.605147 0.2145303

                       Estimate          SE        CV
Average p             0.4281779  0.06414965 0.1498201
N in covered region 210.1929997 35.67070764 0.1697045
```

Additional diagnostic plots:

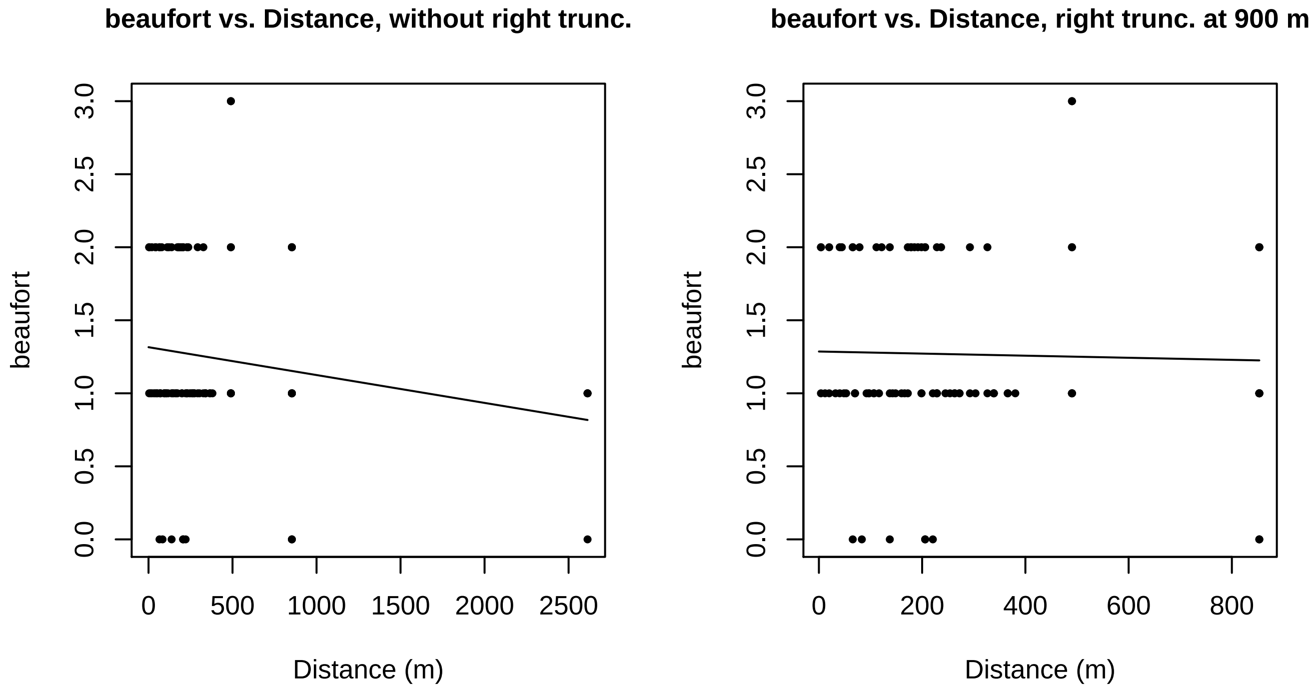


Figure 21: Scatterplots showing the relationship between Beaufort sea state and perpendicular sighting distance, for all sightings (left) and only those not right truncated (right). The line is a simple linear regression.

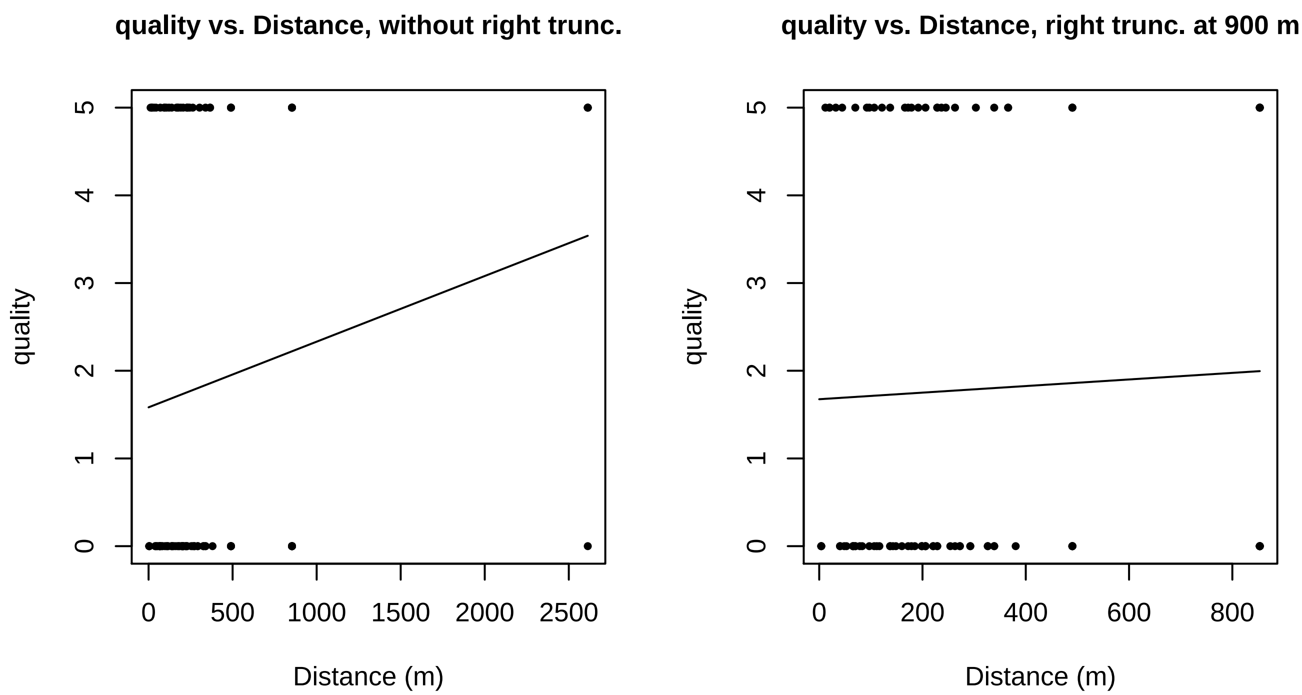


Figure 22: Scatterplots showing the relationship between the survey-specific index of the quality of observation conditions and perpendicular sighting distance, for all sightings (left) and only those not right truncated (right). Low values of the quality index correspond to better observation conditions. The line is a simple linear regression.

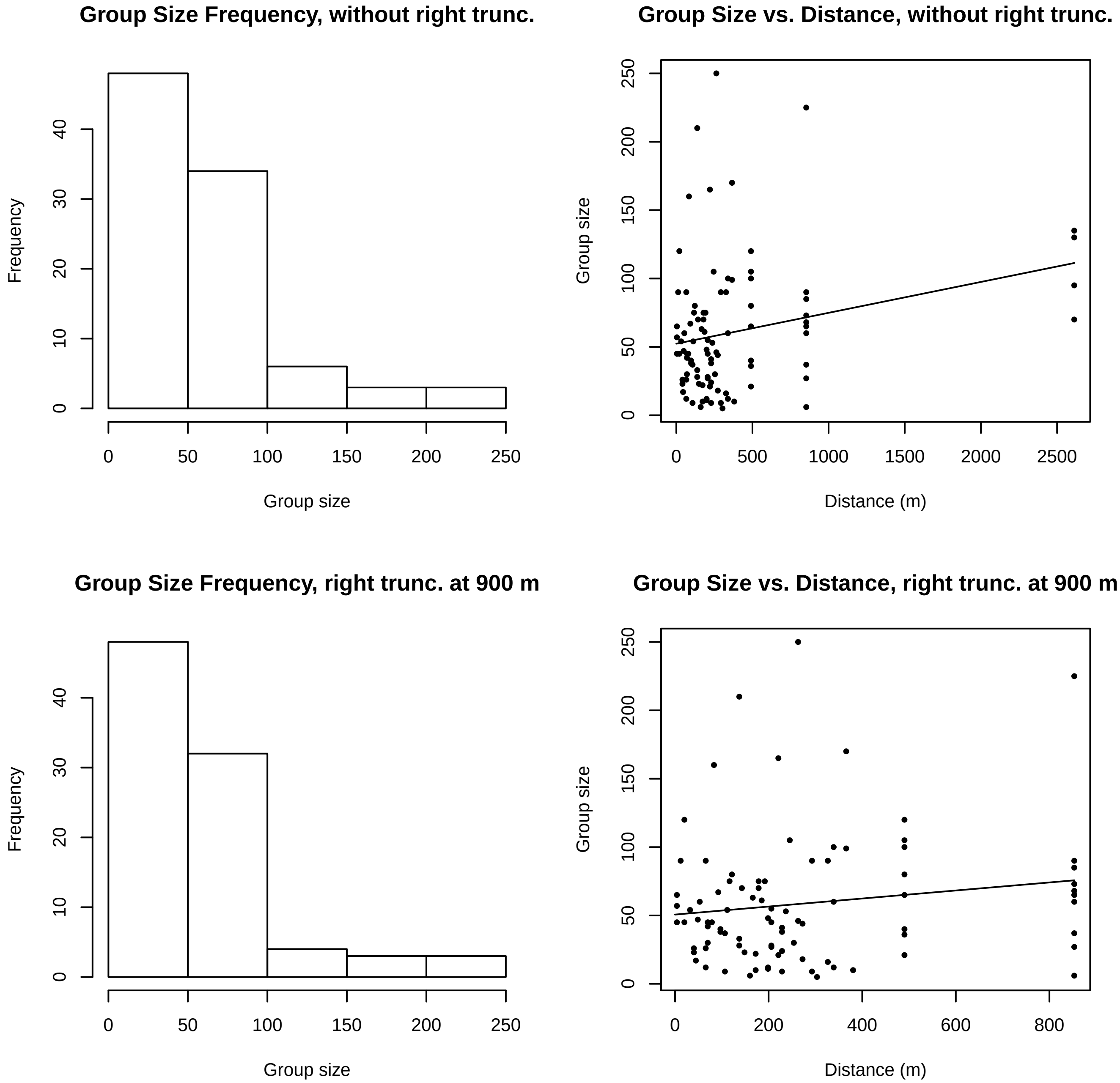


Figure 23: Histograms showing group size frequency and scatterplots showing the relationship between group size and perpendicular sighting distance, for all sightings (top row) and only those not right truncated (bottom row). In the scatterplot, the line is a simple linear regression.

g(0) Estimates

Platform	Surveys	Group Size	*g(0)*	Biases Addressed	Source
Shipboard	All	1-20	0.856	Perception	Barlow and Forney (2007)
		>20	0.970	Perception	Barlow and Forney (2007)
Aerial	All	1-5	0.43	Both	Palka (2006)
		>5	0.960	Both	Carretta et al. (2000)

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

No g(0) estimates were published for any of the shipboard surveys available to us from this region. Instead, we utilized Barlow and Forney's (2007) estimates for delphinids, produced from several years of dual-team surveys that used bigeye binoculars and similar protocols to the surveys in our study. This study provided separate estimates for small and large groups, but pooled sightings of several species together to provide a generic estimate for all delphinids, due to sample-size limitations. To our knowledge, there is no species-specific shipboard g(0) estimate that treats small and large groups separately, so we believe Barlow and Forney (2007) provide the best general-purpose alternative. Their estimate accounted for perception bias but not availability bias; dive times for dolphins are short enough that availability bias is not expected to be significant for dolphins observed from shipboard surveys.

For aerial surveys, we were unable to locate species-specific g(0) estimates in the literature. For small groups, defined here as 1-5 individuals, we used Palka's (2006) estimate of g(0) for groups of 1-5 small cetaceans, estimated from two years of aerial surveys using the Hiby (1999) circle-back method. This estimate accounted for both availability and perception bias, but pooled sightings of several species together to provide a generic estimate for all delphinids, due to sample-size limitations. For large groups, defined here as greater than 5 individuals, Palka (2006) assumed that g(0) was 1. When we discussed this with NOAA SWFSC reviewers, they agreed that it was safe to assume that the availability bias component of g(0) was 1 but insisted that perception bias should be slightly less than 1, because it was possible to miss large groups. We agreed to take a conservative approach and obtained our g(0) for large groups from Carretta et al. (2000), who estimated g(0) for both small and large groups of delphinids. We used Carretta et al.'s g(0) estimate for groups of 1-25 individuals (0.960), rather than their larger one for more than 25 individuals (0.994), to account for the fact that we were using Palka's definition of large groups as those with more than 5 individuals.

Density Models

The pantropical spotted dolphin occurs worldwide in tropical and some sub-tropical oceans (Waring et al. 2013), is the most abundant oceanic (> 200m depth) delphinid in the Gulf of Mexico, and occurs almost exclusively in oceanic waters (Jefferson and Schiro 1997). All of the sightings reported by Gulf of Mexico surveys utilized in our analysis occurred off the continental shelf. A prior habitat analysis of a subset of these suggested that pantropical spotted dolphins were distributed over the lower continental slope and deep Gulf (> 1000m depth), but, unexpectedly, statistical analysis did not reveal that the species' distribution was significantly different than a uniform distribution with respect to any of the variables that were tested except depth (Baumgartner et al. 2001). Consistent with the reported absence of pantropical spotted dolphins from the Gulf of Mexico shelf, we fitted our model to the effort that occurred in off-shelf waters, defined here as those deeper than the 100m isobath.

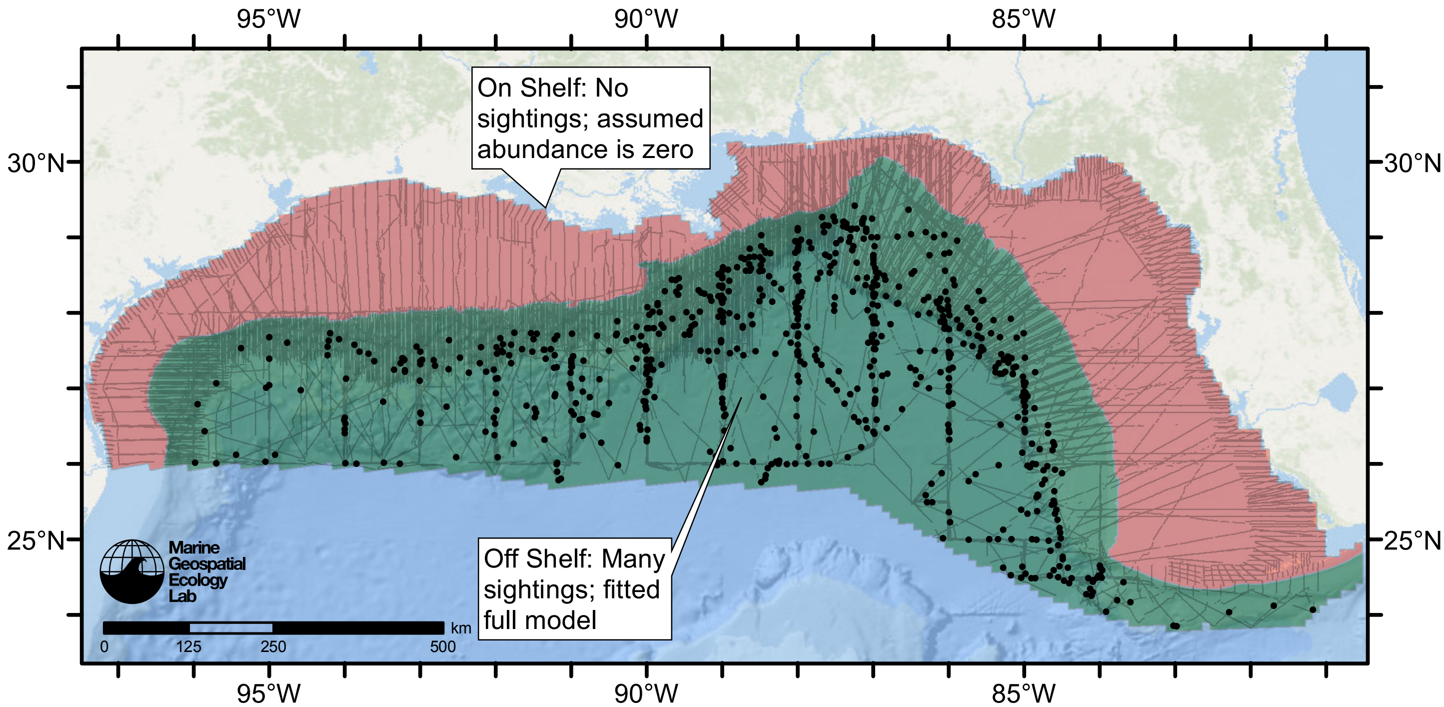


Figure 24: Pantropical spotted dolphin density model schematic. All on-effort sightings are shown, including those that were truncated when detection functions were fitted.

Climatological Model

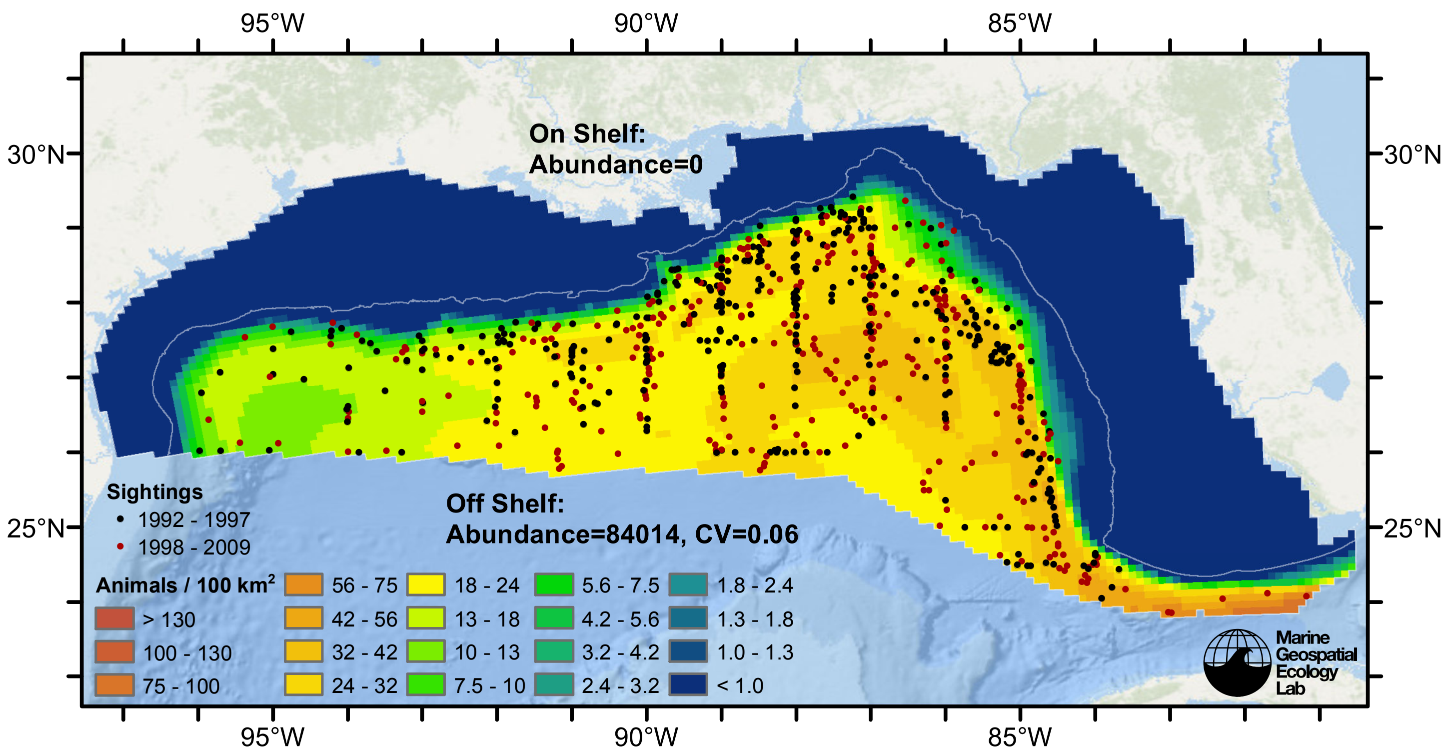


Figure 25: Pantropical spotted dolphin density predicted by the climatological model that explained the most deviance. Pixels are 10x10 km. The legend gives the estimated individuals per pixel; breaks are logarithmic. Abundance for each region was computed by summing the density cells occuring in that region.

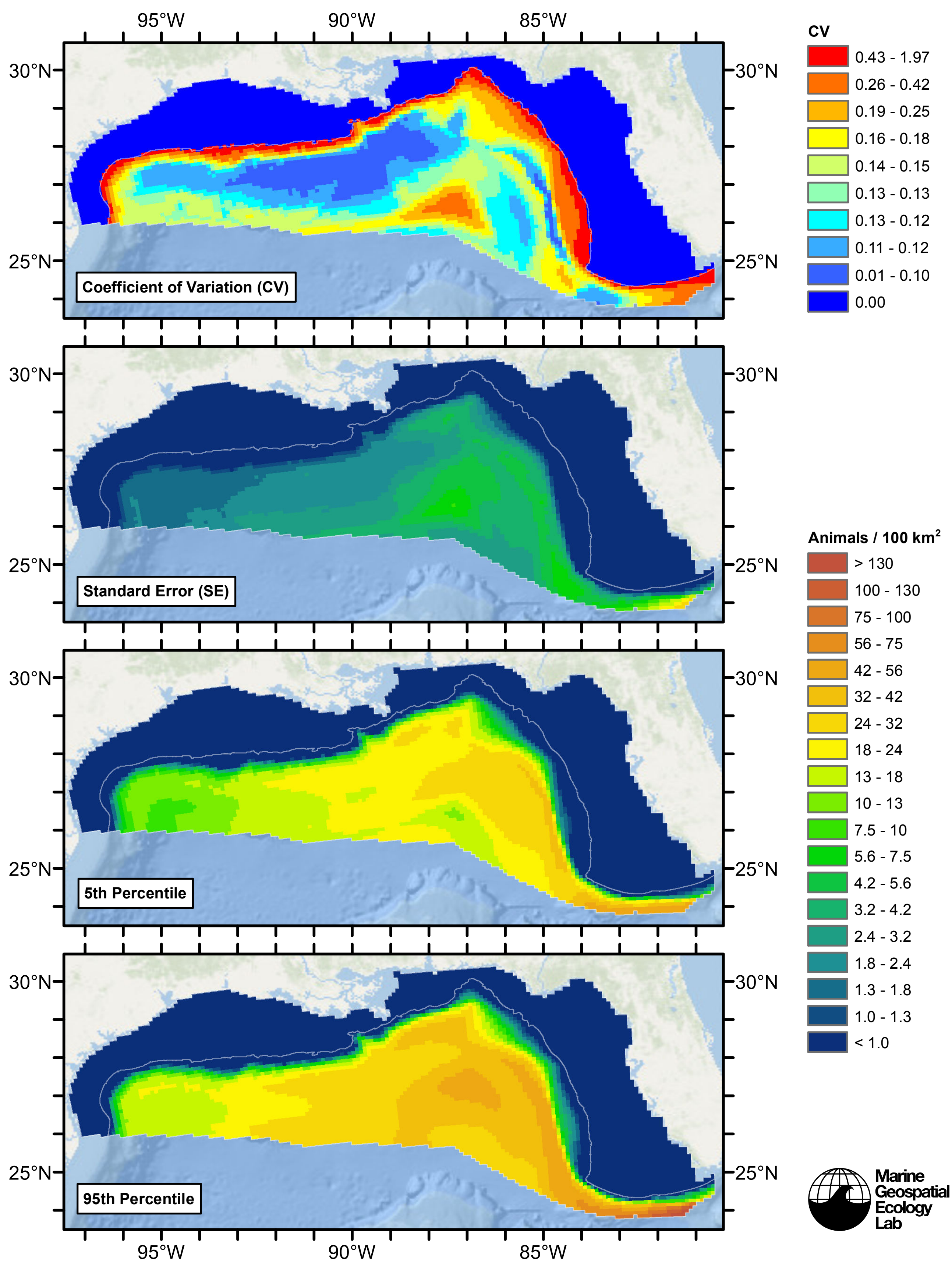


Figure 26: Estimated uncertainty for the climatological model that explained the most deviance. These estimates only incorporate the statistical uncertainty estimated for the spatial model (by the R mgcv package). They do not incorporate uncertainty in the detection functions, g(0) estimates, predictor variables, and so on.

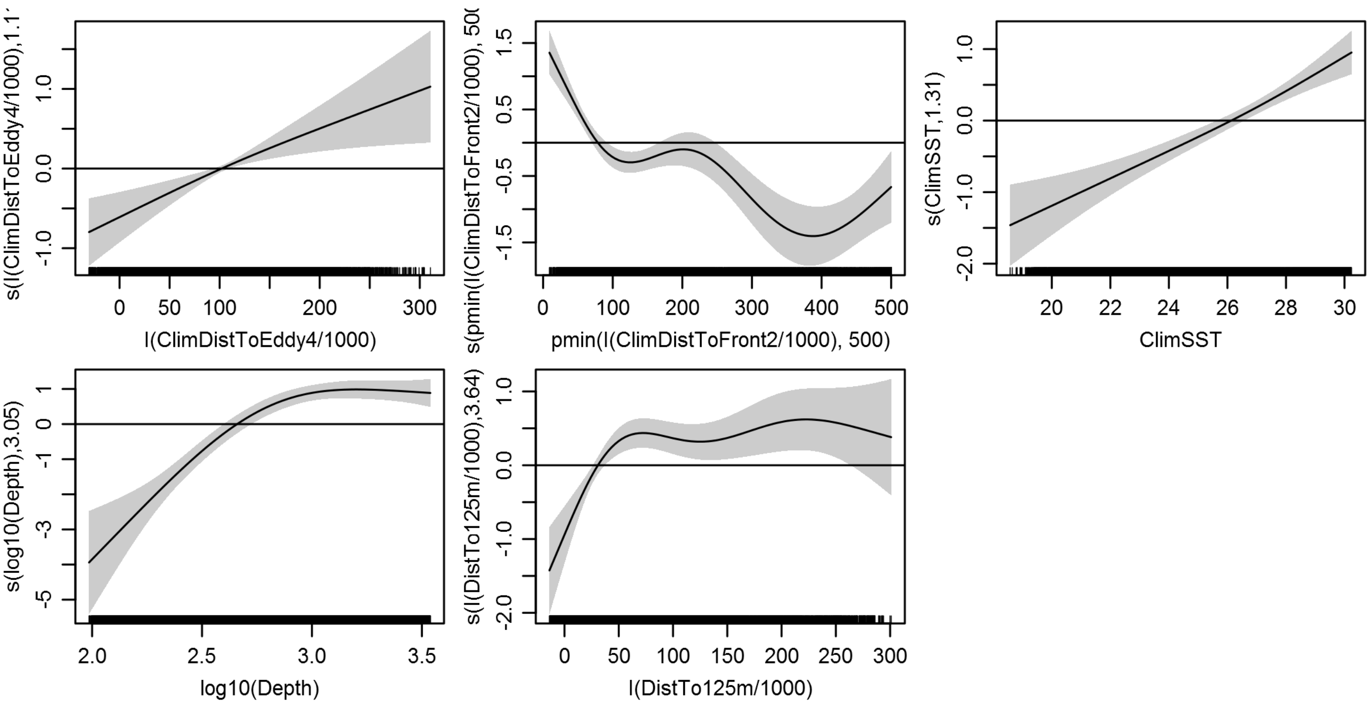
Off Shelf

Statistical output

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


Family: Tweedie(p=1.263)
Link function: log

Formula:
abundance ~ offset(log(area_km2)) + s(log10(Depth), bs = "ts",
    k = 5) + s(I(DistTo125m/1000), bs = "ts", k = 5) + s(ClimSST,
    bs = "ts", k = 5) + s(pmin(I(ClimDistToFront2/1000), 500),
    bs = "ts", k = 5) + s(I(ClimDistToEddy4/1000), bs = "ts",
    k = 5)

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -2.6482     0.1006  -26.32   <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))                       3.046      4 18.728  < 2e-16 ***
s(I(DistTo125m/1000))                 3.639      4  5.955 2.48e-05 ***
s(ClimSST)                            1.310      4 12.195 8.17e-13 ***
s(pmin(I(ClimDistToFront2/1000), 500)) 3.809      4 19.016  < 2e-16 ***
s(I(ClimDistToEddy4/1000))            1.107      4  3.791 6.31e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) =   0.03   Deviance explained = 21.7%
-REML = 6492.1  Scale est. = 154.41    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 [-0.0003520776,0.0008564004]
(score 6492.072 & scale 154.4125).
Hessian positive definite, eigenvalue range [0.1918471,2350.356].
Model rank =  21 / 21

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 3.046   0.797    0.02
s(I(DistTo125m/1000))                 4.000 3.639   0.807    0.16
s(ClimSST)                            4.000 1.310   0.807    0.12
s(pmin(I(ClimDistToFront2/1000), 500)) 4.000 3.809   0.814    0.26
s(I(ClimDistToEddy4/1000))            4.000 1.107   0.803    0.06

Predictors retained during the model selection procedure: Depth, DistTo125m, ClimSST,
ClimDistToFront2, ClimDistToEddy4
```

```
Predictors dropped during the model selection procedure: Slope, ClimTKE
```

Model term plots



Diagnostic plots

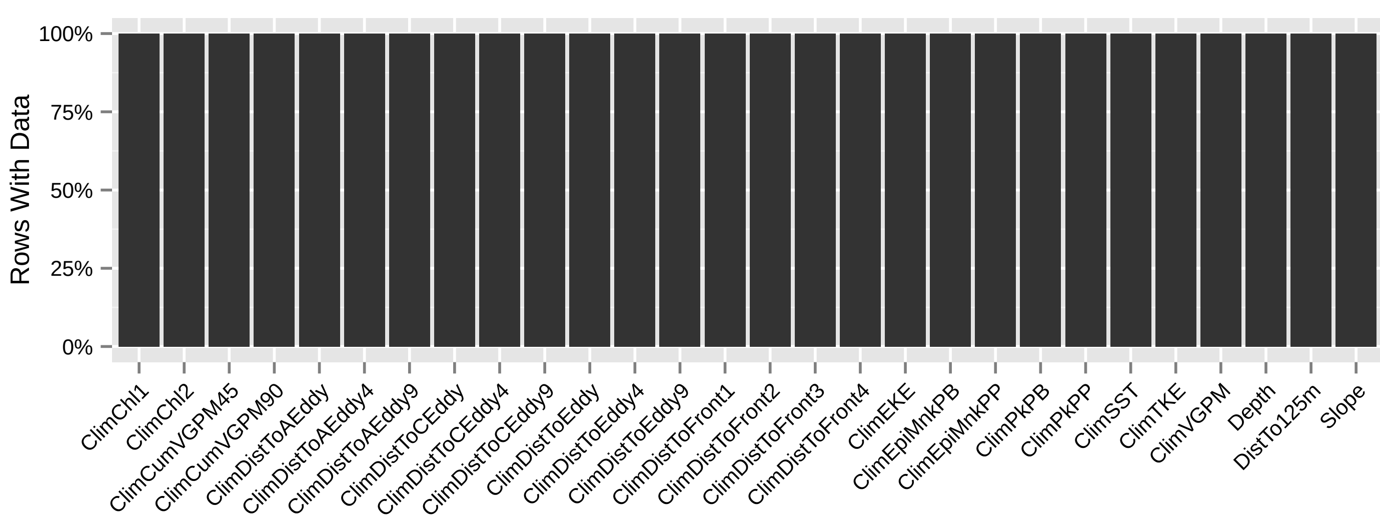


Figure 27: Segments with predictor values for the Pantropical spotted dolphin Climatological model, Off Shelf. This plot is used to assess how many segments would be lost by including a given predictor in a model.

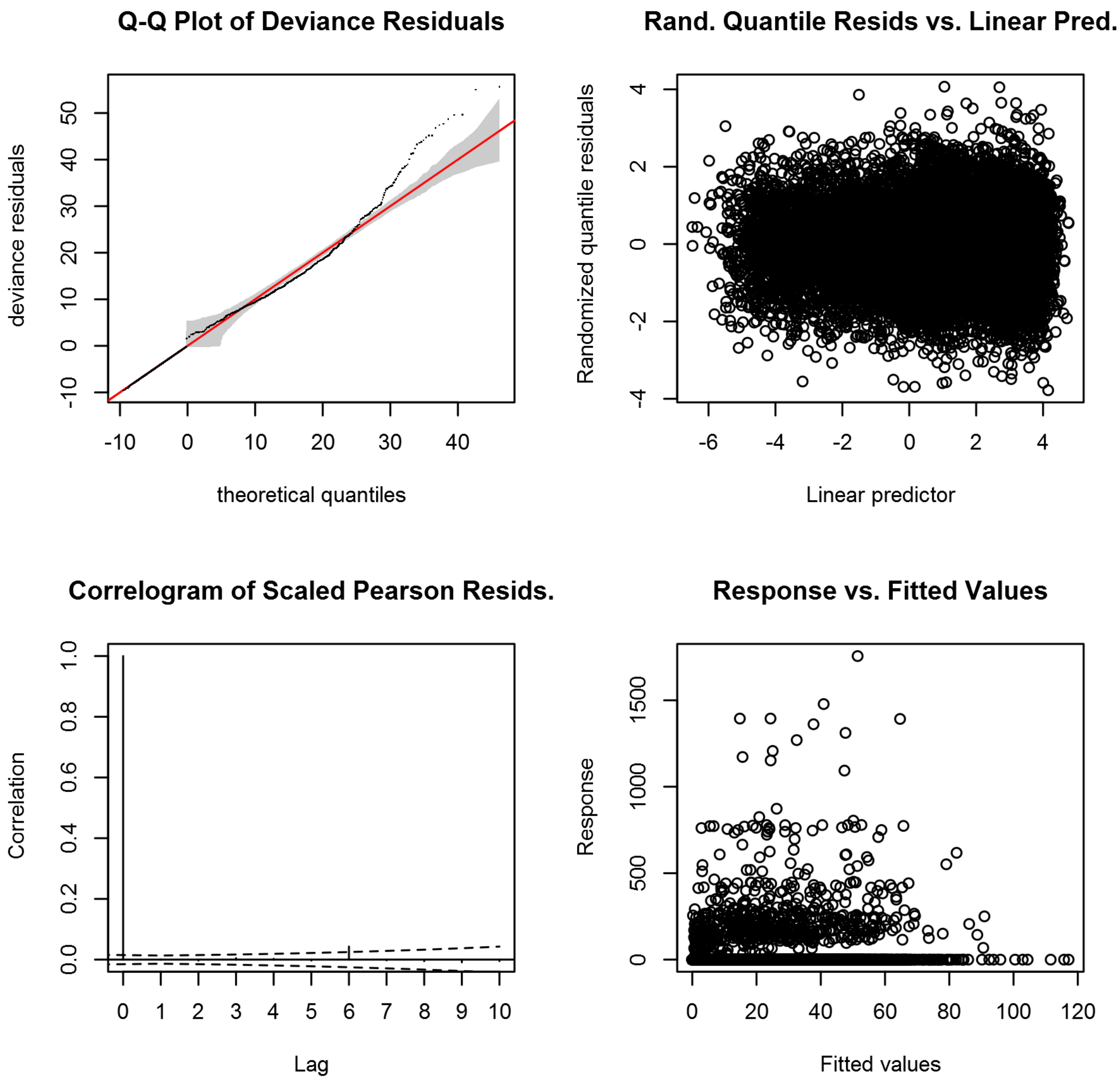


Figure 28: Statistical diagnostic plots for the Pantropical spotted dolphin Climatological model, Off Shelf.

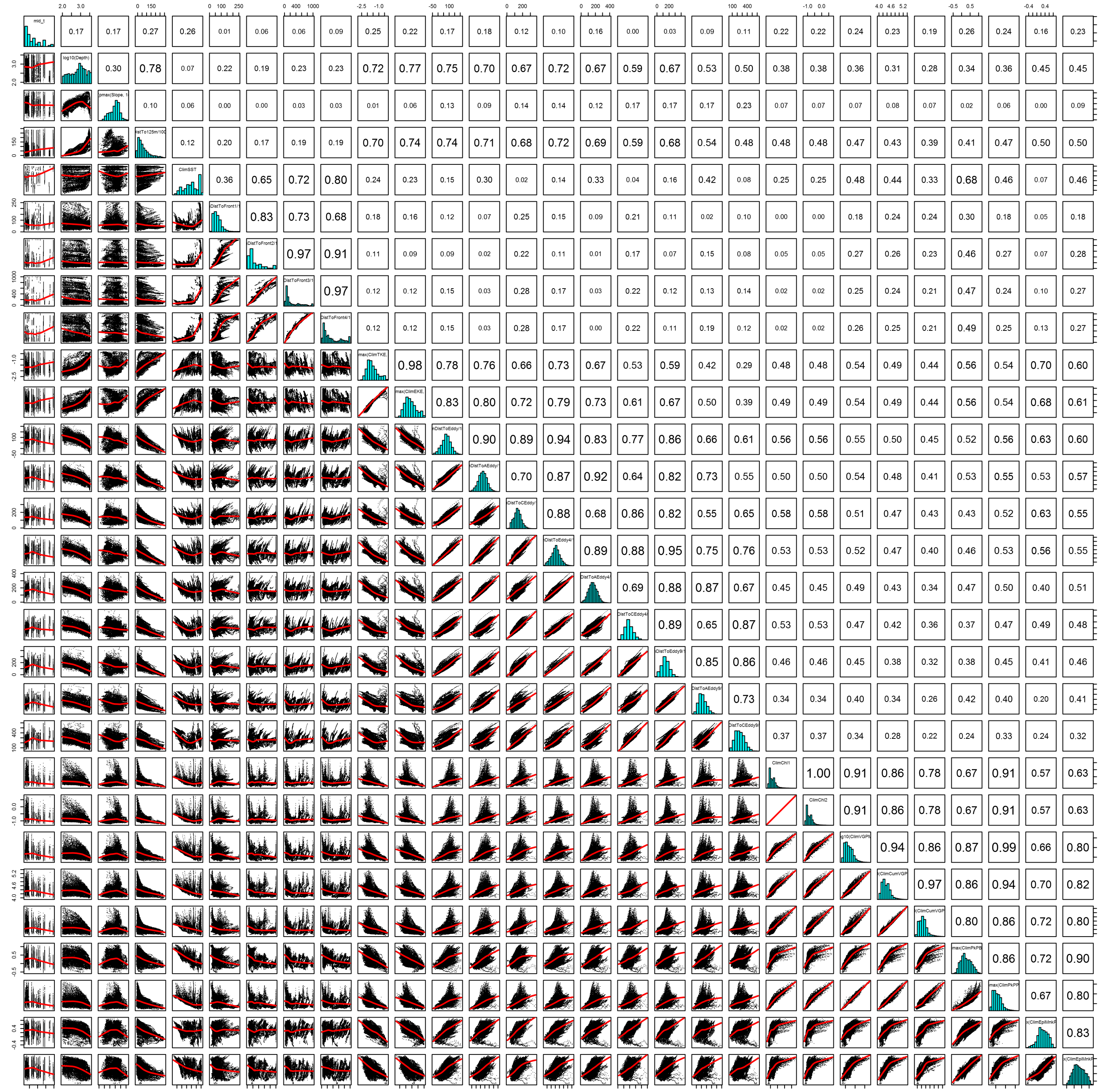


Figure 29: Scatterplot matrix for the Pantropical spotted dolphin Climatological model, Off Shelf. This plot is used to inspect the distribution of predictors (via histograms along the diagonal), simple correlation between predictors (via pairwise Pearson coefficients above the diagonal), and linearity of predictor correlations (via scatterplots below the diagonal). This plot is best viewed at high magnification.

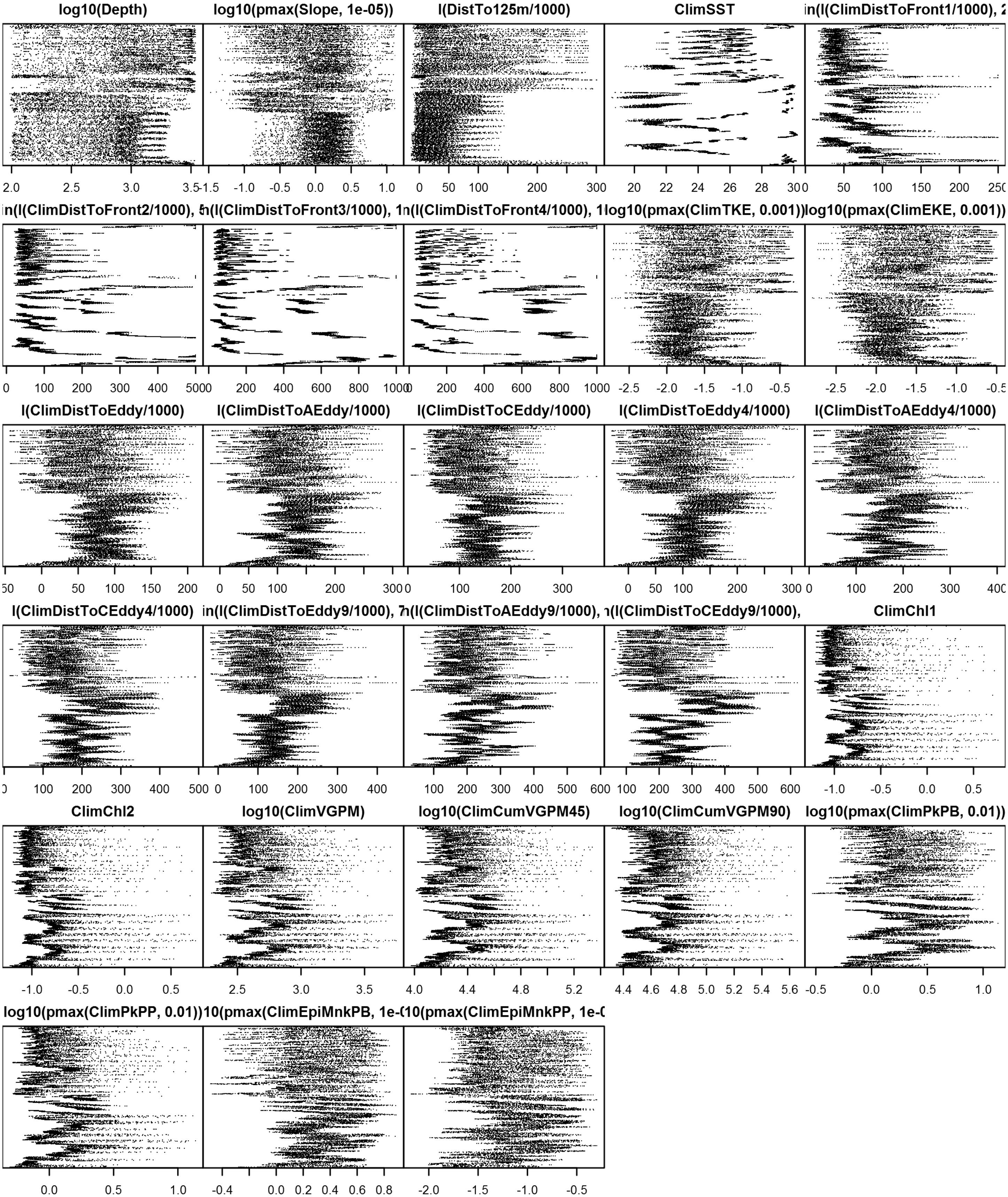


Figure 30: Dotplot for the Pantropical spotted dolphin Climatological model, Off Shelf. This plot is used to check for suspicious patterns and outliers in the data. Points are ordered vertically by transect ID, sequentially in time.

On Shelf

Density assumed to be 0 in this region.

Contemporaneous Model

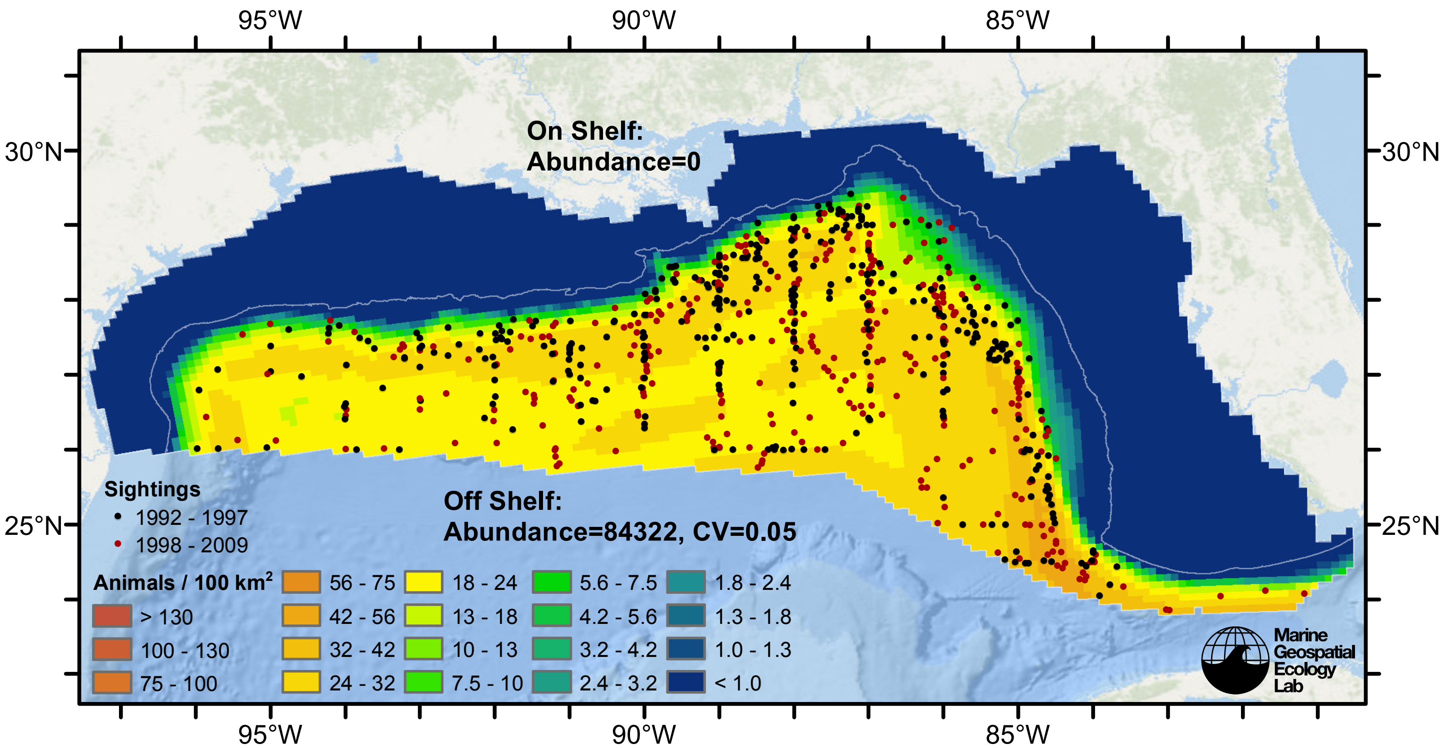


Figure 31: Pantropical spotted dolphin density predicted by the contemporaneous model that explained the most deviance. Pixels are 10x10 km. The legend gives the estimated individuals per pixel; breaks are logarithmic. Abundance for each region was computed by summing the density cells occuring in that region.

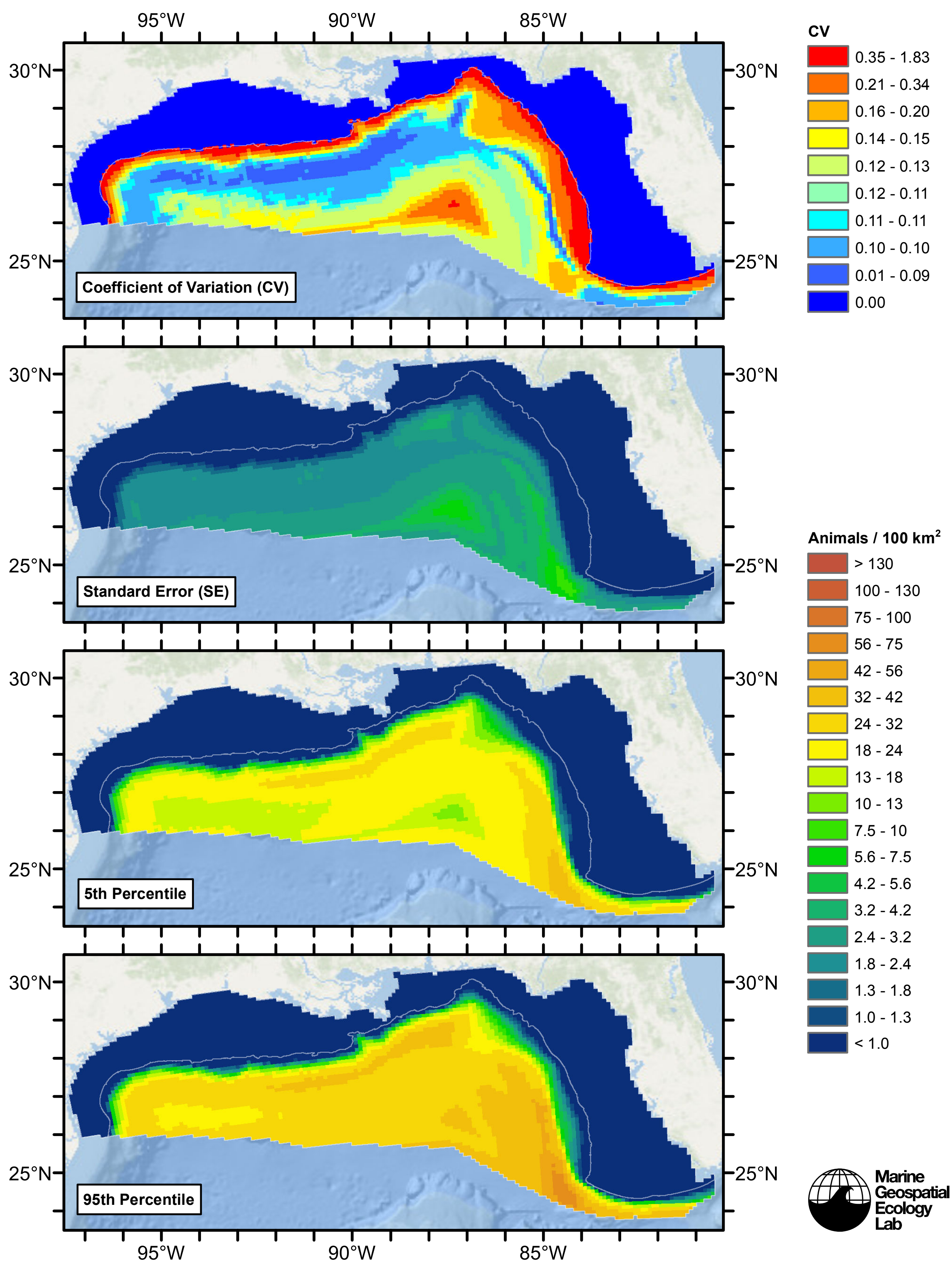


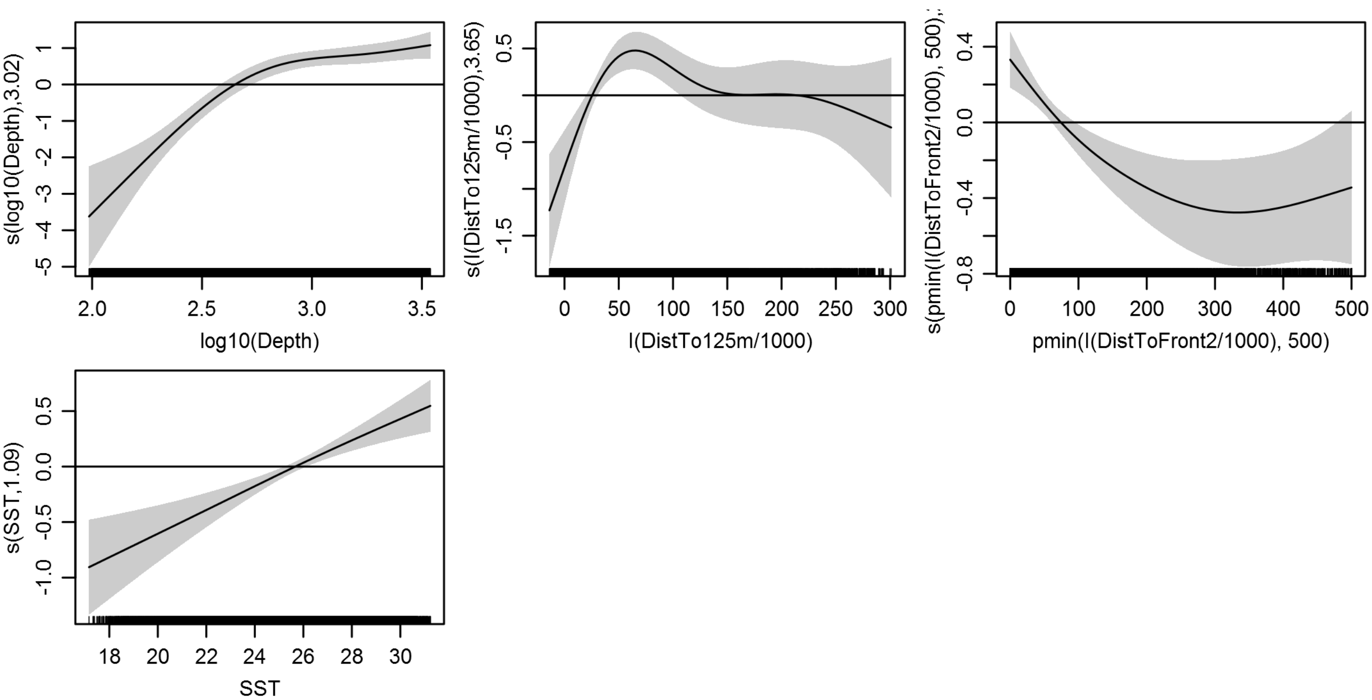
Figure 32: Estimated uncertainty for the contemporaneous model that explained the most deviance. These estimates only incorporate the statistical uncertainty estimated for the spatial model (by the R mgcv package). They do not incorporate uncertainty in the detection functions, g(0) estimates, predictor 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=1.268)
Link function: log

Formula:
abundance ~ offset(log(area_km2)) + s(log10(Depth), bs = "ts",
    k = 5) + s(I(DistTo125m/1000), bs = "ts", k = 5) + s(SST,
    bs = "ts", k = 5) + s(pmin(I(DistToFront2/1000), 500), bs = "ts",
    k = 5)

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.48732    0.09389  -26.49   <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))                  3.023      4 16.097 2.05e-15 ***
s(I(DistTo125m/1000))            3.653      4  6.329 1.48e-05 ***
s(SST)                           1.090      4  6.547 1.73e-07 ***
s(pmin(I(DistToFront2/1000), 500)) 2.136      4  5.813 2.59e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) =  0.0226   Deviance explained = 19.3%
-REML = 6524.2  Scale est. = 156.22    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 9 iterations.
Gradient range [-7.361426e-05,4.419736e-06]
(score 6524.184 & scale 156.2169).
Hessian positive definite, eigenvalue range [0.4448625,2317.164].
Model rank =  17 / 17

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 3.023   0.778    0.00
s(I(DistTo125m/1000))            4.000 3.653   0.822    0.58
s(SST)                           4.000 1.090   0.798    0.04
s(pmin(I(DistToFront2/1000), 500)) 4.000 2.136   0.829    0.82

Predictors retained during the model selection procedure: Depth, DistTo125m, SST, DistToFront2

Predictors dropped during the model selection procedure: Slope
```

Model term plots



Diagnostic plots

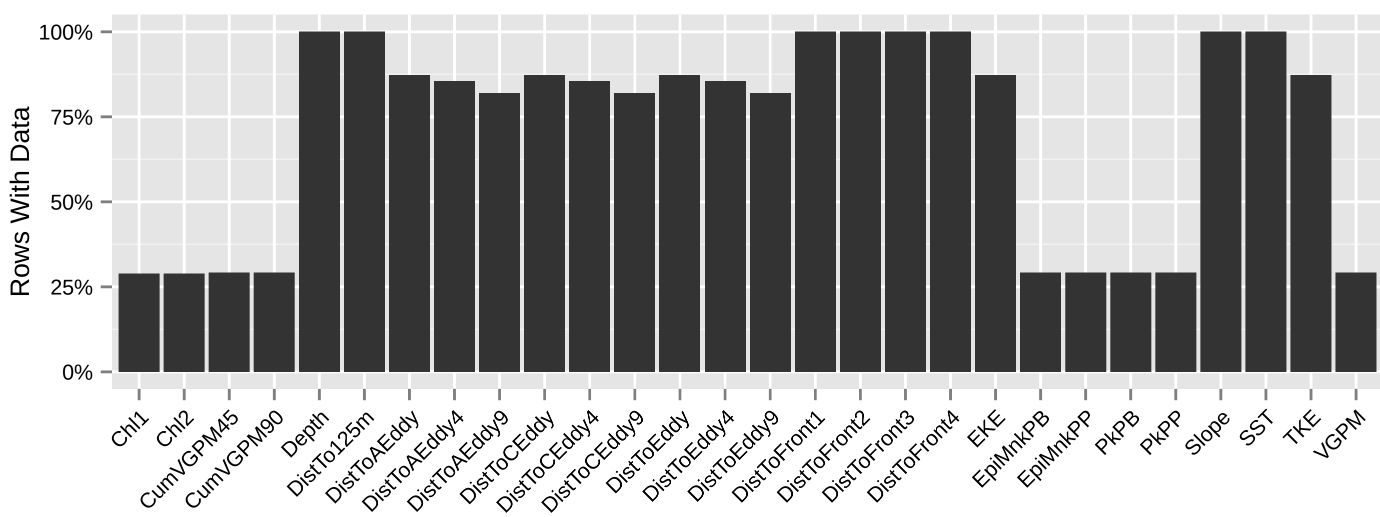


Figure 33: Segments with predictor values for the Pantropical spotted dolphin Contemporaneous model, Off Shelf. This plot is used to assess how many segments would be lost by including a given predictor in a model.

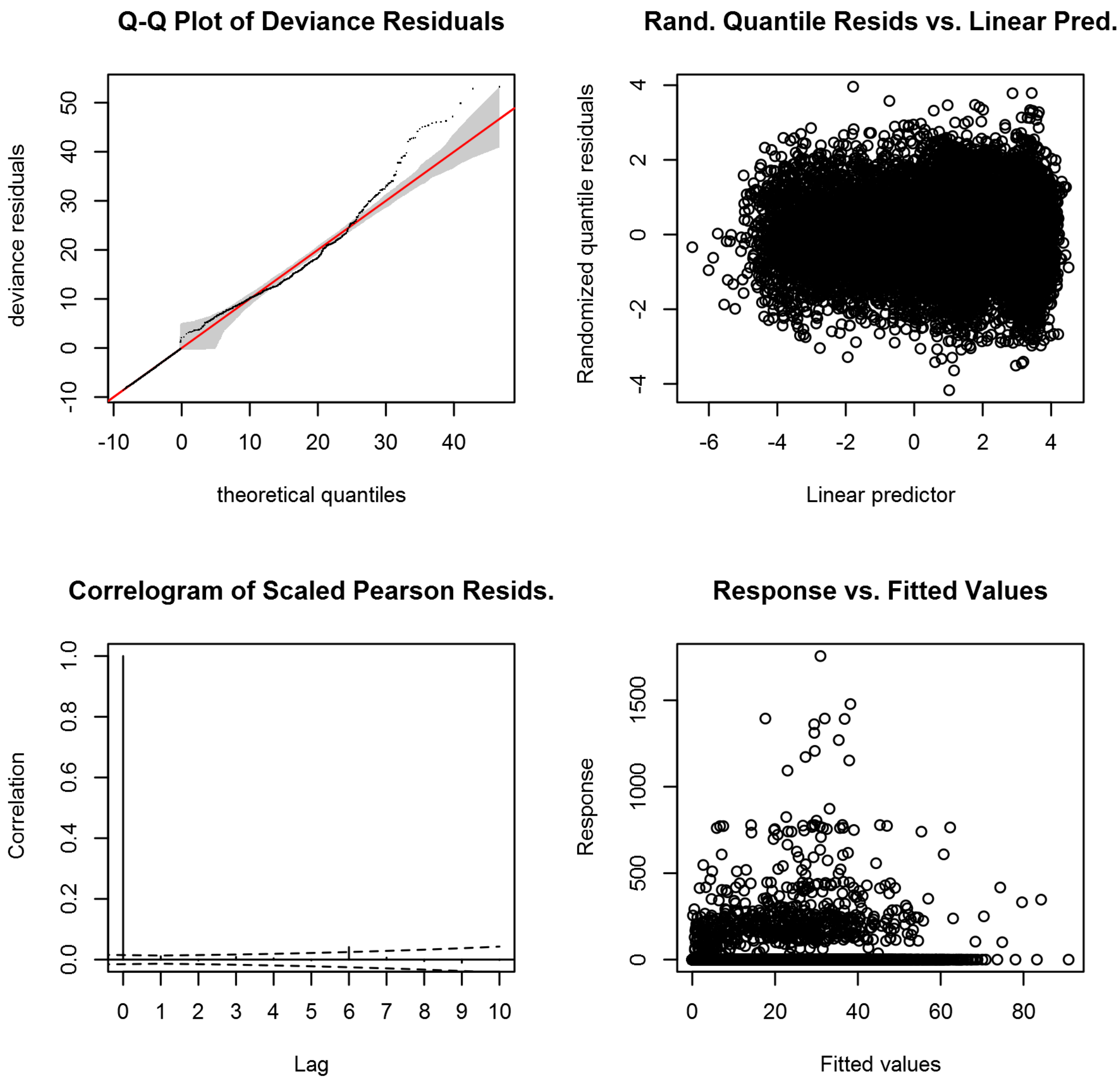


Figure 34: Statistical diagnostic plots for the Pantropical spotted dolphin Contemporaneous model, Off Shelf.

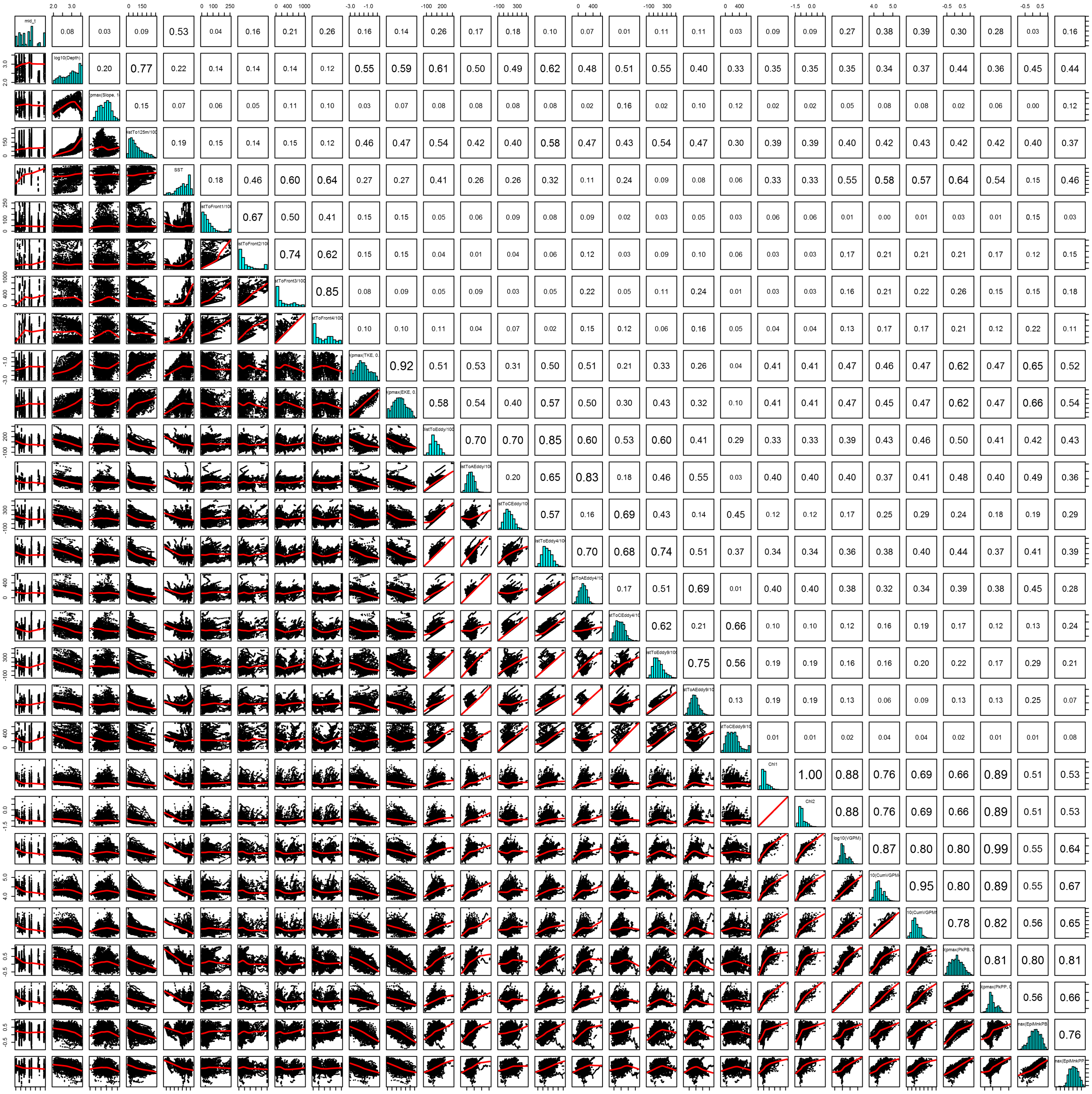


Figure 35: Scatterplot matrix for the Pantropical spotted dolphin Contemporaneous model, Off Shelf. This plot is used to inspect the distribution of predictors (via histograms along the diagonal), simple correlation between predictors (via pairwise Pearson coefficients above the diagonal), and linearity of predictor correlations (via scatterplots below the diagonal). This plot is best viewed at high magnification.

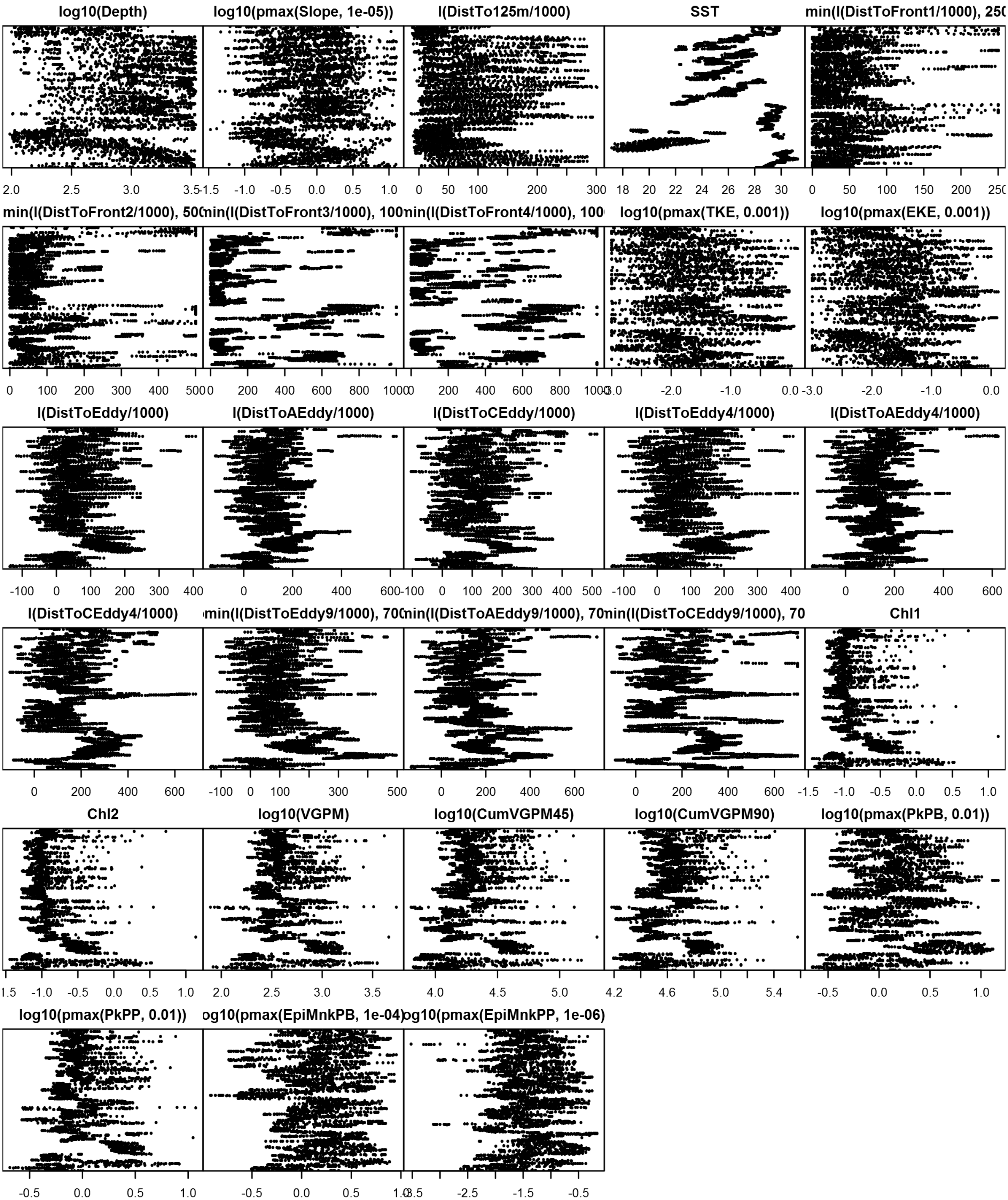


Figure 36: Dotplot for the Pantropical spotted dolphin Contemporaneous model, Off Shelf. This plot is used to check for suspicious patterns and outliers in the data. Points are ordered vertically by transect ID, sequentially in time.

On Shelf

Density assumed to be 0 in this region.

Climatological Same Segments Model

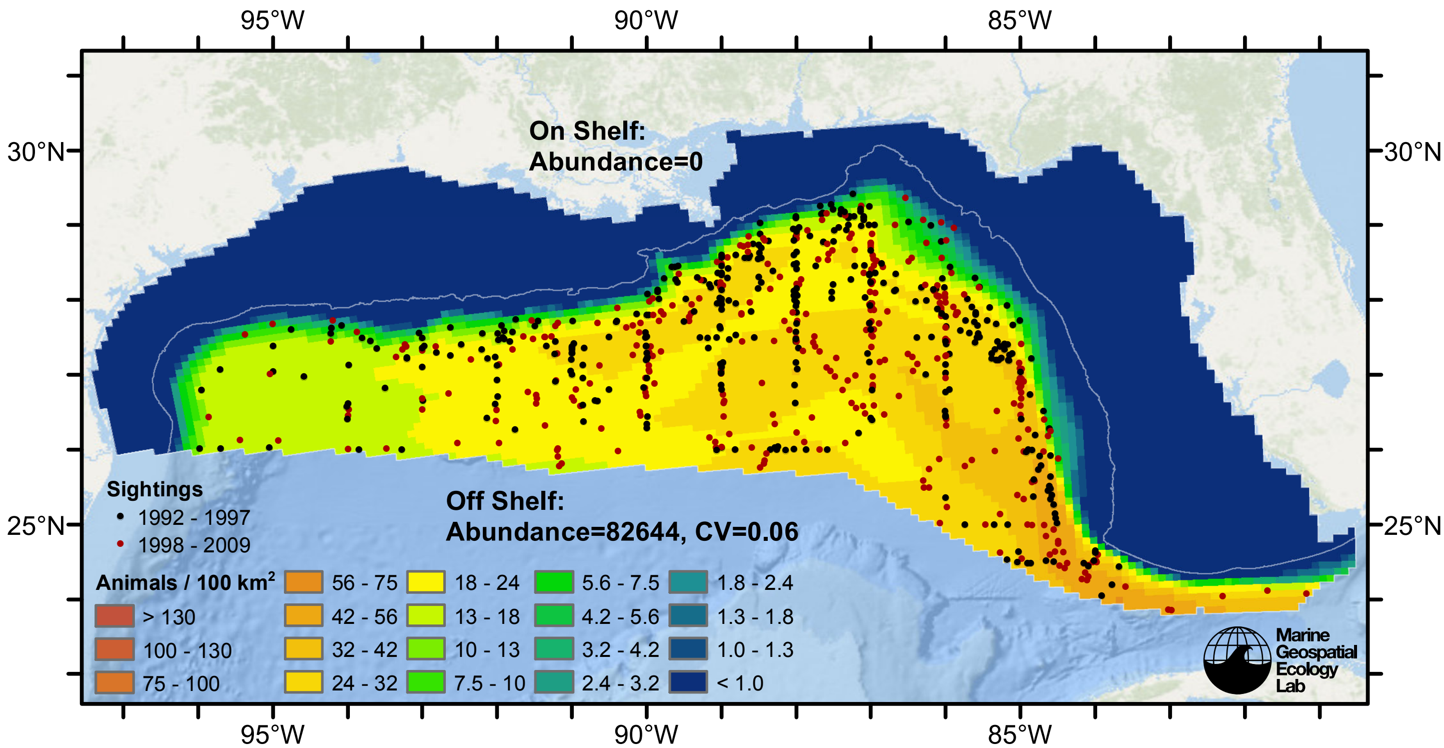


Figure 37: Pantropical spotted dolphin density predicted by the climatological same segments model that explained the most deviance. Pixels are 10x10 km. The legend gives the estimated individuals per pixel; breaks are logarithmic. Abundance for each region was computed by summing the density cells occuring in that region.

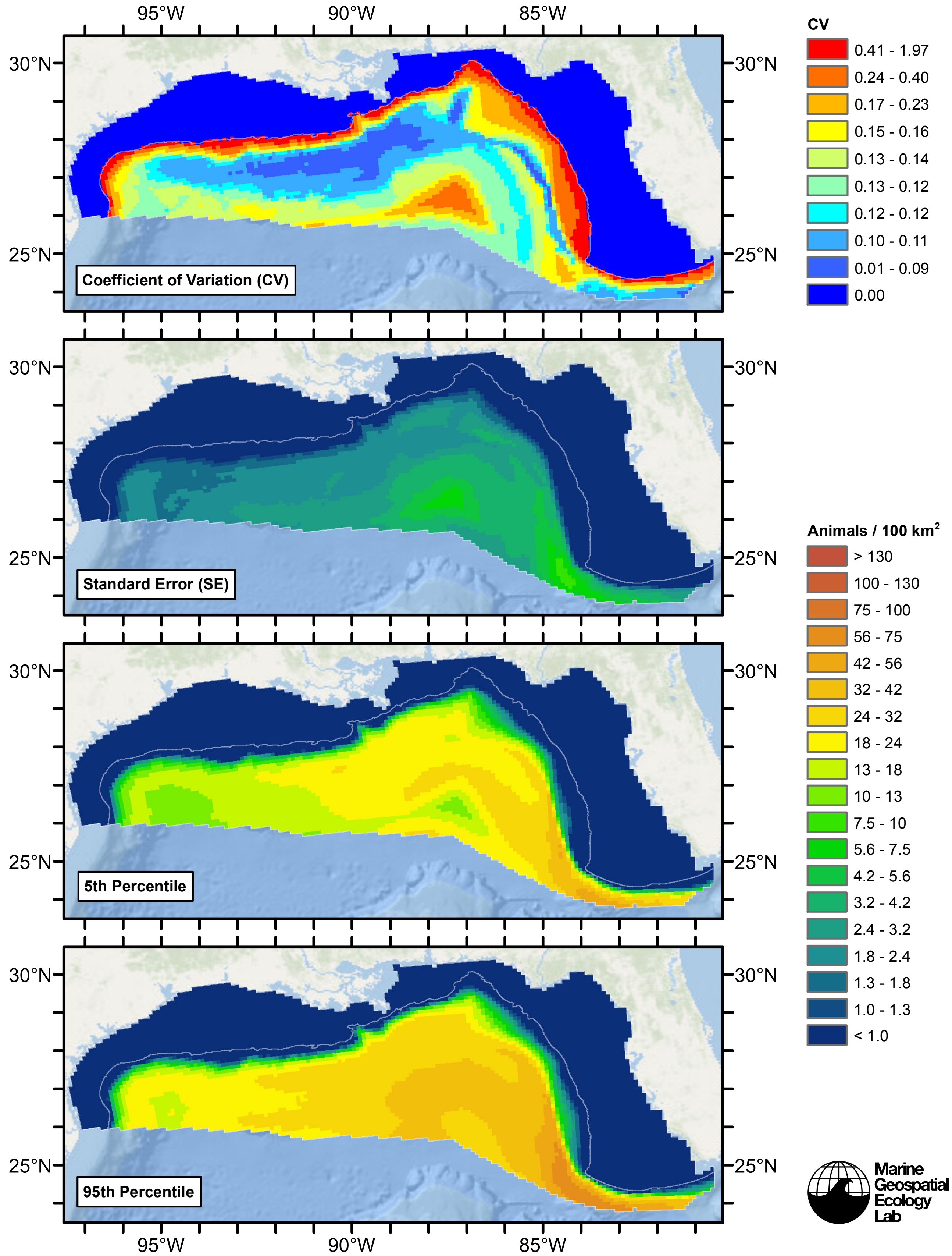


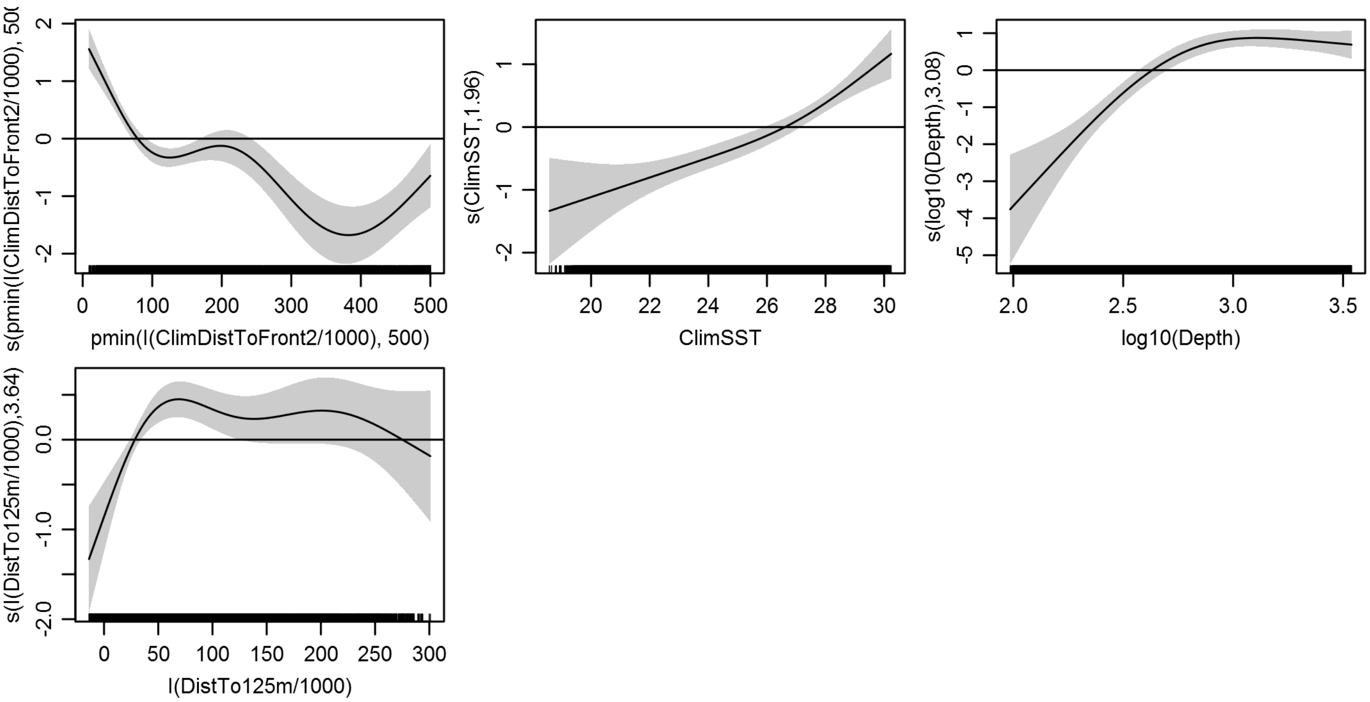
Figure 38: Estimated uncertainty for the climatological same segments model that explained the most deviance. These estimates only incorporate the statistical uncertainty estimated for the spatial model (by the R mgcv package). They do not incorporate uncertainty in the detection functions, g(0) estimates, 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=1.264)
Link function: log

Formula:
abundance ~ offset(log(area_km2)) + s(log10(Depth), bs = "ts",
    k = 5) + s(I(DistTo125m/1000), bs = "ts", k = 5) + s(ClimSST,
    bs = "ts", k = 5) + s(pmin(I(ClimDistToFront2/1000), 500),
    bs = "ts", k = 5)

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -2.6416     0.1012   -26.1   <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))                        3.082      4 16.379 1.48e-15 ***
s(I(DistTo125m/1000))                  3.642      4  5.457 8.37e-05 ***
s(ClimSST)                             1.958      4 12.473 2.49e-13 ***
s(pmin(I(ClimDistToFront2/1000), 500)) 3.853      4 21.981  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) =  0.0288   Deviance explained = 21.4%
-REML = 6497.4  Scale est. = 154.57    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 10 iterations.
Gradient range [-0.000242619,5.965936e-05]
(score 6497.392 & scale 154.5693).
Hessian positive definite, eigenvalue range [0.2651722,2344.734].
Model rank =  17 / 17

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 3.082   0.823    0.19
s(I(DistTo125m/1000))                  4.000 3.642   0.805    0.04
s(ClimSST)                             4.000 1.958   0.812    0.05
s(pmin(I(ClimDistToFront2/1000), 500)) 4.000 3.853   0.827    0.28

Predictors retained during the model selection procedure: Depth, DistTo125m, ClimSST,
ClimDistToFront2

Predictors dropped during the model selection procedure: Slope
```

Model term plots



Diagnostic plots

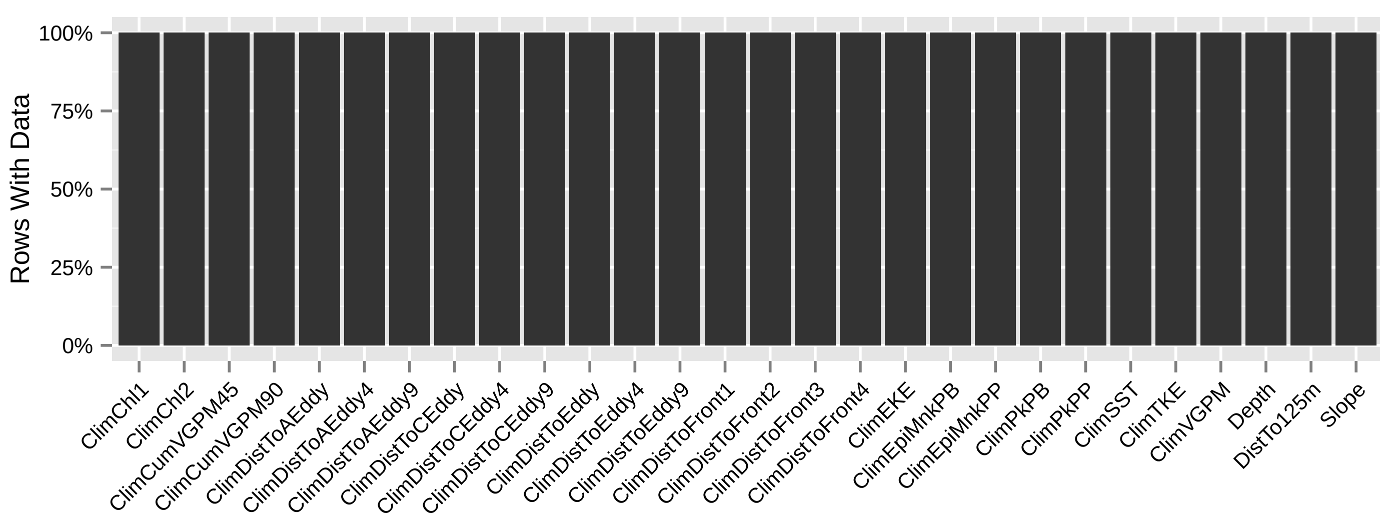


Figure 39: Segments with predictor values for the Pantropical spotted dolphin Climatological model, Off Shelf. This plot is used to assess how many segments would be lost by including a given predictor in a model.

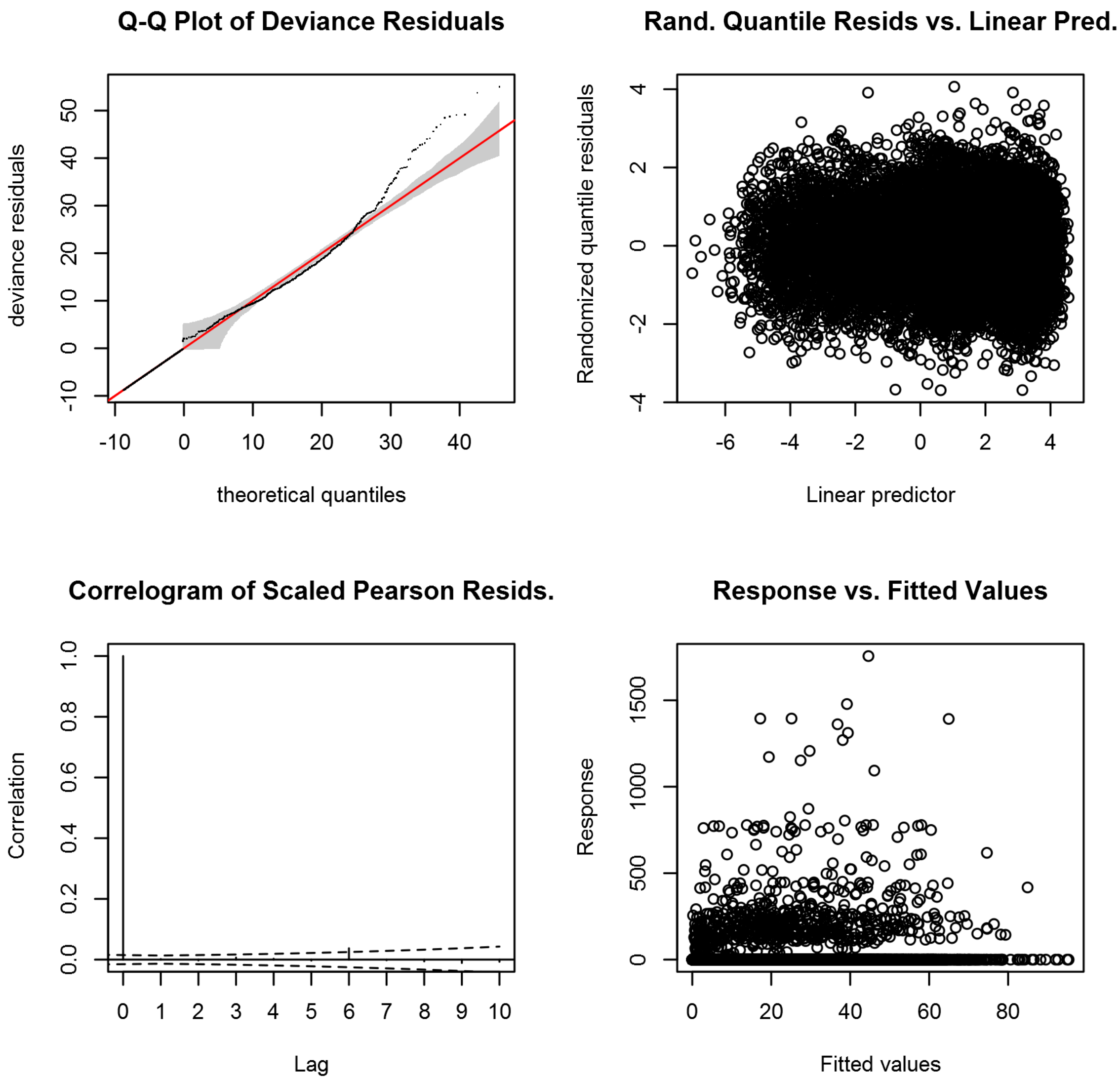


Figure 40: Statistical diagnostic plots for the Pantropical spotted dolphin Climatological model, Off Shelf.

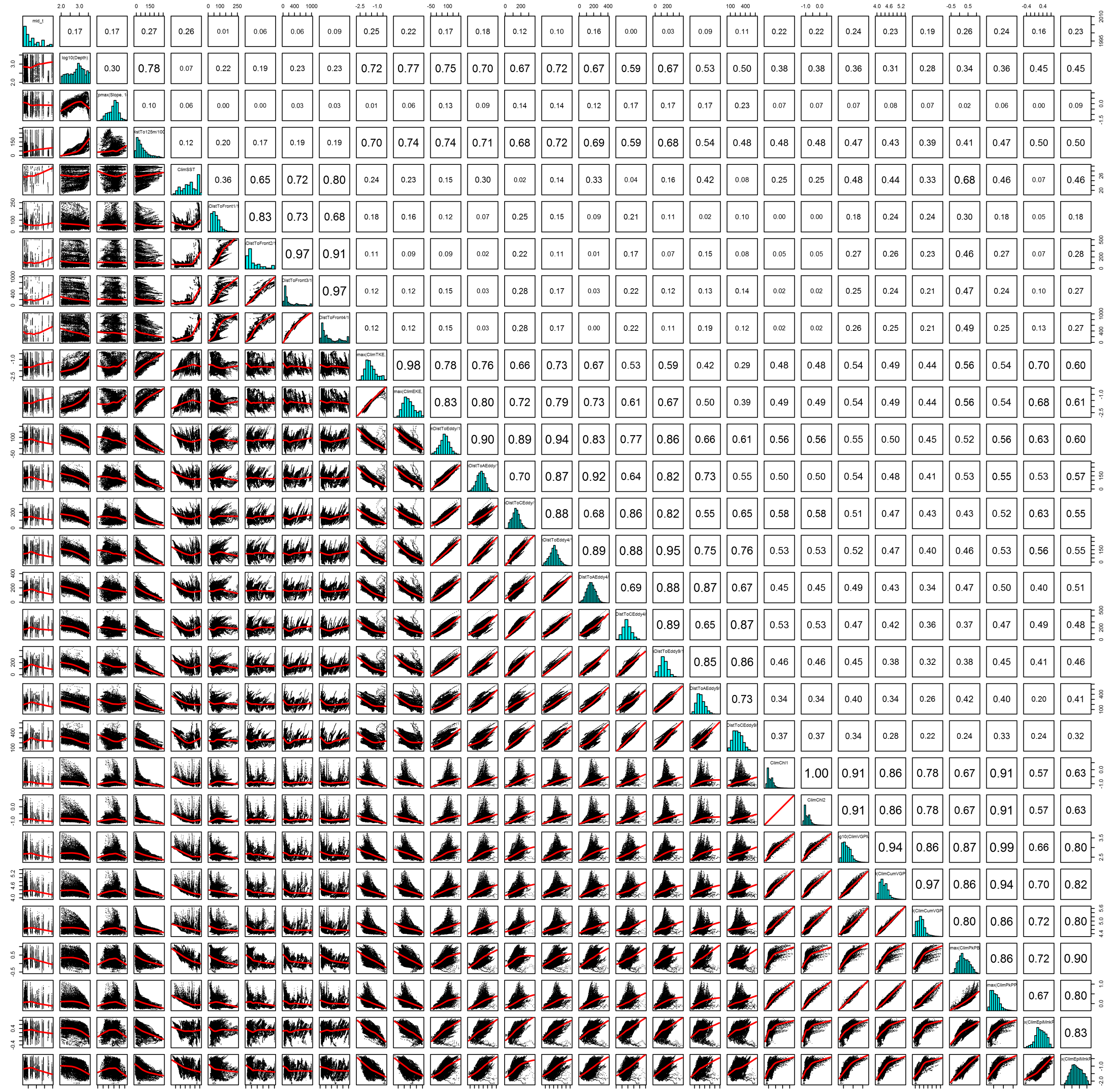


Figure 41: Scatterplot matrix for the Pantropical spotted dolphin Climatological model, Off Shelf. This plot is used to inspect the distribution of predictors (via histograms along the diagonal), simple correlation between predictors (via pairwise Pearson coefficients above the diagonal), and linearity of predictor correlations (via scatterplots below the diagonal). This plot is best viewed at high magnification.

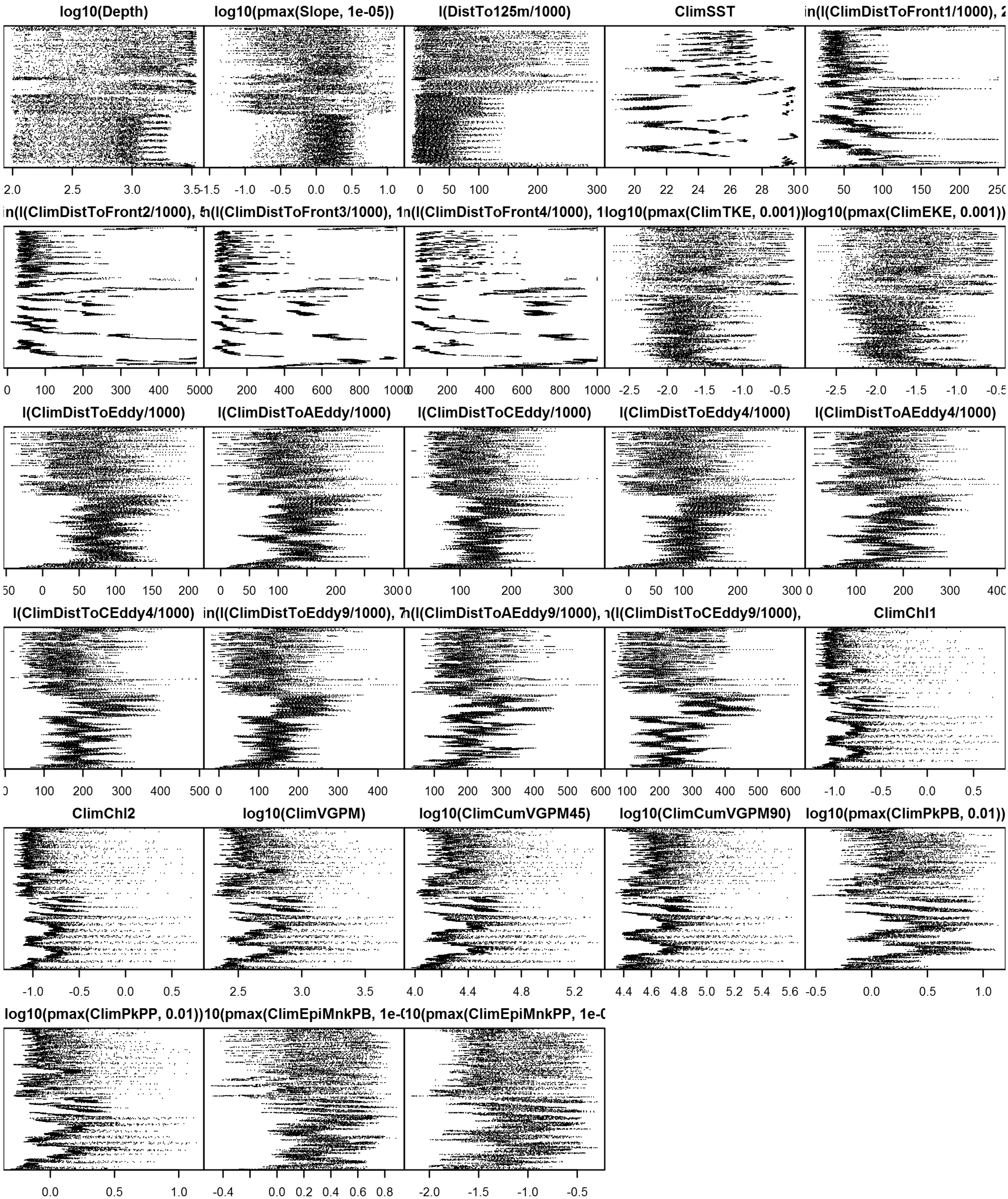


Figure 42: Dotplot for the Pantropical spotted dolphin Climatological model, Off Shelf. This plot is used to check for suspicious patterns and outliers in the data. Points are ordered vertically by transect ID, sequentially in time.

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	18.1			14455		1992-2009
Phys+SST	21.4	19.3	21.4	14455	0.0	1992-2009
Phys+SST+Curr	21.7	18.4	19.5	11840	18.1	1993-2009
Phys+SST+Curr+Prod	21.7	10.6	7.9	4179	71.1	1998-2009

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*	84014	0.06	No	
1992-2009	Contemporaneous model	84322	0.05	No	
1992-2009	Climatological same segments model	82644	0.06	No	
2009	Oceanic waters, Jun-Aug (Waring et al. 2013)	50880	0.27	Yes	Yes
2003-2004	Oceanic waters, Jun-Aug (Mullin 2007)	34067	0.18	Yes	Yes
1996-2001	Oceanic waters, Apr-Jun (Mullin and Fulling 2004)	91321	0.16	Yes	Yes
1991-1994	Oceanic waters, Apr-Jun (Hansen et al. 1995)	31320	0.20	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

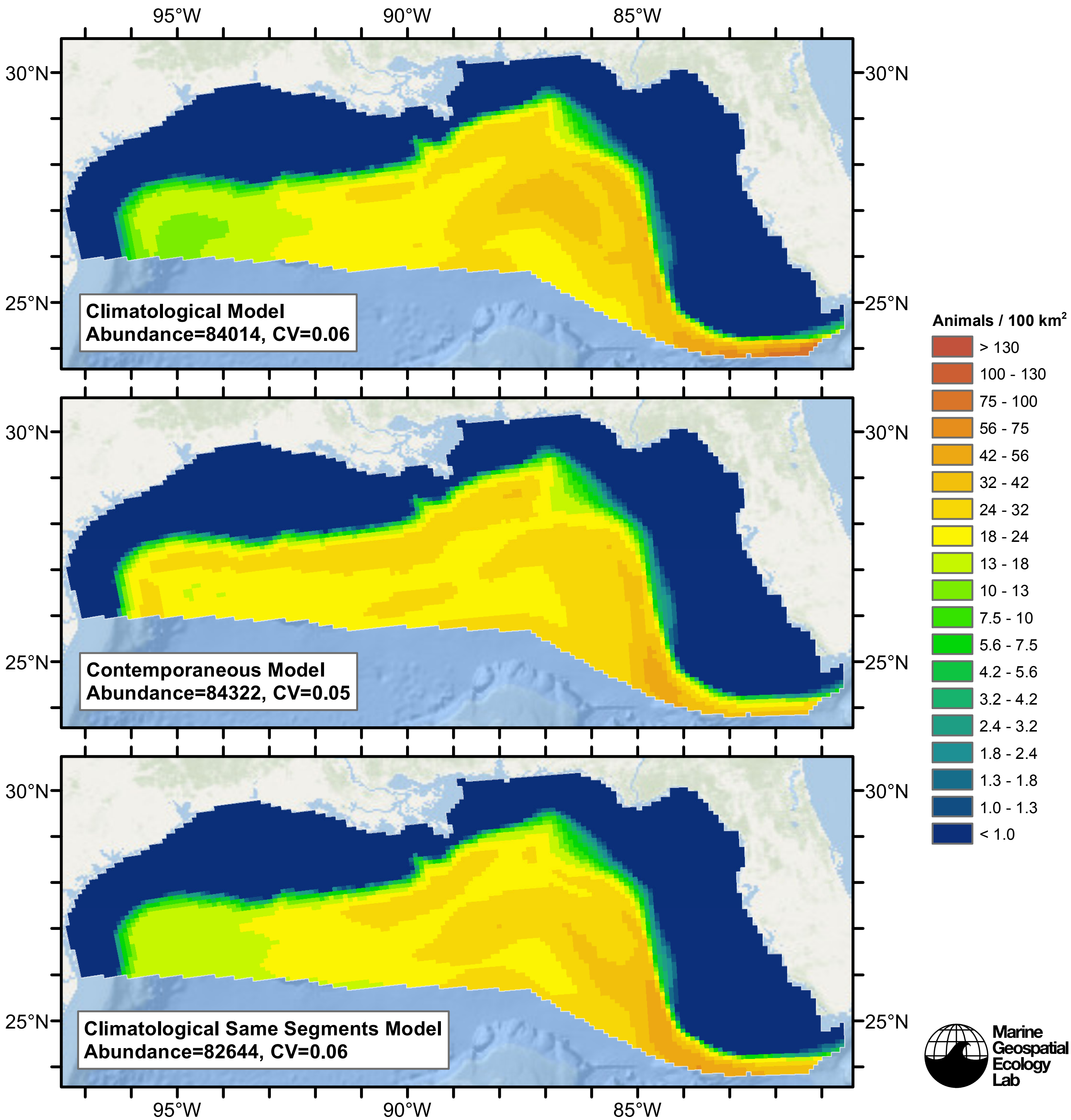


Figure 43: Pantropical spotted dolphin density and abundance predicted by the models that explained the most deviance. Regions inside the study area (white line) where the background map is visible are areas we did not model (see text).

Temporal Variability

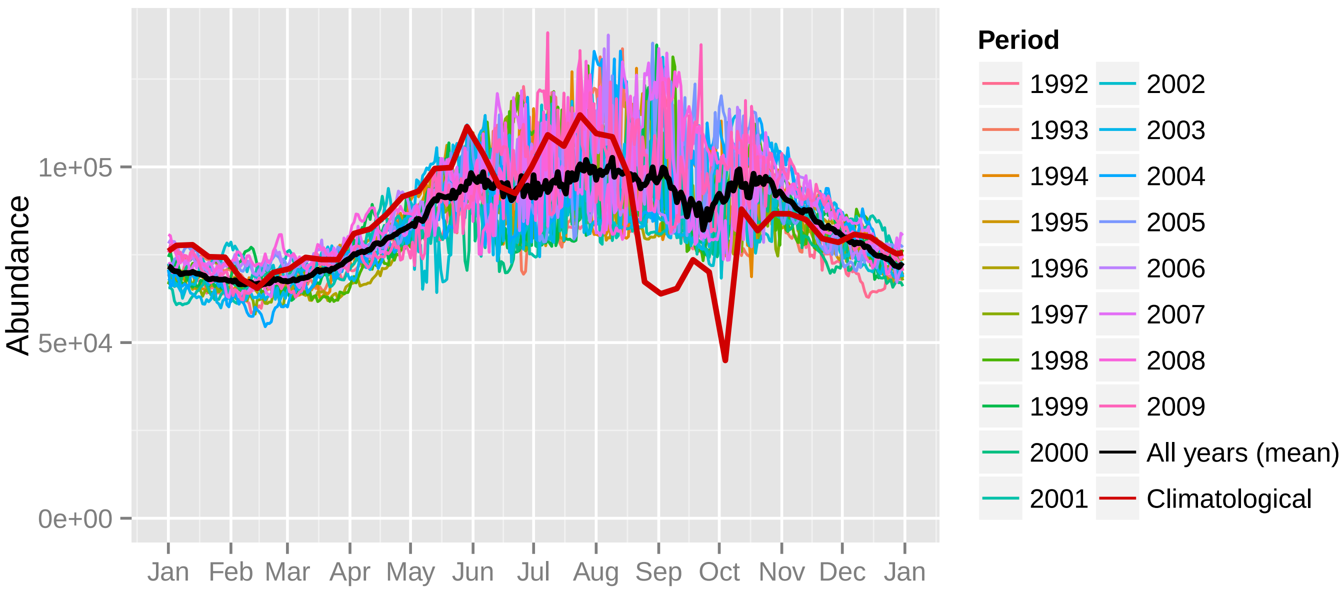


Figure 44: Comparison of Pantropical spotted dolphin abundance predicted at a daily time step for different time periods. Individual years were predicted using contemporaneous models. “All years (mean)” averages the individual years, giving the mean annual abundance of the contemporaneous model. “Climatological” was predicted using the climatological model. The results for the climatological same segments model are not shown.

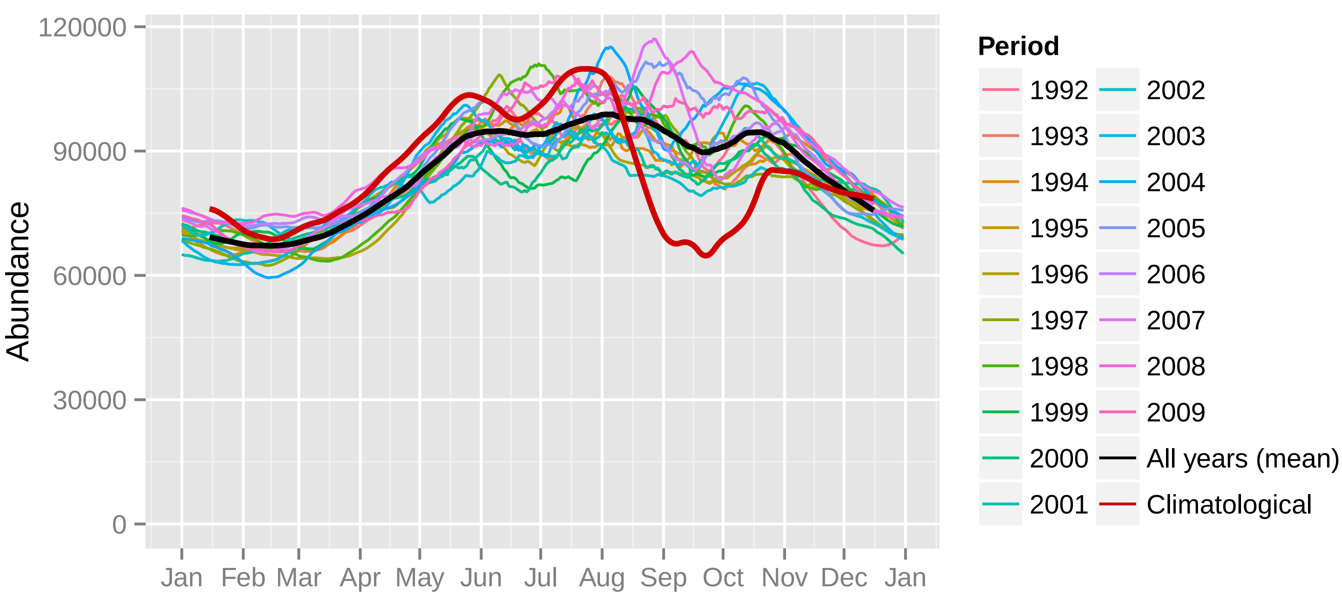
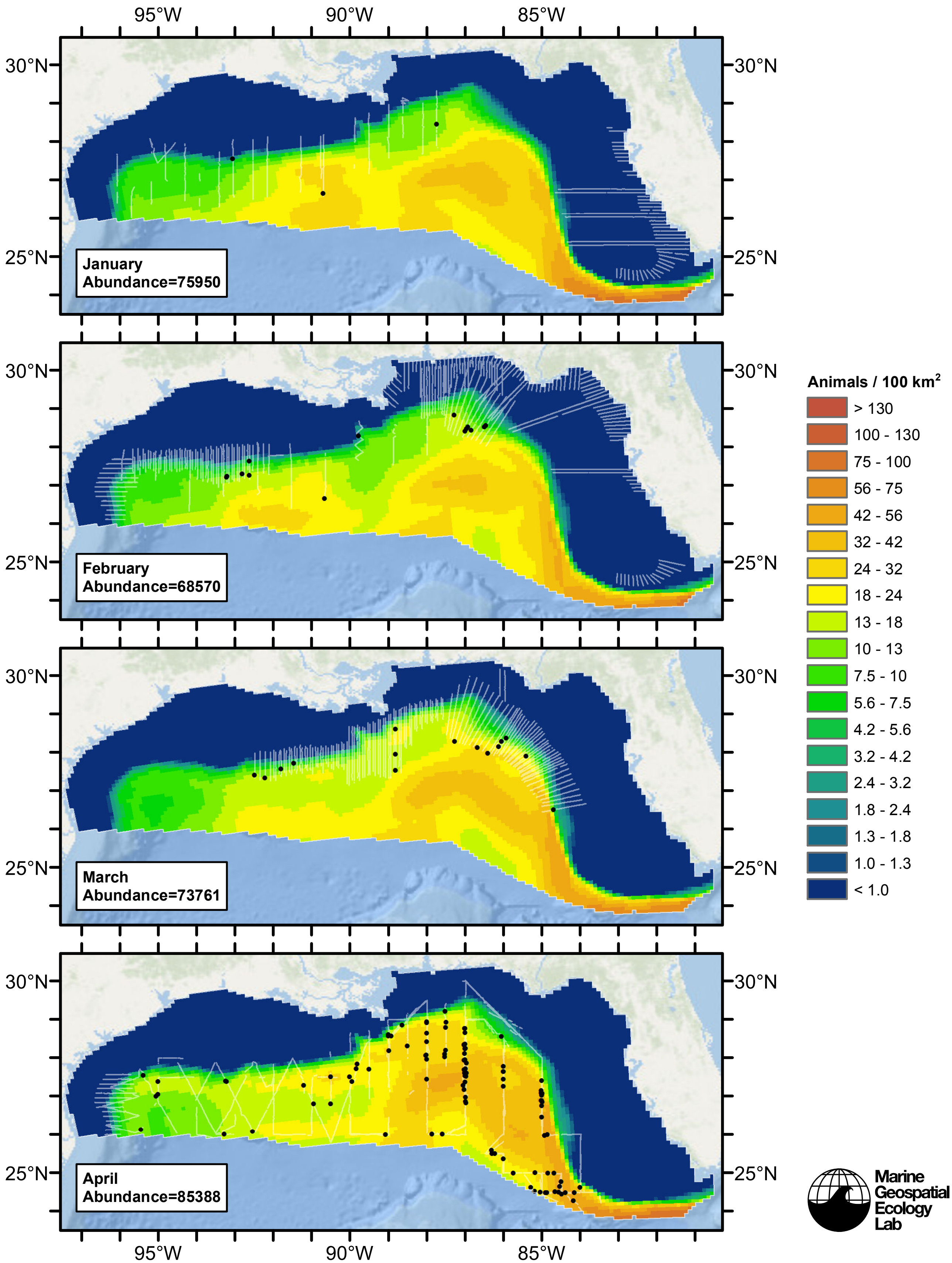
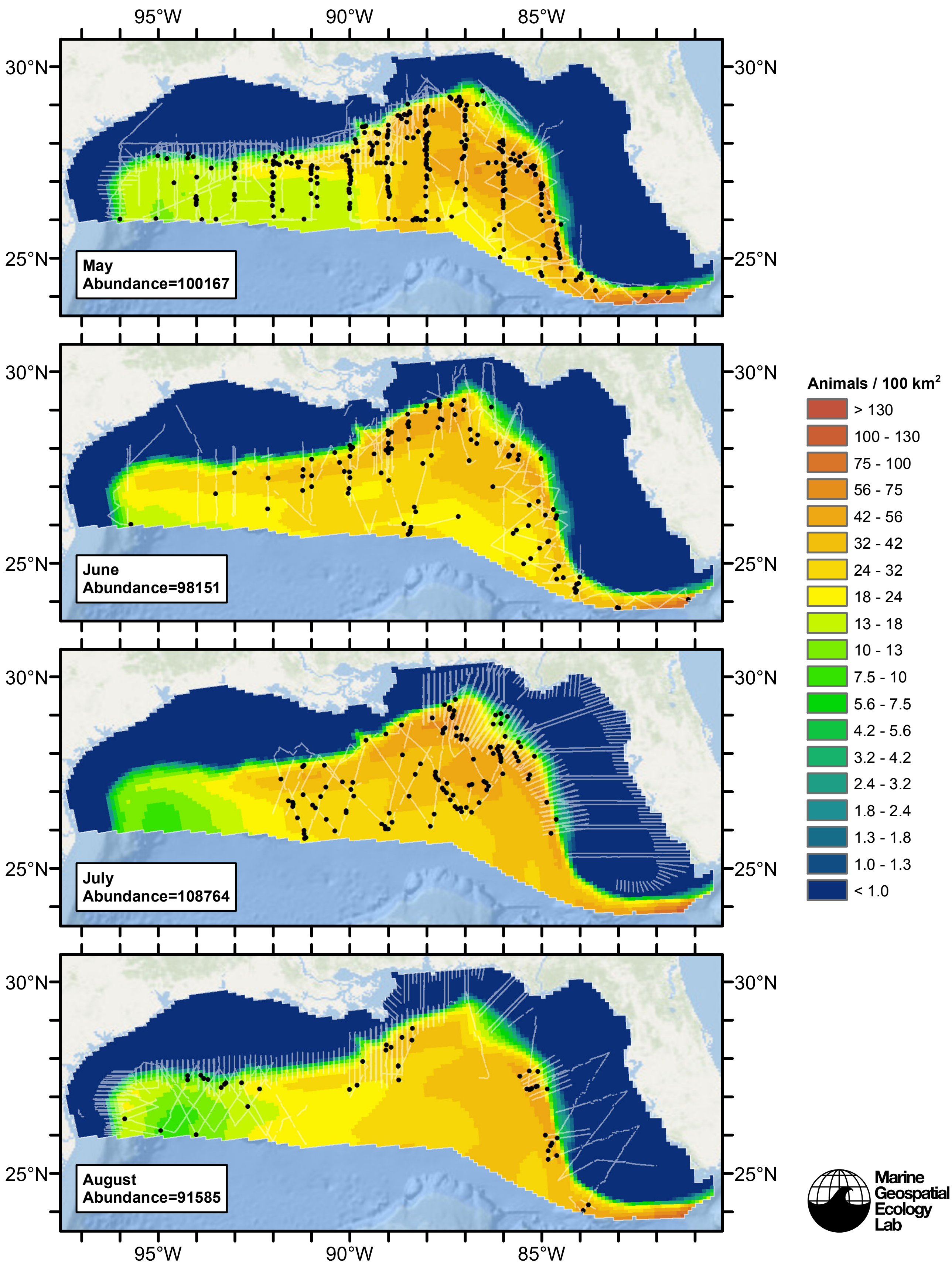


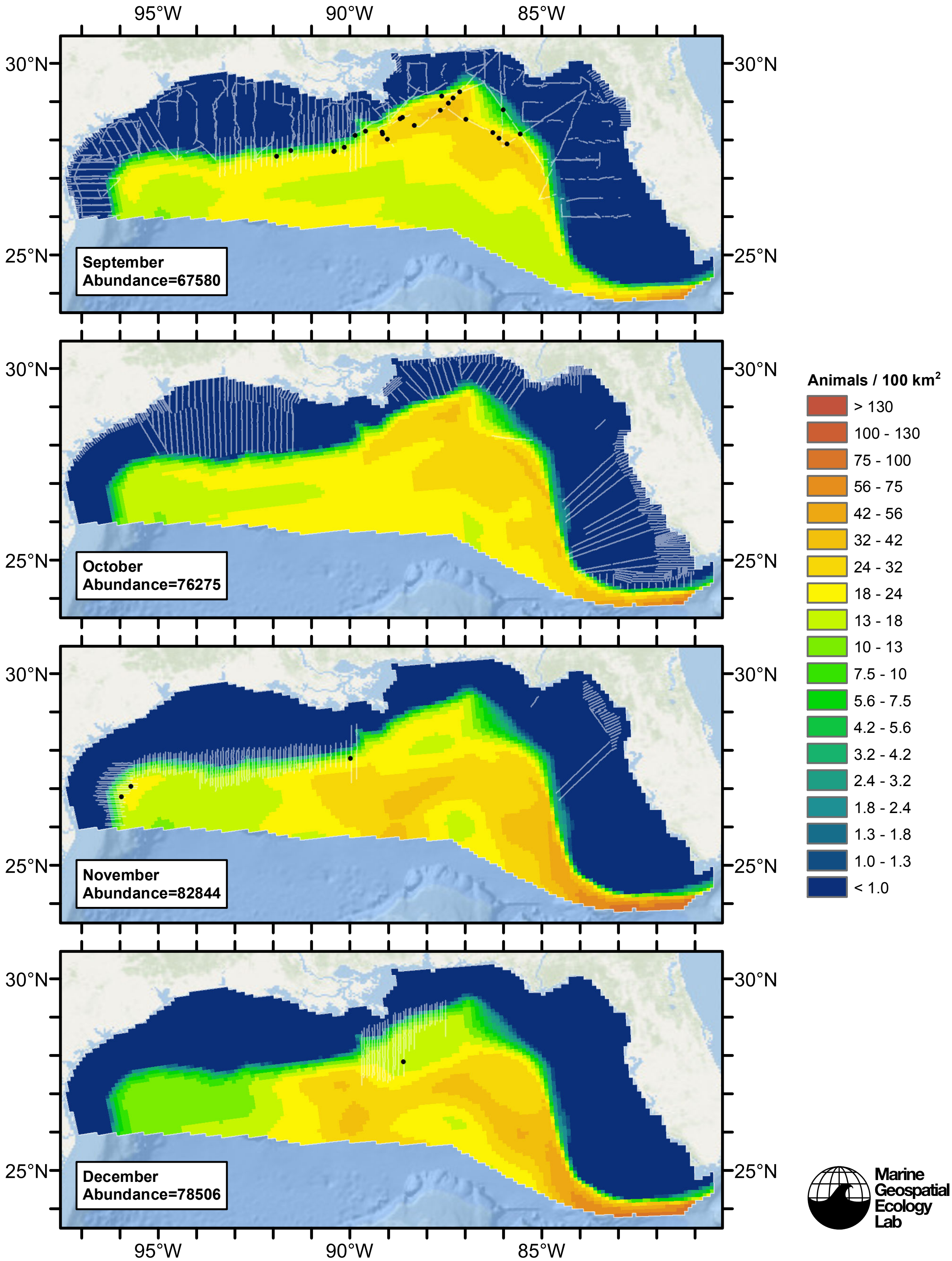
Figure 45: The same data as the preceding figure, but with a 30-day moving average applied.

Climatological Model

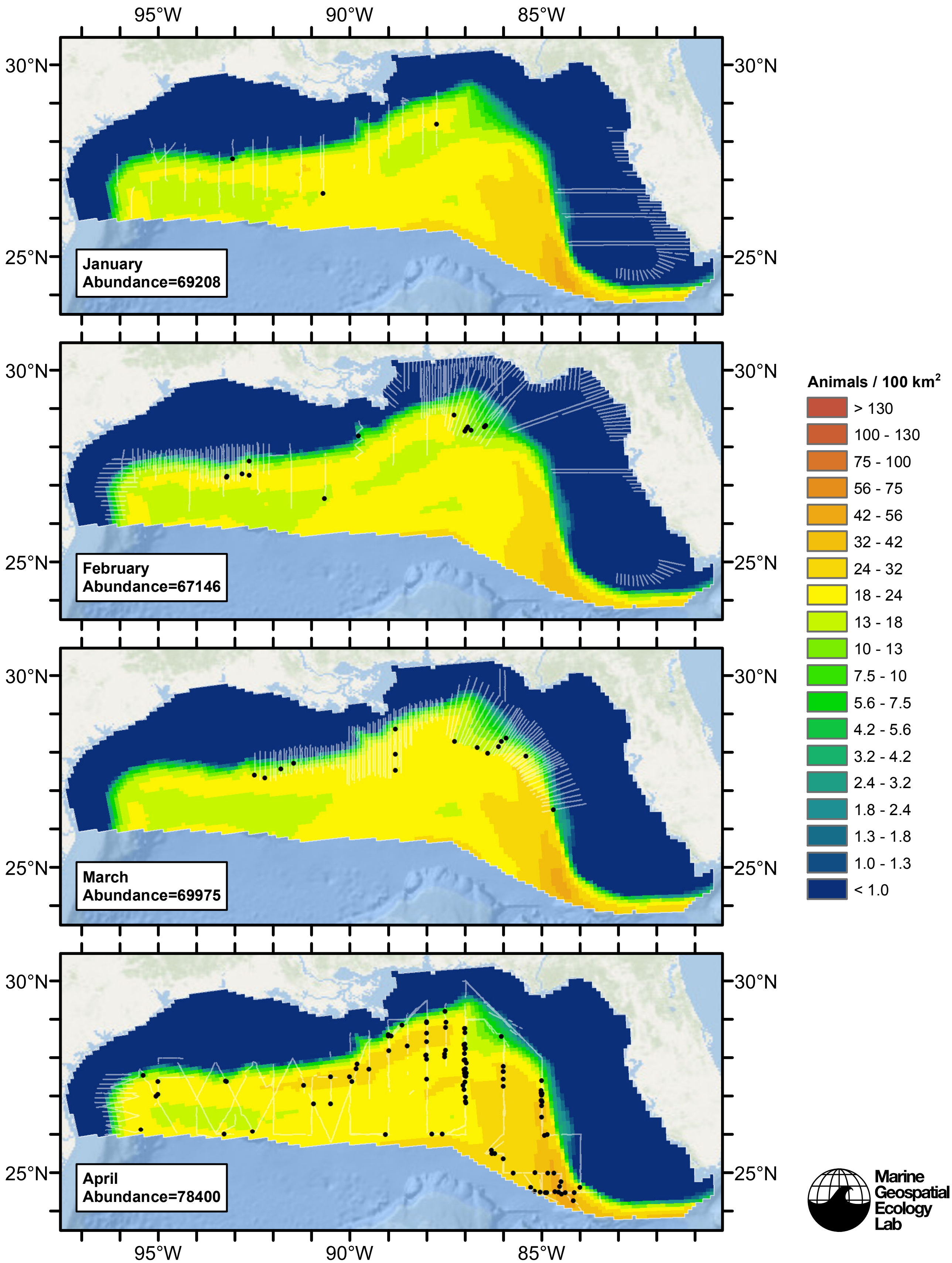


Marine Geospatial Ecology Lab

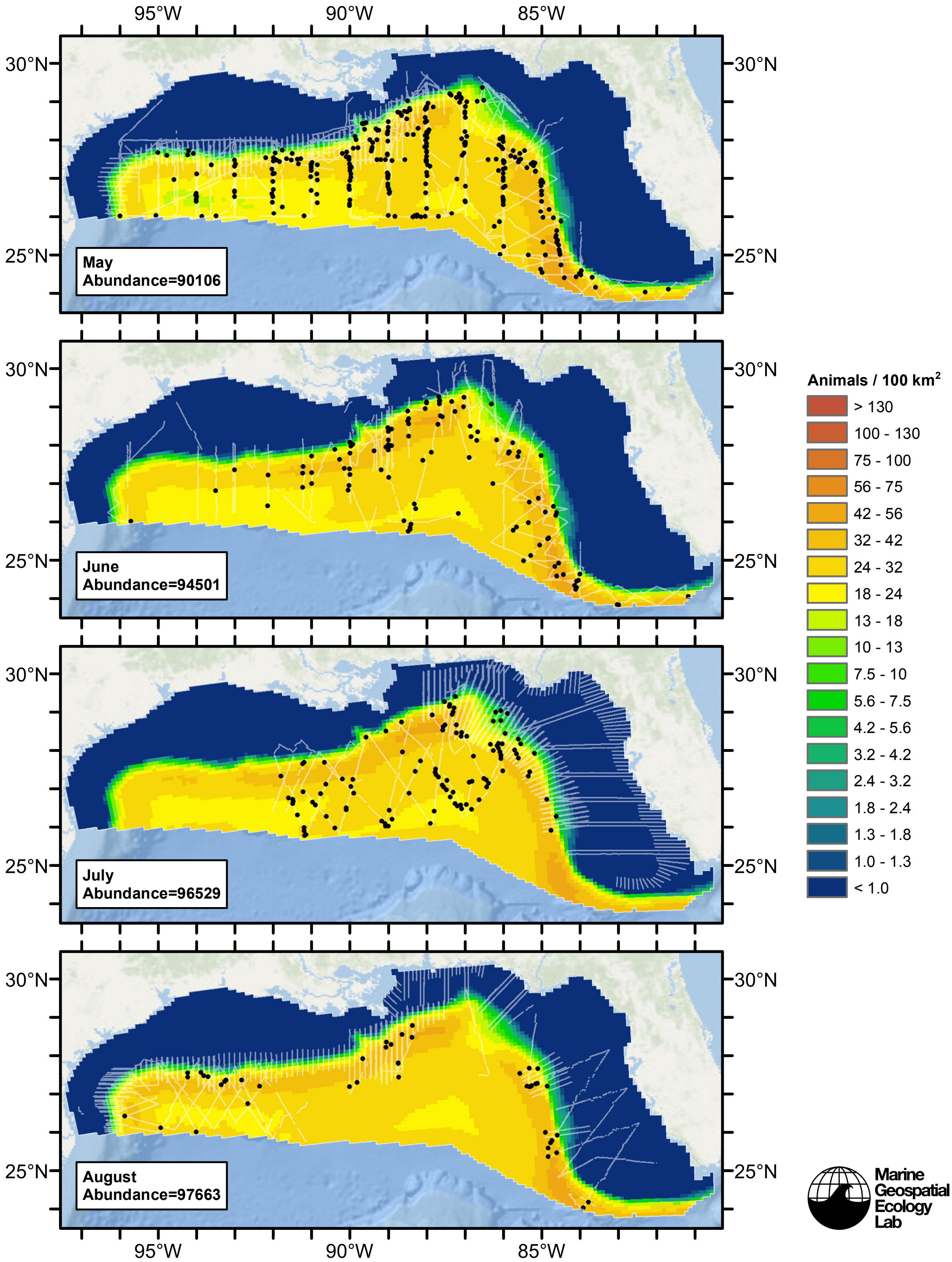

95°W
90°W
85°W
30°N
25°N
May
Abundance=100167
June
Abundance=98151
July
Abundance=108764
August
Abundance=91585
Animals / 100 km²
> 130
100 - 130
75 - 100
56 - 75
42 - 56
32 - 42
24 - 32
18 - 24
13 - 18
10 - 13
7.5 - 10
5.6 - 7.5
4.2 - 5.6
3.2 - 4.2
2.4 - 3.2
1.8 - 2.4
1.3 - 1.8
1.0 - 1.3
< 1.0
Marine Geospatial Ecology Lab


95°W
90°W
85°W
30°N
25°N
September
Abundance=67580
October
Abundance=76275
November
Abundance=82844
December
Abundance=78506
Animals / 100 km²
> 130
100 - 130
75 - 100
56 - 75
42 - 56
32 - 42
24 - 32
18 - 24
13 - 18
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
< 1.0
Marine Geospatial Ecology Lab

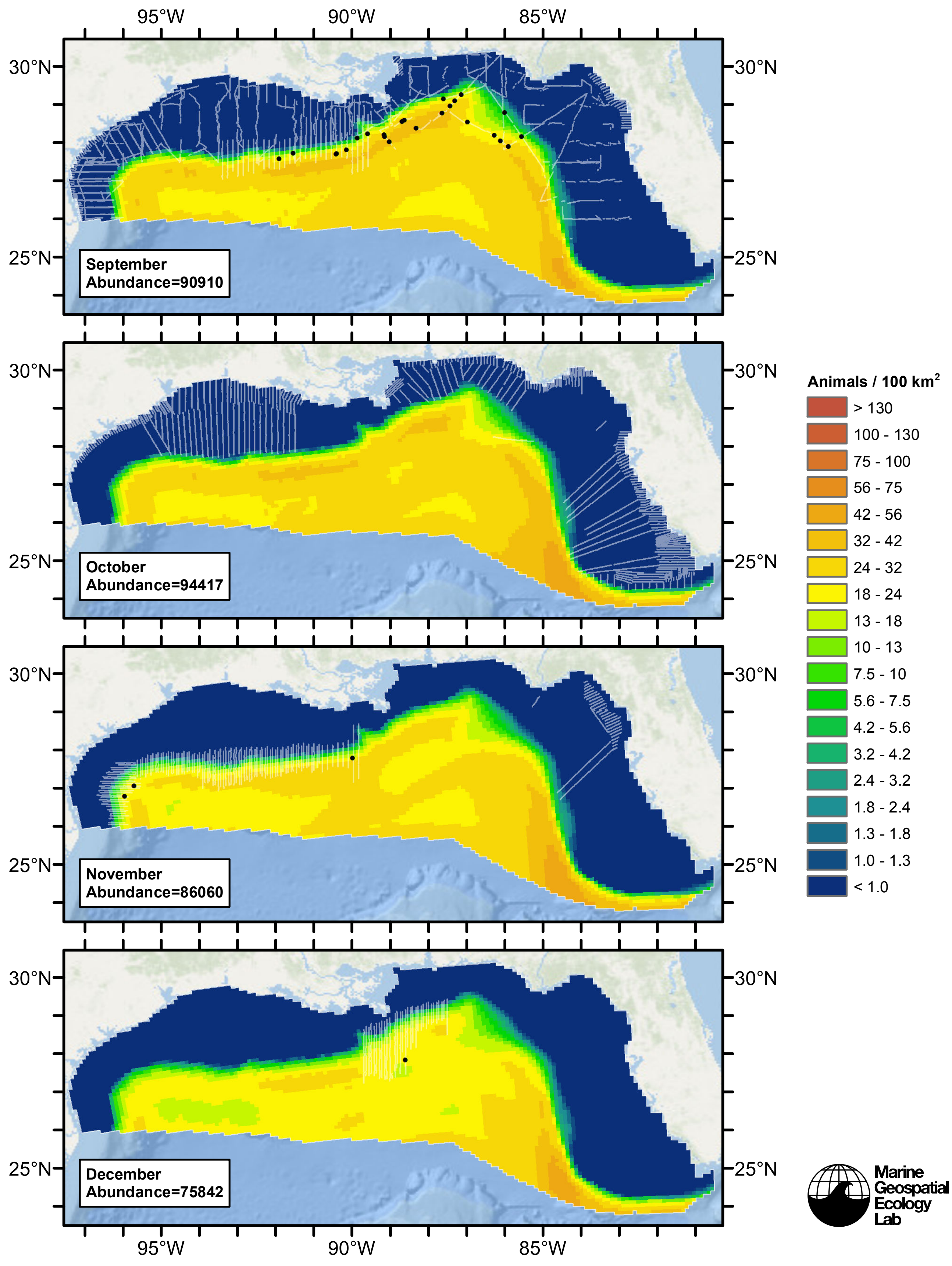
Contemporaneous Model

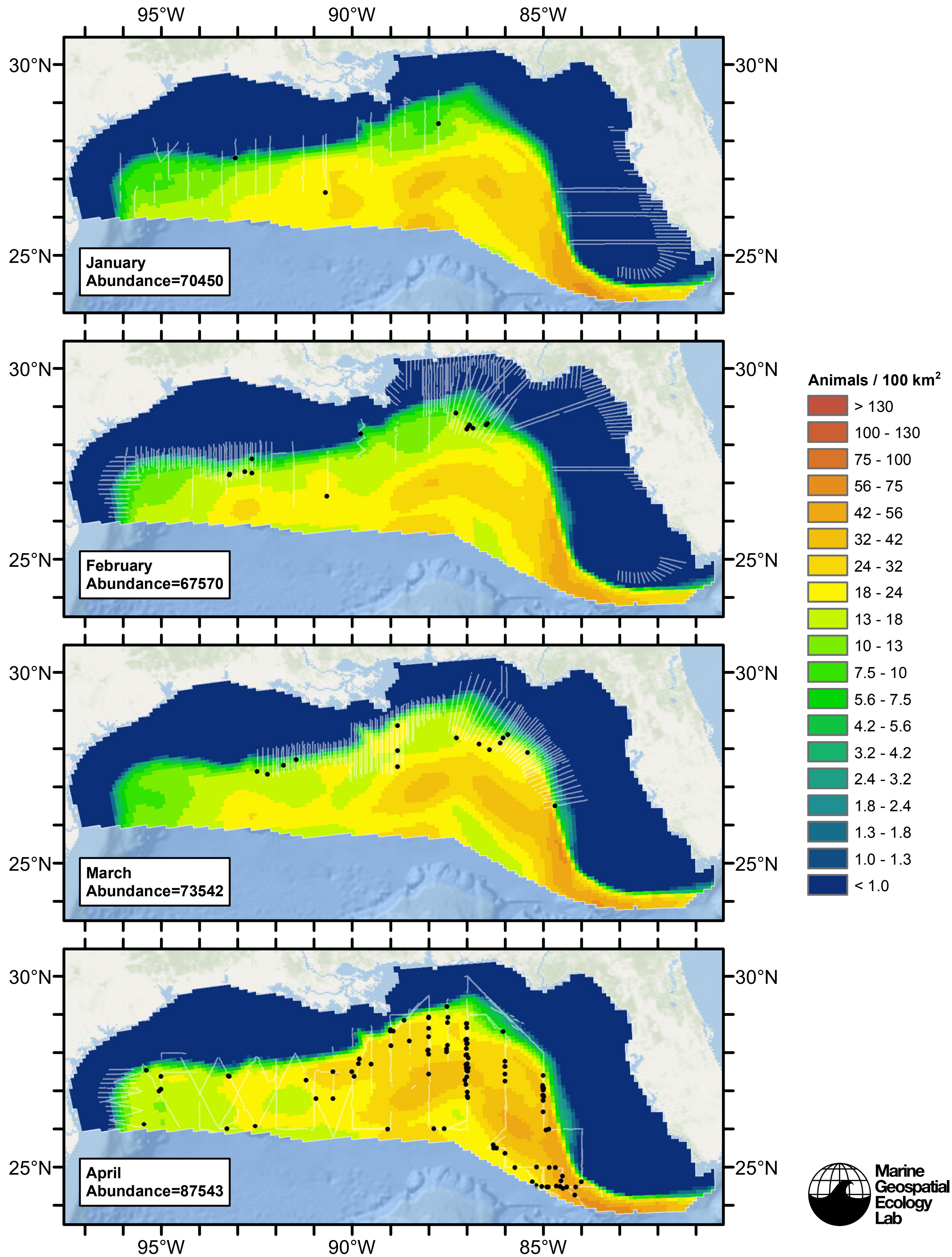


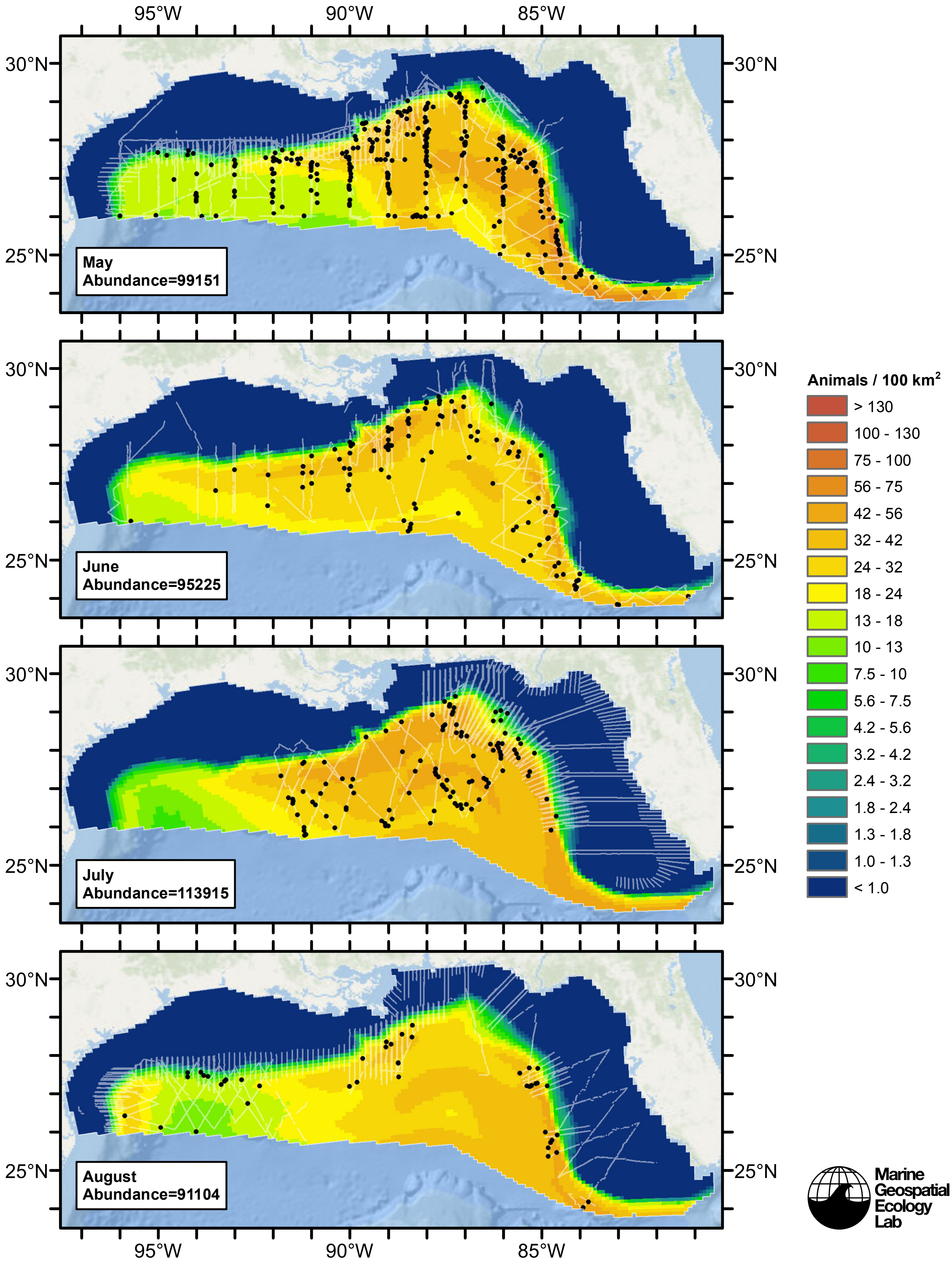
Marine Geospatial Ecology Lab

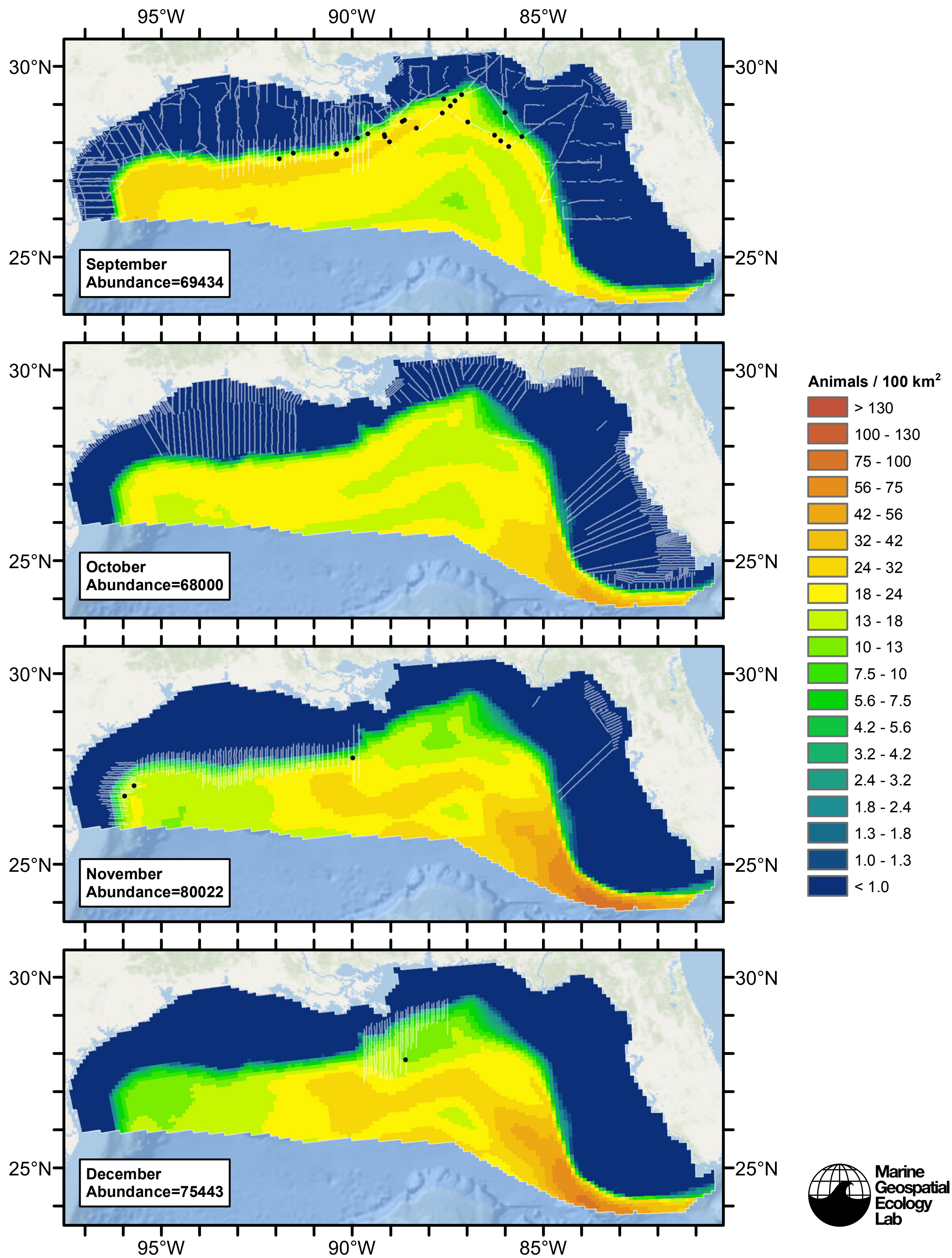


95°W
90°W
85°W
30°N
25°N
May
Abundance=90106
June
Abundance=94501
July
Abundance=96529
August
Abundance=97663
Animals / 100 km²
> 130
100 - 130
75 - 100
56 - 75
42 - 56
32 - 42
24 - 32
18 - 24
13 - 18
10 - 13
7.5 - 10
5.6 - 7.5
4.2 - 5.6
3.2 - 4.2
2.4 - 3.2
1.8 - 2.4
1.3 - 1.8
1.0 - 1.3
< 1.0
Marine Geospatial Ecology Lab


95°W
90°W
85°W
30°N
25°N
September
Abundance=90910
October
Abundance=94417
November
Abundance=86060
December
Abundance=75842
Animals / 100 km²
> 130
100 - 130
75 - 100
56 - 75
42 - 56
32 - 42
24 - 32
18 - 24
13 - 18
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
< 1.0
Marine
Geospatial
Ecology
Lab


95°W
90°W
85°W
30°N
25°N
January
Abundance=70450
February
Abundance=67570
March
Abundance=73542
April
Abundance=87543
Animals / 100 km²
> 130
100 - 130
75 - 100
56 - 75
42 - 56
32 - 42
24 - 32
18 - 24
13 - 18
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
< 1.0
Marine Geospatial Ecology Lab


95°W
90°W
85°W
30°N
25°N
May
Abundance=99151
June
Abundance=95225
July
Abundance=113915
August
Abundance=91104
Animals / 100 km²
> 130
100 - 130
75 - 100
56 - 75
42 - 56
32 - 42
24 - 32
18 - 24
13 - 18
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
< 1.0
Marine
Geospatial
Ecology
Lab



95°W
90°W
85°W
30°N
25°N
September
Abundance=69434
October
Abundance=68000
November
Abundance=80022
December
Abundance=75443
Animals / 100 km²
> 130
100 - 130
75 - 100
56 - 75
42 - 56
32 - 42
24 - 32
18 - 24
13 - 18
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
< 1.0
Marine Geospatial Ecology Lab

Discussion

When models included only physiographic covariates and covariates related to SST and ocean currents, the models fitted to climatological estimates of dynamic predictors explained slightly more deviance than models fitted to contemporaneous estimates. But when covariates related to biological productivity were introduced, the contemporaneous model explained more deviance than the climatological model fitted to the same segments. These models, however, resulted in a loss of 71% of the survey segments, which we consider an unacceptable loss of data for the marginal improvement the contemporaneous model provided over the climatological model. Thus we selected the climatological model fitted to all segments as our best estimate of pantropical spotted dolphin distribution and abundance. In any case, the total abundance estimates yielded by these models differed by less than 2% (Table 15), a statistically insignificant amount, and the spatial predictions were very similar (Fig. 43).

Because the survey effort used as input to this model was biased toward spring and summer and was spatiotemporally patchy (see maps in the Temporal Variability section above), we were not confident that our models could produce realistic predictions at a monthly temporal resolution. This problem affected all species that we modeled in the Gulf of Mexico, and we recommend that year-round average predictions be used for all Gulf of Mexico species.

Our abundance estimate of 84,014 fell within the range of NOAA's estimates, which ranged from a low of 34,067 in 2003-2004 to a high in 91,321 in 1996-2001. We believe the difference between our estimate and NOAA's lower estimates may be related mainly to two factors (assuming pantropical spotted dolphin abundances have not changed). The first factor relates to differences in detection functions. NOAA's detection functions typically had a larger effective strip half width than ours. For example, for NOAA's lowest estimate that occurred in 2003-2004, NOAA's shipboard detection function for Small Dolphins had an effective strip half width (ESHW) of 2336m (Mullin 2007). Our shipboard detection functions had mean ESHWs of 1817m and 1404m. Abundance scales inversely with ESHW, so all else being equal, our abundance estimate is expected to be proportionally larger.

A second factor, applied multiplicatively after detection functions are applied, concerns the g(0) parameter. In all of NOAA's analyses, g(0) was assumed to be 1. In our analysis, for shipboard surveys, which reported 625 of the 719 sightings we utilized, we assumed g(0)=0.856 for groups of 1-20 individuals and g(0)=0.970 for groups of more than 20. Approximately 22% of the sightings were of small groups. Abundance scales inversely with g(0), thus for these 22% of the shipboard sightings, our analysis would scale abundance up by roughly 17%. Scaling due to g(0) occurs after detection functions are applied, thus this increase would occur multiplicatively on top of any relative increase resulting from our detection functions utilizing a smaller ESHW than NOAA's.

References

Barlow J, Forney KA (2007) Abundance and density of cetaceans in the California Current ecosystem. Fish. Bull. 105: 509-526.

Baumgartner MF, Mullin KD, May LN, Leming TD (2001) Cetacean habitats in the northern Gulf of Mexico. Fishery Bulletin 99: 219-239.

Carretta JV, Lowry MS, Stinchcomb CE, Lynn MS, Cosgrove RE (2000) Distribution and abundance of marine mammals at San Clemente Island and surrounding offshore waters: results from aerial and ground surveys in 1998 and 1999. Administrative Report LJ-00-02, available from Southwest Fisheries Science Center, P.O. Box 271, La Jolla, CA USA 92038. 44 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.

Hiby L (1999) The objective identification of duplicate sightings in aerial survey for porpoise. In: Marine Mammal Survey and Assessment Methods (Garner GW, Amstrup SC, Laake JL, Manly BFJ, McDonald LL, Robertson DG, eds.). Balkema, Rotterdam, pp. 179-189.

Jefferson TA, Schiro AJ (1997) Distribution of cetaceans in the offshore Gulf of Mexico. Mammal Rev. 27(1): 27-50.

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.

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.