Mapping the distribution, abundance and risk assessment of marine birds in the Northwest Atlantic

Earvin Balderama
Dept. of Mathematics and Statistics
Loyola University Chicago
Chicago, IL 60660

Beth Gardner
Dept. of Forestry and Environmental Resources
North Carolina State University
Raleigh, NC 27605

Brian J. Reich
Department of Statistics
North Carolina State University
Raleigh, NC 27605

Abstract

Relatively little is known about the distribution, abundance and spatio-temporal variability of marine birds in their habitats off the Eastern coast of the U.S. A number of surveys have been conducted over the last 10-15 years, but analyzing these data in a unified framework is difficult due to the use of different sampling methods, spatial and temporal scales, as well as lack of sampling design. Thus, we incorporate a multi-scale approach to develop models for the space-time distribution and abundance of marine birds to identify potential high-use areas in need of further study. With data taken from past and ongoing survey efforts, we provide relative abundance and density estimates for marine birds over a wide geographical area during multiple years. Due to the excessive zero-counts and extremely large counts exhibited in the data, a double-hurdle model was formulated that includes a negative binomial and a generalized Pareto distribution mixture. Spatial heterogeneity is modeled using a conditional auto-regressive (CAR) prior, and a Fourier basis was used for seasonal variation. We demonstrate our model by creating probability maps that show areas of high-abundance and aggregation for twenty-four species of marine bird.

1 Introduction

One of the fastest-growing segments of the energy market is wind power (American Council On Renewable Energy 2014), particularly wind resources in offshore waters (American Wind Energy Association 2013). The potential impacts of offshore wind facilities on marine bird species are not well understood, and may include exposing birds to increased mortality through turbine collisions, altering their behavior and flight pathways (Drewitt and Langston 2006), and altering the habitat upon which marine birds depend for foraging. One of the first steps in evaluating potential interactions from wind energy facilities is to understand the distribution and abundance of marine birds (Huettman and Diamond 2001; Ford et al. 2004; Garthe and Huppop 2004). This information can be used in identifying sensitive and high-use areas of birds and is thus a key component of marine planning (Huettman and Diamond 2001; Ford et al. 2004; Renner et al. 2013; Winiarski et al. 2014).

Several wind energy facilities are currently being planned for offshore Atlantic waters of the United States. However, relatively little is known about the distribution, abundance and spatio-temporal variability of marine birds in their offshore habitats and it is becoming increasingly necessary to accurately characterize these demographic parameters before assessing the influence

of factors such as offshore energy development on populations (Zipkin et al. 2010; Menza et al. 2012). In response to the urgent need for this information, a number of aerial- and boat-based and individual tracking surveys have been performed in recent years or are underway to determine distribution and abundance of marine birds in both state and federal waters of the U.S. Atlantic Shelf, either in response to proposed offshore wind facilities or in anticipation of future development (e.g., NJ Department of Environmental Protection 2010, Atlantic Sea Duck Project/USGS and BDJV). Ongoing survey efforts, such as the Atlantic Marine Assessment Program for Protected Species (AMAPPS; NOAA 2011), aim to provide abundance and density estimates for marine birds over a wide geographical area during multiple years. Unfortunately, marine bird data often incur a number of issues that pose statistical challenges for modeling the counts. For example, marine bird data often display a high variance to mean ratio due to being zero-inflated, over-dispersed, or both (Zipkin et al. 2014). Additionally, large aggregations commonly reported in marine bird surveys, are often not adequately modeled (Zipkin et al. 2014).

In marine bird research and ecology more generally, count data are often analyzed using Poisson, Quasi-poisson, or negative binomial distributions or zero-inflated versions of these models (Ver Hoef and Boveng 2007; Oppel et al. 2012; McGowan et al. 2013; Renner et al. 2013) Given the level of over-dispersion that can occur in marine bird data, these distributions may not be robust enough for the data, but recent work has shown that other, less commonly used distributions, may be useful in capturing the high variance to mean ratio. Beauchamp (2011) found the power law distribution outperformed the negative binomial for modeling group sizes in seven marine bird species in the Western North Atlantic. Zipkin et al. (2014) compare and contrast a suite of distributions for sea duck data in the Western North Atlantic, finding that the discretized lognormal was the best fit over the geometric, logarithmic, zeta, Poisson, negative binomial, and Yule-Simon distributions for modeling flock sizes. Other methods for trying to account for the large variation in marine bird data is using a Box-Cox hurdle model (Menza et al. 2012), which transforms count data to be more normally distributed. The main goal of these studies was to try and best model the over-dispersion and zero-inflation of marine bird data without throwing data out, truncating, or mis-specifying the distribution, to better understand the main ecological drivers or spatial patterns of marine bird distributions.

Our objectives were to build on this foundation and further examine the extreme counts that arise in marine bird data (e.g., large aggregations, where counts can be 500-2000 birds at one location) and to explicitly account for spatial autocorrelation in marine bird data. By doing so, we aim to create more accurate predictions of sea bird distributions across a large spatial domain. To assess extreme counts, we created two models, one that uses a single hurdle for zero-inflation and one that uses a double hurdle for both zero-inflation and extreme counts. We then formulated the models using a negative binomial or generalized Pareto distribution to compare the performance of these two distributions. For the double hurdle model, we used the negative binomial for the first hurdle and the generalized Pareto for the second hurdle that examines the spatial extremes of distributions. The generalized Pareto distribution is commonly used in the spatial extremes literature (see, e.g., Buishand et al. 2008), but has not been frequently used in ecology. To account for spatial autocorrelation, we incorporated a Gaussian Markov random field into all of the models (Besag et al. 1991). Additionally, select biophysical covariates (e.g., sea surface temperature, chlorophyll-a level, ocean depth) were incorporated in to the models and used as a proxy for the variability in food resources that influence where birds occur. We evaluated model performance, different thresholds for extreme counts, and compared models for 24 different species in the western North Atlantic.

This paper is organized as follows: Section 2 gives details on how the data were collected, and what variables were considered in the study. Section 3 describes the double-hurdle model and the hierarchical Bayesian approach to estimate parameters. These estimates are used to create exposure maps in Section 4 and results are discussed in Section 5.

*This report and digital versions of the maps are available at:
http://d25ripjvlq5c77.cloudfront.net/marine_birds_full.zip

2 Marine bird data

Between 1992 and 2010, a total of 43,701 boat and aerial transects made up an avian data collection effort that spanned the Atlantic coastline from Maine to Florida. We only consider continuous time strip (CTS) surveys in our study. From these CTS surveys, over 2 million individual marine birds representing nearly 200 species were observed in 133,890 separate sightings. For each sighting, the count, species, date and location were recorded (other information is over recorded, but not consistently across surveys). Observations without a specific species name were labeled as "unknown" or "unidentified" and these counts were not included in our analyses.

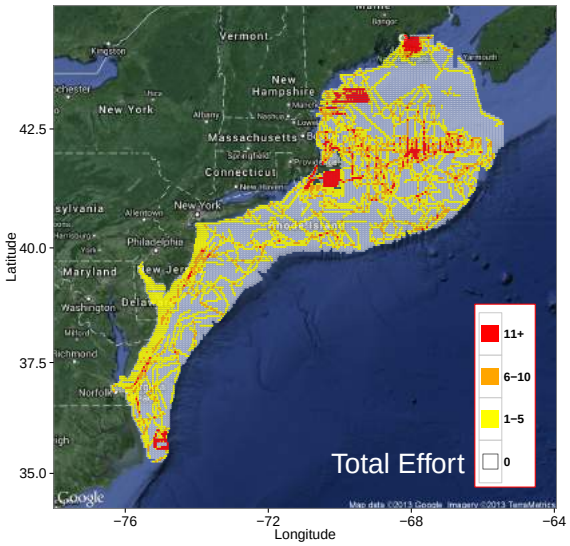

Figure 1: Total amount of survey effort, defined as the number of transects passing through the grid cell.

We created contiguous grid cells, with each cell approximately 4×4 km, over the entire Atlantic region to match the resolution of biophysical covariate information obtained from the National Oceanographic and Atmospheric Administration (NOAA). These covariates, e.g., sea surface temperature, chlorophyll-a level, and ocean depth, were used as a proxy for the variability in food resources that may influence where we expect birds occur (Zipkin et al. 2010). This will allow us to explore the relationships between marine birds and biophysical variables and enhance our understanding of how birds are distributed throughout the marine environment. We focus our analyses on data accumulated from July 2002 to November 2010 (101 months), on 15,984 grid cells which satisfy the following constraints: north of 35.25° latitude, east of −76.5° longitude, and with an ocean depth of no more than 500m. This definition of our study area tries to maintain a low number of grid cells while surrounding most of the survey effort. Because the data are very sparse, observations which occurred in the same month and year were combined to give total monthly counts for each grid cell. Thus, counts may fall into one of 15,984×101=1,614,384 space-time grid cells.

Sea surface temperature and chlorophyll-a concentration were remotely sensed via the NASA Aqua satellite and acquired through Marine Geospatial Ecology Tools (Roberts et al. 2010). Average monthly values of SST and CHL were gathered, then standardized using the overall mean and standard deviation across all 15,984 spatial grid cells and across all sixty-eight months containing at least one survey. The overall mean SST is $13.4°C$ with a standard deviation of $6.5°C$. CHL has mean 1.7mg/m^3 and standard deviation 2.6mg/m^3. Fine-scale resolution bathymetry data were acquired through NOAA's GEODAS Grid Translator, and average depth

*This report and digital versions of the maps are available at:
`http://d25ripjvlq5c77.cloudfront.net/marine_birds_full.zip`

values for each grid cell were determined. The distances to the nearest shore from the centroid of each grid cell were calculated using the `maps` package in R (R Core Team 2014).

There are no detection parameters in our models because of the lack of a universal resampling technique, so survey effort is used as an offset term (see Section 3). Using the start and end locations of each transect, effort is counted in a grid cell for every transect line passing through that cell. A map of total effort is shown in Figure 1. This map combines all surveys in the study. Additional maps of survey effort, including monthly survey maps and corresponding observed point locations of individual species, are found in the appendix. At least one survey within the study area was performed in 68 of the 101 months of data. Out of the 1,614,384 space-time grid cells, only 14,007 had non-zero effort. For these, the average amount of effort was 3.2 transects, with a minimum of 1 and a maximum of 63. Large amounts of effort can be explained from the heavily surveyed areas in and around Nantucket Sound which were carried out in the early 2000's in anticipation of the Cape Wind wind energy project.

We create models for the twenty-four most abundant species in our dataset, each with at least 196 total sightings. Table 1 gives, for each species, the total number of non-zero observations and summary statistics based on the positive count distribution, including the mean, standard deviation, median, maximum, and skewness. Fitted models are discussed in subsequent sections.

3 Methods

3.1 Hurdle models

Zero-inflated models are widely used for count data exhibiting more zeros than would be expected given some typical count distribution (see, e.g., Ridout et al. 1998; Winkelmann 2008). These are simply two-component mixture models that can separate the excess zeros in the data from the underlying (parametric) count distribution (Greene 2007). A common choice for the count distribution is Poisson (Lambert 1992), but is typically replaced with negative binomial if there is over-dispersion in the data (Bliss and Fisher 1953; Greene 1994). Hurdle models are similar to zero-inflated models in that they also account for the high occurrence of zeros in data. The difference is that for hurdle models, *all* zero-count observations are assumed to have been generated by a mechanism separate of the non-zeros (Mullahy 1986; Gurmu 1998; Zorn 1998). For example, when observed counts are i.i.d., the negative binomial hurdle model, denoted h_1, can be written as

$$h_1(y \,|\, \mu, r) = \begin{cases} p & \text{if } y = 0, \\ [1-p] \cdot f_{zt}(y \,|\, \mu, r) & \text{if } y \geq 1, \end{cases} \tag{1}$$

where $p \in (0, 1)$ is the probability of a zero count and $f_{zt}(y \,|\, \mu, r)$ is the zero-truncated negative binomial probability mass function (Sampford 1955), with $\mu > 0$ and $r > 0$ as the mean and dispersion parameters, respectively, of the un-truncated negative binomial distribution.

Surveys are assumed to observe and record any species, so a count of 0 indicates an *observed* 0, not a missing observation, so a hurdle model is reasonable. However, the count data exhibits not only zero-inflation, but also extreme over-dispersion. For example, data on the northern gannet consist of 11,759 zero-counts and 2,248 non-zero counts that range from 1 to 1,775. The mean and median of the non-zero counts are 6.06 and 2, respectively, thus indicating an extremely right-skewed distribution. A simple negative binomial hurdle model provides an inadequate fit due to these extreme counts. However, modeling large observations above a threshold is standard practice in extreme value analysis (Coles et al. 2001). Thus, we propose attaching a second hurdle to accommodate these extremes. Just as in the hurdle model, we assume separate data-generating processes for each of the components of our model. The first hurdle represents the divide between the zero counts and the "typical" non-zero counts. Larger counts can be thought of as coming from a heavy-tailed distribution if a second hurdle is crossed.

Double-hurdle models (Cragg 1971) have been widely used in economics and social sciences (see, e.g., Jones 1989; Labeaga 1999; Miranda 2010; Ricker-Gilbert et al. 2011) to model

Table 1: Summary statistics for the twenty-four species, based on the positive count observations for each species. The "best model" for each species is the result from the model comparison framework in Section 4.1.

Species	#Obs.	Mean (SD)	Median	Max	Skewness	Best model
Atlantic Puffin	248	2.0 (1.5)	1	10	2.2	Double-hurdle
Black-legged Kittiwake	738	4.9 (9.3)	2	105	5.4	Double-hurdle
Black Scoter	540	23.1 (56.0)	6	605	5.3	Double-hurdle
Bonaparte's Gull	376	5.8 (18.3)	2	262	9.6	Double-hurdle
Common Eider	1432	572.1 (3019.8)	15	50025	9.4	GPD-hurdle
Common Loon	1319	3.3 (3.9)	2	40	3.8	Double-hurdle
Common Tern	809	11.8 (51.7)	3	1094	14.6	Double-hurdle
Cory's Shearwater	634	4.9 (16.9)	2	266	10.7	Double-hurdle
Double-crested Cormorant	232	13.9 (43.9)	2	501	7.5	Double-hurdle
Dovekie	550	7.6 (17.6)	3	299	10.1	Double-hurdle
Great Black-backed Gull	3188	4.8 (25.8)	2	1300	40.5	Double-hurdle
Great Shearwater	3195	12.2 (35.8)	4	950	11.5	Double-hurdle
Herring Gull	3249	5.7 (31.0)	2	1300	28.5	Double-hurdle
Laughing Gull	464	3.2 (6.3)	2	88	8.1	Double-hurdle
Leach's Storm-petrel	840	6.2 (21.3)	2	345	9.6	Double-hurdle
Long-tailed Duck	1443	94.0 (432.6)	17	11000	18.3	GPD-hurdle
Northern Fulmar	1330	7.8 (43.2)	2	1352	24.4	Double-hurdle
Northern Gannet	2248	6.1 (41.7)	2	1775	35.3	Double-hurdle
Razorbill	1002	10.7 (19.7)	4	293	5.9	Double-hurdle
Roseate Tern	196	7.1 (16.8)	2	137	5.5	GPD-hurdle
Sooty Shearwater	729	9.4 (38.8)	2	700	11.9	Double-hurdle
Surf Scoter	1135	60.2 (146.3)	15	1400	5.1	Double-hurdle
White-winged Scoter	885	24.5 (73.1)	4	1027	7.6	Double-hurdle
Wilson's Storm-petrel	1790	13.4 (92.2)	2	3061	23.3	Double-hurdle

a consumer's two-part decision-making process. For instance, a consumer must first decide whether or not to purchase an item, and then must decide on the amount to purchase conditional on the decision to purchase an item. In the context of marine bird data, the first decision is attributed to whether or not any number of individuals are observed at a certain site, and the second decision is attributed to whether the number of individuals observed is considered typical or large, conditional on observing a positive number of individuals. In this paper, we consider the generalized Pareto distribution (GPD) to describe those larger-than-typical observations of the data (Hosking and Wallis 1987). If $y \sim \mathsf{GPD}(\psi, \sigma, \xi)$, then y has distribution function

$$G(y \,|\, \psi, \sigma, \xi) = 1 - \left[1 + \xi \cdot \frac{(y - \psi)}{\sigma}\right]^{-1/\xi} \tag{2}$$

for $y \geq \psi$ and $\xi \neq 0$. The GPD is bounded below by the threshold parameter ψ, which defines the second hurdle that must crossed in order for the observation to be considered "large." Thus, $G(y \,|\, \psi, \sigma, \xi) = 0$ for $y < \psi$. Since we are dealing with count data, we use a discretized version of the GPD distribution function, defined by taking the difference between the cumulative densities at $y + 0.5$ and $y - 0.5$. The negative binomial component is now right-truncated at ψ in addition to being left-truncated at 0. It is normalized to maintain separability in the likelihood and ensure that each component is a density. The probability mass function can be expressed as

$$f_t(y) = \frac{f(y)}{F(\psi - 1) - f(0)}, \tag{3}$$

where f and F are the density and distribution functions, respectively, of the un-truncated negative binomial distribution.

For a particular species, let y_{ij} be the number of individuals observed at spatial location i and calendar month j for $i = 1, \ldots, 15984$ and $j = 1, \ldots, 12$. We run the model with only the $N = 14{,}007$ space-time locations of non-zero effort. Later, we use model results to form

predictions on the un-surveyed spatial locations. The double-hurdle model, denoted h_2, has the following form:

$$h_2(y_{ij}\,|\,\boldsymbol{\theta}) = \begin{cases} p_{ij} & \text{if } y_{ij} = 0, \\ \Big[1 - q_{ij}\Big] \cdot \Big[1 - p_{ij}\Big] \cdot f_t(y_{ij}\,|\,\mu_{ij}, r) & \text{if } 1 \le y_{ij} < \psi, \\ q_{ij} \cdot \Big[1 - p_{ij}\Big] \cdot \Big[G\big(y_{ij} + \frac{1}{2}\,|\,\psi, \sigma, \xi\big) - G\big(y_{ij} - \frac{1}{2}\,|\,\psi, \sigma, \xi\big)\Big] & \text{if } y_{ij} \ge \psi, \end{cases} \quad (4)$$

where $\boldsymbol{\theta} = \{p_{ij}, q_{ij}, \mu_{ij}, r, \psi, \sigma, \xi\}$ is the set of model parameters, p_{ij} is the probability that site i in month j is governed by a zero-generating process, and q_{ij} is the probability that site i in month j produces a count at or above the GPD threshold conditional on it being a non-zero count. Notice that when $\psi = \infty$, (4) simplifies to (1), showing that the negative binomial hurdle model is a special case of our double-hurdle model. By setting $\psi = 1$, we remove the second component and model all non-zero counts using a GPD. We define this as a GPD-hurdle model, which may be used as another single-hurdle alternative when the negative binomial distribution is inadequate.

The GPD threshold is usually fixed before the estimation of all other parameters (Coles and Tawn 1996; Reiss et al. 2001), although treating ψ as an additional unknown parameter has been examined (Behrens et al. 2004; Tancredi et al. 2006). When fixing ψ, the choice is often determined empirically, through the use of mean residual life plots (Davis and Resnick 1984; Raqab and Asadi 2008), or by a high percentile level (DuMouchel 1983), as we do in this paper.

3.2 Spatial hierarchical regression

For simplicity in expressing the forthcoming regression equations, denote $\boldsymbol{\mu} = \{\mu_{ij} : z_{ij} > 0\}$ as the $N \times 1$ vector of NB means of the $N = 14{,}007$ space-time locations with non-zero effort. Variation in the mean can be explained through a log-linear spatial regression model

$$\log(\boldsymbol{\mu}) = \log(\mathbf{z}) + \mathbf{X}\boldsymbol{\beta} + \mathbf{S}_{[\mu]}, \quad (5)$$

where the vector $\mathbf{z} = \{z_{ij} : z_{ij} > 0\}$ is the amount of survey effort, used as an offset term, $\mathbf{X}$ is the $N \times (k+1)$ matrix of k site-level covariates plus a column vector of 1's, with corresponding vector of coefficients $\boldsymbol{\beta}$, and $\mathbf{S}_{[\mu]}$ is a Gaussian process that accounts for spatial autocorrelation (discussed in the next subsection). Similarly, we allow the probabilities $\mathbf{p} = \{p_{ij} : z_{ij} > 0\}$ and $\mathbf{q} = \{q_{ij} : z_{ij} > 0\}$ to vary according to the same covariate information, but estimate separate sets of coefficients, denoted as $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$, respectively. For these parameters, we use the logit link such that

$$\text{logit}(\mathbf{p}) = \mathbf{X}\boldsymbol{\gamma} + \mathbf{S}_{[\text{p}]} \qquad \text{and} \qquad \text{logit}(\mathbf{q}) = \mathbf{X}\boldsymbol{\delta}.$$

We do not model the probability vector $\mathbf{q}$ with spatial random effects because initial tests show that large right-tail count values vary independently of spatial location, and crossing the second hurdle already relies on the spatially-varying values of $1 - \mathbf{p}$.

Environmental and biophysical covariate information, as well as temporal random effects, are incorporated in the design matrix $\mathbf{X} = \{\mathbf{1}, \mathbf{x}_1, \ldots, \mathbf{x}_k\}'$. The site-level covariates we include in our models are $\mathbf{x}_1$ = sea surface temperature (SST), $\mathbf{x}_2$ = ocean depth (DEP), $\mathbf{x}_3$ = chlorophyll-a concentration (CHL), and $\mathbf{x}_4$ = distance to shore (DTS). Because marine bird counts across our time window are expected to be periodic due to similar monthly and seasonal patterns year after year, we model temporal variability using a Fourier basis expansion of the calendar month number (MON). This goes into our model as $\mathbf{x}_5 = sin(\frac{\pi}{6} \cdot \text{MON})$ and $\mathbf{x}_6 = cos(\frac{\pi}{6} \cdot \text{MON})$. Covariates were then standardized to have mean 0 and standard deviation 1 before extracting the 14,007 space-time areas for analyses. A sine curve, with a positive coefficient, is concave down from January to June and concave down from July to December, then reconnects with January of the following year to give it its cyclic pattern. The degree of concavity depends on the coefficient. A negative coefficient will reverse the concavity of the curve. The points of inflection is fixed at zero, thus the need to add a cosine function, which has inflection points

at the extrema of the sine function. Combining the sine and cosine curves will yield a flexible model for month-to-month variation. We include only the first sine and cosine pair to avoid the possibility of overfitting.

3.3 Dimension reduction for spatial random effects

Spatial random effects may be modeled explicitly by a Gaussian Markov random field that accounts for spatial variability using the neighborhood structure of sites (Besag 1974; Besag et al. 1991). This is specified by the conditional autoregressive (CAR) prior distribution

$$\pi(\mathbf{S} \,|\, \tau) \propto \tau^{\text{rank}(\mathbf{Q})/2} \exp\Big(-\frac{\tau}{2}\mathbf{S}'\mathbf{Q}\mathbf{S}\Big), \tag{6}$$

where $\mathbf{Q}$ is the CAR precision matrix and τ is a smoothing parameter. For the CAR prior, $\mathbf{Q} = \mathbf{D} - \rho\mathbf{A}$, where $\mathbf{A}$ is the $n \times n$ adjacency matrix with entries $A_{il} = 1$ if site i is a neighbor of site l and 0 otherwise, $\mathbf{D}$ is diagonal with entries $D_{ii} = n_i = \sum_{l=1}^{n} A_{il}$, and $\rho \in [0, 1]$ measures the degree of spatial association with $\rho = 0$ implying no spatial dependence. By letting $\rho = 1$, $\mathbf{Q}$ becomes singular, and this specifies the intrinsic CAR (ICAR) prior, with full conditional distributions

$$S_i \,\Big|\, S_{-i}, \mathbf{A}, \tau^2 \sim \mathsf{Normal}\left(\frac{1}{n_i}\sum_{l\sim i} S_l, \frac{\tau^2}{n_i}\right). \tag{7}$$

When the number of sites n is large, the number of spatial model parameters can flood the model with unnecessary estimation and computing time. In our analysis, the spatial domain consists of n = 15,984 sites, so the covariance matrix has over 255 million entries. The study of spatial generalized linear mixed models (SGLMMs) has developed methods for reducing the dimensionality of spatial random effects by using an eigen-decomposition of the covariance matrix, and using the first $w < n$ eigenvectors that capture the most spatial variability, i.e., those with the largest corresponding eigenvalues (Boots and Tiefelsdorf 2000; Saul et al. 2006).

We can reparameterize the spatial regression model such that

$$\underset{n\times 1}{\mathbf{S}} = \underset{n\times n}{\mathbf{V}} \cdot \underset{n\times 1}{\boldsymbol{\alpha}} \approx \underset{n\times w}{\mathbf{V}} \cdot \underset{w\times 1}{\boldsymbol{\alpha}}, \tag{8}$$

where $\mathbf{V}$ is the set of eigenvectors coming from the eigen-decomposition of $\mathbf{Q} = \mathbf{D} - \mathbf{A}$, and $\boldsymbol{\alpha}$ is the respective set of coefficients. The first five eigenvectors explains a little more than 50% of spatial variability. In order to account for 70% of the variability, 125 eigenvectors are needed, drastically increasing the number of spatial parameters while minimally increasing the amount of variability explained. We use the first fifty eigenvectors to explain 67% of the variability, which seems like a reasonable compromise between computing speed and information, for data with more than 1000 observations. The $w = 50$ eigenvectors can now be treated as spatial random effects in the regression equations, adding only 50 extra regression parameters per equation to be estimated hierarchically. We chose to model the less abundant species (those with less than 1000 observations) with five eigenvectors, as we did not want to overparameterize.

3.4 Bayesian estimation

Parameter estimation under a Bayesian framework is suitable because of the model's hierarchical structure (Gelman et al. 2003). Model parameters were given uninformative priors and updated one-at-a-time using Markov chain Monte Carlo (MCMC). Initial simulation results showed that, due to extreme overdispersion in some species data, the negative binomial dispersion parameter r approached zero, which prevented other parameters from also converging. To remedy this, we set $r = 1$ for all simulations, thus specifying a (truncated) geometric distribution with (untruncated) mean μ_{ij} and variance $\mu_{ij} + \mu_{ij}^2$ (Hilbe 2011). A Gibbs update was used for τ, and Metropolis-Hastings updates were used for all other parameters. Each of the regression coefficients in $\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}$ was given a $\mathsf{Normal}(0, 10^4)$ prior, while $\sigma \sim \log \mathsf{Normal}(5, 100)$ and $\xi \sim \mathsf{Normal}(0, 1)$. The

conditional distribution for $\boldsymbol{\alpha}$ is $\pi(\alpha_j \,|\, \tau) \overset{ind}{\sim} \mathsf{Normal}(0, \tau\nu_j)$, where $(\tau\nu_j)^{-1}$ is the variance and ν_j is the eigenvalue associated with the j^{th} eigenvector in $\mathbf{V}$. For the full conditional distribution of τ, $\pi(\tau \,|\, \boldsymbol{\alpha}) \propto \pi(\boldsymbol{\alpha} \,|\, \tau) \cdot \pi(\tau)$, we let the prior $\pi(\tau) \sim \mathsf{Gamma}(a, b)$ with shape $a = 0.5$ and rate $b = 0.0005$. The Gibbs update for τ is sampling from the posterior distribution

$$\pi(\tau \,|\, \boldsymbol{\alpha}) \sim \mathsf{Gamma}\left(\frac{w-1}{2} + a,\ \frac{1}{2}\sum_{j=1}^{w-1}(\nu_j \alpha_j^2) + b\right). \tag{9}$$

All data analyses were carried out using R (R Core Team 2014). The MCMC sampler for 200,000 iterations with the first 50,000 discarded as burn-in. Trace plots of posterior distributions after thinning were visually inspected for convergence. Trace plots and histograms of posterior distributions of all model parameters, can be found in the appendix. The appendix also includes boxplots and tables that inform of significant environmental covariates for each species.

4 Results

4.1 Model Comparison

For each species, we run hurdle models based on several GPD threshold values, and whether spatial random effects are included within the hierarchical framework or not. We use two common Bayesian model checking metrics. The deviance information criterion (DIC) tests how well the model fits the data based on deviances (Spiegelhalter et al. 2002). The log-pseudo marginal likelihood (LPML) is a model selection criterion using posterior predictive densities (Geisser and Eddy 1979). Smaller values for DIC and larger values of LPML indicates a more desired model. In all cases, with everything else being equal, the inclusion of $\mathbf{S}$ offers a better overall fit than its exclusion from the model. We consider the 97.5th and 99th percentiles for each species count distribution as threshold values, unless any of these values are below five, in which case we simply set $\psi = 5$ accordingly (See Table 2). Although these percentile values may differ from species to species, we are fixing the number of observations attributed to the GPD component. Thus, for each of these models, what is considered a "large" count depends on the species; rounding up to five is warranted by many of the less-abundant species. For instance, the 99th percentile of the Roseate Tern counts is one, while the 99th percentile of the Common Eider counts is 680.

In addition to the 97.5^{th} and 99^{th} percentiles, we consider the cases where $\psi = 1$ and $\psi = \infty$. These threshold values result in special cases of our double-hurdle formulation. Specifically, $\psi = 1$ characterizes a single-hurdle model with a generalized Pareto count distribution, and $\psi = \infty$ characterizes the well-known negative binomial hurdle model. For additional comparison, we also consider a log-normal hurdle, in which a discretized log-normal distributional form is used in place of negative binomial for the case when $\psi = \infty$. Table 2 compares models based on DIC, for each species, according to different threshold values and number of eigenvectors w used for the spatial random effect $\mathbf{S}$. Basing the comparison on the LPML criterion gives the same results.

Highly abundant species (those with $\geq$1000 sightings) are modeled with fifty spatial eigenvectors, while the rest were modeled with five. However, parameter estimates for sea duck models containing fifty eigenvectors did not converge well. Although the spatial model with fifty eigenvectors resulted in a lower DIC and larger LPML than the spatial model with five eigenvectors, these values can not be trusted because of the issue with convergence. We therefore only consider five eigenvectors in the spatial models for Common Eider, Long-tailed Duck, and Surf Scoter, which improved mixing in MCMC simulations.

The motivation for introducing a second hurdle is that the count distribution typically used in zero-inflated and single-hurdle models, such as Poisson or negative binomial, provides an inadequate fit (Zipkin et al. 2014). In order to account for the excessive overdispersion caused

Table 2: DIC values for models according to the form of the non-zero count distribution with corresponding threshold value ψ, and number of spatial eigenvectors w. When $w = 0$, this specifies the non-spatial model, which excludes the spatial random effect $\mathbf{S}$ in the regression of $\mathbf{p}$ and $\boldsymbol{\mu}$. For each species, the value of DIC of the best model is in **bold**.

Species	w	**Model for non-zero counts (ψ)**			
		GPD (1)	Double-hurdle	NB (∞)	log-Normal (∞)
Atlantic Puffin	5	2674	**2660** (5)	2685	2762
	0	2825	2810 (5)	2835	2912
Black-legged Kittiwake	5	7230	**7175** (5); 7190 (6)	7906	7422
	0	7597	7543 (5); 7558 (6)	8382	7837
Black Scoter	5	7114	**7085** (5); 7183 (17)	8344	7305
	0	7323	7273 (5); 7460 (17)	8903	7703
Bonaparte's Gull	5	4139	**4121** (5)	4549	4235
	0	4307	4289 (5)	4747	4403
Common Eider	5	**19140**	20424 (89); 20470 (680)	24807	NA
	0	20749	21127 (89); 22104 (680)	26512	21675
Common Loon	50	11270	10883 (5); **10880** (7)	10935	11133
	0	12092	11730 (5); 11750 (7)	11868	12070
Common Tern	5	9184	**9116** (5); 9203 (10)	10188	9305
	0	9232	9165 (5); 9266 (10)	10483	9476
Cory's Shearwater	5	6717	**6716** (5)	7227	6718
	0	6951	6950 (5)	7746	7052
Double-crested Cormorant	5	3230	**3227** (5)	3621	3389
	0	3230	3236 (5)	3650	3412
Dovekie	5	6220	6186 (5); **6185** (7)	6683	6254
	0	6235	6201 (5); 6200 (7)	6737	6270
Great Black-backed Gull	50	26065	**25903** (7); 26042 (17)	28169	26854
	0	26773	26661 (7); 26901 (17)	29776	NA
Great Shearwater	50	28939	**28383** (23); 28666 (52)	29666	28725
	0	29864	29530 (23); 29920 (52)	31474	29941
Herring Gull	50	26710	**26506** (8); 26758 (20)	28475	27211
	0	27515	27354 (8); 27758 (20)	30873	28422
Laughing Gull	5	4703	**4673** (5)	4871	4752
	0	5089	5060 (5)	5298	5152
Leach's Storm-petrel	5	7767	**7740** (5)	8879	7874
	0	7871	7845 (5)	9106	7999
Long-tailed Duck	5	**17087**	17090 (64); 17272 (200)	18612	17863
	0	18759	18775 (64); 19005 (200)	20870	19664
Northern Fulmar	50	12029	**11801** (5); 11844 (12)	12926	11818
	0	12866	12639 (5); 12769 (12)	14431	12783
Northern Gannet	50	20323	**20268** (6); 20475 (13)	22604	20703
	0	20931	20892 (6); 21228 (13)	23969	21586
Razorbill	50	10392	**10250** (7); 10264 (20)	10266	10698
	0	11566	11429 (7); 11510 (20)	11484	12017
Roseate Tern	5	**2435**	2483 (5)	2577	2498
	0	2535	2583 (5)	2664	2600
Sooty Shearwater	5	7222	7220 (5); **7198** (6)	8703	7373
	0	7264	7262 (5); 7266 (6)	9109	7477
Surf Scoter	5	14398	**14343** (30); 14355 (113)	14970	14783
	0	15010	14977 (30); 15047 (113)	15673	15544
White-winged Scoter	5	10063	**10002** (7); 10005 (31)	10882	10101
	0	10913	10851 (7); 10890 (31)	11743	10989
Wilson's Storm-petrel	50	16287	**16238** (6); 16495 (20)	19807	16592
	0	16700	16672 (6); 17062 (20)	22933	17409

by long right tails, the GPD is introduced as the third component in the double-hurdle likelihood. Observed counts at or above a certain threshold value (ψ) are assumed to have been generated

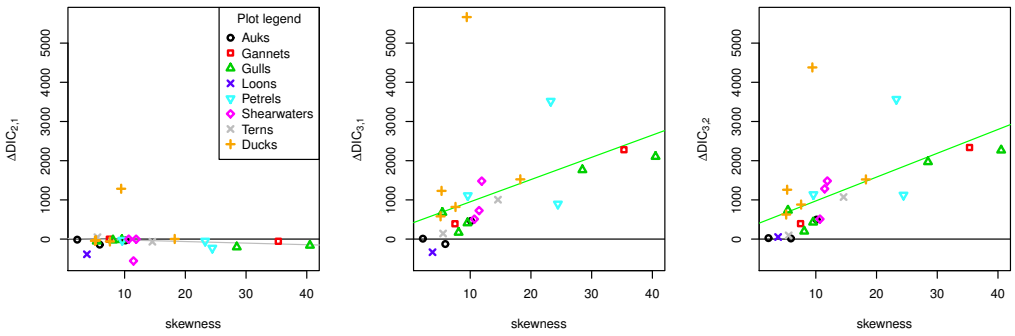

Figure 2: Effect of data skewness on choice of hurdle model. Differences between three types of hurdle models (1. GPD-hurdle, 2. double-hurdle, 3. NB-hurdle) are measured by ΔDIC. Left plot: Positive values support choosing the GPD-hurdle over the double-hurdle. Middle plot: Positive values support choosing the GPD-hurdle over the NB-hurdle. Right plot: Positive values support choosing the double-hurdle over the NB-hurdle.

by the GPD component. However, we found that for some species, it may not be necessary to include a second hurdle. As shown in Table 2, the double-hurdle model does indeed offer a better fit, compared to its single hurdle counterparts, for 21 out of 24 species in our study, with the second hurdle placed near the 97.5^{th} percentile of the distribution. For the other three species (Common Eider, Long-tailed Duck, and Roseate Tern), the best model was the GPD-hurdle.

The log-normal hurdle model never outperformed the best double-hurdle model in each species. Furthermore, the log-normal hurdle model beats the GPD-hurdle for only Common Loon and Great Shearwater. However, it outperformed the NB-hurdle in all cases except Atlantic Puffin and Razorbill, where the log-normal hurdle model offered the worst fit, and Common Loon, where it offered a worse fit than NB-hurdle but a better fit than GPD-hurdle.

Interestingly, if one were to only select from the single-hurdle specifications, the GPD distribution offers a better fit than both negative binomial and log-normal in most cases; for two species (Common Loon and Razorbill), an NB-hurdle is preferred, while for another two species (Great Shearwater and Northern Fulmar), a log-normal hurdle model is preferred.

Skewness plays a role in determining whether a model that includes a GPD component is more desirable than a simple negative binomial hurdle model. In general, the greater the skewness of the data, the less the negative binomial hurdle model is preferred. Figure 2 illustrates this point by plotting the skewness of the (positive) count data versus the difference in DIC values (ΔDIC) between two given models. In Figure 2b and 2c, positive ΔDIC values indicates more preference towards a GPD-hurdle and a double-hurdle model, respectively. A simple linear regression in these two plots show a significant positive relationship between skewness and ΔDIC. Figure 2a shows that skewness does not significantly determine the choice between a GPD-hurdle and a double-hurdle model.

4.2 Significant predictors

The relationship between covariates and the model parameters **p** and **q** are shown in Tables 3 and 4, respectively. The cells in Table 3 are marked with the direction of significance with nonzero-count probability $1 - \mathbf{p}$ (instead of **p**) for better interpretation and comparison with large-count probability **q**. To find significant relationships with **p**, the signs can simply be reversed. For most species, higher sea surface temperature indicated a lower probability of observing at least one individual of that species. However, most gulls and terns seem to have a

Table 3: Significant covariates affecting $1 - \mathbf{p}$, the probability of a non-zero count. Plus signs (+) indicate that an increase in the value of the covariate significantly *increases* occupancy probability, while negative signs (−) indicate that an increase in the value of the covariate significantly *decreases* occupancy probability. Blanks indicate non-significance at the 0.05 level.

Species	γ_1 SST	γ_2 DEP	γ_3 CHL	γ_4 DTS	γ_5 SIN	γ_6 COS
Atlantic Puffin	−	−	−			
Black-legged Kittiwake					+	+
Black Scoter	−	−			−	
Bonaparte's Gull	+	−		−		+
Common Eider	−	−	+	−	−	+
Common Loon	−		−		−	−
Common Tern		−		−	−	−
Cory's Shearwater	−	−	−	+	−	−
Double-crested Cormorant	−		+	−	−	−
Dovekie		+	−	+	+	+
Great Black-backed Gull						
Great Shearwater	−	−	−	+	−	−
Herring Gull		−	+		−	+
Laughing Gull	+	−	−	−		
Leach's Storm-petrel	−	+	−	+	−	−
Long-tailed Duck	−	−	+		−	
Northern Fulmar	−	−	−	+	−	
Northern Gannet	−	−			−	+
Razorbill	−			+	+	+
Roseate Tern	+	−		−		−
Sooty Shearwater	−		−	+	−	−
Surf Scoter	−	−		+	−	−
White-winged Scoter	−	−	+		−	
Wilson's Storm-petrel	−		−	+	−	−

different effect. For Bonaparte's Gull, Laughing Gull and Roseate Tern, cooler temperatures indicated a higher probability of observing an individual from these species. Sea surface temperature was not found to significantly affect the probability of observing Black-legged Kittiwake, Common Tern, Dovekie, Great Black-back Gull or Herring Gull.

Depth has a negative relationship with $1-\mathbf{p}$ for most species (see Table 3, γ_2 column), which suggests a higher probability of occupancy in shallower waters. This was expected given that many species such as surf scoters, white-winged scoters, and common eider are generally thought to prefer coastal areas with shallower depths (Guillemette et al. 1993; Lewis et al. 2008). Only Dovekie, which winters far out at sea, and Leach's Storm-petrel, which often feeds far at sea, were found to have a positive relationship with ocean depth.

Almost all species exhibit significant temporal variation in occupancy probability, with Atlantic Puffin and Laughing Gull as exceptions. The periodic pattern can be inferred from the signs of the coefficients, γ_5 and γ_6, of the sine and cosine variables, respectively. Species with a significant negative sine coefficient, such as the Long-tailed Duck, will generally have peak contribution to occupancy probability during the summer or early autumn months. If in addition the cosine coefficient is significantly positive, as in the case for Northern Gannet, the peak is delayed towards late autumn. Positive sine coefficients suggests that occupancy probability is at its peak during the winter months. Figure 3 illustrate the differences in temporal variation for certain combinations of γ_5 and γ_6. For the Black-legged Kittiwake species, we see that with all other variables fixed, occupancy probability in a given area is highest during December and January and lowest during June and July. For the Wilson's Storm-petrel species, the signs of these coefficients are both negative, thus having the opposite effect: high occupancy probability in June–July and low occupancy probability in December–January.

Table 4 shows which covariates were found to have a significant effect on the extreme counts. While sea surface temperature seems to have either the same or no effect on the probability of

Table 4: Significant covariates affecting the parameter **q**, the conditional probability of a large, GPD-distributed count, given the count is greater than zero. The definition of a large count differs from to species to species, and is only valid for double-hurdle models. Plus signs indicate that an increase in the values of the covariate significantly *increases* the conditional probability of observing a large count, given the count is non-zero, while negative signs (−) indicate that an increase in the values of the covariate significantly *decreases* the conditional probability of observing a large count, given the count is non-zero. Blanks indicate non-significance at the 0.05 level. (*Common Eider, Long-tailed Duck, and Roseate Tern are modeled by GPD-hurdle so does not contain the parameters $\boldsymbol{\mu}$ and **q**.)

Species	δ_1 **SST**	δ_2 **DEP**	δ_3 **CHL**	δ_4 **DTS**	δ_5 **SIN**	δ_6 **COS**
Atlantic Puffin						−
Black-legged Kittiwake	−	−			−	+
Black Scoter	+		+		+	+
Bonaparte's Gull	+				+	+
Common Loon	−	−		+	−	−
Common Tern		−			−	
Cory's Shearwater		−				
Double-crested Cormorant			+			+
Dovekie						
Great Black-backed Gull					−	−
Great Shearwater	−			+	−	−
Herring Gull	−			−	−	
Laughing Gull				−		
Leach's Storm-petrel	−		−		−	−
Northern Fulmar	−				−	
Northern Gannet	+	−		−		+
Razorbill			+	−		+
Sooty Shearwater					+	
Surf Scoter		−	+	+		+
White-winged Scoter		−		+		+
Wilson's Storm-petrel	−				−	−

any positive count and the probability of a large positive count for most species, both the Black Scoter and Northern Gannet had the opposite effect. Like most other species, the probability of observing at least one bird from these two species is higher with cooler temperatures. However, given at least one, the probability that it is a large count governed by GPD is higher with warmer temperatures. This kind of opposing effect is also true for surf scoters when it relates to chlorophyll level, and razorbills when it relates to distance to shore. The probability of observing at least one razorbill is higher away from the shore, but larger counts are less probable away from the shore given a razorbill observation.

4.3 Distribution maps

Monthly distribution maps of $1 - \hat{p}_{ij}$ (estimated occupancy probability) and $[1 - \hat{p}_{ij}] \times \hat{q}_{ij}$ (estimated large count probability) may be useful for offshore spatial planning, and are shown in the appendix. These maps of exposure probabilities offer insight to seasonal migration trends and aggregations of individual species.

Another useful display is to combine the monthly maps to create a single-map representation of exposure probabilities. That is, we could map $1-\hat{p}_i^* = 1-\prod_j \hat{p}_{ij}$, which gives the (estimated) probability of observing at least one individual at a given location during the year. Figure 4 gives estimates of $1 - \hat{p}_i^*$, using the best model for the Northern Gannet species as an example. The median estimate serves as our best-guess estimate, accompanied by maps of the 5^{th} and

*This report and digital versions of the maps are available at:
http://d25ripjvlq5c77.cloudfront.net/marine_birds_full.zip

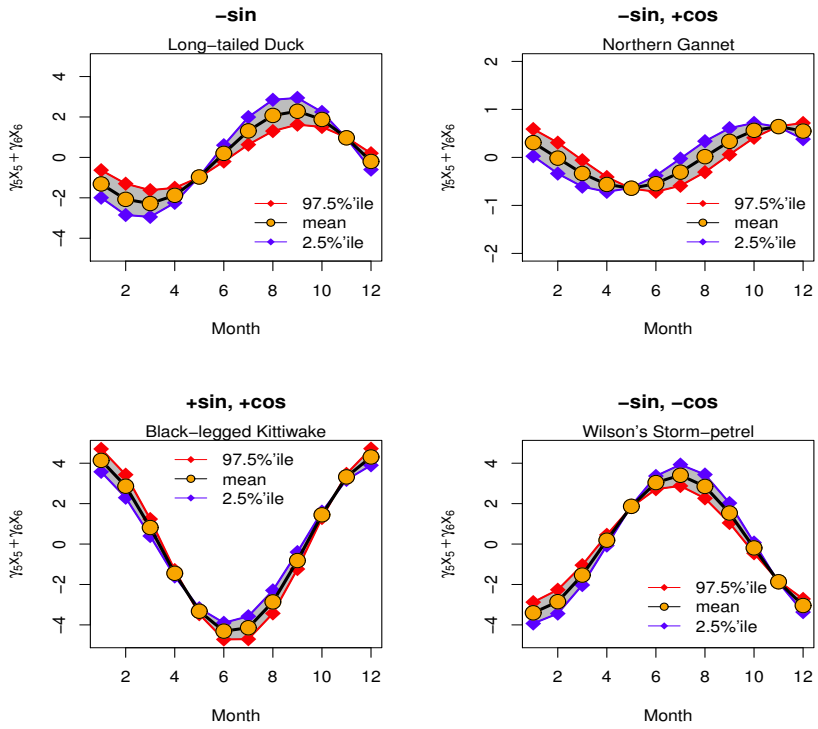

Figure 3: Effect of month-to-month variation on occupancy probability for select species with different combinations of signs of the sine and cosine variables. The average values of the coefficients γ_5 and γ_6 are plotted along with 95% bounds.

95^{th} percentiles. Mapping the low and high values of a 90% interval helps to gauge uncertainty in our parameter estimates. We can also map the estimated large count probability $\hat{q}_i^* = 1 - \prod_j \{1 - [1 - \hat{p}_{ij}] \times \hat{q}_{ij}\}$, which gives the probability of observing at least one large count during the year. Maps of $\hat{q}_i^*$ for the Northern Gannet are shown in Figure 5. Because the Northern Gannet is a highly abundant species, occupancy probabilities in Figure 4 are very high throughout the Atlantic coast (with the exception of Delaware Bay). In Figure 5, large count probabilities seem to follow the same general pattern as Figure 4, but areas of high large-count probabilities, ranging between 73% and 92% probability in some locations, are more revealing. The maps show distinct hotspots at the Virginia/North Carolina border (home of several national wildlife refuge areas), the southern border of Long Island, and off the coast of Maine around Mt. Desert Island (home of Acadia National Park). Probabilities above different threshold values may be mapped, regardless of the value of ψ, to reveal corresponding hotspots of species aggregation. For example, we can create a map for Northern Gannet that shows the probability of observing a group of ten or more individuals, using the model with $\psi = 6$.

5 Discussion

We proposed a three-component mixture model that accounts for both the extreme right tail of the observed distribution and the high frequency of zero-counts. Our method avoids the need for censoring or deleting the large and rare counts, but instead uses all available information from the data into one unified modeling strategy. We also explicitly included spatial auto-correlation in the model as well as biophysical covariates. Our results show that for marine bird data, which are often zero-inflated and highly over-dispersed, it is desirable to use a double-hurdle model

*This report and digital versions of the maps are available at:
`http://d25ripjvlq5c77.cloudfront.net/marine_birds_full.zip`

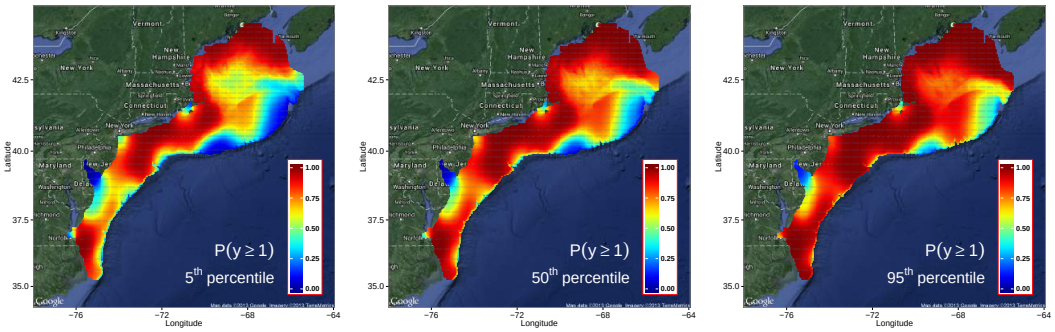

Figure 4: Northern Gannet: Maps of the probability of observing at least one individual during the year, using the best model for Northern Gannet, i.e., spatial model with fifty eigenvectors and $\psi = 6$. The median estimate of $1 - \hat{p}_i^*$ is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

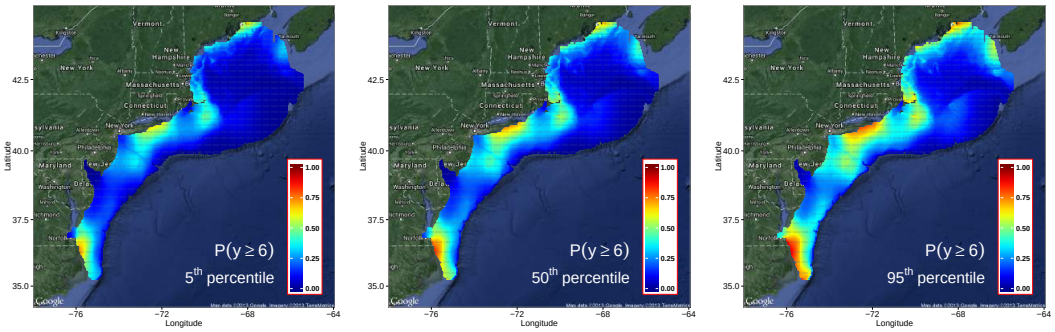

Figure 5: Northern Gannet: Maps of the probability of observing at least one large count of individuals during the year, using the best model for Northern Gannet, i.e., spatial model with fifty eigenvectors and $\psi = 6$. The median estimate of $\hat{q}_i^*$ is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

that fits a generalized Pareto distribution to the large counts. We found that the double-hurdle model performed the best for all species we analyzed, except three: Common Eider, Long-tailed Duck, and Roseate Tern. For those species, the GPD-hurdle was the better model, according to DIC. In all cases, the GPD outperformed the negative binomial, suggesting that distributions with a large right tail are often more appropriate for marine bird data.

These results are consistent with more recent work modeling group and flock sizes in marine bird data. Beauchamp (2011) used an information theoretic approach to compare a truncated power law, Poisson, exponential, and the negative binomial distributions for seven marine bird species separated by male and female. In this case, Beauchamp (2011) removed the zeros from the dataset and thus truncated all distributions to account for the removed zeros. The truncated power law distribution was selected as the top model according to AIC in ten cases and truncated negative binomial in the other four cases. Similarly, Zipkin et al. (2014) used a marked point process to separate number of flocks from flock size in their dataset. They found that the negative binomial worked well for number of flocks in an area, but that the discretized lognormal performed best for the flock size (Zipkin et al. 2014). Included in this analysis were a suite of distributions: discretized lognormal, geometric, logarithmic, zeta, Poisson, negative

*This report and digital versions of the maps are available at:
`http://d25ripjvlq5c77.cloudfront.net/marine_birds_full.zip`

binomial, and Yule-Simon. We note that the zeta distribution used by Zipkin et al. (2014) is a form of the generalized Pareto distribution (Clauset et al. 2009), but uses only a scale parameter as opposed to our parameterization which has two parameters. The added flexibility provided by this extra parameter is key to adequately modeling the extreme counts. While count data in ecology are often analyzed using Poisson or negative binomial distributions or zero-inflated versions (Ver Hoef and Boveng 2007; Oppel et al. 2012; McGowan et al. 2013; Renner et al. 2013), group size data have been fit with other distributions (Niwa 2003; Ma et al. 2011). It appears from these studies, as well as our results, that using counts and aggregations (or number of flocks and flock size) together in one analysis, requires combining multiple distributions to deal with the zero-inflation and large counts.

In addition to showing that the double-hurdle model fit the data well in our study, we were also able to model covariate relationships for all three components of the model—zero-inflation, typical abundance, and extreme counts. Other work has included covariate information into marine bird models (e.g., Zipkin et al. 2010; Kinlan et al. 2012; Johnston et al. 2014; Winiarski et al. 2014); however, ours is the first approach at incorporating covariate information to predict the extreme counts. In doing so, we can separate the potential influence of a standard predictor (say SST) and its effect on abundance versus extreme counts (i.e., aggregations). This approach could be very useful for predicting hotspots, when hotspots are defined as aggregates of marine birds. The inclusion of a sine and cosine covariate also allows for temporal variation in a cyclic pattern, common for marine birds that are utilizing the marine environment in different ways across the year.

While the double-hurdle model improved fit for most species, the single-hurdle model was preferred for three species: common eider, long-tailed duck, and roseate tern. This requires further study. One might expect that these species would be less abundant and have fewer extreme counts than other species. While this describes the roseate turn, the long-tailed duck and common eider are both abundant and have several extremely high counts. Understanding the characteristics of the distribution for which the single-hurdle model is sufficient would be useful in future modeling efforts.

Components of the model can be improved upon in many different ways. Some species maps suffer from edge effects, especially along the continental shelf, where there is a sudden and extreme drop in ocean depth, and many zero observations far away from nonzero observations. Cells near the edge naturally have more uncertainty because they have fewer neighbors providing spatial information. This effect is less pronounced near the coast where non-zero counts are more common, but remain prominent along all the edges of our spatial domain. Cressie (1993) lists several approaches to handle the conditional specification of edge cells, but remarks that dealing with edge cells is an ongoing research problem.

The variation and anomalies in our parameter estimates (see Section 4.2) for the NB and GPD components may be due to the fact that we held these relationships constant (for example, our parameterization does not allow sea surface temperature to have a different effect on relative abundance in January as it does in June). Due to data constraints, we were not able to allow parameters to vary by season or month. We know certain species are behaving differently throughout the year, nesting, staging, migrating, etc., and thus we would expect that the covariate effects may vary as a result of these different patterns in behavior or that the covariate effects should vary spatially. We also did have data on the true transect path for all surveys, and we approximated the amount of effort within each grid cell in the state space by first drawing a straight line between the given start and end coordinates of transects, then only counting effort for each instance a transect intersects a grid cell (See appendix). However, one may also consider counting effort using either the amount of time (if available) or the length of the transect that lie within the grid cell boundaries, which may decrease bias in the parameter estimates by including a more specific offset for survey effort.

Spatial random effects have been shown to be confounded with the fixed-effect covariates $\mathbf{X}$ in spatial regression models (Hodges and Reich 2010). The methods developed by Reich et al. (2006); Hughes and Haran (2013) have addressed this issue by using a restricted form of the spatial random effect instead. Eigenvectors from a restricted form of the ICAR model, for

example, can be used as spatial covariates that are not confounded with $\mathbf{X}$. Studying the extent of spatial confounding in the present models and possibly applying these methods is a subject of future research.

The notion of risk can be interpreted in several different ways. In this paper, we have presented what are commonly called "exposure" maps using measures of density, rather than measures of impact on individual birds or species as a direct result of offshore energy developments. Our results, however, may be used to assess which populations are at risk of being impacted. We can only assume that areas of high use and aggregation are considered high priorities for ocean planners. An alternative view of risk is based on the magnitude of a potential impact on a species or community of species (Burthe et al. 2014), instead of simply being in the presence of a potential impact. The idea is that certain species may be more vulnerable to habitat change or substantial population decreases than other species. Robinson Willmott et al. (2013) built upon the works of Garthe and Huppop (2004), Desholm (2009), Furness and Wade (2012), and Furness et al. (2013) to rank a species' collision vulnerability and displacement vulnerability using relative sensitivity scores. While some measures used to compute these scores are based on observable data, other measures are based solely on expert opinion. Because of the uncertain nature of sensitivity scores and the questions surrounding the methods used to calculate them, we did not incorporate vulnerability indices in our current models, although they can easily be included in future extensions of our maps. A specific application may be to use such indices as weights when the outputs of several individual species models are combined and a map associated with the entire group is produced, assuming that each species behave and react similarly to environmental factors. Combining species from the same family or functional group, for example, common terns and roseate terns, may satisfy this assumption (Goyert et al. 2014). Combined species maps are under current investigation. For now, we decided to create monthly exposure maps that combined model outputs of all twenty-four species we have analyzed in this paper. These maps can be found in the Appendix. Our main focus in this paper, however, is the construction of individual species exposure maps. When used as weights, vulnerability indices may still prove useful when determining which individual species map to consider more than others.

Overall, our modeling approach extends marine bird survey analyses to include extreme counts, while still explicitly incorporating spatial correlation and biophysical and temporal covariates. We demonstrate the need for more flexible distributions when analyzing marine bird data which often display a very high variance to mean ratio.

Acknowledgements

We thank the North Atlantic Landscape Conservation Cooperative and Wildlife Management Institute for funding. We also thank the Fish and Wildlife Service for access to the North Atlantic Seabird Compendium, and all those that have contributed their data to this resource. Thanks to Allison Sussman for data support throughout this project. We thank Brian Kinlan, Dick Veit, Andrew Gilbert, Timothy Jones, and Melanie Steinkamp for their ideas, proposal writing, project support, and consulting, and thanks to Scott Schwenk for comments on an earlier draft of this report.

Notes

This report was revised on May 16, 2016, to make the following corrections: Earlier versions of this report mistakenly referred to a couple of species data as "Black-capped Petrel" (BCPE) and "Black Tern" (BLTE). The correct species names that correspond to these data are "Black Scoter" (BLSC) and "Black-legged Kittiwake" (BLKI), respectively. These corrections are reflected in the updated tables, figures 2 and 3, appendices, and some inline references in the main body of the report.

References

American Council On Renewable Energy (2014), "The outlook for renewable energy in America," 1600 K St. NW, Suite 650, Washington, DC 20006.

American Wind Energy Association (2013), "Offshore wind: significant investment in research and development," .

Beauchamp, G. (2011), "Fit of aggregation models to the distribution of group sizes in Northwest Atlantic seabirds," *Marine Ecology Progress Series*, 425, 261–268.

Behrens, C. N., Lopes, H. F., and Gamerman, D. (2004), "Bayesian analysis of extreme events with threshold estimation," *Statistical Modelling*, 4, 227–244.

Besag, J. (1974), "Spatial interaction and the statistical analysis of lattice systems," *Journal of the Royal Statistical Society. Series B (Methodological)*, 192–236.

Besag, J., York, J., and Mollié, A. (1991), "Bayesian image restoration, with two applications in spatial statistics," *Annals of the Institute of Statistical Mathematics*, 43, 1–20.

Bliss, C. I. and Fisher, R. A. (1953), "Fitting the negative binomial distribution to biological data," *Biometrics*, 9, 176–200.

Boots, B. and Tiefelsdorf, M. (2000), "Global and local spatial autocorrelation in bounded regular tessellations," *Journal of Geographical Systems*, 2, 319–348.

Buishand, T., de Haan, L., and Zhou, C. (2008), "On spatial extremes: with application to a rainfall problem," *The Annals of Applied Statistics*, 624–642.

Burthe, S. J., Wanless, S., Newell, M. A., Butler, A., and Daunt, F. (2014), "Assessing the vulnerability of the marine bird community in the western North Sea to climate change and other anthropogenic impacts," *Marine Ecology Progress Series*, 507, 277–295.

Clauset, A., Shalizi, C. R., and Newman, M. E. (2009), "Power-law distributions in empirical data," *SIAM Review*, 51, 661–703.

Coles, S., Bawa, J., Trenner, L., and Dorazio, P. (2001), *An introduction to statistical modeling of extreme values*, vol. 208, Springer.

Coles, S. G. and Tawn, J. A. (1996), "A Bayesian analysis of extreme rainfall data," *Applied Statistics*, 463–478.

Cragg, J. G. (1971), "Some statistical models for limited dependent variables with application to the demand for durable goods," *Econometrica: Journal of the Econometric Society*, 829–844.

Cressie, N. (1993), *Statistics for Spatial Data: Wiley Series in Probability and Mathematical Statistics*, Wiley-Interscience New York.

Davis, R. and Resnick, S. (1984), "Tail estimates motivated by extreme value theory," *The Annals of Statistics*, 1467–1487.

Desholm, M. (2009), "Avian sensitivity to mortality: Prioritising migratory bird species for assessment at proposed wind farms," *Journal of Environmental Management*, 90, 2672–2679.

Drewitt, A. L. and Langston, R. H. W. (2006), "Assessing impacts of windfarms on birds," *Ibis*, 148, 29–42.

DuMouchel, W. H. (1983), "Estimating the stable index α in order to measure tail thickness: a critique," *The Annals of Statistics*, 1019–1031.

Ford, R. G., Casey, D. G., Keiper, C. A., Spear, L. B., and Balance, L. T. (2004), "The biogeographic patterns of seabirds in the central portion of the California current," *Marine Ornithology*, 32, 77–96.

Furness, B. and Wade, H. (2012), "Vulnerability of Scottish seabirds to offshore wind turbines," *MacArthur Green Ltd report. 39pp*.

Furness, R. W., Wade, H. M., and Masden, E. A. (2013), "Assessing vulnerability of marine bird populations to offshore wind farms," *Journal of Environmental Management*, 119, 56–66.

Garthe, S. and Huppop, O. (2004), "Scaling possible adverse effects of marine wind farms on seabirds: developing and applying a vulnerability index," *Journal of Applied Ecology*, 41, 724–734.

Geisser, S. and Eddy, W. F. (1979), "A predictive approach to model selection," *Journal of the American Statistical Association*, 74, 153–160.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003), *Bayesian data analysis*, CRC Press.

Goyert, H. F., Manne, L. L., and Veit, R. R. (2014), "Facilitative interactions among the pelagic community of temperate migratory terns, tunas and dolphins," *Oikos*.

Greene, W. (2007), *Functional form and heterogeneity in models for count data*, Now Publishers Inc.

Greene, W. H. (1994), "Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models," in *Economics Working Papers*, NYU Stern School of Business, EC-94-10.

Guillemette, M., Himmelman, J. H., Barette, C., and Reed, A. (1993), "Habitat selection by common eiders in winter and its interaction with flock size," *Canadian Journal of Zoology*, 71, 1259–1266.

Gurmu, S. (1998), "Generalized hurdle count data regression models," *Economics Letters*, 58, 263–268.

Hilbe, J. M. (2011), *Negative binomial regression*, Cambridge University Press.

Hodges, J. S. and Reich, B. J. (2010), "Adding spatially-correlated errors can mess up the fixed effect you love," *The American Statistician*, 64, 325–334.

Hosking, J. R. and Wallis, J. R. (1987), "Parameter and quantile estimation for the generalized Pareto distribution," *Technometrics*, 29, 339–349.

Huettman, F. and Diamond, A. W. (2001), "Seabird colony locations and environmental determination of seabird distribution: a spatially explicit breeding seabird model for the Northwest Atlantic," *Ecological Modelling*, 141, 261–298.

Hughes, J. and Haran, M. (2013), "Dimension reduction and alleviation of confounding for spatial generalized linear mixed models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75, 139–159.

Johnston, A., Thaxter, C. B., Austin, G. E., Cook, A. S., Humphreys, E. M., Still, D. A., Mackay, A., Irvine, R., Webb, A., and Burton, N. H. (2014), "Modelling the abundance and distribution of marine birds accounting for uncertain species identification," *Journal of Applied Ecology*.

Jones, A. M. (1989), "A double-hurdle model of cigarette consumption," *Journal of Applied Econometrics*, 4, 23–39.

Kinlan, B. P., Zipkin, E., O'Connell, A., and Caldow, C. (2012), "Statistical analyses to support guidelines for marine avian sampling: final report," Tech. Rep. NOAA Technical Memorandum NOS NCCOS 158, NOAA.

Labeaga, J. M. (1999), "A double-hurdle rational addiction model with heterogeneity: estimating the demand for tobacco," *Journal of econometrics*, 93, 49–72.

Lambert, D. (1992), "Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing," *Technometrics*, 34, 1–14.

Lewis, T. L., Esler, D., and Boyd, W. S. (2008), "Foraging behavior of Surf Scoters (Melanitta perspicillata) and White-winged Scoters (M. fusca) in relation to clam density: inferring food availability and habitat quality," *The Auk*, 125, 149–157.

Ma, Q., Johansson, A., and Sumpter, D. J. (2011), "A first principles derivation of animal group size distributions," *Journal of Theoretical Biology*, 283, 35–43.

McGowan, J., Hines, E., Elliott, M., Howar, J., Dransfield, A., Nur, N., and Jahncke, J. (2013), "Using Seabird Habitat Modeling to Inform Marine Spatial Planning in Central California's National Marine Sanctuaries," *PloS ONE*, 8, e71406.

Menza, C., Kinlan, B. P., Dorfman, D. S., Poti, M., and Caldow, C. (2012), "A Biogeographic Assessment of Seabirds, Deep Sea Corals and Ocean Habitats of the New York Bight: Science to Support Offshore Spatial Planning," NOAA Technical Memorandum NOS NCCOS 141, NOAA National centers for coastal Ocean Science, Silver Spring, MD.

Miranda, A. (2010), "A double-hurdle count model for completed fertility data from the developing world," Department of Quantitative Social Science, Institute of Education.

Mullahy, J. (1986), "Specification and testing of some modified count data models," *Journal of Econometrics*, 33, 341–365.

Niwa, H.-S. (2003), "Power-law versus exponential distributions of animal group sizes," *Journal of Theoretical Biology*, 224, 451–457.

Oppel, S., Meirinho, A., Ramírez, I., Gardner, B., O'Connell, A. F., Miller, P. I., and Louzao, M. (2012), "Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds," *Biological Conservation*, 156, 94–104.

R Core Team (2014), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

Raqab, M. Z. and Asadi, M. (2008), "On the mean residual life of records," *Journal of Statistical Planning and Inference*, 138, 3660–3666.

Reich, B. J., Hodges, J. S., and Zadnik, V. (2006), "Effects of Residual Smoothing on the Posterior of the Fixed Effects in Disease-Mapping Models," *Biometrics*, 62, 1197–1206.

Reiss, R.-D., Thomas, M., and Reiss, R. (2001), *Statistical analysis of extreme values*, Springer.

Renner, M., Parrish, J. K., Piatt, J. F., Kuletz, K. J., Edwards, A. E., and Hunt Jr, G. L. (2013), "Modeled distribution and abundance of a pelagic seabird reveal trends in relation to fisheries," *Marine Ecology Progress Series*, 484, 259–277.

Ricker-Gilbert, J., Jayne, T. S., and Chirwa, E. (2011), "Subsidies and crowding out: A double-hurdle model of fertilizer demand in Malawi," *American Journal of Agricultural Economics*, 93, 26–42.

Ridout, M., Demétrio, C. G., and Hinde, J. (1998), "Models for count data with many zeros," in *Proceedings of the XIXth International Biometric Conference*, vol. 19, pp. 179–192.

Roberts, J. J., Best, B. D., Dunn, D. C., Treml, E. A., and Halpin, P. N. (2010), "Marine Geospatial Ecology Tools: An integrated framework for ecological geoprocessing with ArcGIS, Python, R, MATLAB, and C++," *Environmental Modelling & Software*, 25, 1197–1207.

Robinson Willmott, J., Forcey, G., and Kent, A. (2013), "The Relative Vulnerability of Migratory Bird Species to Offshore Wind Energy Projects on the Atlantic Outer Continental Shelf: An Assessment Method and Database," *Final Report to the US Department of the Interior, Bureau of Ocean Energy Management, Office of Renewable Energy Programs. OCS Study BOEM*, 207, 275.

Sampford, M. (1955), "The truncated negative binomial distribution," *Biometrika*, 58–69.

Saul, L. K., Weinberger, K. Q., Ham, J. H., Sha, F., and Lee, D. D. (2006), "Spectral methods for dimensionality reduction," *Semisupervised learning*, 293–308.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002), "Bayesian measures of model complexity and fit," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64, 583–639.

Tancredi, A., Anderson, C., and O'Hagan, A. (2006), "Accounting for threshold uncertainty in extreme value estimation," *Extremes*, 9, 87–106.

Ver Hoef, J. M. and Boveng, P. L. (2007), "Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data?" *Ecology*, 88, 2766–2772.

Winiarski, K. J., Miller, D. L., Paton, P. W., and McWilliams, S. R. (2014), "A spatial conservation prioritization approach for protecting marine birds given proposed offshore wind energy development," *Biological Conservation*, 169, 79–88.

Winkelmann, R. (2008), *Econometric analysis of count data*, Springer, 5th ed.

Zipkin, E. F., Gardner, B., Gilbert, A. T., O'Connell, A. F., Royle, J. A., and Silverman, E. D. (2010), "Distribution patterns of wintering sea ducks in relation to the North Atlantic Oscillation and local environmental characteristics," *Oecologia*, 163, 893–902.

Zipkin, E. F., Leirness, J. B., Kinlan, B. P., O'Connell, A. F., and Silverman, E. D. (2014), "Fitting statistical distributions to sea duck count data: implications for survey design and abundance estimation," *Statistical Methodology*, 17, 67–81.

Zorn, C. J. (1998), "An analytic and empirical examination of zero-inflated and hurdle Poisson specifications," *Sociological Methods & Research*, 26, 368–400.

A Maps

A.1 Survey effort

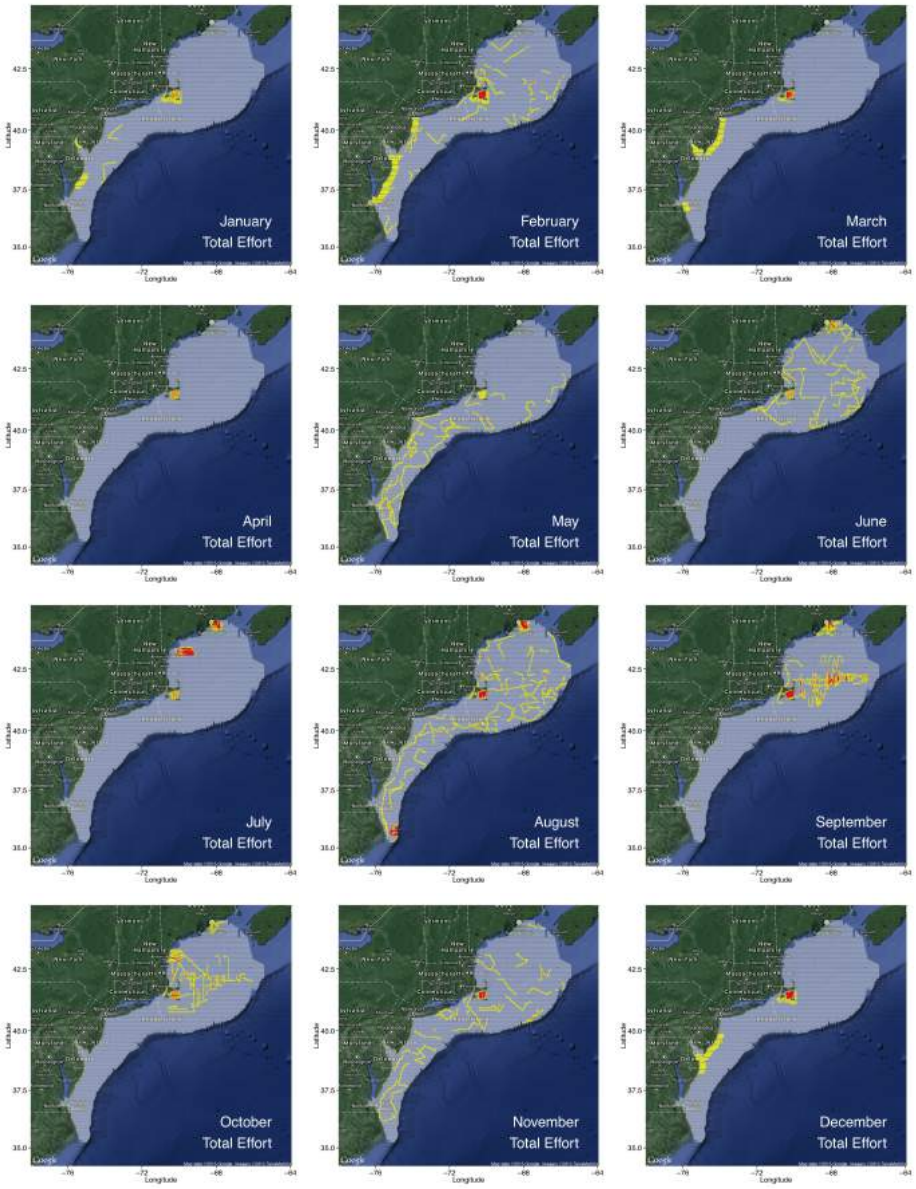

Figure 6: Total amount of effort for each calendar month. Effort is the number of times a spatial grid cell is entered by survey transect lines.

A.2 Combined species exposure maps

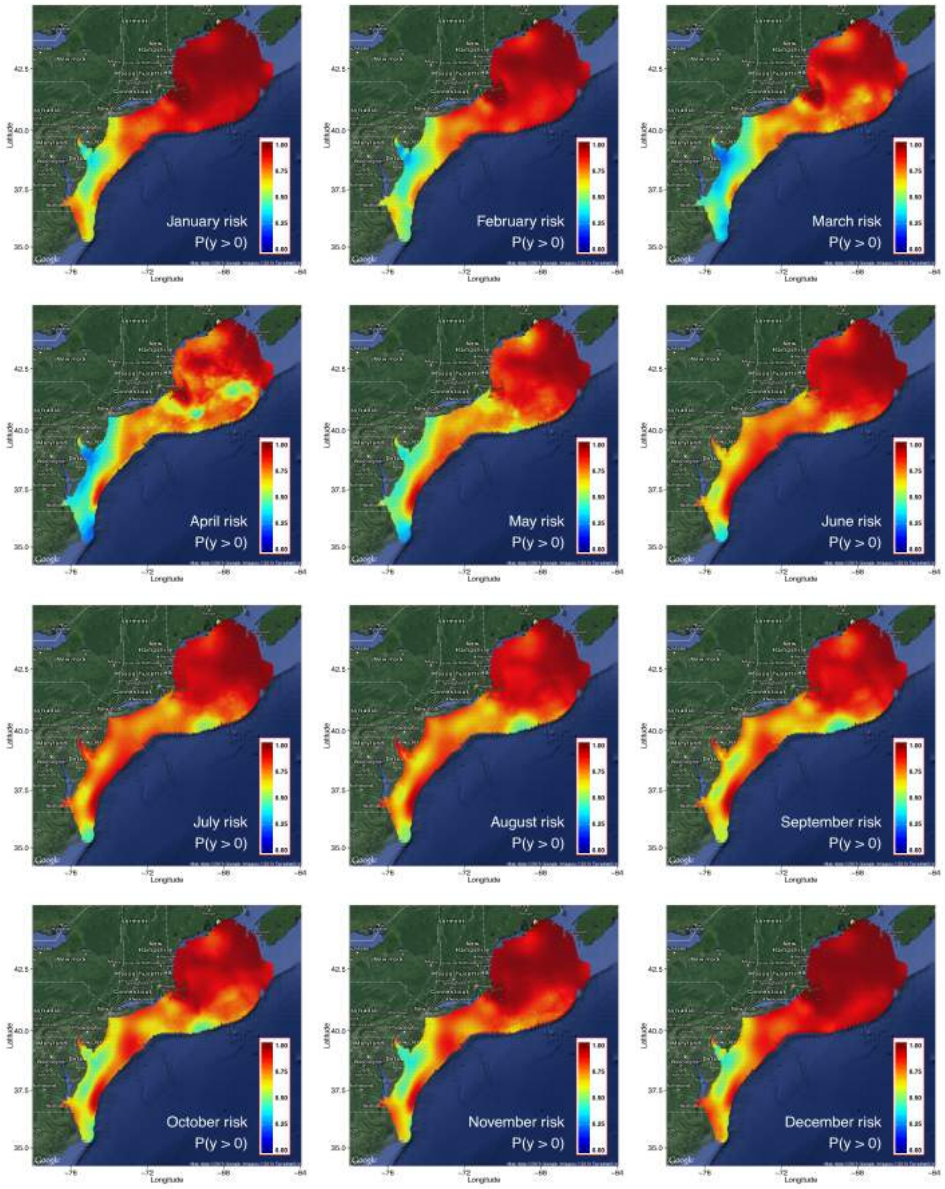

Figure 7: Probability of observing at least one individual during each month, including sea ducks (common eider, long-tailed duck, surf scoter, white-winged scoter).

*This report and digital versions of the maps are available at:
`http://d25ripjvlq5c77.cloudfront.net/marine_birds_full.zip`

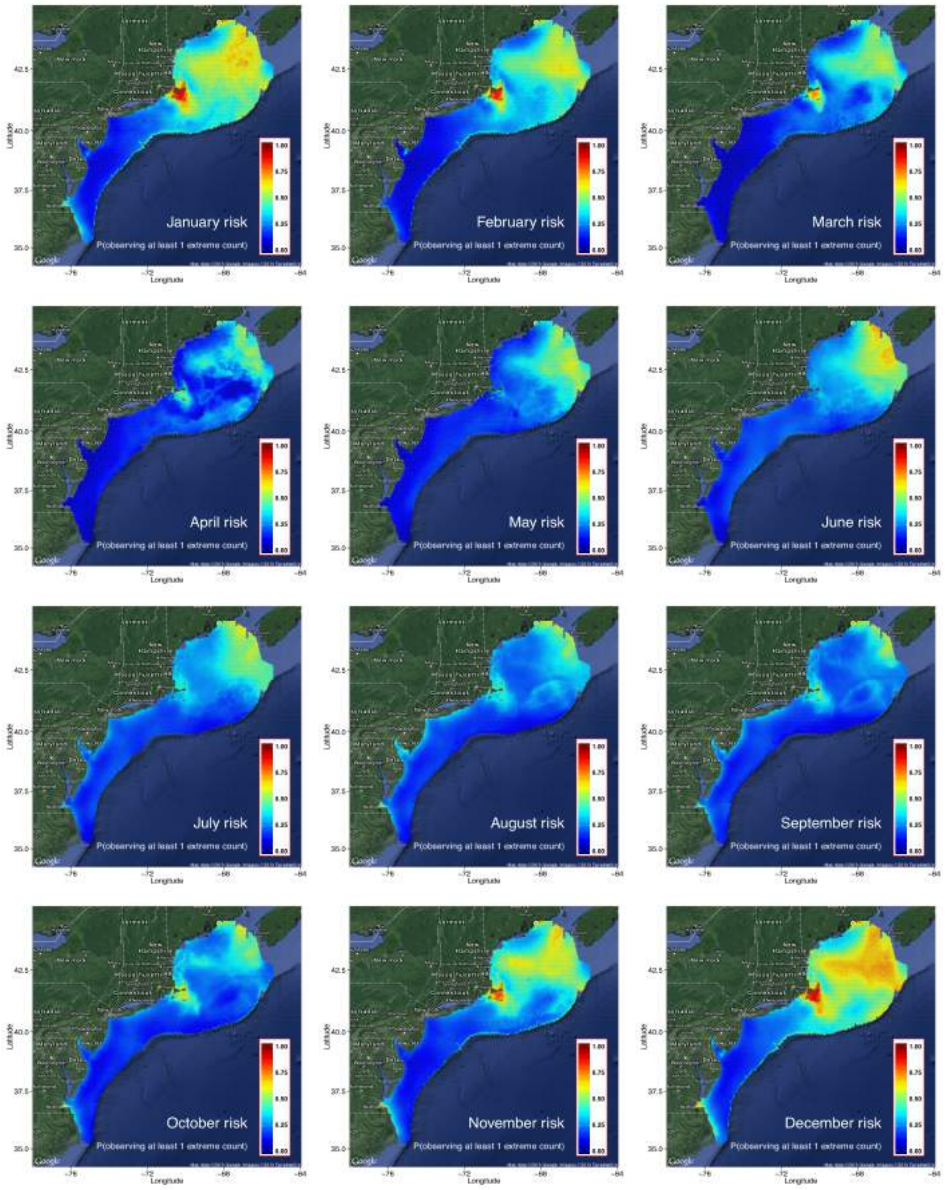

Figure 8: Probability of observing at least one extreme count during each month, including sea ducks (common eider, long-tailed duck, surf scoter, white-winged scoter).

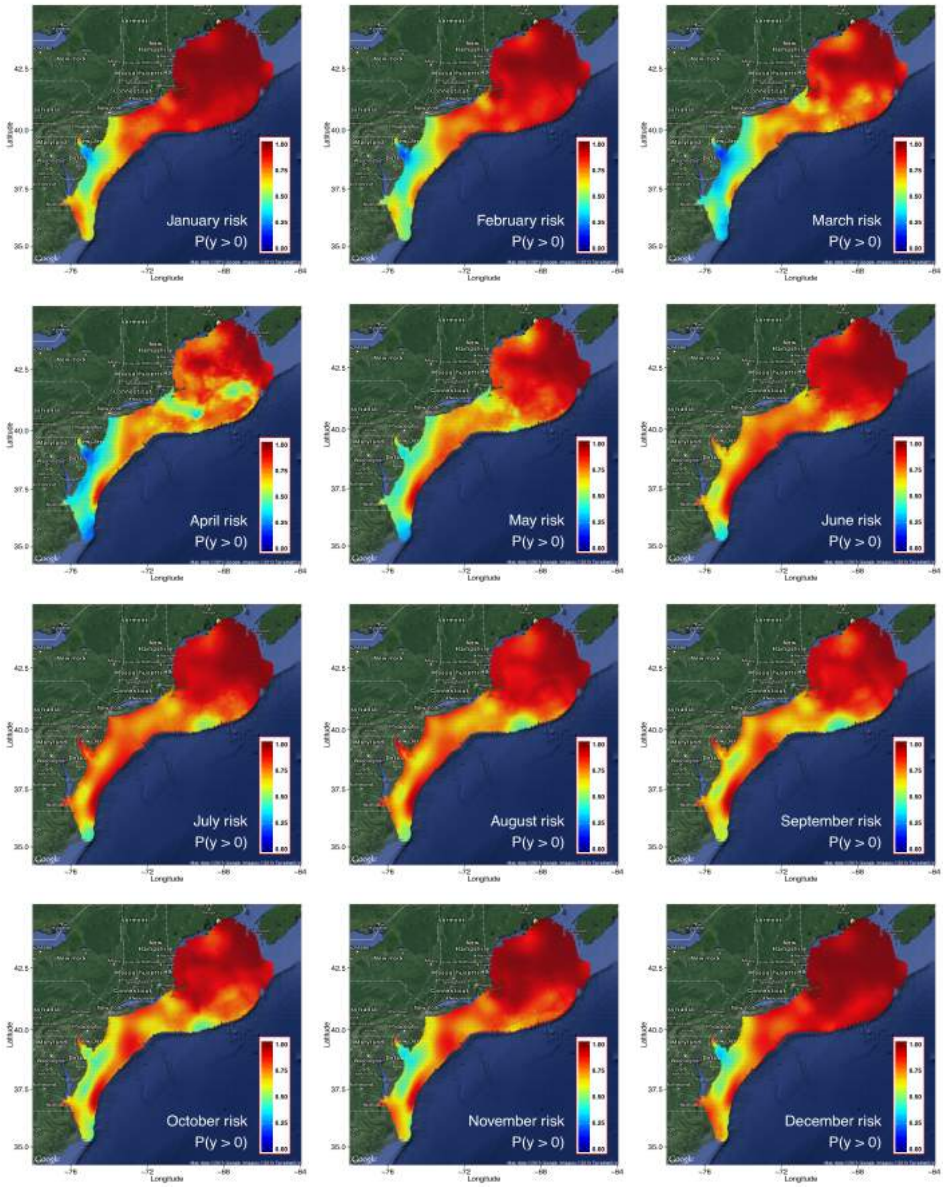

Figure 9: Probability of observing at least one individual during each month during each month, not including sea ducks (common eider, long-tailed duck, surf scoter, white-winged scoter).

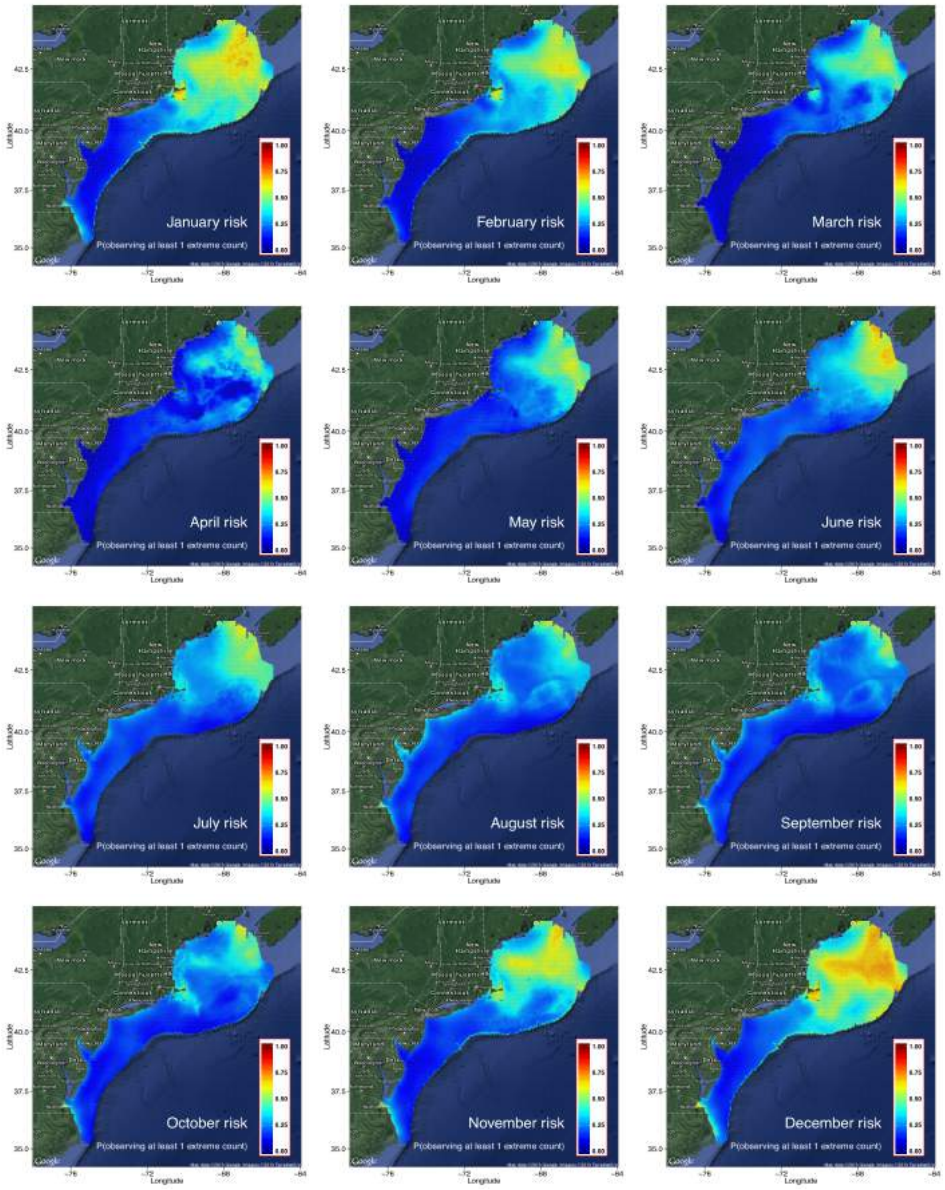

Figure 10: Probability of observing at least one extreme count during each month, not including sea ducks (common eider, long-tailed duck, surf scoter, white-winged scoter).

*This report and digital versions of the maps are available at:
`http://d25ripjvlq5c77.cloudfront.net/marine_birds_full.zip`

A.3 Atlantic Puffin

A.3.1 One-year summary

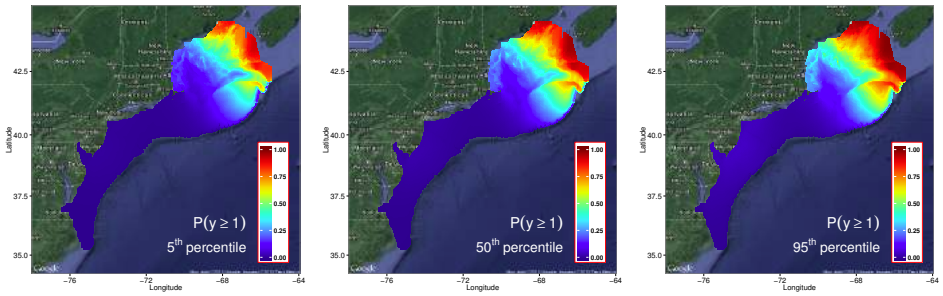

Figure 11: Atlantic Puffin: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

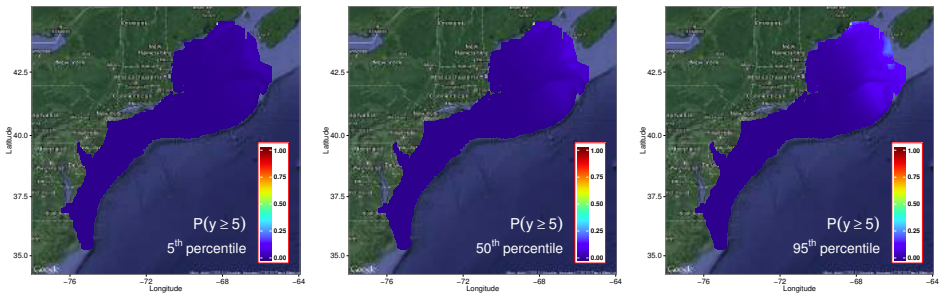

Figure 12: Atlantic Puffin: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.3.2 Monthly observations

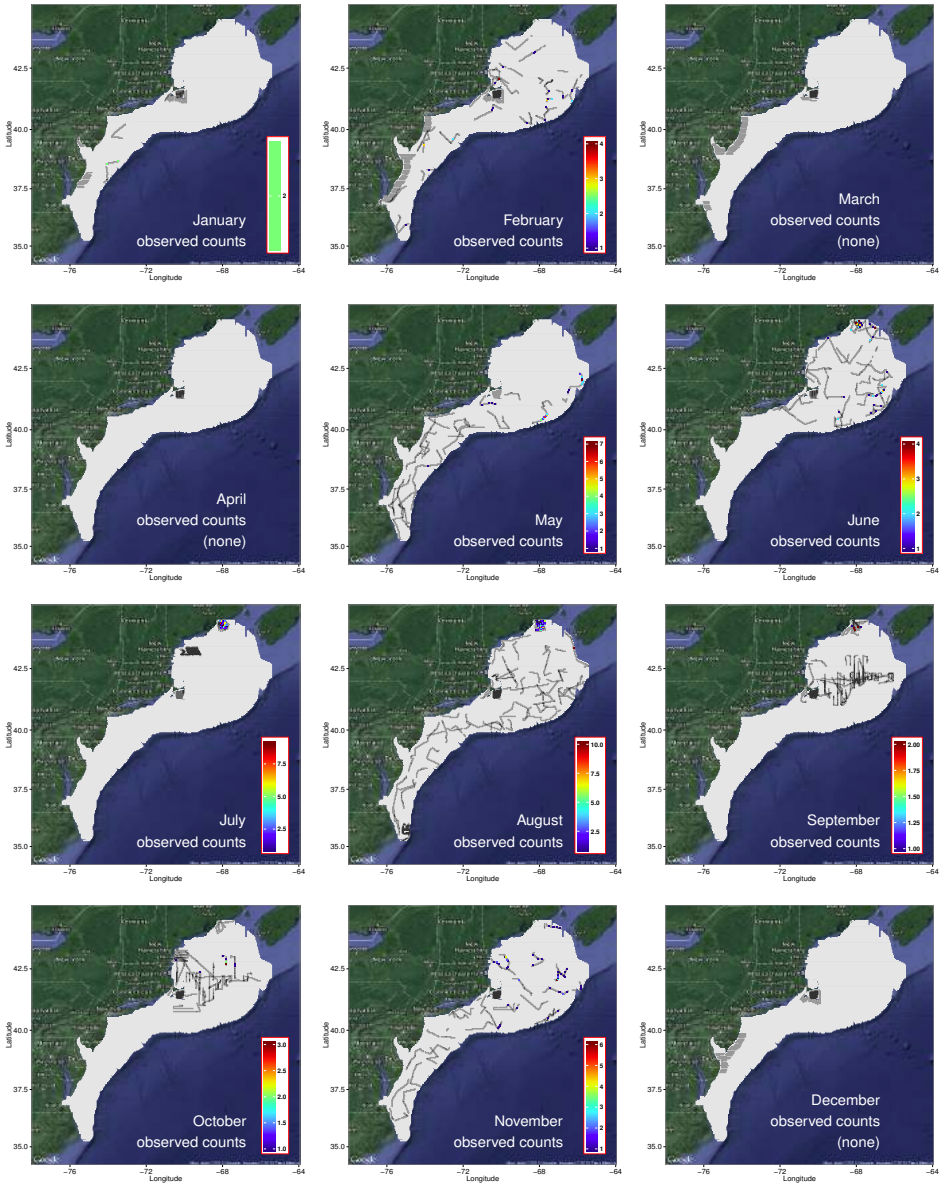

Figure 13: Atlantic Puffin: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.3.3 Monthly exposure maps

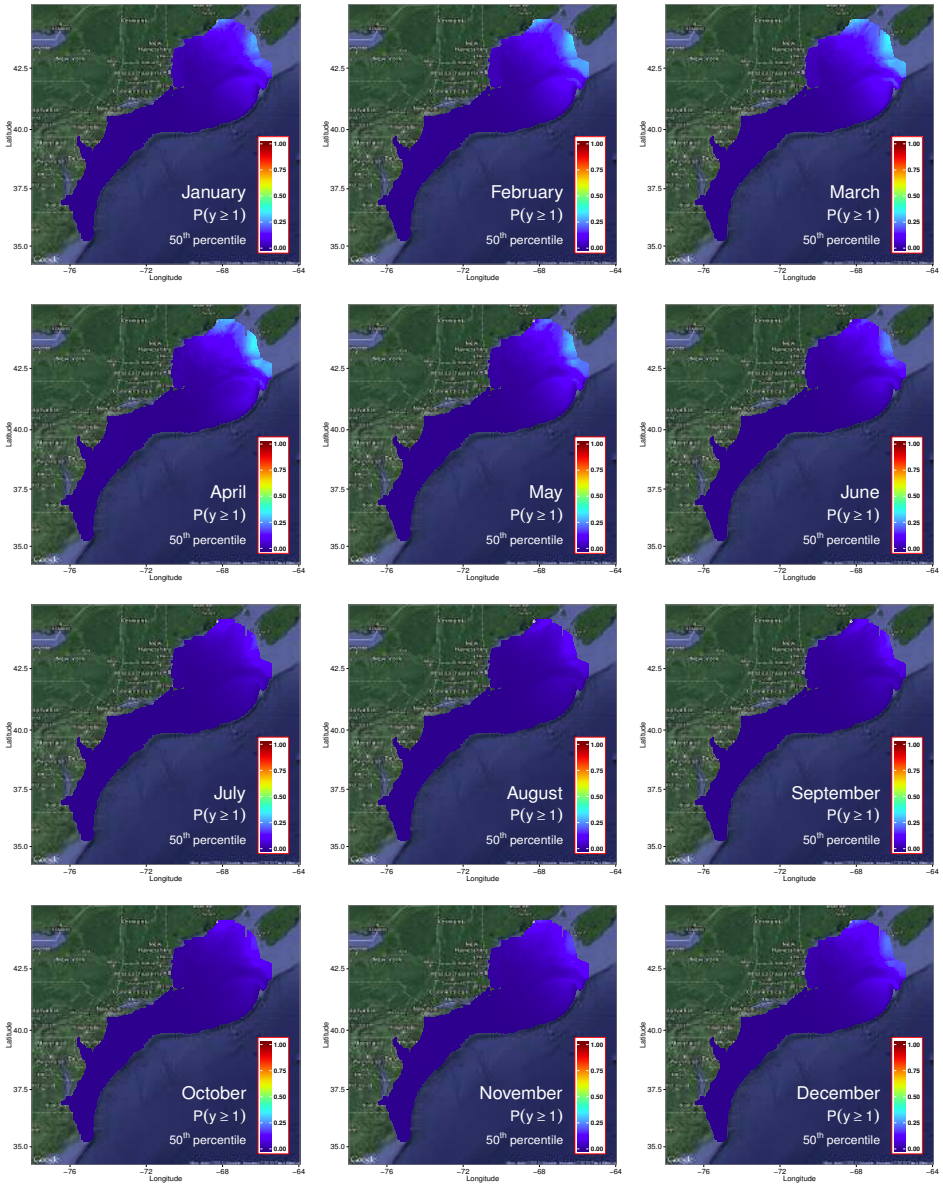

Figure 14: Atlantic Puffin: Probability of observing at least one individual during each month.

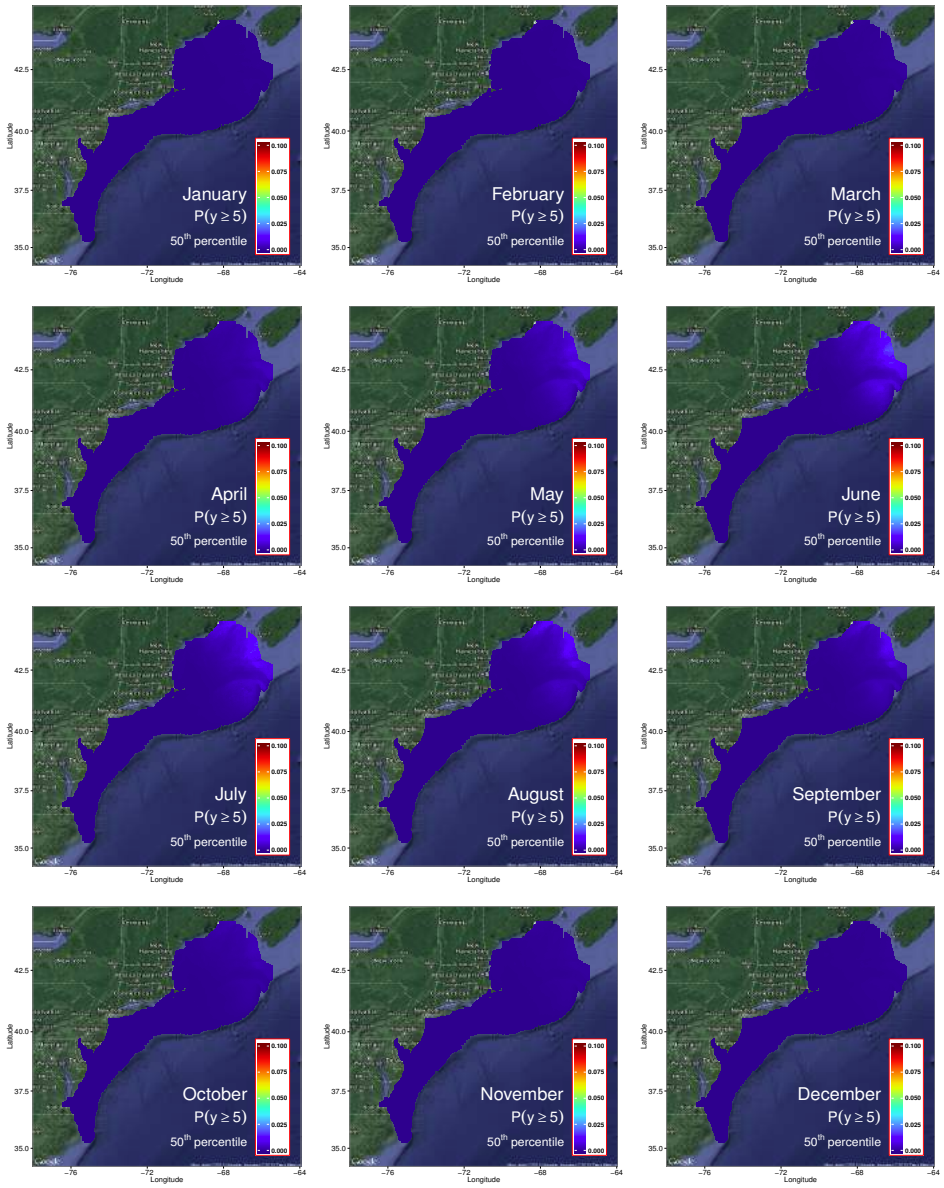

Figure 15: Atlantic Puffin: Probability of observing a large count during each month.

A.4 Black Scoter

A.4.1 One-year summary

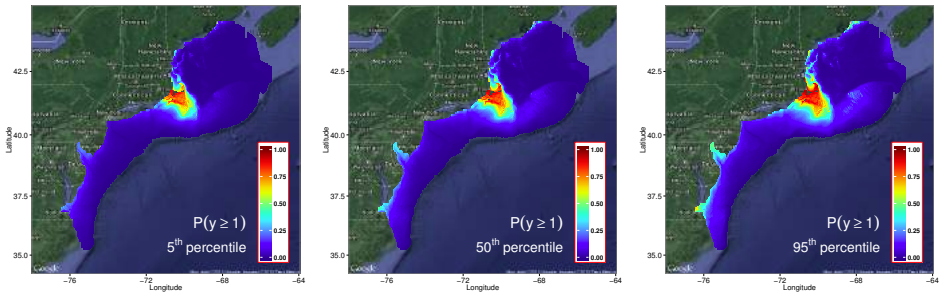

Figure 16: Black Scoter: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

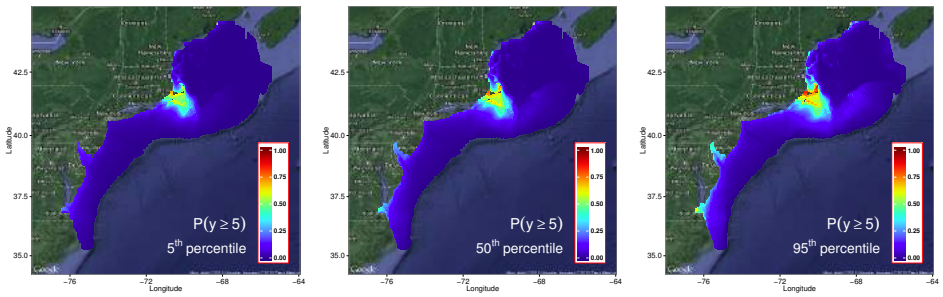

Figure 17: Black Scoter: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.4.2 Monthly observations

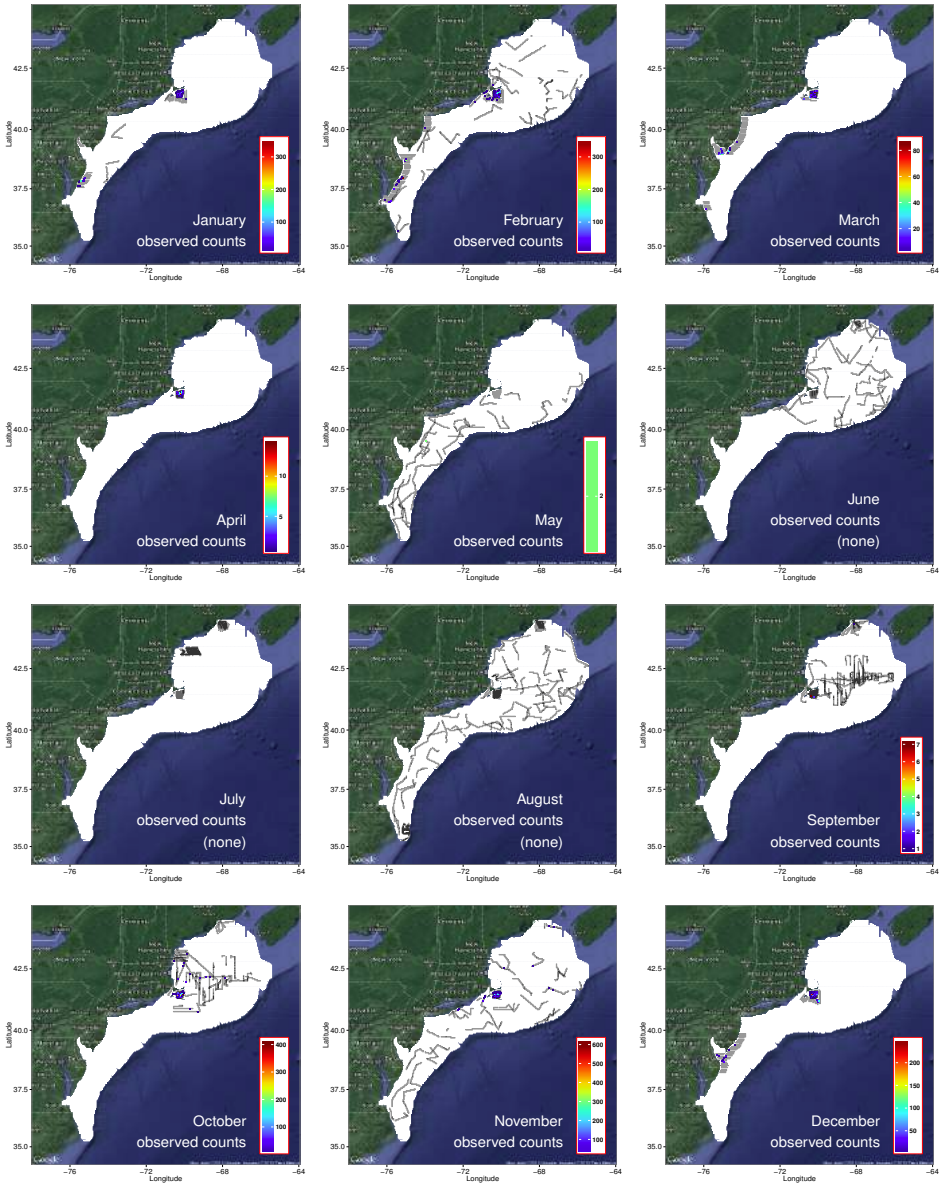

Figure 18: Black Scoter: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.4.3 Monthly exposure maps

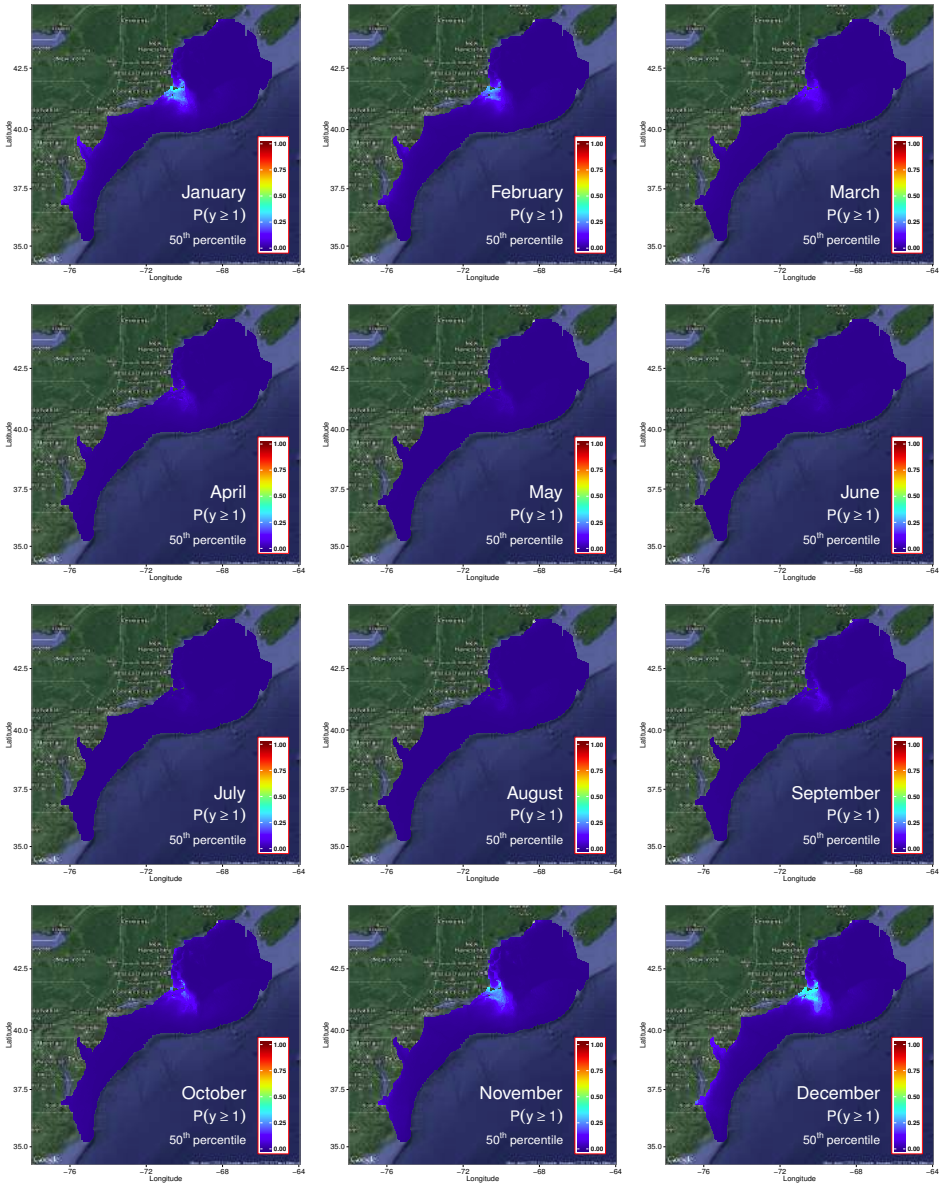

Figure 19: Black Scoter: Probability of observing at least one individual during each month.

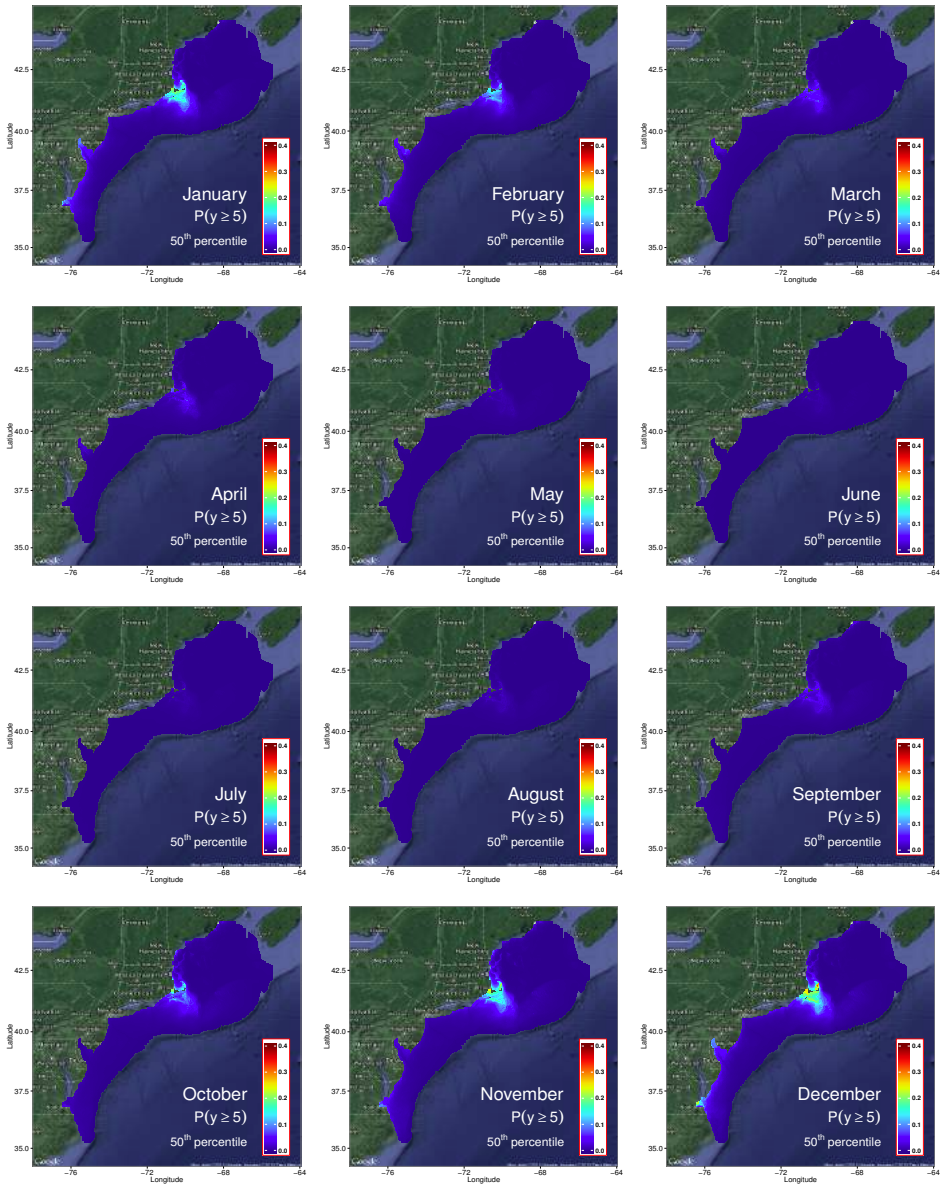

Figure 20: Black Scoter: Probability of observing a large count during each month.

A.5 Black-legged Kittiwake

A.5.1 One-year summary

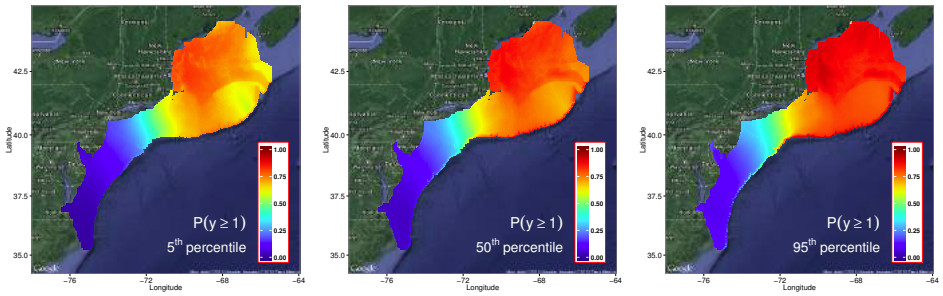

Figure 21: Black-legged Kittiwake: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

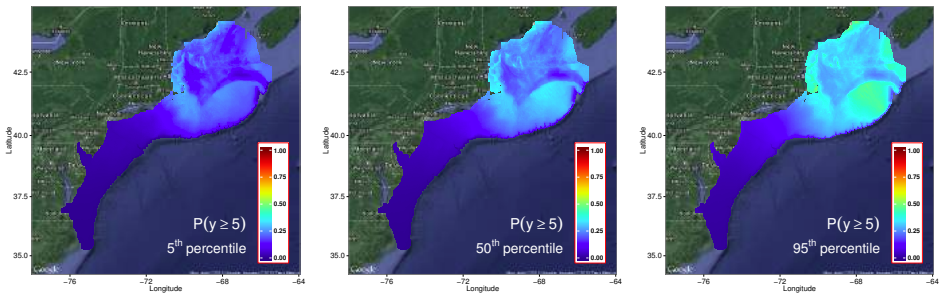

Figure 22: Black-legged Kittiwake: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.5.2 Monthly observations

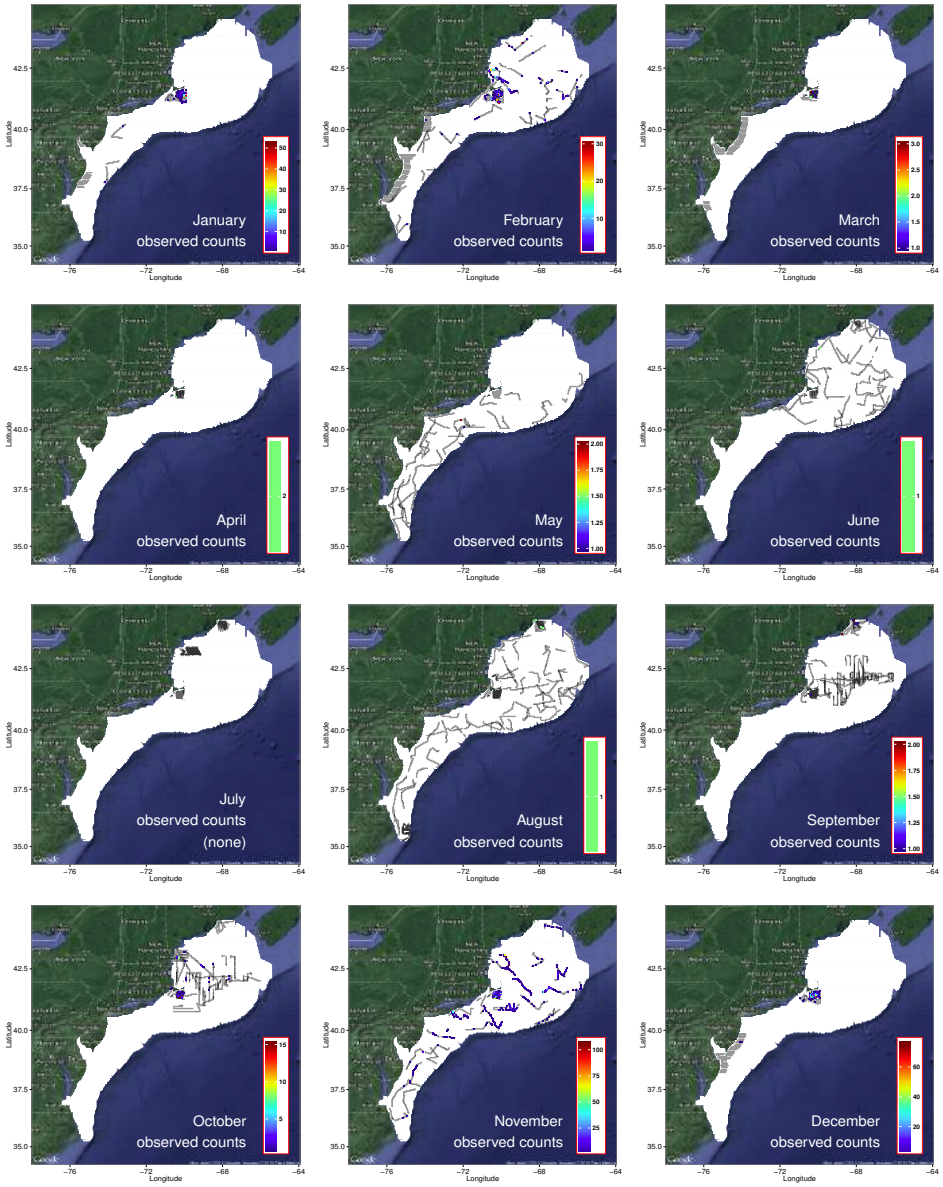

Figure 23: Black-legged Kittiwake: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.5.3 Monthly exposure maps

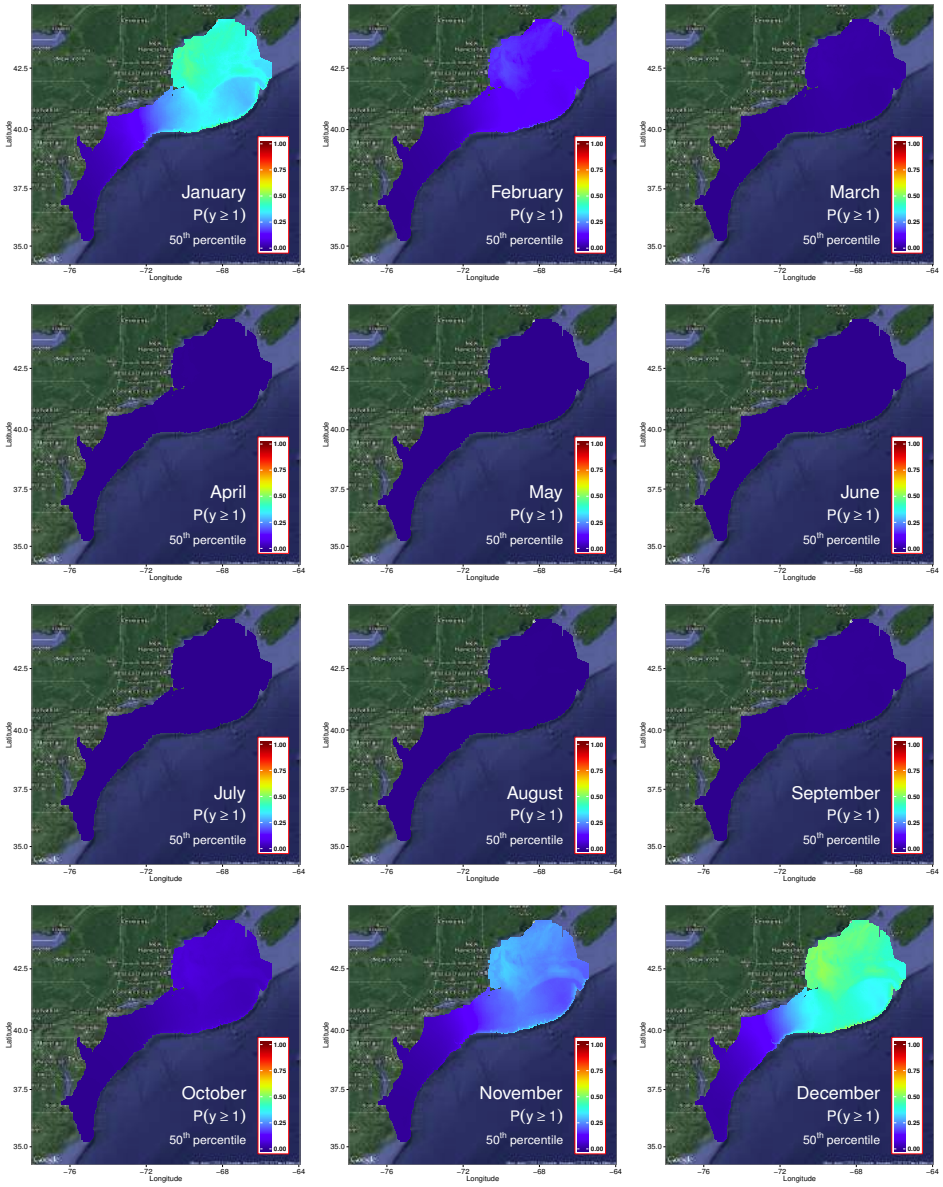

Figure 24: Black-legged Kittiwake: Probability of observing at least one individual during each month.

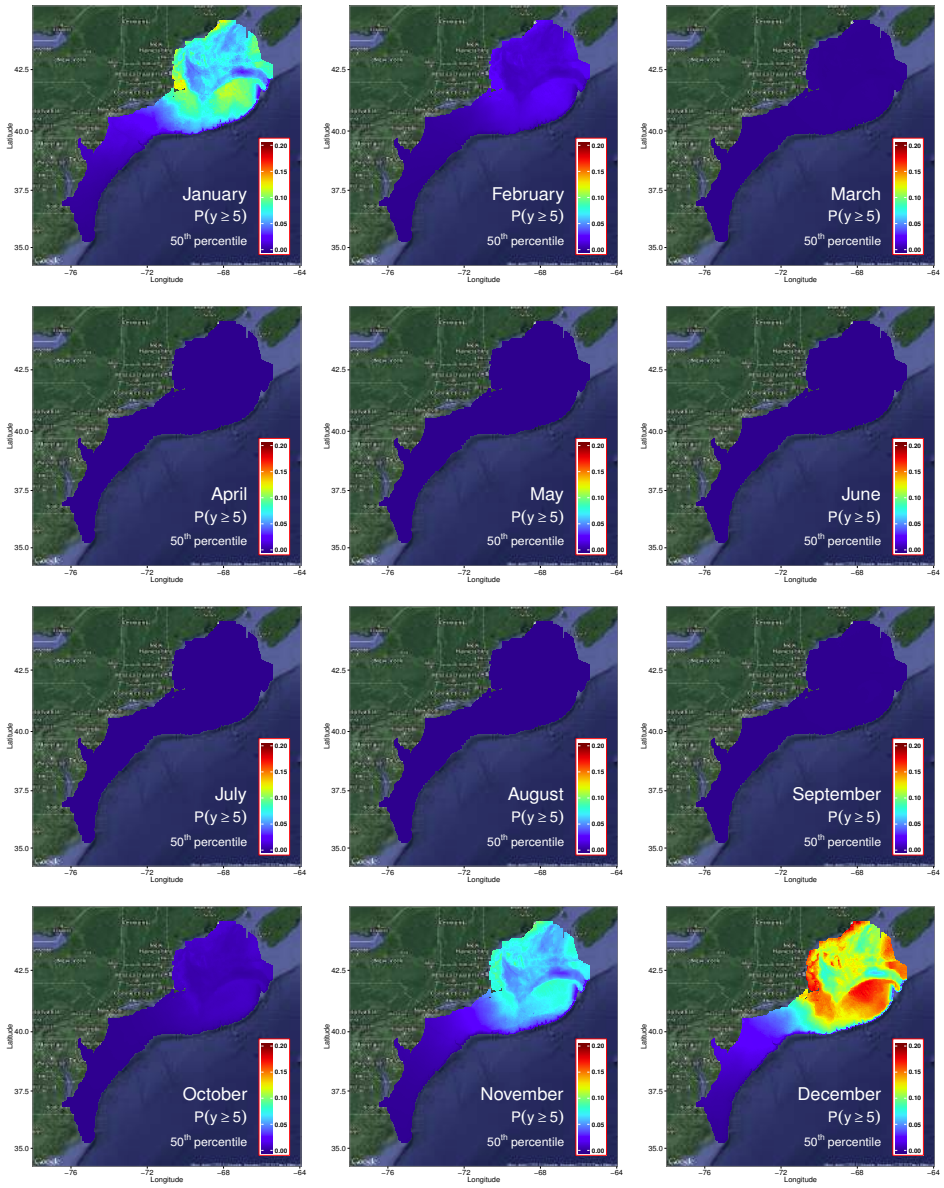

Figure 25: Black-legged Kittiwake: Probability of observing a large count during each month.

A.6 Bonaparte's Gull

A.6.1 One-year summary

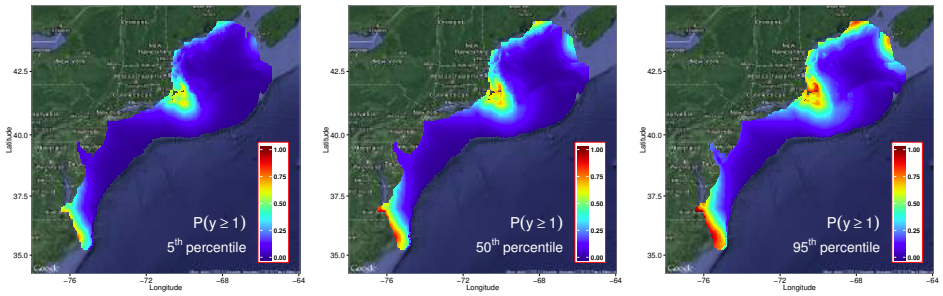

Figure 26: Bonaparte's Gull: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

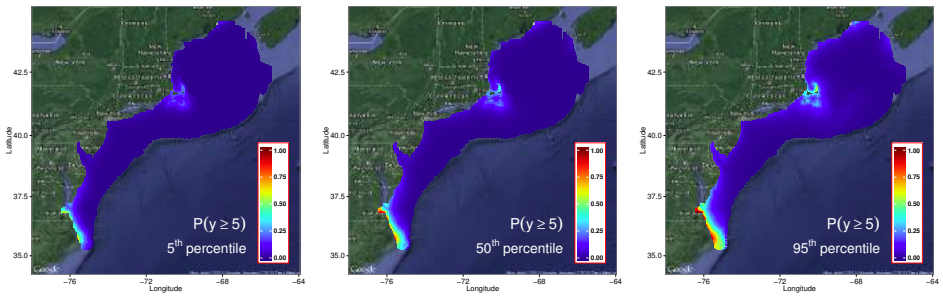

Figure 27: Bonaparte's Gull: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.6.2 Monthly observations

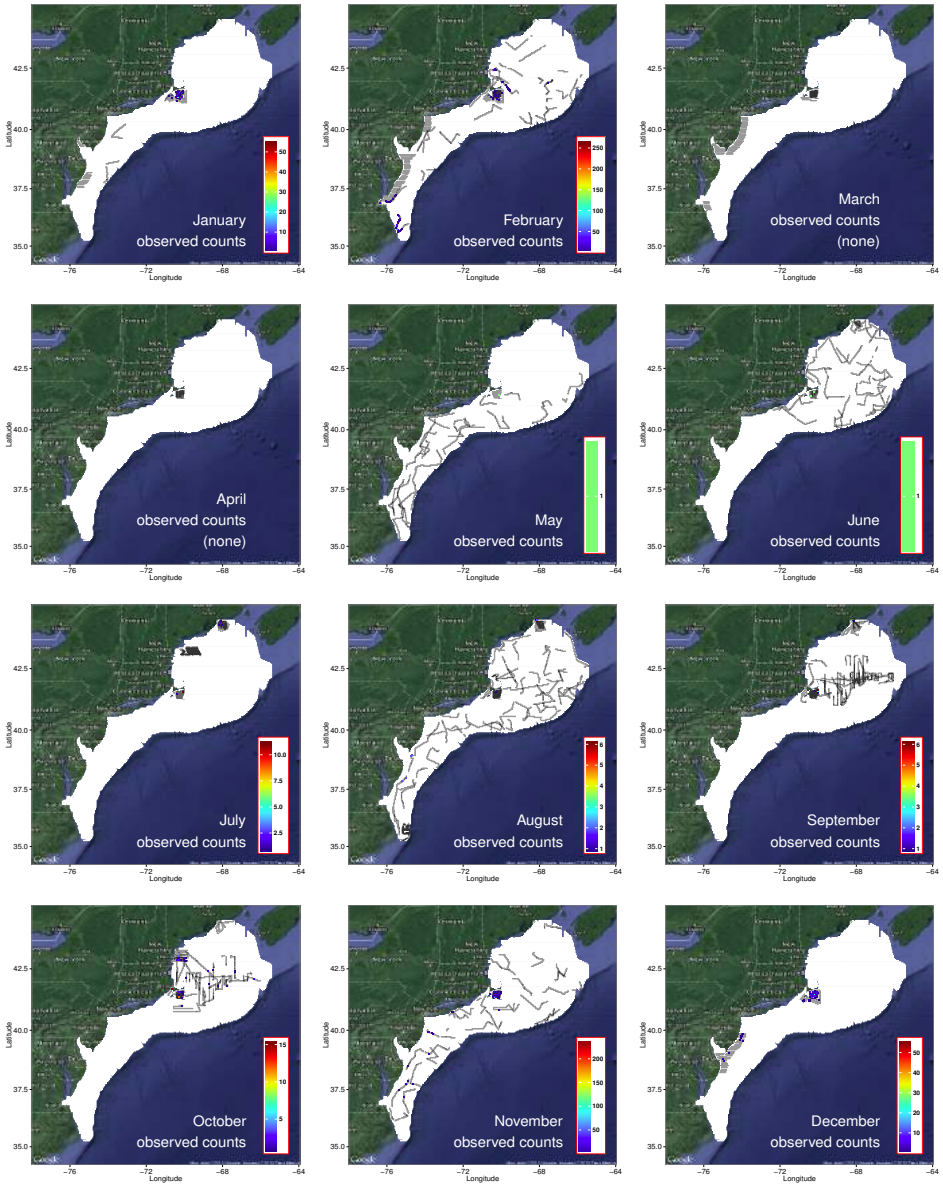

Figure 28: Bonaparte's Gull: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.6.3 Monthly exposure maps

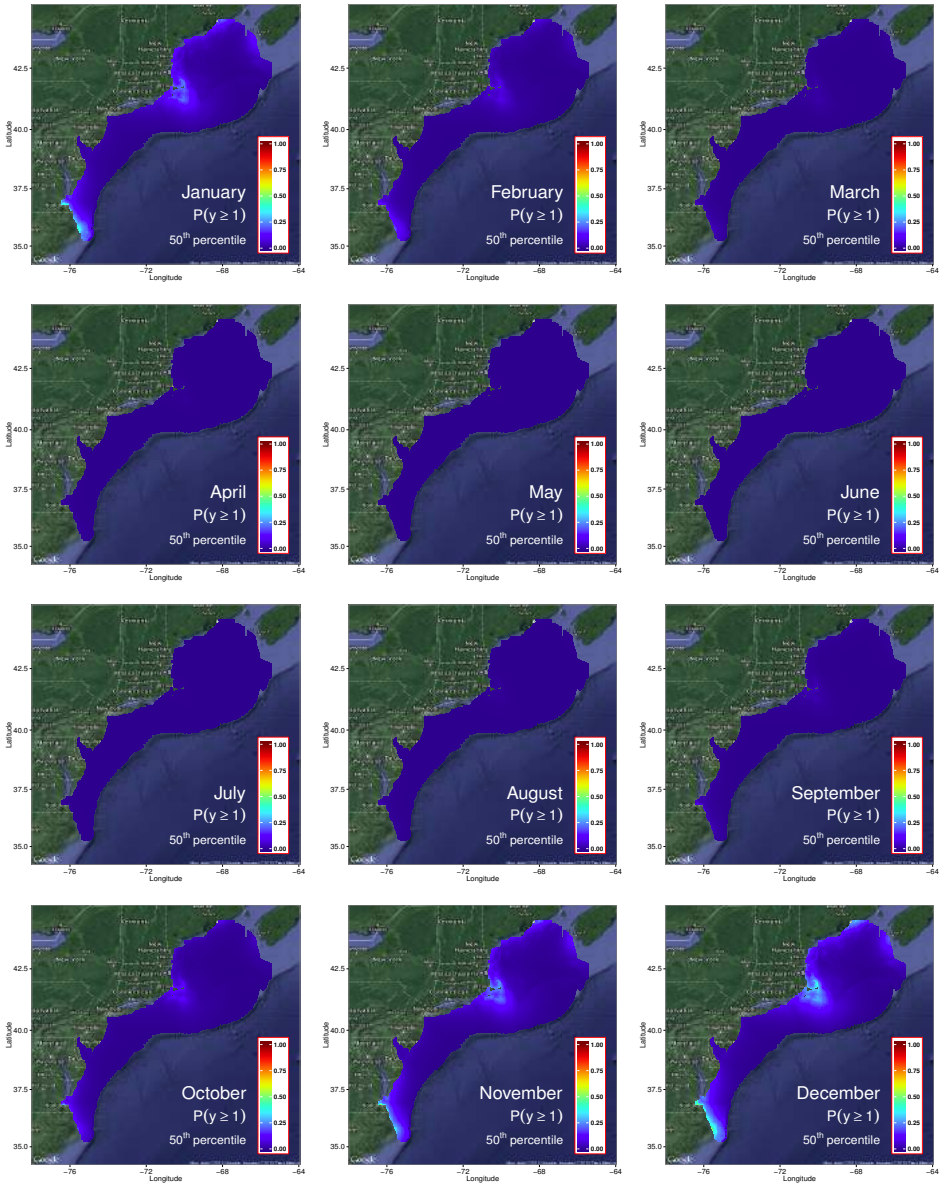

Figure 29: Bonaparte's Gull: Probability of observing at least one individual during each month.

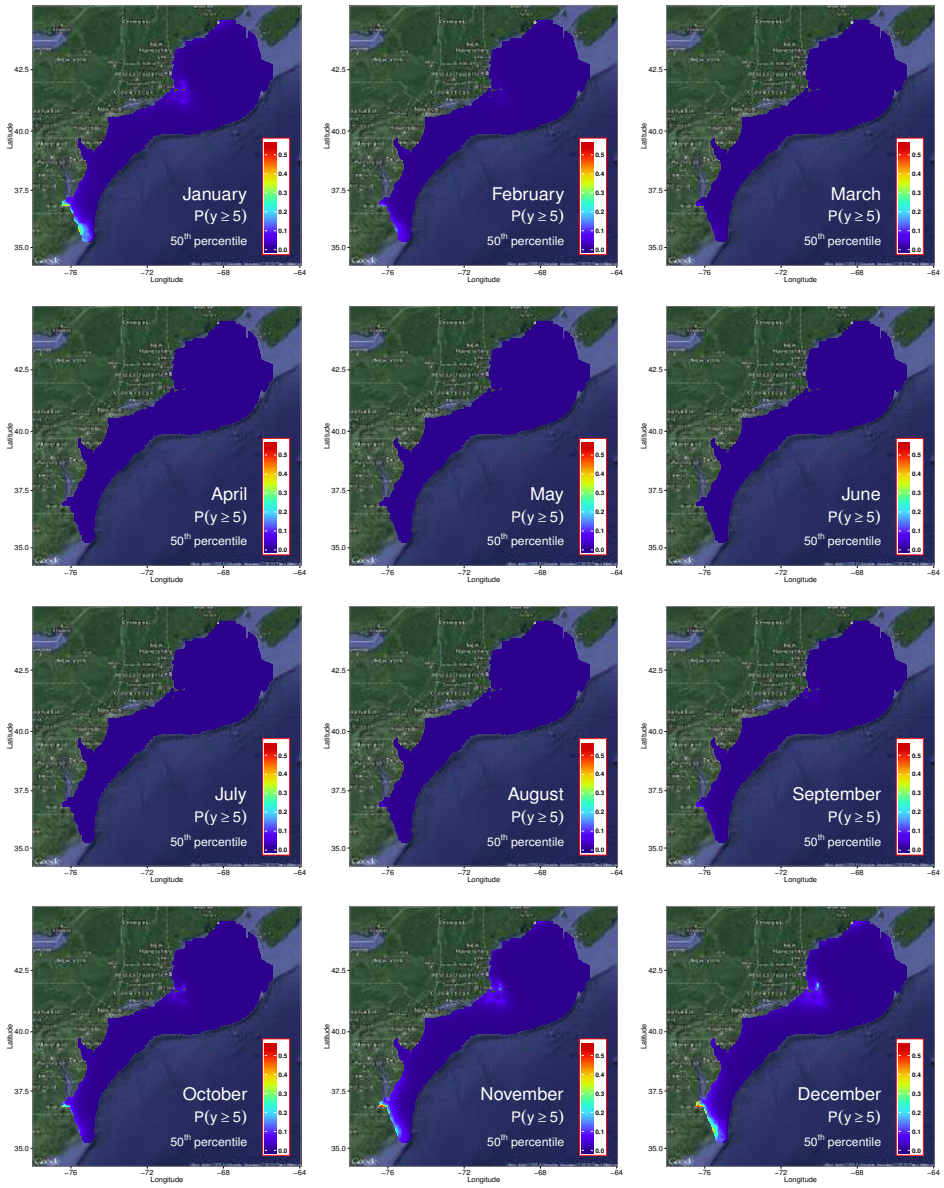

Figure 30: Bonaparte's Gull: Probability of observing a large count during each month.

A.7 Common Eider

A.7.1 One-year summary

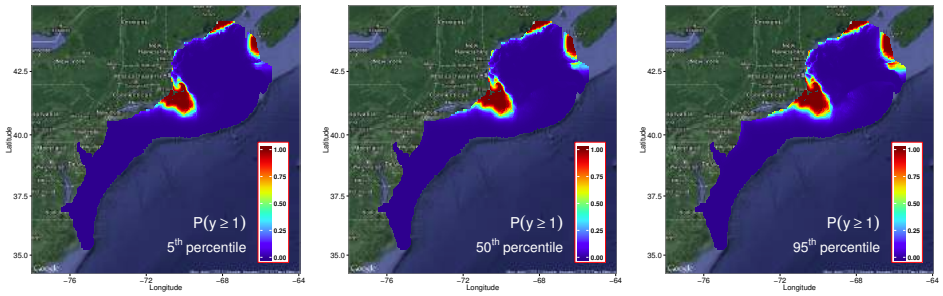

Figure 31: Common Eider: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.7.2 Monthly observations

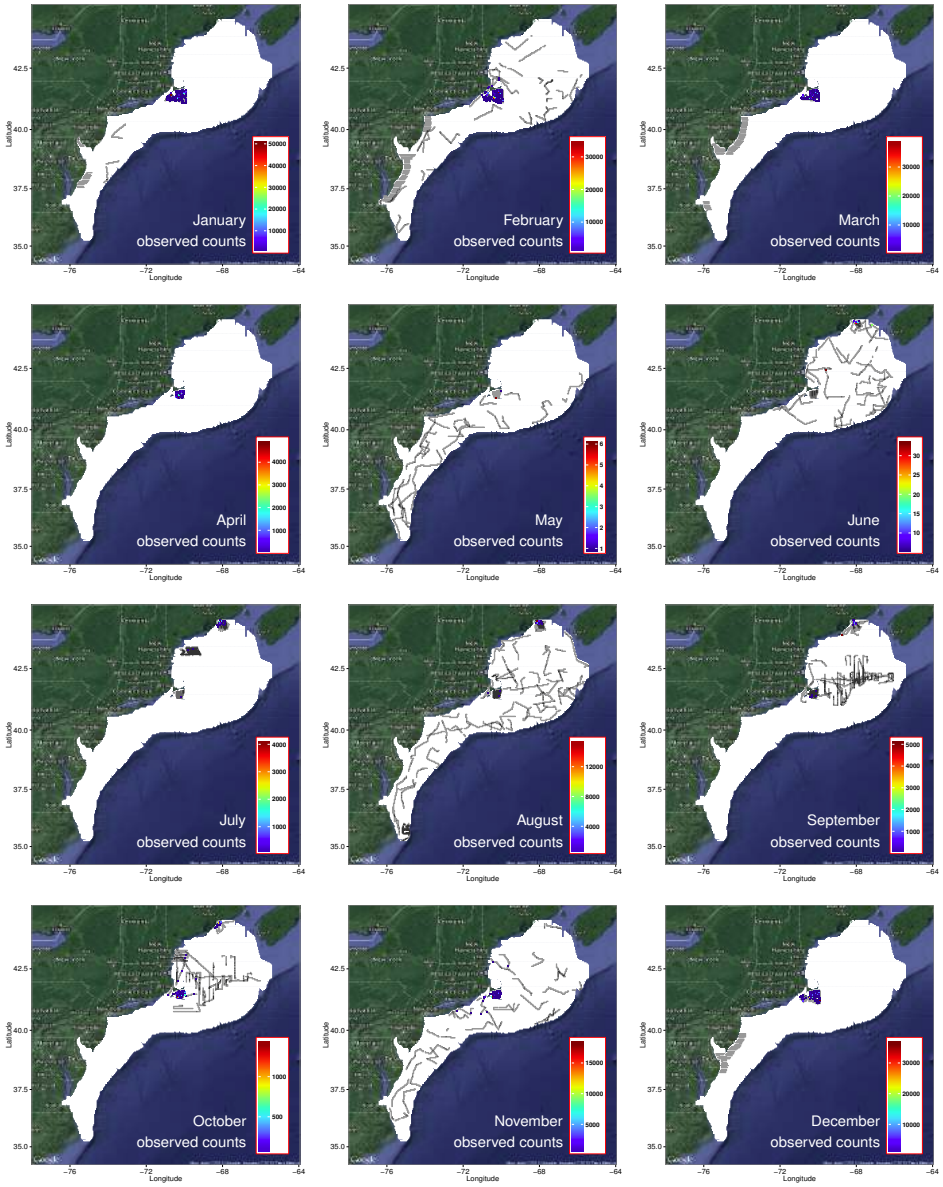

Figure 32: Common Eider: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.7.3 Monthly exposure maps

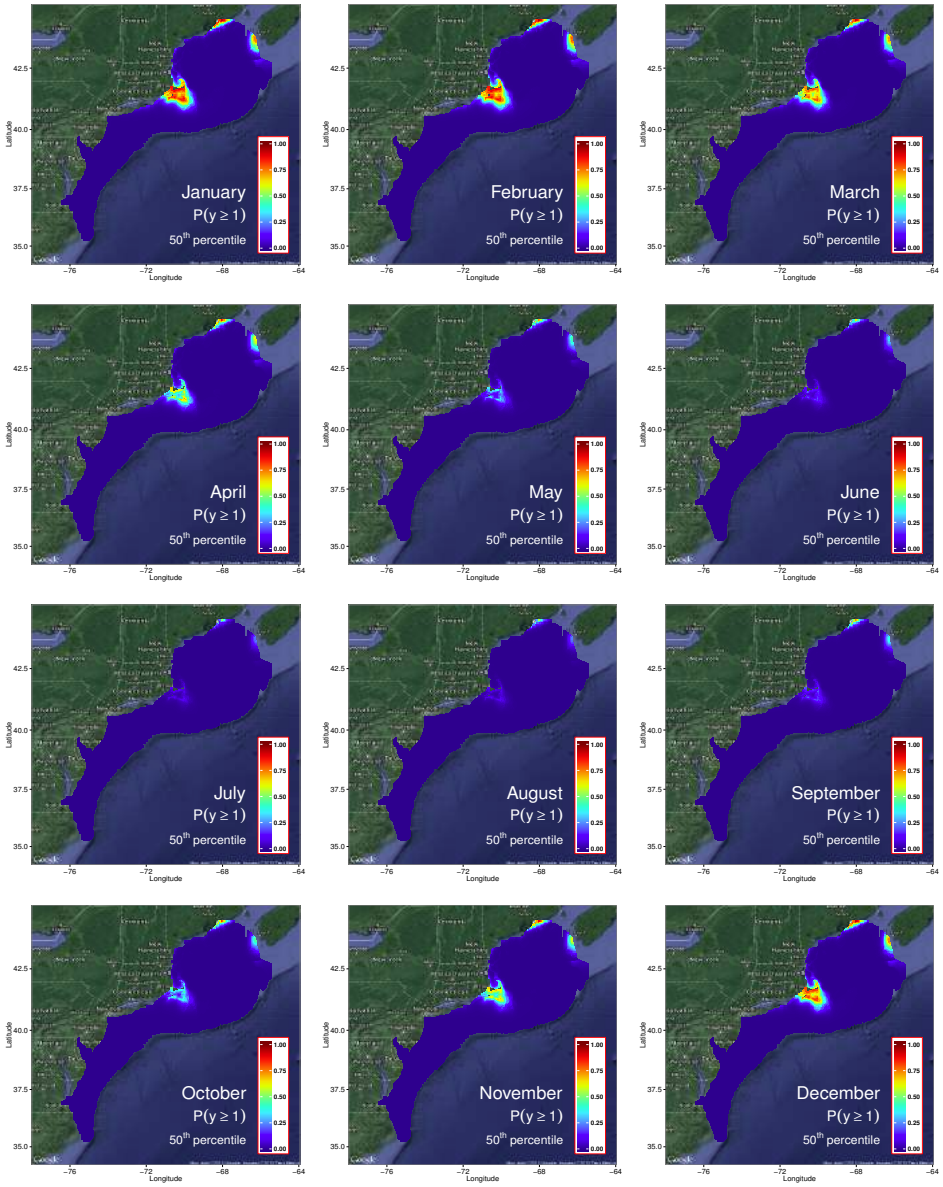

Figure 33: Common Eider: Probability of observing at least one individual during each month.

A.8 Common Loon

A.8.1 One-year summary

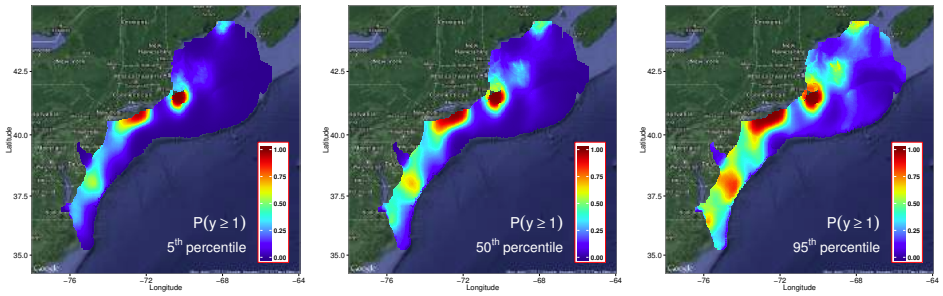

Figure 34: Common Loon: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

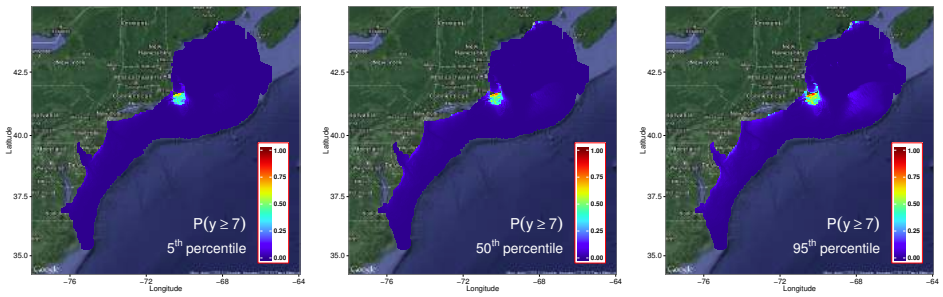

Figure 35: Common Loon: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.8.2 Monthly observations

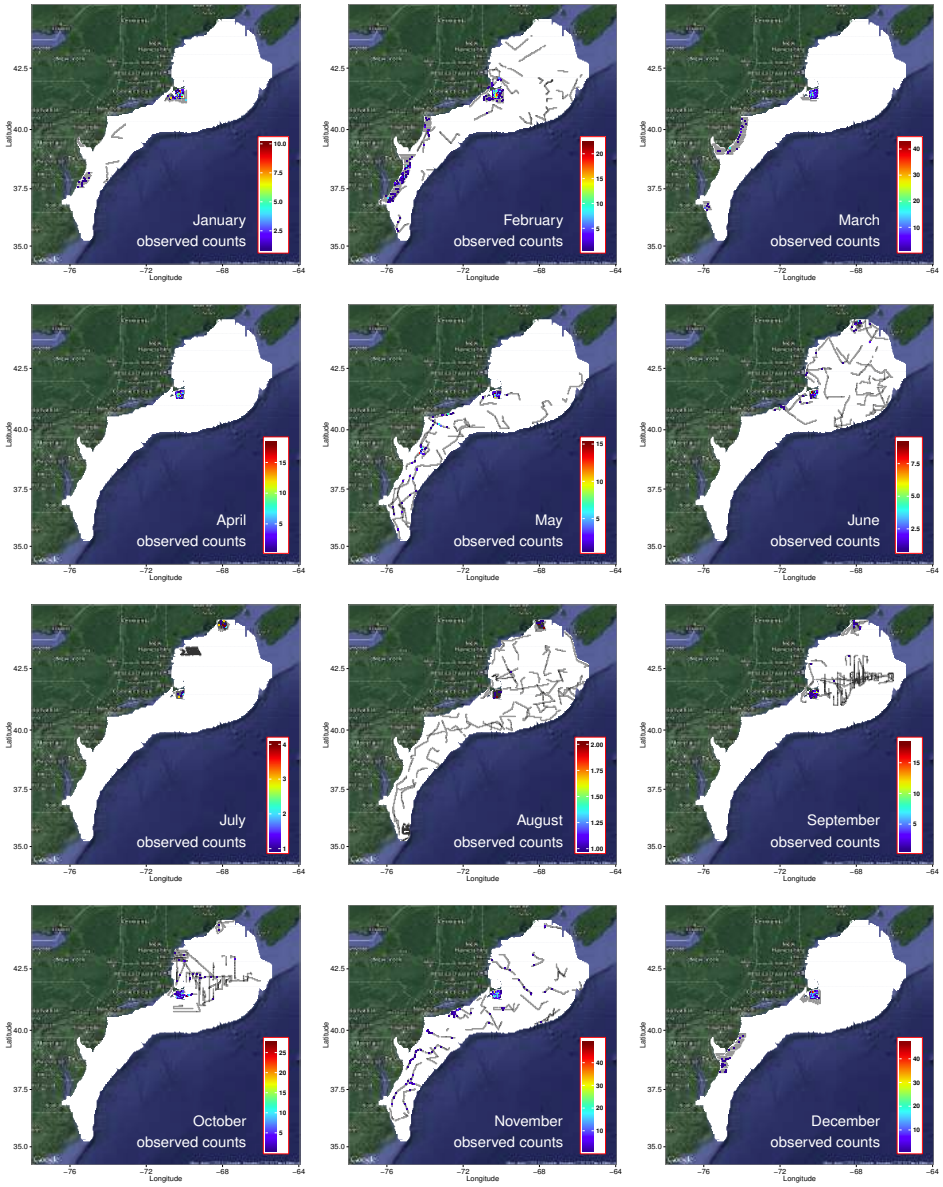

Figure 36: Common Loon: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.8.3 Monthly exposure maps

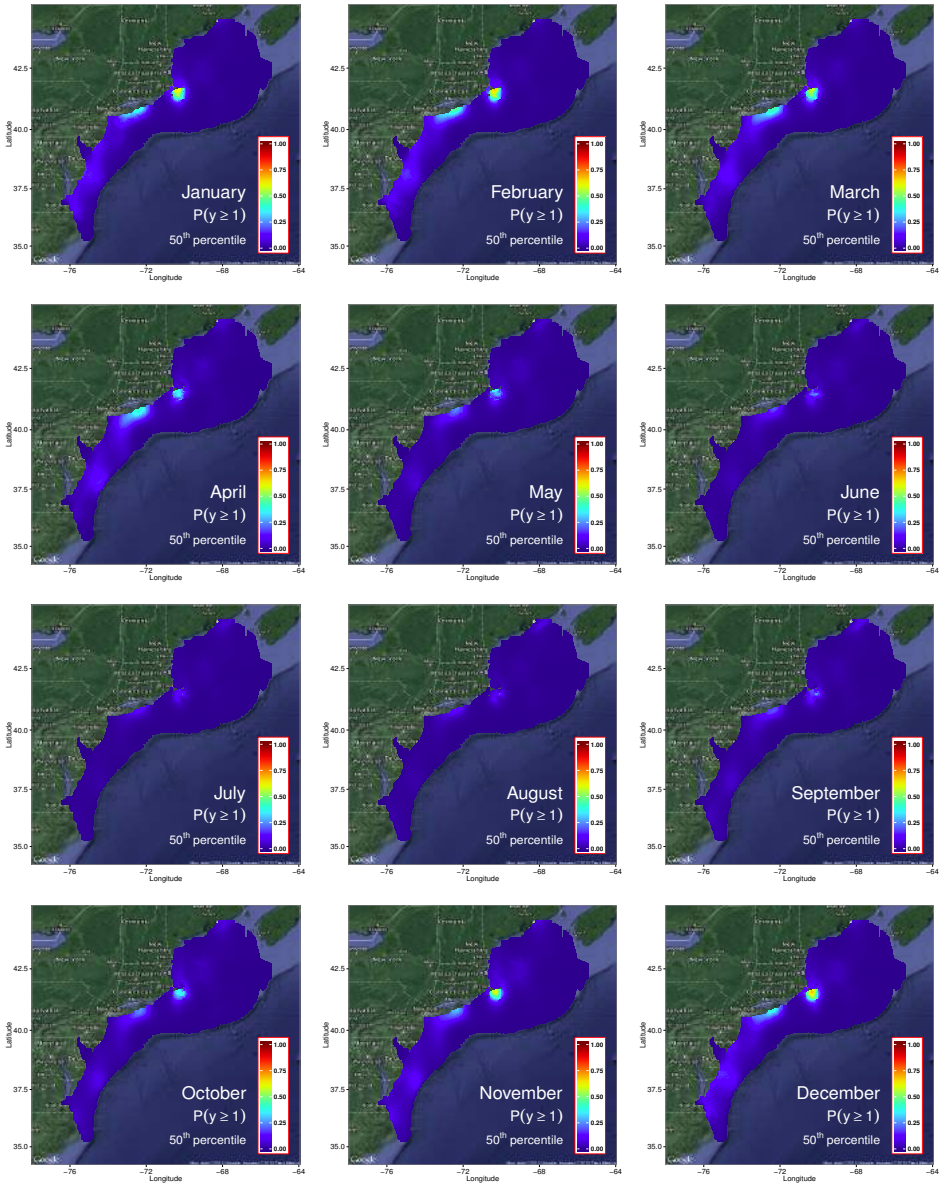

Figure 37: Common Loon: Probability of observing at least one individual during each month.

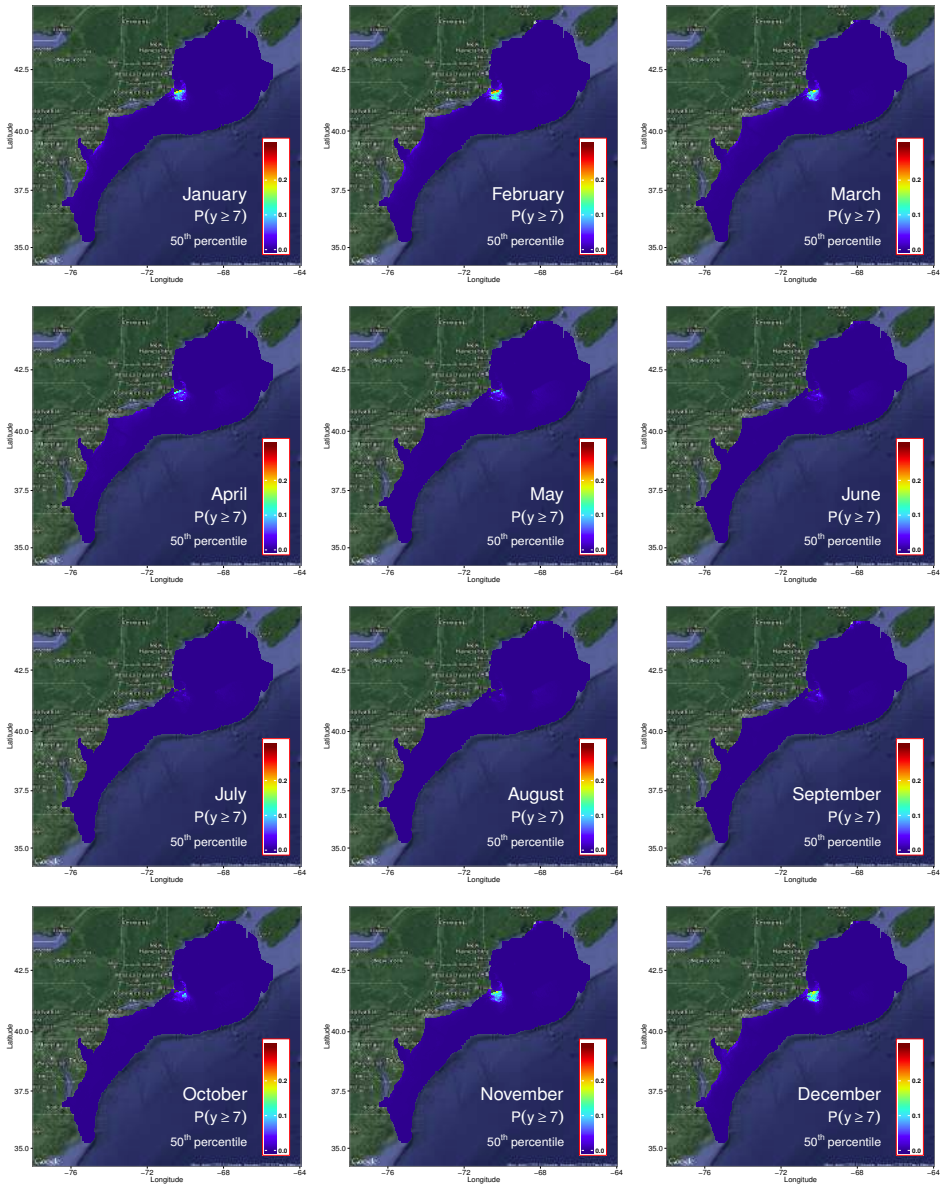

Figure 38: Common Loon: Probability of observing a large count during each month.

A.9 Common Tern

A.9.1 One-year summary

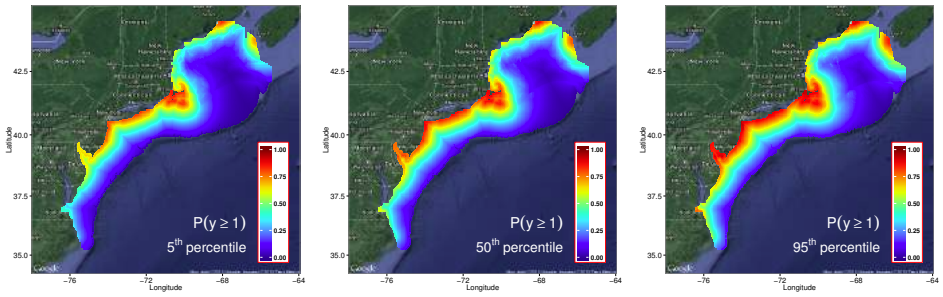

Figure 39: Common Tern: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

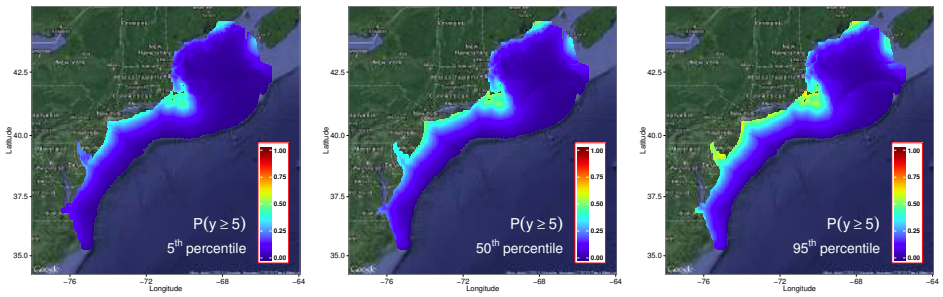

Figure 40: Common Tern: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.9.2 Monthly observations

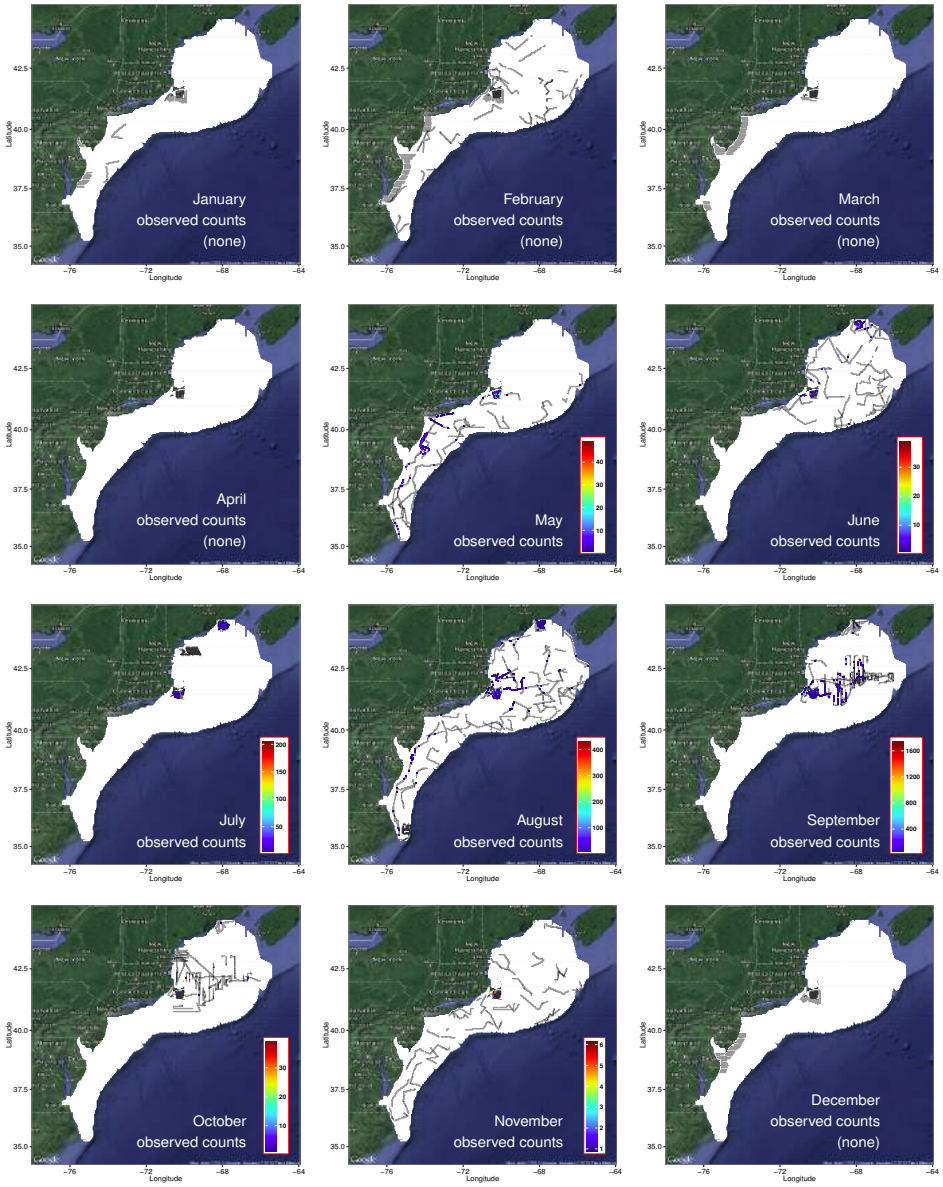

Figure 41: Common Tern: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.9.3 Monthly exposure maps

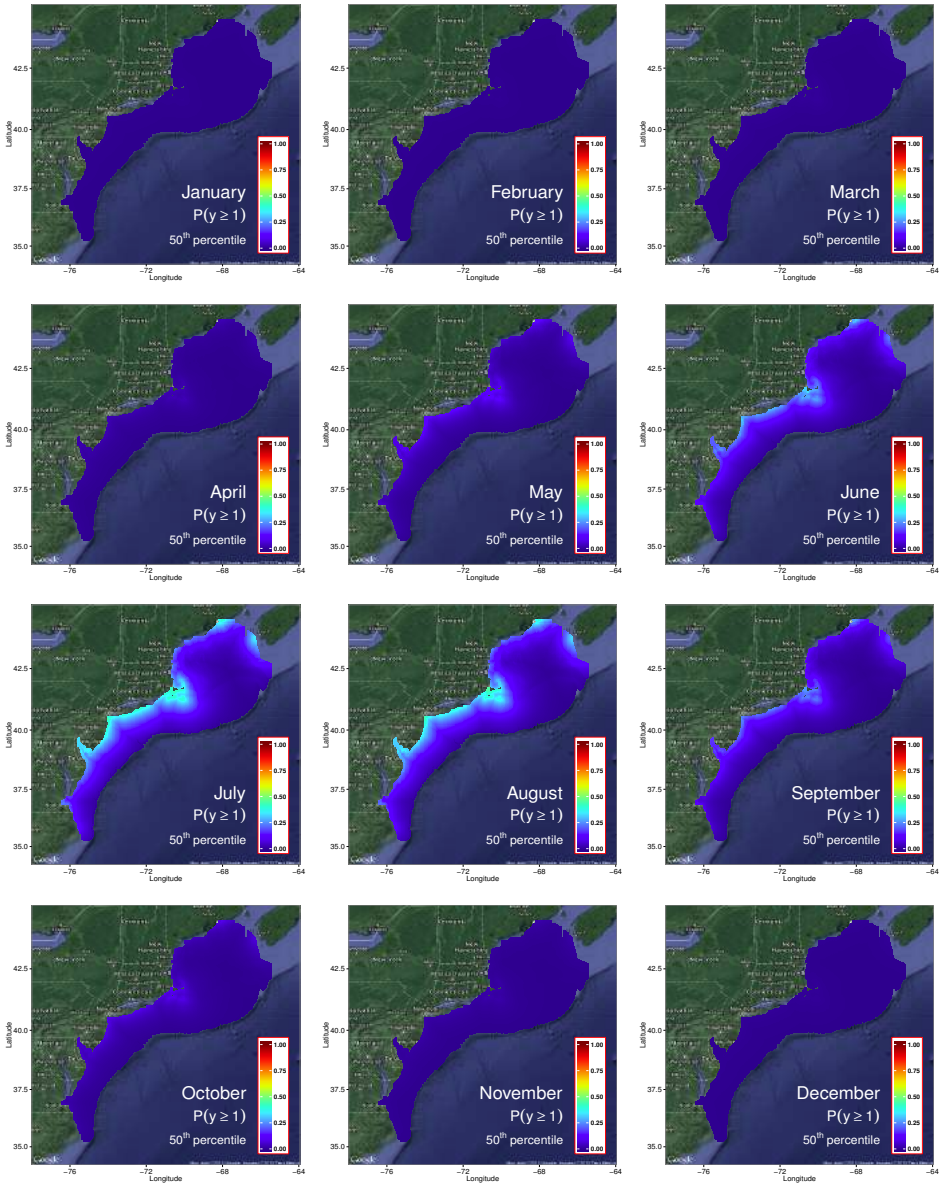

Figure 42: Common Tern: Probability of observing at least one individual during each month.

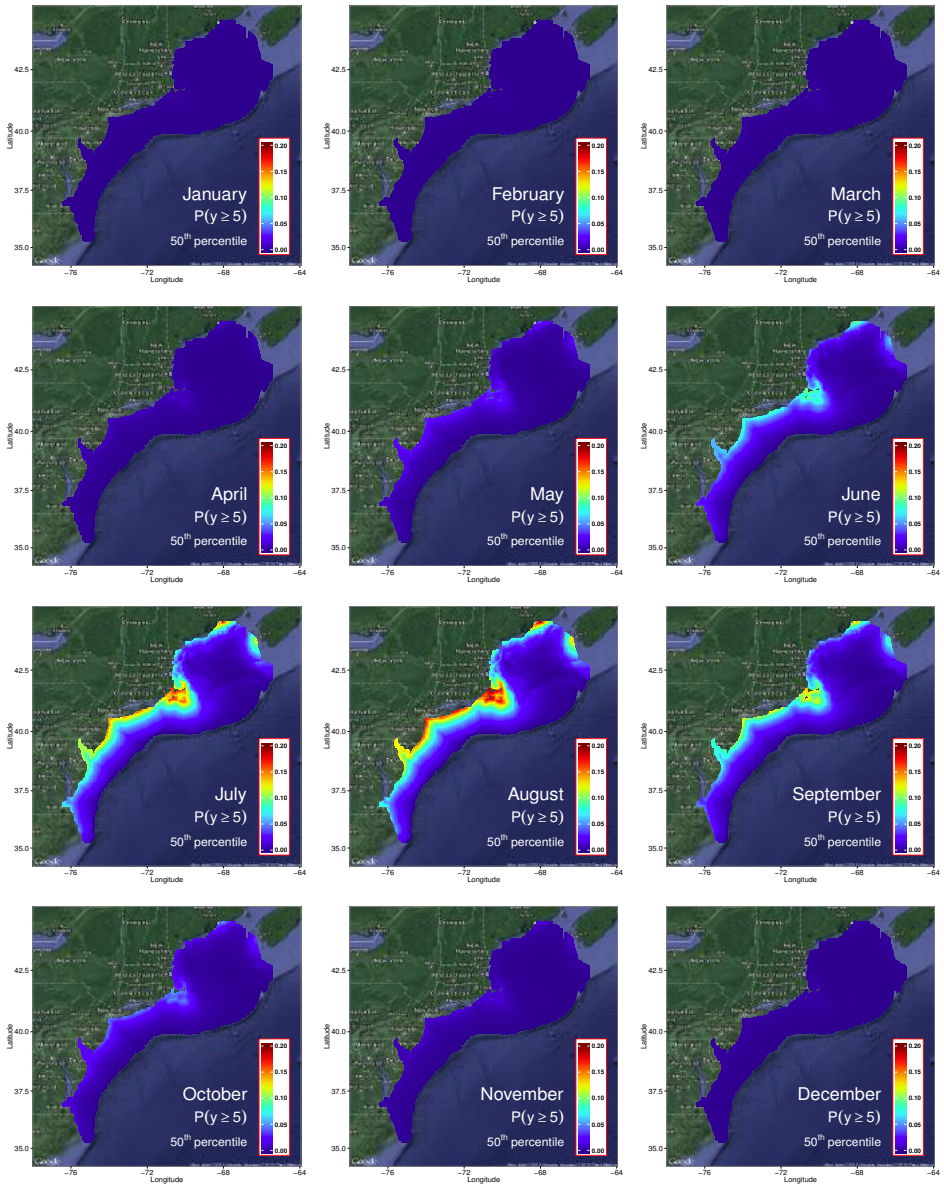

Figure 43: Common Tern: Probability of observing a large count during each month.

A.10 Cory's Shearwater

A.10.1 One-year summary

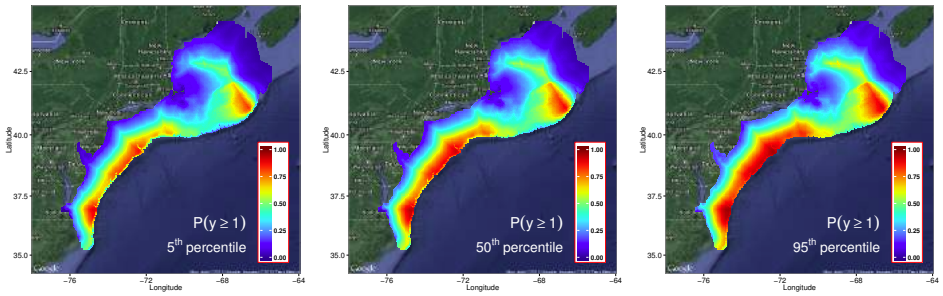

Figure 44: Cory's Shearwater: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

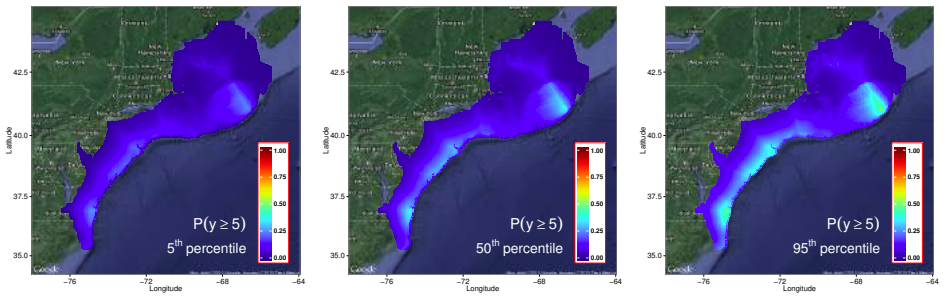

Figure 45: Cory's Shearwater: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.10.2 Monthly observations

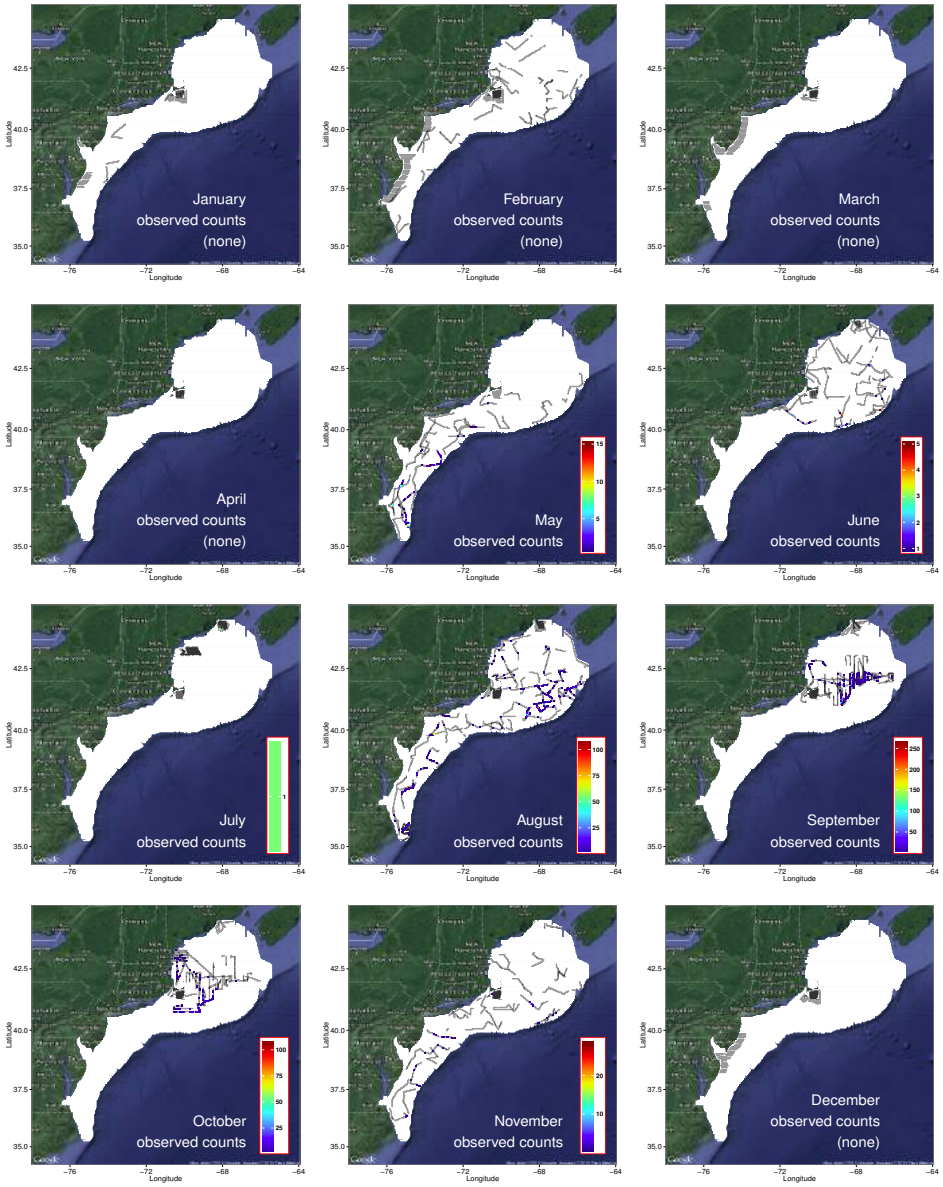

Figure 46: Cory's Shearwater: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.10.3 Monthly exposure maps

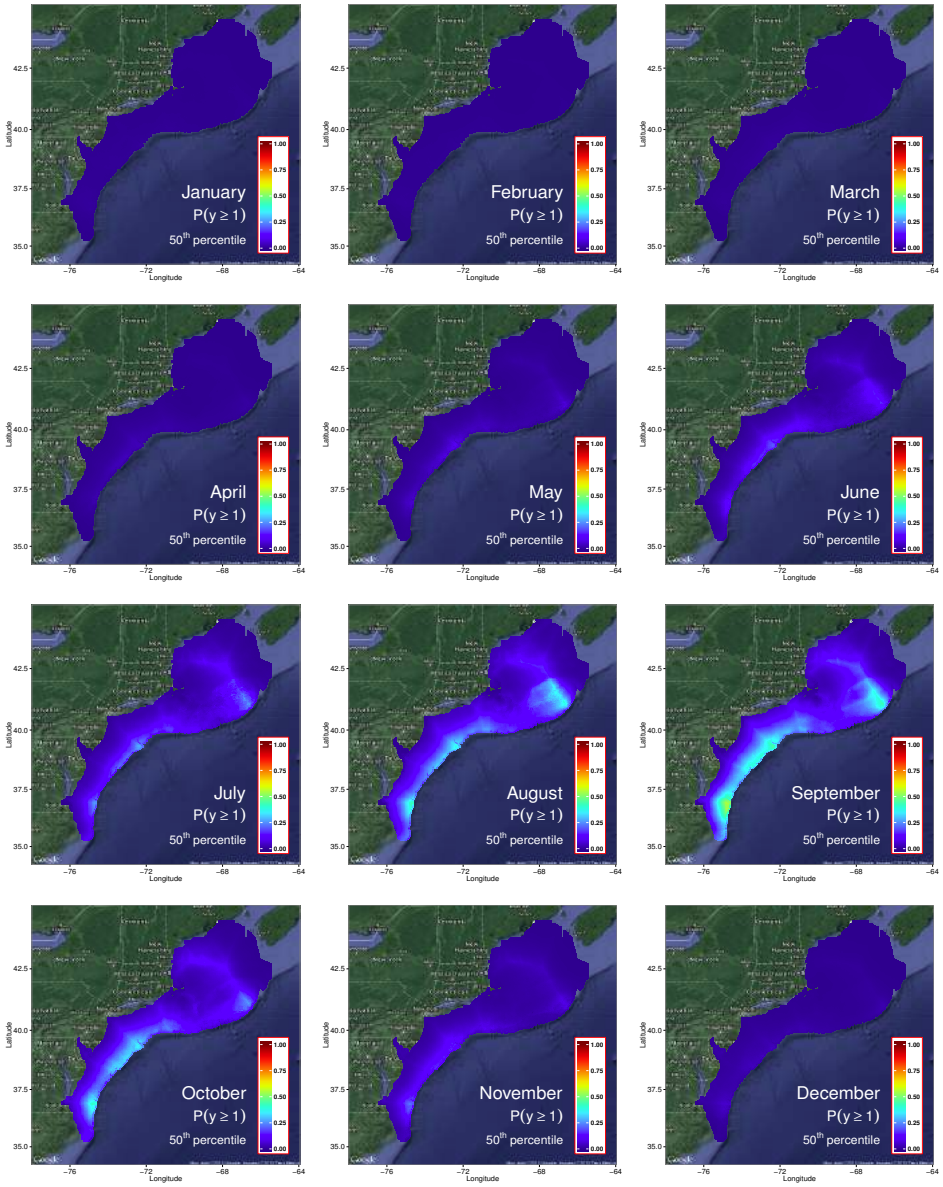

Figure 47: Cory's Shearwater: Probability of observing at least one individual during each month.

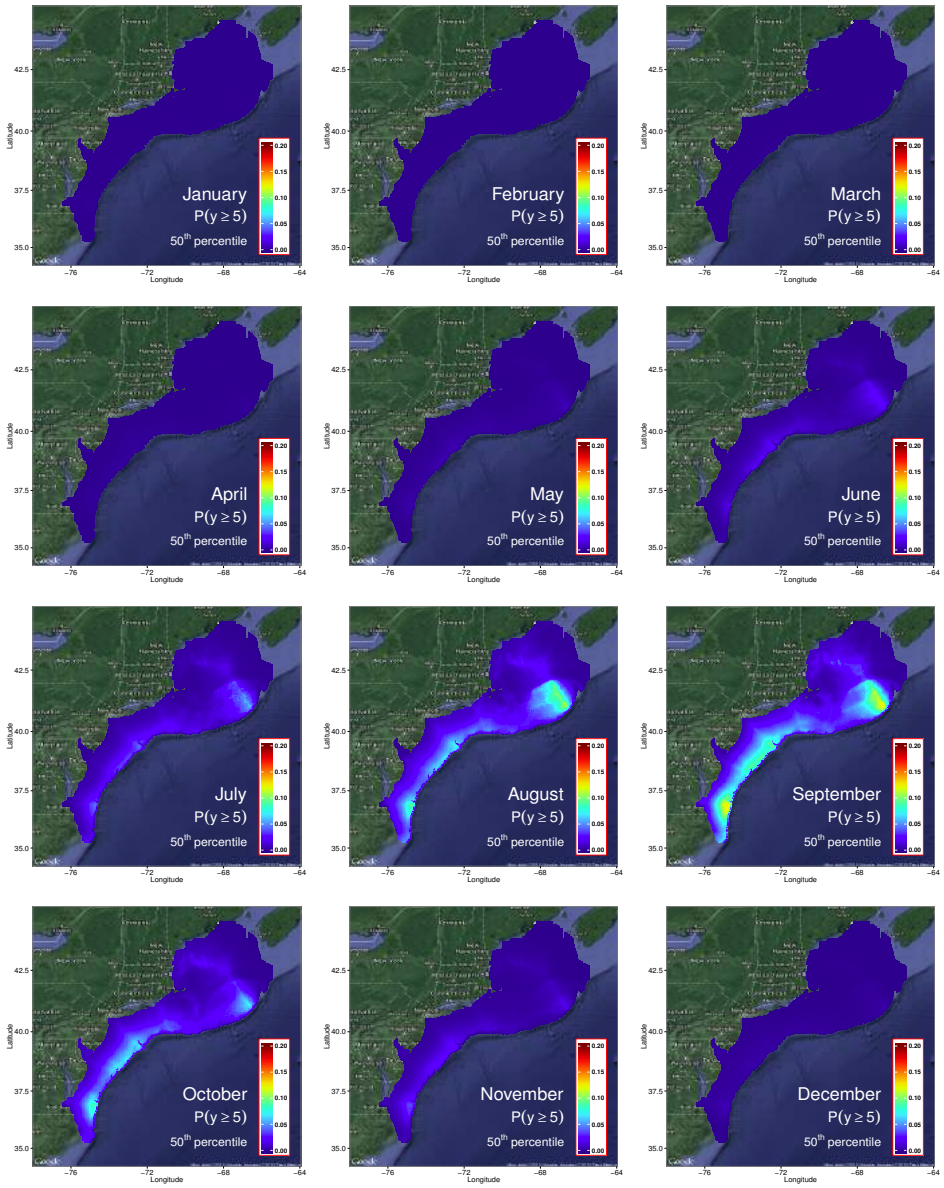

Figure 48: Cory's Shearwater: Probability of observing a large count during each month.

A.11 Double-crested Cormorant

A.11.1 One-year summary

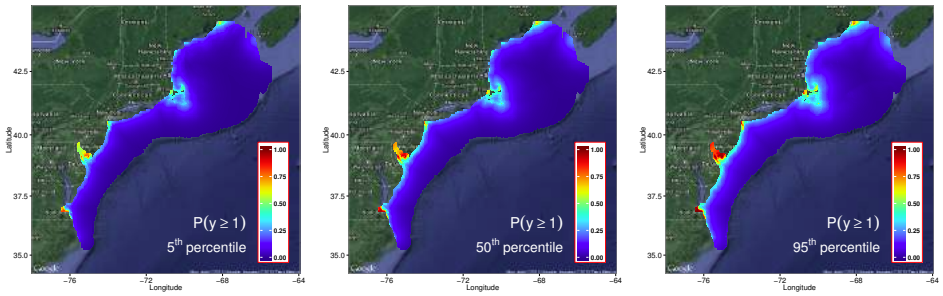

Figure 49: Double-crested Cormorant: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

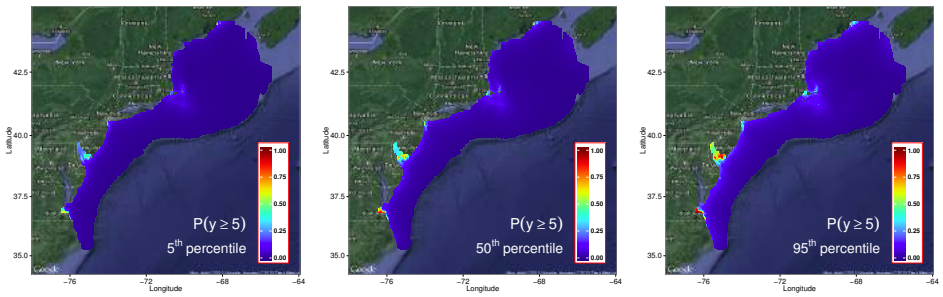

Figure 50: Double-crested Cormorant: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.11.2 Monthly observations

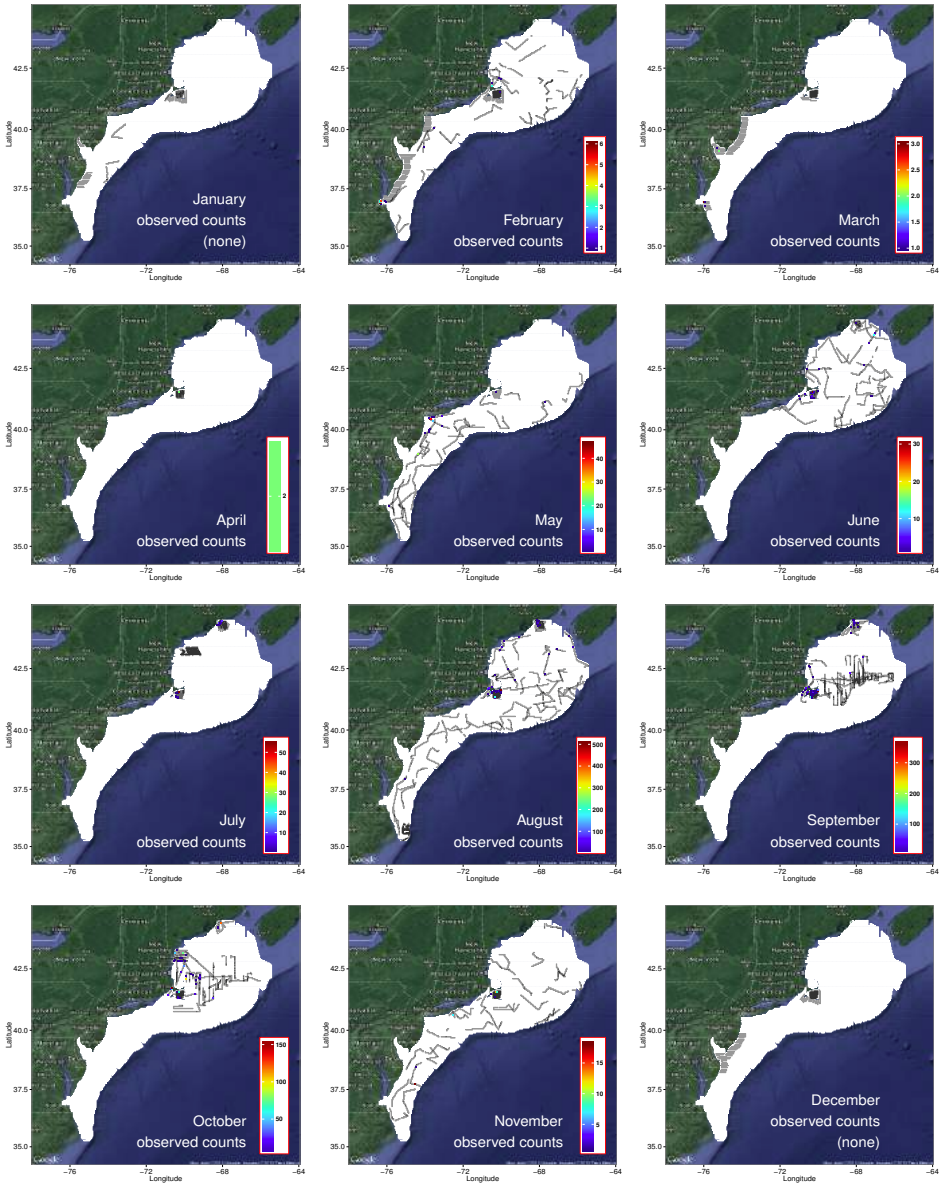

Figure 51: Double-crested Cormorant: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.11.3 Monthly exposure maps

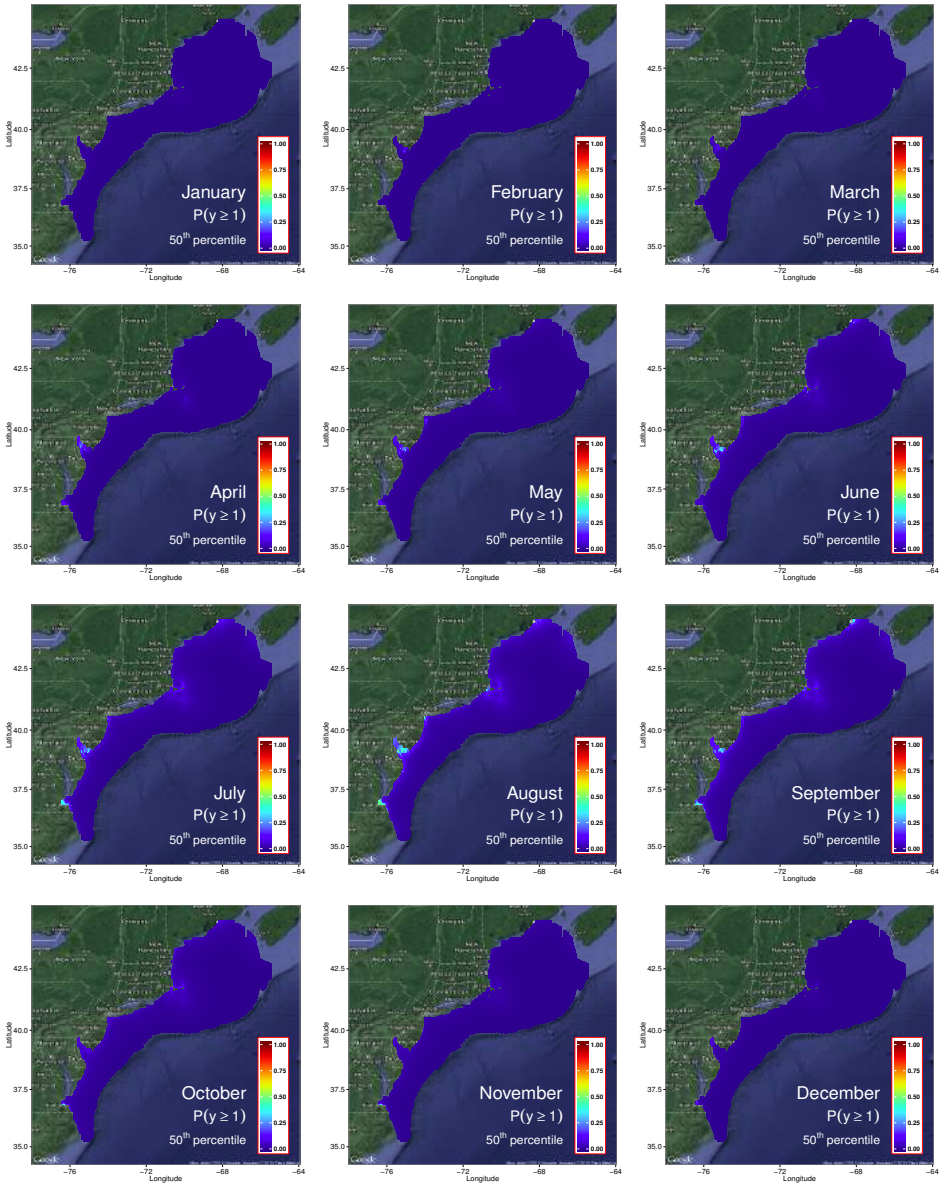

Figure 52: Double-crested Cormorant: Probability of observing at least one individual during each month.

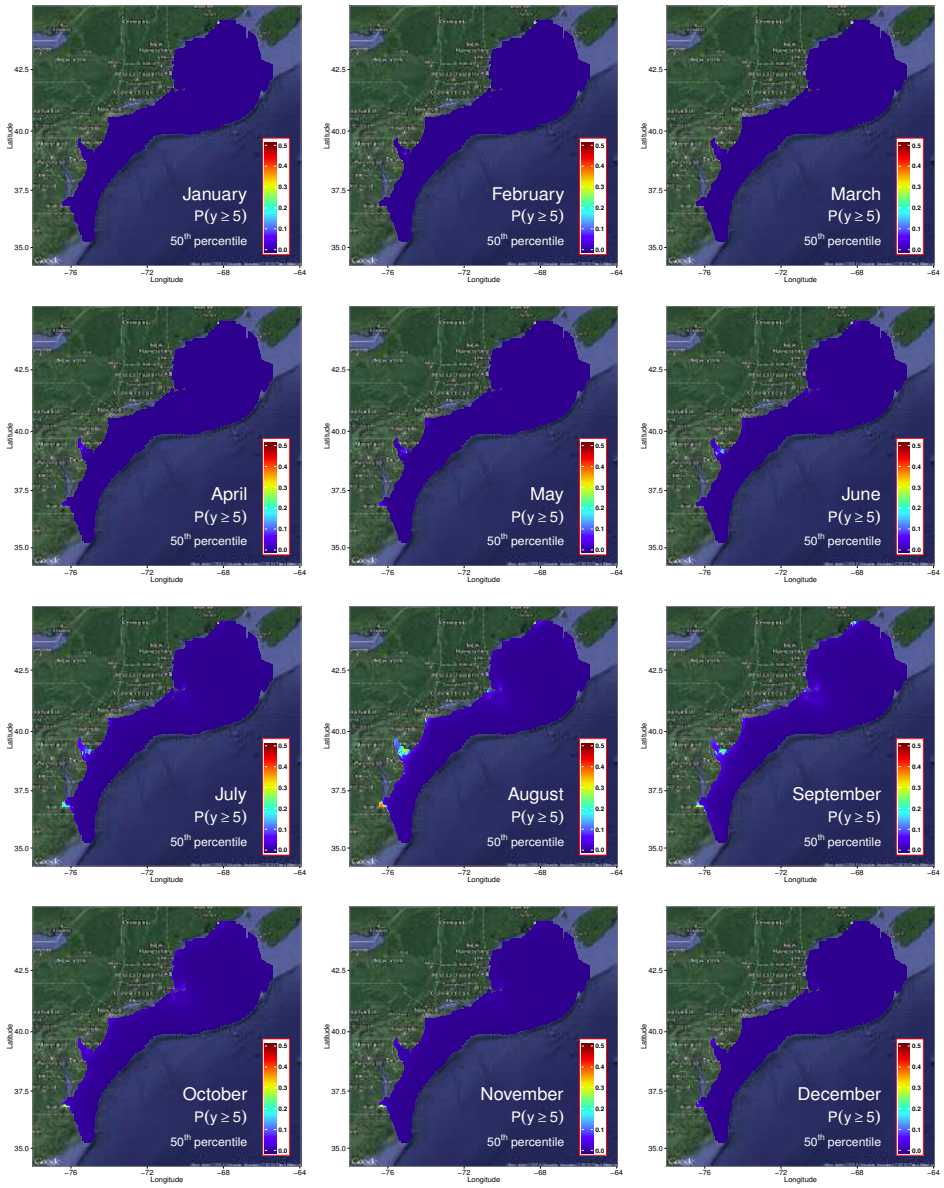

Figure 53: Double-crested Cormorant: Probability of observing a large count during each month.

A.12 Dovekie

A.12.1 One-year summary

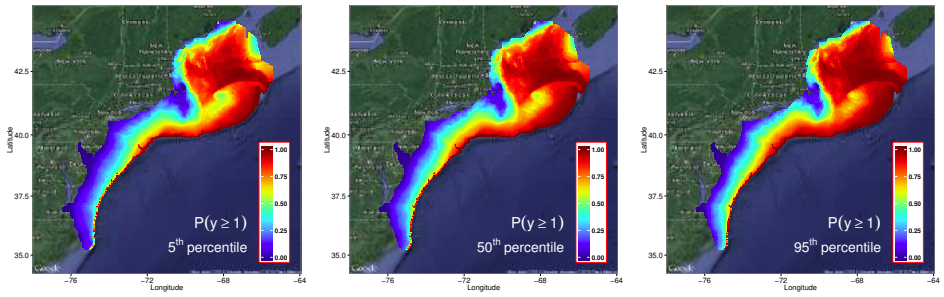

Figure 54: Dovekie: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

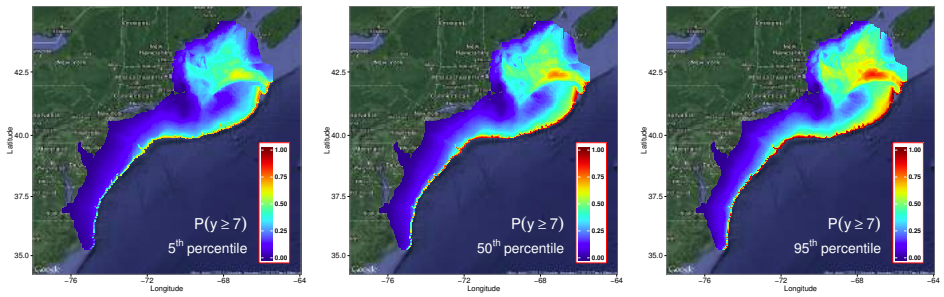

Figure 55: Dovekie: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.12.2 Monthly observations

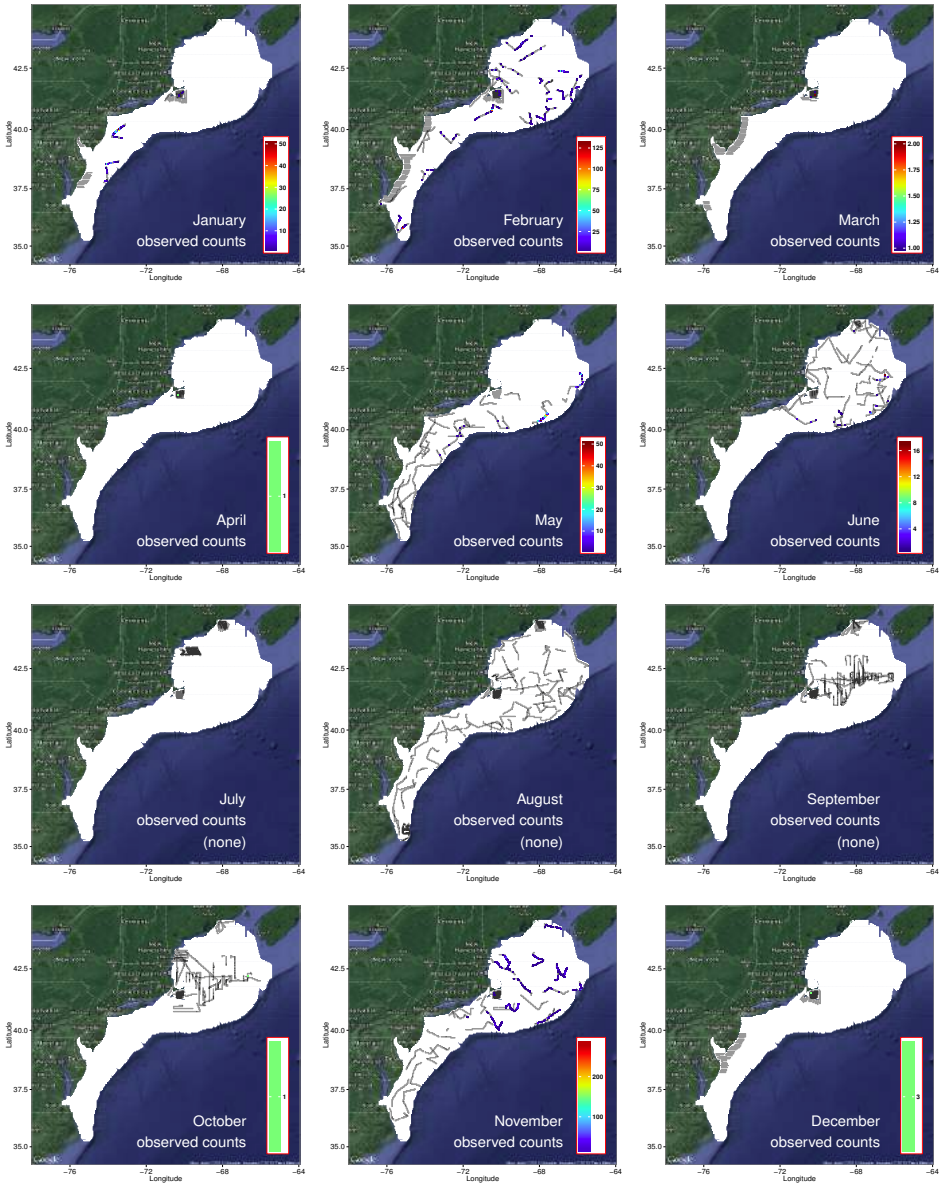

Figure 56: Dovekie: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.12.3 Monthly exposure maps

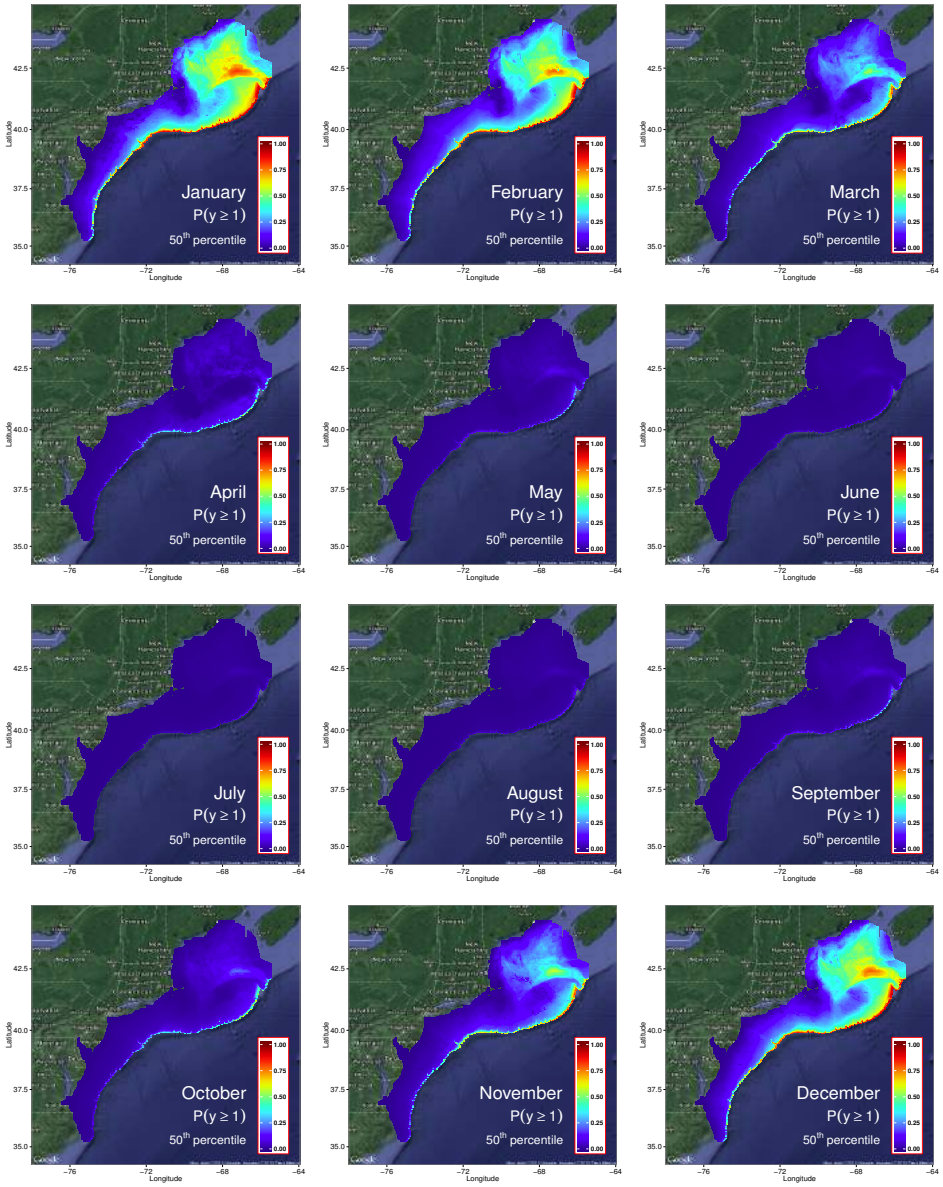

Figure 57: Dovekie: Probability of observing at least one individual during each month.

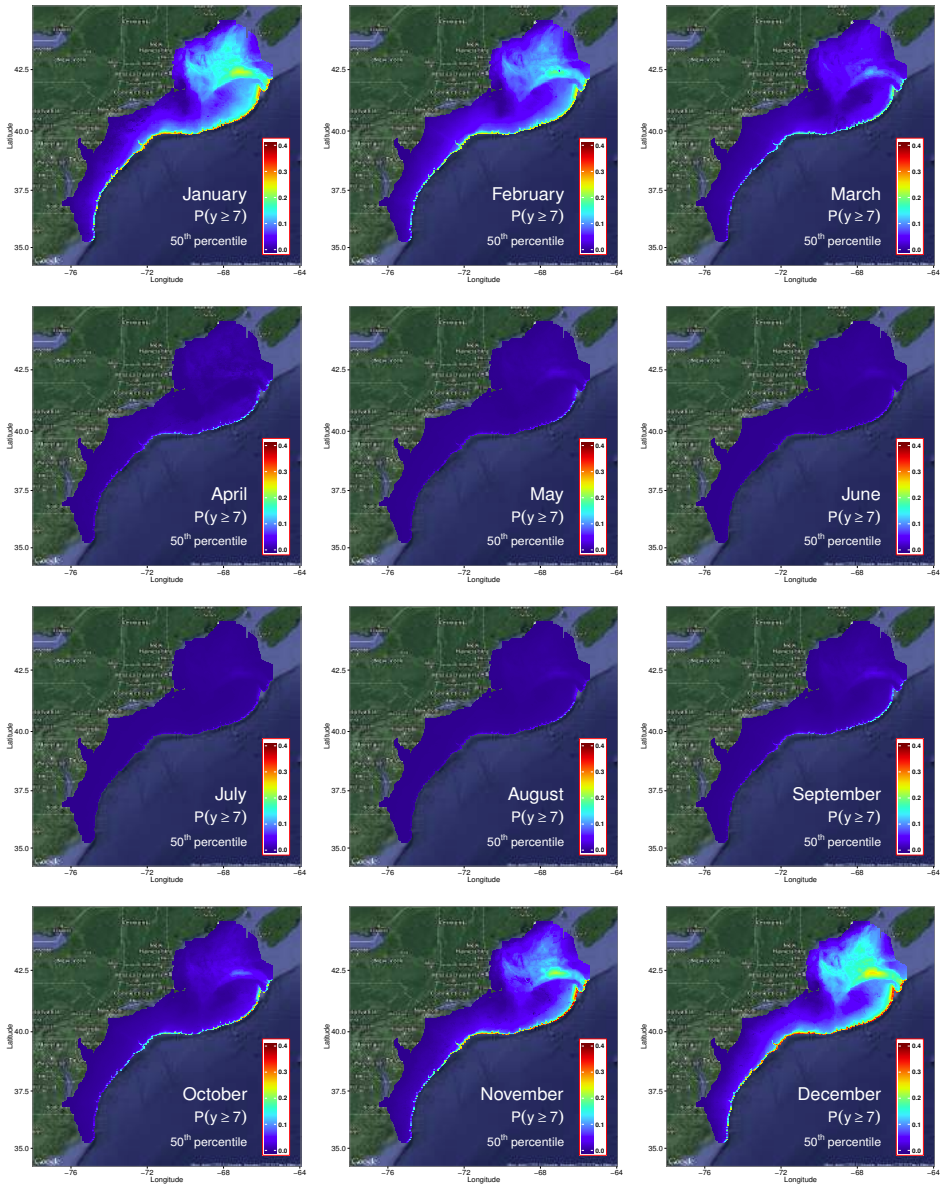

Figure 58: Dovekie: Probability of observing a large count during each month.

A.13 Great Black-backed Gull

A.13.1 One-year summary

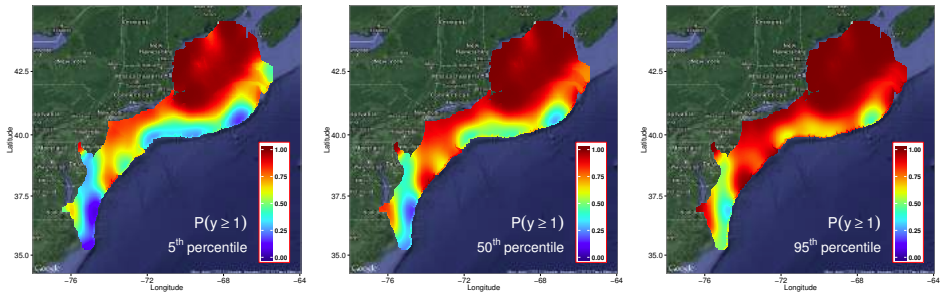

Figure 59: Great Black-backed Gull: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

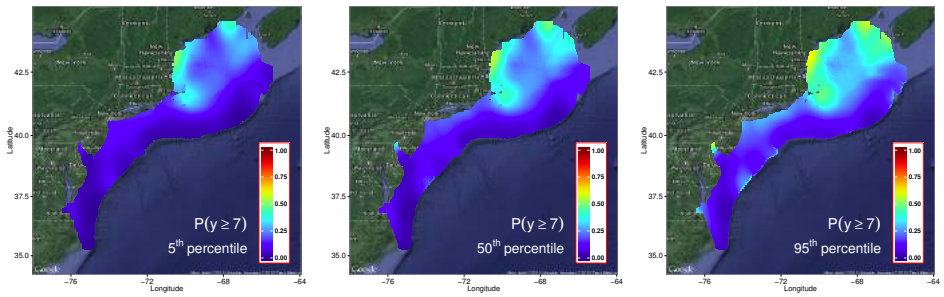

Figure 60: Great Black-backed Gull: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.13.2 Monthly observations

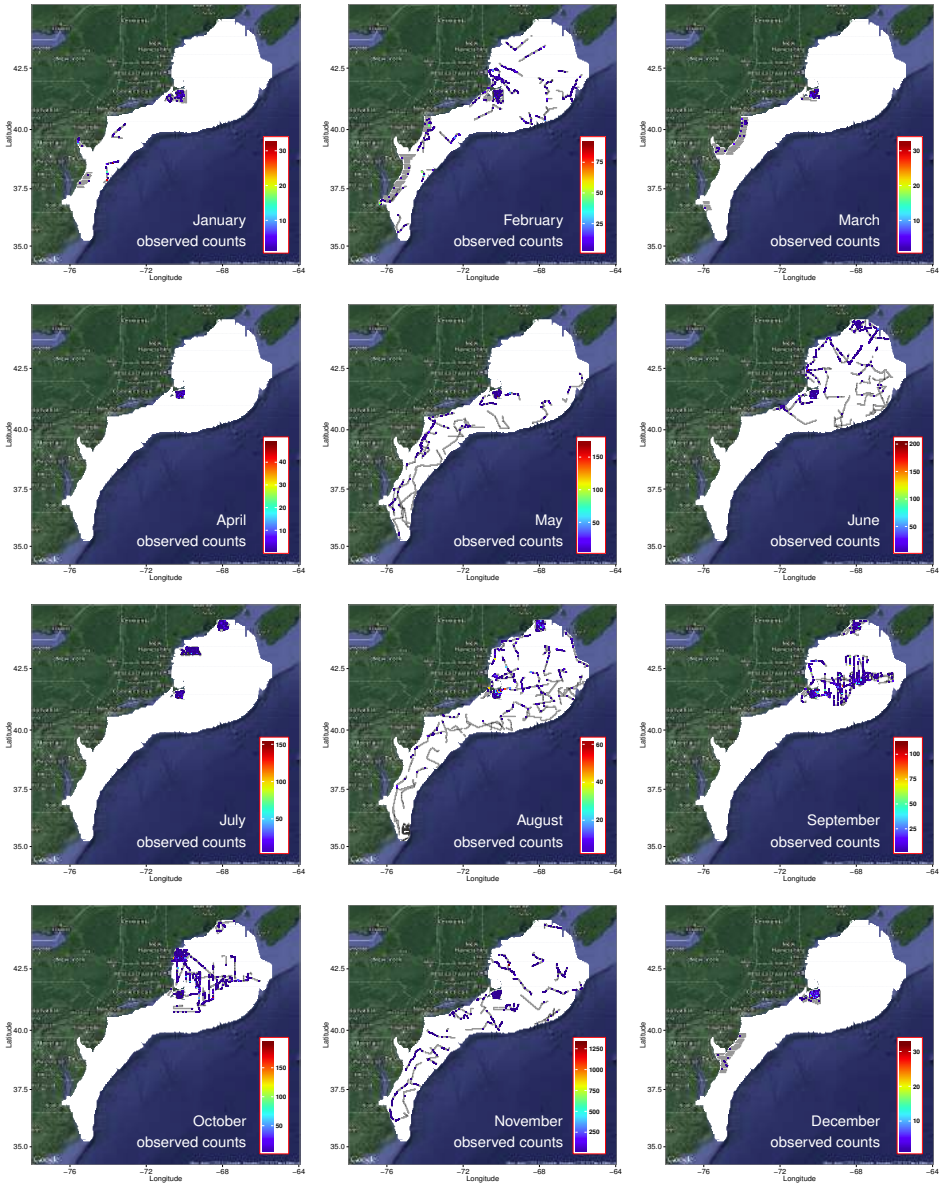

Figure 61: Great Black-backed Gull: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.13.3 Monthly exposure maps

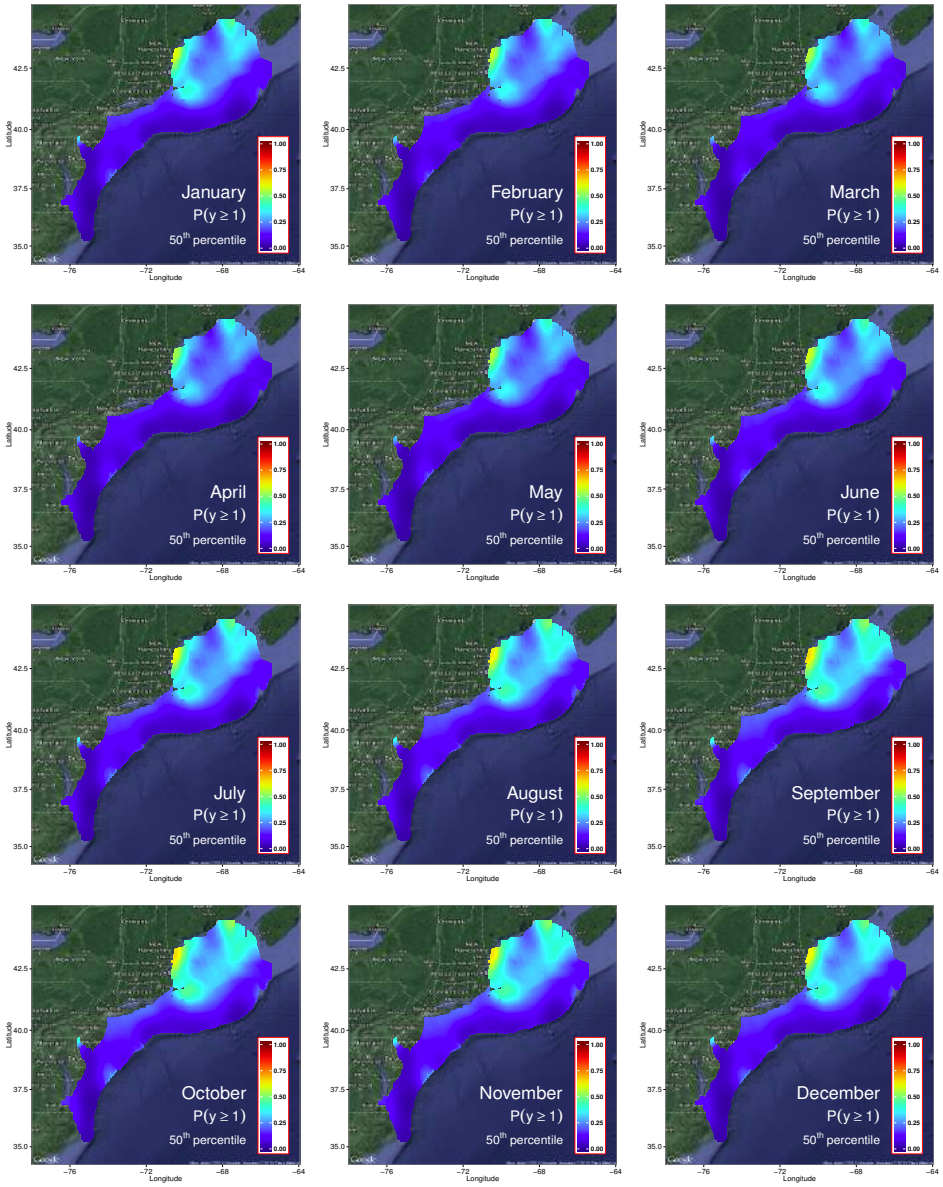

Figure 62: Great Black-backed Gull: Probability of observing at least one individual during each month.

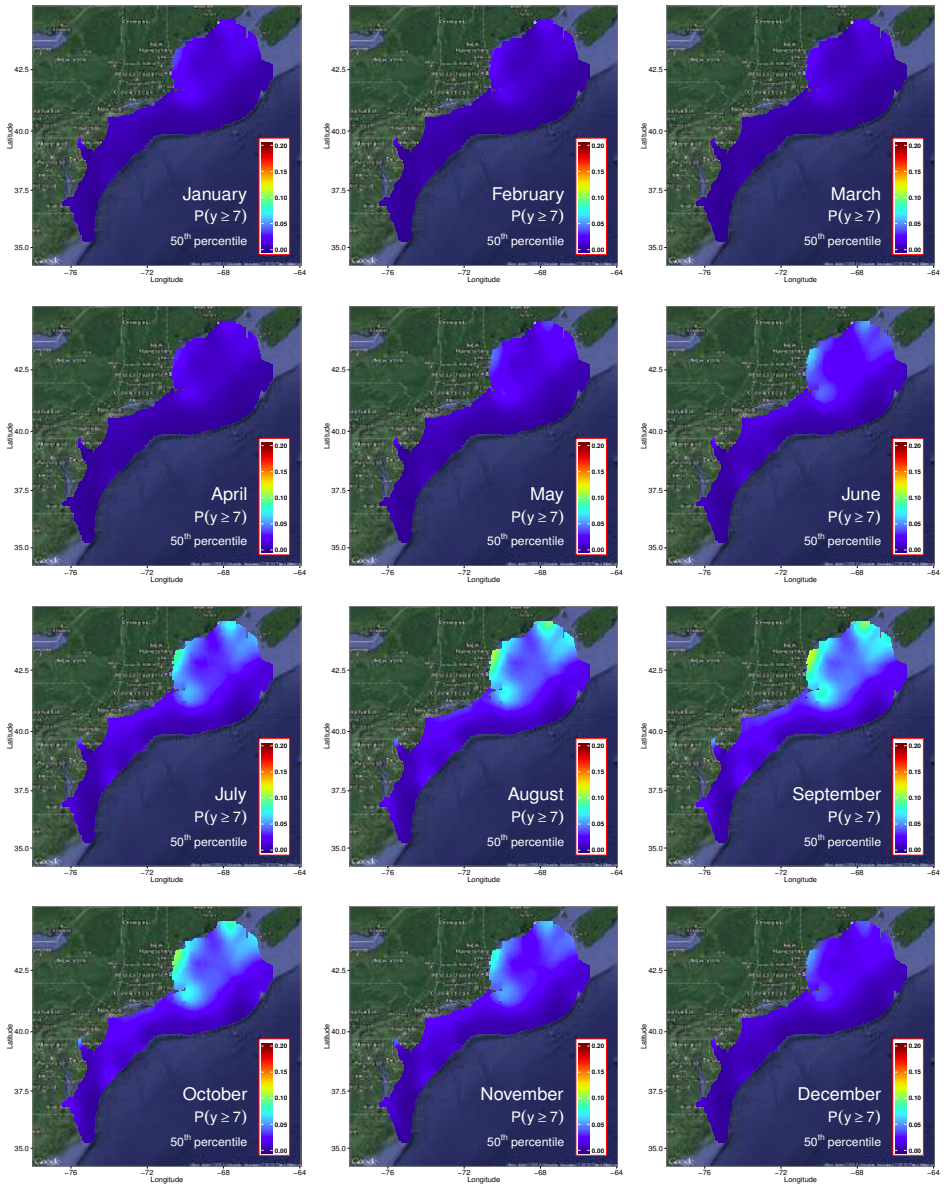

Figure 63: Great Black-backed Gull: Probability of observing a large count during each month.

A.14 Great Shearwater

A.14.1 One-year summary

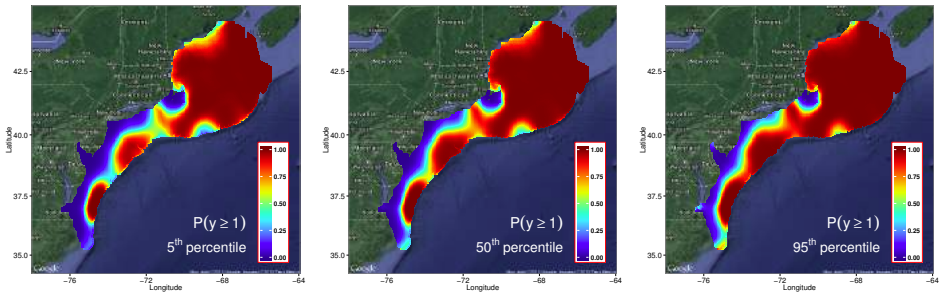

Figure 64: Great Shearwater: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

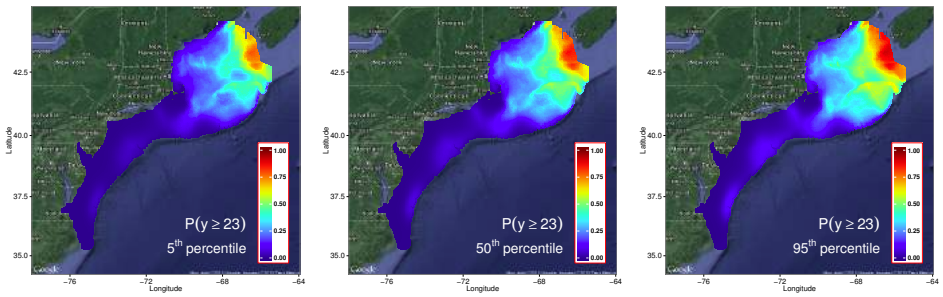

Figure 65: Great Shearwater: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.14.2 Monthly observations

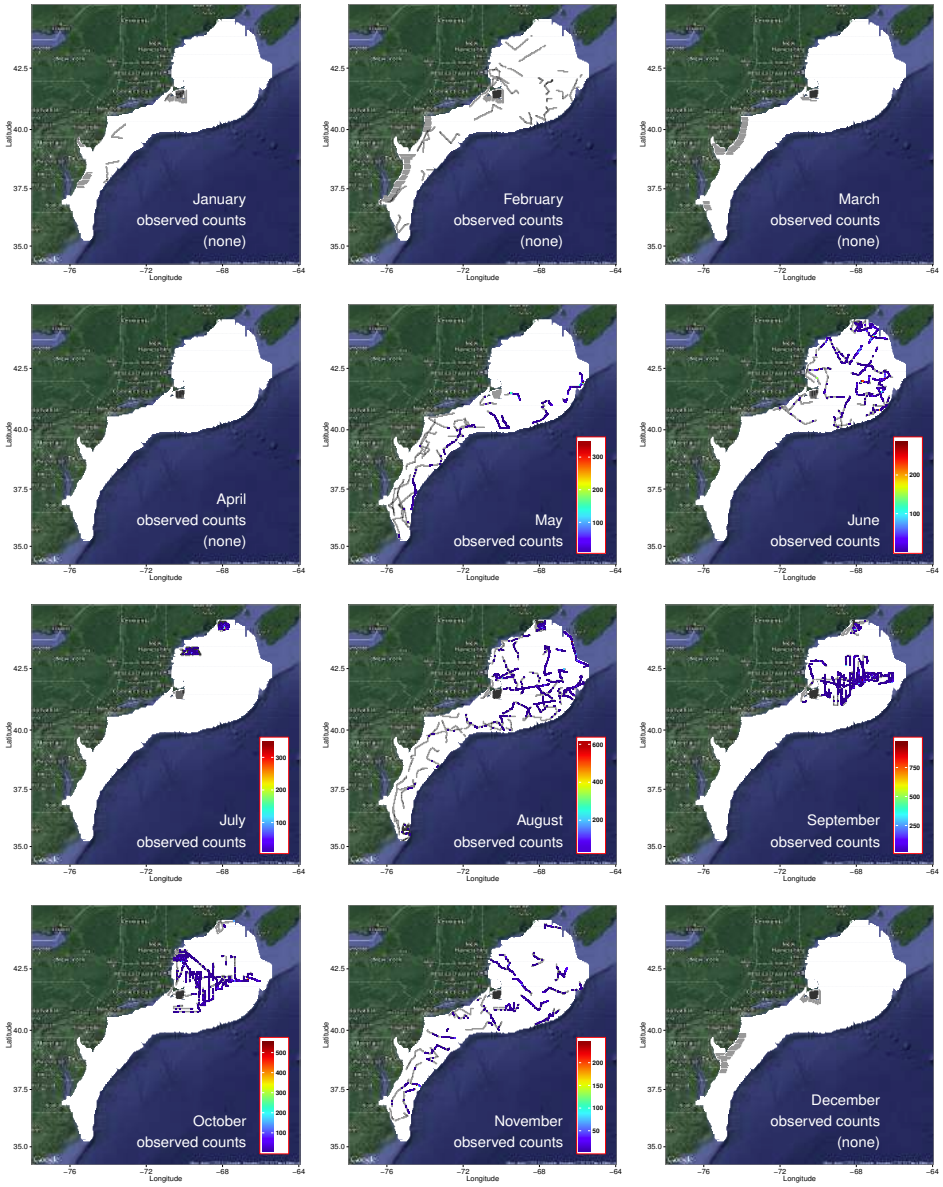

Figure 66: Great Shearwater: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.14.3 Monthly exposure maps

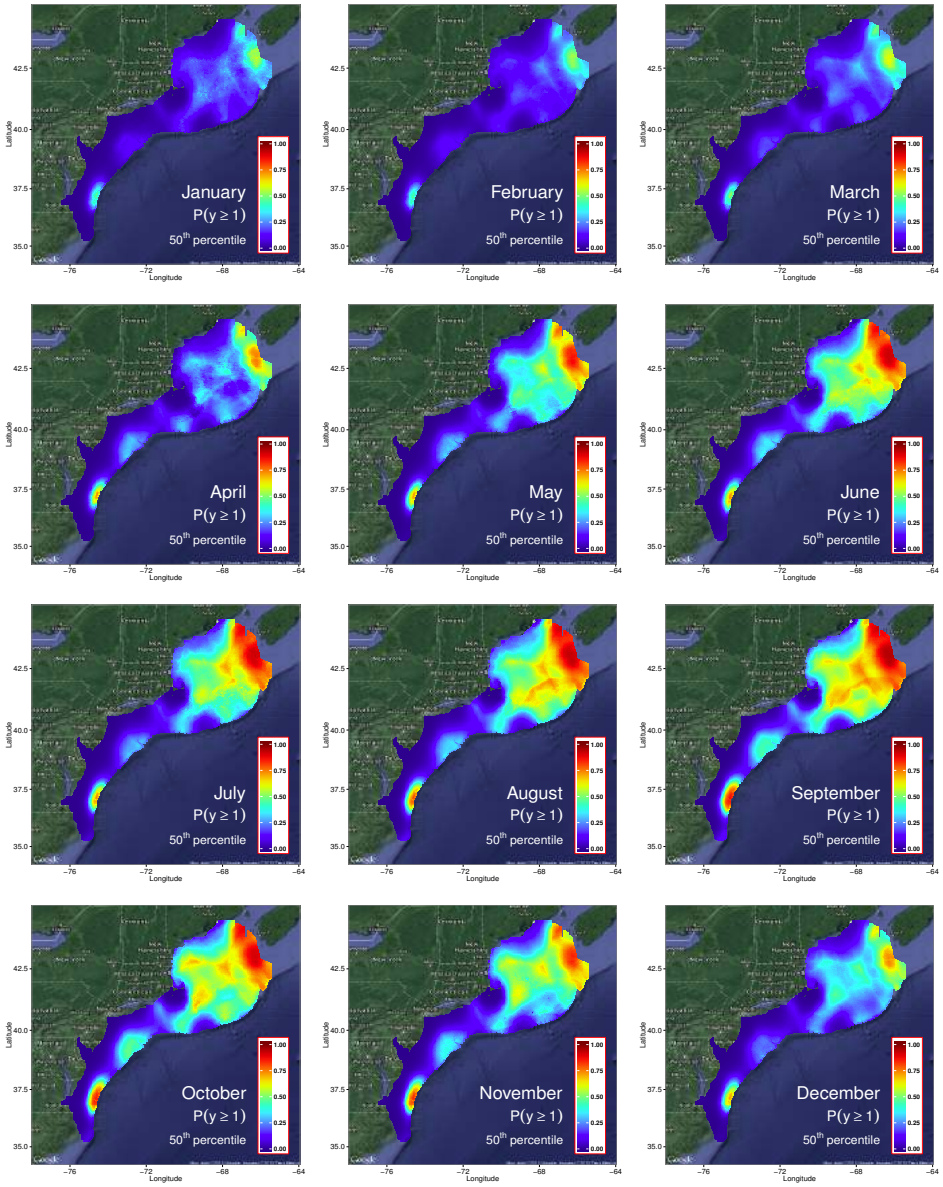

Figure 67: Great Shearwater: Probability of observing at least one individual during each month.

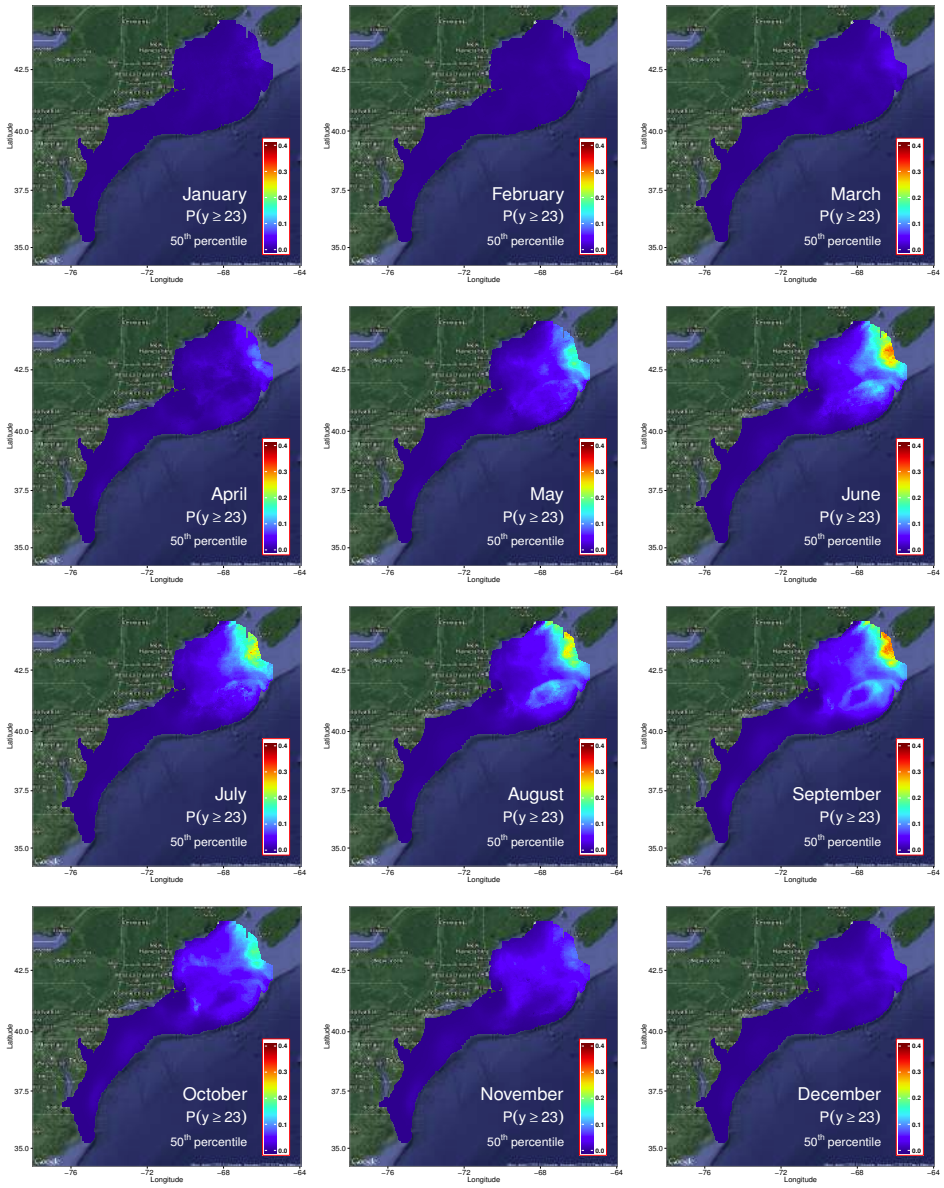

Figure 68: Great Shearwater: Probability of observing a large count during each month.

A.15 Herring Gull

A.15.1 One-year summary

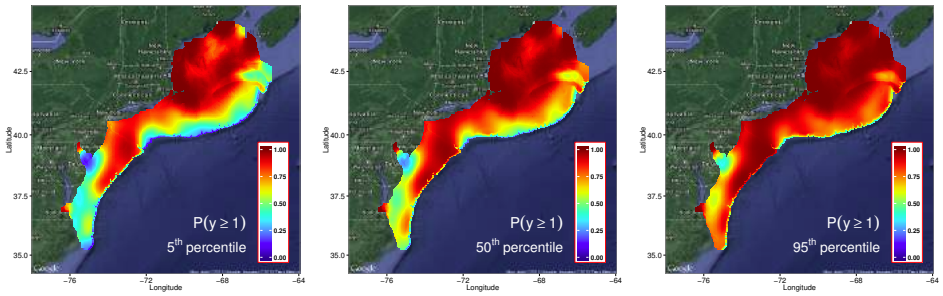

Figure 69: Herring Gull: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

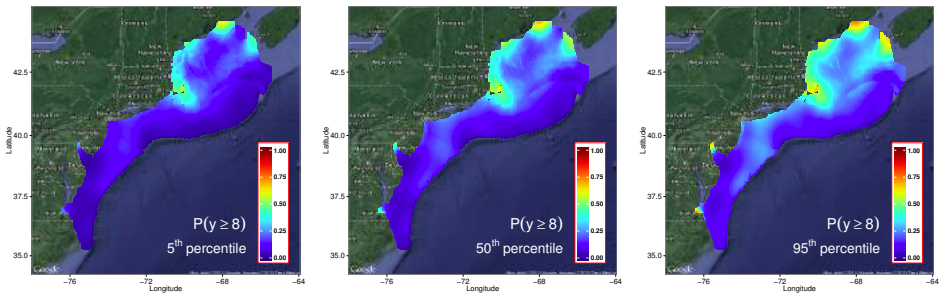

Figure 70: Herring Gull: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.15.2 Monthly observations

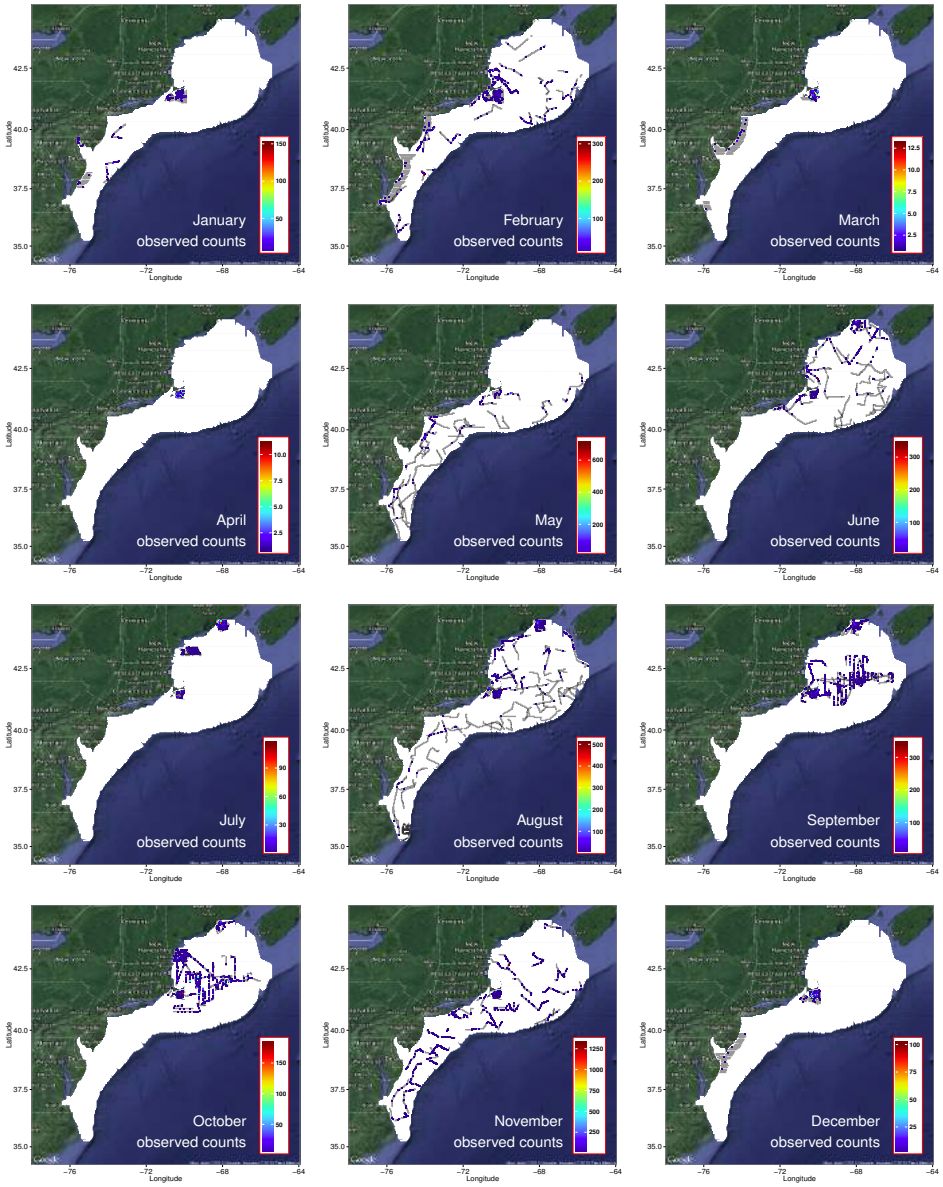

Figure 71: Herring Gull: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.15.3 Monthly exposure maps

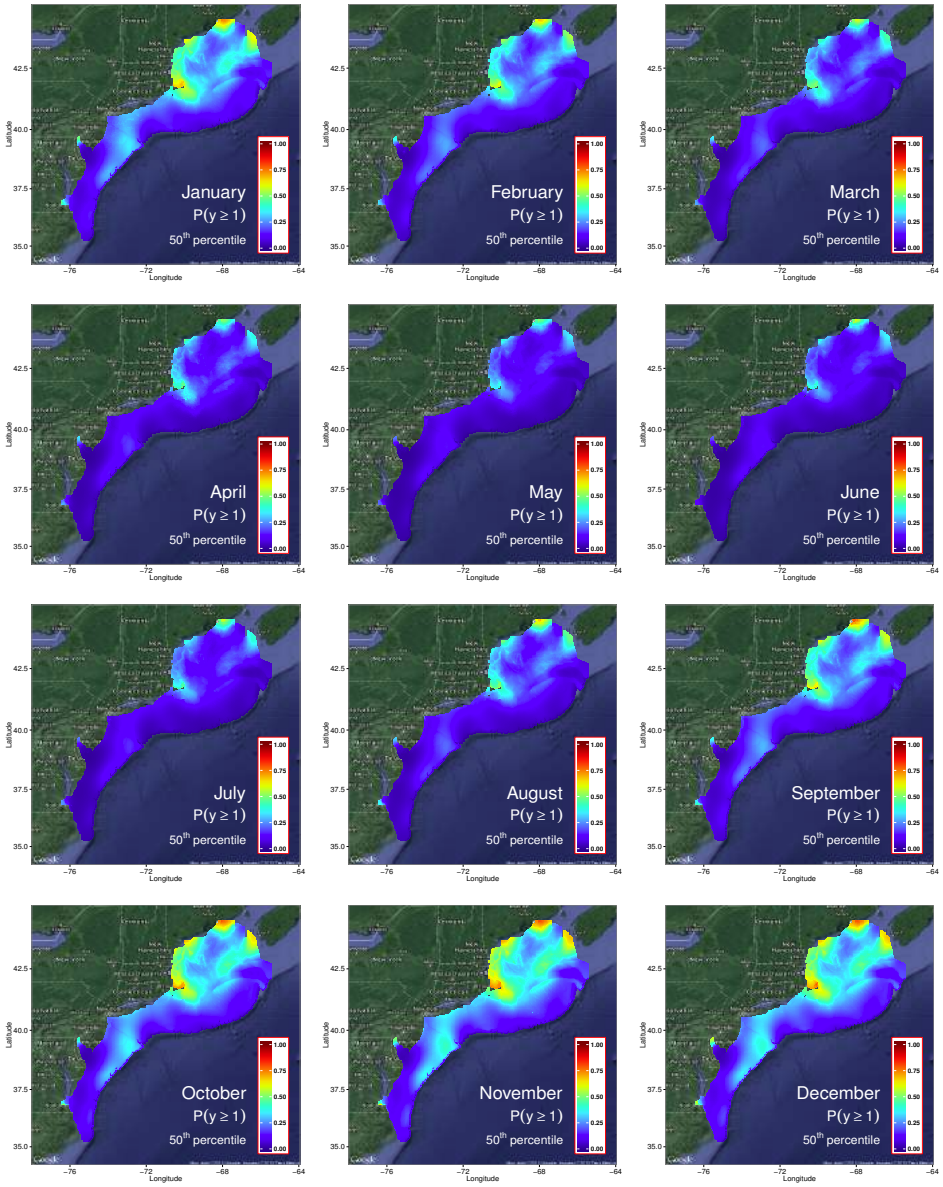

Figure 72: Herring Gull: Probability of observing at least one individual during each month.

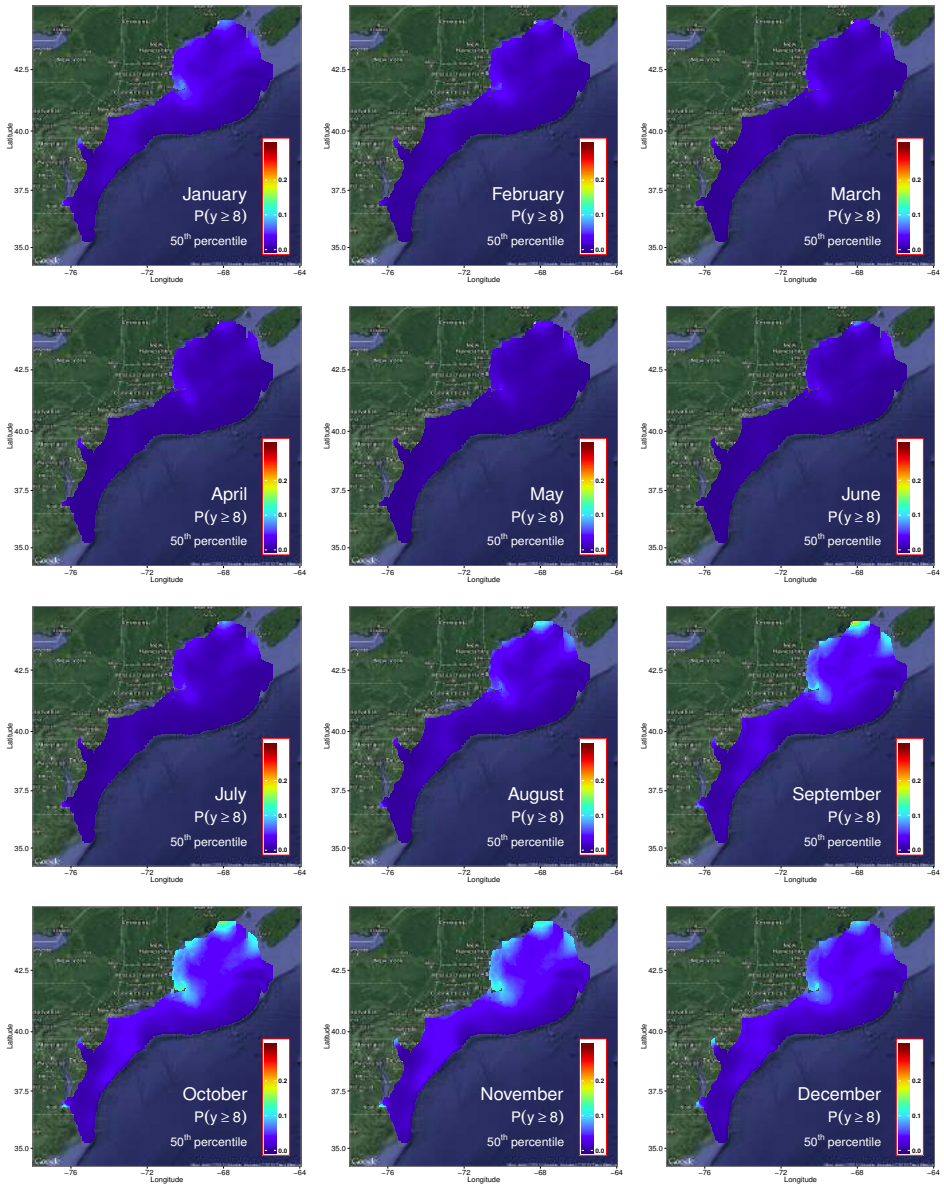

Figure 73: Herring Gull: Probability of observing a large count during each month.

A.16 Laughing Gull

A.16.1 One-year summary

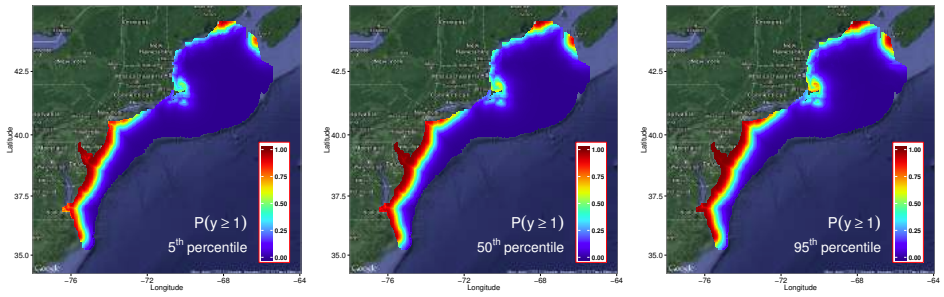

Figure 74: Laughing Gull: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

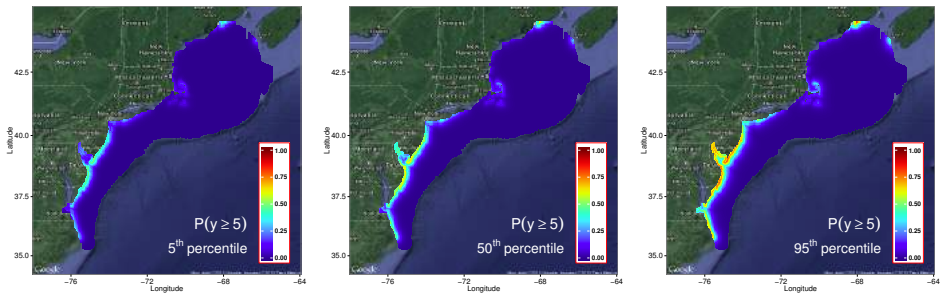

Figure 75: Laughing Gull: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.16.2 Monthly observations

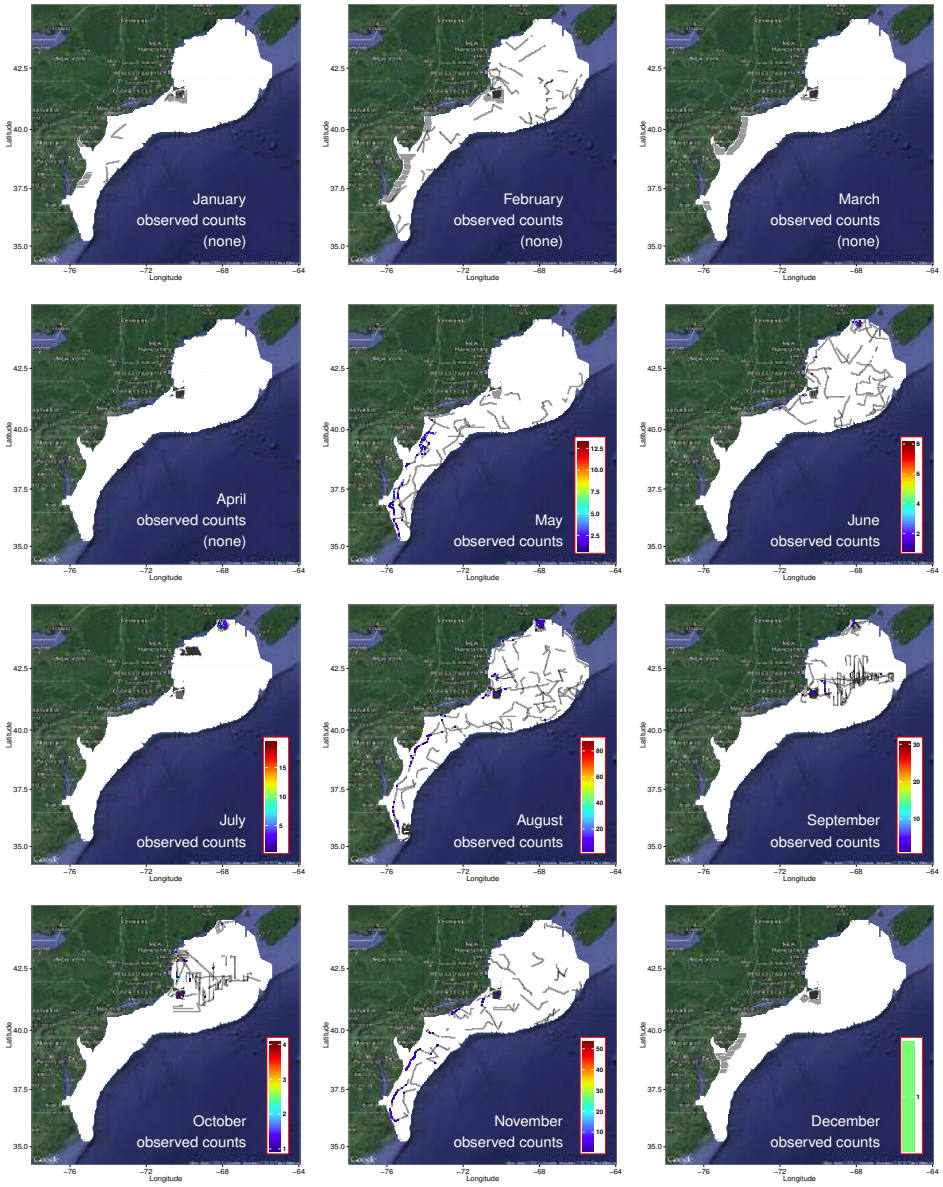

Figure 76: Laughing Gull: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.16.3 Monthly exposure maps

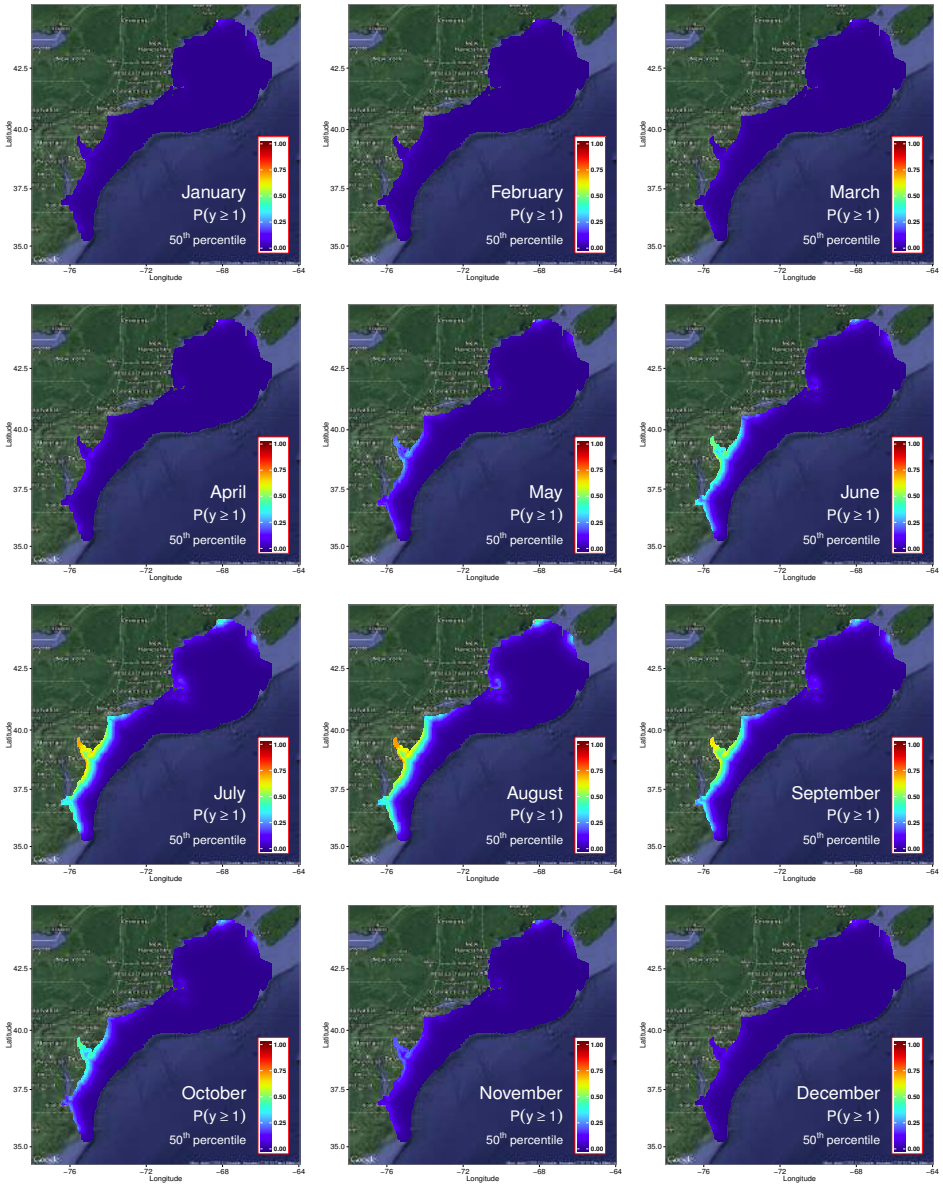

Figure 77: Laughing Gull: Probability of observing at least one individual during each month.

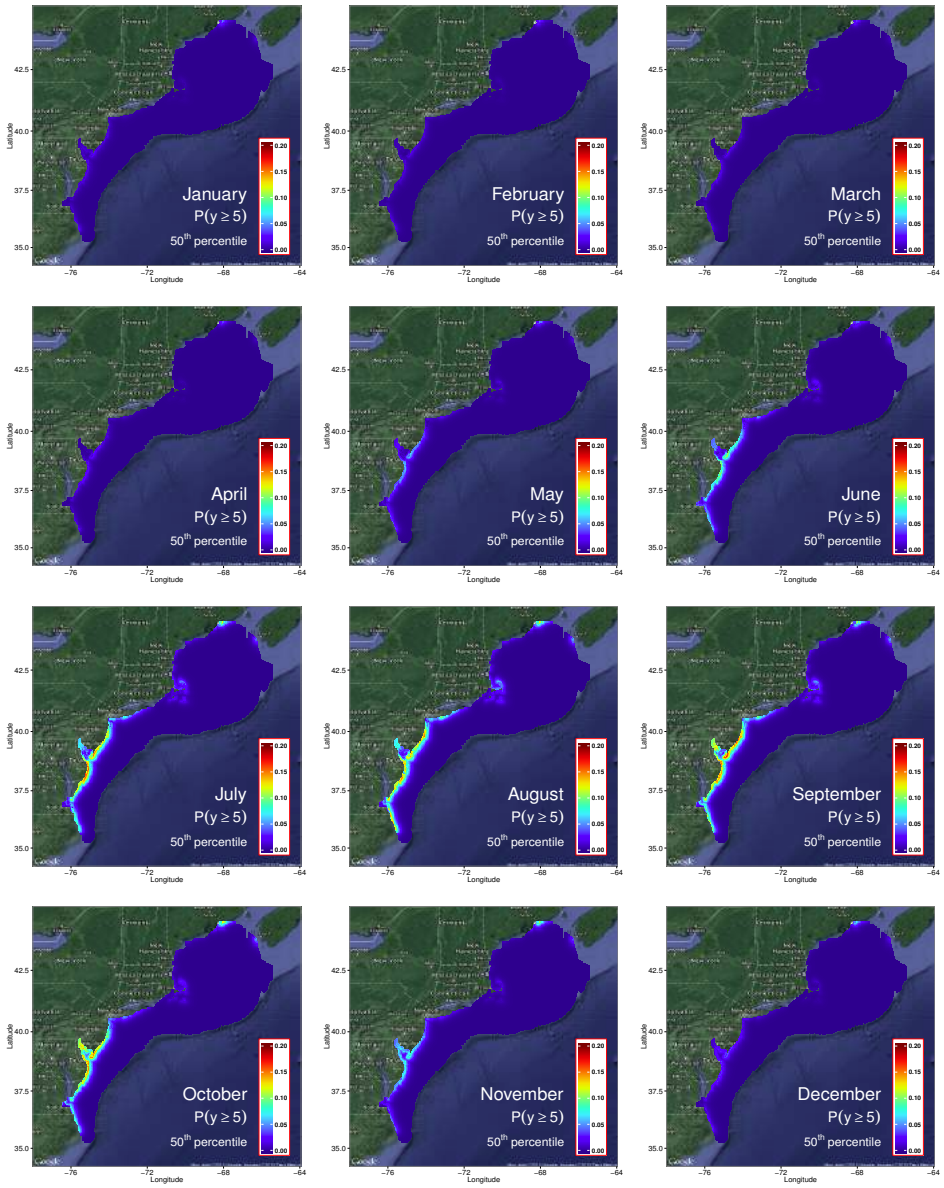

Figure 78: Laughing Gull: Probability of observing a large count during each month.

A.17 Leach's Storm-petrel

A.17.1 One-year summary

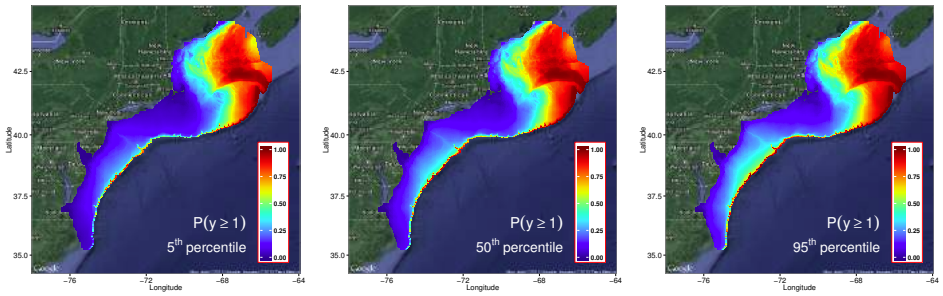

Figure 79: Leach's Storm-petrel: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

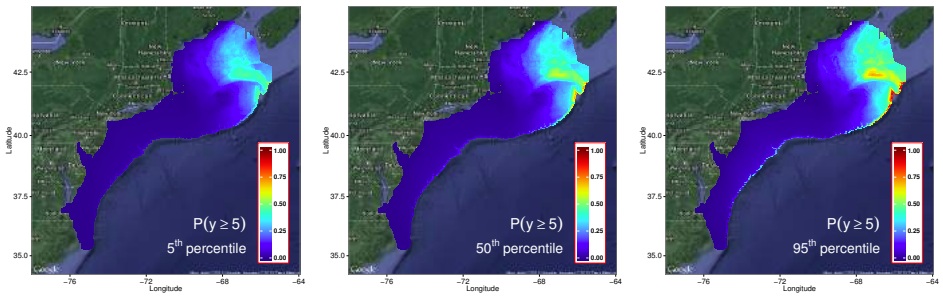

Figure 80: Leach's Storm-petrel: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.17.2 Monthly observations

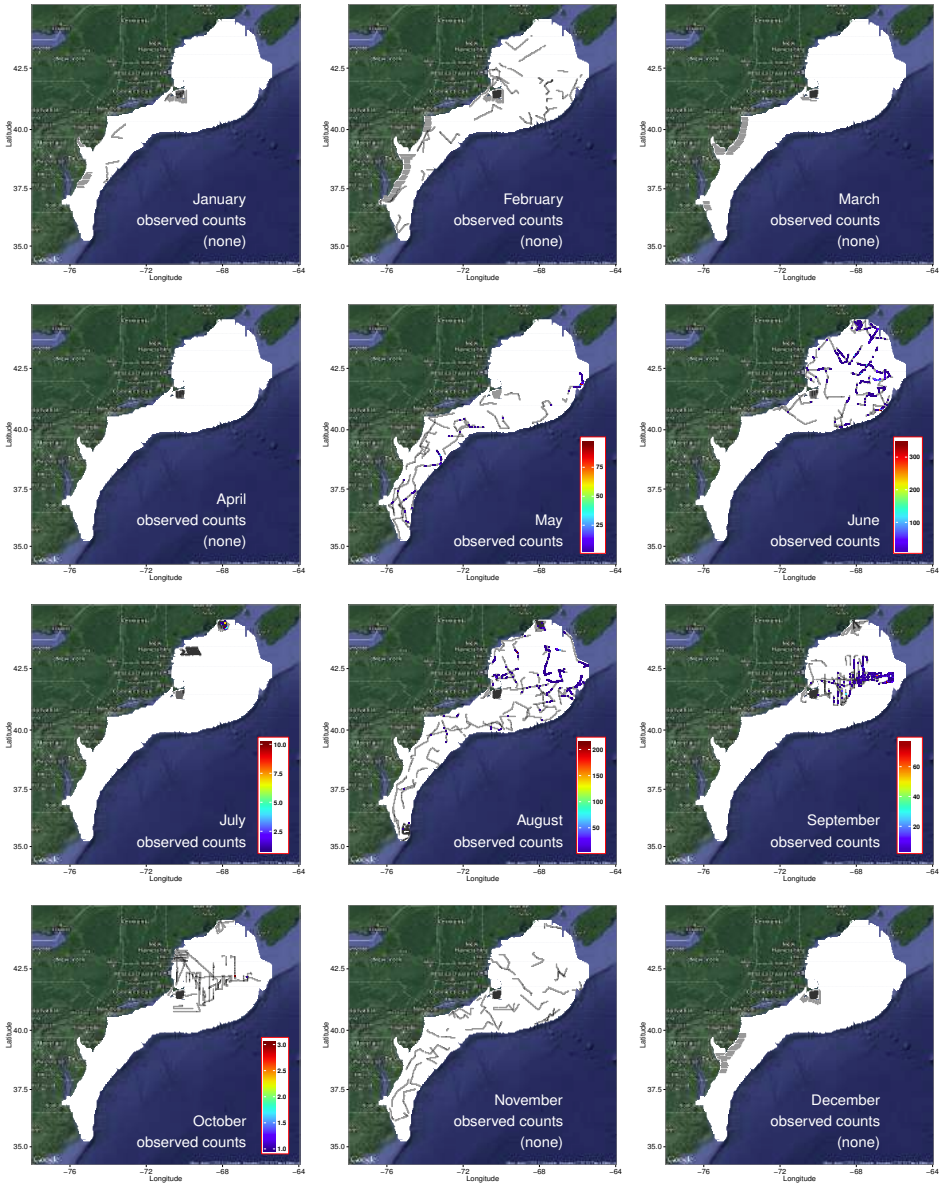

Figure 81: Leach's Storm-petrel: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.17.3 Monthly exposure maps

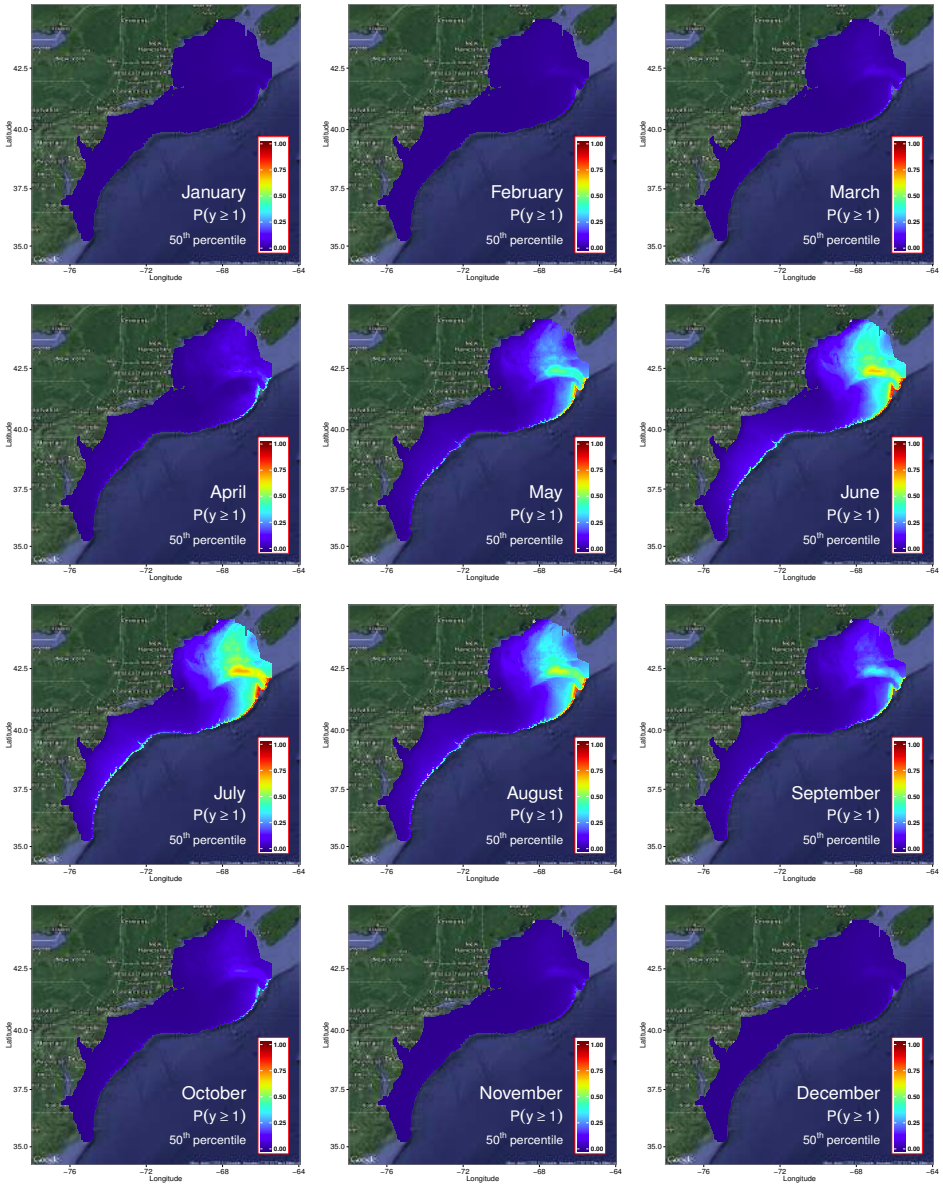

Figure 82: Leach's Storm-petrel: Probability of observing at least one individual during each month.

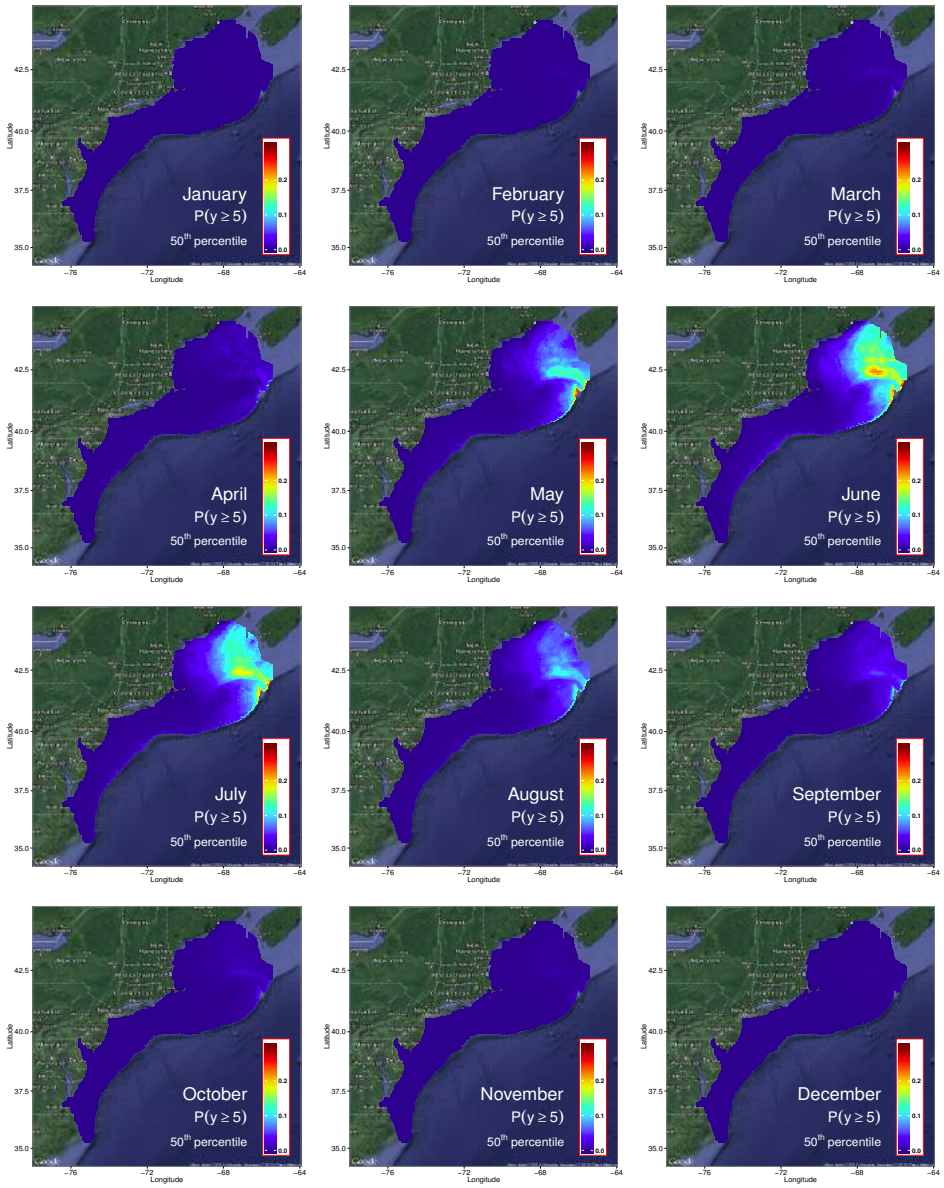

Figure 83: Leach's Storm-petrel: Probability of observing a large count during each month.

A.18 Long-tailed Duck

A.18.1 One-year summary

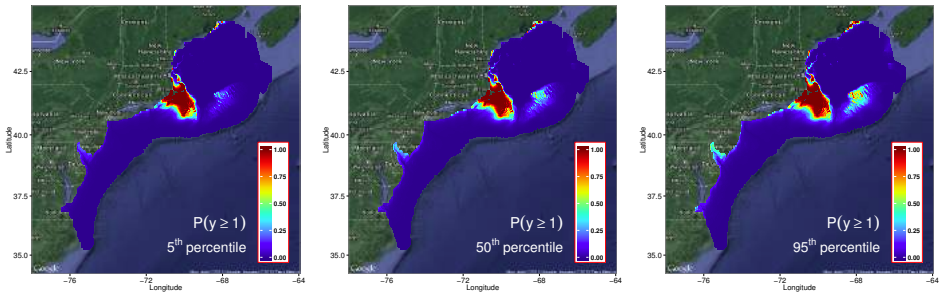

Figure 84: Long-tailed Duck: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.18.2 Monthly observations

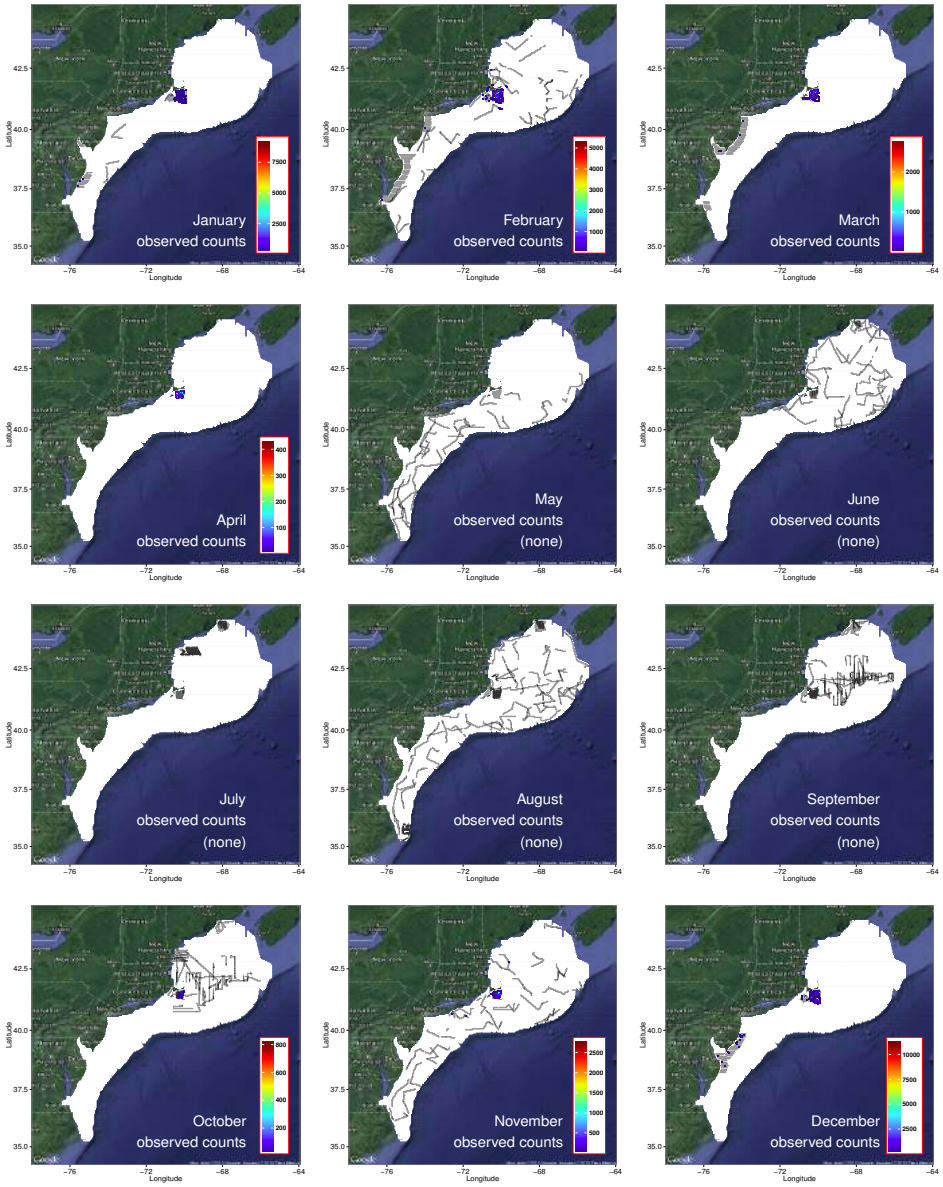

Figure 85: Long-tailed Duck: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.18.3 Monthly exposure maps

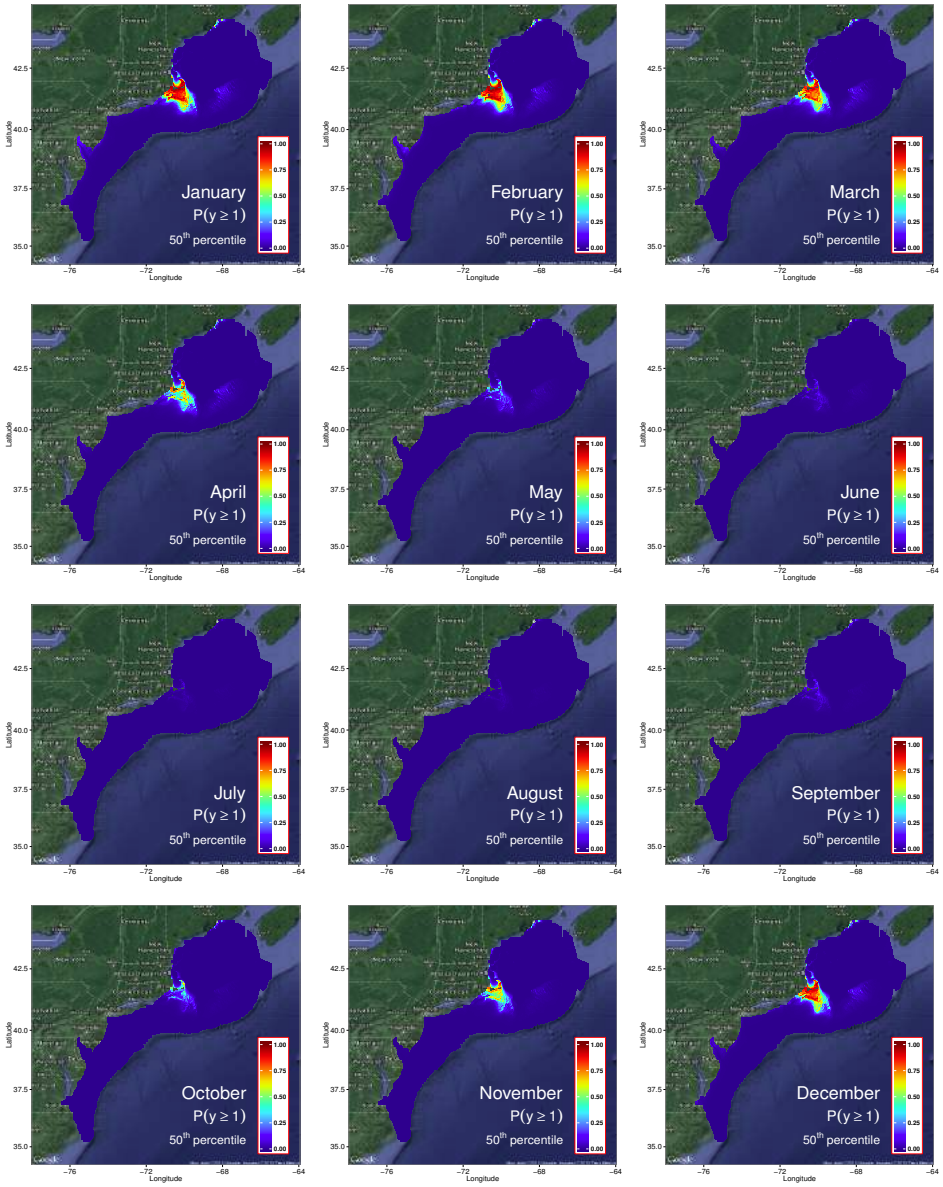

Figure 86: Long-tailed Duck: Probability of observing at least one individual during each month.

A.19 Northern Fulmar

A.19.1 One-year summary

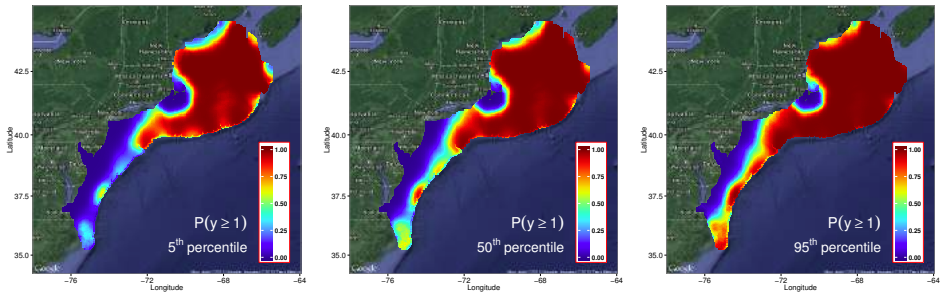

Figure 87: Northern Fulmar: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

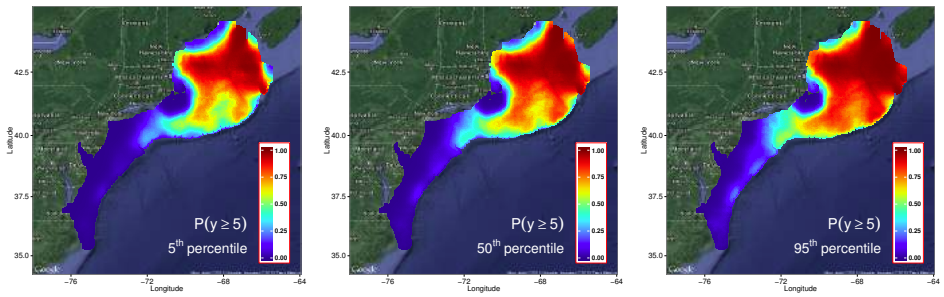

Figure 88: Northern Fulmar: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.19.2 Monthly observations

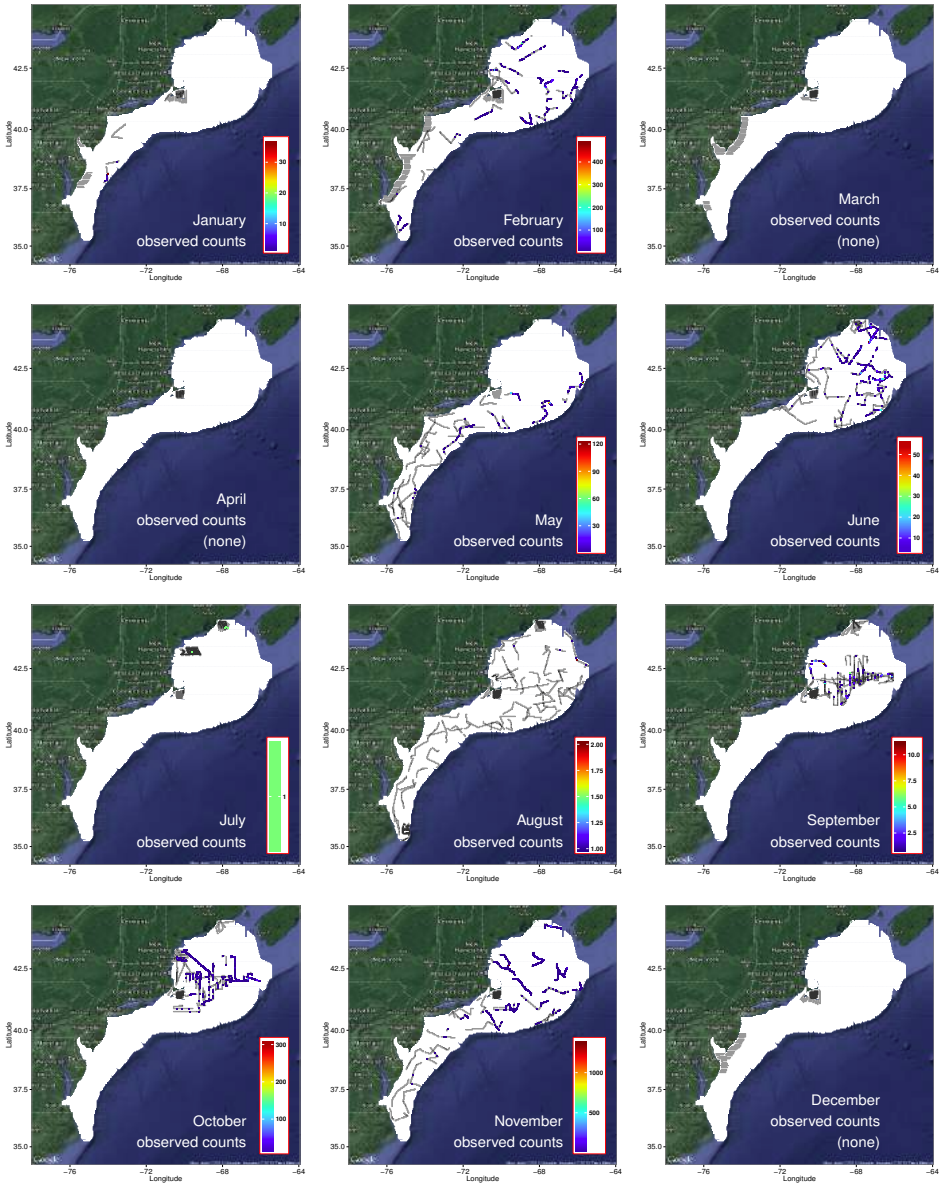

Figure 89: Northern Fulmar: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.19.3 Monthly exposure maps

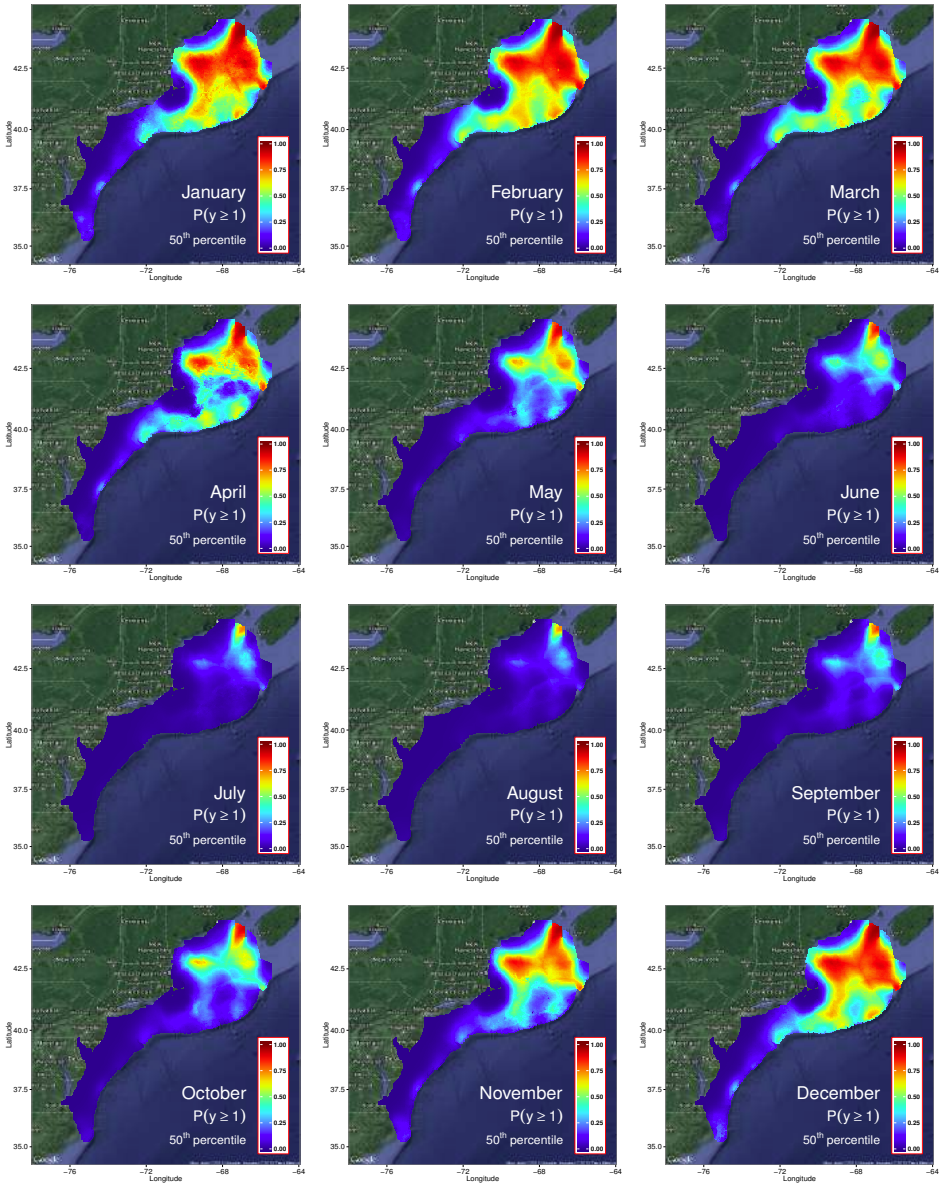

Figure 90: Northern Fulmar: Probability of observing at least one individual during each month.

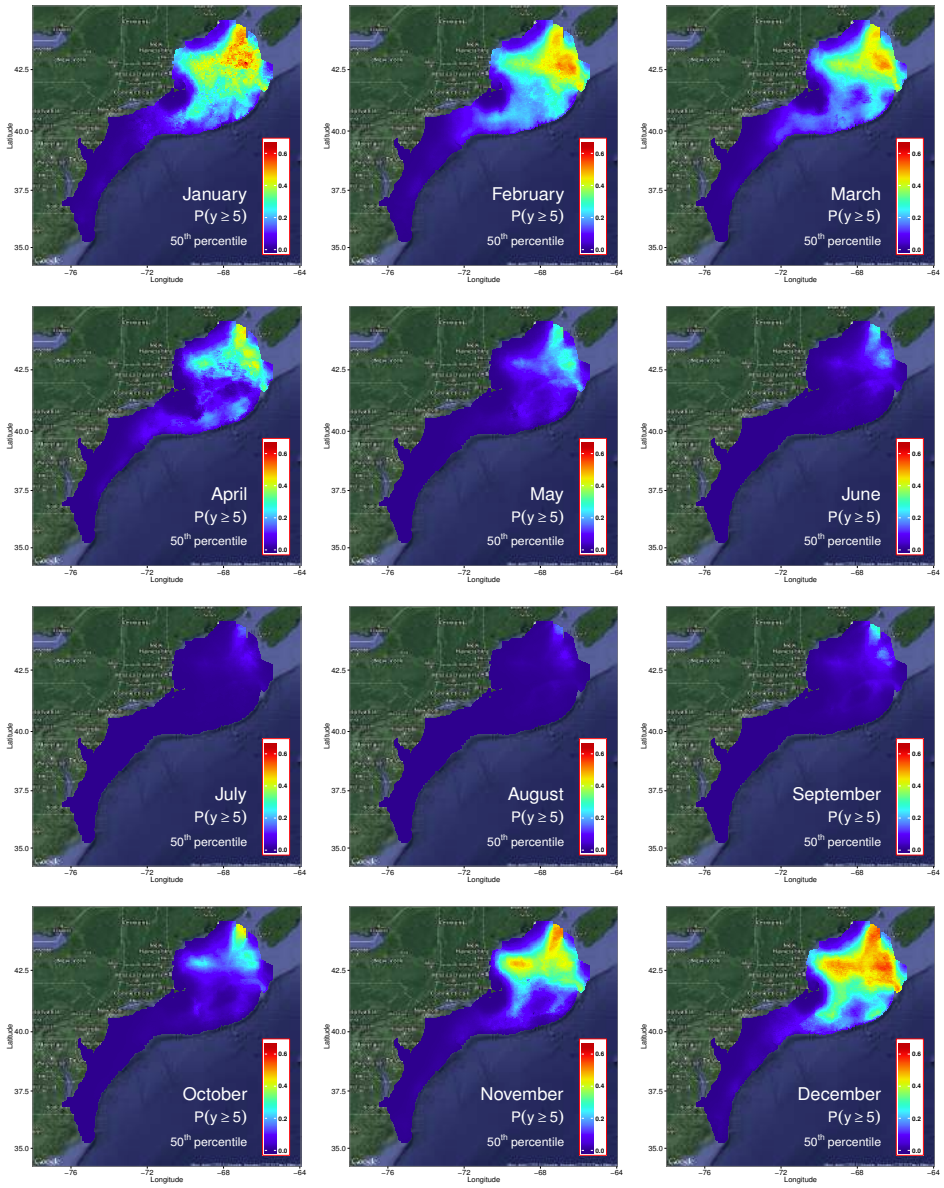

Figure 91: Northern Fulmar: Probability of observing a large count during each month.

A.20 Northern Gannet

A.20.1 One-year summary

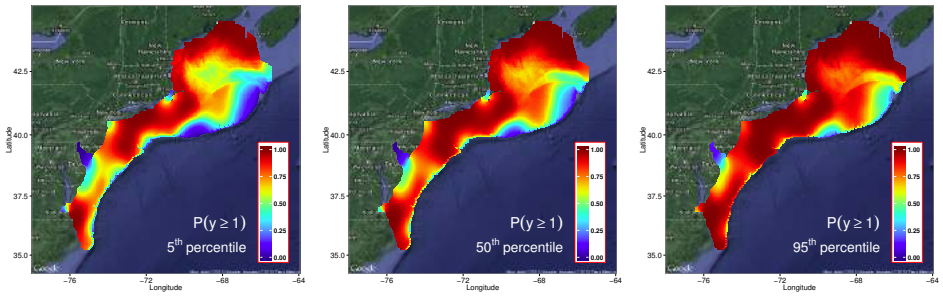

Figure 92: Northern Gannet: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

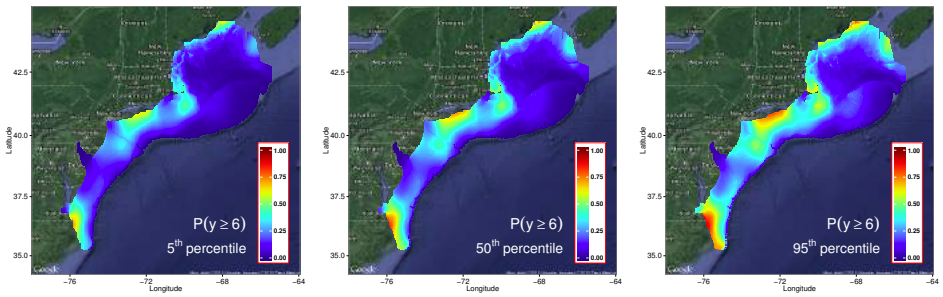

Figure 93: Northern Gannet: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.20.2 Monthly observations

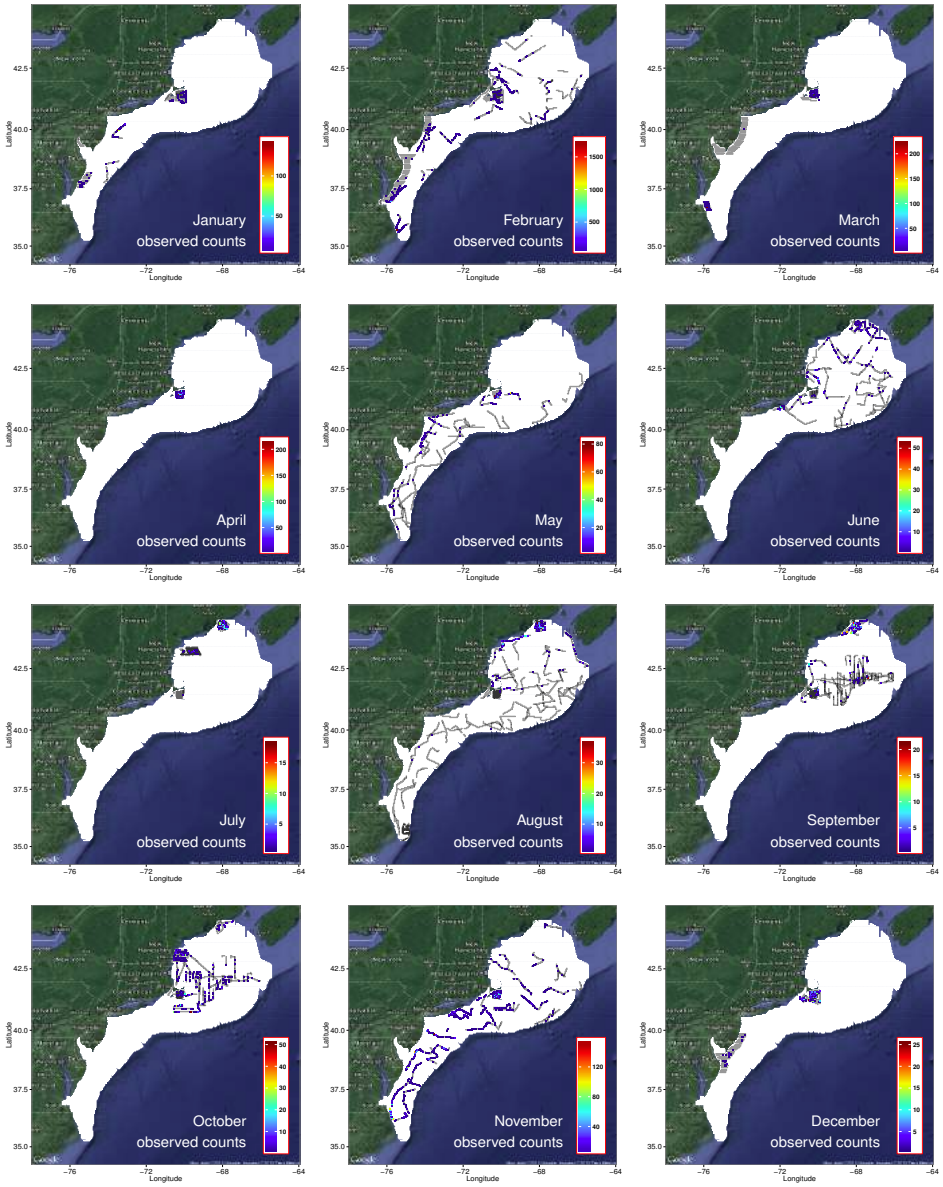

Figure 94: Northern Gannet: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.20.3 Monthly exposure maps

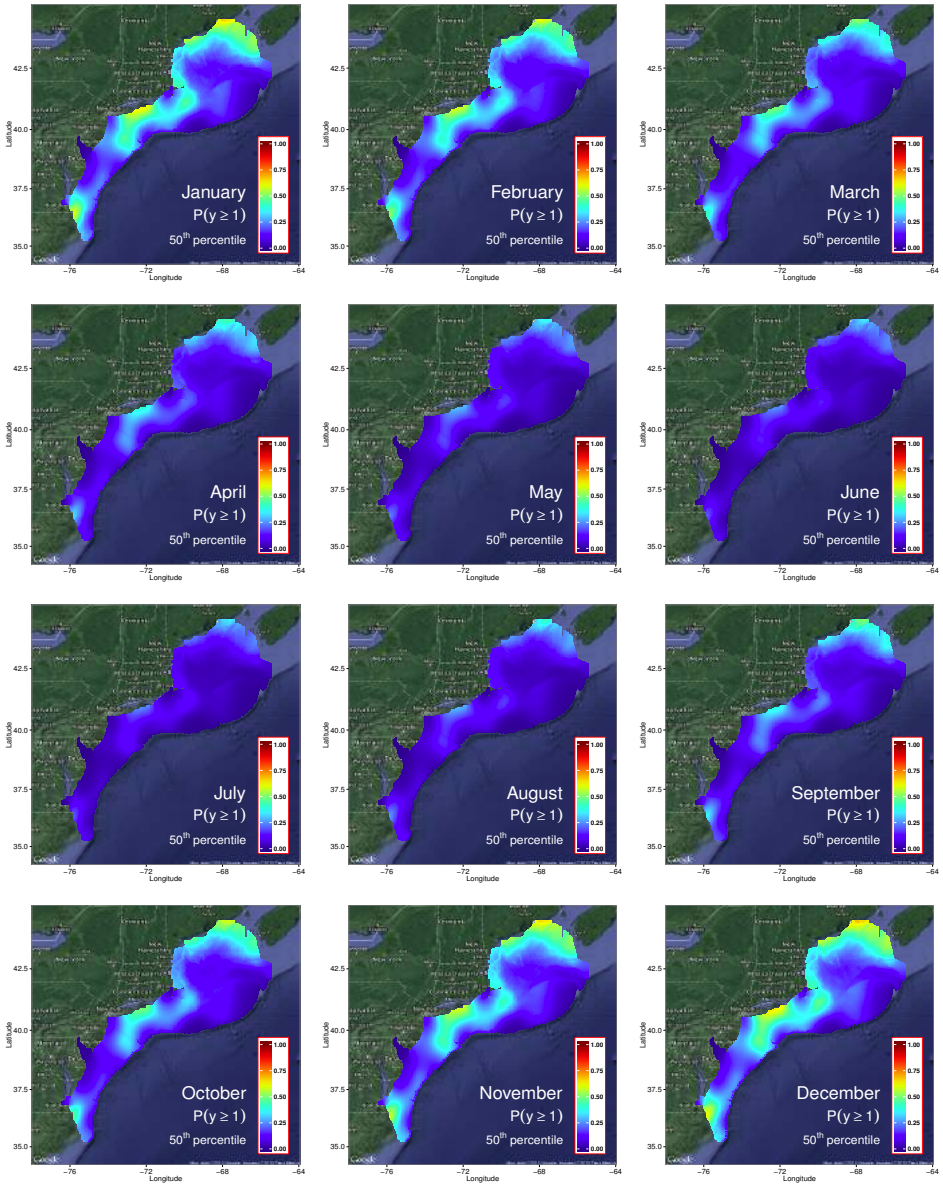

Figure 95: Northern Gannet: Probability of observing at least one individual during each month.

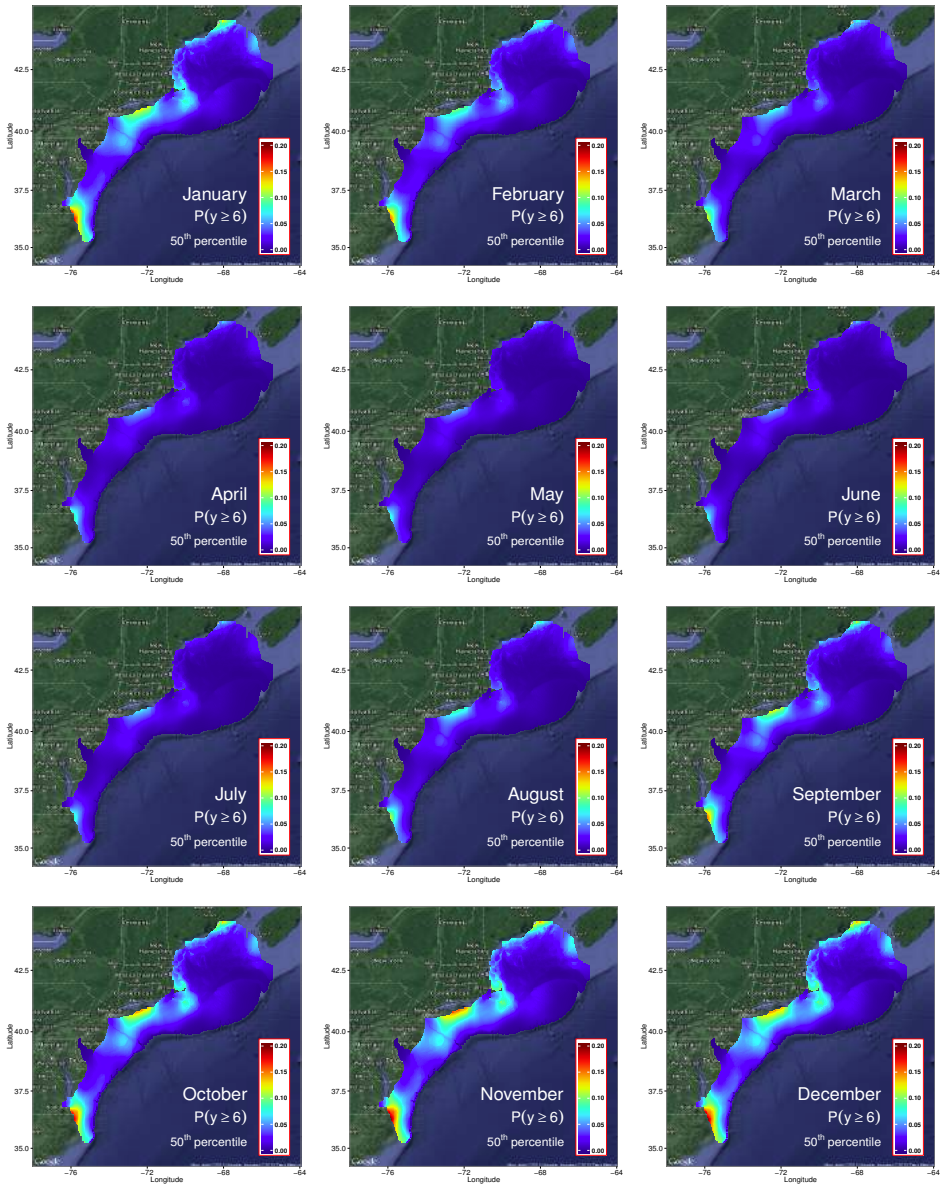

Figure 96: Northern Gannet: Probability of observing a large count during each month.

A.21 Razorbill

A.21.1 One-year summary

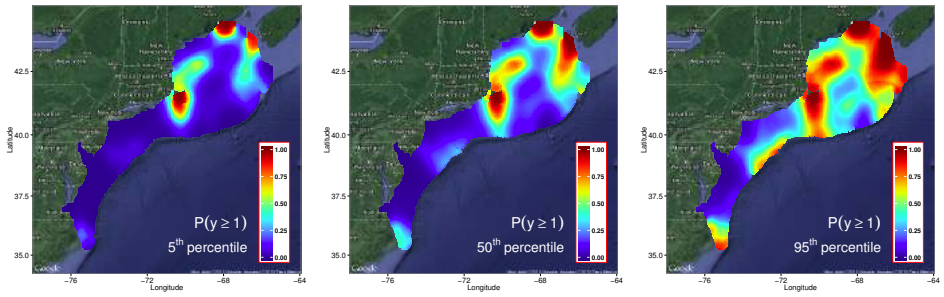

Figure 97: Razorbill: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

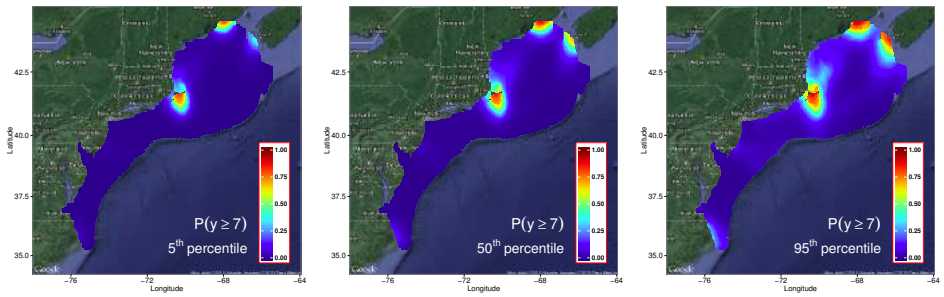

Figure 98: Razorbill: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.21.2 Monthly observations

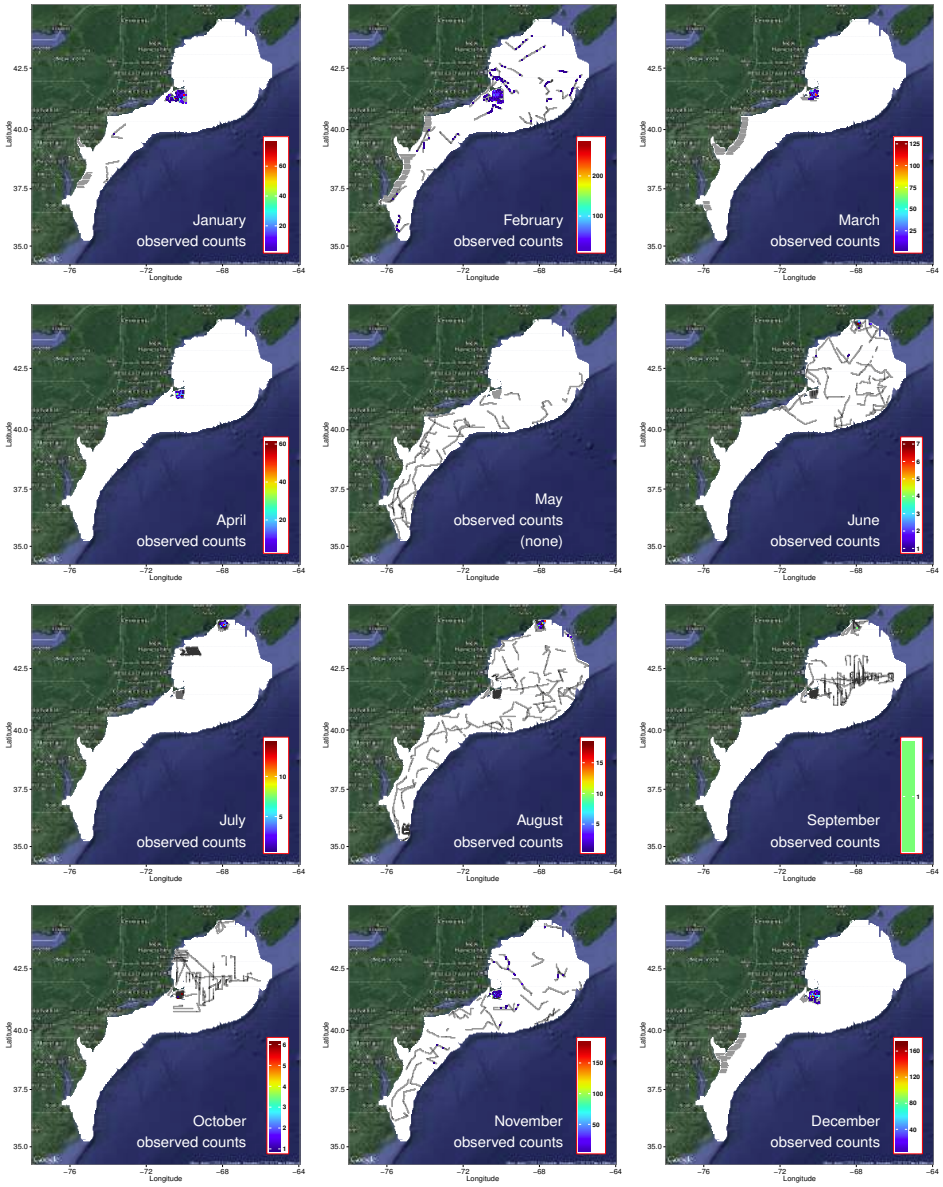

Figure 99: Razorbill: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.21.3 Monthly exposure maps

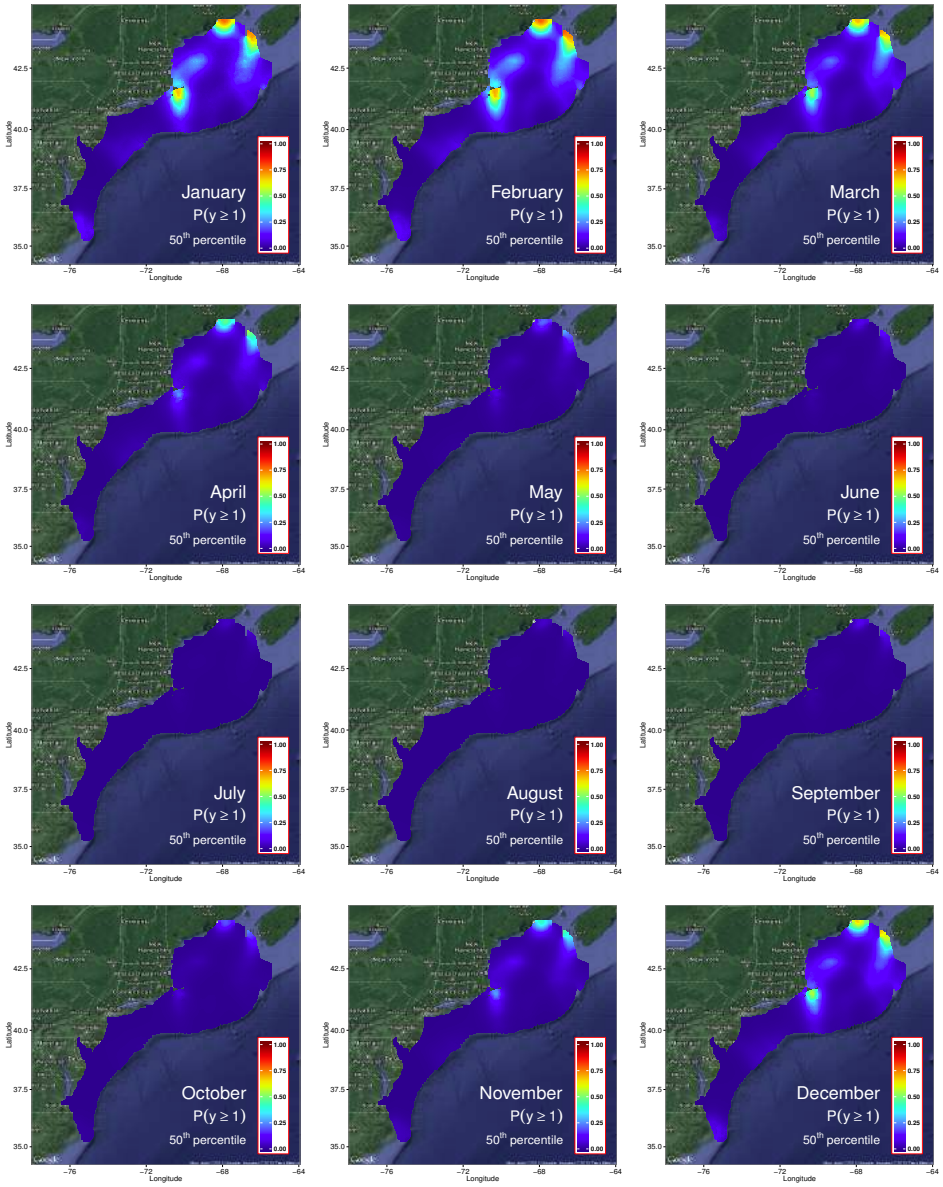

Figure 100: Razorbill: Probability of observing at least one individual during each month.

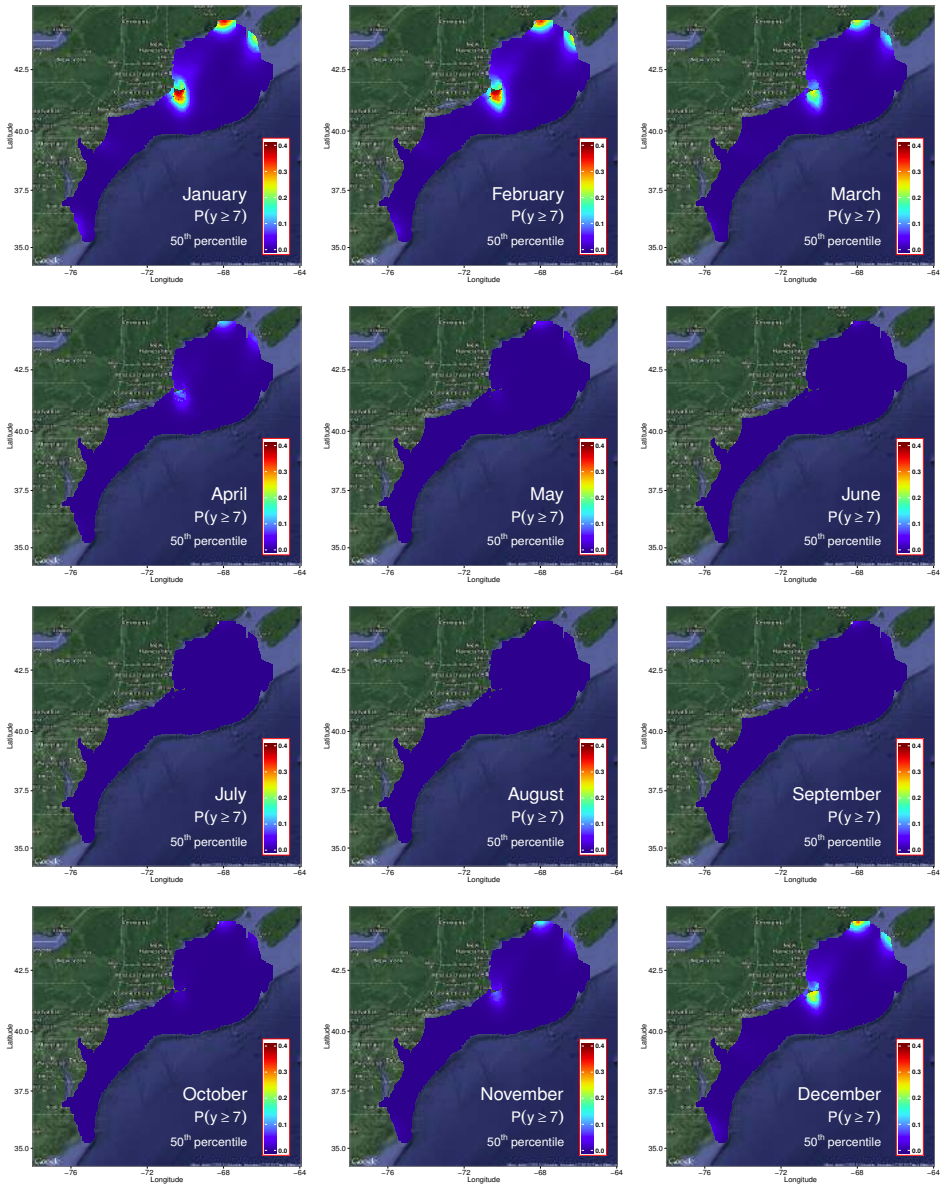

Figure 101: Razorbill: Probability of observing a large count during each month.

A.22 Roseate Tern

A.22.1 One-year summary

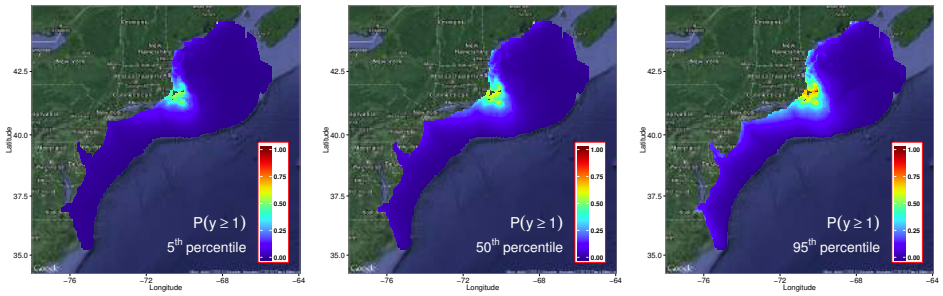

Figure 102: Roseate Tern: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.22.2 Monthly observations

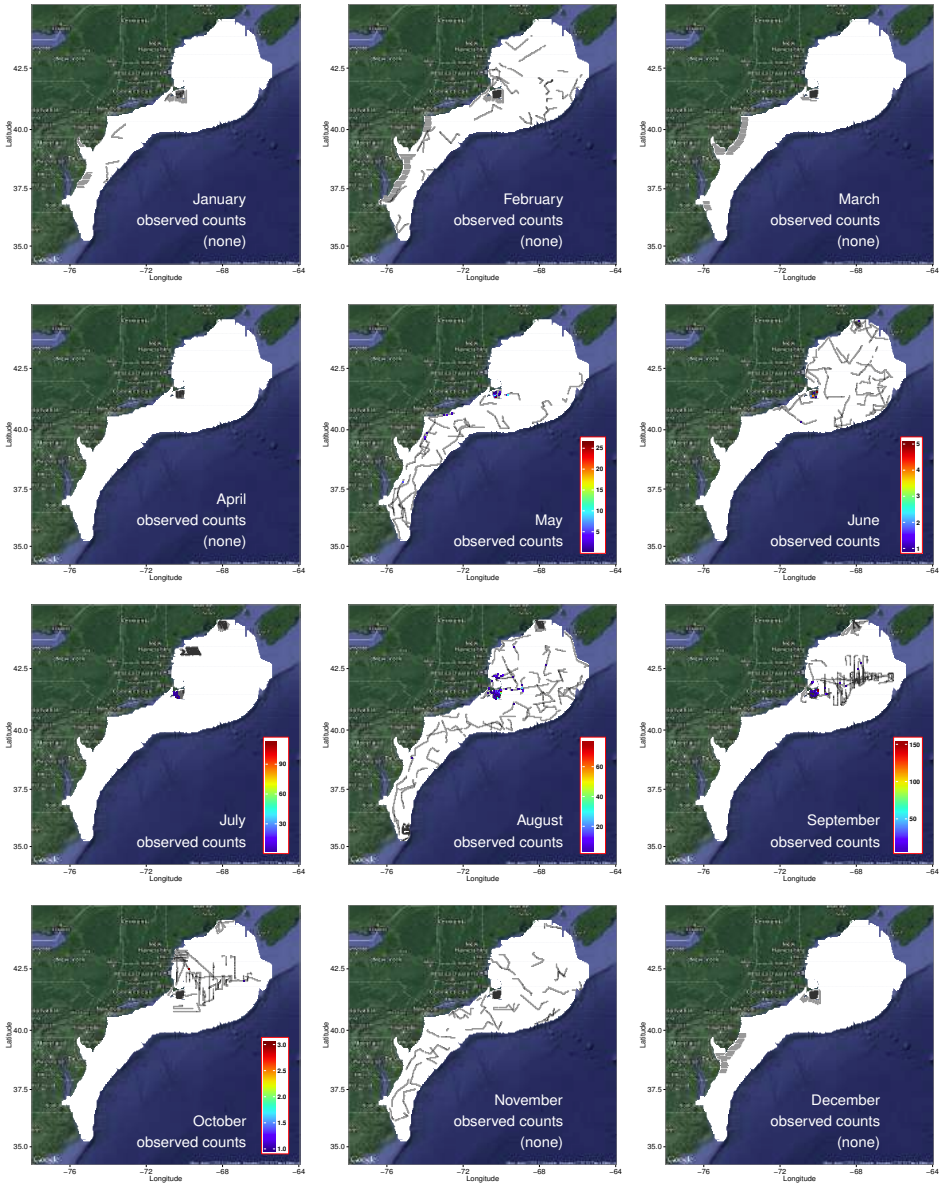

Figure 103: Roseate Tern: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.22.3 Monthly exposure maps

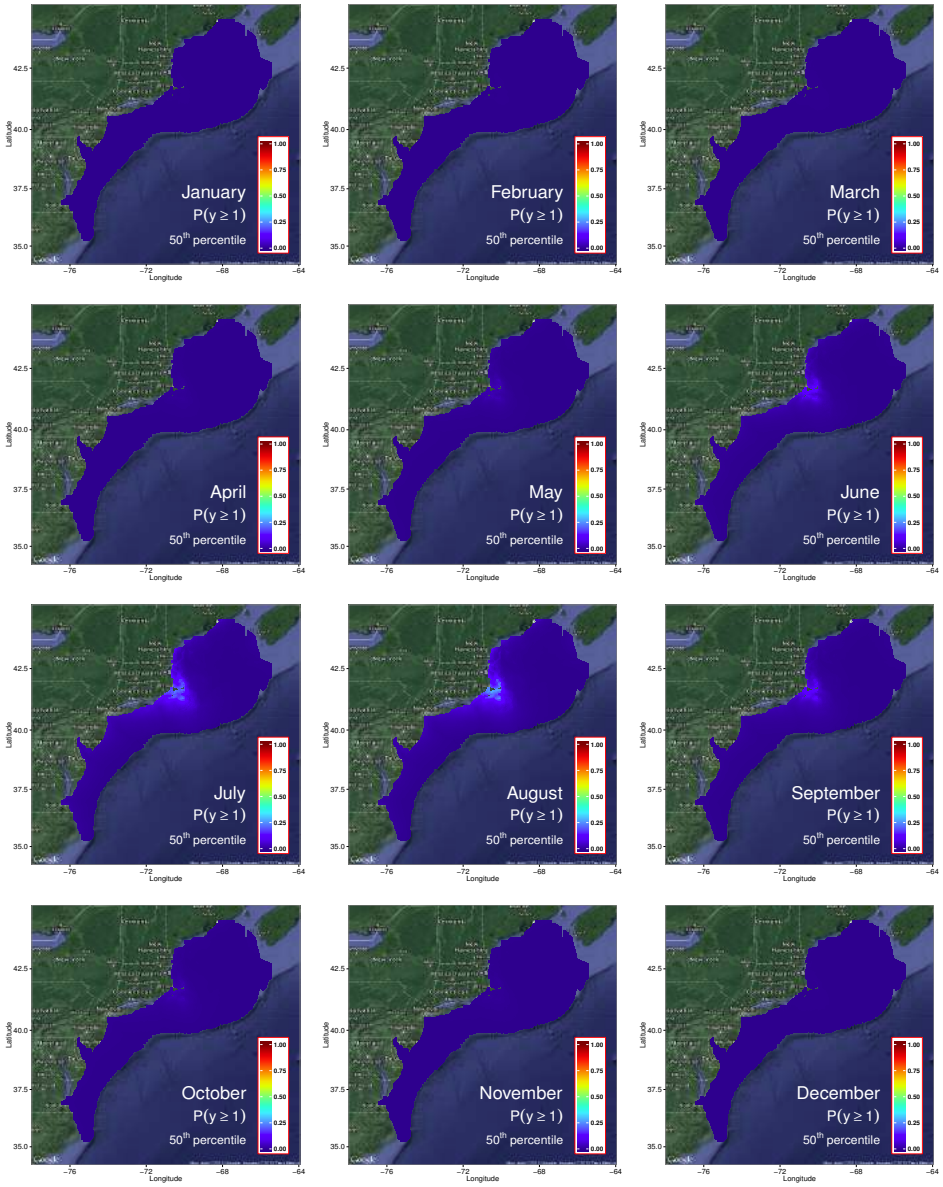

Figure 104: Roseate Tern: Probability of observing at least one individual during each month.

A.23 Sooty Shearwater

A.23.1 One-year summary

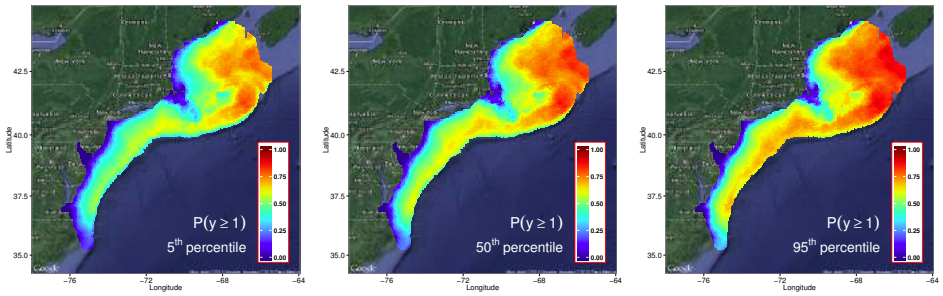

Figure 105: Sooty Shearwater: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

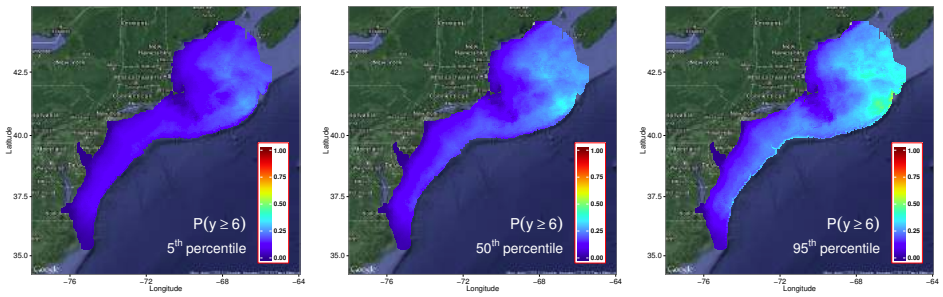

Figure 106: Sooty Shearwater: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.23.2 Monthly observations

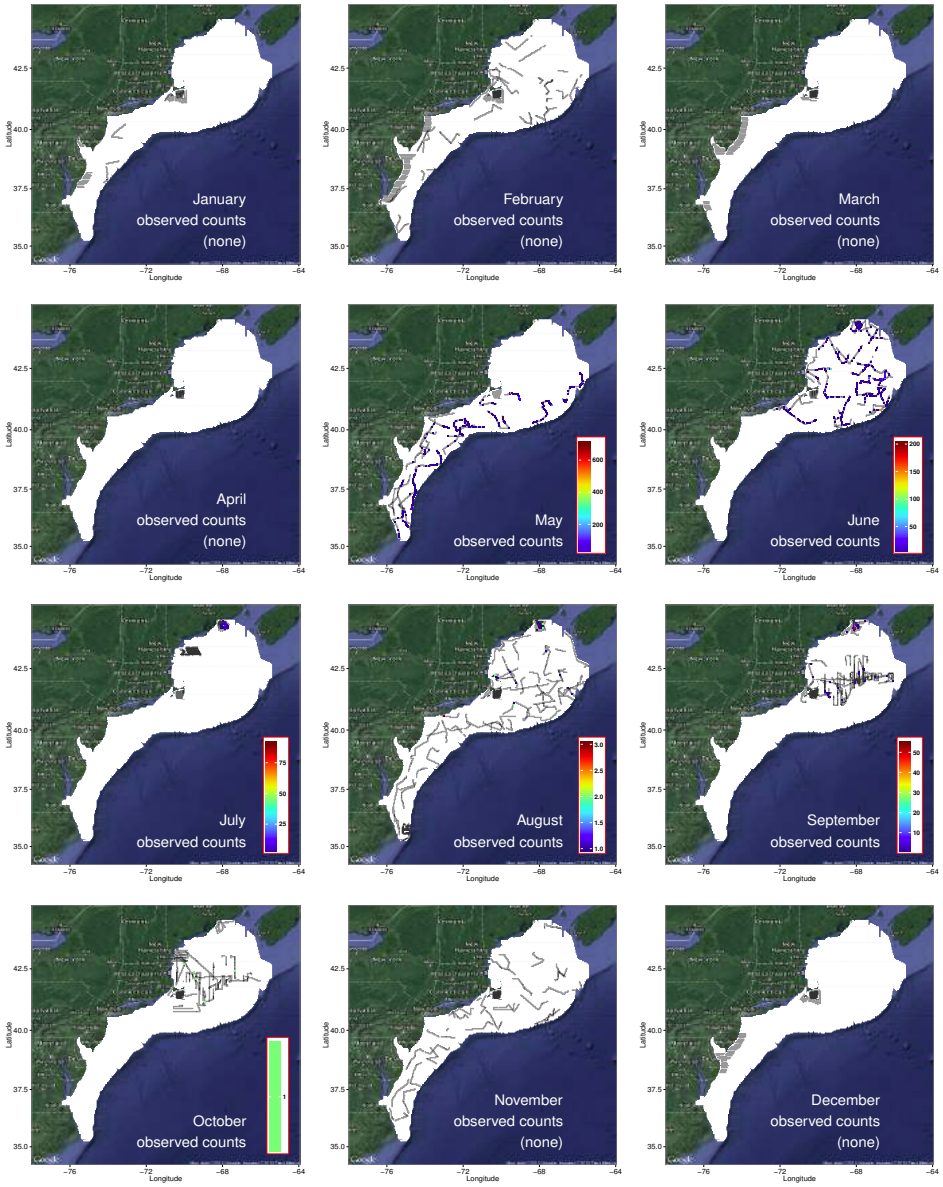

Figure 107: Sooty Shearwater: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.23.3 Monthly exposure maps

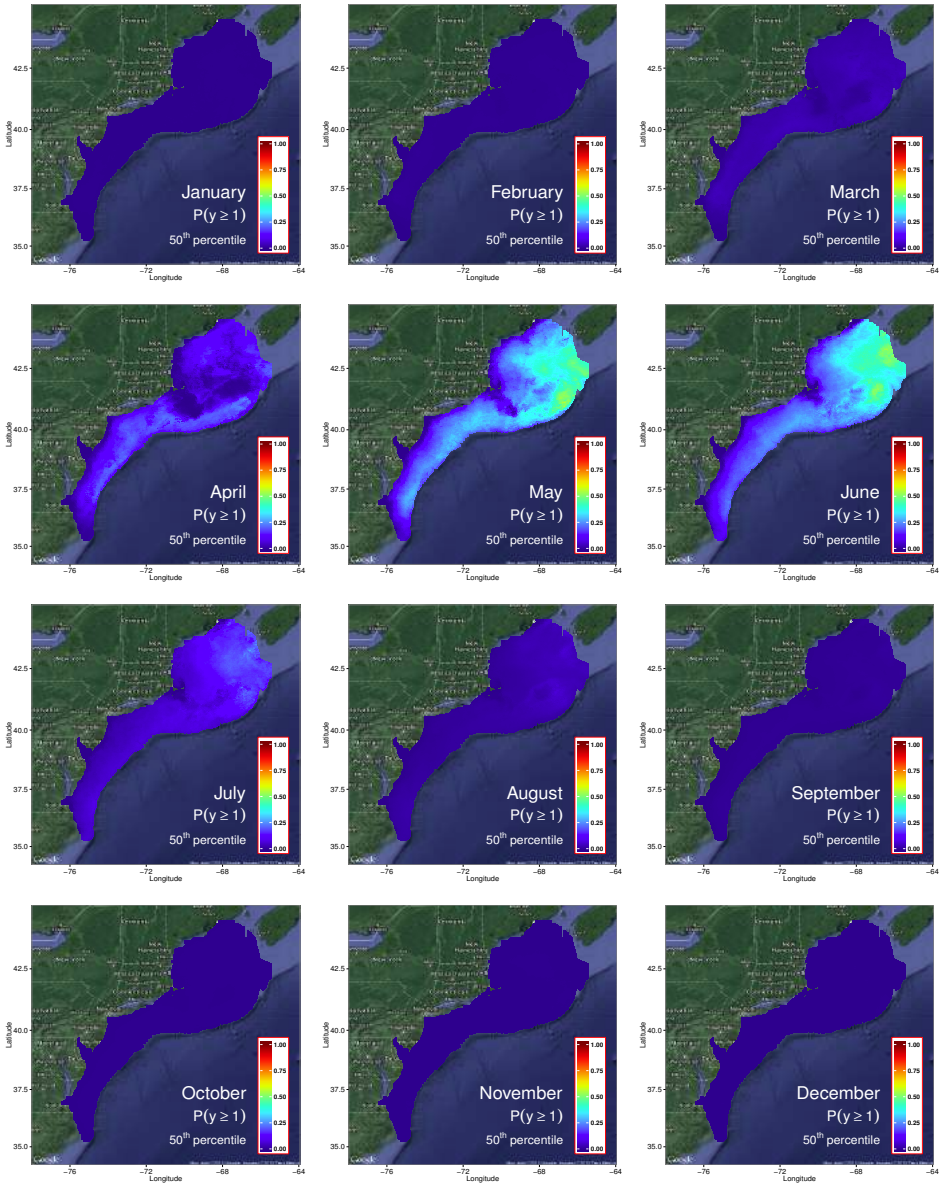

Figure 108: Sooty Shearwater: Probability of observing at least one individual during each month.

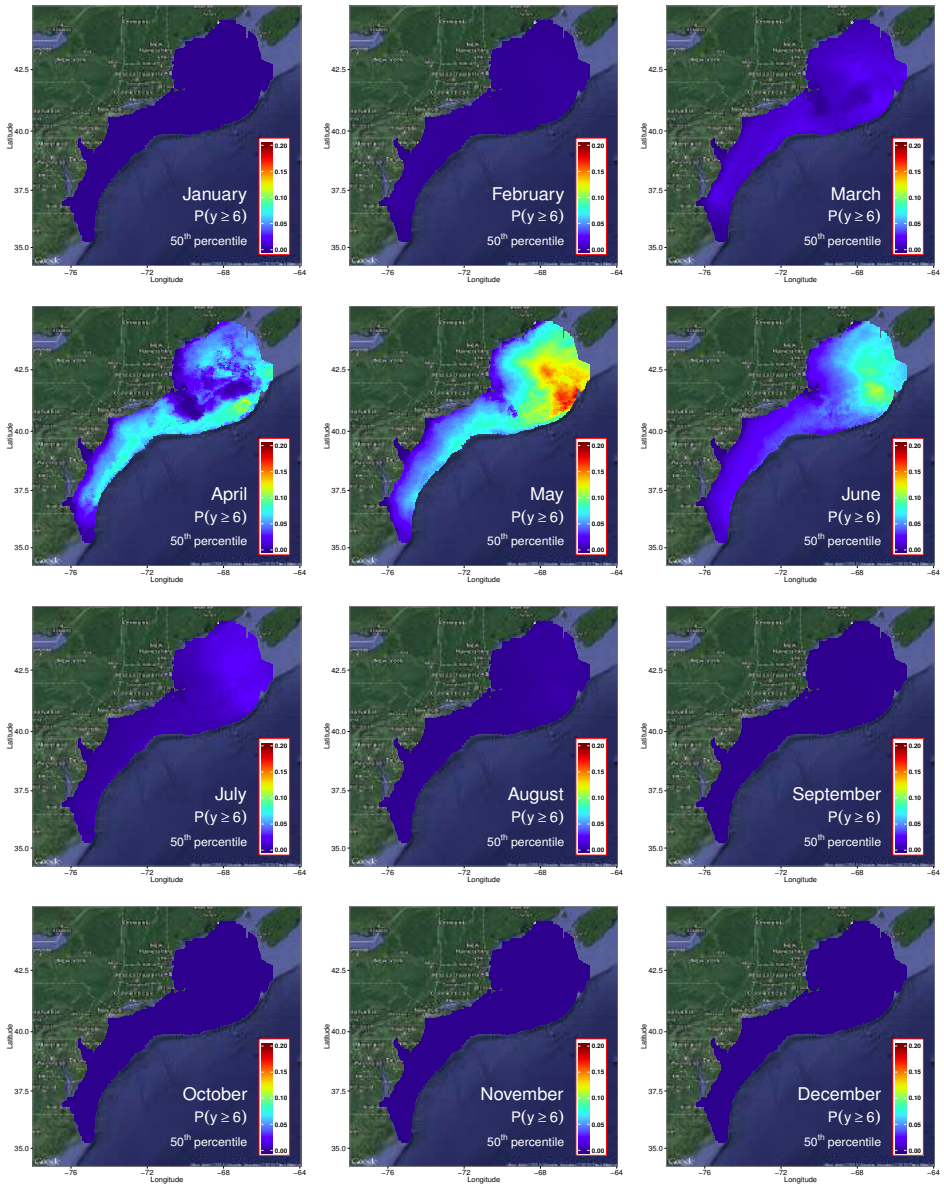

Figure 109: Sooty Shearwater: Probability of observing a large count during each month.

A.24 Surf Scoter

A.24.1 One-year summary

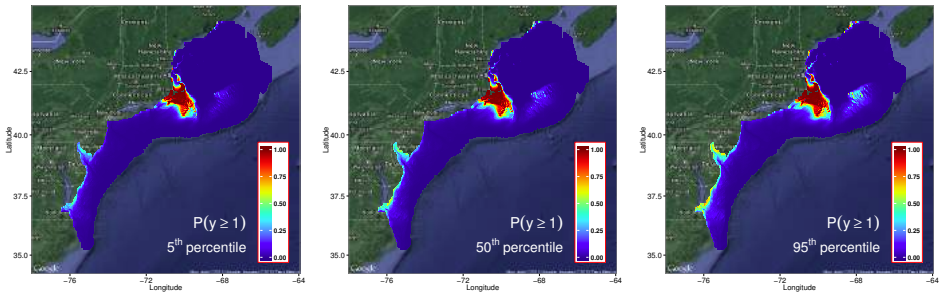

Figure 110: Surf Scoter: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

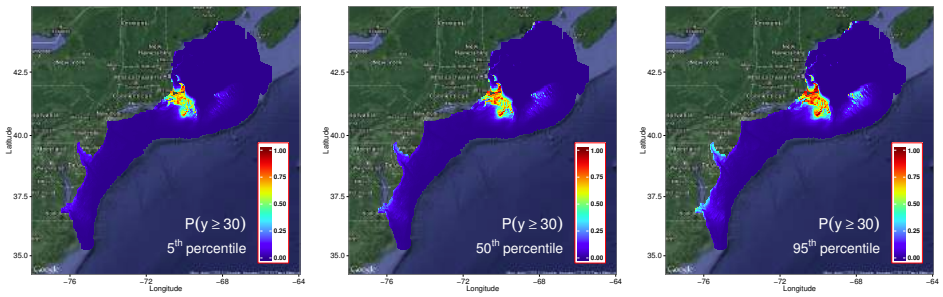

Figure 111: Surf Scoter: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.24.2 Monthly observations

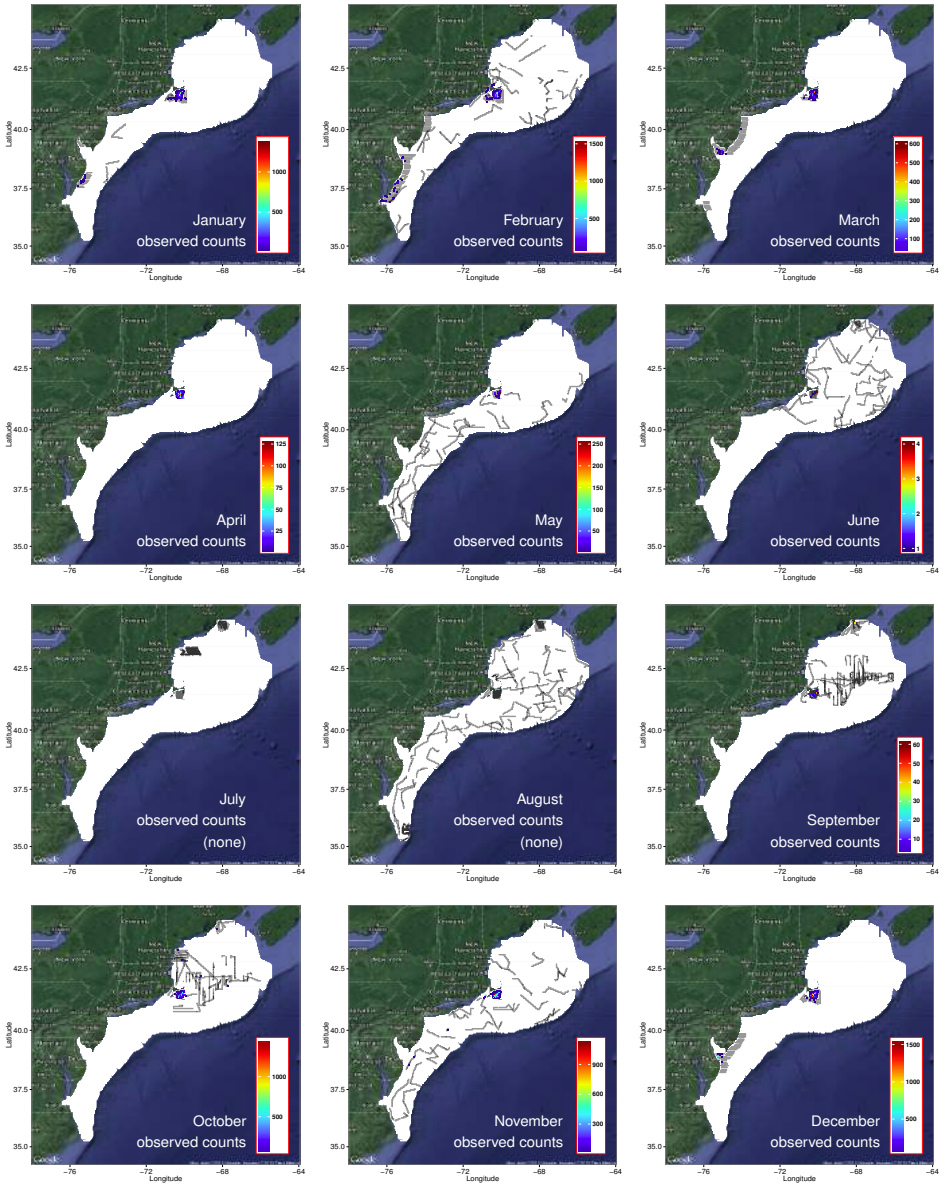

Figure 112: Surf Scoter: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.24.3 Monthly exposure maps

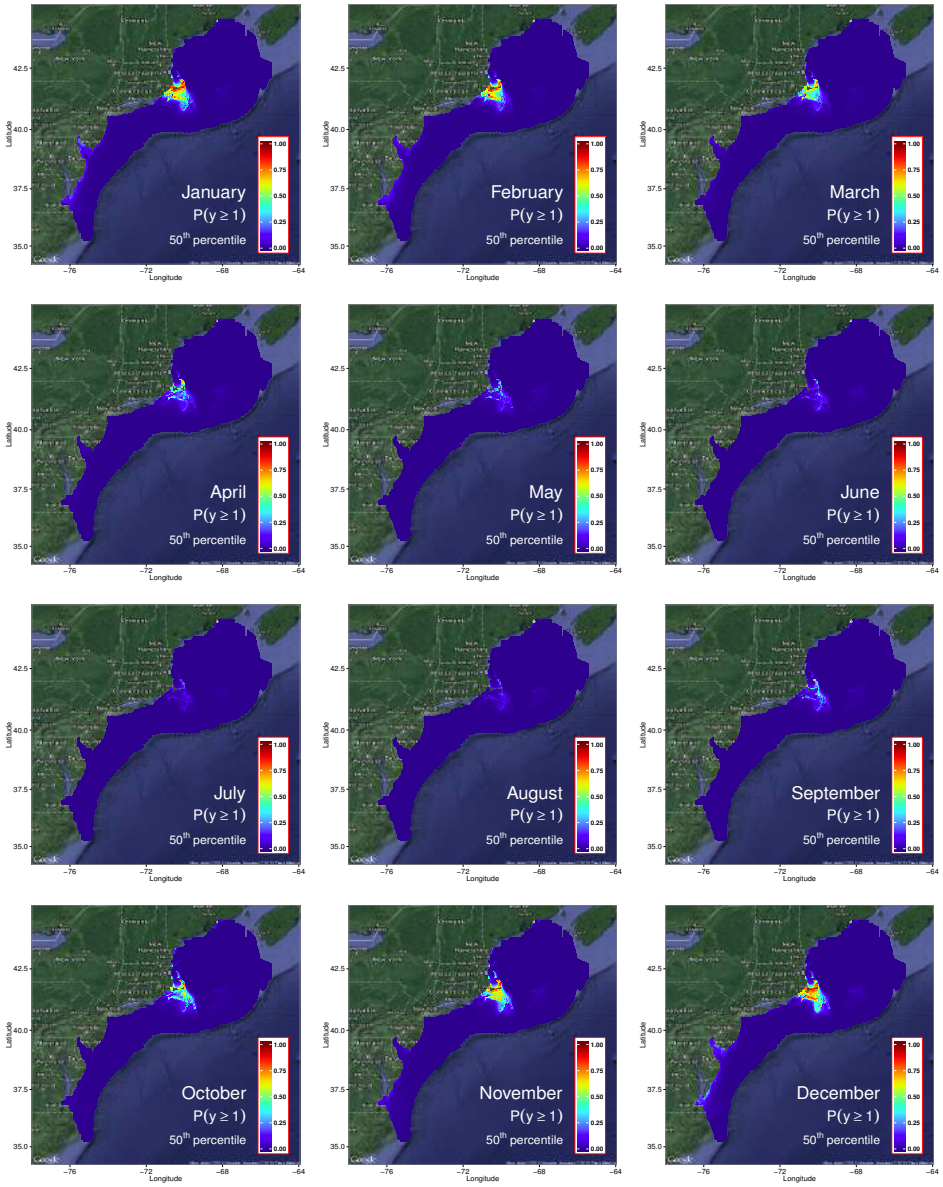

Figure 113: Surf Scoter: Probability of observing at least one individual during each month.

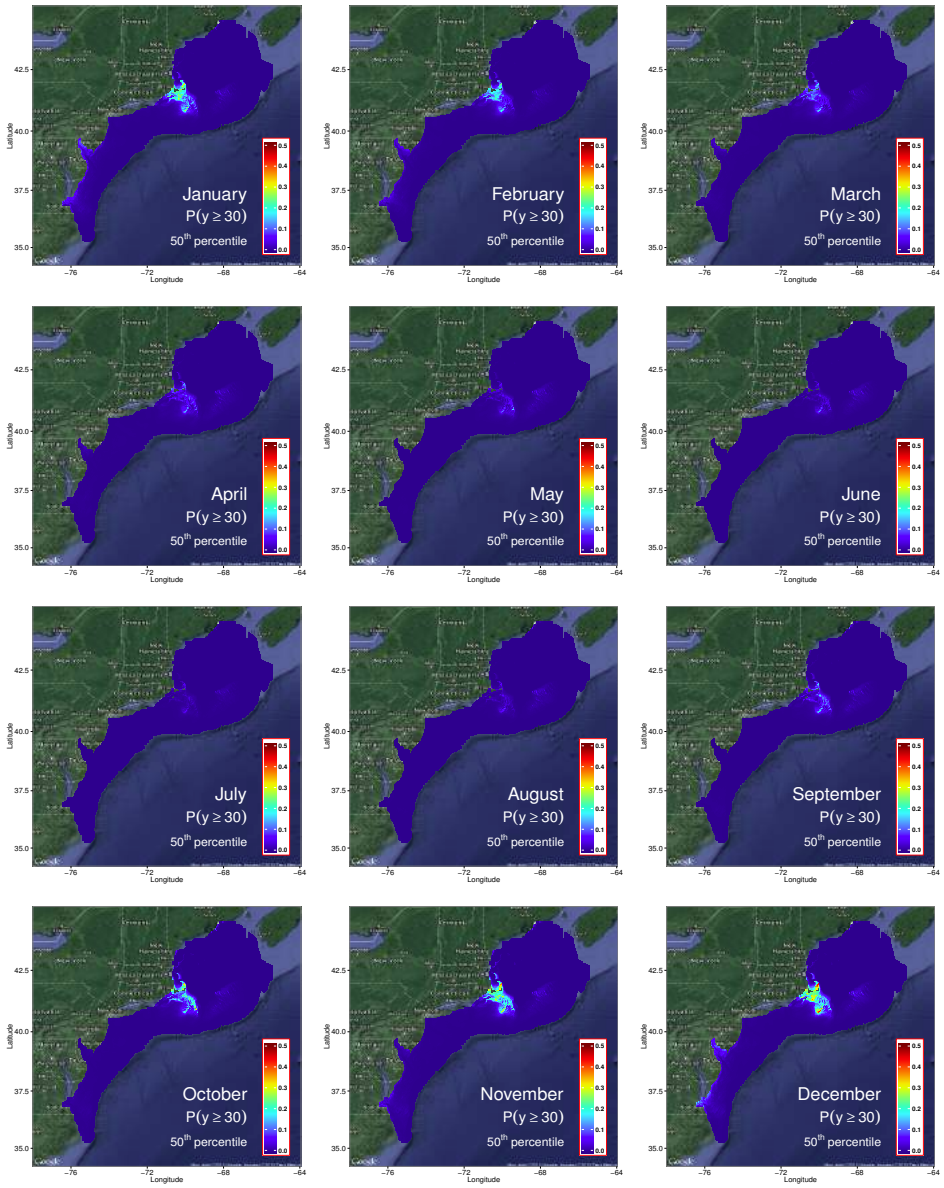

Figure 114: Surf Scoter: Probability of observing a large count during each month.

A.25 White-winged Scoter

A.25.1 One-year summary

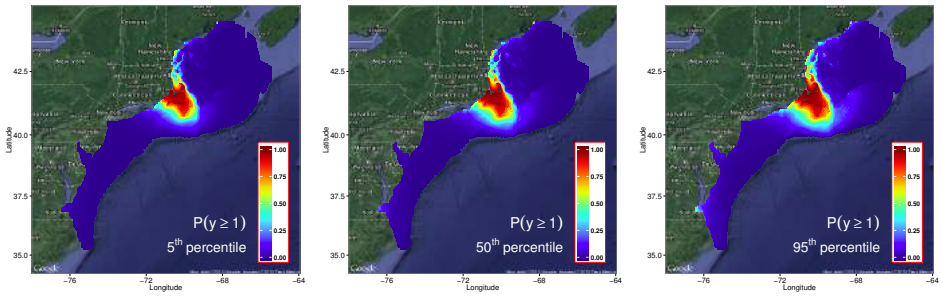

Figure 115: White-winged Scoter: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

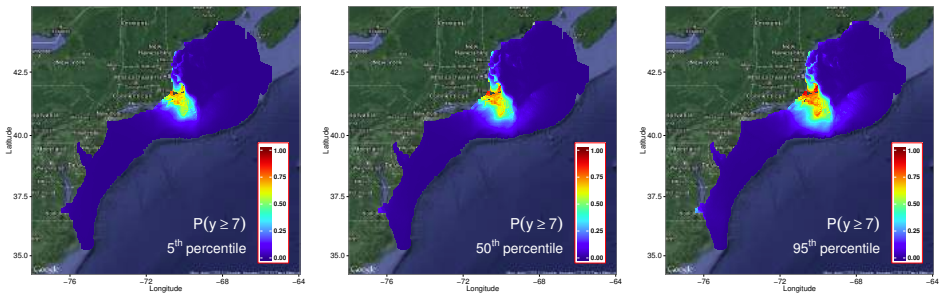

Figure 116: White-winged Scoter: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.25.2 Monthly observations

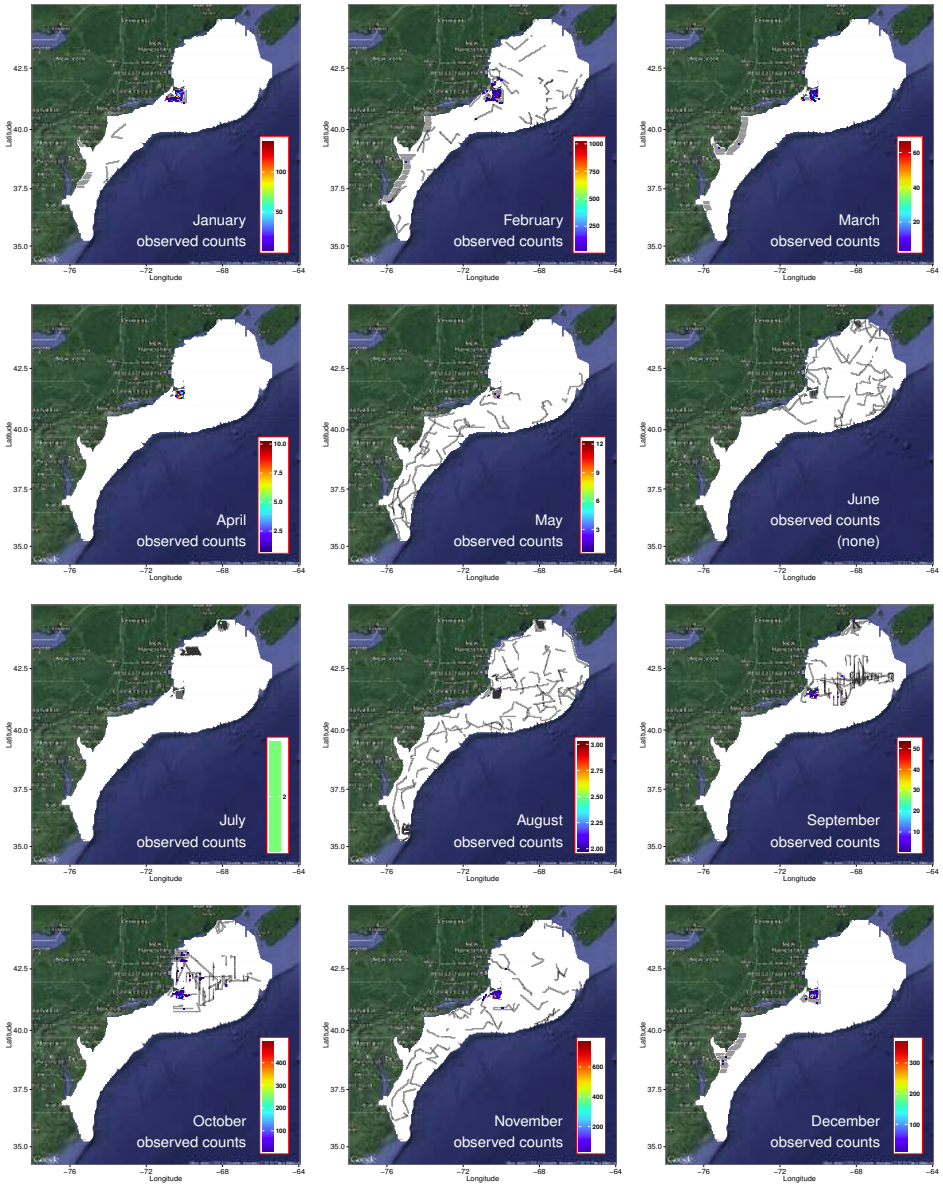

Figure 117: White-winged Scoter: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.25.3 Monthly exposure maps

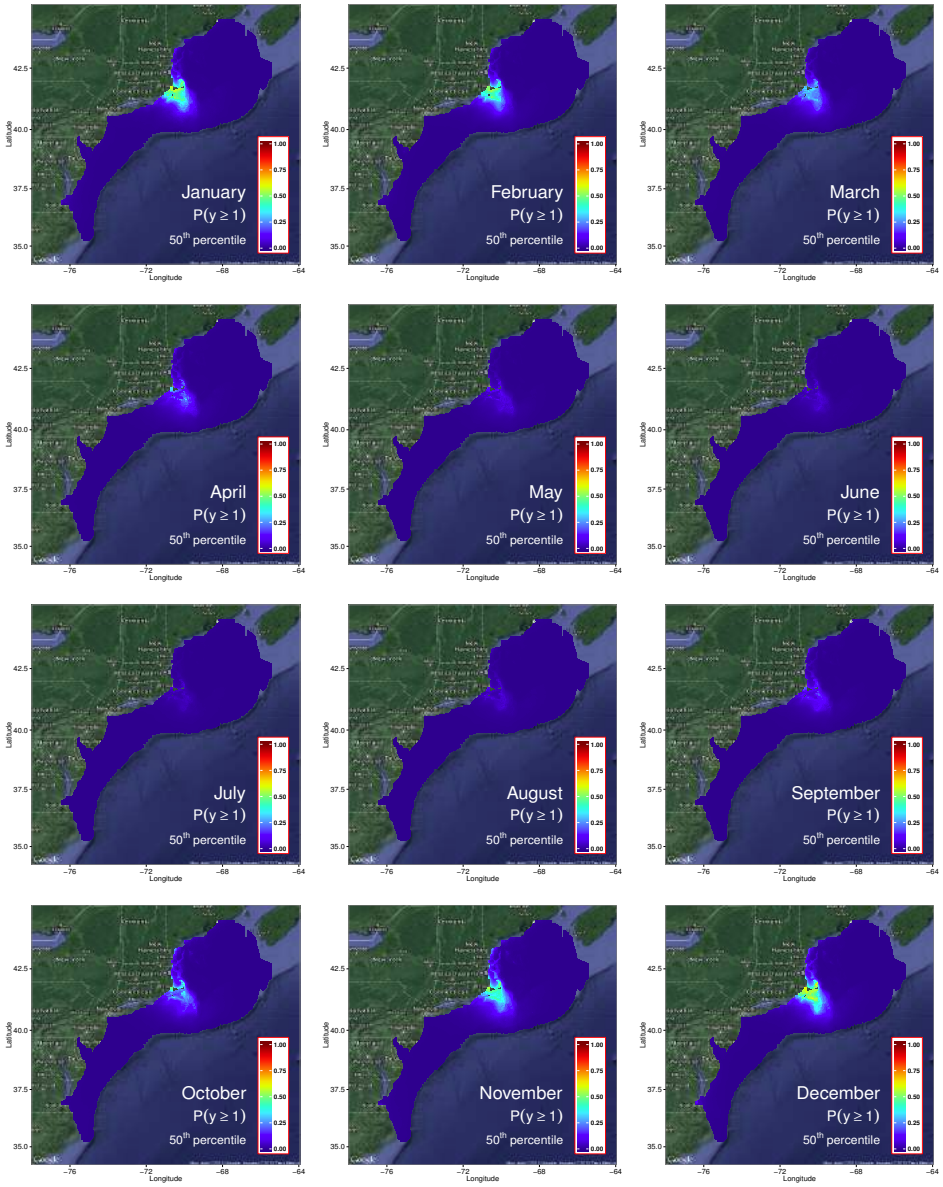

Figure 118: White-winged Scoter: Probability of observing at least one individual during each month.

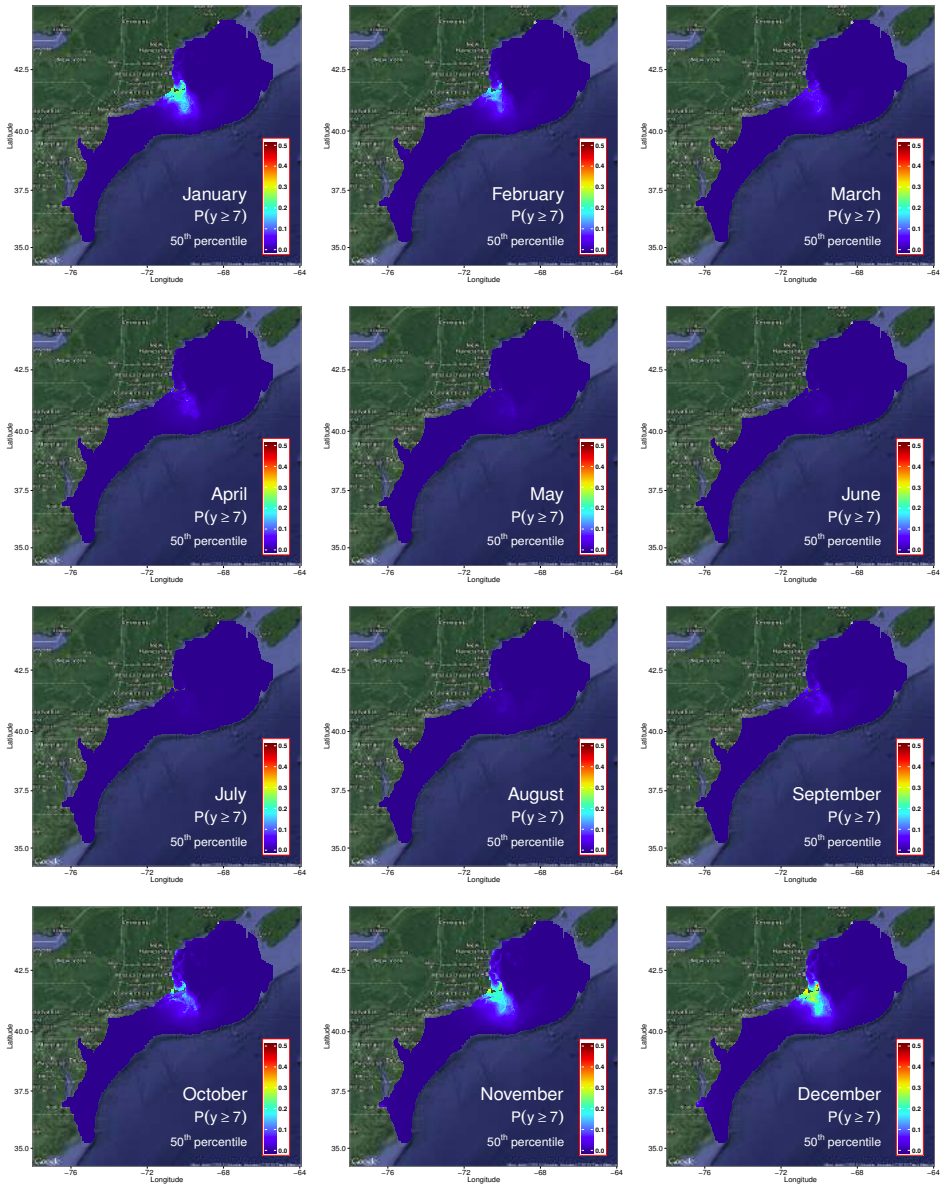

Figure 119: White-winged Scoter: Probability of observing a large count during each month.

A.26 Wilson's Storm-petrel

A.26.1 One-year summary

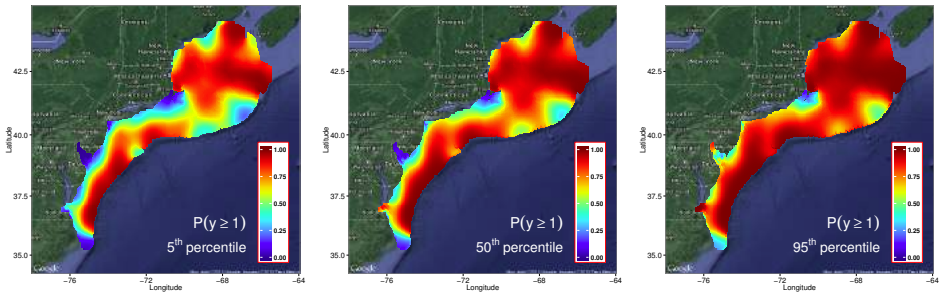

Figure 120: Wilson's Storm-petrel: Select quantiles of the estimated probability of observing at least one individual during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

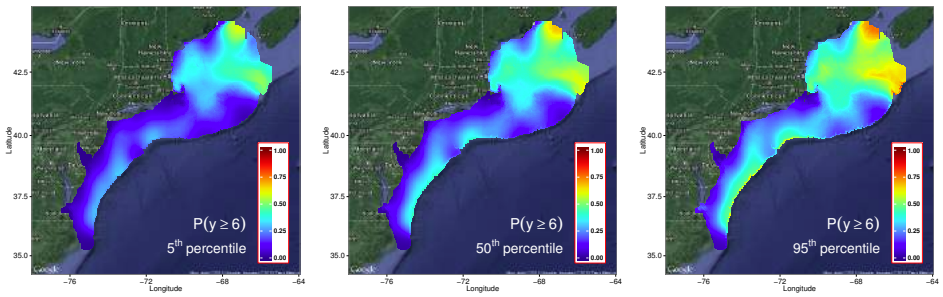

Figure 121: Wilson's Storm-petrel: Select quantiles of the estimated probability of observing at least one large count of individuals during the year. The median estimate is presented along with the 5^{th} and 95^{th} percentiles to show uncertainty in parameter estimates.

A.26.2 Monthly observations

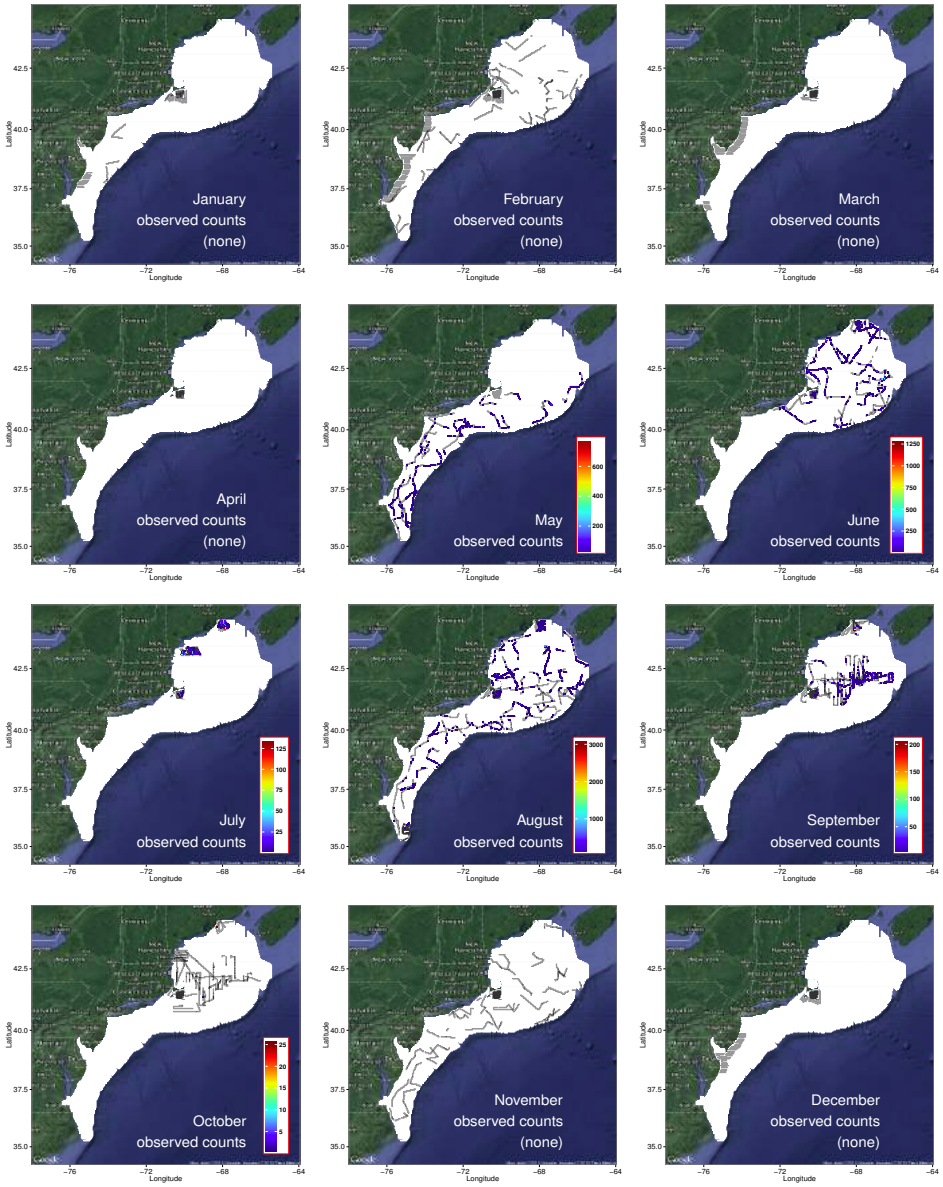

Figure 122: Wilson's Storm-petrel: Monthly maps of observations and survey effort. Areas of survey effort are colored grey. Observations in a grid cell are colored according to the total count for that calendar month.

A.26.3 Monthly exposure maps

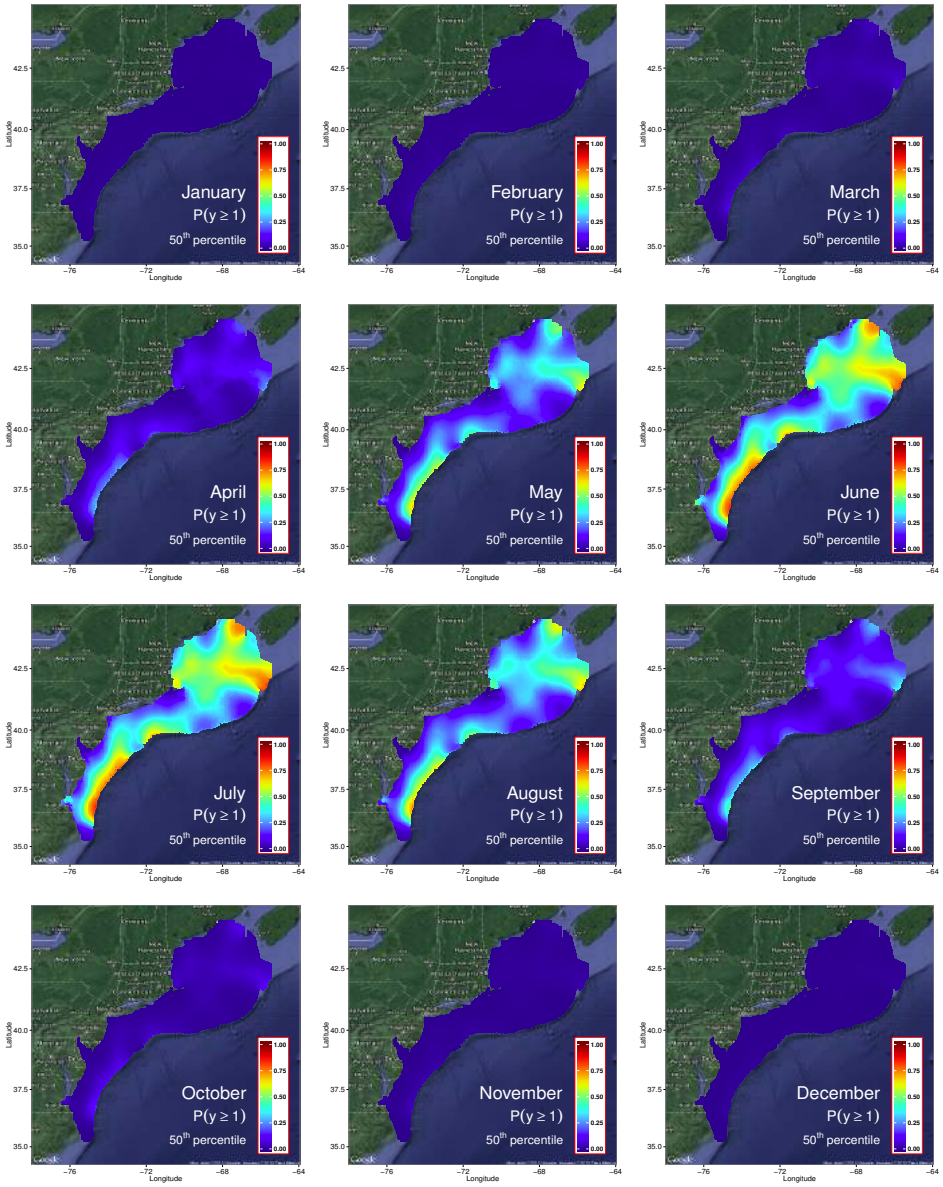

Figure 123: Wilson's Storm-petrel: Probability of observing at least one individual during each month.

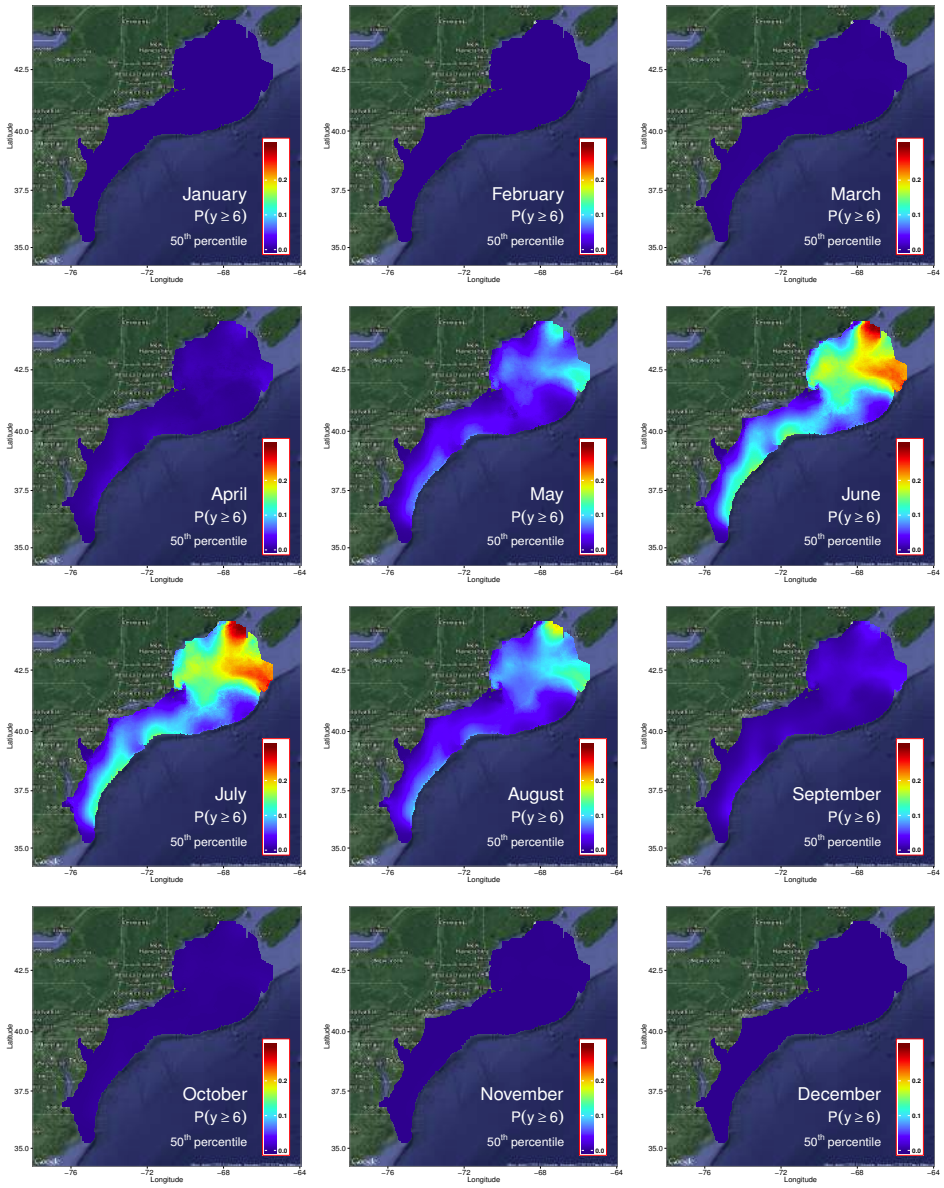

Figure 124: Wilson's Storm-petrel: Probability of observing a large count during each month.

B Diagnostics

B.1 Atlantic Puffin

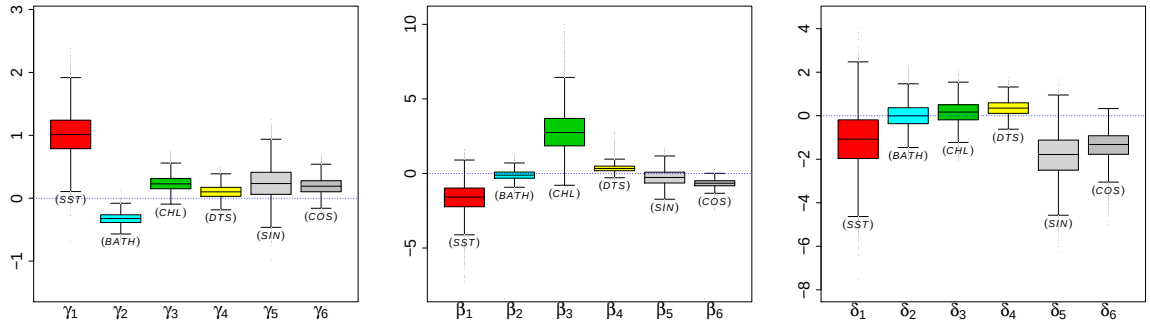

Figure 125: Atlantic Puffin: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	5.4165 (0.2121)	−2.227 (0.543)	−4.95168 (1.16135)
$\mathbf{x}_1$	Sea surface temp.	1.0150 (0.3339)	−1.643 (0.981)	−1.10335 (1.31728)
$\mathbf{x}_2$	Bathymetry	−0.3237 (0.0914)	−0.109 (0.310)	0.00564 (0.55216)
$\mathbf{x}_3$	Chlorophyll	0.2370 (0.1186)	2.822 (1.397)	0.14312 (0.52840)
$\mathbf{x}_4$	Distance to shore	0.1019 (0.1061)	0.344 (0.250)	0.35238 (0.36254)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	0.2380 (0.2590)	−0.282 (0.548)	−1.83878 (1.02867)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	0.1914 (0.1301)	−0.667 (0.254)	−1.37750 (0.64139)

Table 5: Atlantic Puffin: Posterior summary of regression coefficients.

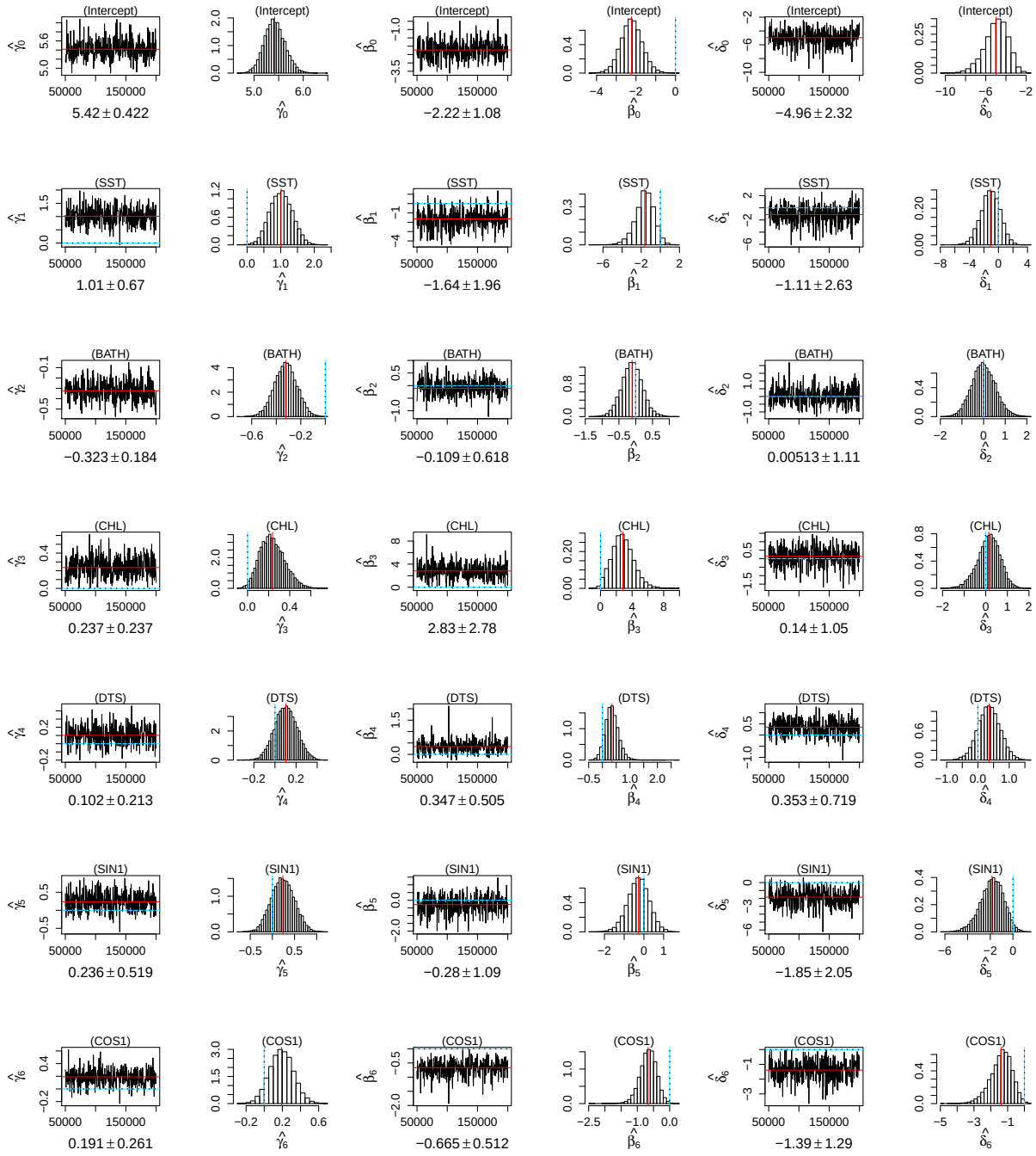

Figure 126: Atlantic Puffin: Traceplots and histograms of the posterior distributions of regression coefficients.

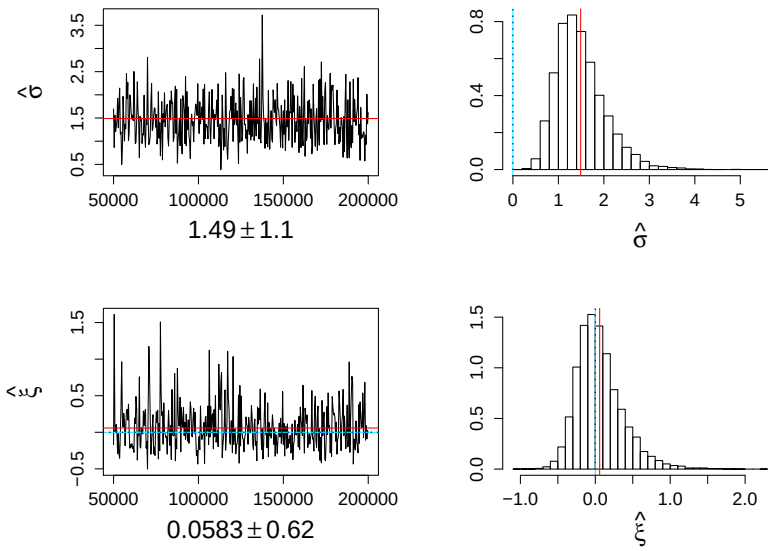

Figure 127: Atlantic Puffin: Traceplots and histograms of the posterior distributions of GPD parameters.

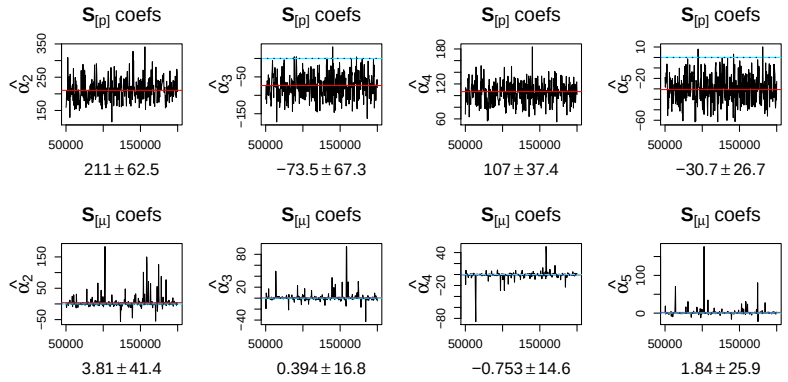

Figure 128: Atlantic Puffin: Traceplots and histograms of the posterior distributions of α coefficients.

B.2 Black-legged Kittiwake

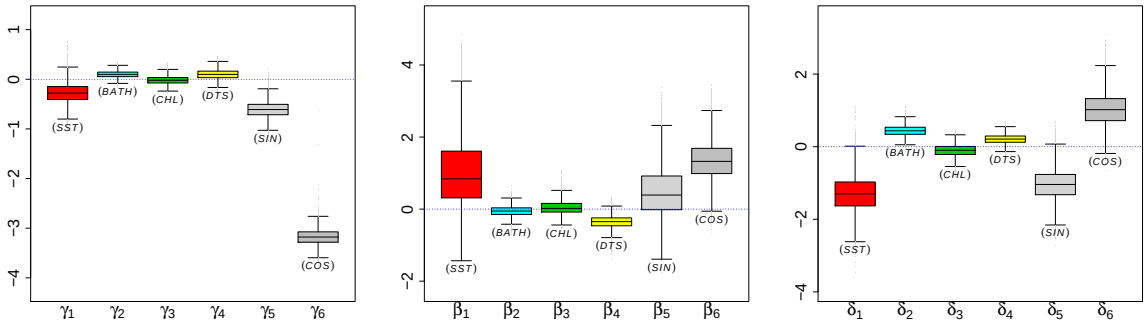

Figure 129: Black-legged Kittiwake: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	5.2470 (0.1570)	−2.0551 (0.4271)	−3.421 (0.455)
$\mathbf{x}_1$	Sea surface temp.	−0.2770 (0.1980)	1.0293 (0.9887)	−1.302 (0.499)
$\mathbf{x}_2$	Bathymetry	0.0987 (0.0672)	−0.0568 (0.1371)	0.440 (0.144)
$\mathbf{x}_3$	Chlorophyll	−0.0178 (0.0799)	0.0531 (0.1958)	−0.109 (0.166)
$\mathbf{x}_4$	Distance to shore	0.0982 (0.0973)	−0.3609 (0.1753)	0.207 (0.127)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	−0.6090 (0.1569)	0.4982 (0.7139)	−1.042 (0.424)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	−3.1787 (0.1586)	1.3481 (0.5337)	1.027 (0.456)

Table 6: Black-legged Kittiwake: Posterior summary of regression coefficients.

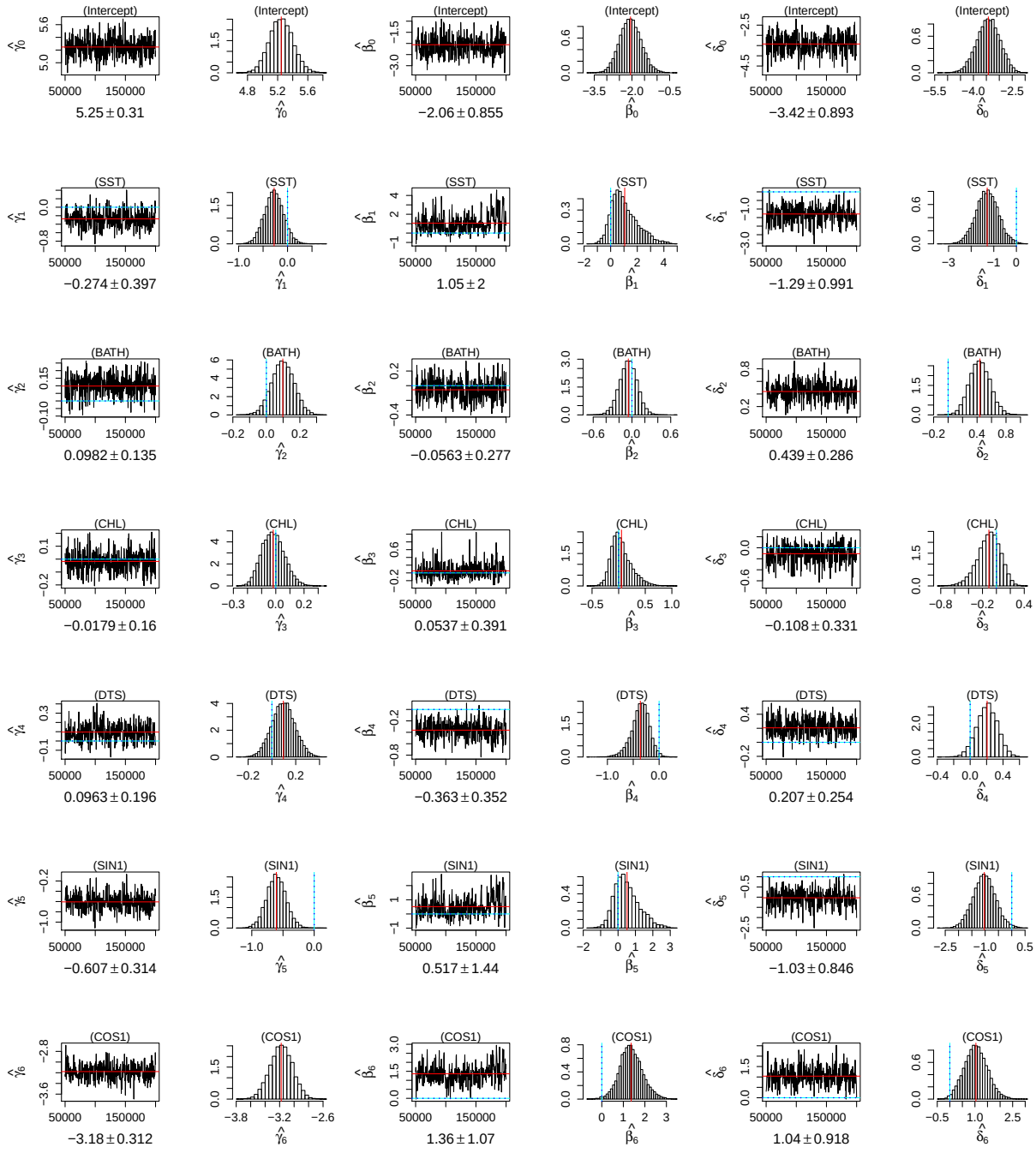

Figure 130: Black-legged Kittiwake: Traceplots and histograms of the posterior distributions of regression coefficients.

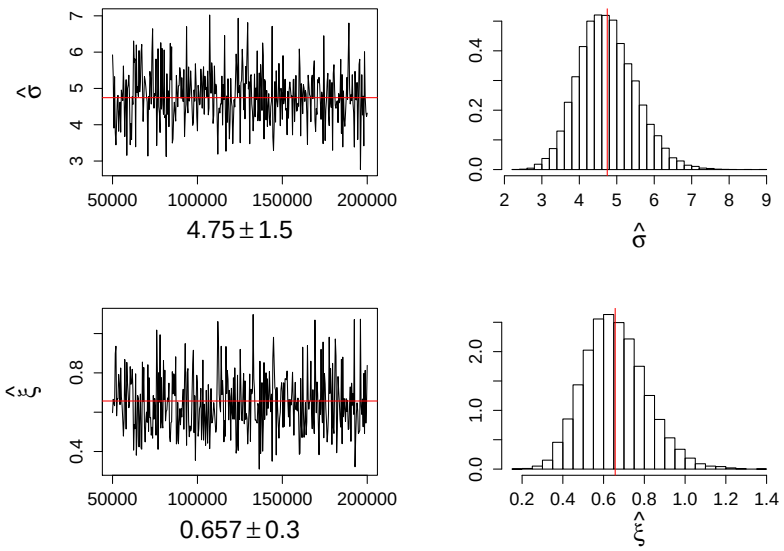

Figure 131: Black-legged Kittiwake: Traceplots and histograms of the posterior distributions of GPD parameters.

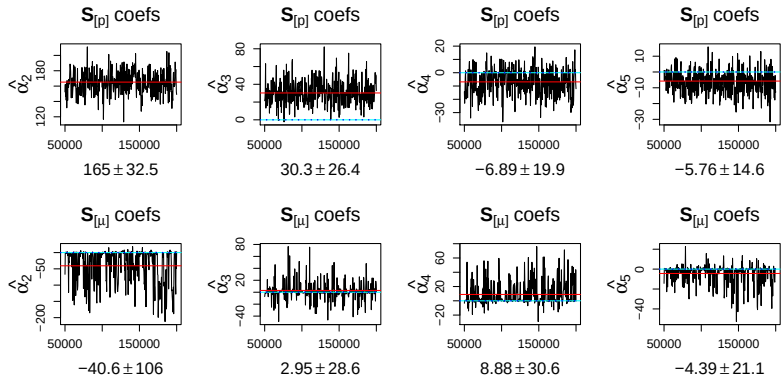

Figure 132: Black-legged Kittiwake: Traceplots and histograms of the posterior distributions of α coefficients.

B.3 Black Scoter

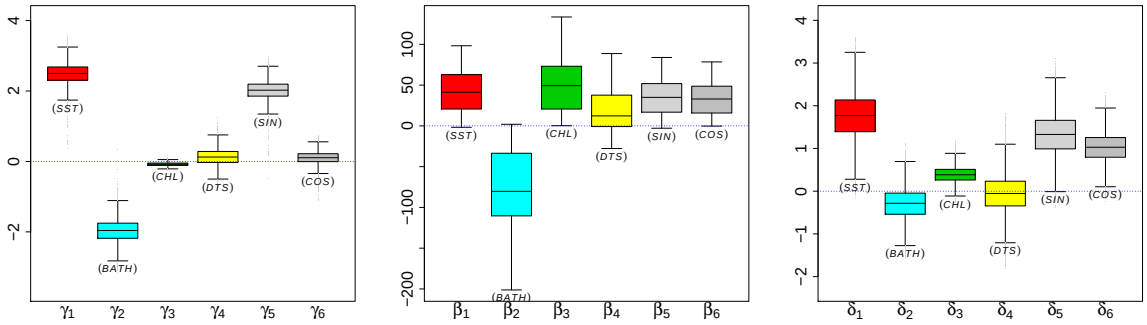

Figure 133: Black Scoter: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	7.3622 (0.2995)	102.7 (64.7)	0.7504 (0.5326)
$\mathbf{x}_1$	Sea surface temp.	2.4967 (0.2930)	41.2 (25.0)	1.7711 (0.5528)
$\mathbf{x}_2$	Bathymetry	−1.9671 (0.3164)	− 82.4 (54.4)	−0.2898 (0.3711)
$\mathbf{x}_3$	Chlorophyll	−0.0777 (0.0492)	49.0 (32.8)	0.3877 (0.1881)
$\mathbf{x}_4$	Distance to shore	0.1259 (0.2311)	18.9 (24.4)	−0.0532 (0.4458)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	2.0244 (0.2661)	34.3 (20.9)	1.3297 (0.4959)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	0.1078 (0.1730)	32.2 (19.7)	1.0284 (0.3430)

Table 7: Black Scoter: Posterior summary of regression coefficients.

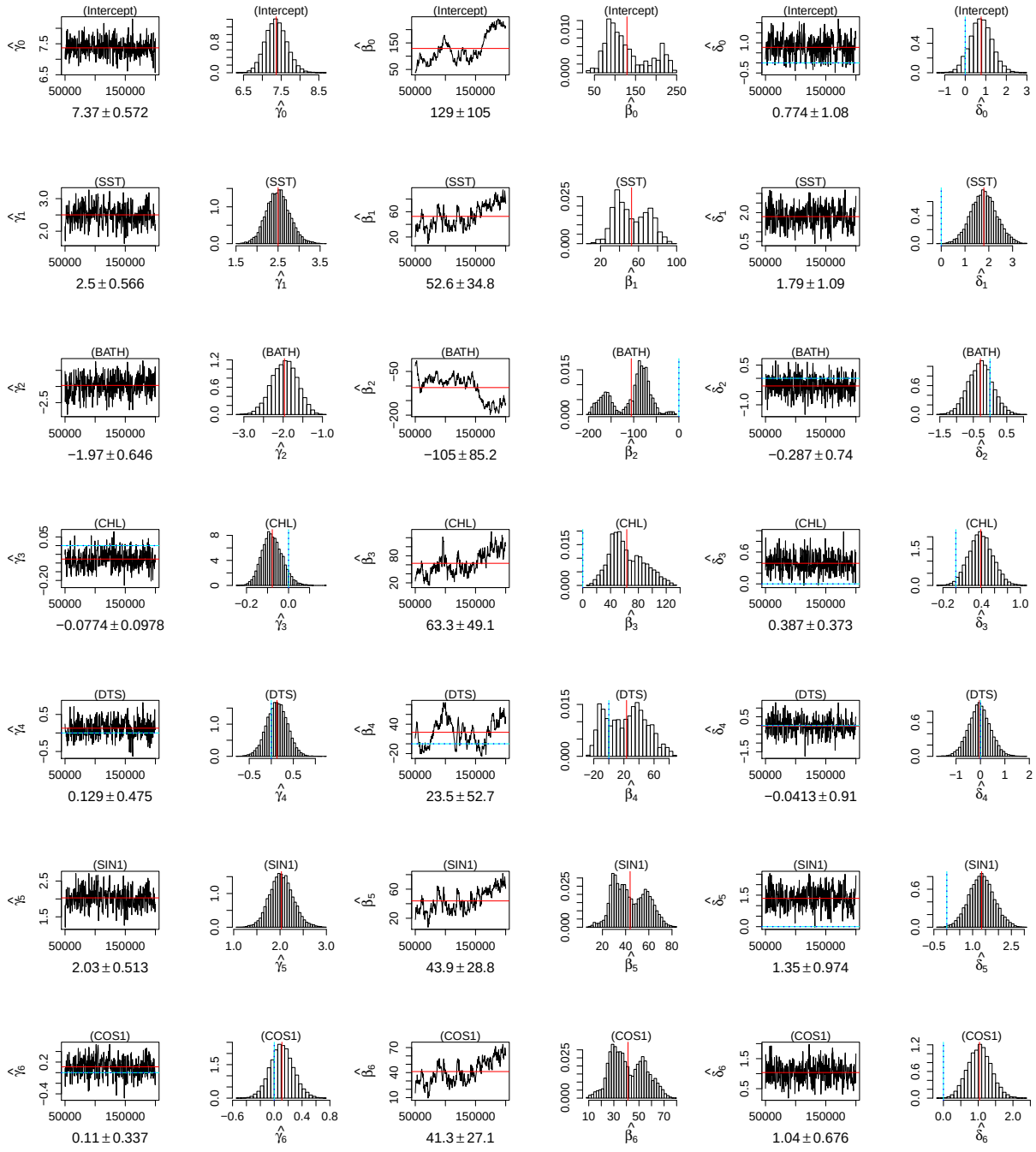

Figure 134: Black Scoter: Traceplots and histograms of the posterior distributions of regression coefficients.

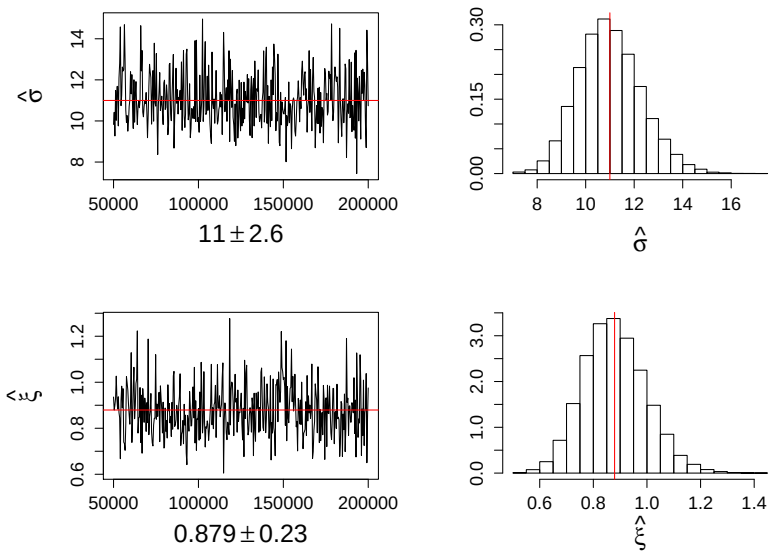

Figure 135: Black Scoter: Traceplots and histograms of the posterior distributions of GPD parameters.

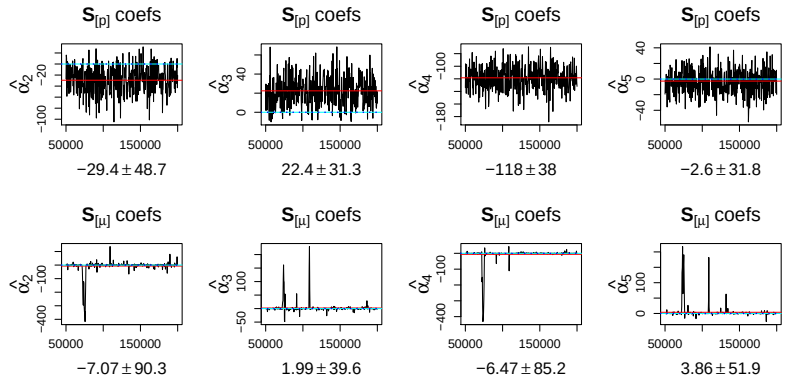

Figure 136: Black Scoter: Traceplots and histograms of the posterior distributions of α coefficients.

B.4 Bonaparte's Gull

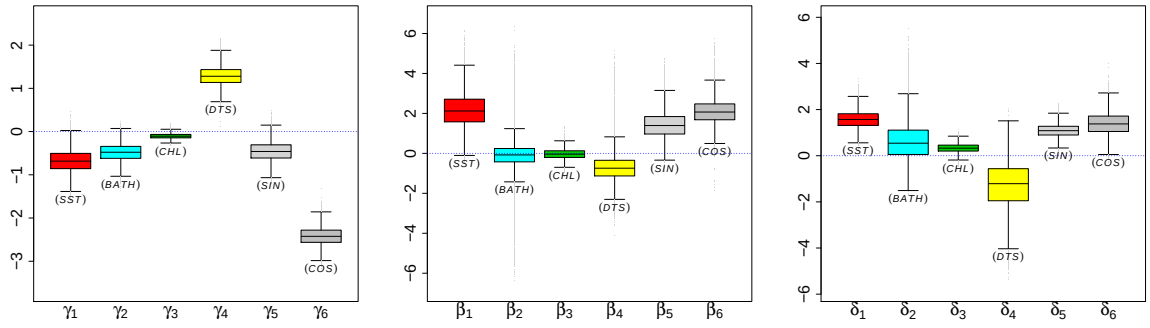

Figure 137: Bonaparte's Gull: Boxplots of the posterior distributions of regression coefficients.

Predictor		$\hat{\boldsymbol{\gamma}}$ mean (sd)	$\hat{\boldsymbol{\beta}}$ mean (sd)	$\hat{\boldsymbol{\delta}}$ mean (sd)
1	Intercept	6.4087 (0.2246)	−1.8401 (0.7944)	−3.998 (1.081)
$\mathbf{x}_1$	Sea surface temp.	−0.6776 (0.2649)	2.1662 (0.8393)	1.571 (0.373)
$\mathbf{x}_2$	Bathymetry	−0.4827 (0.2041)	−0.1145 (0.8258)	0.617 (0.834)
$\mathbf{x}_3$	Chlorophyll	−0.1036 (0.0599)	−0.0323 (0.2547)	0.330 (0.192)
$\mathbf{x}_4$	Distance to shore	1.2849 (0.2227)	−0.6912 (0.7806)	−1.291 (1.003)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	−0.4545 (0.2291)	1.4238 (0.6593)	1.090 (0.285)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	−2.4208 (0.2074)	2.1039 (0.6383)	1.395 (0.493)

Table 8: Bonaparte's Gull: Posterior summary of regression coefficients.

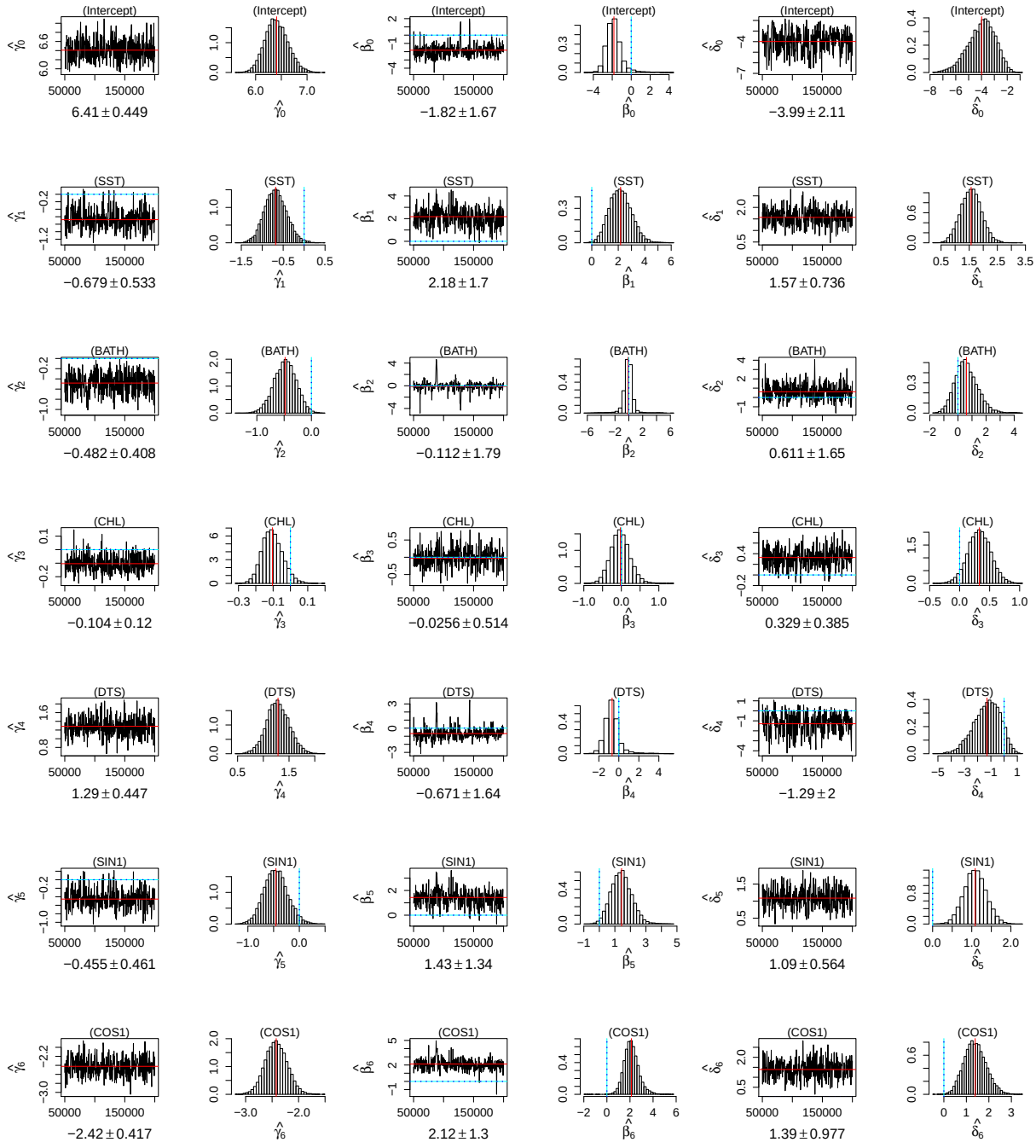

Figure 138: Bonaparte's Gull: Traceplots and histograms of the posterior distributions of regression coefficients.

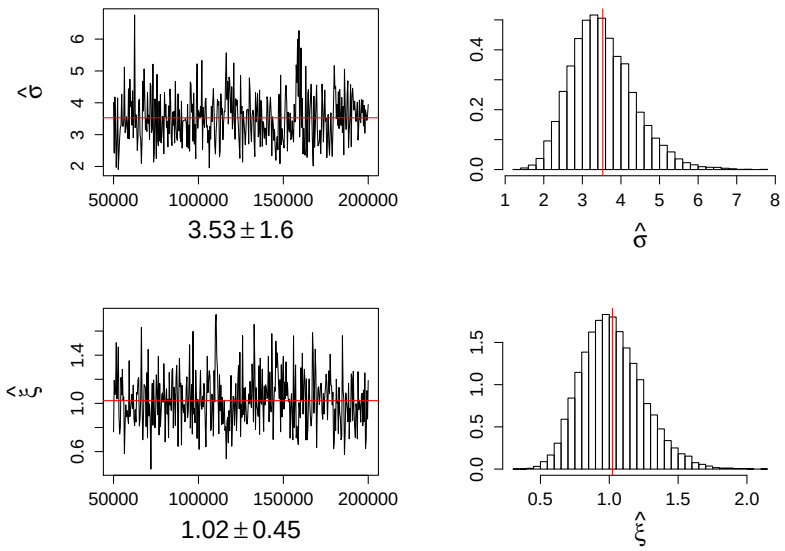

Figure 139: Bonaparte's Gull: Traceplots and histograms of the posterior distributions of GPD parameters.

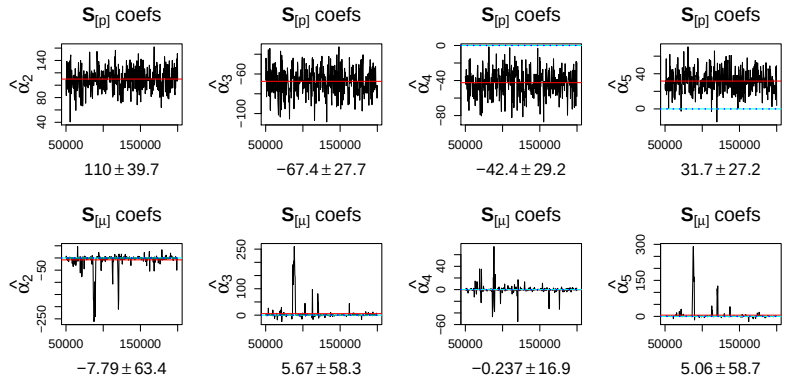

Figure 140: Bonaparte's Gull: Traceplots and histograms of the posterior distributions of α coefficients.

B.5 Common Eider

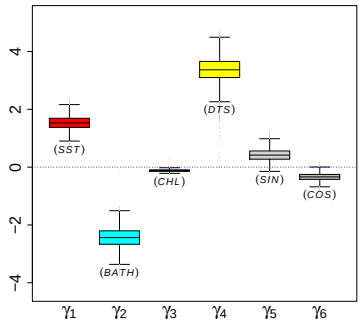

Figure 141: Common Eider: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\gamma}$
Predictor		mean (sd)
1	Intercept	11.7839 (1.0184)
$\mathbf{x}_1$	Sea surface temp.	1.5345 (0.2314)
$\mathbf{x}_2$	Bathymetry	− 2.4340 (0.3365)
$\mathbf{x}_3$	Chlorophyll	− 0.1174 (0.0357)
$\mathbf{x}_4$	Distance to shore	3.3705 (0.4239)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	0.4201 (0.2098)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	− 0.3364 (0.1270)

Table 9: Common Eider: Posterior summary of regression coefficients.

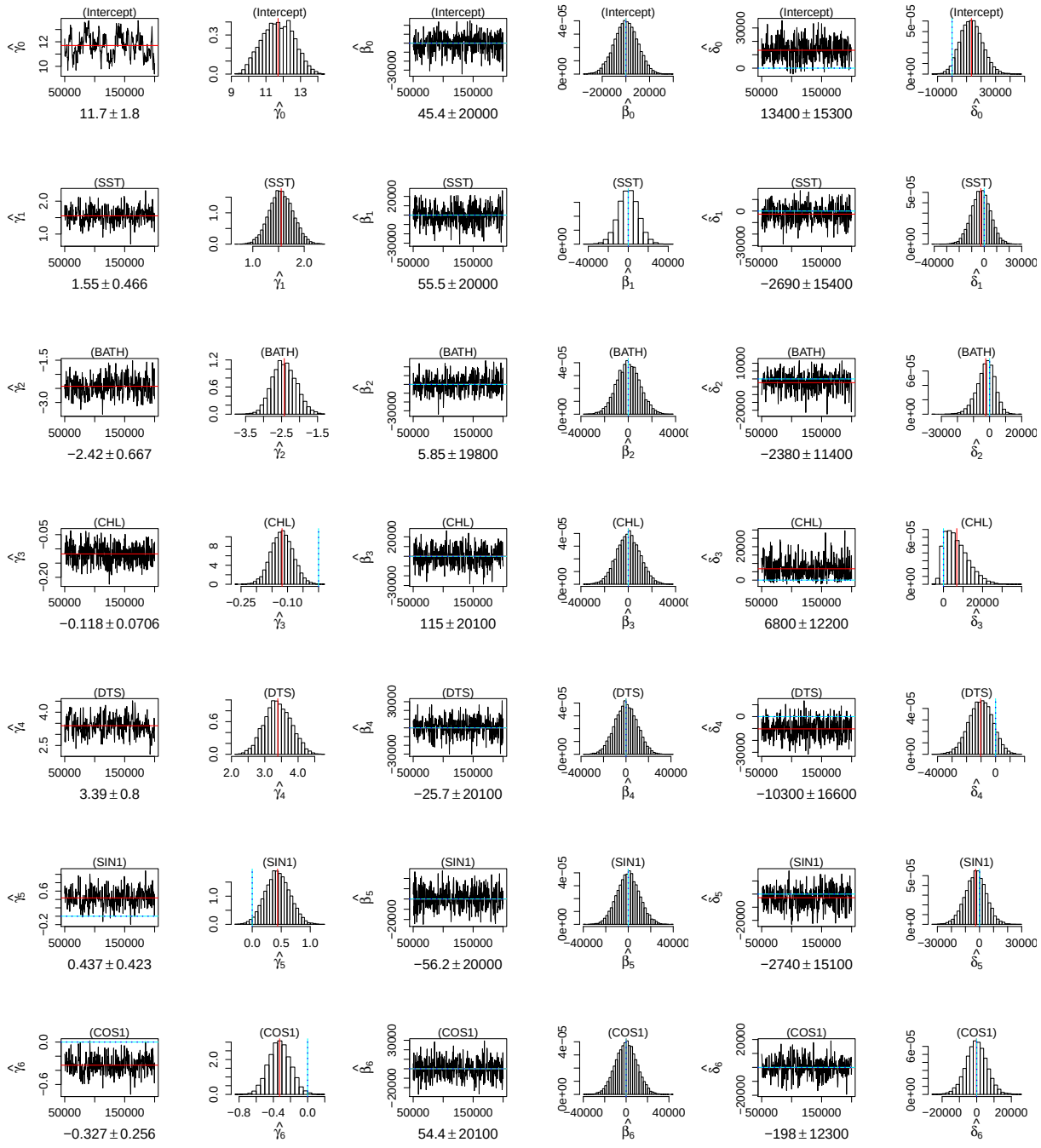

Figure 142: Common Eider: Traceplots and histograms of the posterior distributions of regression coefficients.

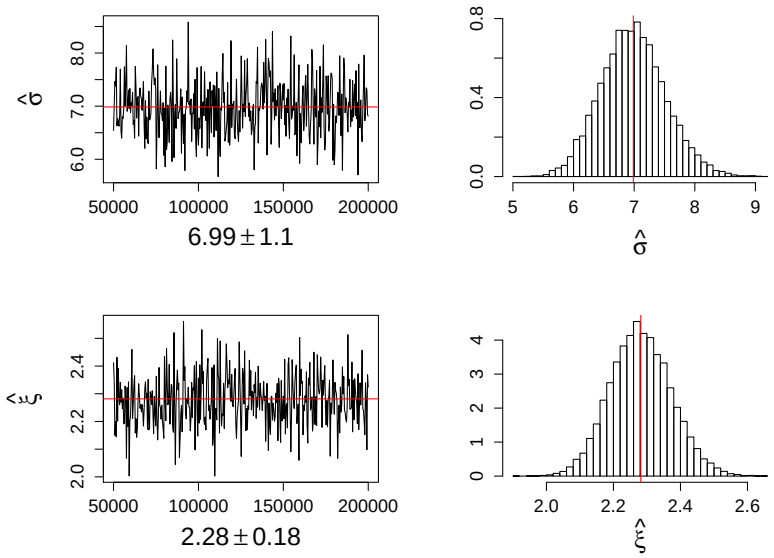

Figure 143: Common Eider: Traceplots and histograms of the posterior distributions of GPD parameters.

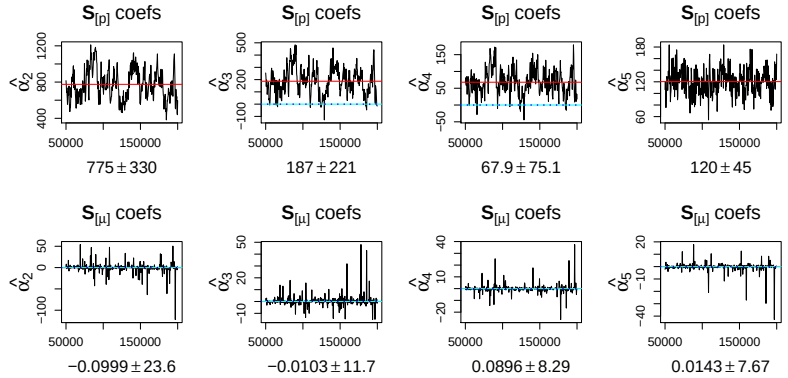

Figure 144: Common Eider: Traceplots and histograms of the posterior distributions of α coefficients.

B.6 Common Loon

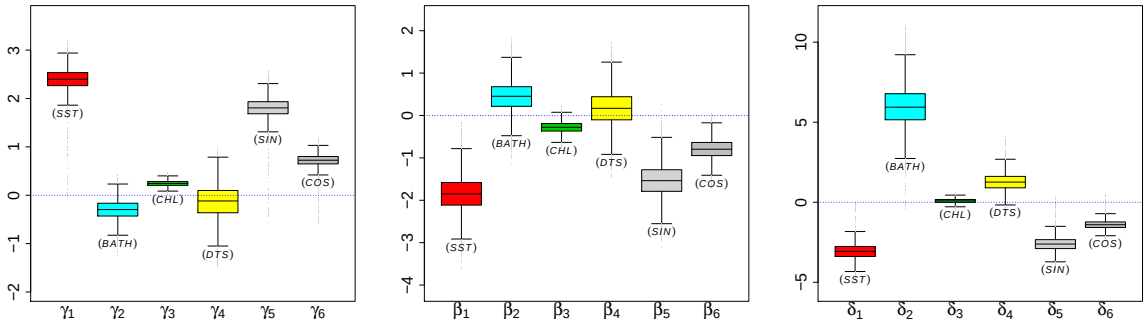

Figure 145: Common Loon: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	5.3249 (0.2120)	−2.602 (0.332)	−8.4029 (1.3827)
$\mathbf{x}_1$	Sea surface temp.	2.3978 (0.2274)	−1.845 (0.399)	−3.0746 (0.4772)
$\mathbf{x}_2$	Bathymetry	−0.2986 (0.1959)	0.451 (0.350)	5.9934 (1.2826)
$\mathbf{x}_3$	Chlorophyll	0.2438 (0.0585)	−0.281 (0.131)	0.0758 (0.1382)
$\mathbf{x}_4$	Distance to shore	−0.1318 (0.3416)	0.163 (0.420)	1.2539 (0.5563)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	1.8060 (0.2130)	−1.533 (0.377)	−2.6139 (0.4264)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	0.7234 (0.1284)	−0.792 (0.227)	−1.4024 (0.2633)

Table 10: Common Loon: Posterior summary of regression coefficients.

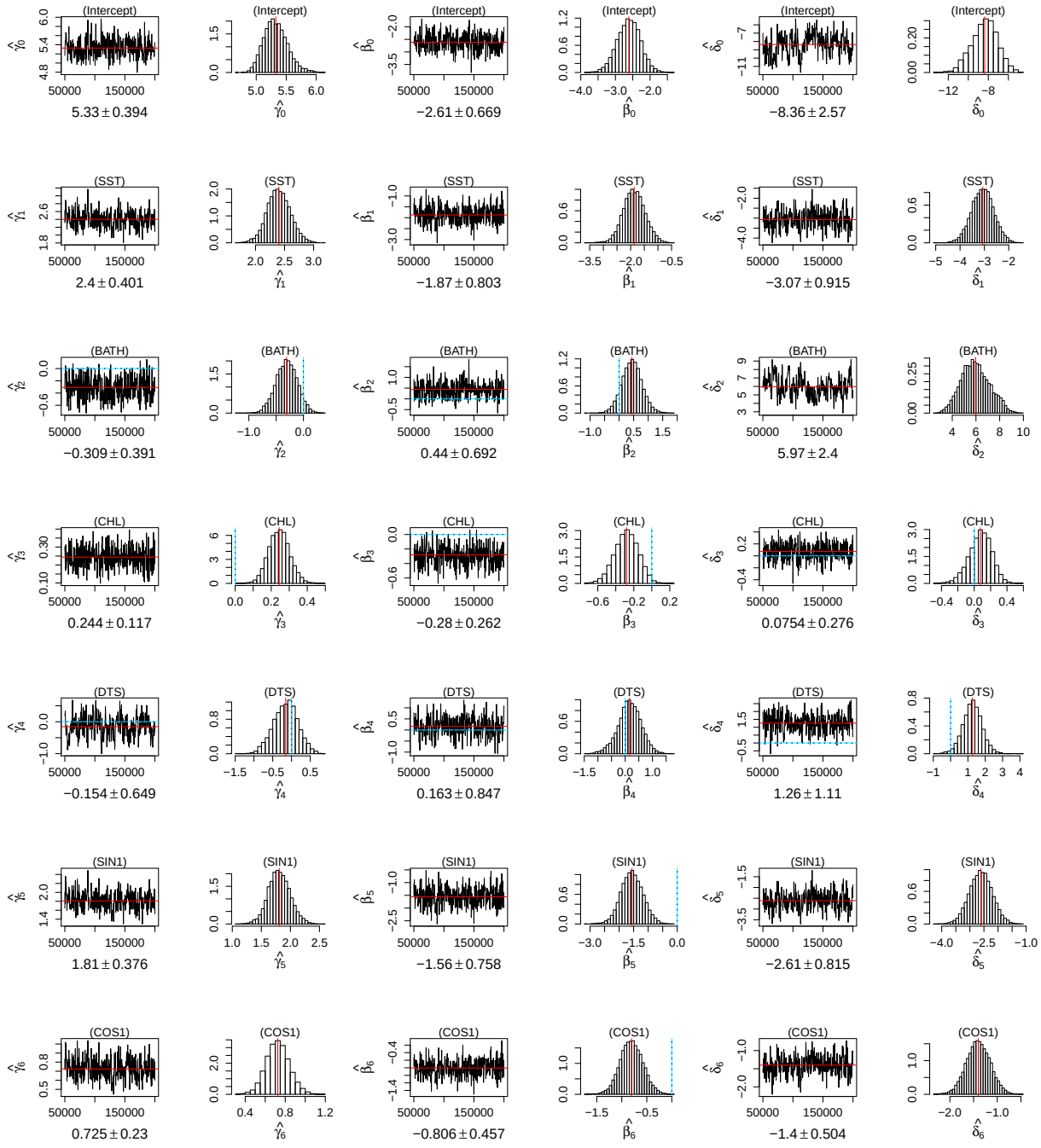

Figure 146: Common Loon: Traceplots and histograms of the posterior distributions of regression coefficients.

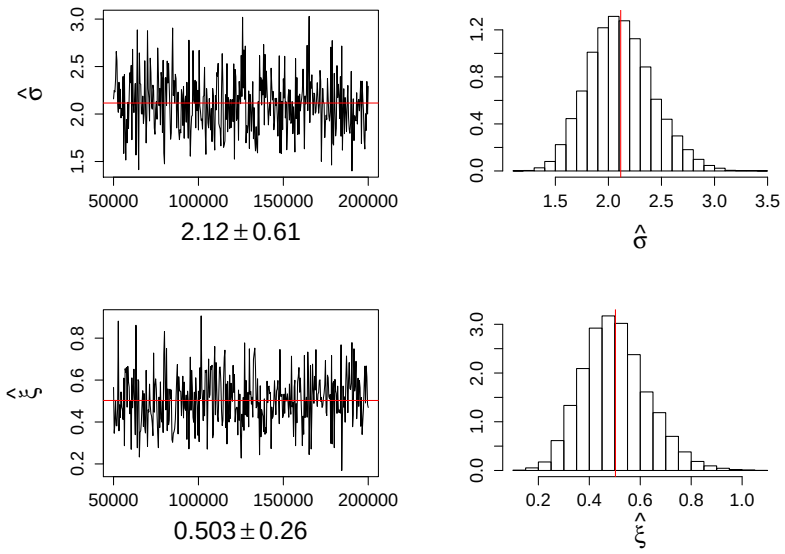

Figure 147: Common Loon: Traceplots and histograms of the posterior distributions of GPD parameters.

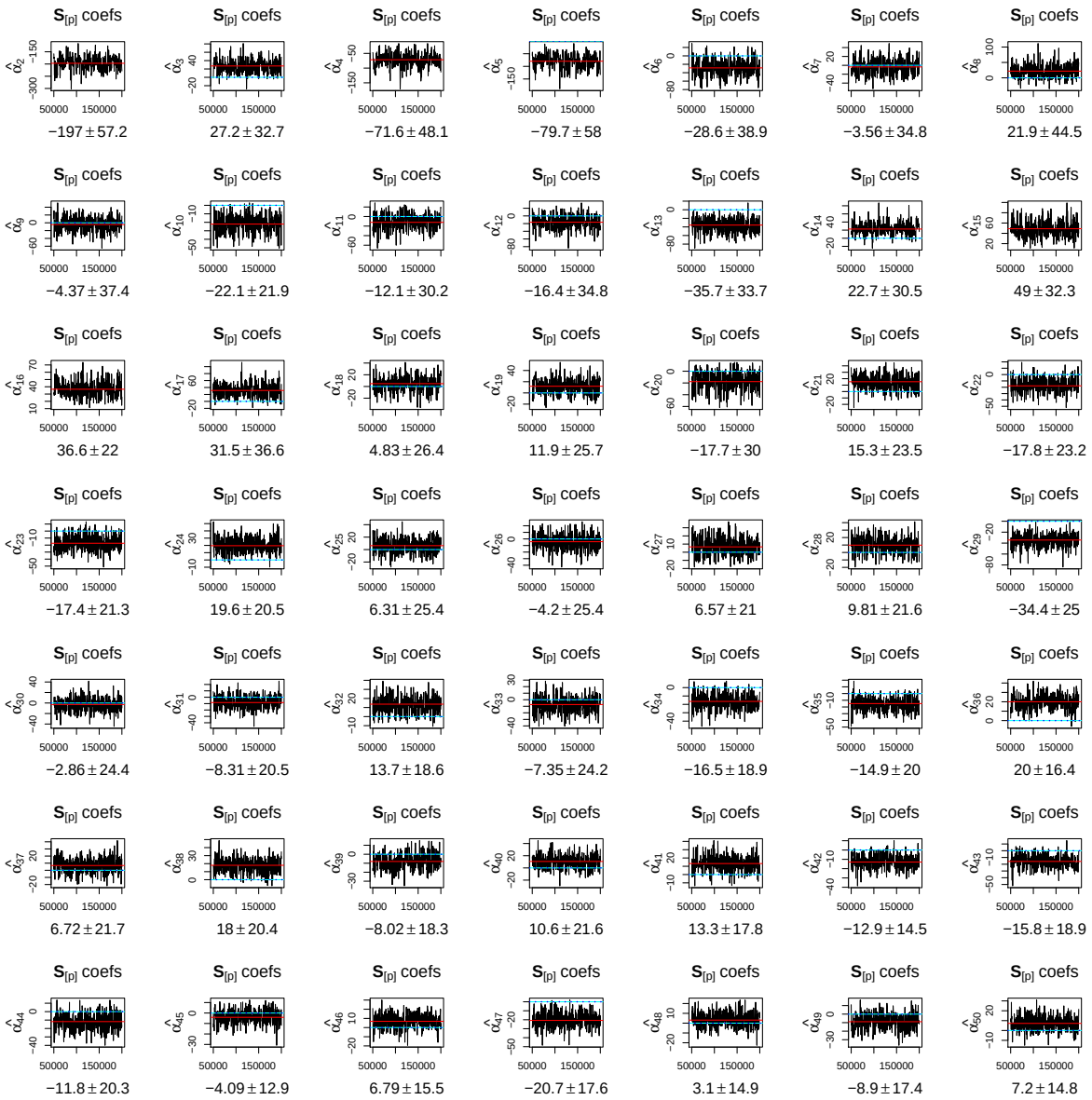

Figure 148: Common Loon: Traceplots and histograms of the posterior distributions of α coefficients in the spatial regression of $\mathbf{p}$.

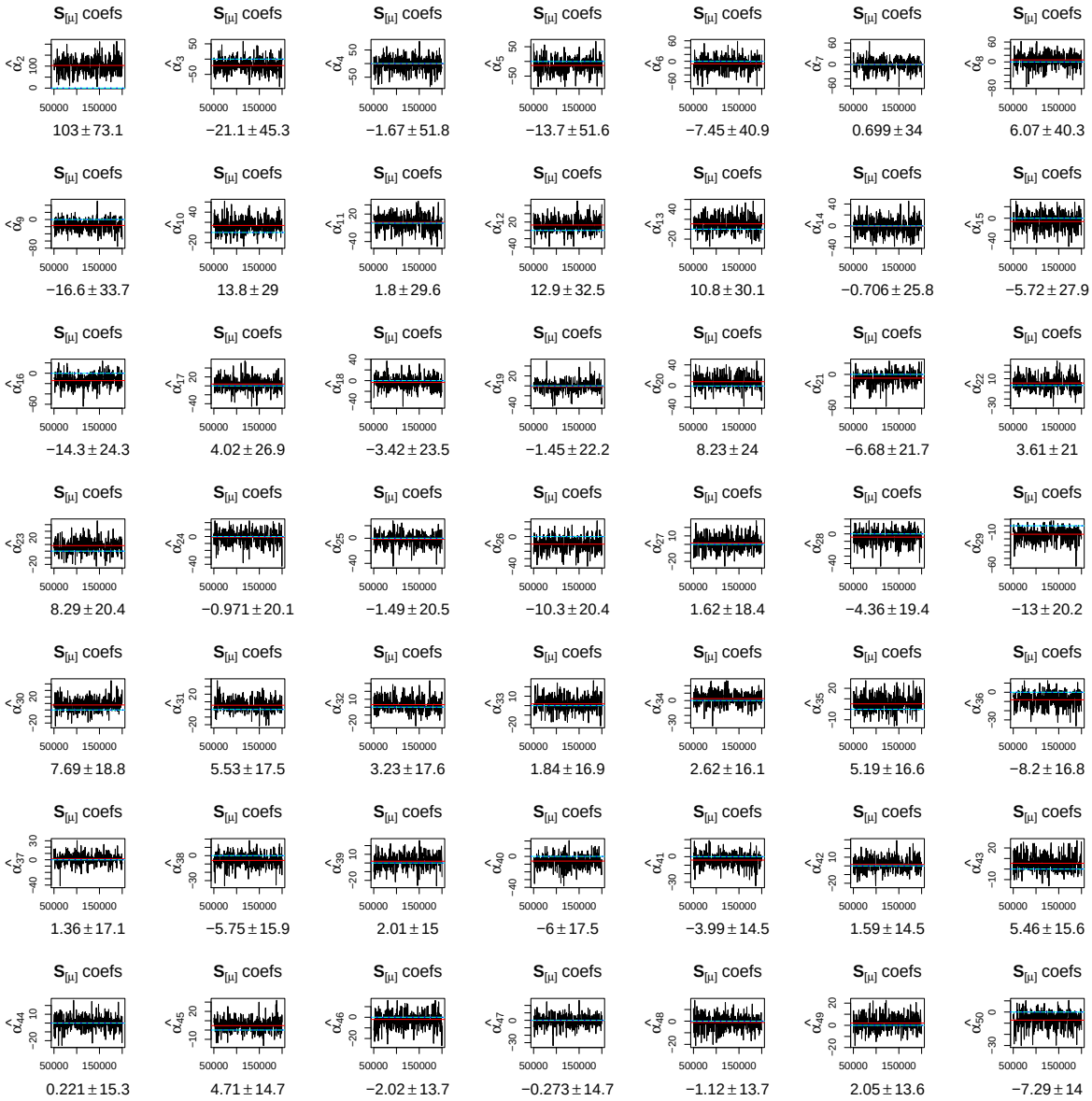

Figure 149: Common Loon: Traceplots and histograms of the posterior distributions of α coefficients in the spatial regression of $\mathbf{m}$.

B.7 Common Tern

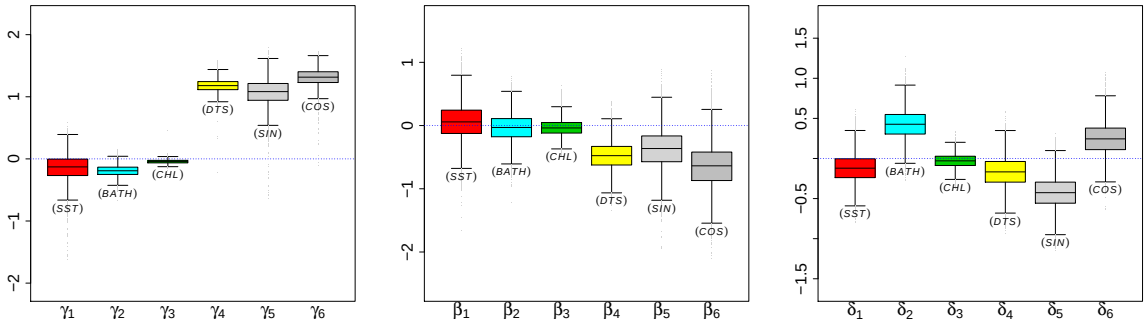

Figure 150: Common Tern: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	4.8563 (0.1413)	−2.0463 (0.3716)	−1.0923 (0.2597)
$\mathbf{x}_1$	Sea surface temp.	−0.1393 (0.2021)	0.0564 (0.2824)	−0.1210 (0.1750)
$\mathbf{x}_2$	Bathymetry	−0.1923 (0.0872)	−0.0359 (0.2174)	0.4273 (0.1835)
$\mathbf{x}_3$	Chlorophyll	−0.0436 (0.0309)	−0.0338 (0.1266)	−0.0296 (0.0872)
$\mathbf{x}_4$	Distance to shore	1.1805 (0.0978)	−0.4773 (0.2178)	−0.1673 (0.1935)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	1.0732 (0.2104)	−0.3732 (0.3108)	−0.4256 (0.1969)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	1.3127 (0.1344)	−0.6442 (0.3463)	0.2459 (0.2003)

Table 11: Common Tern: Posterior summary of regression coefficients.

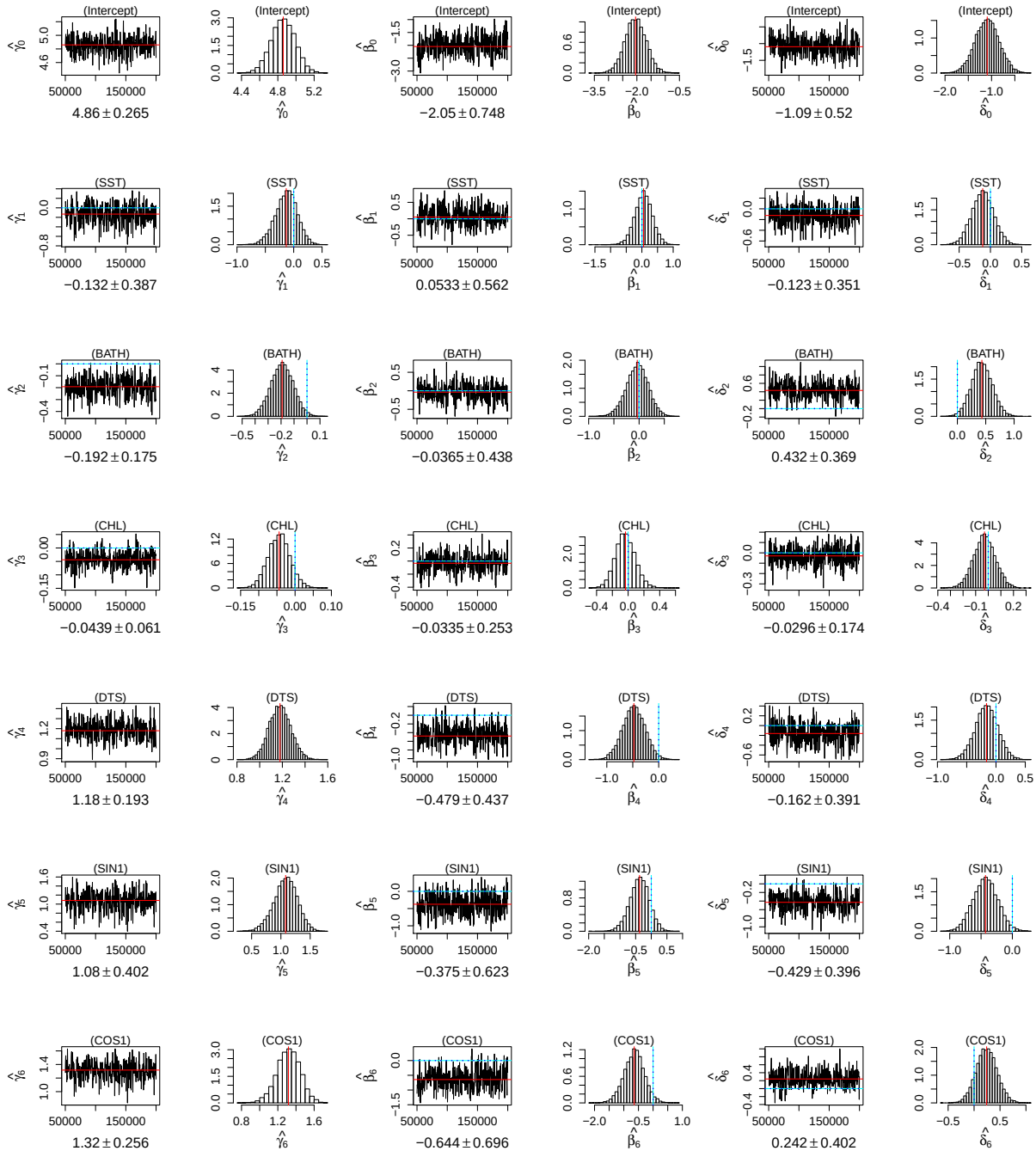

Figure 151: Common Tern: Traceplots and histograms of the posterior distributions of regression coefficients.

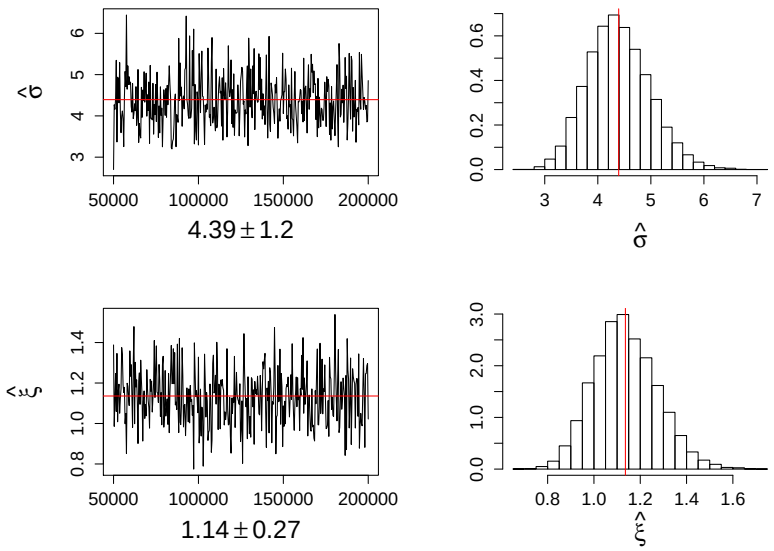

Figure 152: Common Tern: Traceplots and histograms of the posterior distributions of GPD parameters.

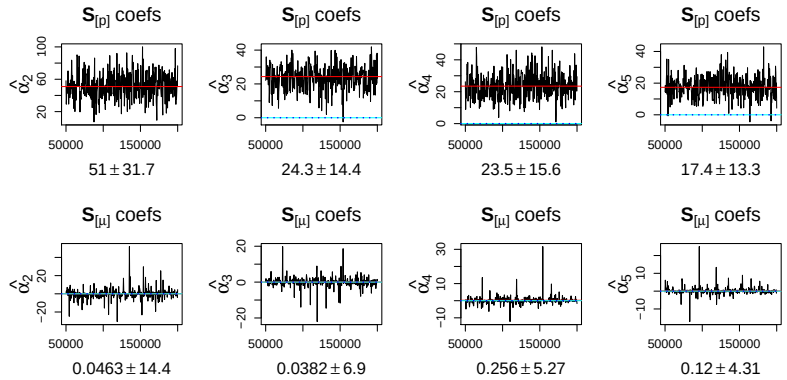

Figure 153: Common Tern: Traceplots and histograms of the posterior distributions of α coefficients.

B.8 Cory's Shearwater

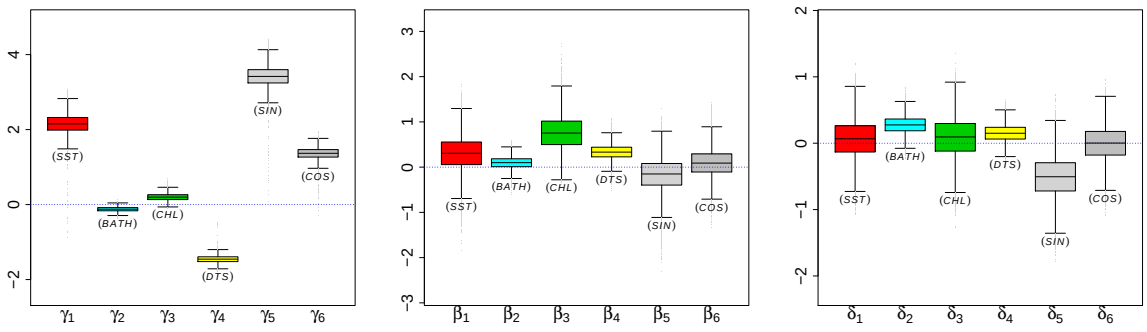

Figure 154: Cory's Shearwater: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	4.8365 (0.1548)	−1.7005 (0.2713)	−2.09507 (0.29452)
$\mathbf{x}_1$	Sea surface temp.	2.1517 (0.2672)	0.2897 (0.4084)	0.06501 (0.29595)
$\mathbf{x}_2$	Bathymetry	−0.1245 (0.0617)	0.0952 (0.1332)	0.27751 (0.13175)
$\mathbf{x}_3$	Chlorophyll	0.2006 (0.0965)	0.7699 (0.3848)	0.08684 (0.31286)
$\mathbf{x}_4$	Distance to shore	−1.4565 (0.0994)	0.3358 (0.1634)	0.15121 (0.13167)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	3.4191 (0.2817)	−0.1822 (0.4080)	−0.50951 (0.31479)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	1.3656 (0.1574)	0.0926 (0.3127)	−0.00296 (0.26718)

Table 12: Cory's Shearwater: Posterior summary of regression coefficients.

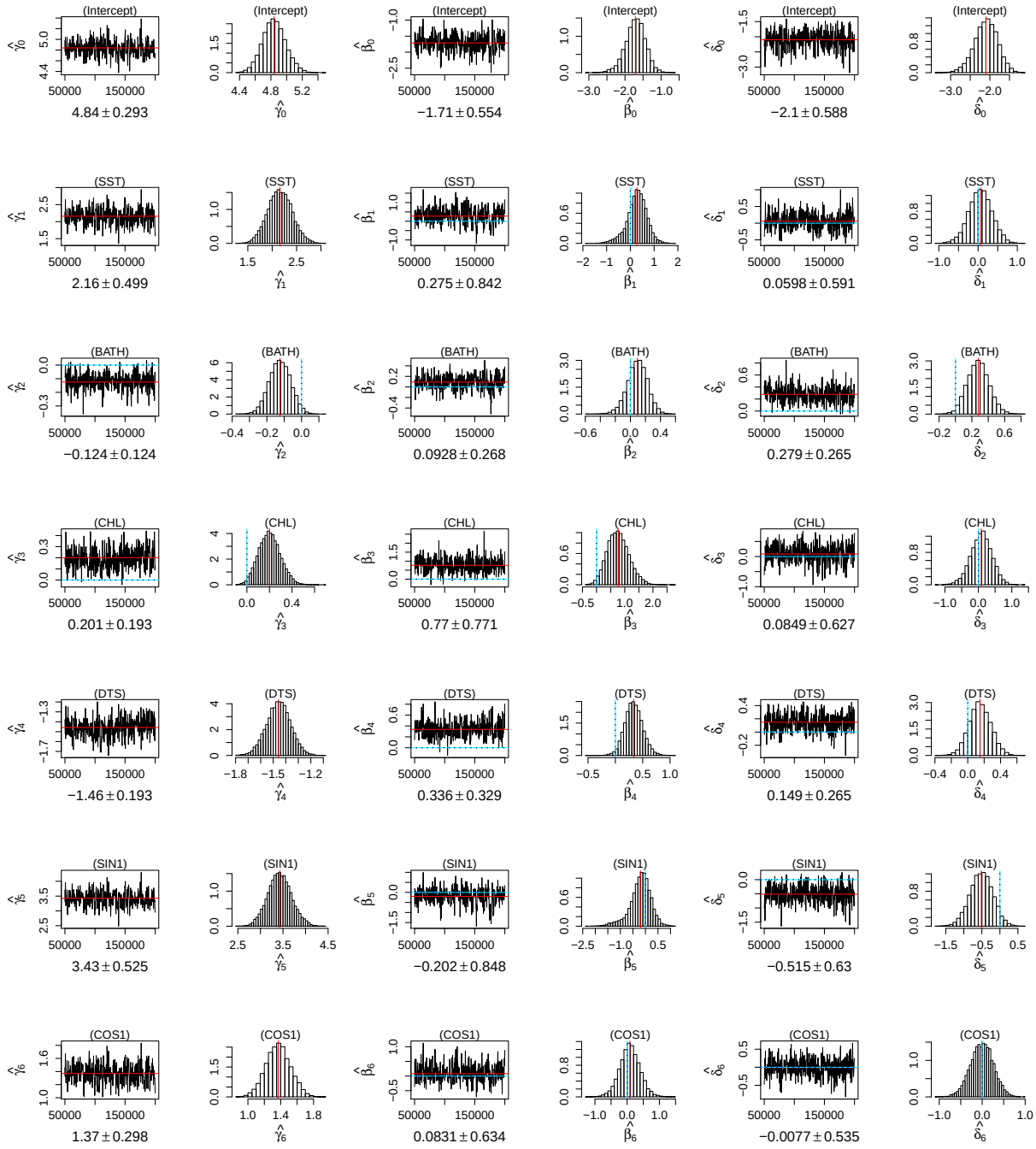

Figure 155: Cory's Shearwater: Traceplots and histograms of the posterior distributions of regression coefficients.

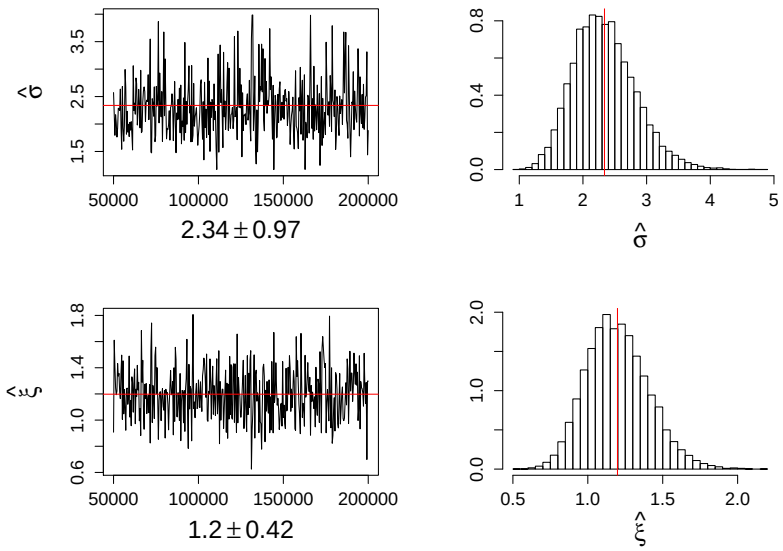

Figure 156: Cory's Shearwater: Traceplots and histograms of the posterior distributions of GPD parameters.

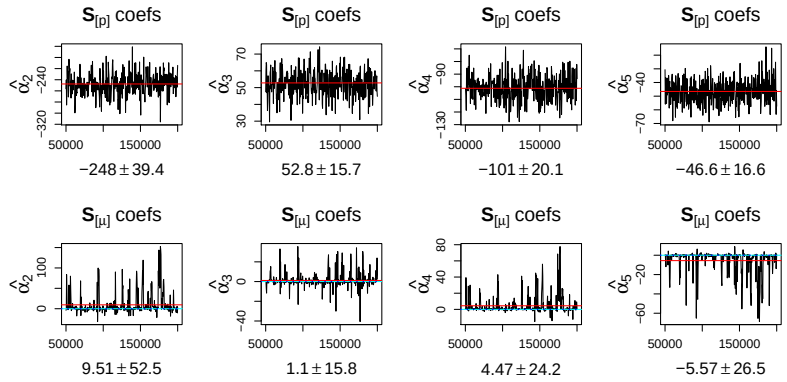

Figure 157: Cory's Shearwater: Traceplots and histograms of the posterior distributions of α coefficients.

B.9 Double-crested Cormorant

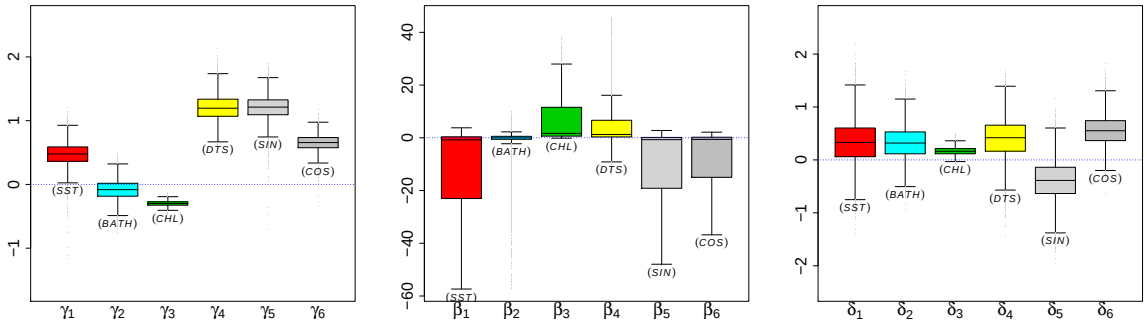

Figure 158: Double-crested Cormorant: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	5.9798 (0.2073)	− 3.231 (4.918)	−1.3324 (0.4252)
$\mathbf{x}_1$	Sea surface temp.	0.4724 (0.1756)	−10.095 (14.882)	0.3348 (0.4038)
$\mathbf{x}_2$	Bathymetry	−0.0847 (0.1513)	− 0.811 (5.130)	0.3232 (0.3058)
$\mathbf{x}_3$	Chlorophyll	−0.2991 (0.0400)	5.899 (7.239)	0.1668 (0.0734)
$\mathbf{x}_4$	Distance to shore	1.2060 (0.1986)	4.244 (6.758)	0.4026 (0.3710)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	1.2067 (0.1808)	− 8.556 (12.553)	−0.3901 (0.3653)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	0.6546 (0.1234)	− 6.598 (9.610)	0.5542 (0.2800)

Table 13: Double-crested Cormorant: Posterior summary of regression coefficients.

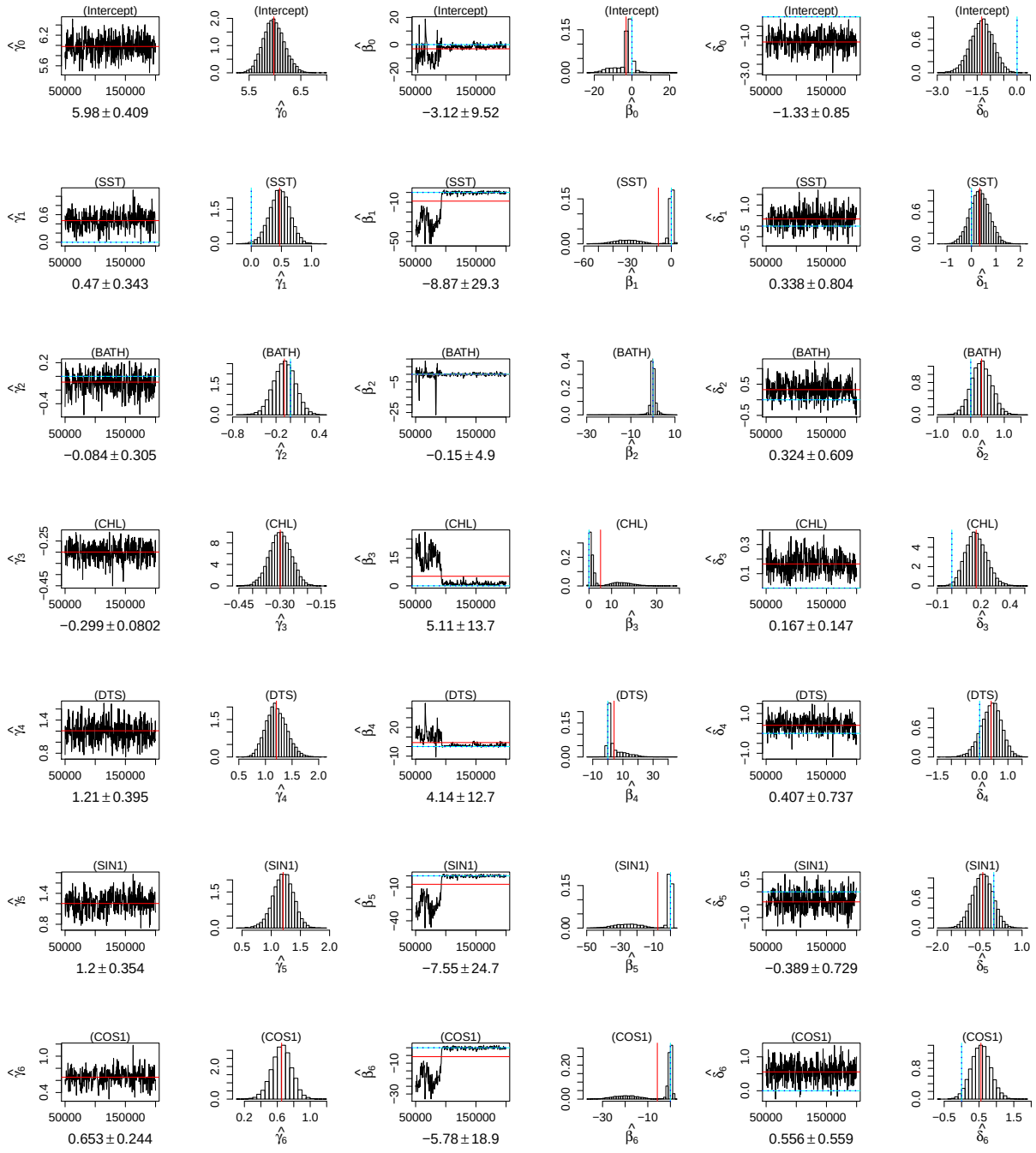

Figure 159: Double-crested Cormorant: Traceplots and histograms of the posterior distributions of regression coefficients.

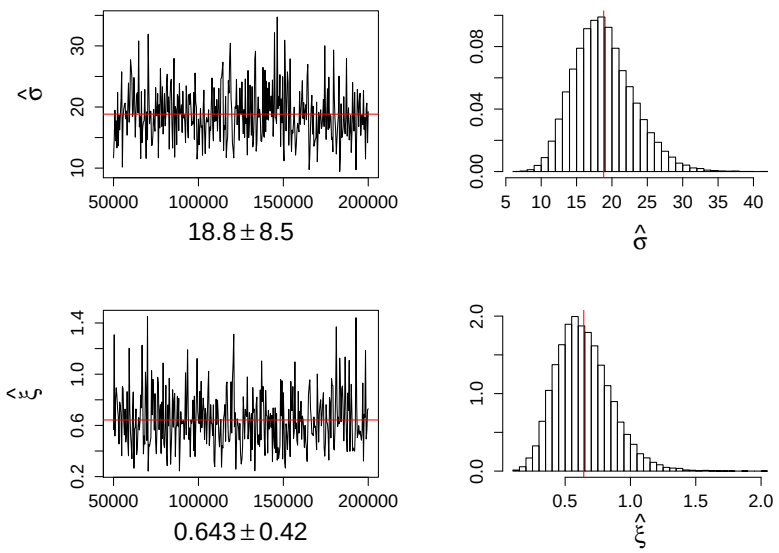

Figure 160: Double-crested Cormorant: Traceplots and histograms of the posterior distributions of GPD parameters.

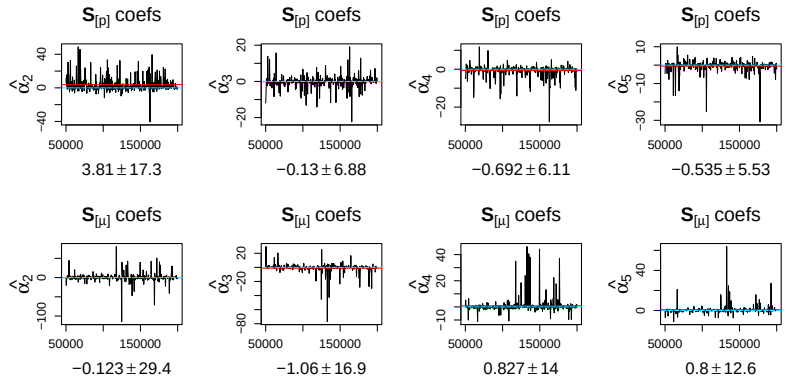

Figure 161: Double-crested Cormorant: Traceplots and histograms of the posterior distributions of α coefficients.

B.10 Dovekie

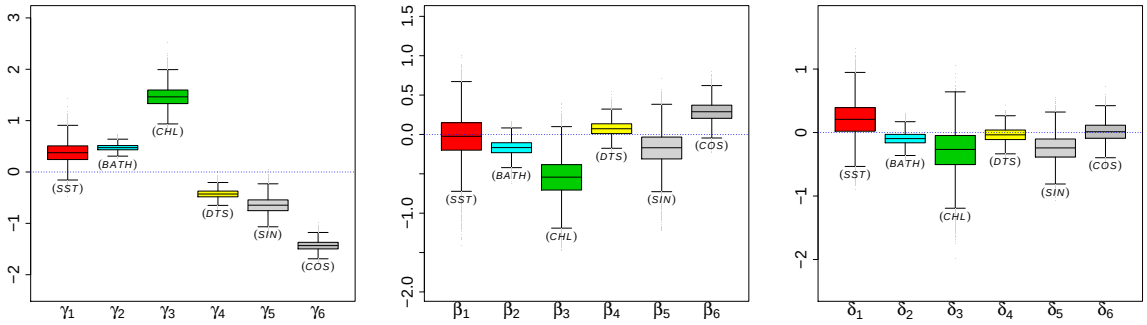

Figure 162: Dovekie: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	3.8384 (0.1077)	−0.6962 (0.1685)	−0.8811 (0.2097)
$\mathbf{x}_1$	Sea surface temp.	0.3763 (0.1979)	−0.0325 (0.2736)	0.2057 (0.2787)
$\mathbf{x}_2$	Bathymetry	0.4741 (0.0606)	−0.1697 (0.0943)	−0.0952 (0.0985)
$\mathbf{x}_3$	Chlorophyll	1.4661 (0.1957)	−0.5467 (0.2412)	−0.2871 (0.3405)
$\mathbf{x}_4$	Distance to shore	−0.4266 (0.0839)	0.0740 (0.0945)	−0.0357 (0.1103)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	−0.6467 (0.1546)	−0.1769 (0.2148)	−0.2441 (0.2130)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	−1.4333 (0.0959)	0.2873 (0.1264)	0.0130 (0.1517)

Table 14: Dovekie: Posterior summary of regression coefficients.

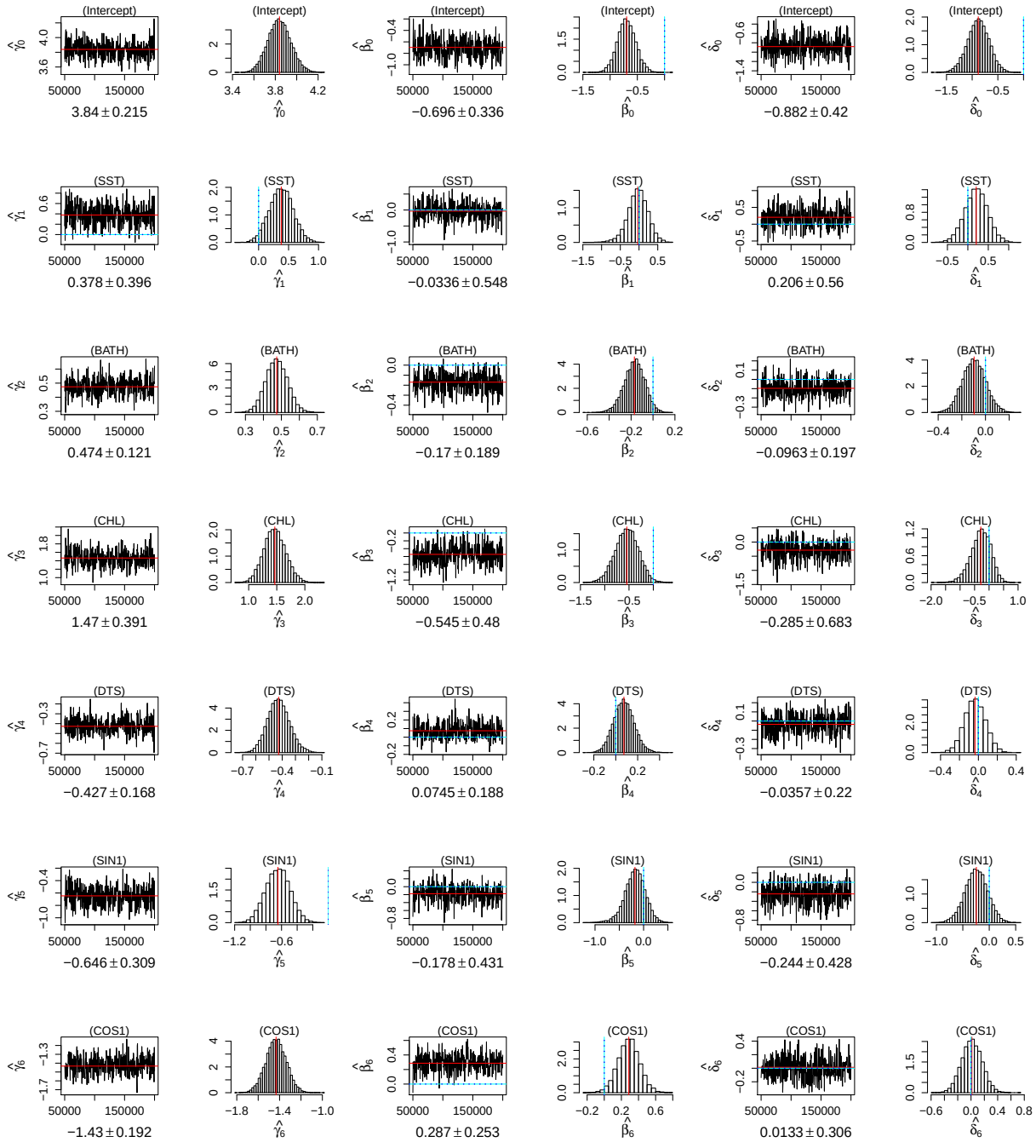

Figure 163: Dovekie: Traceplots and histograms of the posterior distributions of regression coefficients.

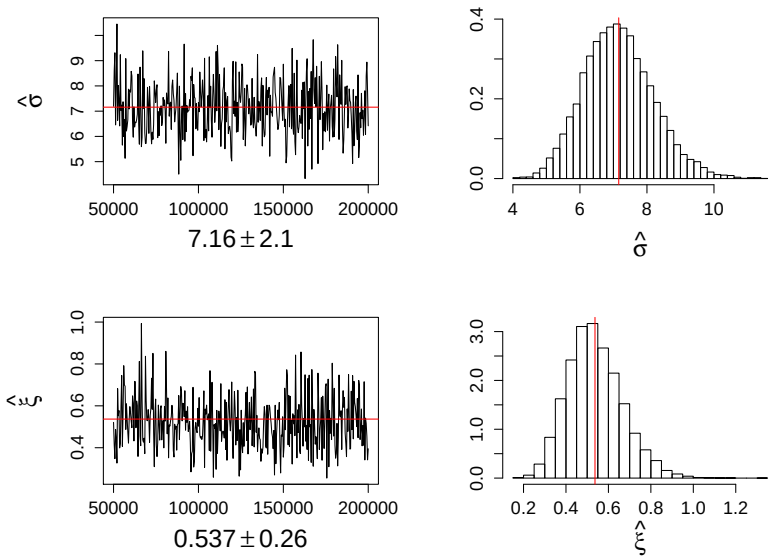

Figure 164: Dovekie: Traceplots and histograms of the posterior distributions of GPD parameters.

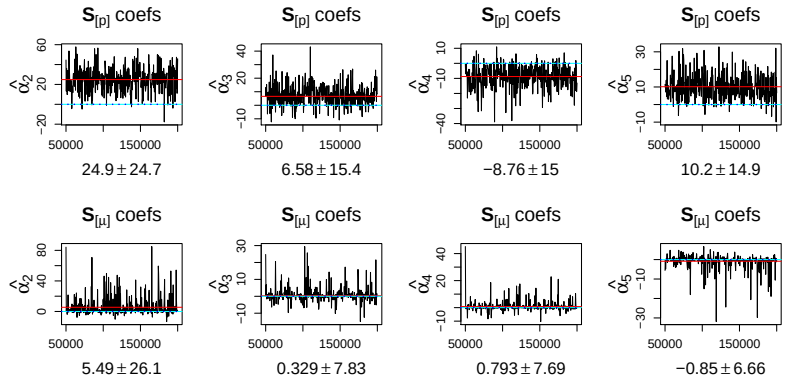

Figure 165: Dovekie: Traceplots and histograms of the posterior distributions of α coefficients.

B.11 Great Black-backed Gull

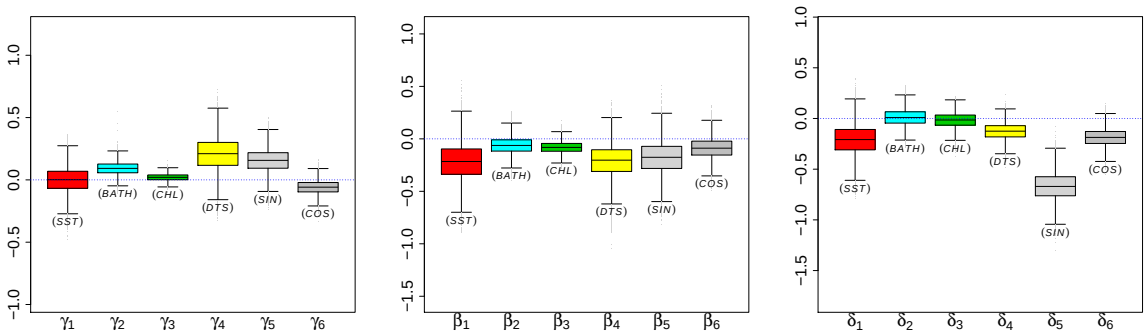

Figure 166: Great Black-backed Gull: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	1.792484 (0.047306)	−1.4222 (0.0809)	−2.46331 (0.11367)
$\mathbf{x}_1$	Sea surface temp.	0.000553 (0.101452)	−0.2168 (0.1813)	−0.20837 (0.15020)
$\mathbf{x}_2$	Bathymetry	0.091939 (0.052646)	−0.0627 (0.0806)	0.00978 (0.08271)
$\mathbf{x}_3$	Chlorophyll	0.020854 (0.028638)	−0.0802 (0.0561)	−0.01722 (0.07419)
$\mathbf{x}_4$	Distance to shore	0.207011 (0.136627)	−0.2124 (0.1602)	−0.12658 (0.08221)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	0.155596 (0.092165)	−0.1763 (0.1588)	−0.66940 (0.13765)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	−0.058917 (0.055651)	−0.0881 (0.0999)	−0.18700 (0.08757)

Table 15: Great Black-backed Gull: Posterior summary of regression coefficients.

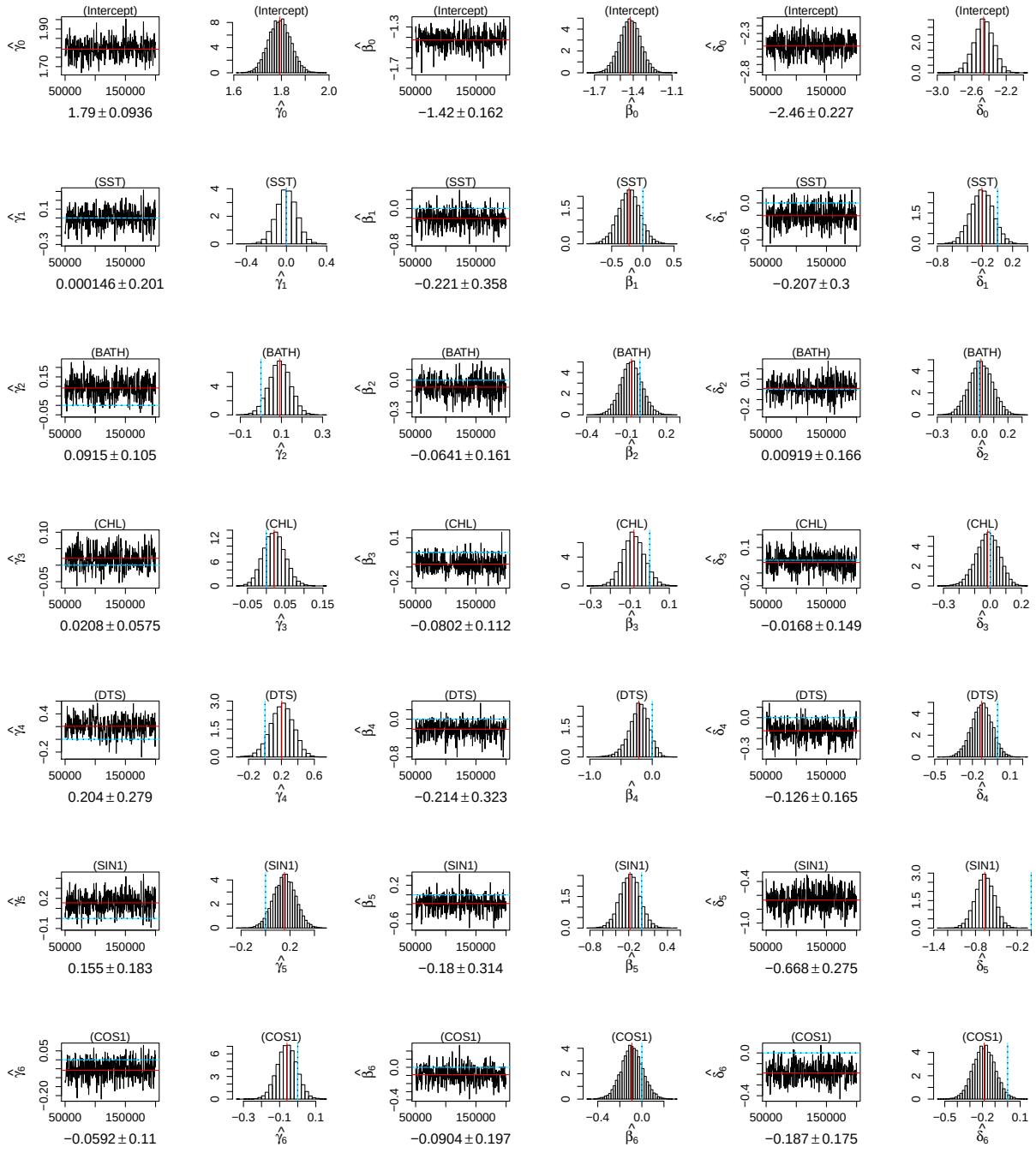

Figure 167: Great Black-backed Gull: Traceplots and histograms of the posterior distributions of regression coefficients.

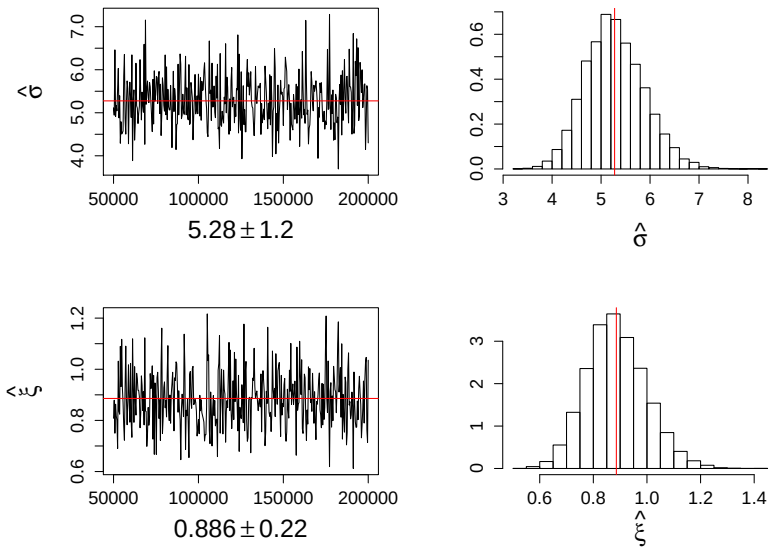

Figure 168: Great Black-backed Gull: Traceplots and histograms of the posterior distributions of GPD parameters.

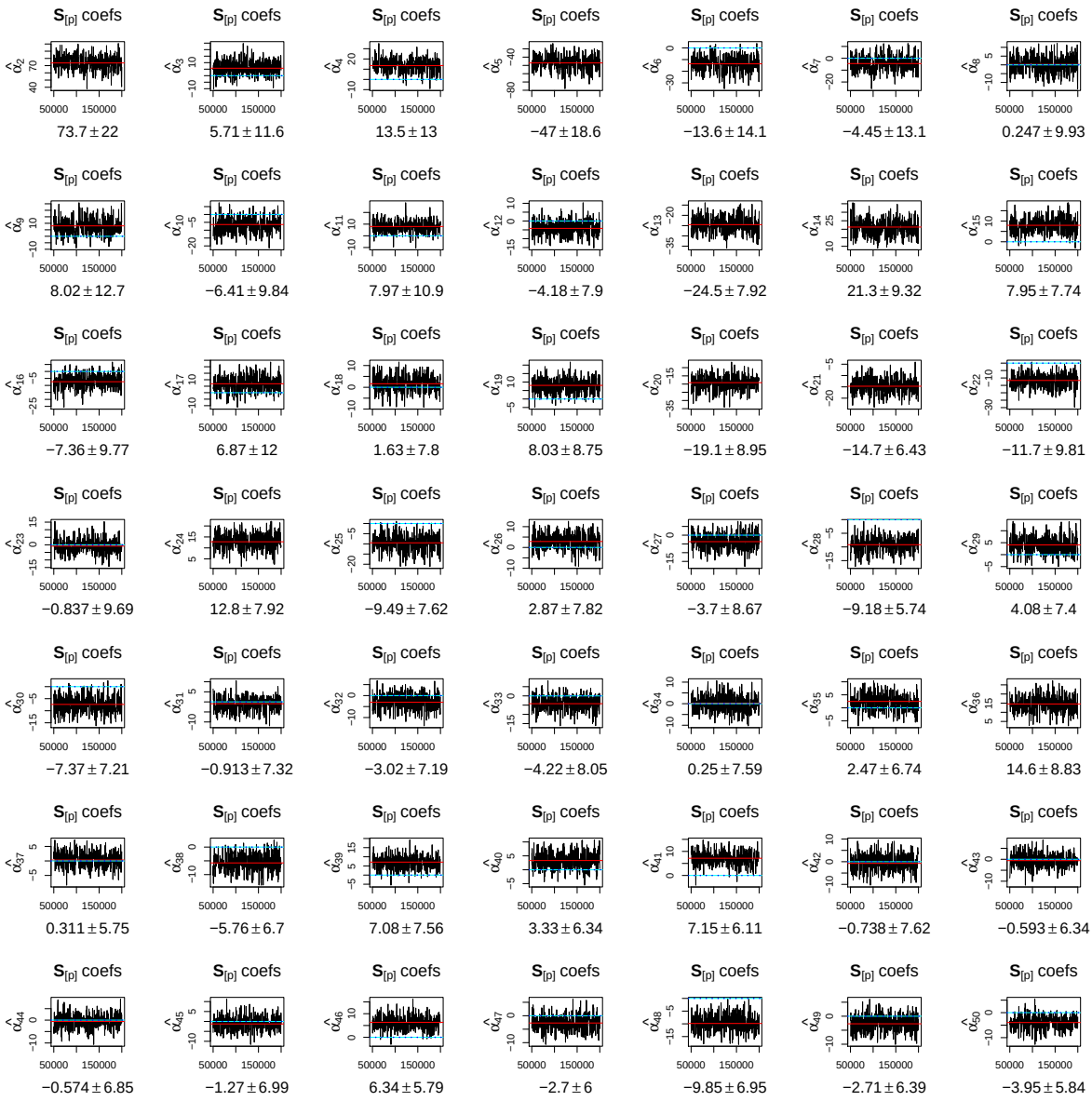

Figure 169: Great Black-backed Gull: Traceplots and histograms of the posterior distributions of α coefficients in the spatial regression of **p**.

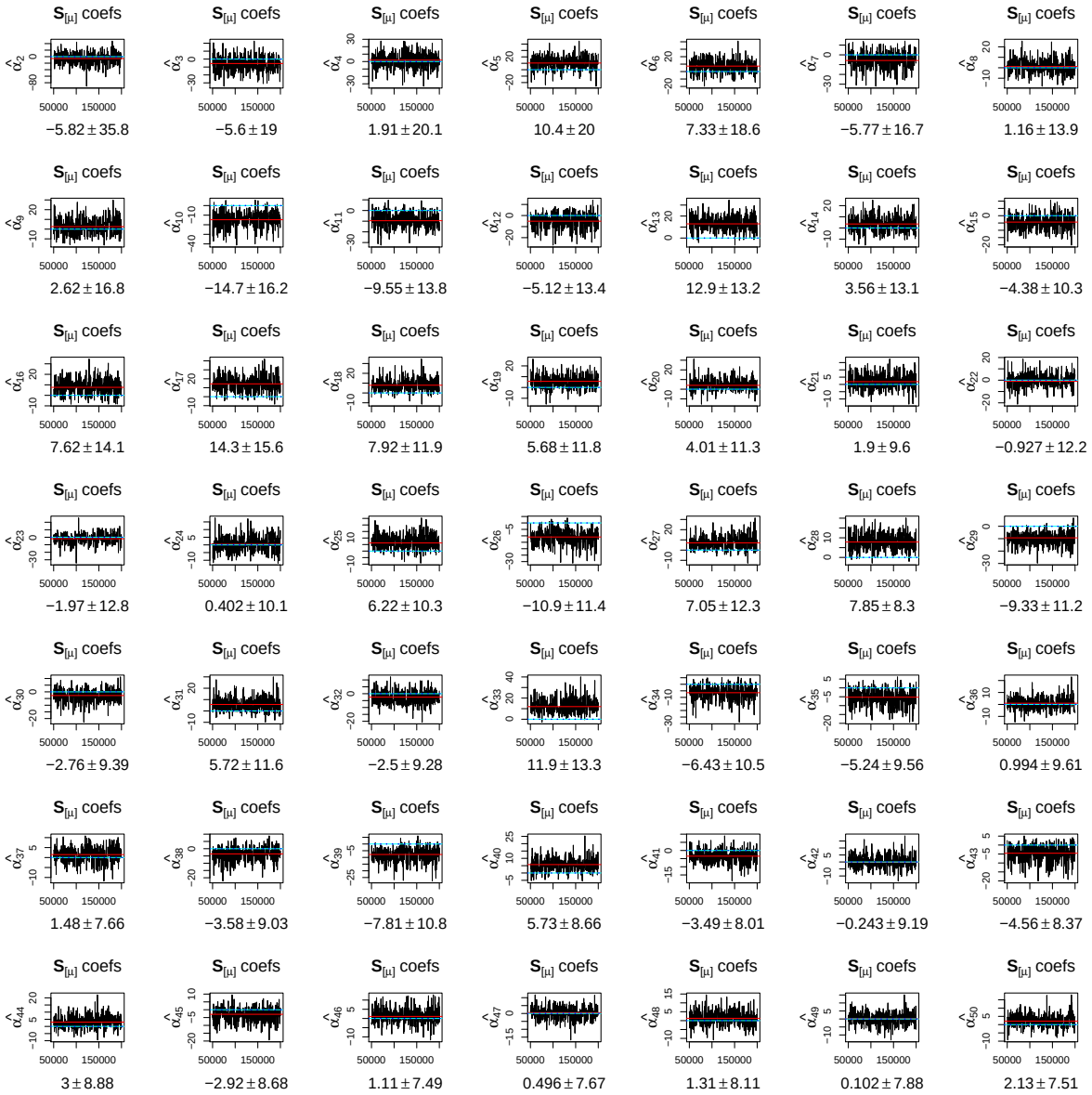

Figure 170: Great Black-backed Gull: Traceplots and histograms of the posterior distributions of α coefficients in the spatial regression of $\mathbf{m}$.

B.12 Great Shearwater

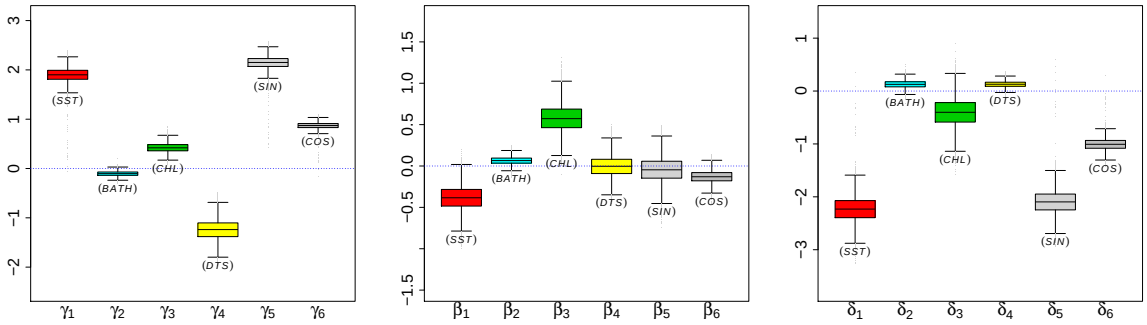

Figure 171: Great Shearwater: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	2.2633 (0.0754)	−0.17217 (0.09540)	−3.7825 (0.2192)
$\mathbf{x}_1$	Sea surface temp.	1.8978 (0.1469)	−0.38429 (0.14877)	−2.2328 (0.2446)
$\mathbf{x}_2$	Bathymetry	−0.1042 (0.0504)	0.06521 (0.04568)	0.1269 (0.0714)
$\mathbf{x}_3$	Chlorophyll	0.4222 (0.0933)	0.57833 (0.16789)	−0.4051 (0.2715)
$\mathbf{x}_4$	Distance to shore	−1.2425 (0.2053)	−0.00507 (0.12979)	0.1283 (0.0578)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	2.1475 (0.1299)	−0.04532 (0.15193)	−2.0958 (0.2311)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	0.8704 (0.0660)	−0.12953 (0.07314)	−1.0079 (0.1121)

Table 16: Great Shearwater: Posterior summary of regression coefficients.

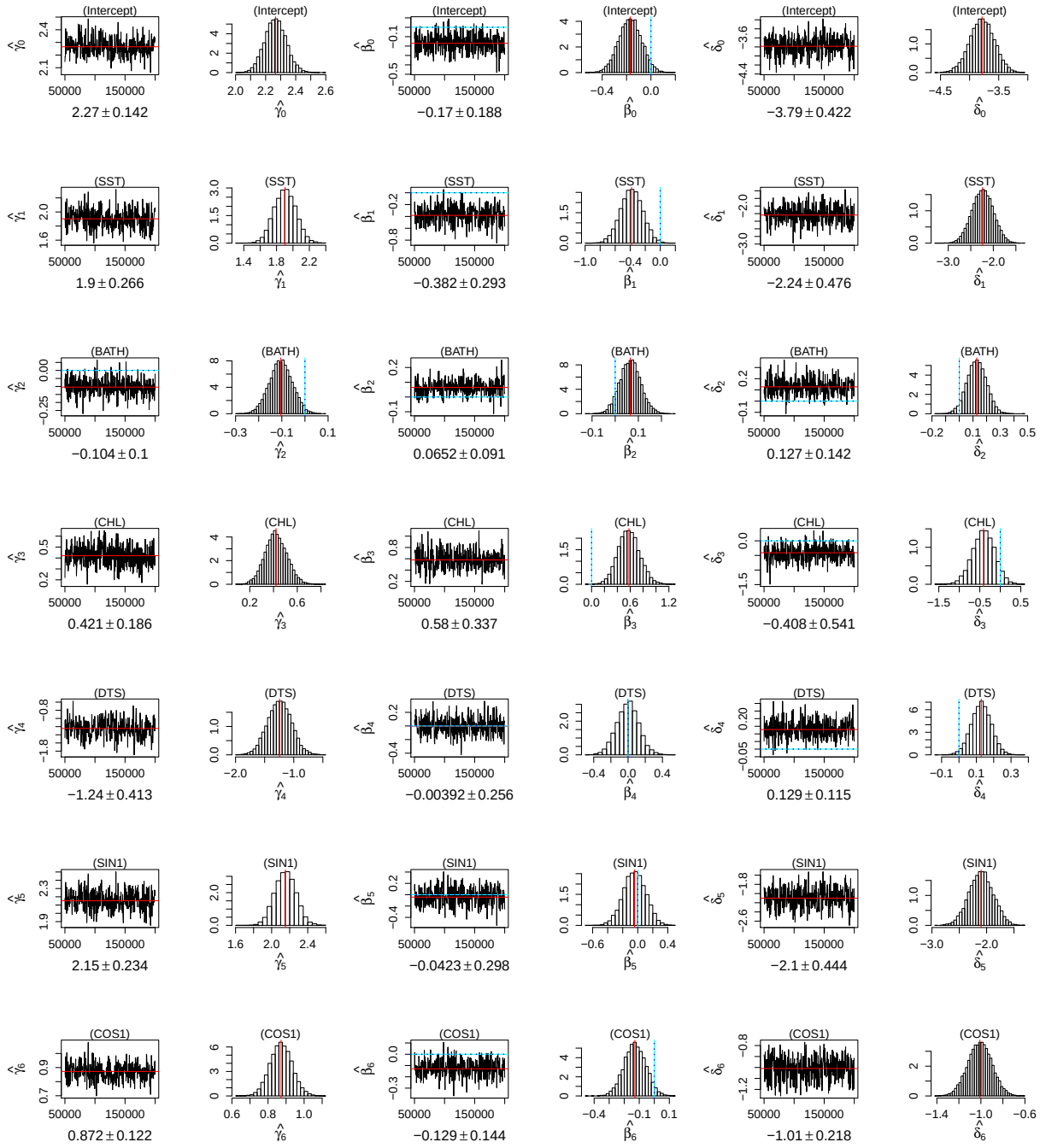

Figure 172: Great Shearwater: Traceplots and histograms of the posterior distributions of regression coefficients.

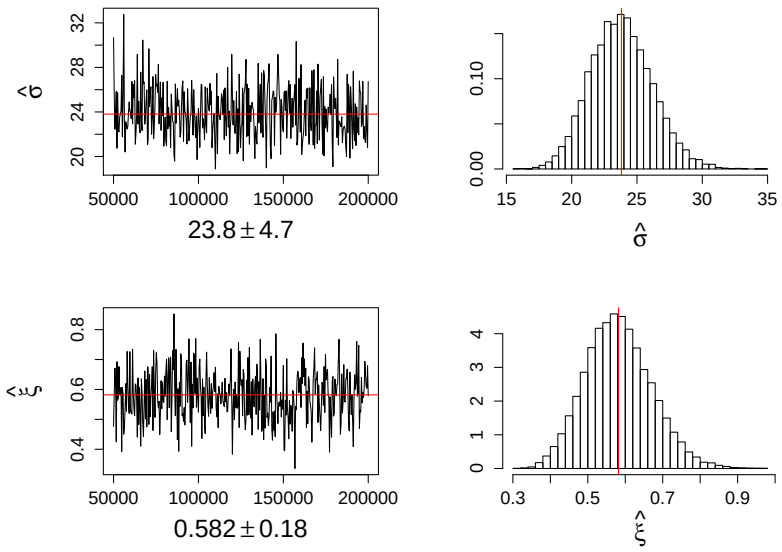

Figure 173: Great Shearwater: Traceplots and histograms of the posterior distributions of GPD parameters.

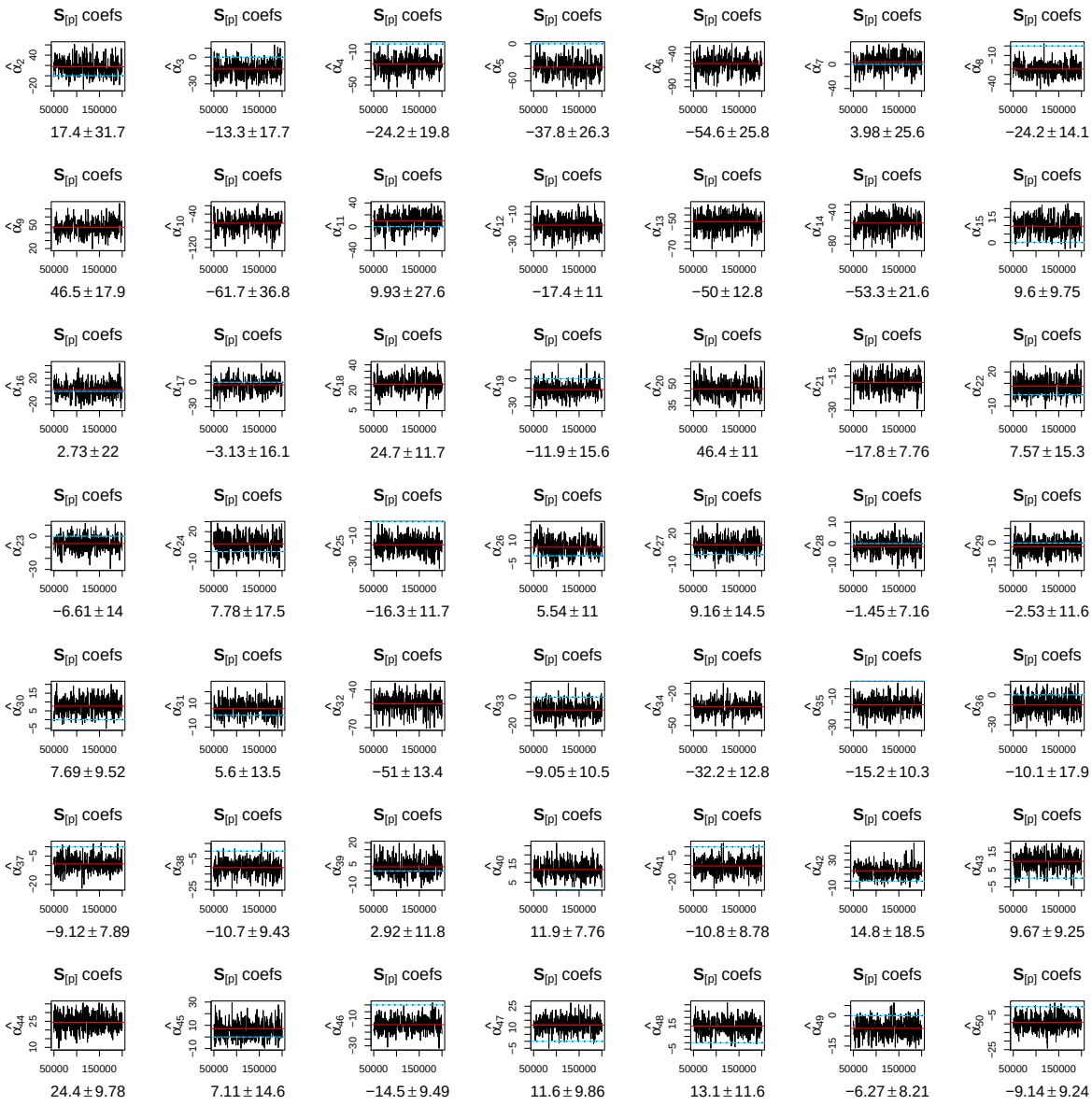

Figure 174: Great Shearwater: Traceplots and histograms of the posterior distributions of α coefficients in the spatial regression of **p**.

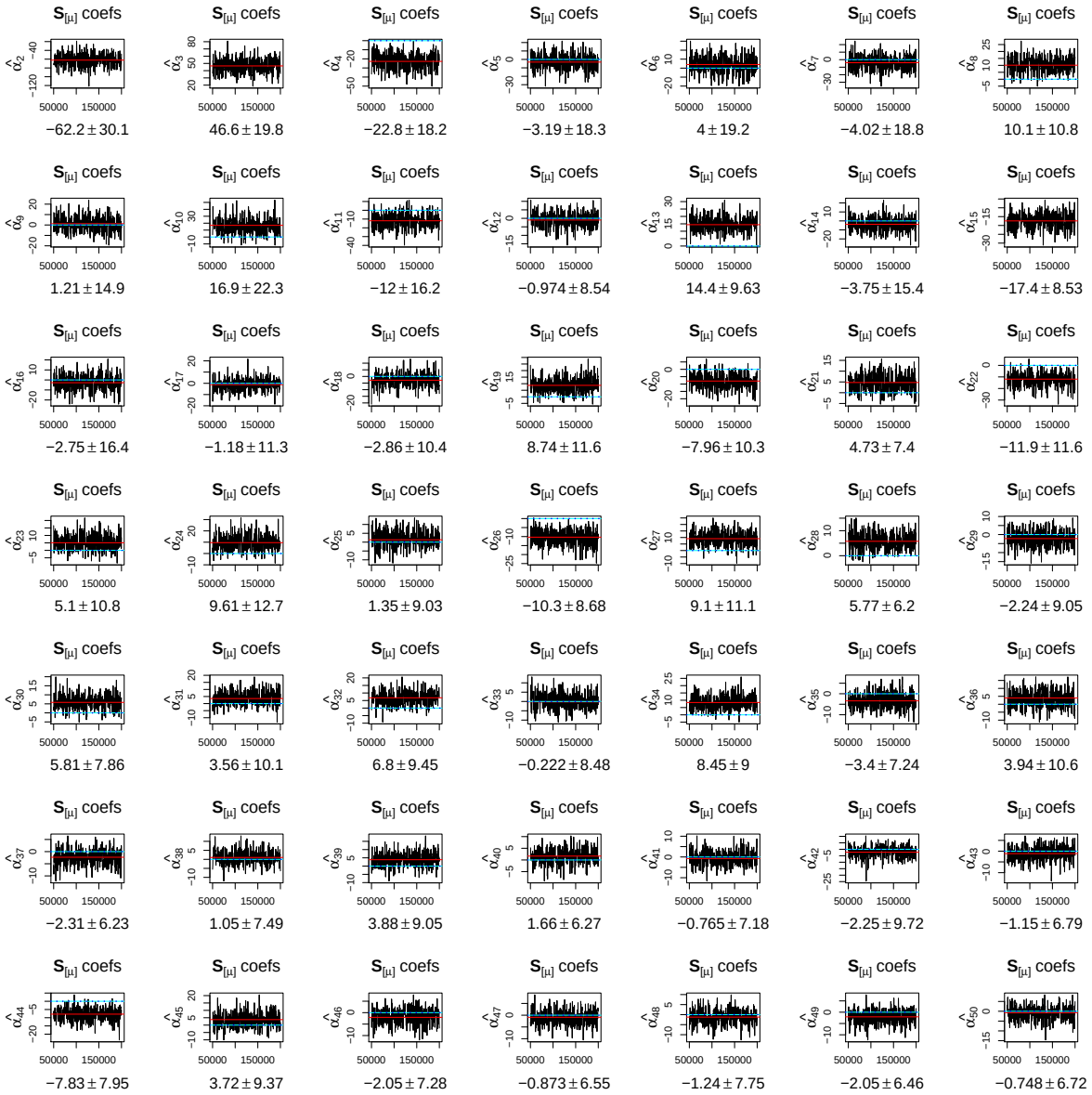

Figure 175: Great Shearwater: Traceplots and histograms of the posterior distributions of α coefficients in the spatial regression of $\mathbf{m}$.

B.13 Herring Gull

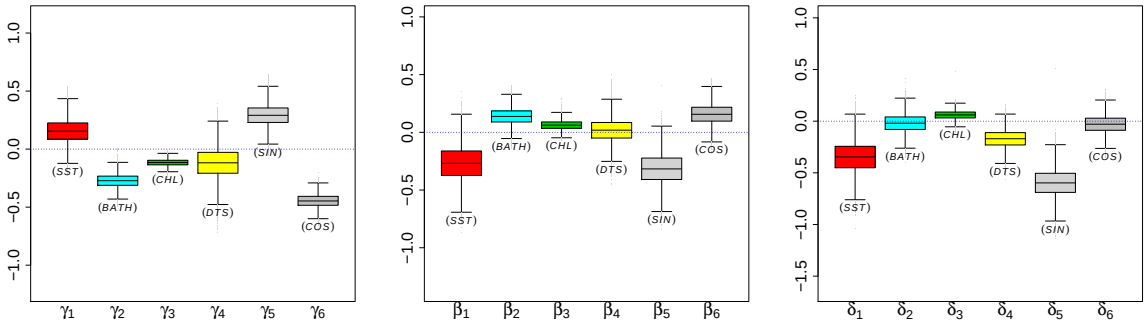

Figure 176: Herring Gull: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	1.8257 (0.0468)	−1.3691 (0.0708)	−2.5146 (0.1179)
$\mathbf{x}_1$	Sea surface temp.	0.1557 (0.1058)	−0.2685 (0.1549)	−0.3460 (0.1541)
$\mathbf{x}_2$	Bathymetry	−0.2725 (0.0589)	0.1370 (0.0708)	−0.0195 (0.0891)
$\mathbf{x}_3$	Chlorophyll	−0.1161 (0.0297)	0.0639 (0.0415)	0.0595 (0.0429)
$\mathbf{x}_4$	Distance to shore	−0.1197 (0.1391)	0.0175 (0.1027)	−0.1714 (0.0885)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	0.2922 (0.0953)	−0.3164 (0.1360)	−0.5962 (0.1381)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	−0.4460 (0.0578)	0.1565 (0.0883)	−0.0295 (0.0872)

Table 17: Herring Gull: Posterior summary of regression coefficients.

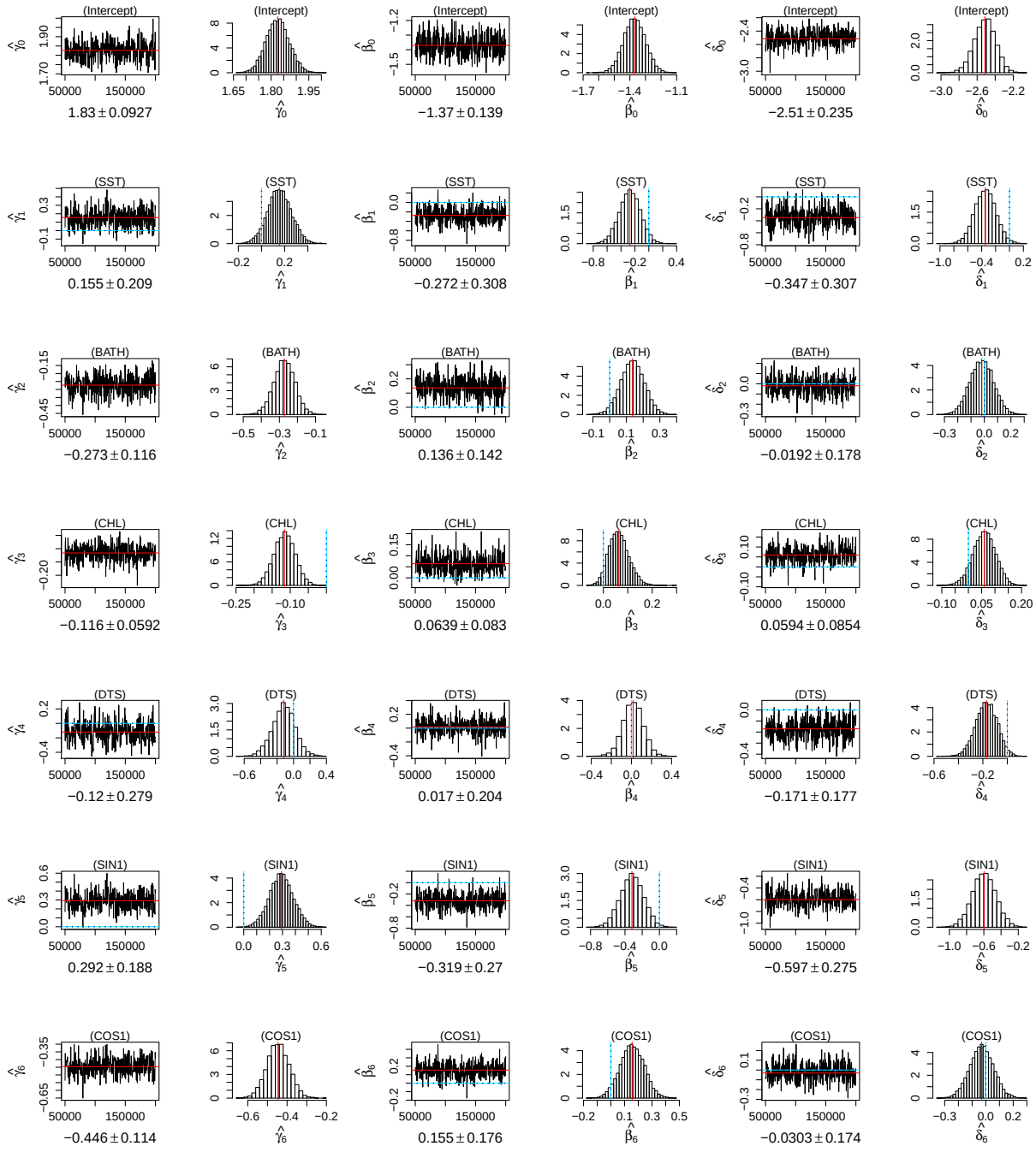

Figure 177: Herring Gull: Traceplots and histograms of the posterior distributions of regression coefficients.

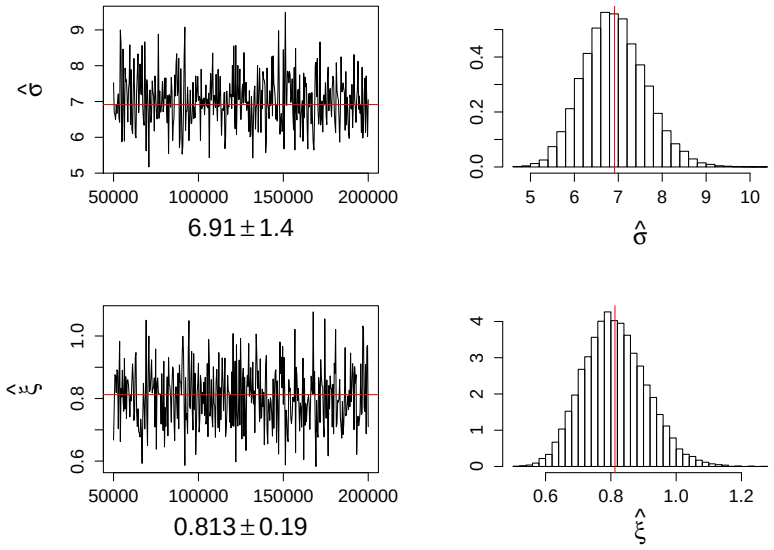

Figure 178: Herring Gull: Traceplots and histograms of the posterior distributions of GPD parameters.

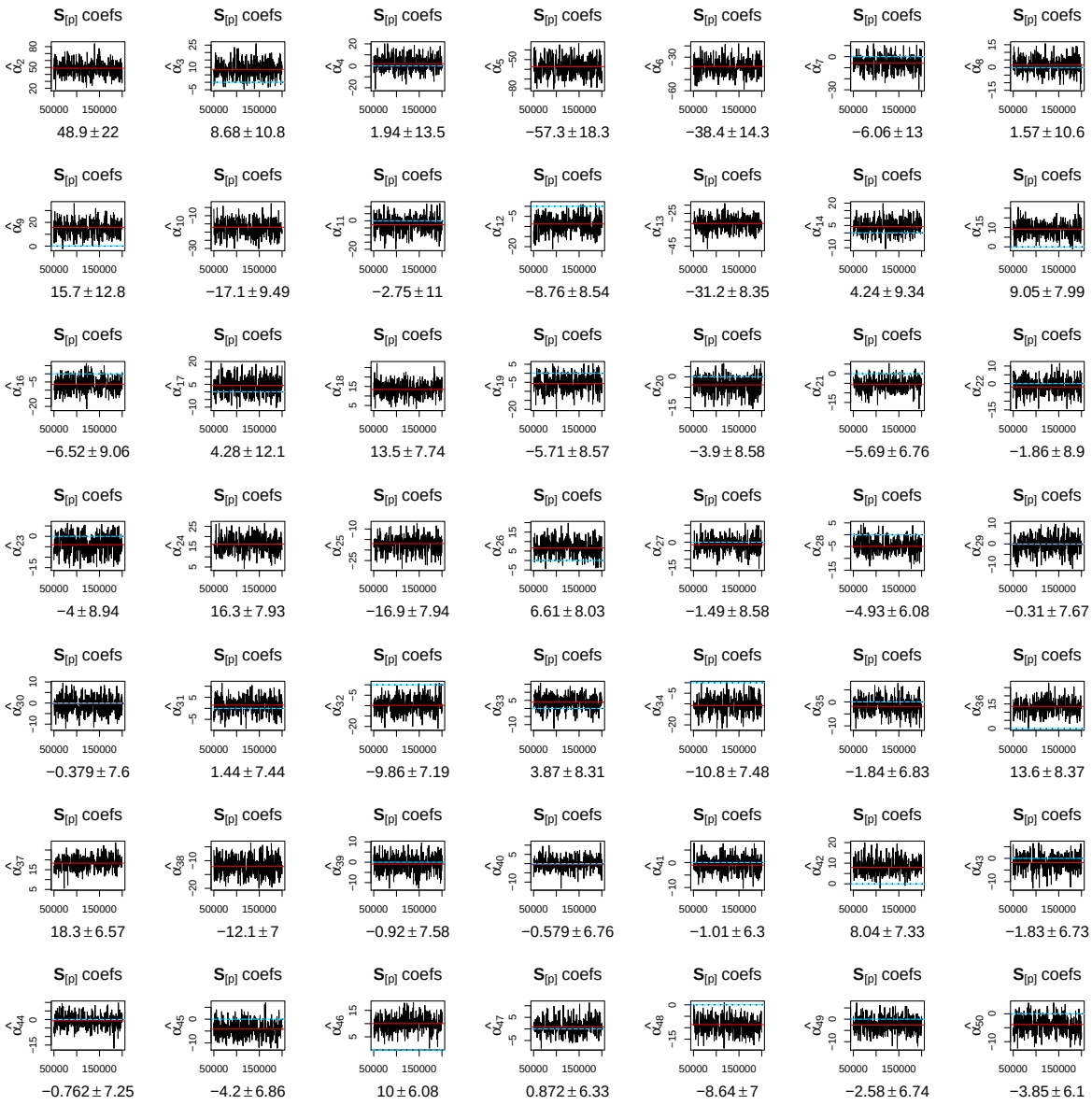

Figure 179: Herring Gull: Traceplots and histograms of the posterior distributions of α coefficients in the spatial regression of **p**.

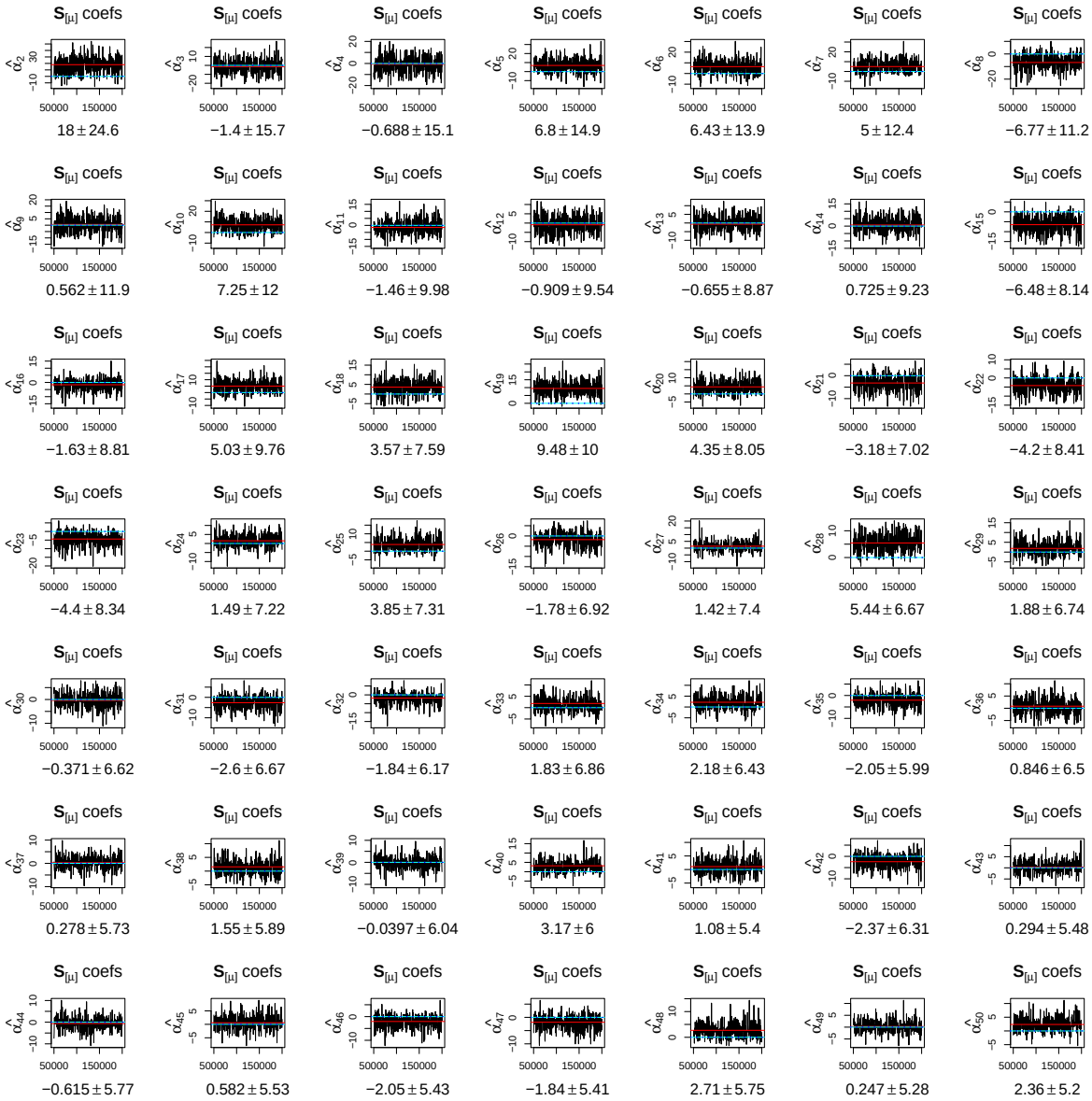

Figure 180: Herring Gull: Traceplots and histograms of the posterior distributions of α coefficients in the spatial regression of $\mathbf{m}$.

B.14 Laughing Gull

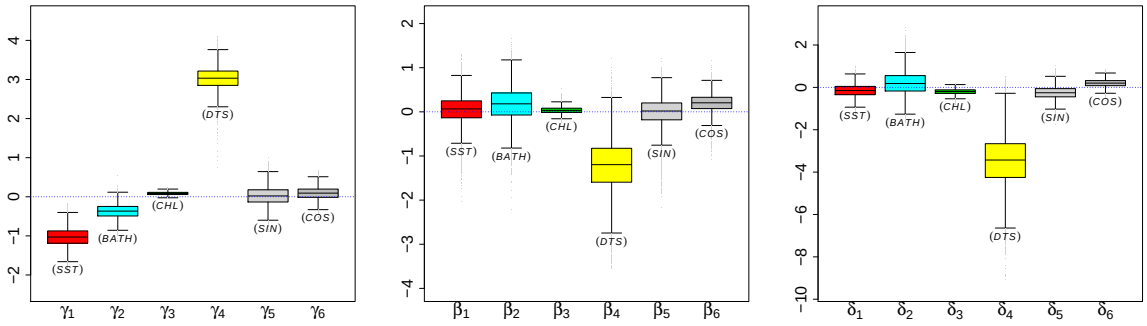

Figure 181: Laughing Gull: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	6.4414 (0.2844)	−3.00498 (0.60006)	−5.926 (1.293)
$\mathbf{x}_1$	Sea surface temp.	−1.0283 (0.2325)	0.04454 (0.31316)	−0.152 (0.292)
$\mathbf{x}_2$	Bathymetry	−0.3703 (0.1808)	0.17107 (0.38078)	0.203 (0.540)
$\mathbf{x}_3$	Chlorophyll	0.0860 (0.0413)	0.03711 (0.07450)	−0.204 (0.127)
$\mathbf{x}_4$	Distance to shore	3.0330 (0.2837)	−1.22633 (0.57915)	−3.471 (1.205)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	0.0228 (0.2305)	−0.00696 (0.32143)	−0.255 (0.287)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	0.0930 (0.1556)	0.19583 (0.20737)	0.200 (0.178)

Table 18: Laughing Gull: Posterior summary of regression coefficients.

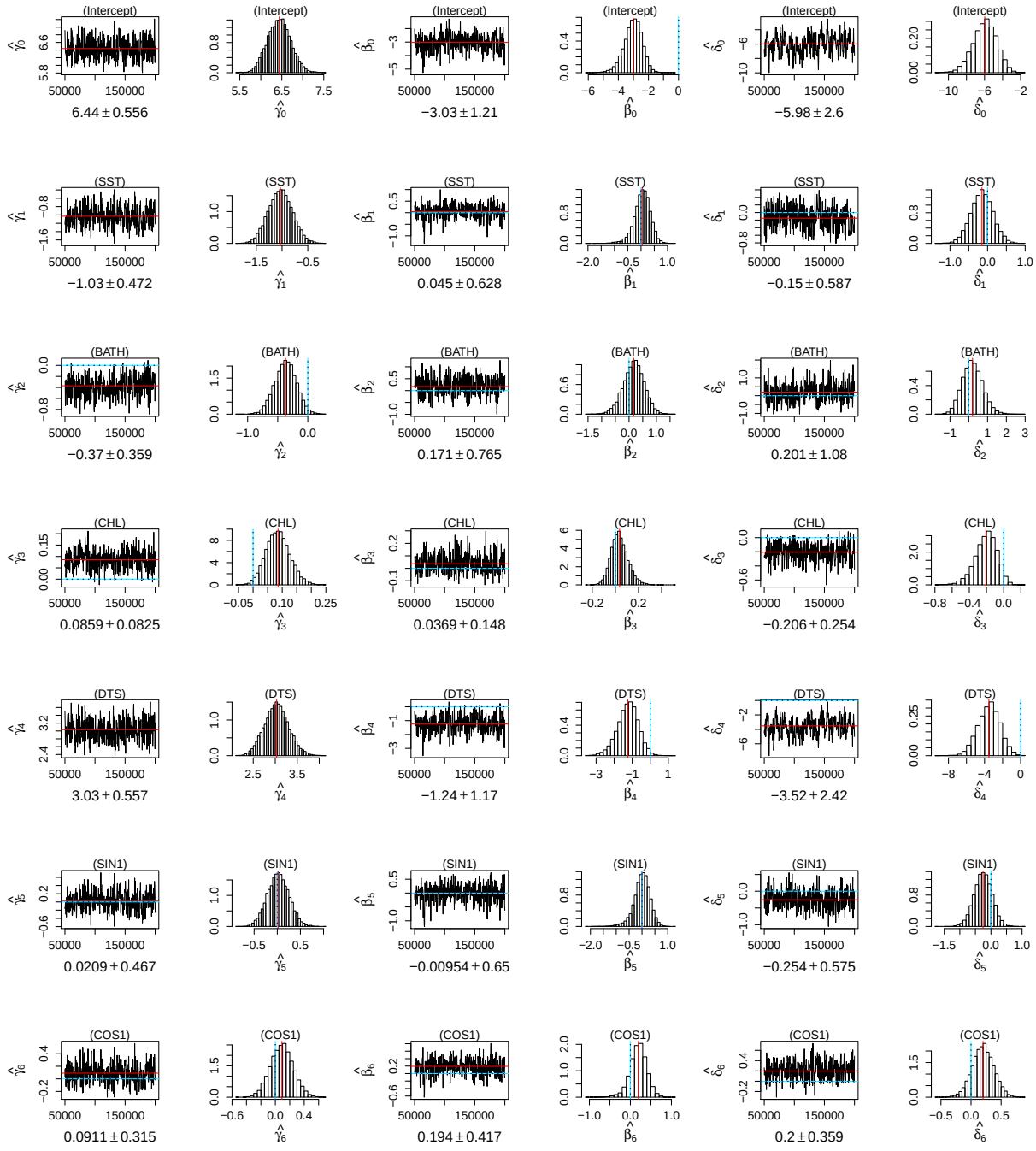

Figure 182: Laughing Gull: Traceplots and histograms of the posterior distributions of regression coefficients.

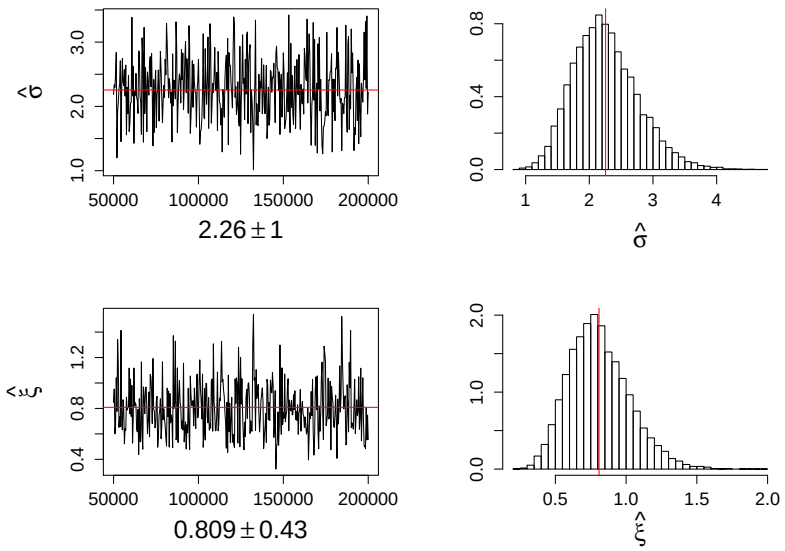

Figure 183: Laughing Gull: Traceplots and histograms of the posterior distributions of GPD parameters.

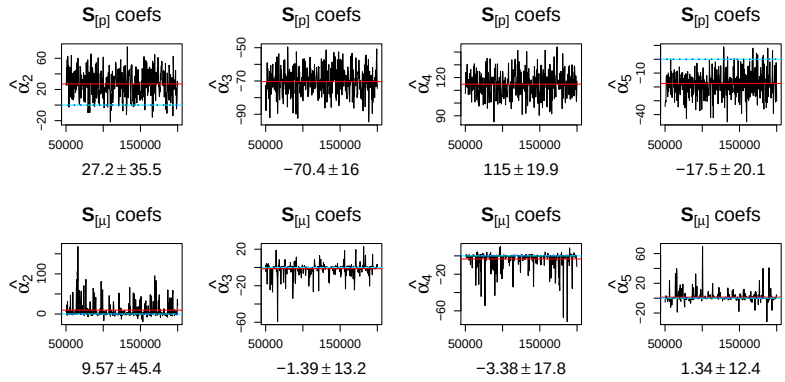

Figure 184: Laughing Gull: Traceplots and histograms of the posterior distributions of α coefficients.

B.15 Leach's Storm-petrel

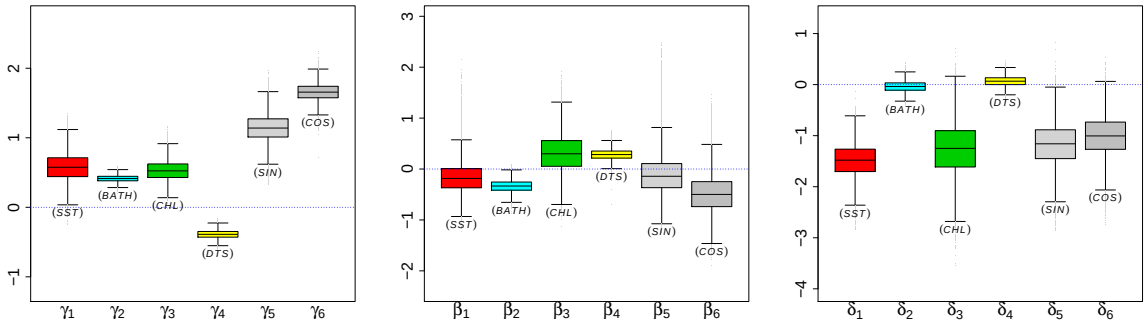

Figure 185: Leach's Storm-petrel: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	4.4069 (0.1278)	−2.0376 (0.4620)	−3.1322 (0.5753)
$\mathbf{x}_1$	Sea surface temp.	0.5782 (0.2033)	−0.1545 (0.3342)	−1.4889 (0.3309)
$\mathbf{x}_2$	Bathymetry	0.4133 (0.0482)	−0.3367 (0.1187)	−0.0367 (0.1059)
$\mathbf{x}_3$	Chlorophyll	0.5284 (0.1446)	0.3128 (0.3767)	−1.2674 (0.5213)
$\mathbf{x}_4$	Distance to shore	−0.3885 (0.0608)	0.2833 (0.1058)	0.0676 (0.1007)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	1.1432 (0.1976)	−0.0931 (0.4270)	−1.1751 (0.4207)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	1.6600 (0.1237)	−0.4769 (0.3894)	−1.0068 (0.3945)

Table 19: Leach's Storm-petrel: Posterior summary of regression coefficients.

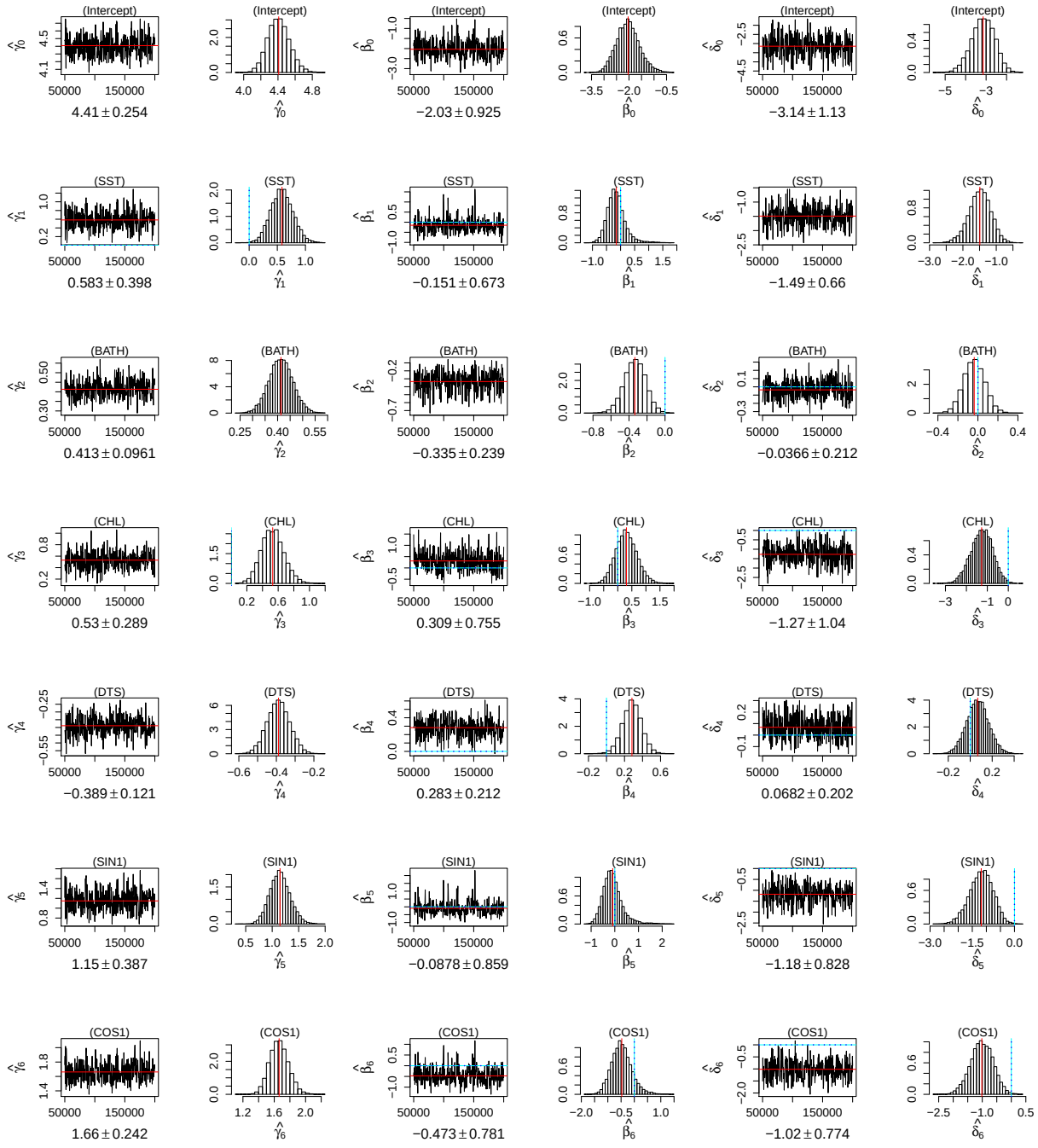

Figure 186: Leach's Storm-petrel: Traceplots and histograms of the posterior distributions of regression coefficients.

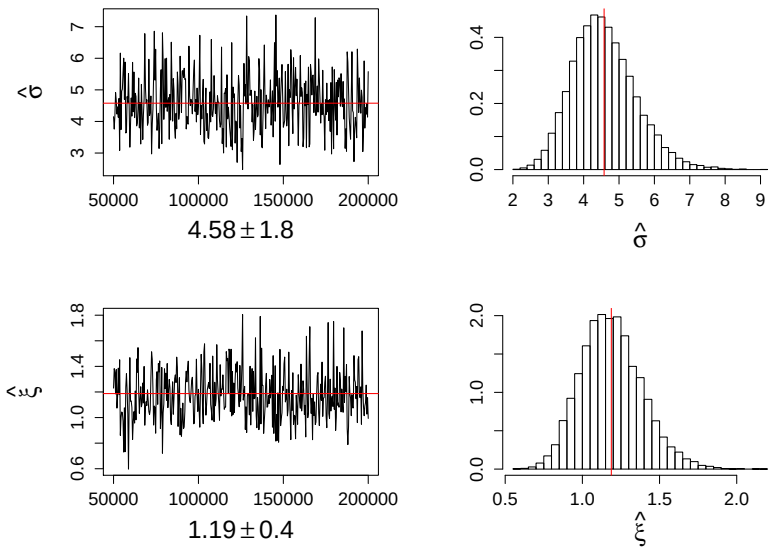

Figure 187: Leach's Storm-petrel: Traceplots and histograms of the posterior distributions of GPD parameters.

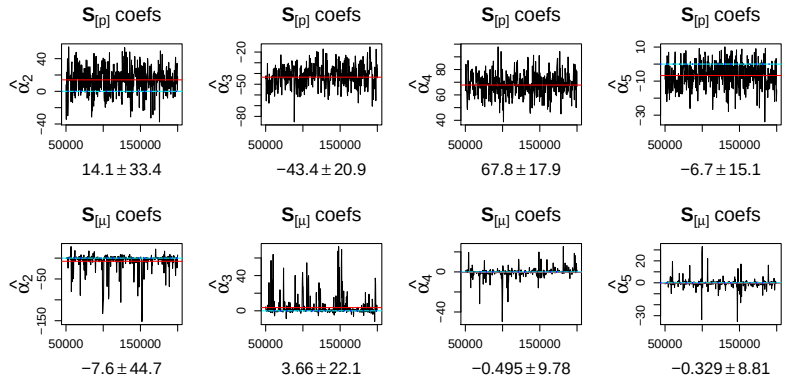

Figure 188: Leach's Storm-petrel: Traceplots and histograms of the posterior distributions of α coefficients.

B.16 Long-tailed Duck

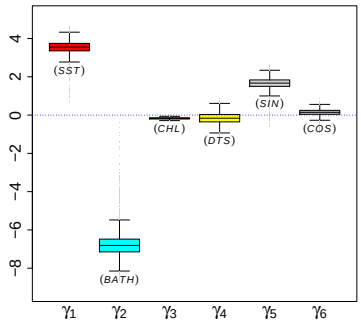

Figure 189: Long-tailed Duck: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\gamma}$
Predictor		mean (sd)
1	Intercept	12.7158 (0.5877)
$\mathbf{x}_1$	Sea surface temp.	3.5501 (0.3076)
$\mathbf{x}_2$	Bathymetry	− 6.7983 (0.5515)
$\mathbf{x}_3$	Chlorophyll	− 0.1777 (0.0425)
$\mathbf{x}_4$	Distance to shore	− 0.1655 (0.2878)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	1.6684 (0.2620)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	0.1421 (0.1577)

Table 20: Long-tailed Duck: Posterior summary of regression coefficients.

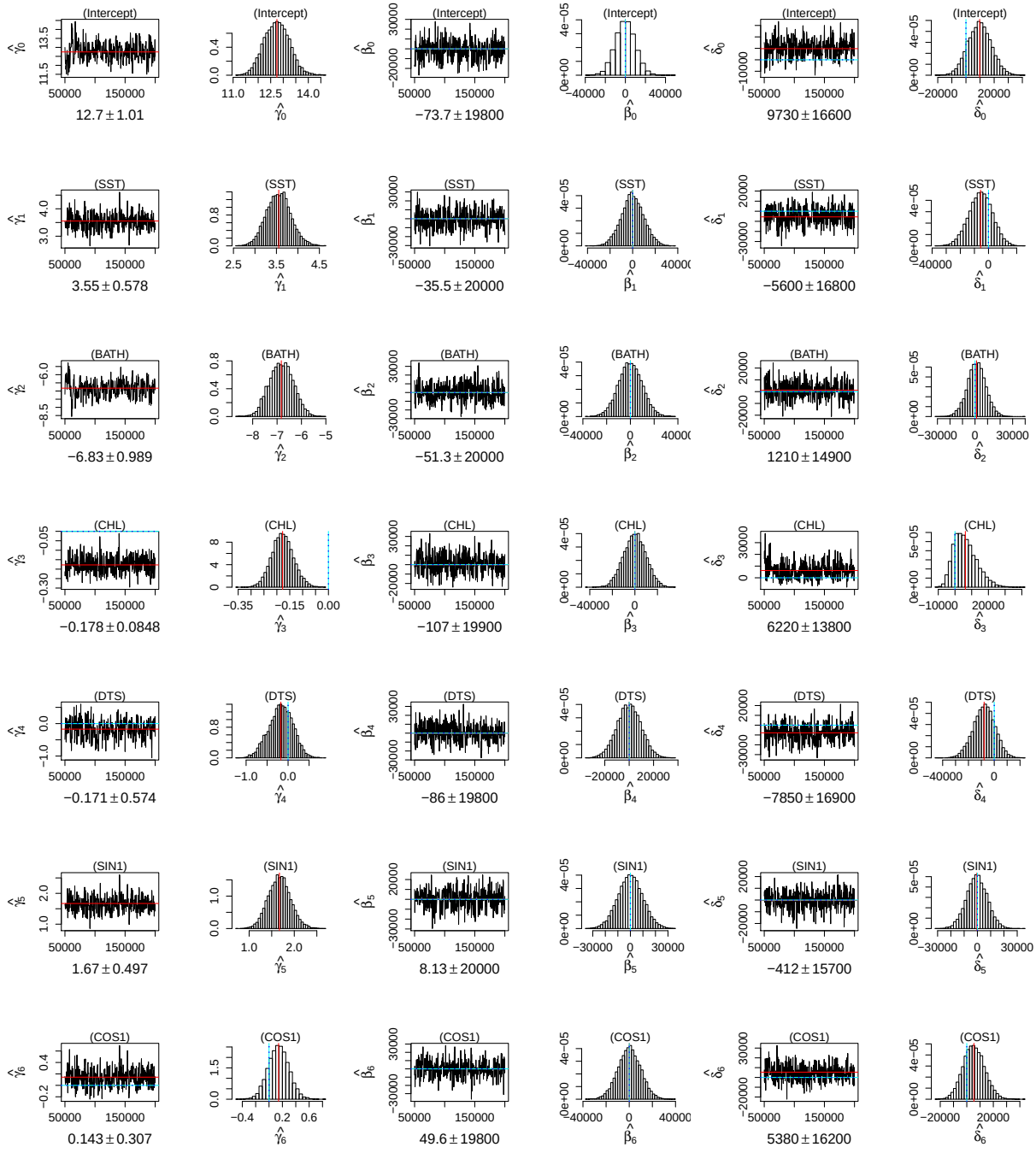

Figure 190: Long-tailed Duck: Traceplots and histograms of the posterior distributions of regression coefficients.

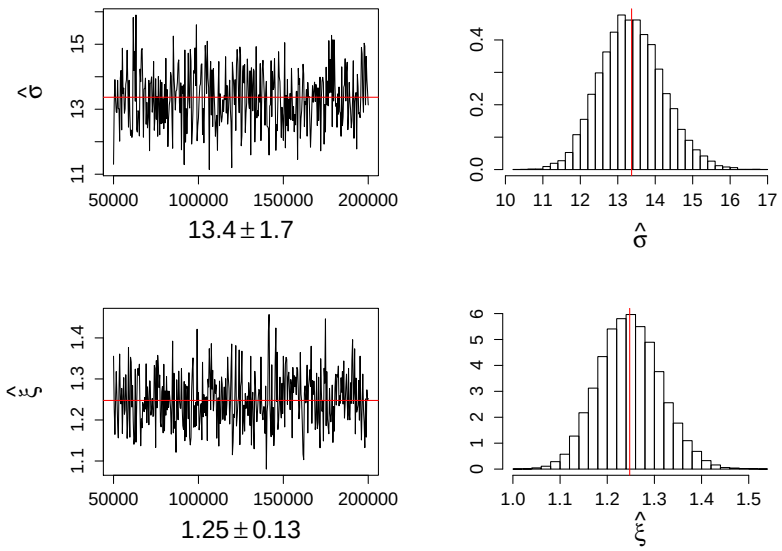

Figure 191: Long-tailed Duck: Traceplots and histograms of the posterior distributions of GPD parameters.

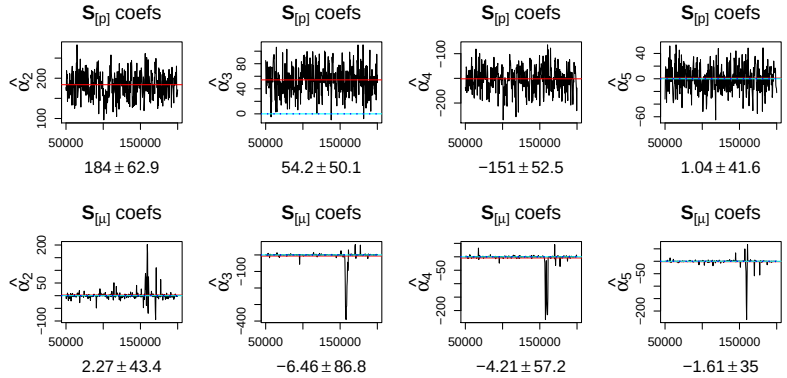

Figure 192: Long-tailed Duck: Traceplots and histograms of the posterior distributions of α coefficients.

B.17 Northern Fulmar

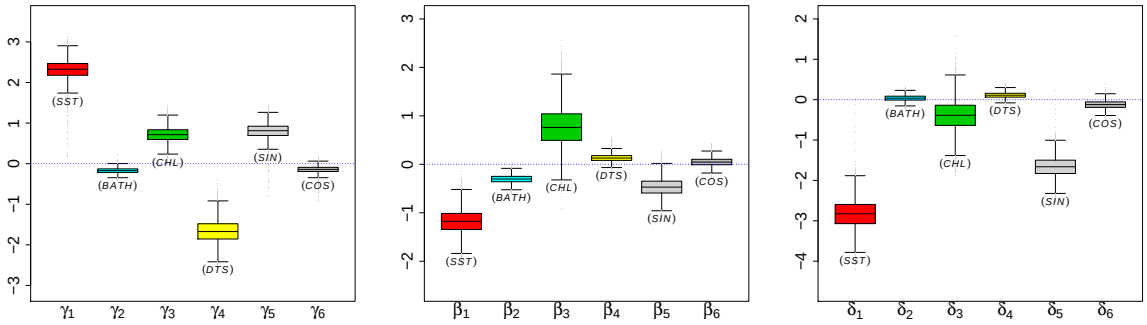

Figure 193: Northern Fulmar: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	3.3184 (0.1155)	−1.5018 (0.1409)	−2.4646 (0.2137)
$\mathbf{x}_1$	Sea surface temp.	2.3183 (0.2261)	−1.1788 (0.2461)	−2.8353 (0.3543)
$\mathbf{x}_2$	Bathymetry	−0.1719 (0.0660)	−0.3048 (0.0817)	0.0386 (0.0723)
$\mathbf{x}_3$	Chlorophyll	0.7176 (0.1784)	0.7723 (0.4074)	−0.3905 (0.3745)
$\mathbf{x}_4$	Distance to shore	−1.6642 (0.2806)	0.1306 (0.0750)	0.1113 (0.0703)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	0.8057 (0.1768)	−0.4701 (0.1799)	−1.6658 (0.2465)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	−0.1425 (0.0789)	0.0470 (0.0840)	−0.1221 (0.1001)

Table 21: Northern Fulmar: Posterior summary of regression coefficients.

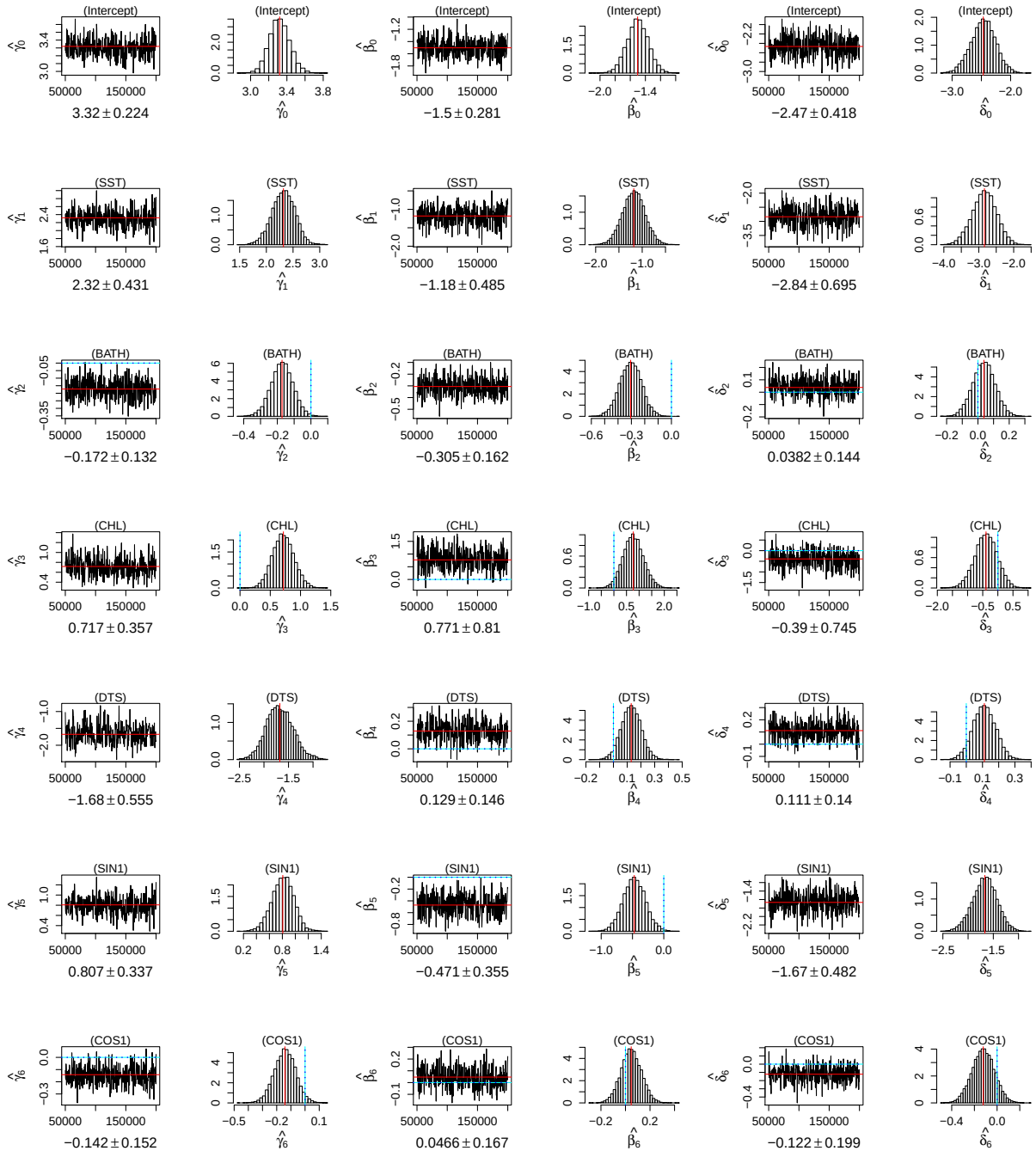

Figure 194: Northern Fulmar: Traceplots and histograms of the posterior distributions of regression coefficients.

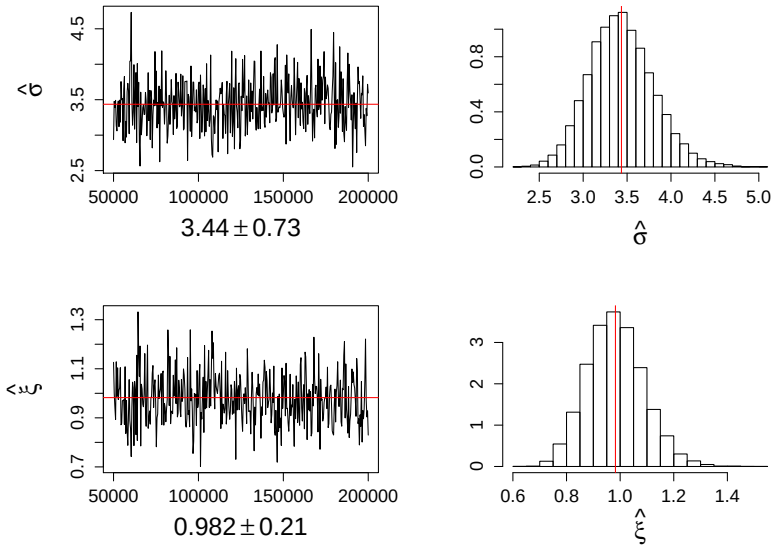

Figure 195: Northern Fulmar: Traceplots and histograms of the posterior distributions of GPD parameters.

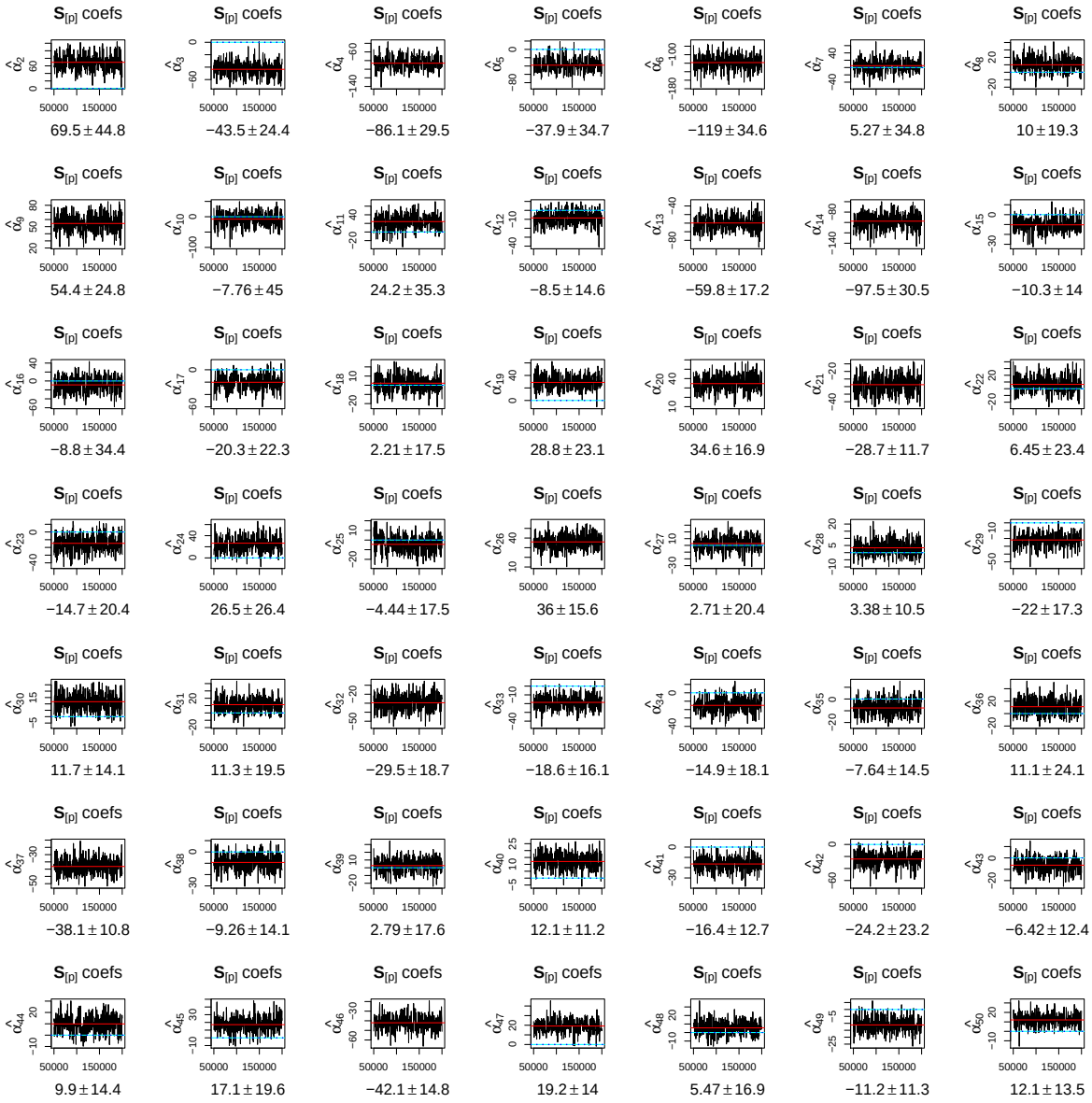

Figure 196: Northern Fulmar: Traceplots and histograms of the posterior distributions of α coefficients in the spatial regression of **p**.

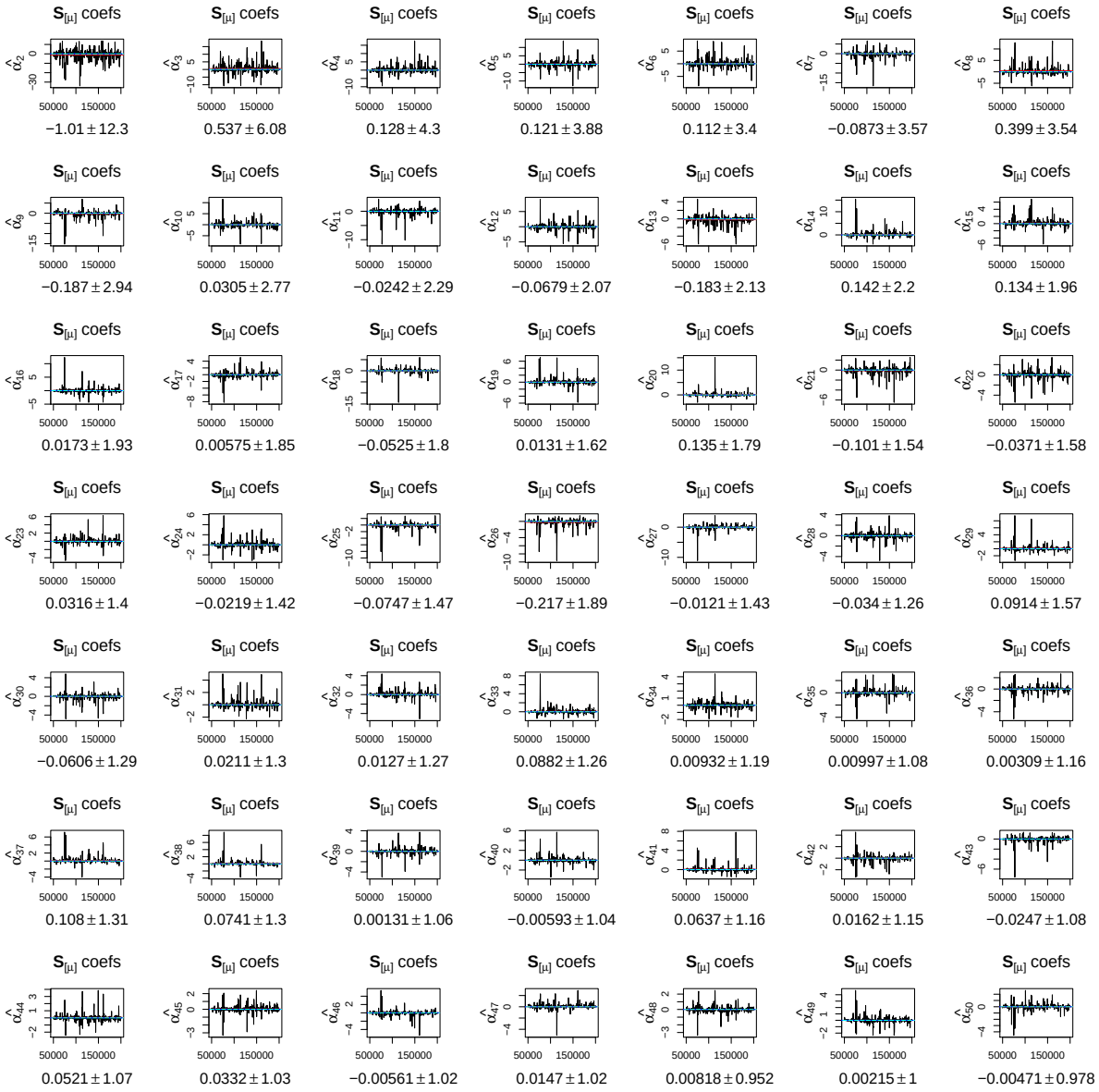

Figure 197: Northern Fulmar: Traceplots and histograms of the posterior distributions of α coefficients in the spatial regression of **m**.

B.18 Northern Gannet

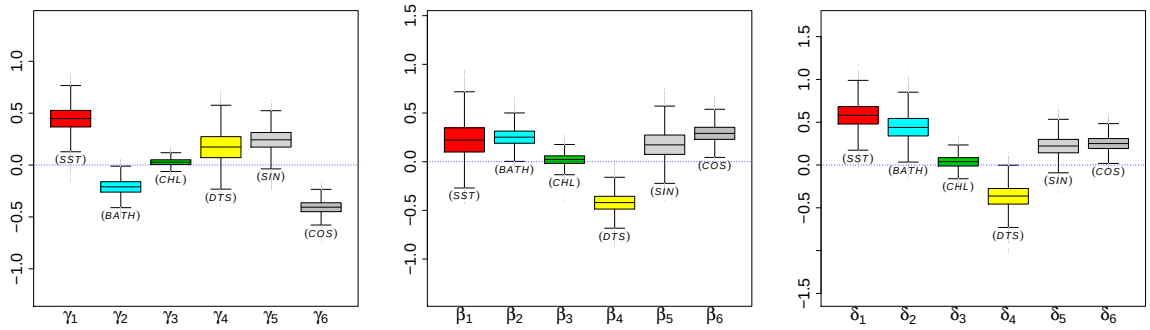

Figure 198: Northern Gannet: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	2.0702 (0.0524)	−1.3562 (0.0794)	−2.1358 (0.1313)
$\mathbf{x}_1$	Sea surface temp.	0.4463 (0.1193)	0.2240 (0.1857)	0.5793 (0.1523)
$\mathbf{x}_2$	Bathymetry	−0.2108 (0.0740)	0.2523 (0.0926)	0.4421 (0.1516)
$\mathbf{x}_3$	Chlorophyll	0.0286 (0.0338)	0.0221 (0.0582)	0.0367 (0.0742)
$\mathbf{x}_4$	Distance to shore	0.1716 (0.1492)	−0.4216 (0.0995)	−0.3663 (0.1352)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	0.2426 (0.1046)	0.1758 (0.1484)	0.2202 (0.1160)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	−0.4065 (0.0643)	0.2921 (0.0934)	0.2513 (0.0852)

Table 22: Northern Gannet: Posterior summary of regression coefficients.

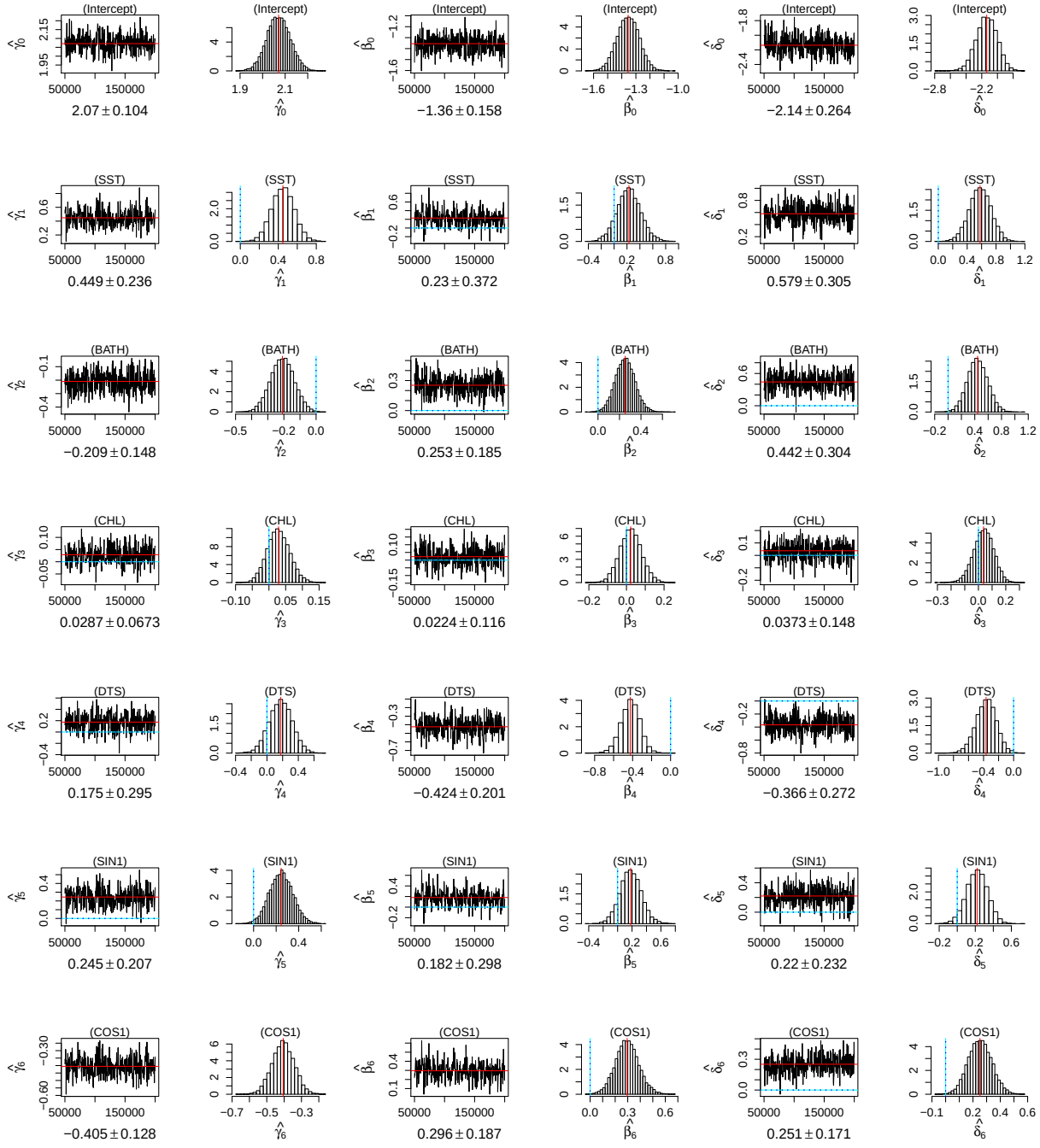

Figure 199: Northern Gannet: Traceplots and histograms of the posterior distributions of regression coefficients.

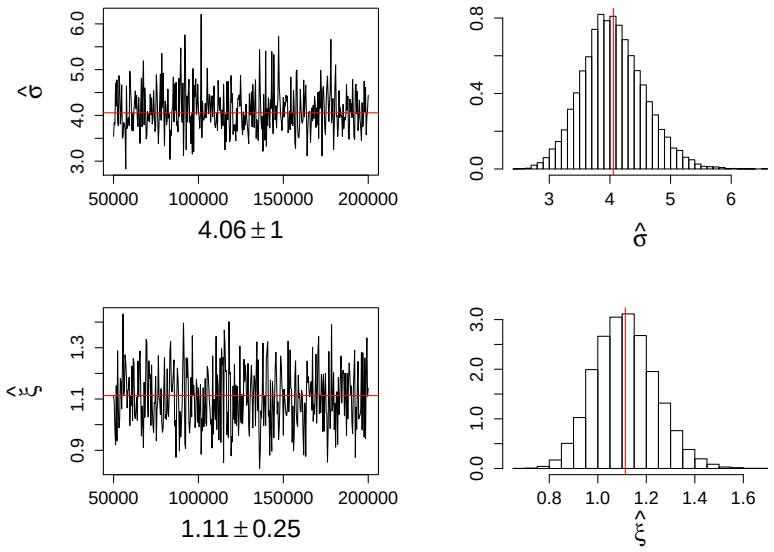

Figure 200: Northern Gannet: Traceplots and histograms of the posterior distributions of GPD parameters.

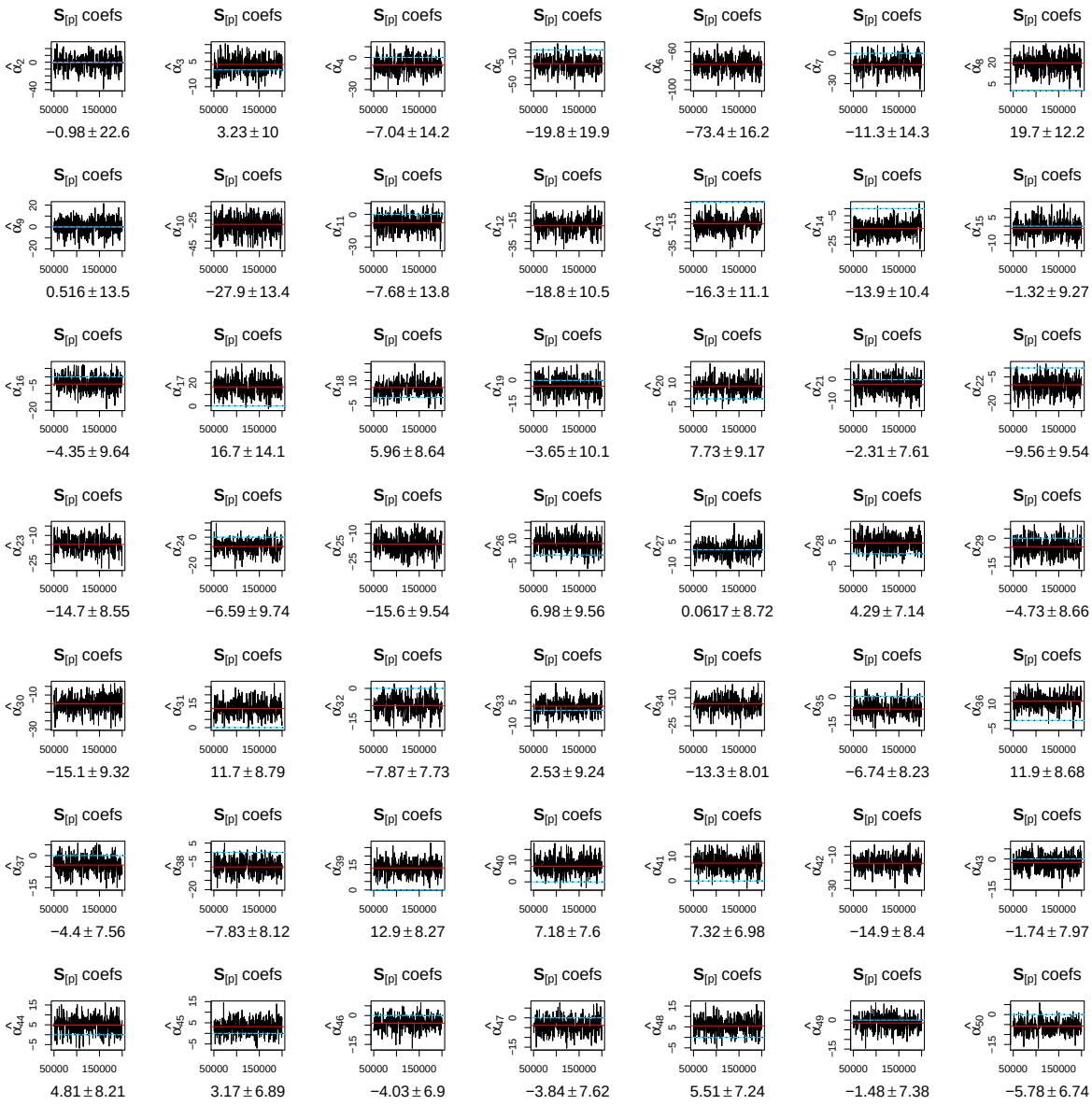

Figure 201: Northern Gannet: Traceplots and histograms of the posterior distributions of α coefficients in the spatial regression of $\mathbf{p}$.

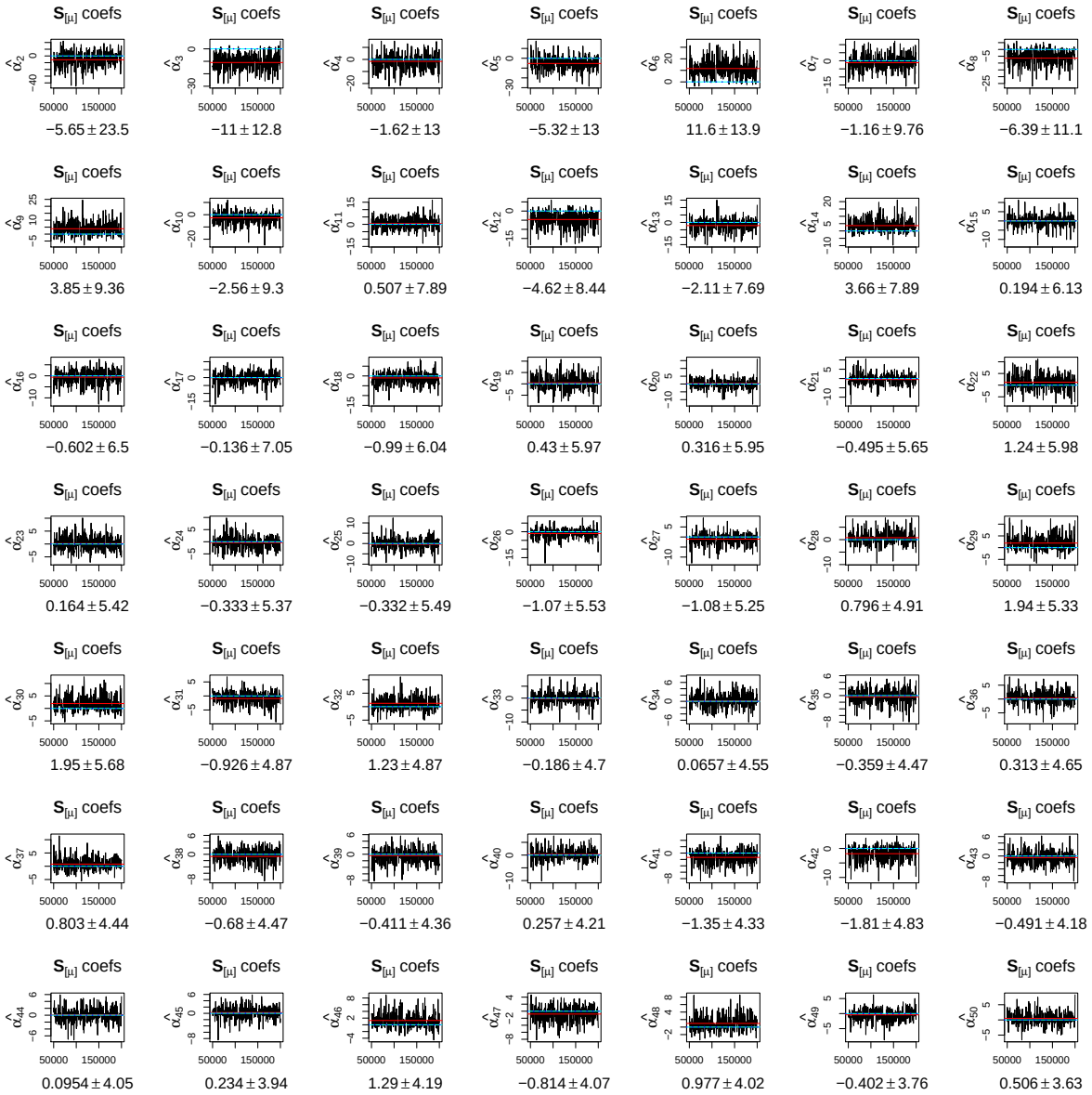

Figure 202: Northern Gannet: Traceplots and histograms of the posterior distributions of α coefficients in the spatial regression of $\mathbf{m}$.

B.19 Razorbill

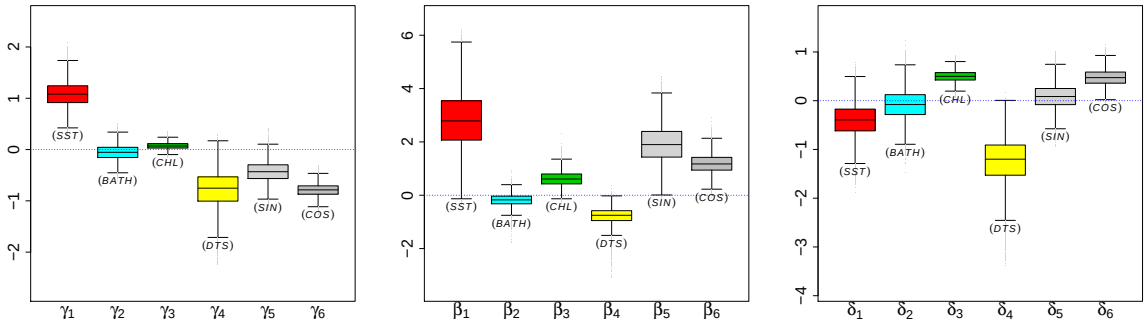

Figure 203: Razorbill: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	5.4164 (0.1860)	0.315 (0.773)	−3.2558 (0.4747)
$\mathbf{x}_1$	Sea surface temp.	1.0798 (0.2396)	2.820 (1.033)	−0.4005 (0.3387)
$\mathbf{x}_2$	Bathymetry	−0.0569 (0.1477)	−0.183 (0.231)	−0.0788 (0.3070)
$\mathbf{x}_3$	Chlorophyll	0.0717 (0.0620)	0.618 (0.282)	0.5000 (0.1130)
$\mathbf{x}_4$	Distance to shore	−0.7792 (0.3597)	−0.785 (0.309)	−1.2351 (0.4619)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	−0.4325 (0.2001)	1.925 (0.693)	0.0838 (0.2491)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	−0.7896 (0.1214)	1.187 (0.355)	0.4750 (0.1700)

Table 23: Razorbill: Posterior summary of regression coefficients.

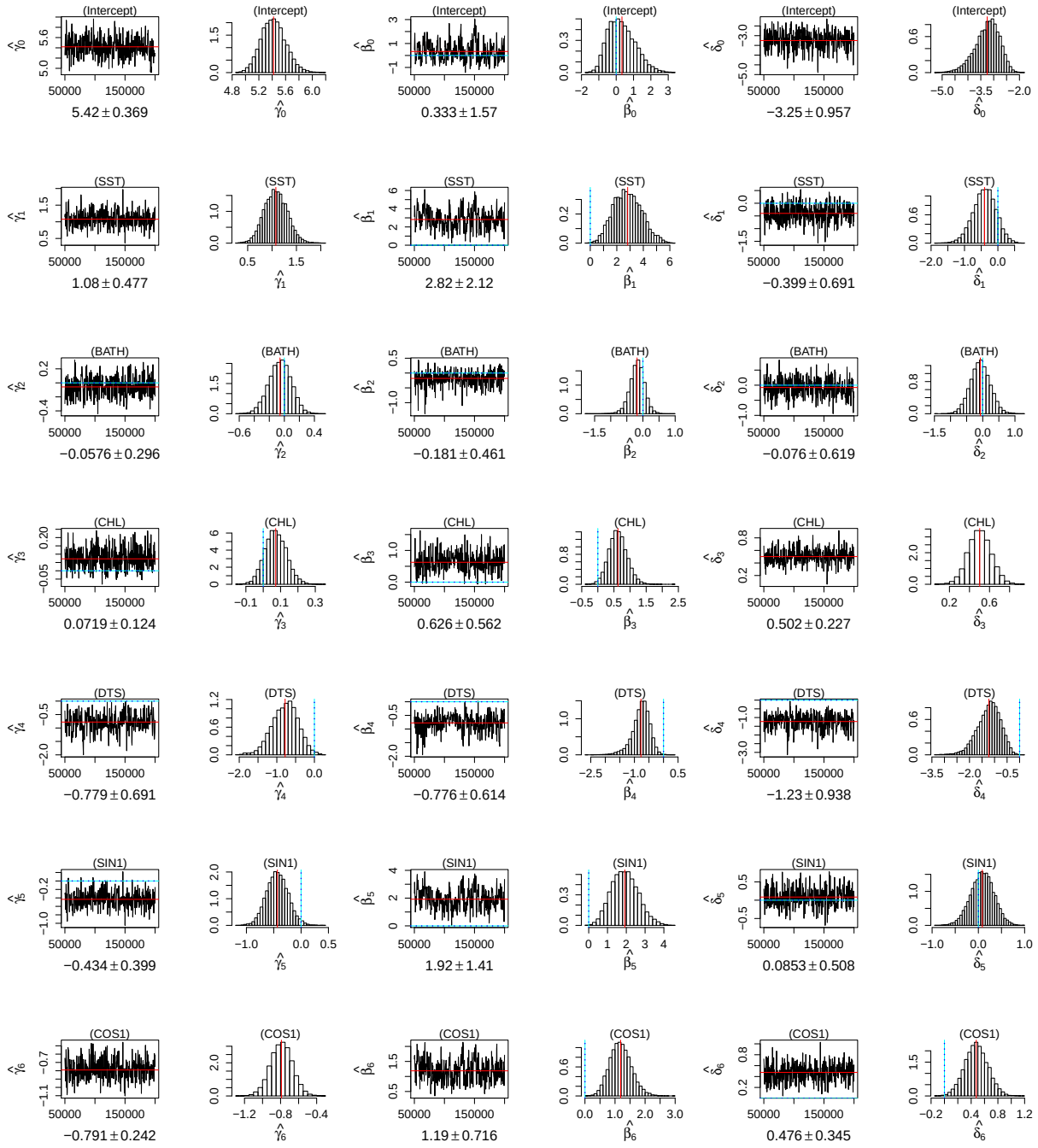

Figure 204: Razorbill: Traceplots and histograms of the posterior distributions of regression coefficients.

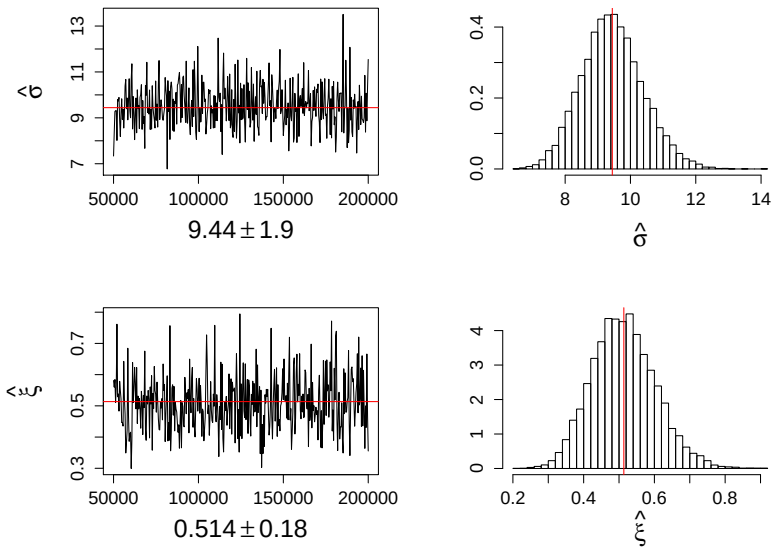

Figure 205: Razorbill: Traceplots and histograms of the posterior distributions of GPD parameters.

Figure 206: Razorbill: Traceplots and histograms of the posterior distributions of α coefficients in the spatial regression of $\mathbf{p}$.

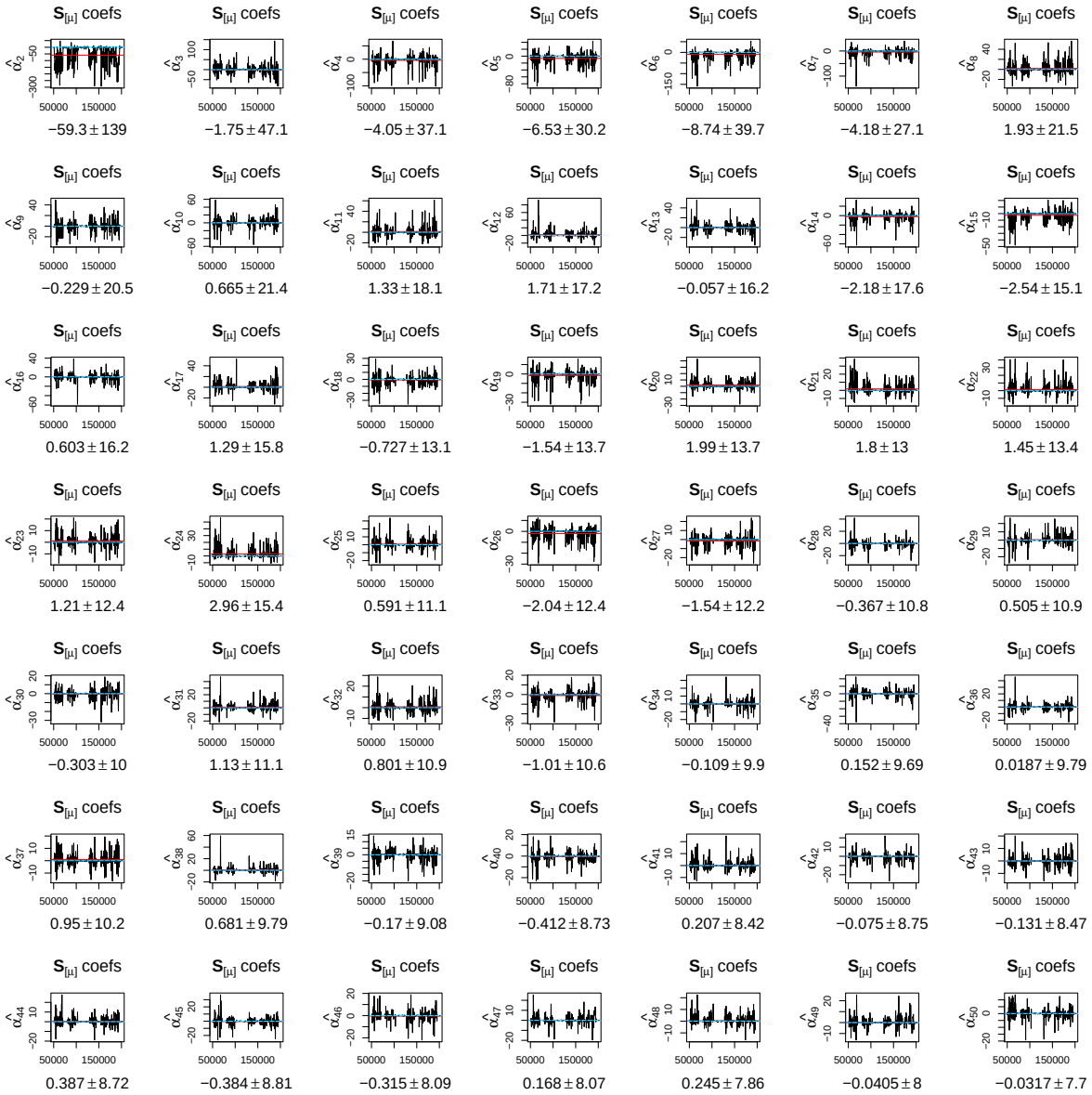

Figure 207: Razorbill: Traceplots and histograms of the posterior distributions of α coefficients in the spatial regression of $\mathbf{m}$.

B.20 Roseate Tern

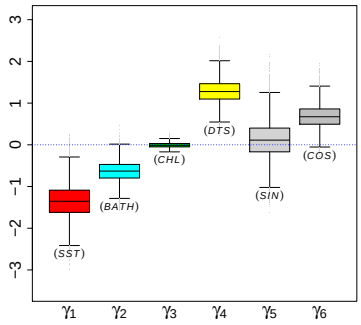

Figure 208: Roseate Tern: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\gamma}$
Predictor		mean (sd)
1	Intercept	7.43340 (0.37155)
$\mathbf{x}_1$	Sea surface temp.	−1.34921 (0.39881)
$\mathbf{x}_2$	Bathymetry	−0.63382 (0.24193)
$\mathbf{x}_3$	Chlorophyll	−0.00714 (0.06098)
$\mathbf{x}_4$	Distance to shore	1.28812 (0.27227)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	0.12395 (0.42975)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	0.68069 (0.27116)

Table 24: Roseate Tern: Posterior summary of regression coefficients.

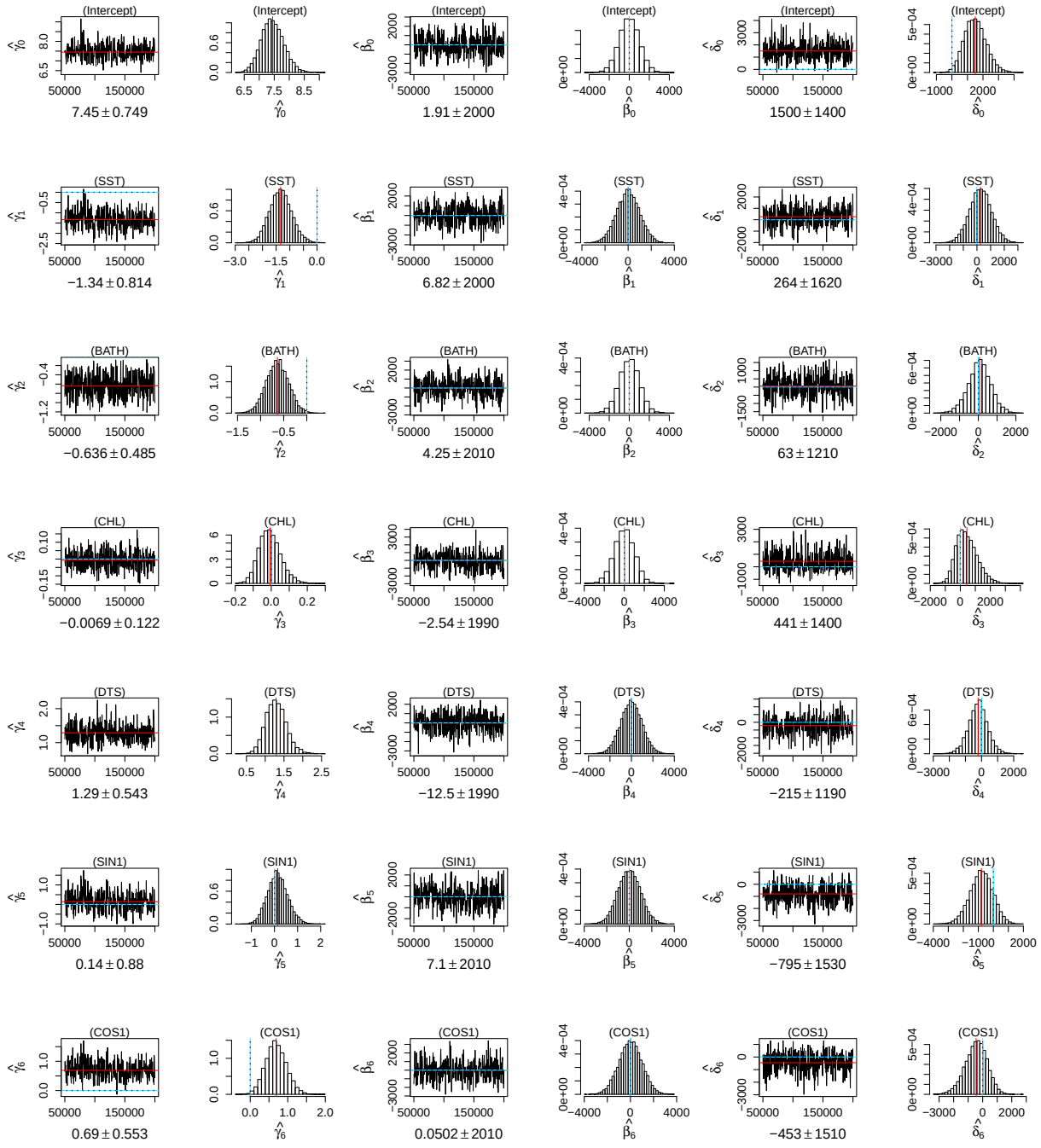

Figure 209: Roseate Tern: Traceplots and histograms of the posterior distributions of regression coefficients.

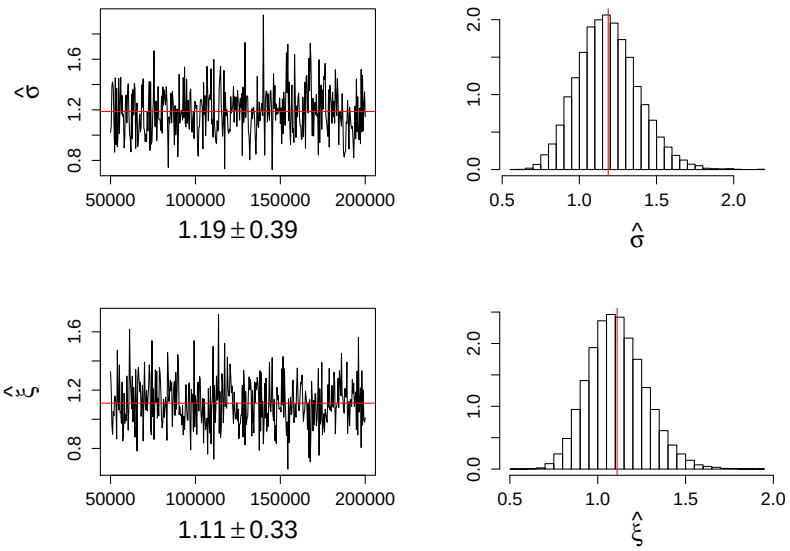

Figure 210: Roseate Tern: Traceplots and histograms of the posterior distributions of GPD parameters.

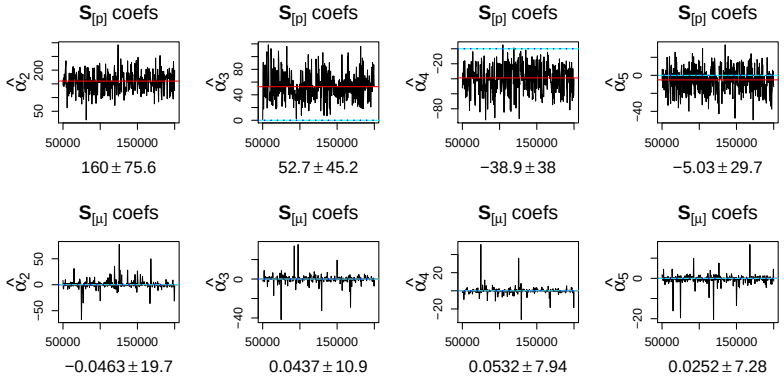

Figure 211: Roseate Tern: Traceplots and histograms of the posterior distributions of α coefficients.

B.21 Sooty Shearwater

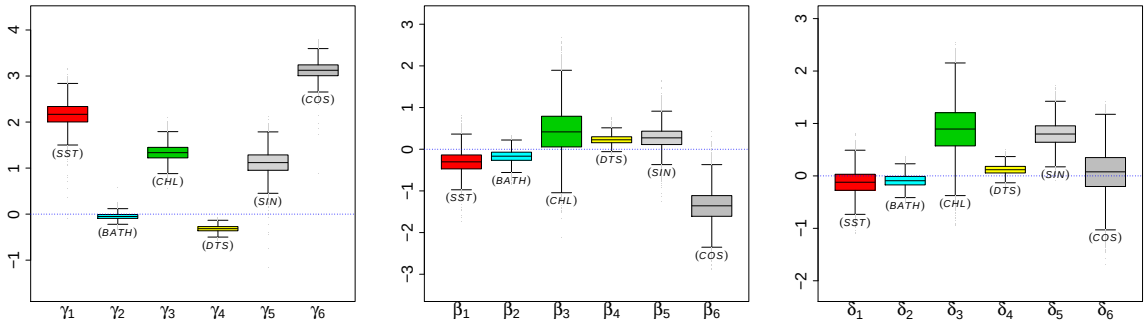

Figure 212: Sooty Shearwater: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	5.1122 (0.1722)	−2.543 (0.450)	−1.3150 (0.4899)
$\mathbf{x}_1$	Sea surface temp.	2.1690 (0.2463)	−0.306 (0.252)	−0.1226 (0.2270)
$\mathbf{x}_2$	Bathymetry	−0.0516 (0.0641)	−0.171 (0.145)	−0.0905 (0.1200)
$\mathbf{x}_3$	Chlorophyll	1.3375 (0.1706)	0.430 (0.547)	0.8861 (0.4701)
$\mathbf{x}_4$	Distance to shore	−0.3169 (0.0676)	0.229 (0.107)	0.1182 (0.0923)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	1.1184 (0.2473)	0.271 (0.247)	0.7982 (0.2327)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	3.1270 (0.1757)	−1.363 (0.375)	0.0639 (0.4110)

Table 25: Sooty Shearwater: Posterior summary of regression coefficients.

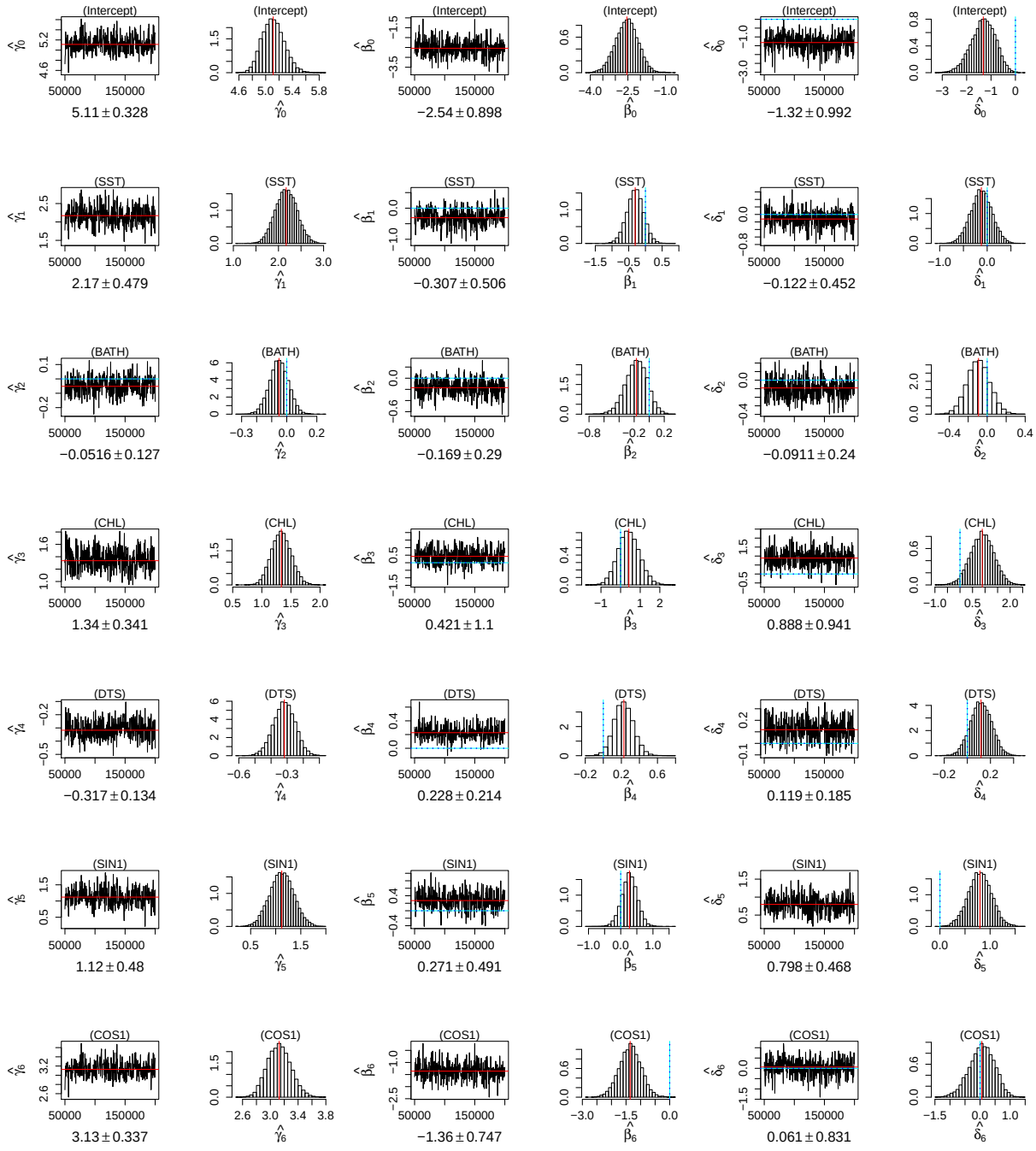

Figure 213: Sooty Shearwater: Traceplots and histograms of the posterior distributions of regression coefficients.

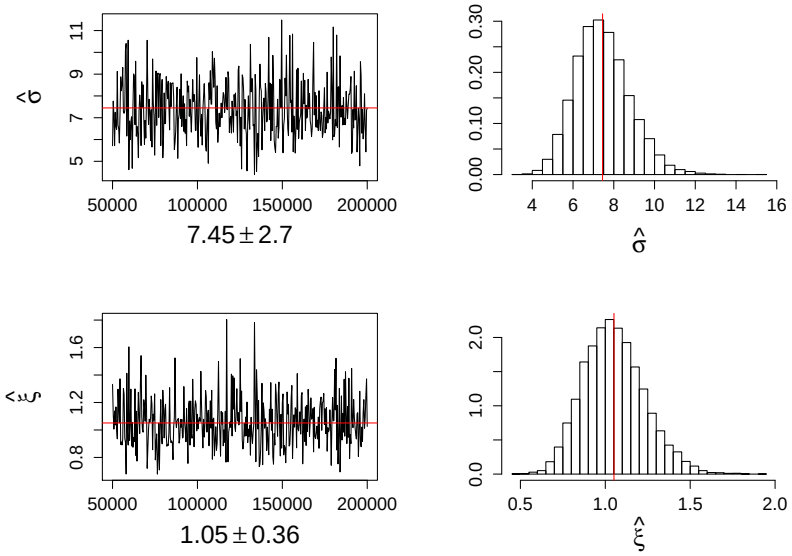

Figure 214: Sooty Shearwater: Traceplots and histograms of the posterior distributions of GPD parameters.

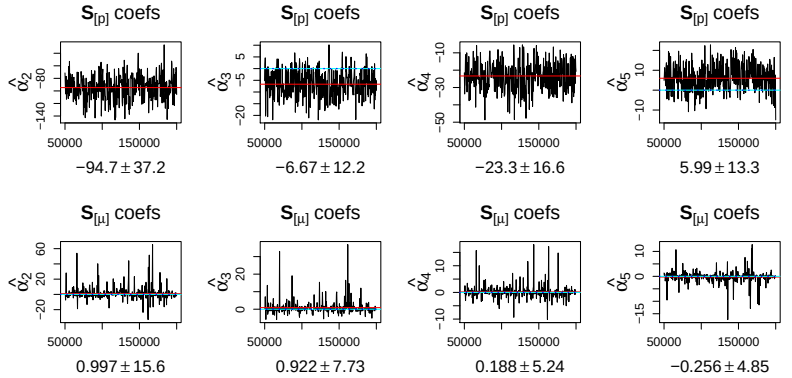

Figure 215: Sooty Shearwater: Traceplots and histograms of the posterior distributions of α coefficients.

B.22 Surf Scoter

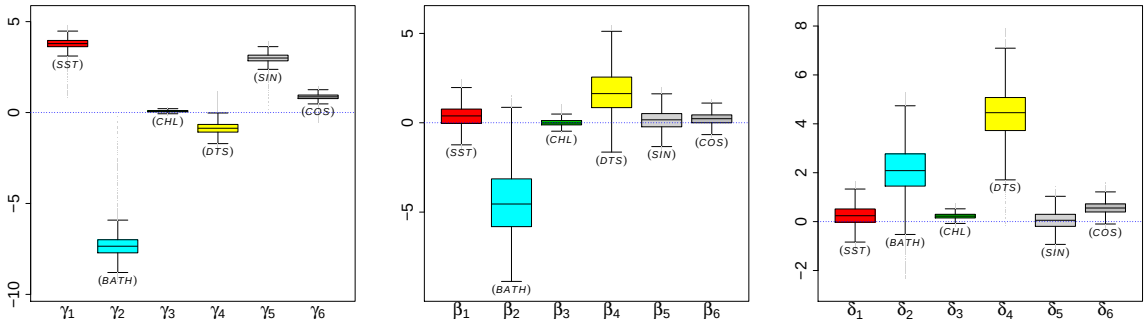

Figure 216: Surf Scoter: Boxplots of the posterior distributions of regression coefficients.

		$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\delta}}$
Predictor		mean (sd)	mean (sd)	mean (sd)
1	Intercept	12.3842 (0.6176)	9.6965 (2.8037)	2.0719 (1.1479)
$\mathbf{x}_1$	Sea surface temp.	3.7876 (0.2887)	0.3664 (0.5782)	0.2435 (0.4089)
$\mathbf{x}_2$	Bathymetry	− 7.3428 (0.6223)	−4.3712 (1.9668)	2.0912 (1.0349)
$\mathbf{x}_3$	Chlorophyll	0.0738 (0.0528)	0.0149 (0.1911)	0.2236 (0.1116)
$\mathbf{x}_4$	Distance to shore	− 0.8630 (0.3388)	1.7037 (1.1655)	4.3810 (1.0315)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	2.9948 (0.2637)	0.1484 (0.5304)	0.0529 (0.3721)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	0.8623 (0.1585)	0.2216 (0.3173)	0.5593 (0.2454)

Table 26: Surf Scoter: Posterior summary of regression coefficients.

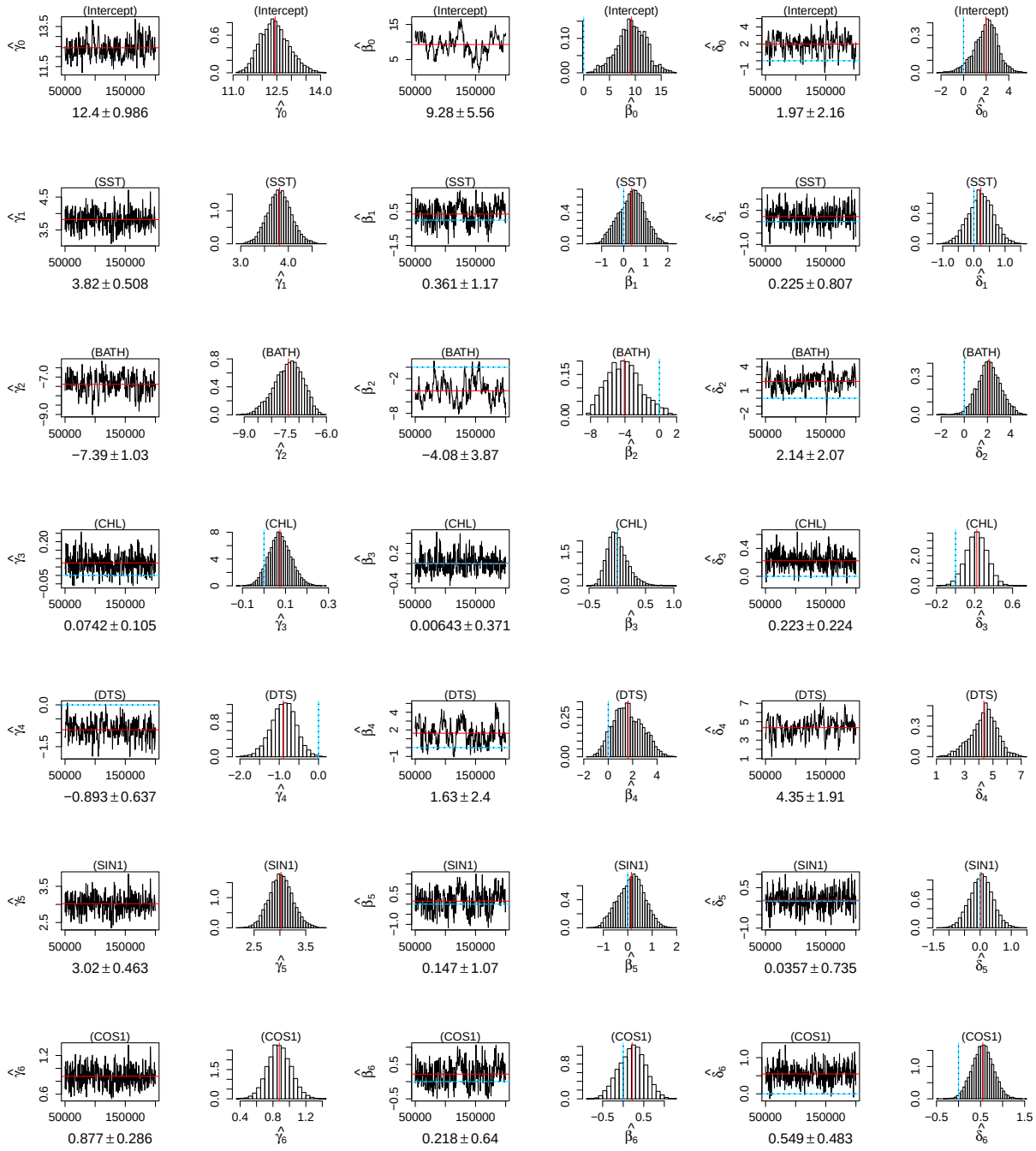

Figure 217: Surf Scoter: Traceplots and histograms of the posterior distributions of regression coefficients.

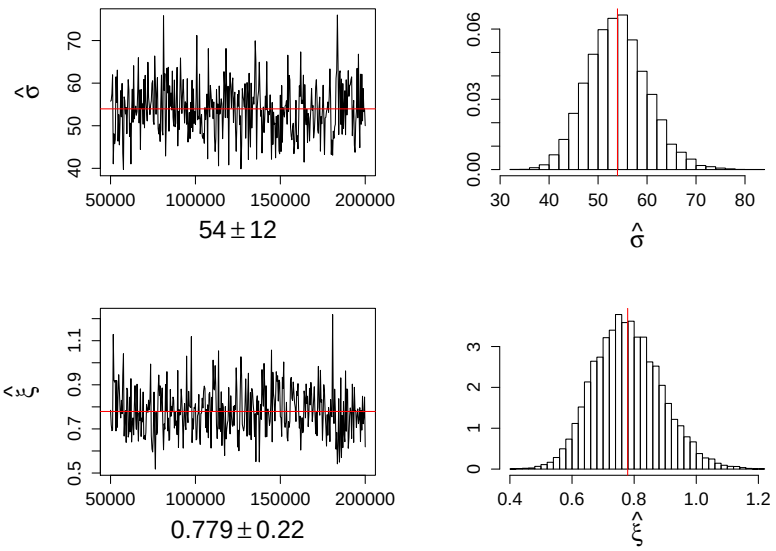

Figure 218: Surf Scoter: Traceplots and histograms of the posterior distributions of GPD parameters.

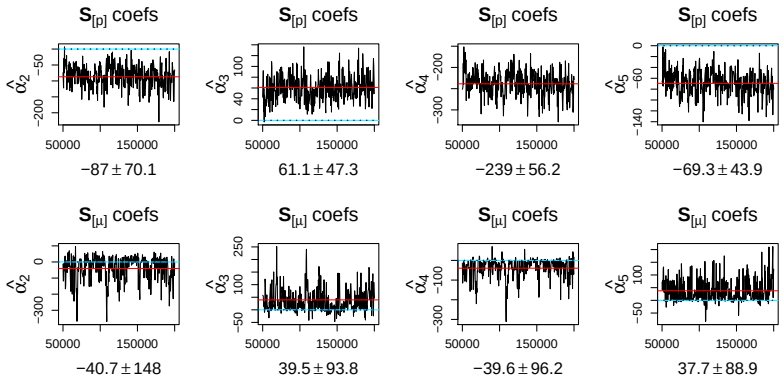

Figure 219: Surf Scoter: Traceplots and histograms of the posterior distributions of α coefficients.

B.23 White-winged Scoter

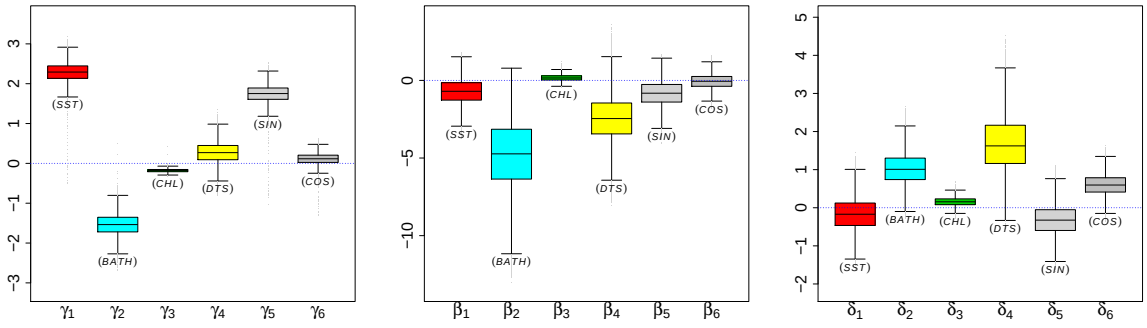

Figure 220: White-winged Scoter: Boxplots of the posterior distributions of regression coefficients.

Predictor		$\hat{\gamma}$ mean (sd)	$\hat{\beta}$ mean (sd)	$\hat{\delta}$ mean (sd)
1	Intercept	7.9584 (0.3338)	1.6556 (2.1087)	−0.124 (0.731)
$\mathbf{x}_1$	Sea surface temp.	2.2872 (0.2562)	−0.7188 (0.8435)	−0.170 (0.435)
$\mathbf{x}_2$	Bathymetry	−1.5406 (0.2797)	−4.8671 (2.3324)	1.036 (0.426)
$\mathbf{x}_3$	Chlorophyll	−0.1811 (0.0419)	0.1692 (0.2060)	0.157 (0.115)
$\mathbf{x}_4$	Distance to shore	0.2725 (0.2712)	−2.4370 (1.5845)	1.694 (0.769)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	1.7466 (0.2320)	−0.8431 (0.8456)	−0.325 (0.401)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	0.1132 (0.1454)	−0.0631 (0.4717)	0.601 (0.275)

Table 27: White-winged Scoter: Posterior summary of regression coefficients.

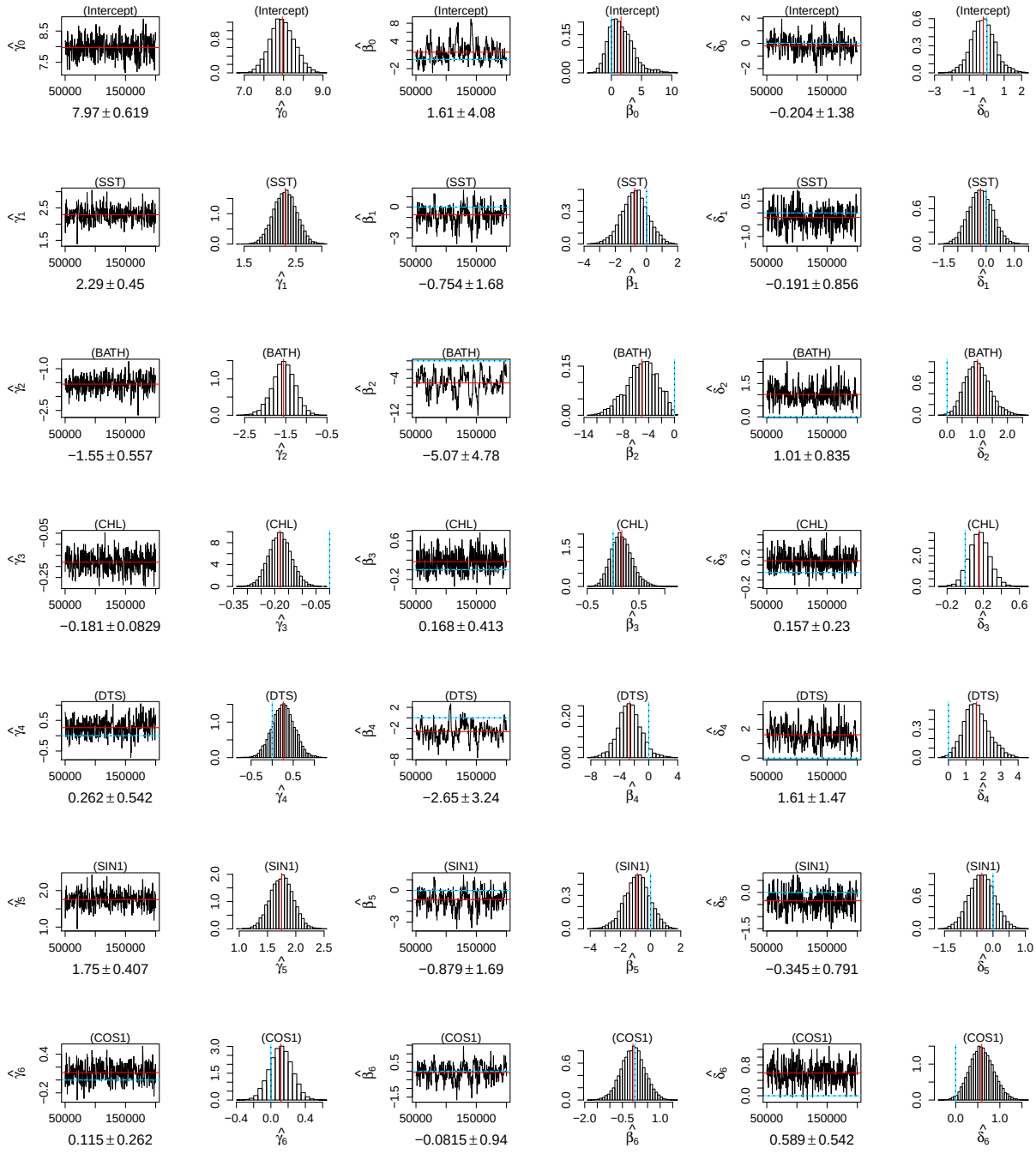

Figure 221: White-winged Scoter: Traceplots and histograms of the posterior distributions of regression coefficients.

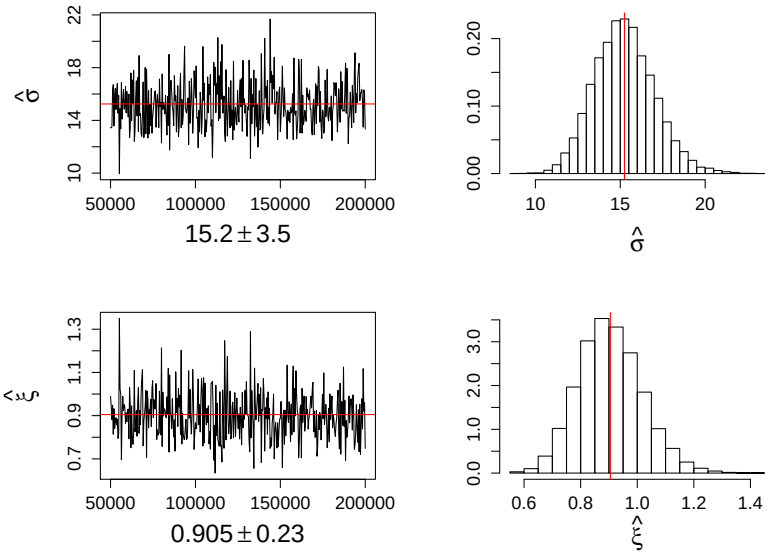

Figure 222: White-winged Scoter: Traceplots and histograms of the posterior distributions of GPD parameters.

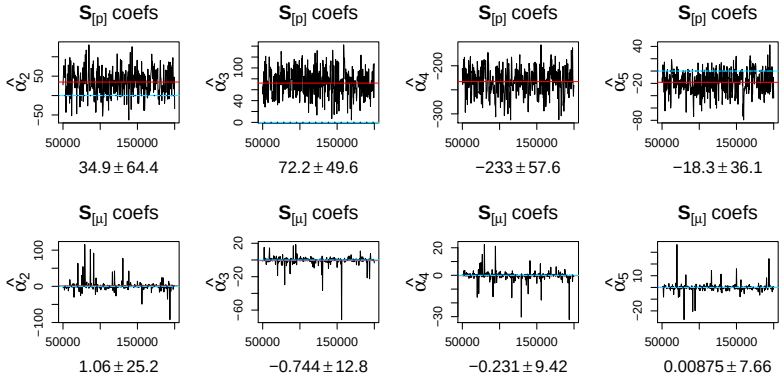

Figure 223: White-winged Scoter: Traceplots and histograms of the posterior distributions of α coefficients.

B.24 Wilson's Storm-petrel

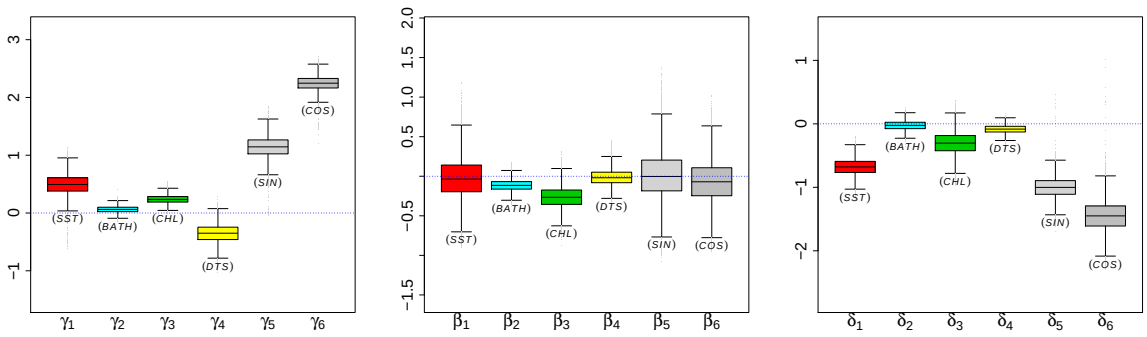

Figure 224: Wilson's Storm-petrel: Boxplots of the posterior distributions of regression coefficients.

Predictor		$\hat{\boldsymbol{\gamma}}$ mean (sd)	$\hat{\boldsymbol{\beta}}$ mean (sd)	$\hat{\boldsymbol{\delta}}$ mean (sd)
1	Intercept	3.4174 (0.1146)	−1.1373 (0.2532)	−2.8834 (0.2955)
$\mathbf{x}_1$	Sea surface temp.	0.4920 (0.1806)	−0.0212 (0.2565)	−0.6793 (0.1300)
$\mathbf{x}_2$	Bathymetry	0.0615 (0.0565)	−0.1153 (0.0703)	−0.0250 (0.0742)
$\mathbf{x}_3$	Chlorophyll	0.2372 (0.0715)	−0.2644 (0.1365)	−0.3068 (0.1742)
$\mathbf{x}_4$	Distance to shore	−0.3528 (0.1620)	−0.0119 (0.1013)	−0.0836 (0.0664)
$\mathbf{x}_5$	$sin(\frac{\pi}{6} \cdot Month)$	1.1426 (0.1899)	0.0150 (0.2938)	−1.0021 (0.1622)
$\mathbf{x}_6$	$cos(\frac{\pi}{6} \cdot Month)$	2.2454 (0.1259)	−0.0671 (0.2631)	−1.4502 (0.2414)

Table 28: Wilson's Storm-petrel: Posterior summary of regression coefficients.

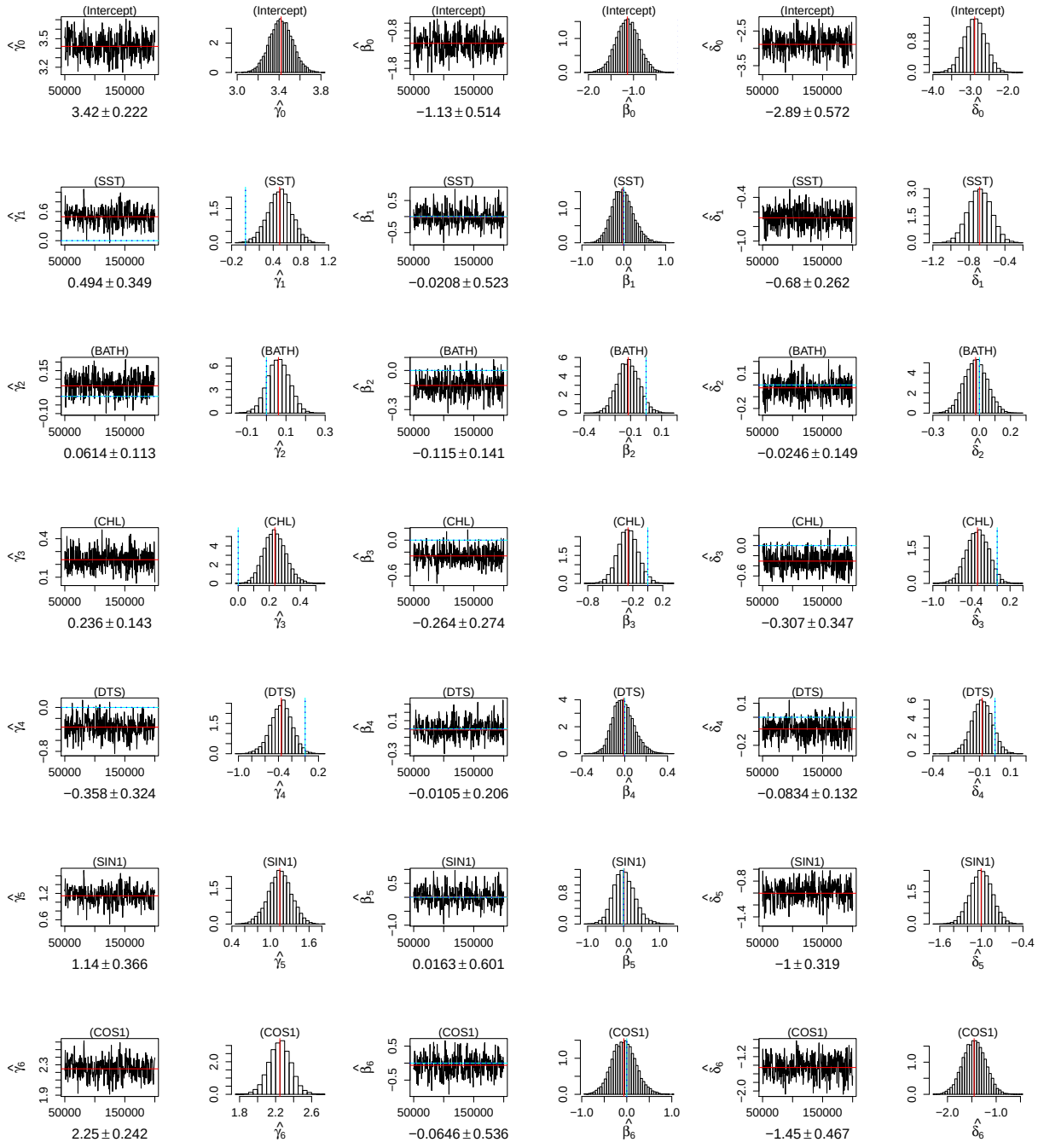

Figure 225: Wilson's Storm-petrel: Traceplots and histograms of the posterior distributions of regression coefficients.

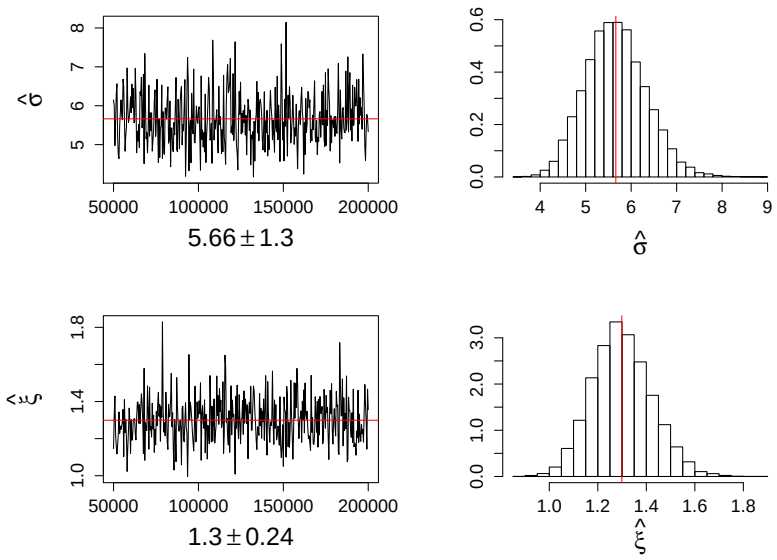

Figure 226: Wilson's Storm-petrel: Traceplots and histograms of the posterior distributions of GPD parameters.

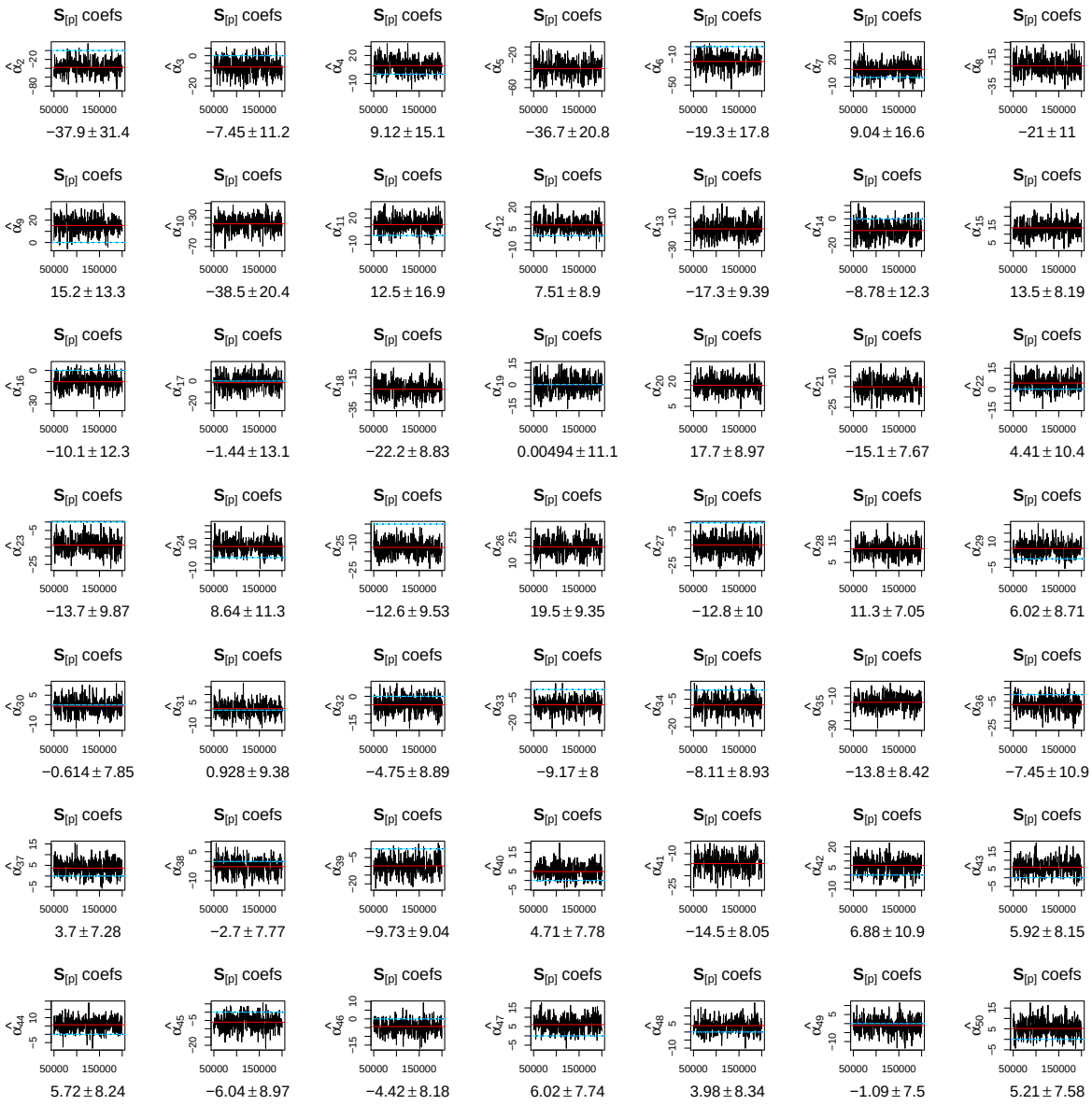

Figure 227: Wilson's Storm-petrel: Traceplots and histograms of the posterior distributions of α coefficients in the spatial regression of **p**.

Figure 228: Wilson's Storm-petrel: Traceplots and histograms of the posterior distributions of α coefficients in the spatial regression of $\mathbf{m}$.