Hierarchical Models for Estimating Seabird Distributions in the Northwest Atlantic

Beth Gardner[1, 2], Allan F. O'Connell, Jr.[2], & Andrew T. Gilbert[2]

[1]North Carolina State University

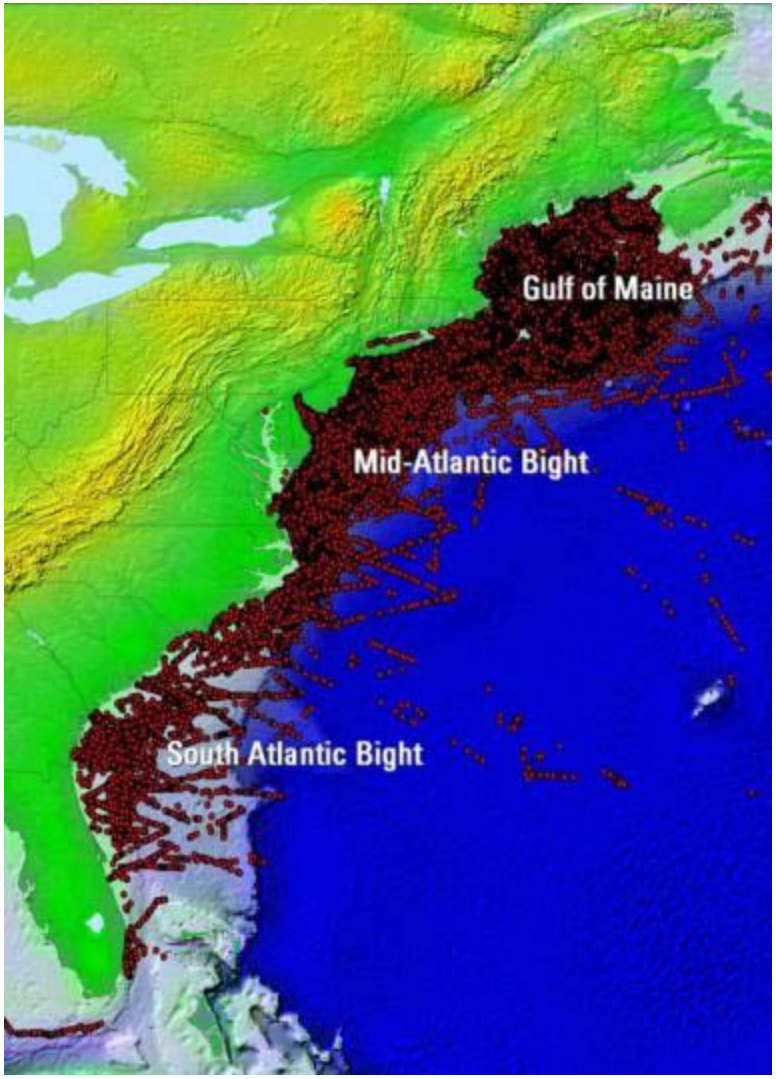
[2]USGS Patuxent Wildlife Research Center

Background

- To model and predict the distribution of selected seabird species in the NW Atlantic
  - gather seabird data and biotic and abiotic variables
  - develop models for seabirds using available data and environmental variables

Seabird Occurrence Database

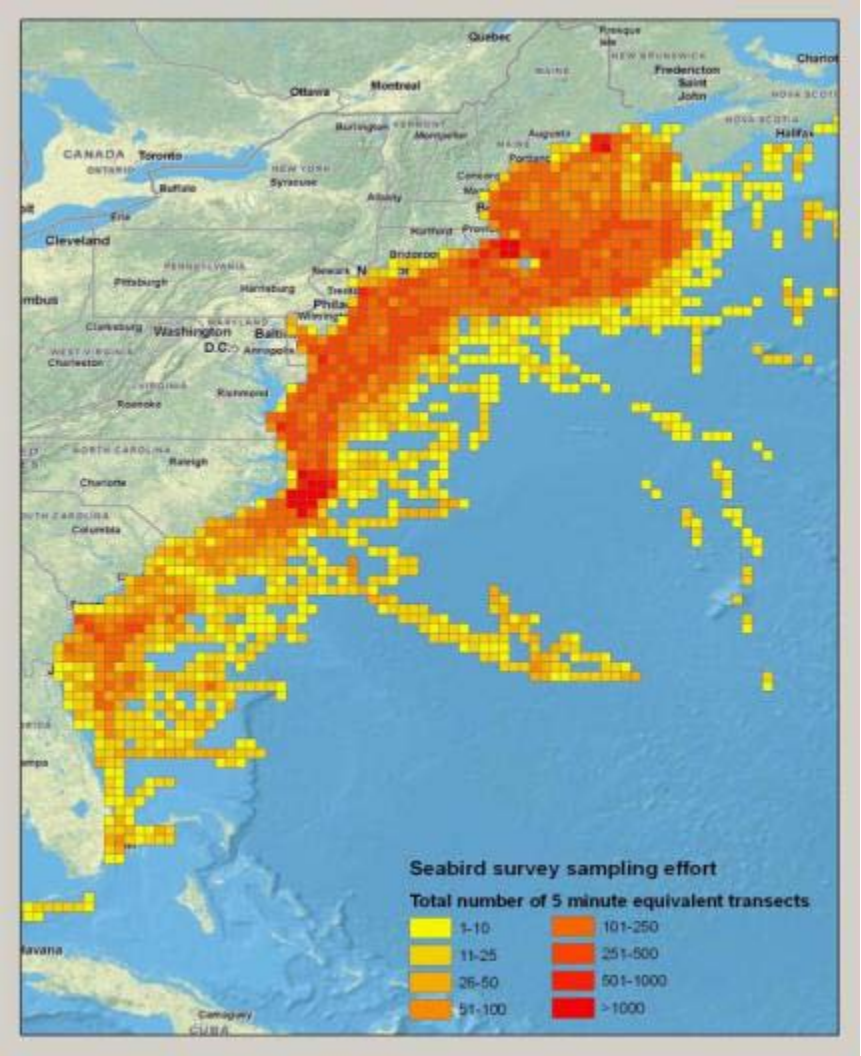
- >250,000 seabird observation from U.S. Atlantic waters (>100k from Canada in PIROP)
- most data collected from 1978 through 2009
- Data collected using a mix of scientific and non-scientific methods



Standardized Effort

- We standardized survey effort to account for both discrete-time and continuous-strip surveys.

- Standardized 5-minute equivalents in each cell.



Overview

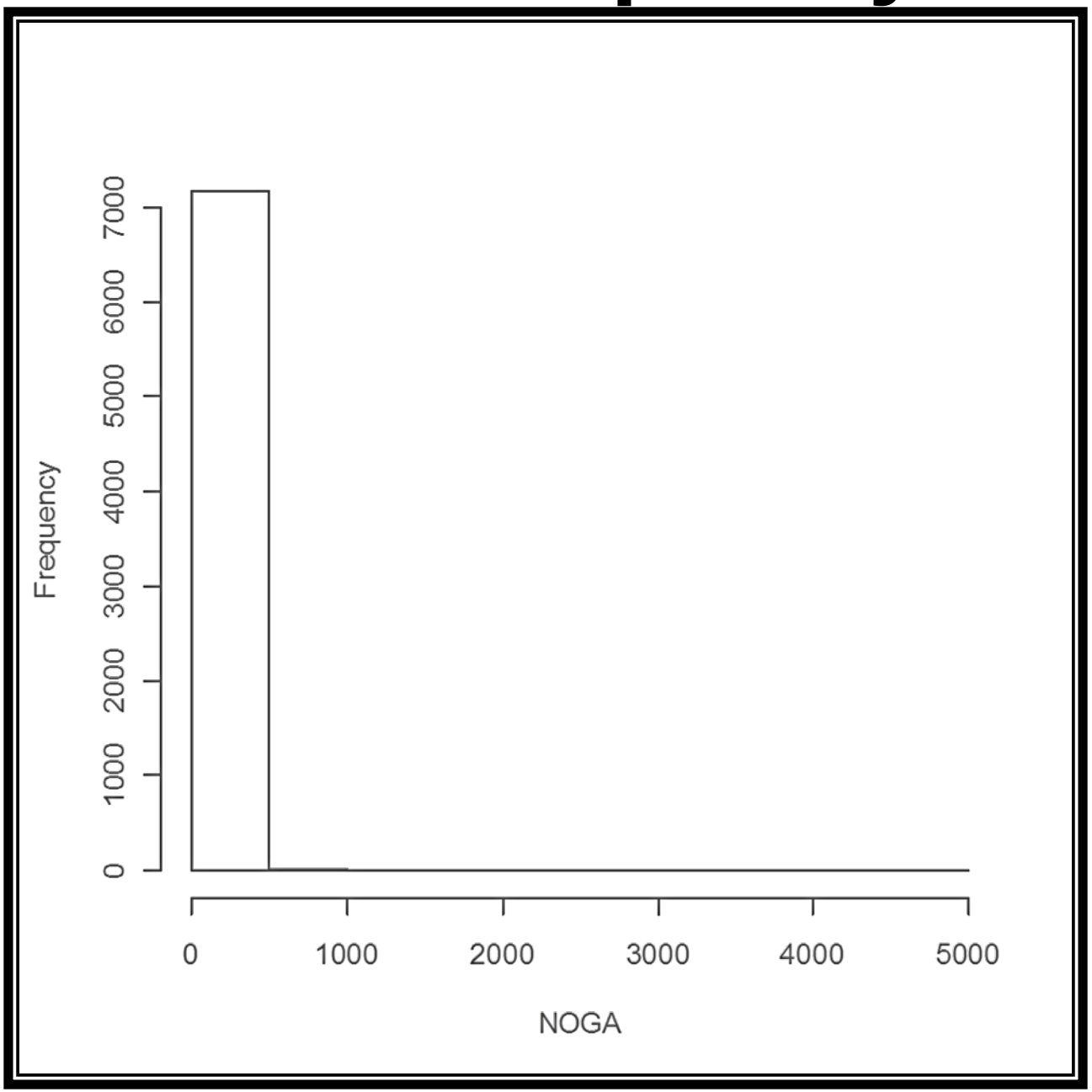
- ~ 70 sea bird species in the data base that are typically found in the Atlantic
  - 10 – 15 of particular interest

- Demonstrate two modeling techniques
  - Broad species distribution mapping
  - Community occupancy modeling

Spatial Poisson Regression

- Poisson observation model (commonly used for count data) with explicit spatial correlation and covariates

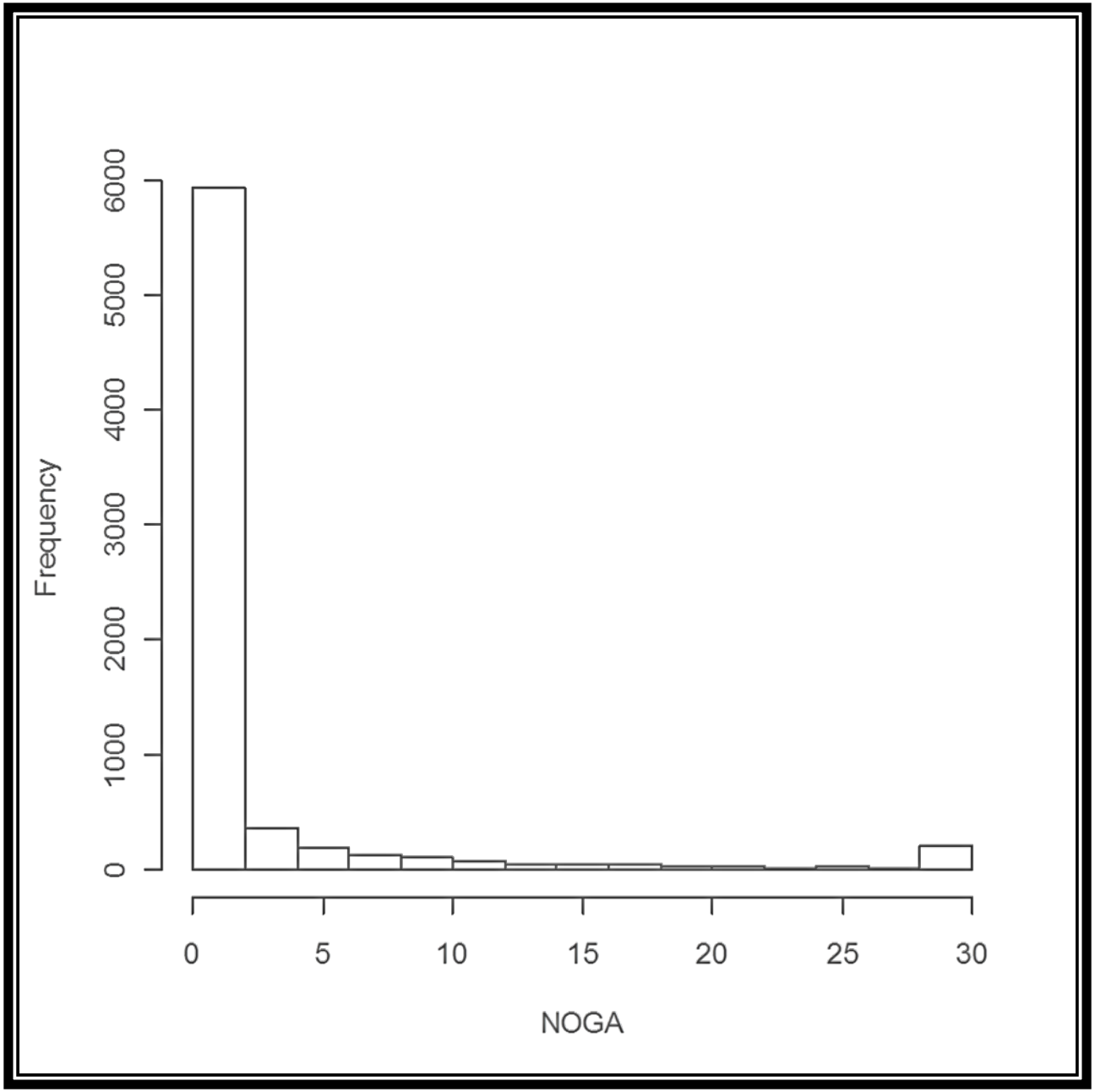
Count Frequency



Truncated Data

- Set any count over 30 to be equal to 30
- This improves fitting a distribution to the data
- Possible aggregations are eliminated by this truncation

Truncated Count Frequency



Poisson Model

$$y_{ijt} \sim Poisson(\lambda_{ijt})$$

where y_{ij} is the observation at grid cell *i* during survey *j* in year *t*

The log-linear full model is:

$$\log(\lambda_{ijt}) = \log(effort_{ij}) + a + \alpha_2[Month_{ij}] + \beta_1 * SST_{ij} + \beta_2 * Chlor_{ij} + \beta_3 * NAO_t + Z_i$$

Breaking Down the Model

$$\log(\lambda_{ijt}) = \log(effort_{ij}) + a + \alpha_2[Month_{ij}]$$

Where:

- effort is an offset term
- a is the intercept
- a_2 is a seasonal effect (month)

Breaking Down the Model

$$\beta_1 * SST_{ij} + \beta_2 * Chlor_{ij} + \beta_3 * NAO_t + z_i$$

Where:

- SST is sea surface temperature
- Chlor is chlorophyll a
- NAO is the North Atlantic Oscillation
- z_i is a spatial random effect

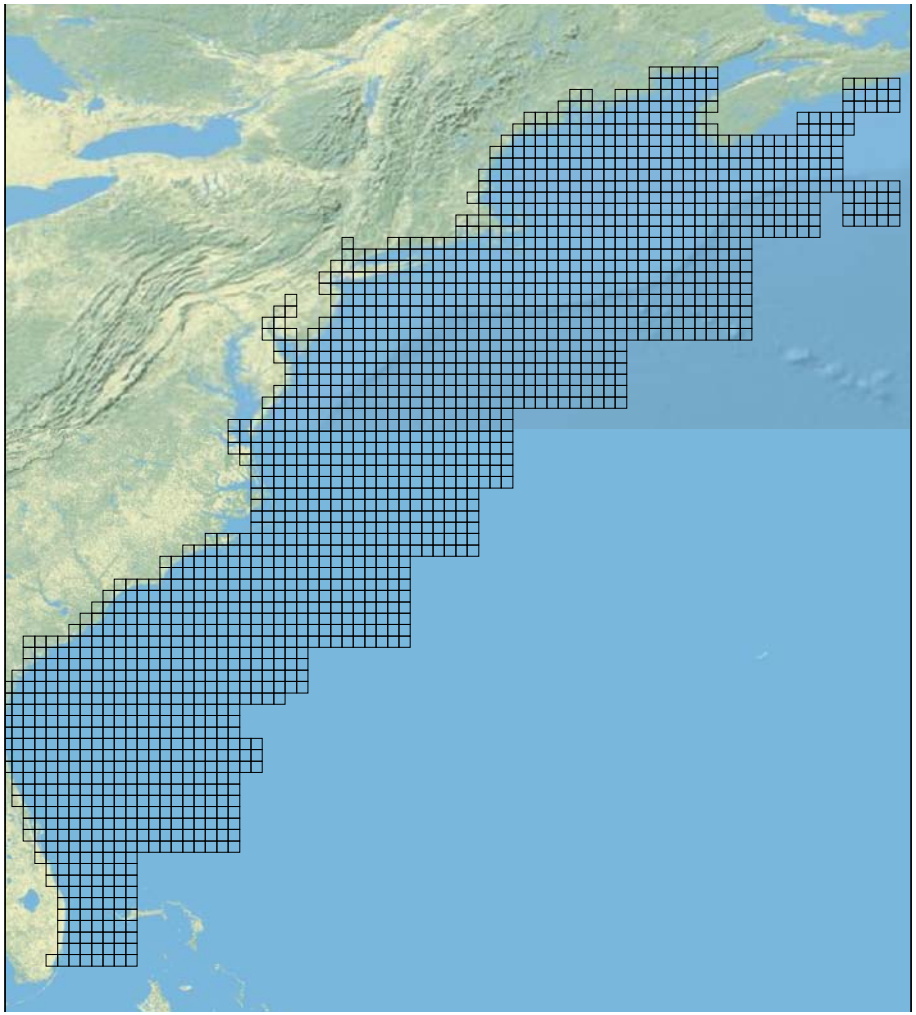
Spatial Random Effect

- Conditional autoregression (CAR)

$$[z_1, z_2, \ldots, z_n] \sim Normal(0, \sigma^2 (\mathrm{I} - \rho \mathrm{W})^{-1})$$

- Useful when data have a discrete support (e.g. areal measurements – counties, strata, etc.)

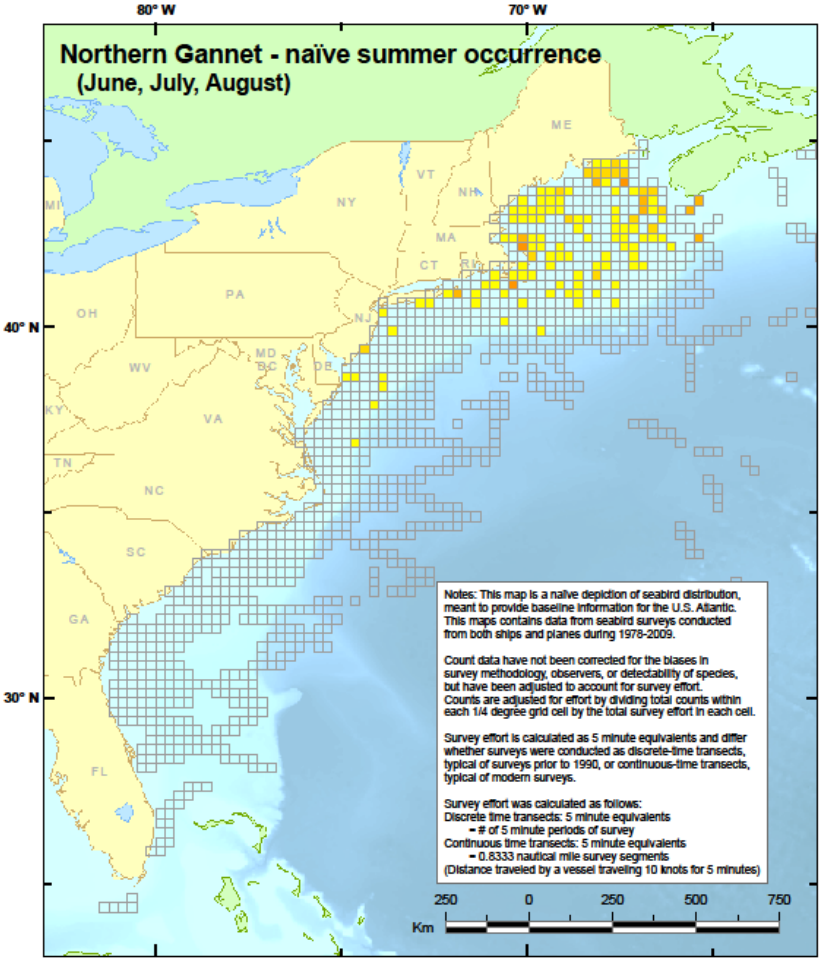
Spatial Random Effect



- ¼ decimal degree size
- 1,909 grid cells
- 1,066 sampled at least once

Bayesian Framework

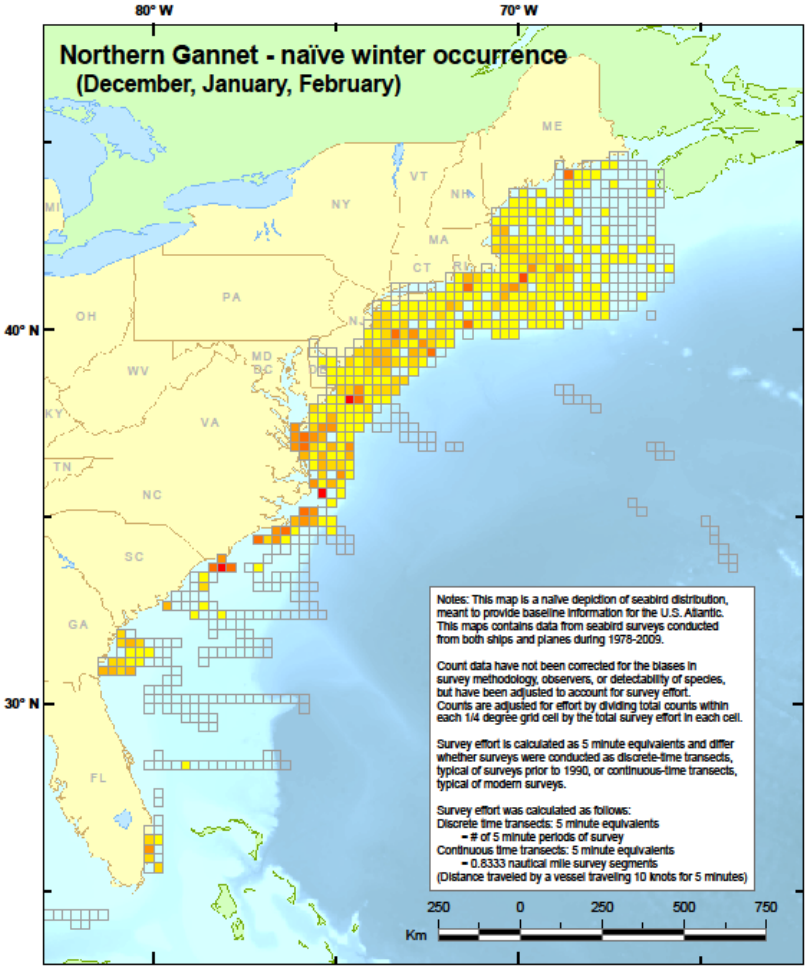
- We used a Bayesian framework for the analysis
  - Used all uninformative priors on the covariate parameters
- Model was implemented in WinBUGS
- We visually checked the chains for convergence and all Rhat values < 1.1



USGS
science for a changing world
A. Gilbert, PWRC, USGS
Map produced: 8/5/2010

Explanation - effort-adjusted counts

0	0.1 - 0.25	0.5 - 1	2 - 5
0 - 0.1	0.25 - 0.5	1 - 2	>5

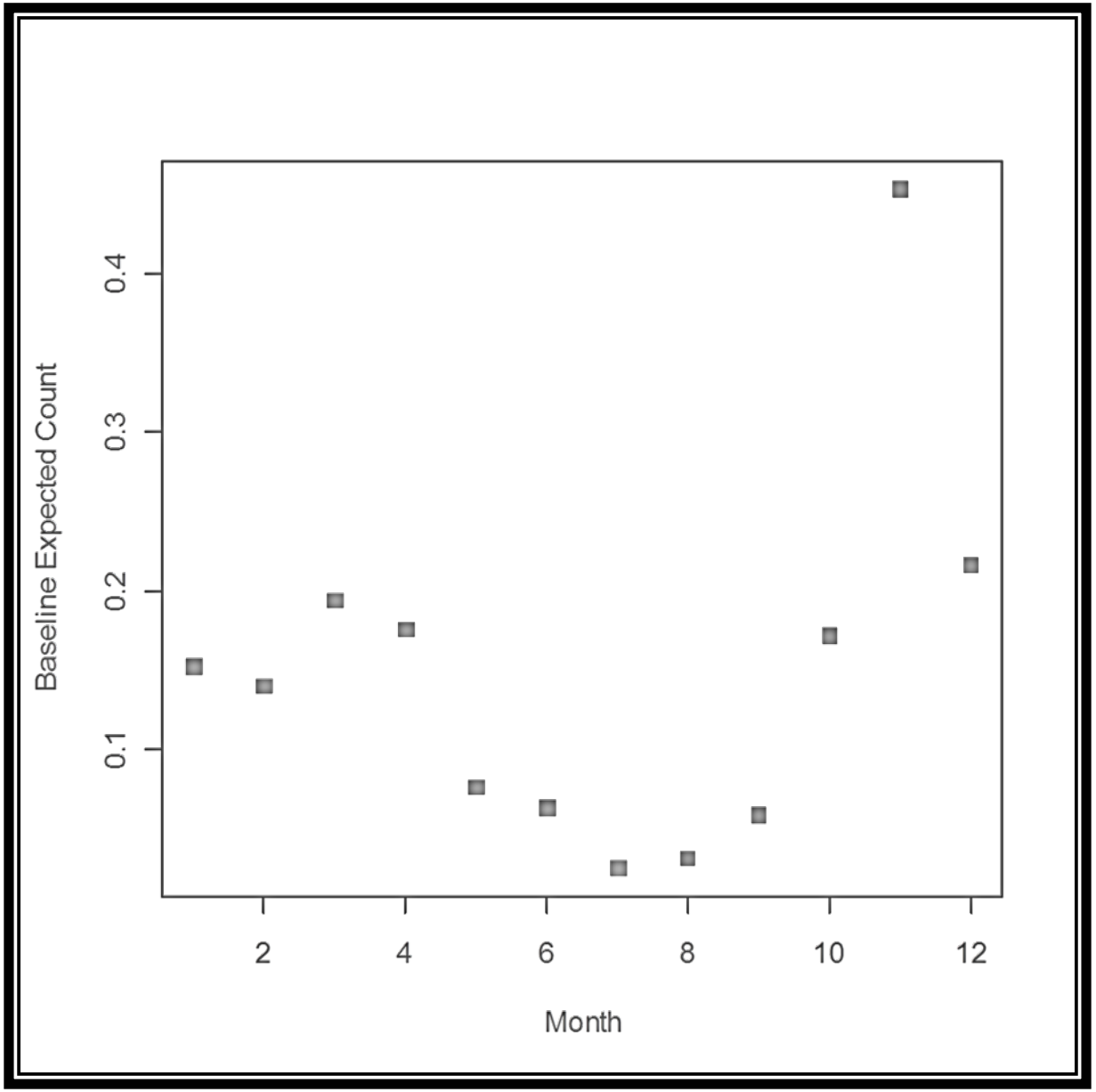


USGS
science for a changing world
A. Gilbert, PWRC, USGS
Map produced: 8/5/2010

Explanation - effort-adjusted counts

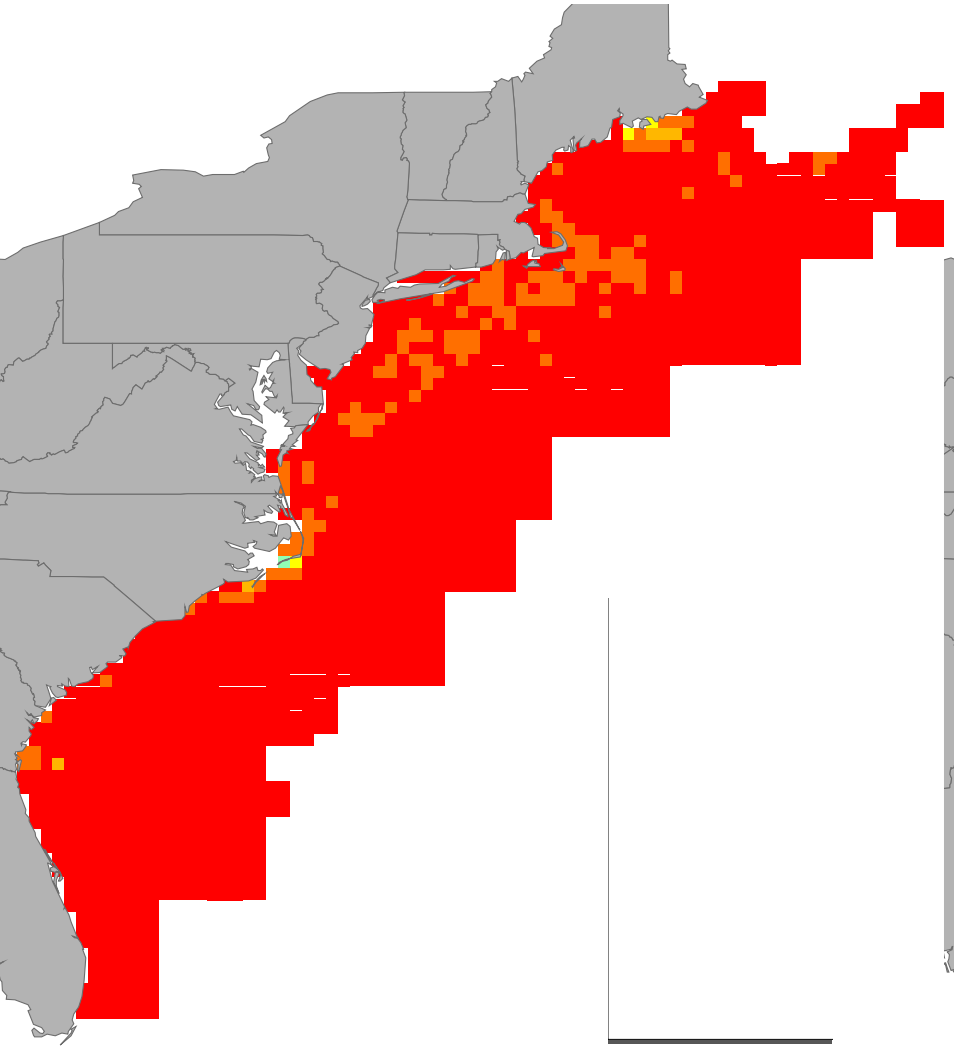
0	0.1 - 0.25	0.5 - 1	2 - 5
0 - 0.1	0.25 - 0.5	1 - 2	5 - 12.4

NOGA – Monthly Effect

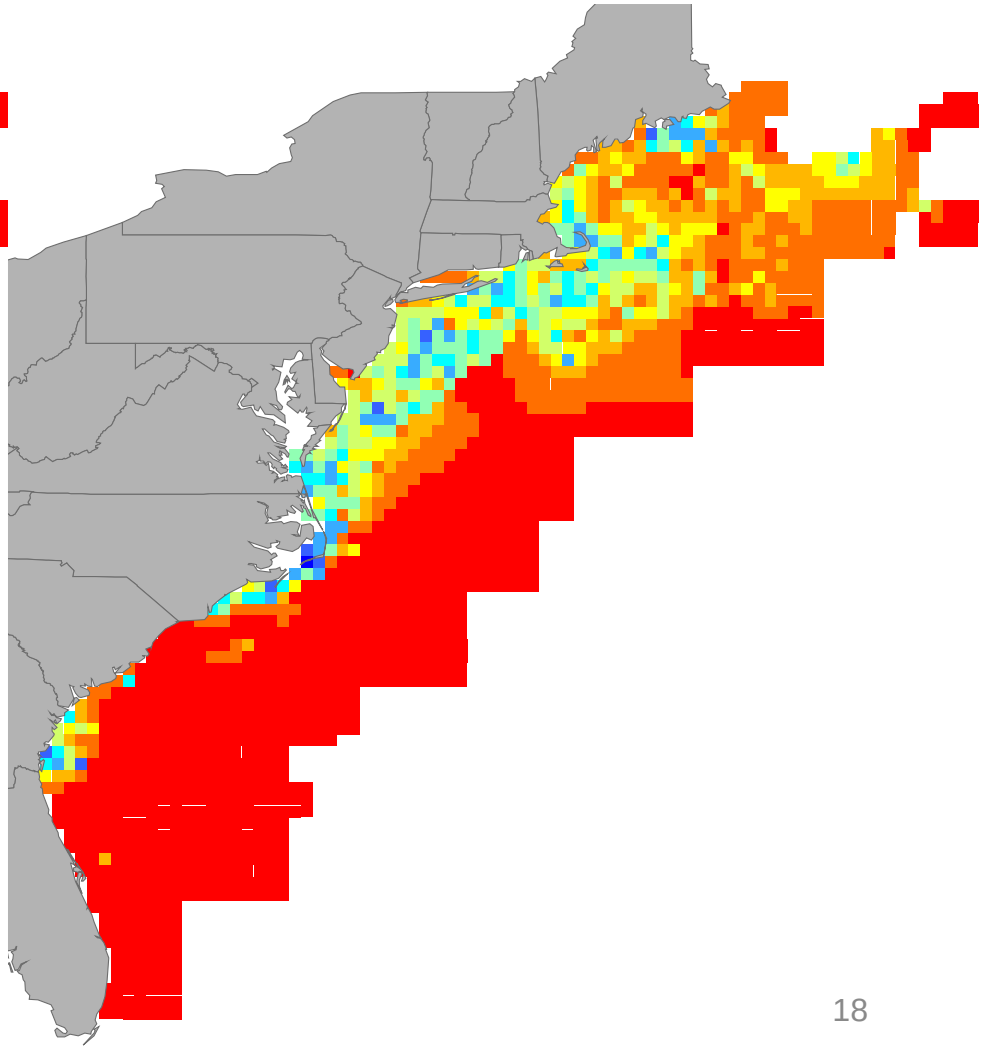


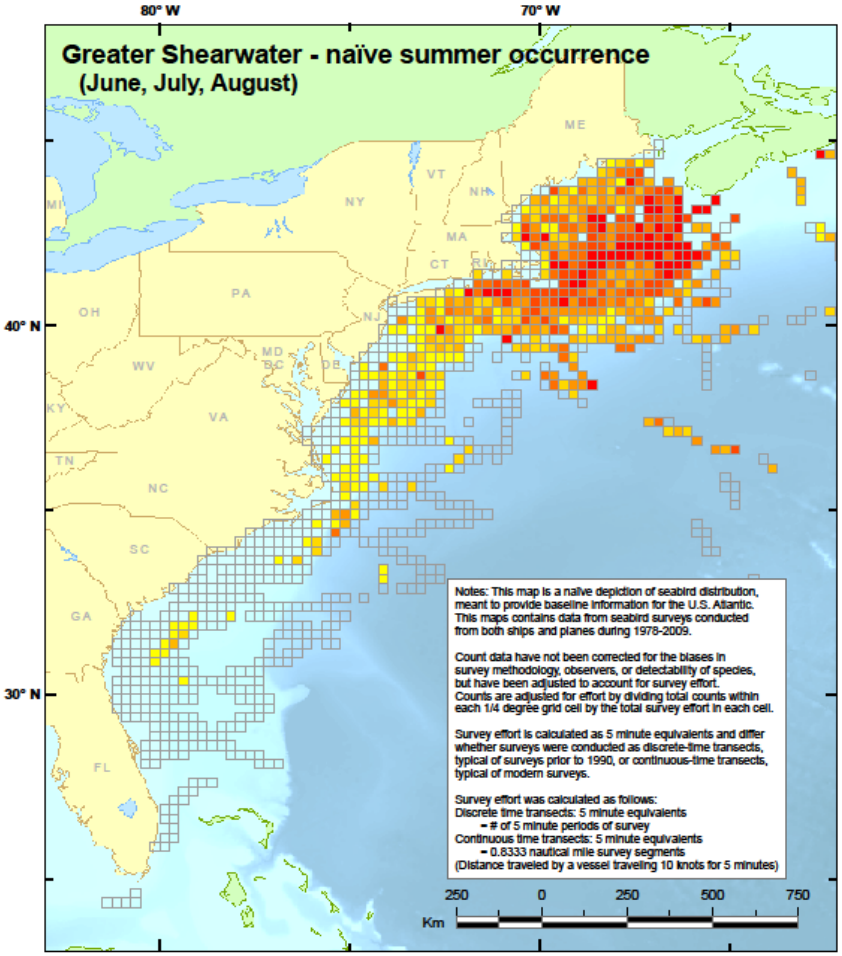
Results – NOGA

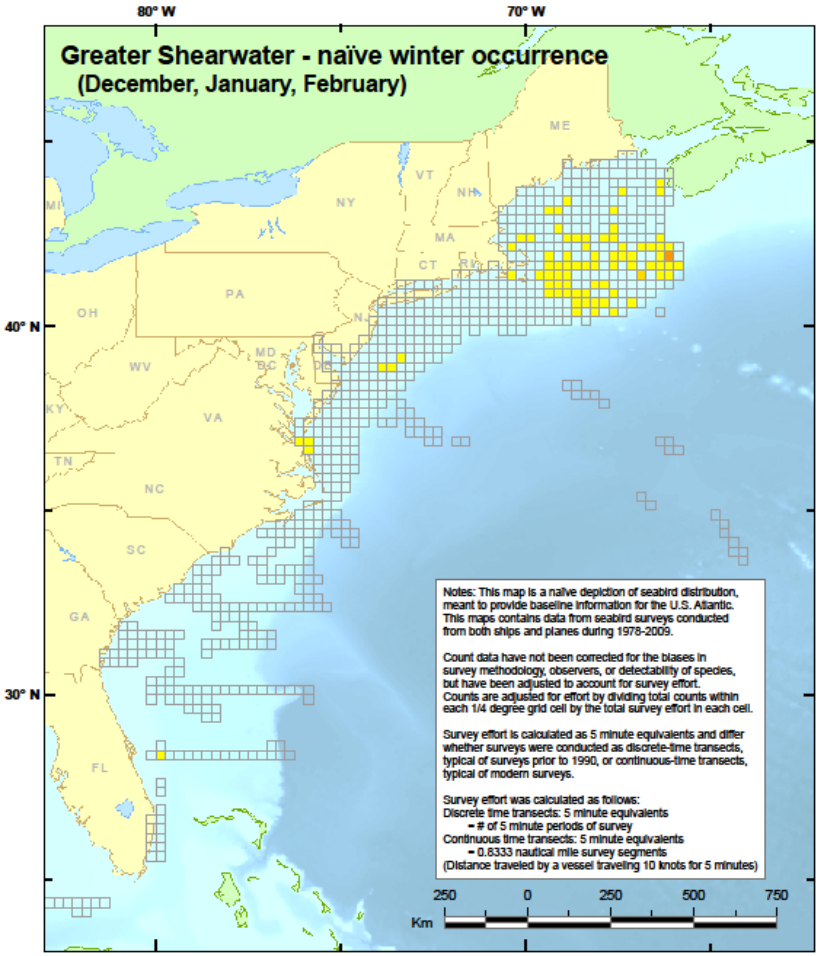
Summer



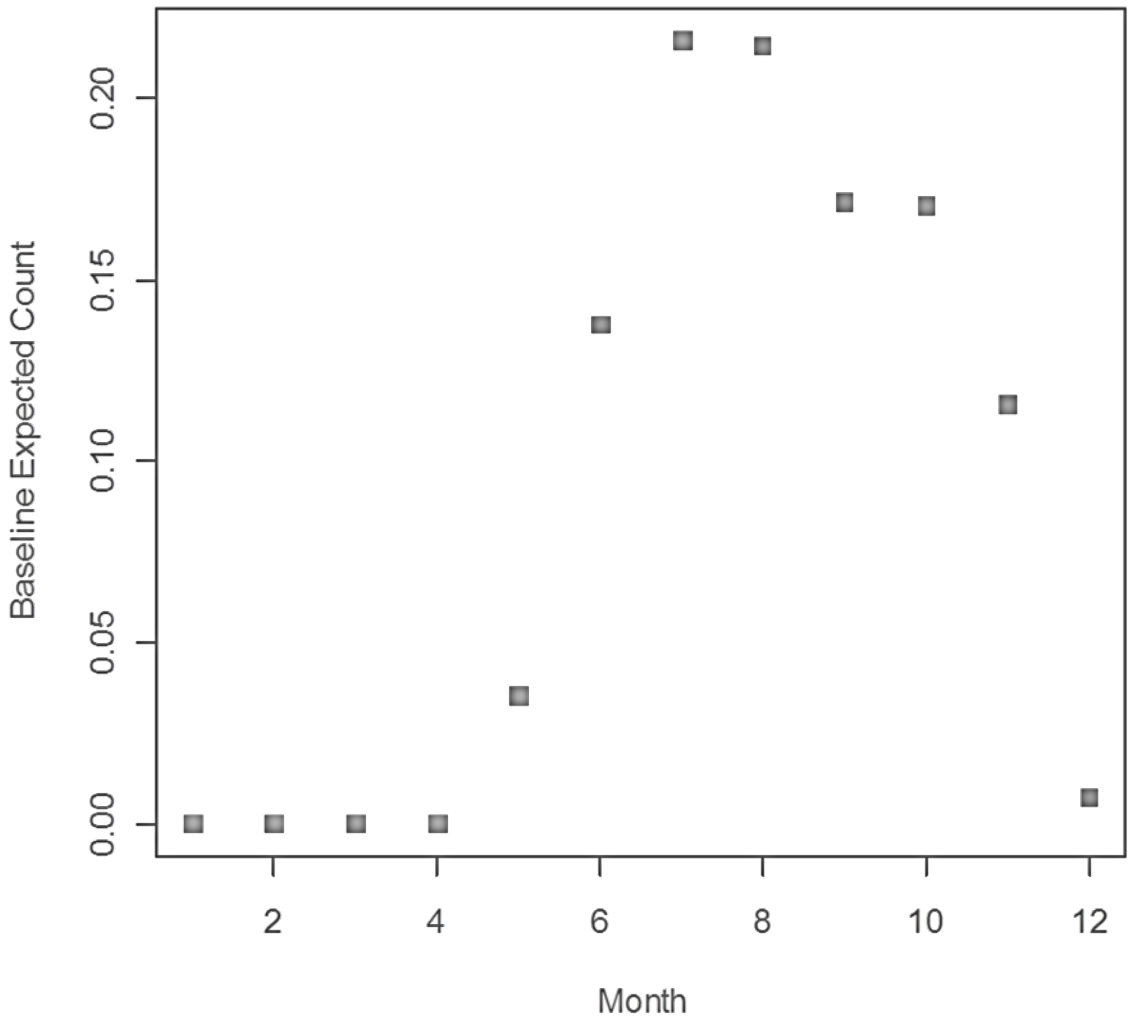
Winter




Greater Shearwater - naïve summer occurrence
(June, July, August)
80° W
70° W
40° N
30° N
ME
VT
NH
NY
MI
MA
CT
RI
PA
OH
NJ
MD
DE
WV
VA
KY
TN
NC
SC
GA
FL
Notes: This map is a naïve depiction of seabird distribution, meant to provide baseline information for the U.S. Atlantic. This maps contains data from seabird surveys conducted from both ships and planes during 1978-2009.
Count data have not been corrected for the biases in survey methodology, observers, or detectability of species, but have been adjusted to account for survey effort. Counts are adjusted for effort by dividing total counts within each 1/4 degree grid cell by the total survey effort in each cell.
Survey effort is calculated as 5 minute equivalents and differ whether surveys were conducted as discrete-time transects, typical of surveys prior to 1990, or continuous-time transects, typical of modern surveys.
Survey effort was calculated as follows:
Discrete time transects: 5 minute equivalents
= # of 5 minute periods of survey
Continuous time transects: 5 minute equivalents
= 0.8333 nautical mile survey segments
(Distance traveled by a vessel traveling 10 knots for 5 minutes)
250 0 250 500 750
Km
USGS
science for a changing world
A. Gilbert, PWRC, USGS
Map produced: 8/5/2010
Explanation - effort-adjusted counts
0
0 - 0.1
0.1 - 0.25
0.25 - 0.5
0.5 - 1
1 - 2
2 - 5
5 - 276.5

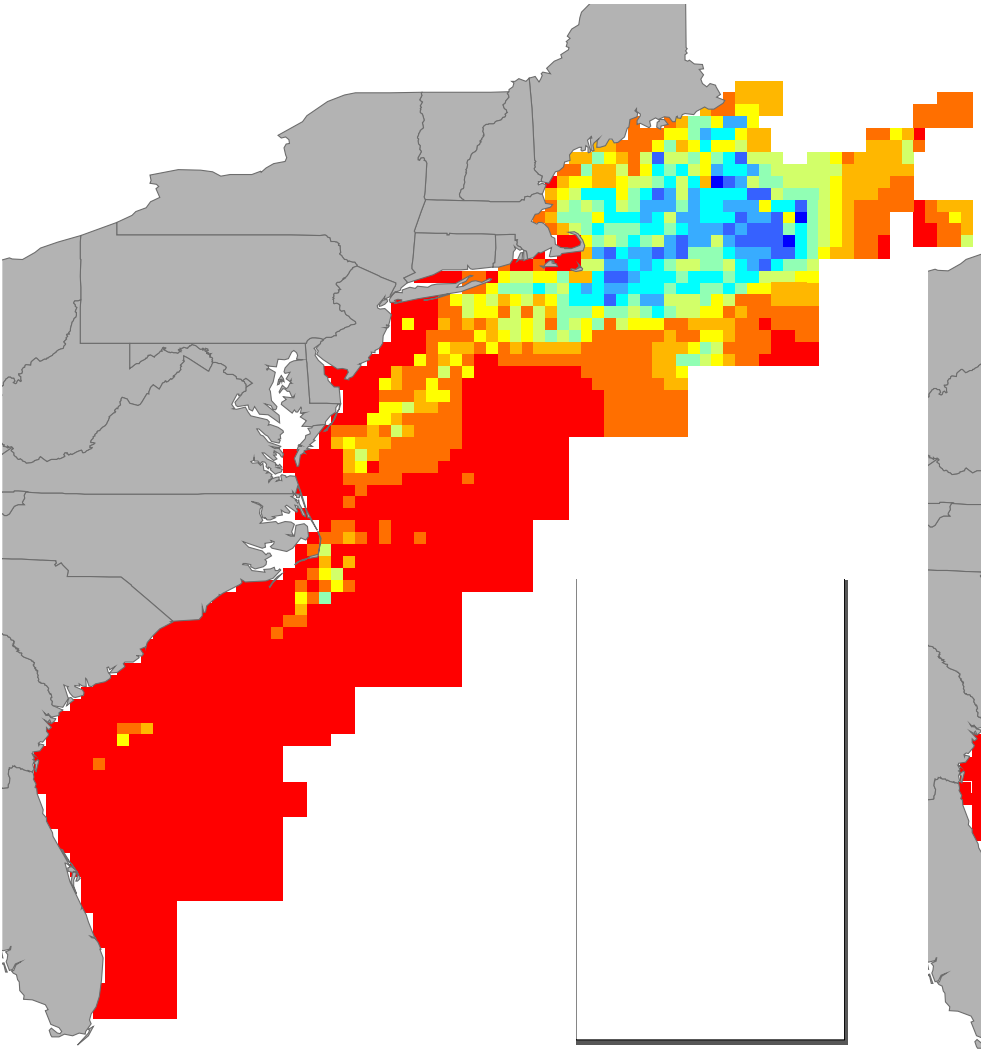

Greater Shearwater - naïve winter occurrence
(December, January, February)
80° W
70° W
40° N
30° N
ME
VT
NH
NY
MI
MA
CT
RI
PA
OH
NJ
MD
DE
WV
VA
KY
TN
NC
SC
GA
FL
Notes: This map is a naïve depiction of seabird distribution, meant to provide baseline information for the U.S. Atlantic. This maps contains data from seabird surveys conducted from both ships and planes during 1978-2009.
Count data have not been corrected for the biases in survey methodology, observers, or detectability of species, but have been adjusted to account for survey effort. Counts are adjusted for effort by dividing total counts within each 1/4 degree grid cell by the total survey effort in each cell.
Survey effort is calculated as 5 minute equivalents and differ whether surveys were conducted as discrete-time transects, typical of surveys prior to 1990, or continuous-time transects, typical of modern surveys.
Survey effort was calculated as follows:
Discrete time transects: 5 minute equivalents
= # of 5 minute periods of survey
Continuous time transects: 5 minute equivalents
= 0.8333 nautical mile survey segments
(Distance traveled by a vessel traveling 10 knots for 5 minutes)
250 0 250 500 750
Km
USGS
science for a changing world
A. Gilbert, PWRC, USGS
Map produced: 8/5/2010
Explanation - effort-adjusted counts
0
0 - 0.1
0.1 - 0.25
0.25 - 0.5
0.5 - 1
1 - 2
2 - 5
>5

GRSH – Monthly Effect

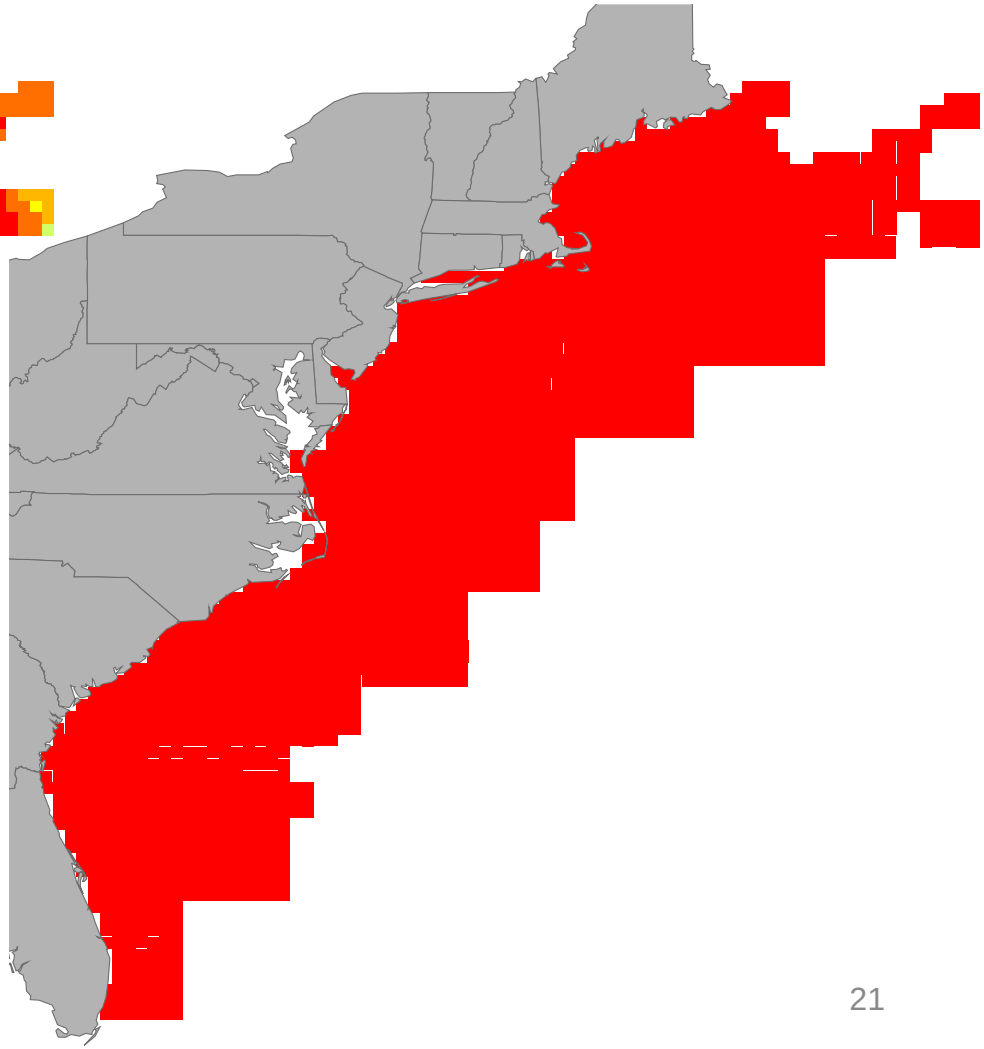


Results – GRSH

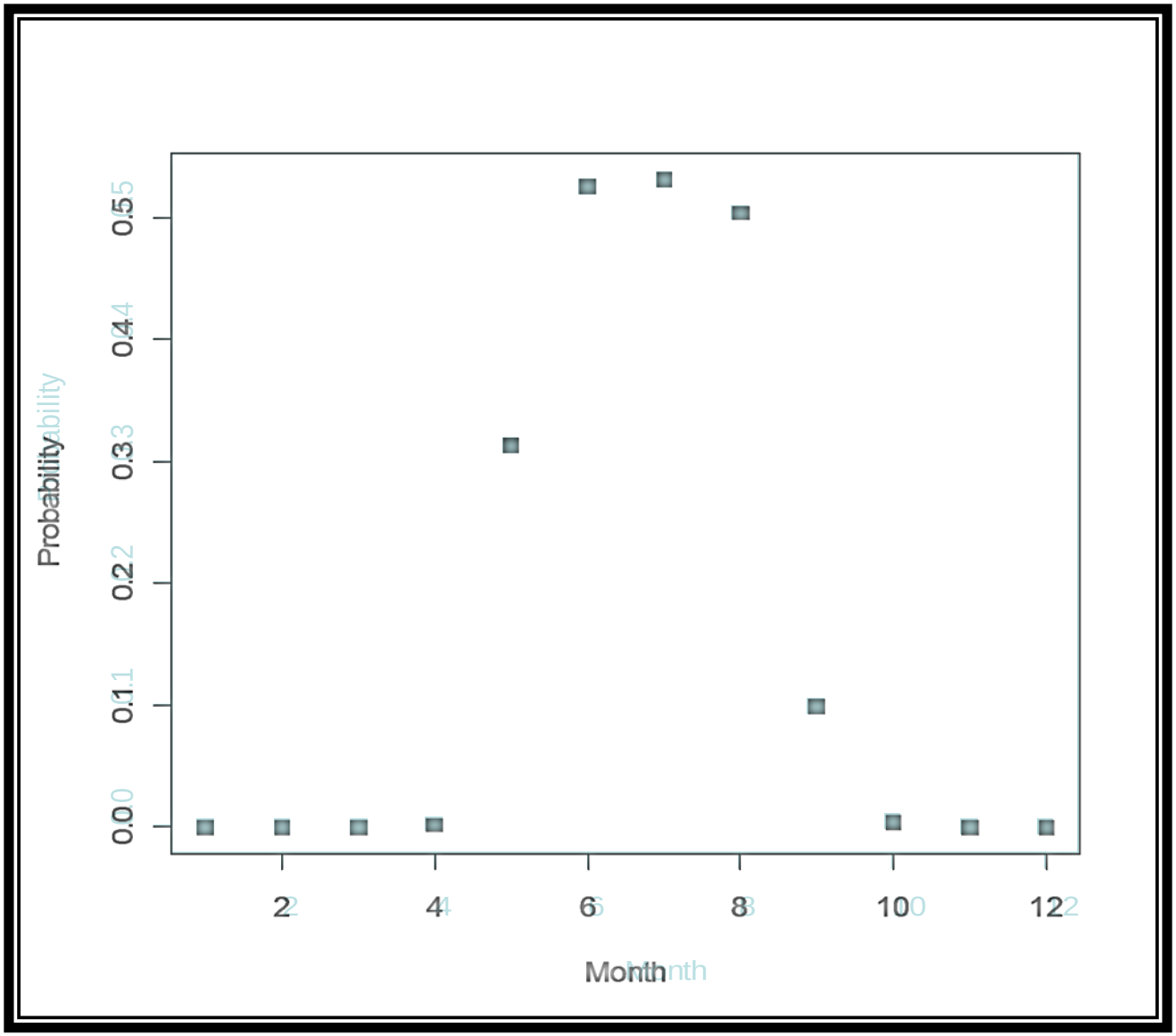
Summer



Winter

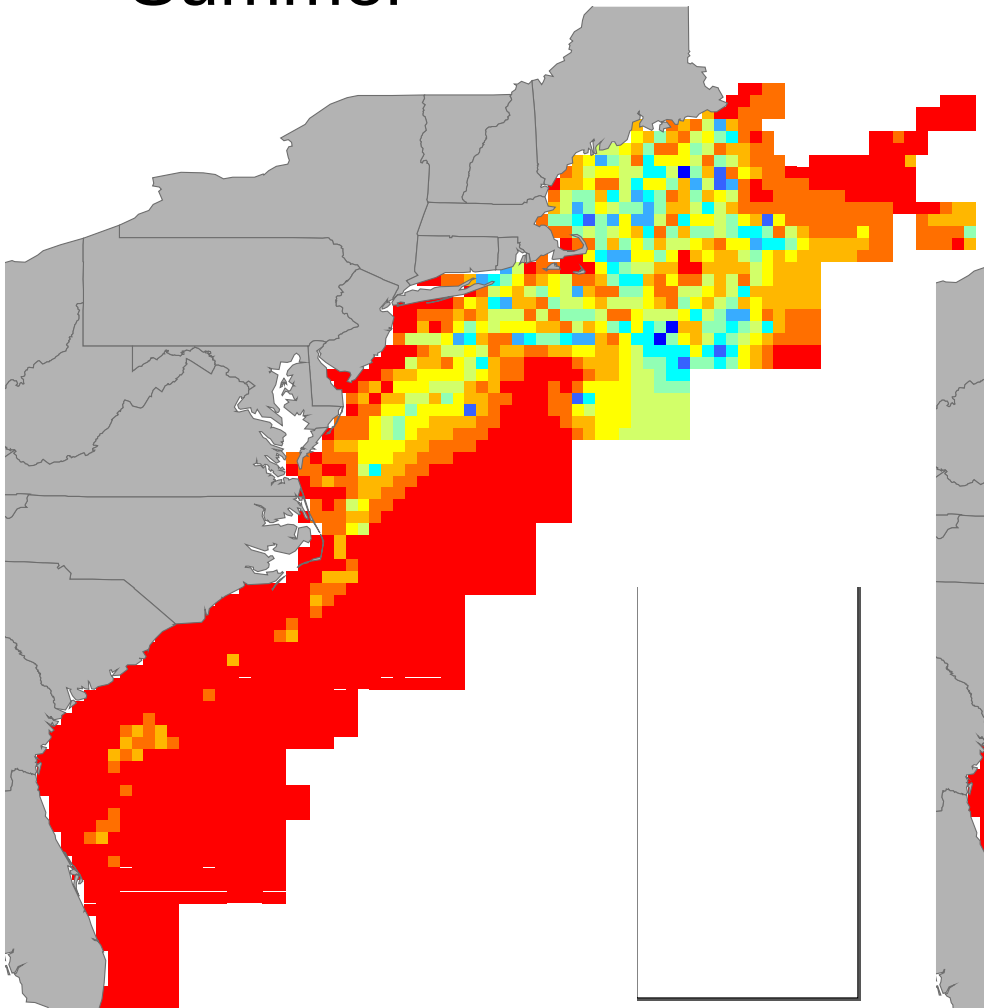


WISP – Monthly Effect

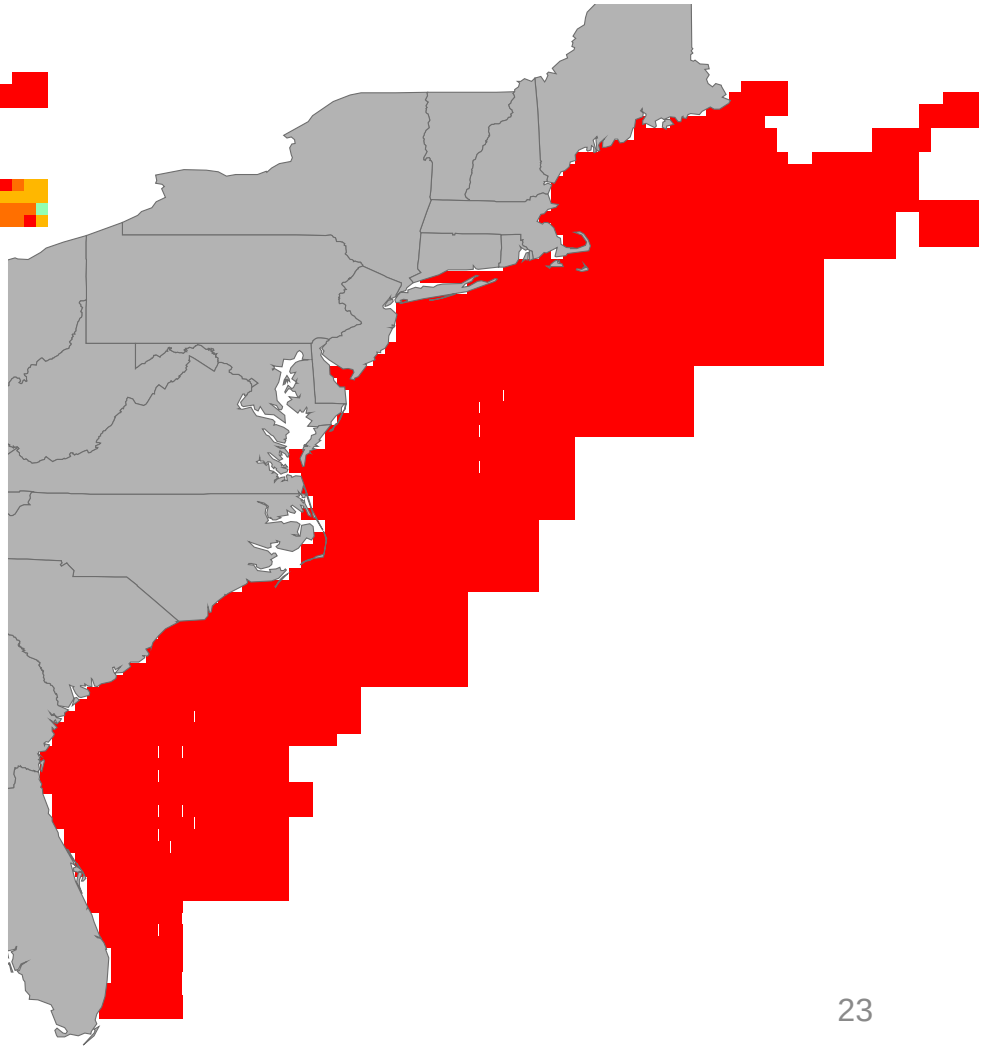


Results – WISP

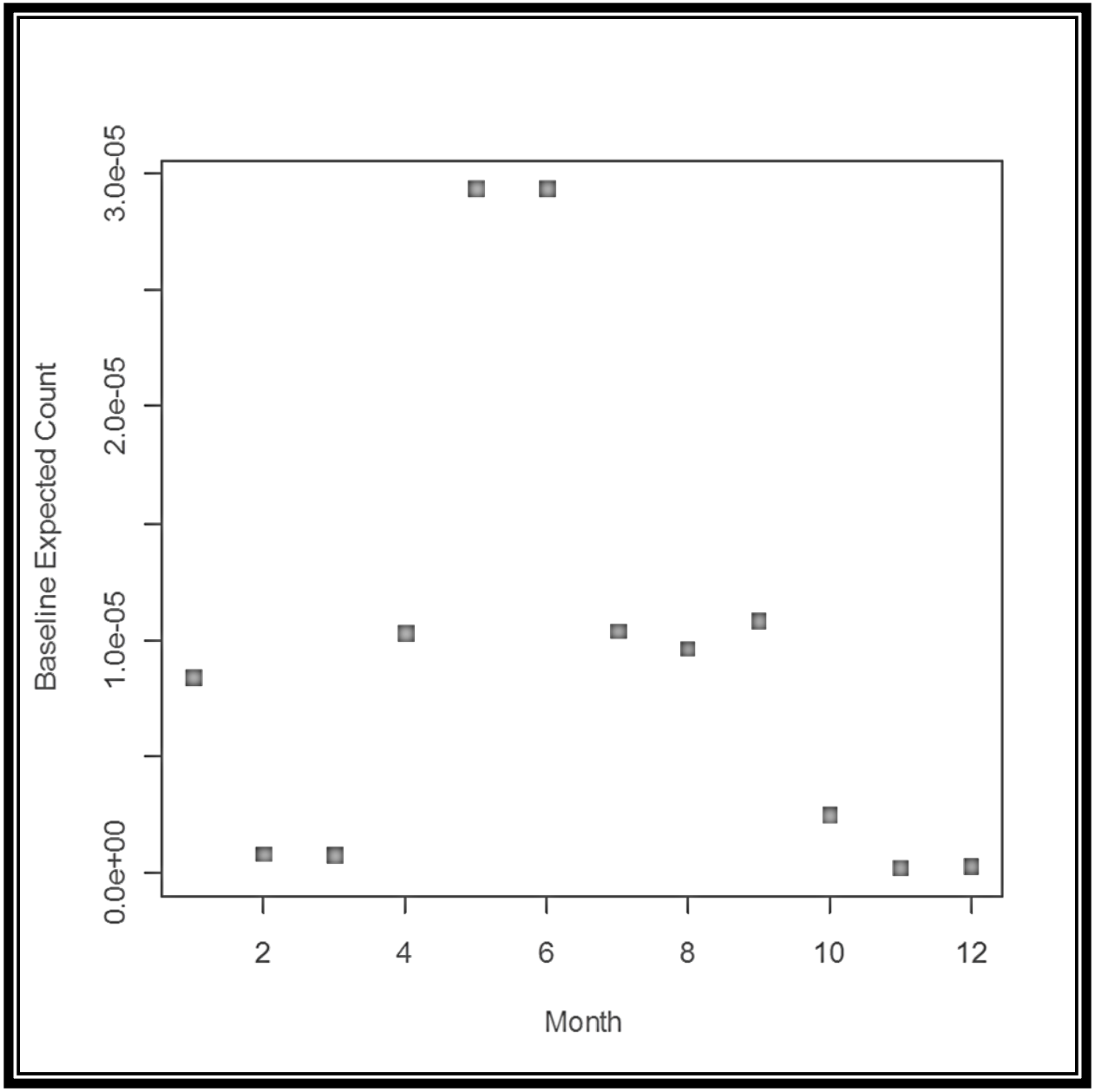
Summer



Winter

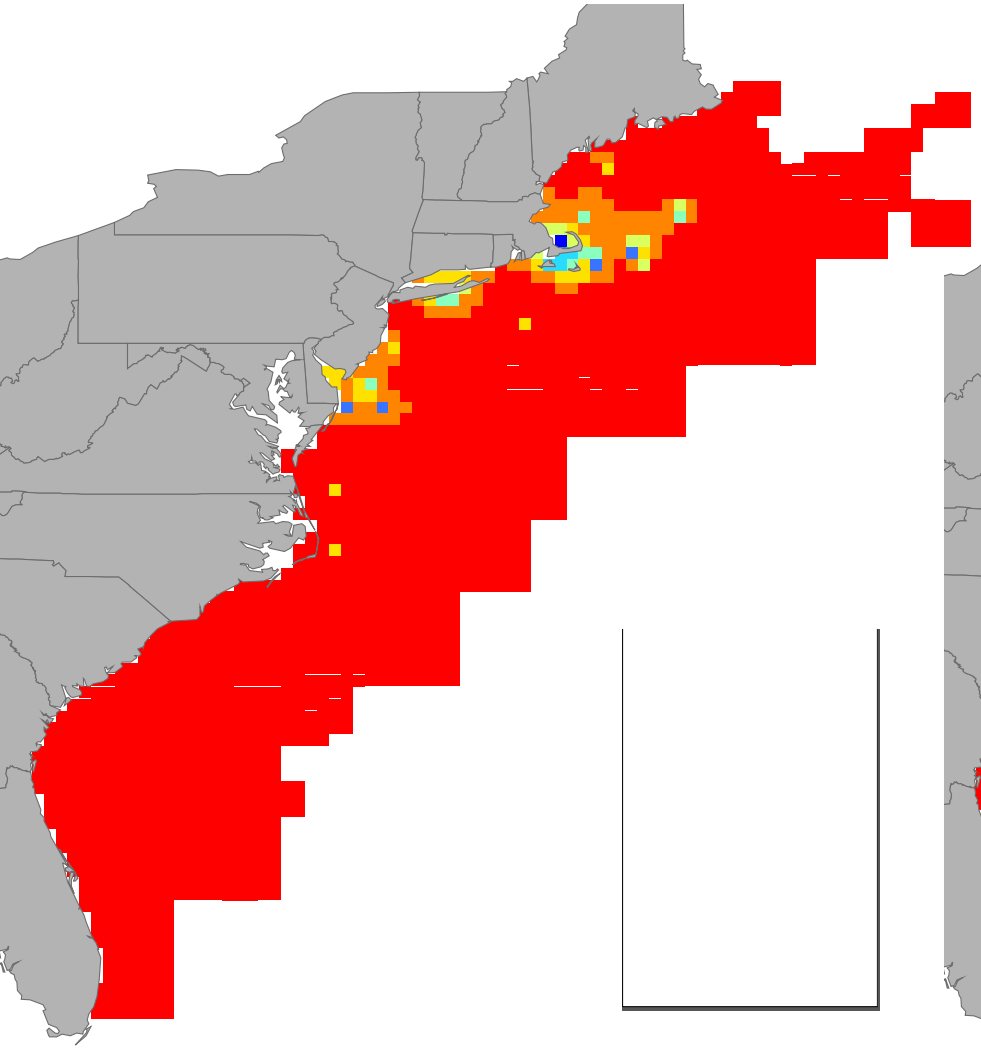


ROST – Monthly Effect

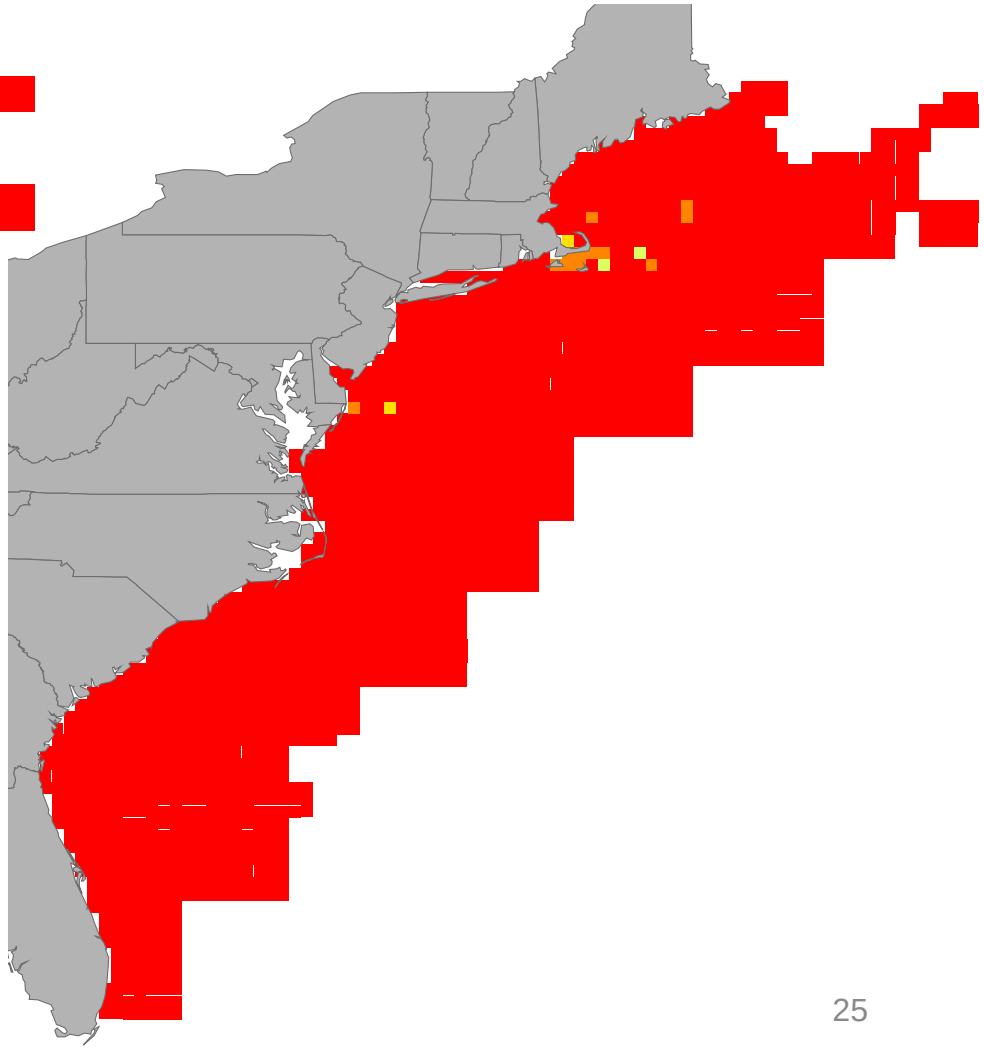


Results – ROST

Summer



Winter



Results – Covariates

	NOGA	GRSH	WISP	ROST
NAO	0.05	0.03	0.04	0.38
SST	-0.77	-0.96	-0.30	2.53
Chlor	0.03	-0.02	-0.03	0.15

* All shown values are different from 0

* We likely need to incorporate SST squared term, spatial by month effect, and possibly NAO lag

Summary

- Modeled the species distribution patterns with explicit spatial correlation and with estimations of uncertainty for 9 species so far

- Provided insight into covariates that are affecting distributional patterns

- Combine individual species maps to determine if there are community patterns

Shortcomings

- General lack of sea bird data collected in a rigorous way across broad areas
- Data are combined over many years from various survey types
- Detection of individuals, or even species, cannot be accounted for with such data

Occupancy Models

- If we look at areas with repeated aerial surveys, we can estimate detection and species richness through the use of site-occupancy models

- This allows us to understand the probability of detecting a species given that it is present

- We expect that detection is very different among species

Massachusetts Audubon

- Nantucket Sound
- Conducted 54 aerial surveys from August 2002 to March 2006
- Multiple transects flown at 500 ft above sea level
- Area of ~1,400 km^2
- 23 sea bird species were observed (there were other shore birds, ducks, etc. that were not included in this analysis)

Nantucket Sound

Community Models

- Create a detection array, **X**, where x_{ijk}=1 if species *i* is detected at site *j* during the *k*th sampling occasion

- We link species-specific parameters through a hierarchical model with community level parameters

Simple Occupancy Models

1. Model x_{ijk} as a Bernoulli random variable:

$$x_{ijk} \sim Bern(p_{ijk}\, z_{ijk})$$

- where p_{ijk} is the detection probability of species *i* at site *j* for the *k*th survey, given that species *i* was present at site *j*
- we included Julian date and effort in the model of detection

Simple Occupancy Models

2. Model occupancy, z_{ijk} such that:

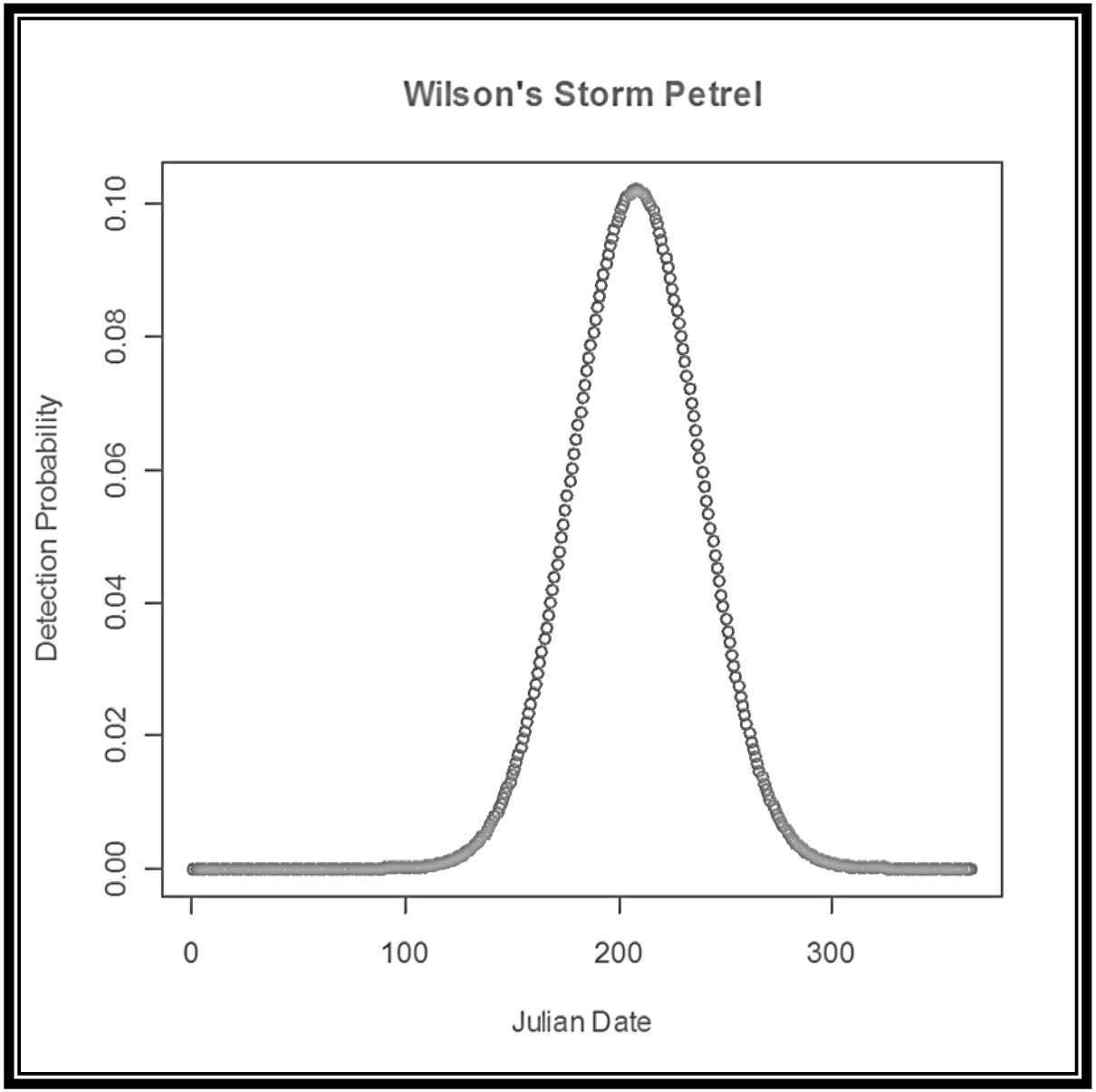
$$z_{ijk} \sim Bern(\Psi_{ij})$$

– where Ψ_{ij} is the probability of occurrence

Bayesian Analysis

- We used a Bayesian framework for the analysis
  - Used all uninformative priors on the covariate parameters
- Model was implemented in WinBUGS
- We visually checked the chains for convergence and all Rhat values

Detectability



Detectability

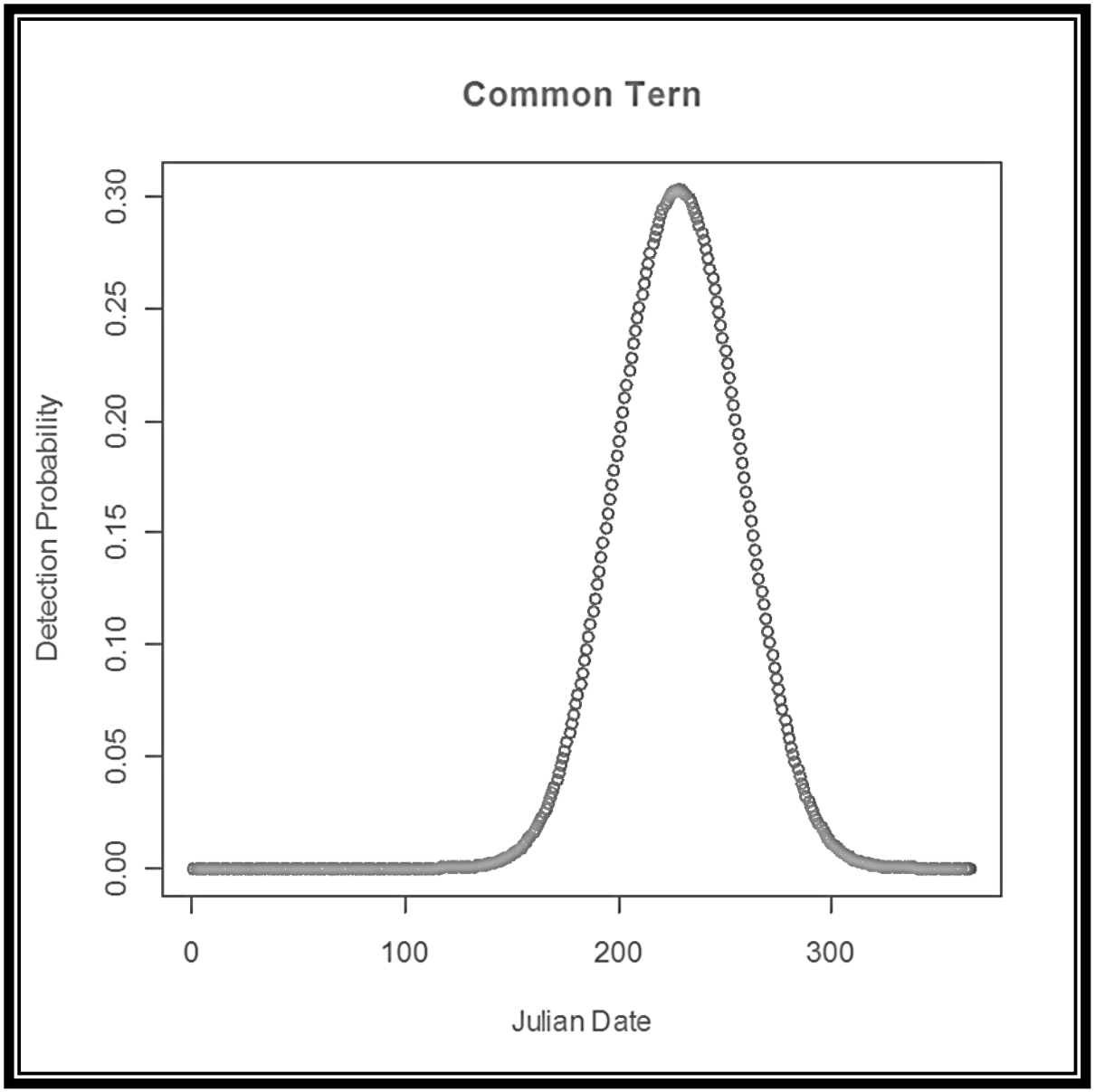


Photo by Andreas Trepte

Detectability

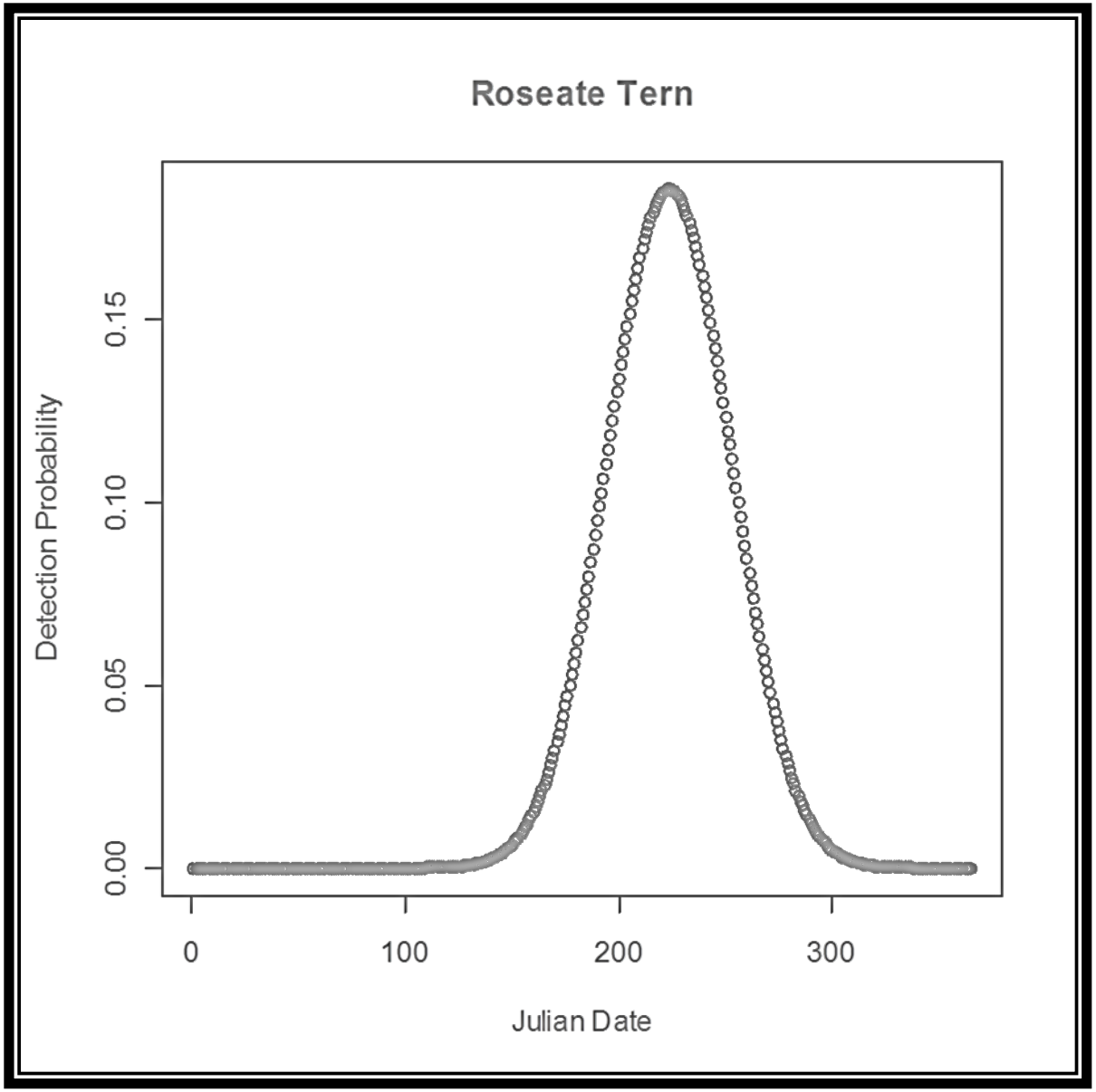


Photo by stfbfc

Detectability

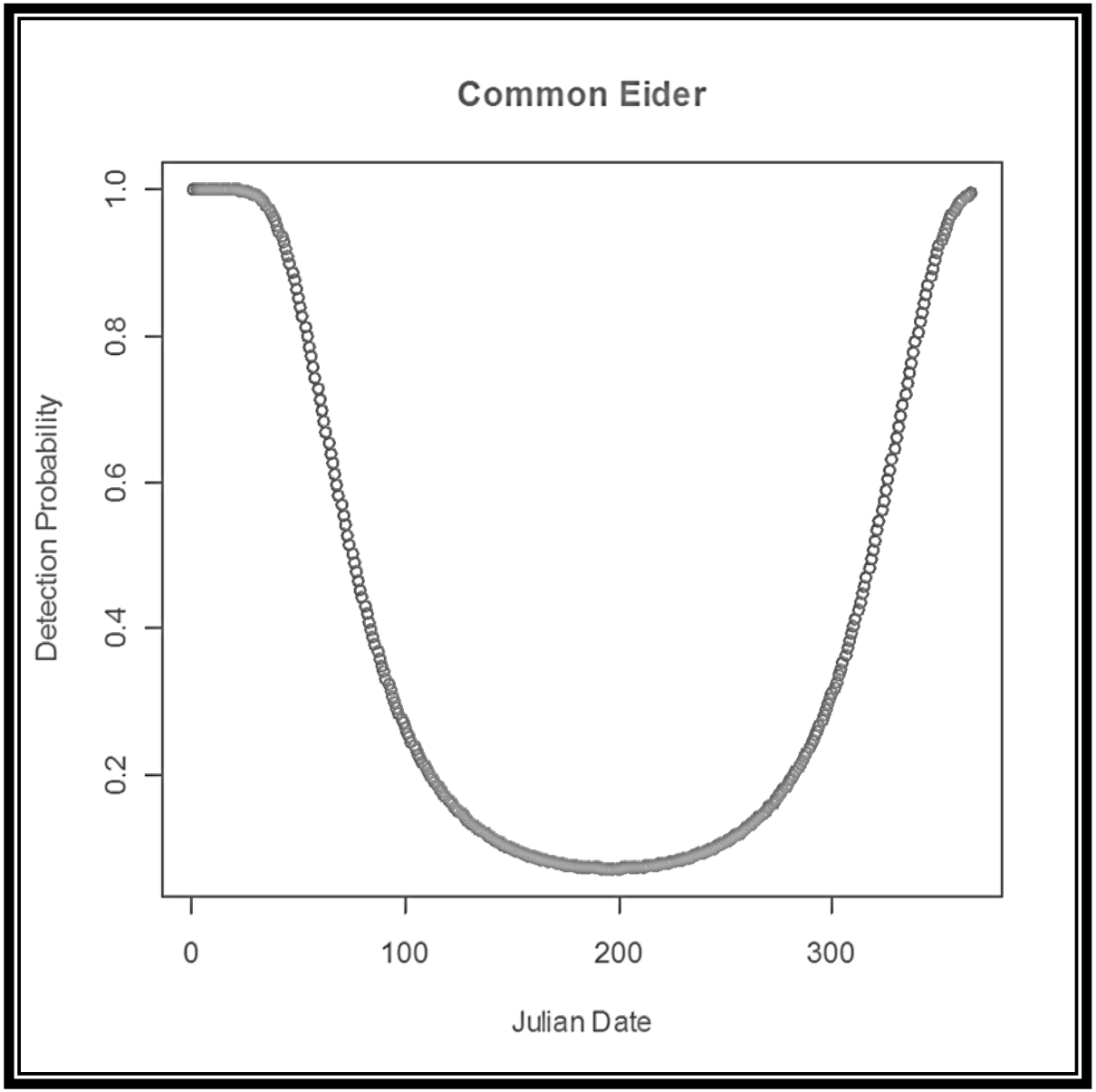


Photo by Paul Joyce

Species Richness

Summary

- Detection was very different by species
- Detection varied by Julian date for all species
- Occupancy was different by species and with more data, we can model by season or date as well

Implications

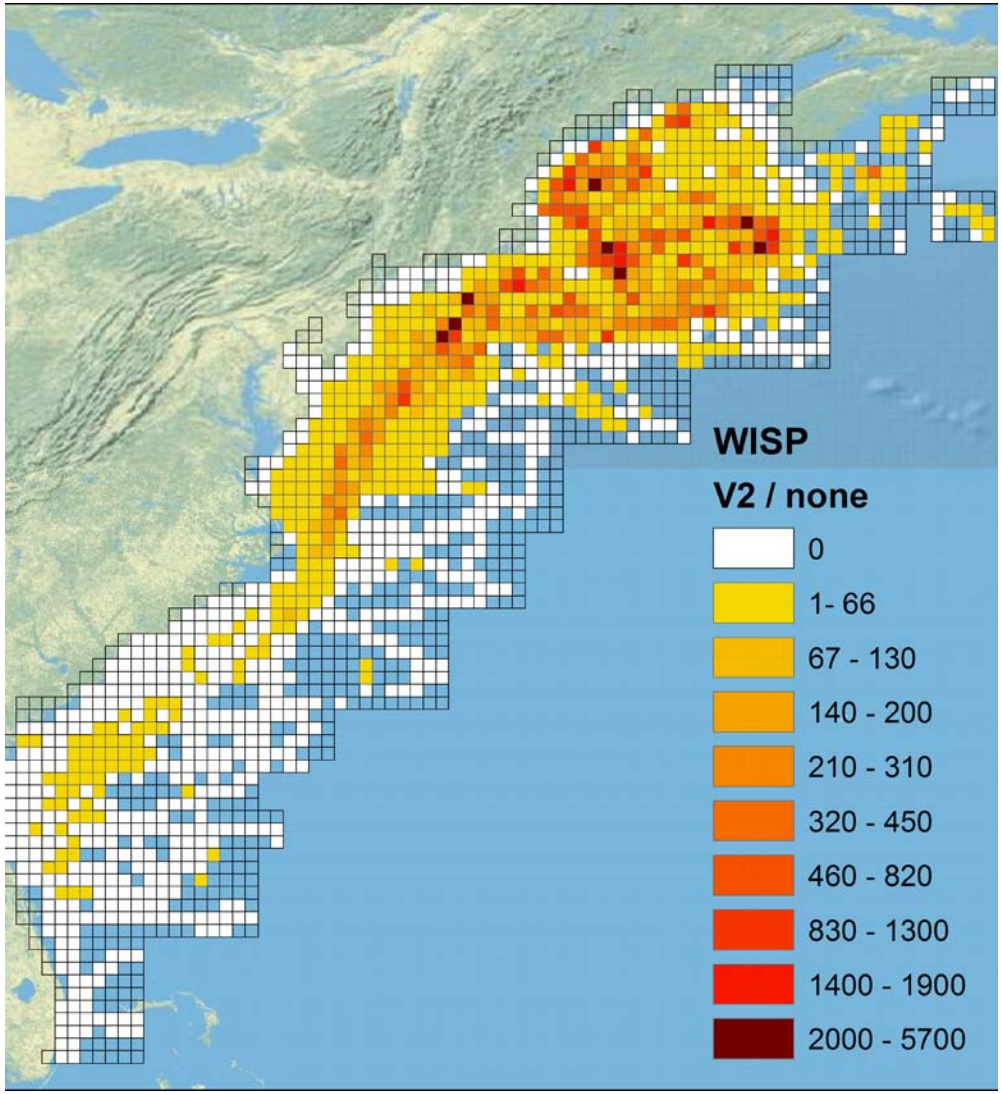
- Results indicate that some species, like Roseate Terns, maybe difficult to detect from aerial surveys.

- Here, in order to have a 95% chance of seeing Roseate Terns when they occupy an area, you would need to conduct at least 15 repeated surveys.

Summary

- More data to improve models at the broad scale
- Detection of species by different survey types is important to estimating occupancy and hence abundance
- More work to be done in combining such information to improve our understanding of sea bird dynamics

Acknowledgments

- Bureau of Ocean Energy Management, Regulation and Enforcement
- U.S. Fish and Wildlife Service, Division of Migratory Birds
- U.S. Geological Survey

Wilson’s Storm-petrel Count