

# MODELING SEABIRD DISTRIBUTIONS OFF THE PACIFIC COAST OF WASHINGTON

## FINAL REPORT

June 2015

### Prepared for:

Washington State Department of Natural Resources  
111 Washington Street, SE  
Olympia, WA 98504-7027

### Contributors:

Charles Menza<sup>1</sup>, Jeffery Leirness<sup>1</sup>, Tim White<sup>1</sup>, Arliss Winship<sup>1</sup>, Brian Kinlan<sup>1</sup>, Jeannette E. Zamon<sup>2</sup>, Lisa Ballance<sup>3</sup>, Elizabeth Becker<sup>3</sup>, Karin Forney<sup>3</sup>, Josh Adams<sup>4</sup>, David Pereksta<sup>5</sup>, Scott Pearson<sup>6</sup>, John Pierce<sup>6</sup>, Liam Antrim<sup>7</sup>, Nancy Wright<sup>7</sup>, and Ed Bowlby<sup>7</sup>

### Address for Correspondence:

NOAA, National Centers for Coastal Ocean Science  
1305 East West Highway, SSMC IV  
Silver Spring, MD 20910

### Deliverable under:

NOS Agreement Code: MOA-2013-038 (Annex #002)/8963



## **Contributor Affiliations**

<sup>1</sup> National Centers for Coastal Ocean Science, NOAA/NOS

<sup>2</sup> Point Adams Research Station, NOAA/NMFS/NWFSC

<sup>3</sup> Marine Mammal and Turtle Division, NOAA/NMFS/SWFSC

<sup>4</sup> Pacific Coastal and Marine Science Center, USGS/WERC

<sup>5</sup> Pacific OCS Region, Bureau of Ocean Energy Management

<sup>6</sup> Wildlife Science Division, Washington Department of Fish and Wildlife

<sup>7</sup> Olympic Coast National Marine Sanctuary, NOAA/NOS

## **Acknowledgements**

Many scientists contributed to the development of this report and the corresponding seabird models it describes. Contributors provided seabird sighting data and guidance in seabird ecology and predictive modeling. Together, the team was able to merge data and disciplines, creating new information that no single organization could accomplish alone. In addition to the contributors identified on the front page of this report, we received valuable advice on climate indices from Bill Peterson at the Northwest Fisheries Science Center, and we received the frequency of chlorophyll peaks index data set from Rob Suryan at Oregon State University.

# Contents

About this report .....	1
Summary .....	2
Introduction.....	2
Model Development.....	3
Model Results .....	16
Discussion .....	38
References .....	40
Appendix A: Seabird survey program descriptions .....	43
Appendix B: Processing steps for environmental predictors .....	44
Appendix C: Model performance metrics (full model assemblage).....	47
Appendix D: Select marginal and residual plots.....	48
Appendix E: Variable importance figures .....	58

## About this report

The following report is a deliverable for contract MOA-2013-038 (Annex #002)/8963 between the Washington State Department of Natural Resources and the National Centers for Coastal Ocean Science (NCCOS). Two additional reports under this contract will provide the state with 1) information to prioritize seafloor mapping and 2) an assessment of marine mammal datasets available within coastal and offshore waters (Kracker and Menza 2015). A preceding report, prepared under a different contract by NCCOS for the state, provided a summary of existing seabird datasets (Menza et al. 2014).

NCCOS science provides coastal managers the information and tools they need to balance society's environmental, social, and economic goals. NCCOS is the primary coastal science arm within the National Oceanic and Atmospheric Administration (NOAA) National Ocean Service. NCCOS works directly with managers, industry, regulators, and scientists to deliver relevant, timely, and accurate scientific information and tools.

This report supports the NOAA Coastal Zone Management Program, a voluntary partnership between the federal government and U.S. coastal and Great Lakes states and territories authorized by the Coastal Zone Management Act (CZMA) of 1972 to address national coastal issues. The act provides the basis for protecting, restoring, and responsibly developing our nation's diverse coastal communities and resources. To meet the goals of the CZMA, the national program takes a comprehensive approach to coastal resource management—balancing the often competing and occasionally conflicting demands of coastal resource use, economic development, and conservation. A wide range of issues are addressed through the program, including coastal development, water quality, public access, habitat protection, energy facility siting, ocean governance and planning, coastal hazards, and climate change. Accurate maps of seabird distributions are an important tool for making informed management decisions that affect all of these issues.

For more information contact:

Charles Menza

National Centers for Coastal Ocean Science

Center for Coastal Monitoring and Assessment

(301) 713-3028 x107

[charles.menza@noaa.gov](mailto:charles.menza@noaa.gov)



## Summary

This report presents seasonal distribution maps of selected seabird species off the Pacific coast of Washington. Maps were created to support state-led marine spatial planning by identifying ecologically important areas for seabirds. Seabird distribution maps were constructed by predicting relative density across the seascape using associative models linking at-sea seabird observations with environmental covariates. Seabird observations were compiled from federal and state monitoring programs with data between 2000 and 2013. Environmental variables were processed from long-term archival satellite, oceanographic, and hydrographic databases. All models show good to excellent performance based on model performance diagnostics and expert review. Although this work was completed to support marine spatial planning by the state of Washington, these data will benefit other organizations and other purposes including assessments of marine sanctuary condition, ecosystem health, coastal hazard impacts, and climate change.

## Introduction

Washington depends on a healthy coastal and marine ecosystem to maintain a thriving economy and vibrant communities. Marine ecosystems support critical habitats for wildlife and a growing number of ocean activities, such as fishing, transportation, aquaculture, recreation, and energy production. Planners, policy makers, and resource managers are being challenged to sustainably balance multiple ocean uses and environmental conservation mandates in a finite space and with limited information. This balancing act can and should be supported by spatial planning.

Marine spatial planning (MSP) is a planning process that enables integrated, forward-looking, and consistent decision making on the human uses of the oceans and coasts (Ehler and Douvère 2009). It can improve marine resource management by planning for human uses in locations that reduce conflict among different activities, and supports a balance among social, economic, and ecological benefits received from ocean resources.

In March 2010, the Washington state legislature enacted a marine spatial planning law (RCW §43.372) to address resource use conflicts in waters off Washington. In 2011, a report to the legislature and a workshop on human use data provided guidance for the marine spatial planning process. In 2012, the governor amended the existing law to focus funding on mapping and ecosystem assessments for Washington's Pacific coast and the legislature provided \$3.7 million in the FY 13-15 biennium to begin marine spatial planning off Washington's coast. The funds were appropriated through the Washington Department of Natural Resources Marine Resources Stewardship Account with coordination among the State Ocean Caucus, the four Coastal Treaty Tribes, four coastal Marine Resource Committees, and the Washington Coastal Marine Advisory Council.

Seabirds are a conspicuous and ecologically important component of the marine ecosystem off Washington. They are typically long-lived, move over broad spatial ranges, and feed at a variety of trophic levels (Schreiber and Burger 2001). As such, they are responsive to changes in the marine and coastal environment, and can be useful indicators of environmental change. Seabirds are also important to coastal economies and provide direct eco-tourism benefits to coastal communities through recreational bird-watching opportunities (U.S. Fish and Wildlife Service 2013).

Many seabird populations have decreased in recent decades (Palczny et al. 2015). They are threatened by impacts of various ocean activities such as coastal development, fishing, resource extraction, and renewable energy development (Croxall et al. 2012). Given changes in population numbers and their economic and ecological importance, all seabird species off Washington are subject to conservation requirements under the Migratory Bird Treaty Act and some under the U.S. Endangered Species Act. Other species are listed in Washington State's list of species of concern (Washington Department of Fish and Wildlife 2015). To minimize the potential for adverse effects of ocean uses on seabirds, coastal zone managers need information on seabird distribution, abundance, and movements.

# Model Development

## Species selected for modeling

This report focuses on developing maps for seven seabird species (Table 1): Marbled Murrelet (*Brachyramphus marmoratus*), Tufted Puffin (*Fratercula cirrhata*), Common Murre (*Uria aalge*), Black-footed Albatross (*Phoebastria nigripes*), Northern Fulmar (*Fulmarus glacialis*), Pink-footed Shearwater (*Puffinus creatopus*) and Sooty Shearwater (*Puffinus griseus*). The seven species were chosen by the Washington Department of Ecology and the Washington Department of Fish and Wildlife (WDFW), because they were species of management concern or representative of different seabird life history patterns. The Marbled Murrelet is listed as threatened by the state and U.S., the Tufted Puffin is listed as endangered by the state, and the Common Murre is listed as a candidate species of concern by the state (Washington Department of Fish and Wildlife 2015, U.S. Fish and Wildlife Service 2005). The state requested models for Short-tailed Albatross as well, but we were unable to model the species due to insufficient sightings. Important habitats, life history patterns, population changes over time, and threats for each of the modeled species are well referenced by Schreiber and Burger (2001) and U.S. Fish and Wildlife Service (2005). In addition, Washington has several active research programs that examine trends in seabird populations (<http://wdfw.wa.gov/conservation/research/projects/seabird/>).

Table 1: Seabird species chosen for modeling. Segments represent modeling analysis units and are discussed more fully in the Seabird observation processing section. Typically segments are 3 kilometer long transects.

Common name	Scientific name	Family	Species code	Number of segments with sightings		
				Total	April to October	November to March
Marbled Murrelet	<i>Brachyramphus marmoratus</i>	Alcidae	mamu	1,632	1,625	7
Tufted Puffin	<i>Fratercula cirrhata</i>	Alcidae	tupu	1,744	1,738	6
Common Murre	<i>Uria aalge</i>	Alcidae	comu	6,938	6,533	405
Black-footed Albatross	<i>Phoebastria nigripes</i>	Diomedidae	bfal	508	421	87
Northern Fulmar	<i>Fulmarus glacialis</i>	Procellariidae	nofu	475	463	12
Pink-footed Shearwater	<i>Puffinus creatopus</i>	Procellariidae	pfsh	611	611	0
Sooty Shearwater	<i>Puffinus griseus</i>	Procellariidae	sosh	2,611	2,586	25

## Study area

The geographic scope of this study extends north to south from Cape Flattery to Cape Disappointment, and east to west from the Pacific coast of Washington to approximately the 700 fathom (~1300 m) depth contour (Figure 1). This area covers all of the continental shelf adjacent to Washington and the upper portion of the continental slope. It includes the entire Olympic Coast National Marine Sanctuary (OCNMS), and Flattery Rocks, Quillayute Needles and Copalis National Wildlife Refuges. It excludes Willapa Bay and Grays Harbor, the Strait of Juan de Fuca, the Lower Columbia River Estuary, and the Salish Sea. The 700 fathom depth contour limit was proposed by the WDFW to delimit the marine area with the most human activity (Washington Department of Ecology 2014).

The study area is used by at least 100 different species of marine birds and shorebirds (OCNMS 2008). Birds use the area for a variety of purposes including foraging within nearshore and offshore habitats, breeding among isolated islands, and migrating between breeding and wintering areas outside of the study area.

## Reference Map

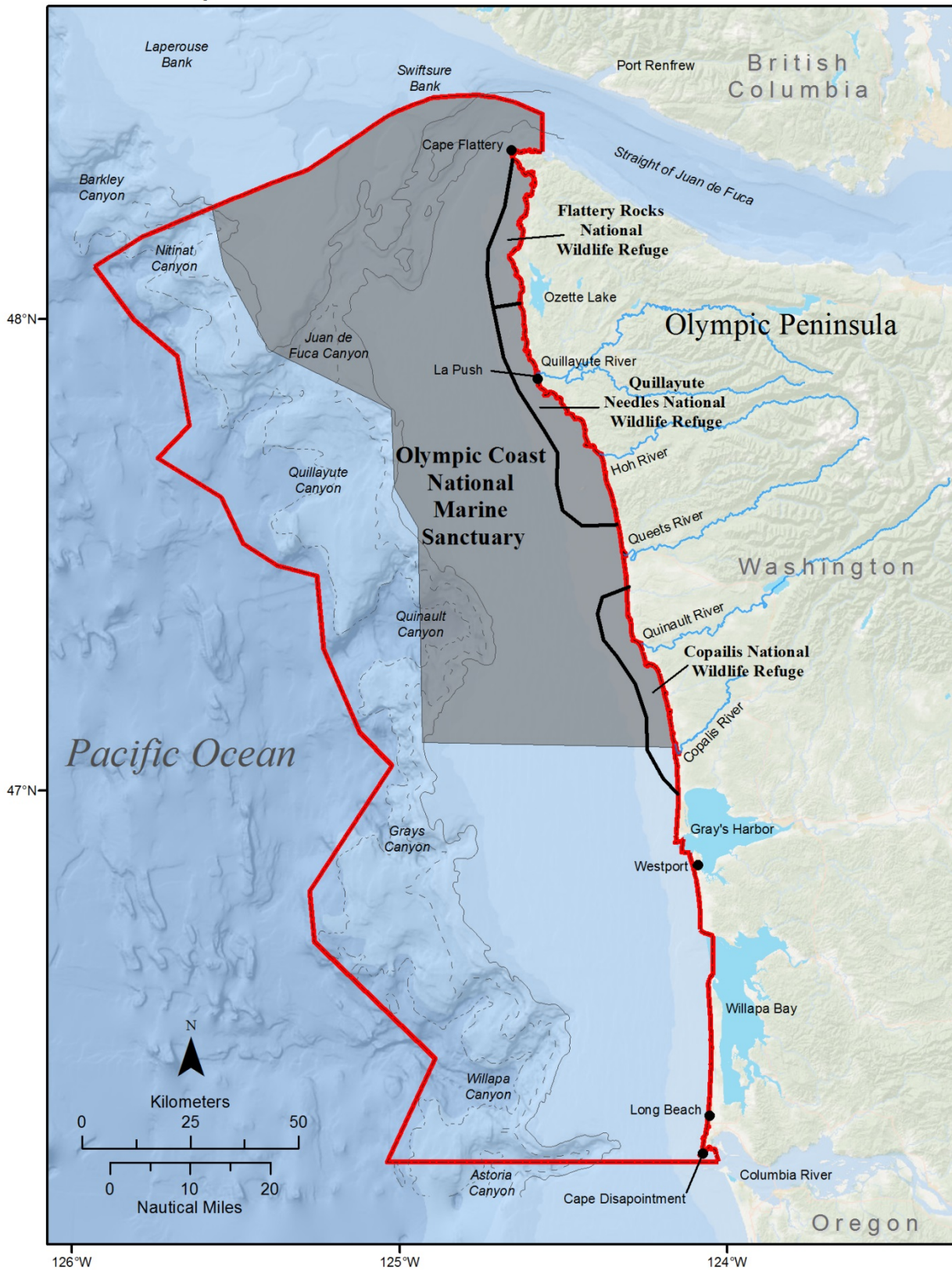


Figure 1: Map of the study area (red line) used to model seabird distributions. Olympic Coast National Marine Sanctuary is designated by the gray shading. 200, 500, and 1000 m isobath contours are shown as solid, large dashed, and small dashed gray lines, respectively. This map serves as a reference map of locations for this report.

### Seabird data sets

Species distributions were modeled using a compilation of at-sea observations collected by eight seabird survey programs (Table 2). Each program collected spatially-explicit seabird observations within a sampling domain which overlapped the study area, and in some cases extended well beyond the study area. Most programs collected data over multiple years, although not necessarily consecutively. The programs include sightings from small boats, large vessels, and fixed-wing aircraft. In most cases, survey programs collected observations within strip transects. Taken together, the spatial distribution of survey programs represents a discontinuous patchwork of observation effort in the study area (Figure 2). Additional information for each survey program is provided in Appendix A.

Table 2: Survey programs used for modeling seabird distributions and the corresponding number of segments across all months (Total), and two time periods.

Seabird survey program	Number of segments		
	Total	April to October	November to March
Pacific Continental Shelf Environmental Assessment (PaCSEA)*	1,040	791	249
ORegon, CALifornia, and WASHINGTON Line-transect Expedition (ORCAWALE) marine mammal survey	194	194	0
Collaborative Survey of Cetacean Abundance and the Pelagic Ecosystem (CSCAPE)	459	459	0
Seasonal Olympic Coast National Marine Sanctuary Surveys	758	758	0
Annual Olympic Coast National Marine Sanctuary Surveys	517	517	0
Northwest Fisheries Science Center (NWFSC) Northern California Current Seabird Surveys	1,548	869	679
Northwest Forest Plan Marbled Murrelet Effectiveness Monitoring Program	6,684	6,684	0
Pacific Coast Winter Seaduck Survey	287	0	287
Totals	11,487	10,272	1,215

\*PaCSEA segments were not used to model Tufted Puffin or Sooty Shearwater because they were not identified to species during these surveys.

Pelagic transects were typically longer than 25 km and ran perpendicular to shore or followed a sawtooth pattern across the continental shelf. Nearshore transects typically ran parallel to shore. Transects ranged from 25 km to hundreds of km long and were spaced from 5 to 125 km apart.

A subset of observations within each survey program was used for analysis. The subset included observations extending 10 km outside the outer perimeter of the study area, and observations collected from 2000 to 2013. Several of the survey programs obtained data prior to 2000, but we chose to focus analysis on data collected after 2000 to ensure results reflected relatively recent spatial patterns.



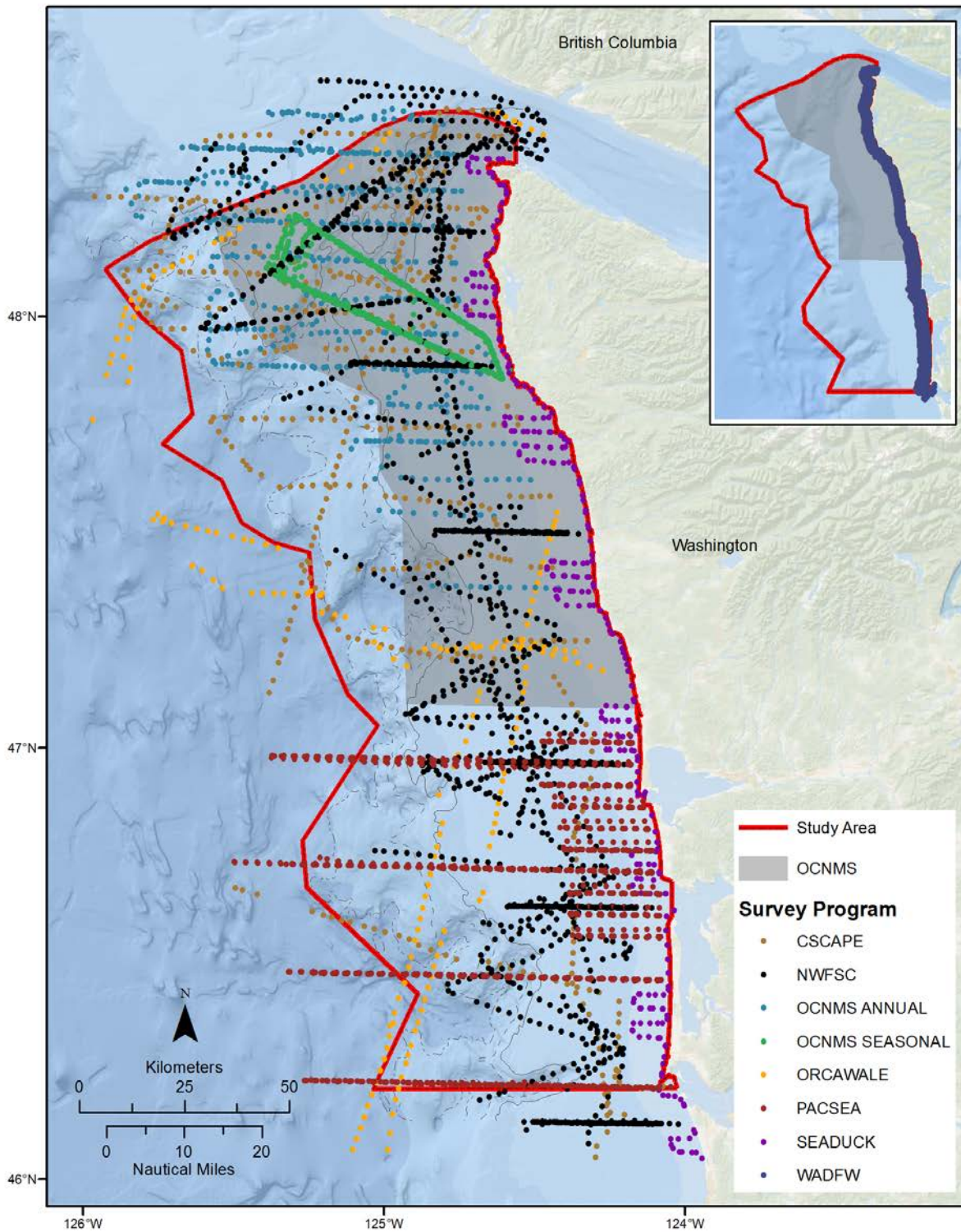


Figure 2: Spatial distribution of seabird survey programs, represented by the location of processed transect segment coordinates. Segments from the Northwest Forest Plan Marbled Murrelet Effectiveness Monitoring Program are shown separately in the inset panel, because they obscure visibility of other program data. Olympic Coast National Marine Sanctuary is designated by gray shading. 200, 500, and 1000 m isobath contours are shown as solid, large dashed, and small dashed gray lines, respectively.

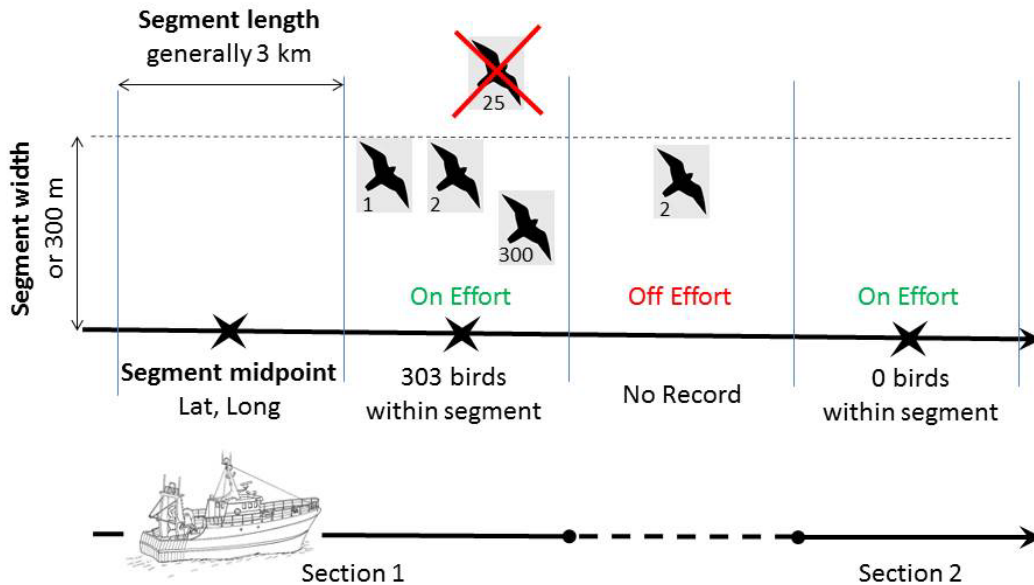


Figure 3: Schematic of the segmentation process used to partition seabird observations along transects.

A series of custom-built processing routines written in R (R Core Team 2015) were used to extract and reformat observations within survey transects. Transects were divided into a series of mutually-exclusive sections and subsequent segments according to changes in observation effort and sea state (Figure 3). Only transect sections with observers “on effort” were selected for segmentation.

We used a modified version of a processing routine written by Karin Forney (Southwest Fisheries Science Center [SWFSC]) and Elizabeth Becker (SWFSC) to divide each transect section into discrete segments. The routine divides sections into predominantly 3 km segments, but also includes a subroutine to use the remainder of a section after 3 km segments are identified. The routine first determines the number of 3 km segments which fit within each “on effort” section and the length leftover after segmentation. If the remainder was greater than or equal to 1.5 km long, the corresponding subsegment was assigned to a randomly selected location along the transect. If the remainder was less than 1.5 km, the subsegment was added to a randomly selected segment. The routine randomizes the location of transect breakpoints, but does not change the location of observations. The randomization was necessary to ensure smaller subsegments did not always occur at the end of a transect.

For all transects, except those which were part of the PaCSEA and NWFSC Northern California Current Seabird Surveys (Table 2), linearly interpolated track lines were generated from geographic positions provided with observations. The track lines consisted of regularly spaced coordinates at spatial resolutions of approximately 10 m, and were used to identify segment coordinates. The midpoint of each segment was calculated from the average of x- and y-coordinates within the corresponding segment track line. Track lines were not generated for PaCSEA and NWFSC Northern California Current Seabird Surveys because the data were provided processed into 3 km segments with midpoint coordinates.

Observations were divided into two non-overlapping time periods based on month of observation: April to October and November to March. These two seasons generally correspond to the major upwelling and downwelling oceanographic seasons off the coast of Washington, respectively (Mann and Lazier 2006). We expected the spatial distribution of seabirds to differ between these two time periods, due to changes in physical water properties, weather, seabird life history, and primary productivity.

All birds of the same species within a segment regardless of bird size or life history stage were summed to produce species counts per segment. The resulting metric, counts per segment, was entered into models with an

area offset calculated as the product of segment strip width and segment length. Consequently, model output is a measure of relative density defined as counts per square kilometer (sq. km). Most survey programs used a strip width of either 150, 200 or 300 m, but when a strip width was not defined, only observations within 300 m of the ship were analyzed.

Predicted densities are considered relative, not absolute measures of abundance per unit area. Relative density is sufficient to achieve our objective of identifying species-specific areas of higher density. Our models assume the effect of detection on predicted relative density for a given species is consistent across the study area; however, we expect that sighting condition (such as sea state) affects detection and future models could be improved by incorporating this measure.

### **Spatiotemporal coverage of seabird data**

The distributions of processed segments were explored by month, year, and distance to shore to examine broad scale spatiotemporal patterns of the modeled seabird data. These patterns provide context for interpreting the final model maps.

Seabird observations were most often collected in the upwelling season, especially May to July, and close to shore (Table 3). There are many fewer observations at the beginning and end of the upwelling period (April, and August to October), and during the entire downwelling period (November to March). The downwelling period has no data before 2008 and the vast majority of data was collected in March, at the end of the season. Across all seasons the number of observations in the study area has generally increased over time, mainly because of an increase in offshore effort over time. Plots of each species' observed densities averaged by month are provided in Appendix D.

### **Environmental predictors**

A wide range of predictor variables were used to model variation in the number of birds sighted per transect segment and to predict the relative density of birds throughout the study area (Table 4). Predictor variables fell into one of six categories: survey, temporal, geographic, topographic, physical oceanographic, and biological. Appendix B provides detailed processing steps.

Survey predictor variables were designed to account for variation in counts arising from heterogeneity in the type of survey platform (e.g., boat or fixed-wing aircraft), characteristics of the survey platform (e.g., observation height), observer identity and expertise, species focus, and sighting conditions. These factors influence the probability that individual birds will be detected and correctly identified to the species level. Of these factors, only the type of survey platform was consistently recorded in all datasets, and thus was directly usable as a predictor variable. We attempted to account for the effects of the remaining factors through two random-effect predictor variables representing survey identity (ID) and transect ID, respectively. The exact definition of transect ID differed somewhat between datasets, but unique transect IDs generally represented pre-defined survey transects or individual days of effort.

Temporal predictor variables were designed to account for variation in the numbers of birds within the study area over time. Julian day was used to account for changes within a season (e.g., arising from migratory movements in and out of the study area), and year was used to account for changes across years (e.g., arising from changes in population abundance or distributional shifts). Effects of Julian day and year were modeled as smooth continuous changes over time. Four climate indices (Table 4) were also included as temporal predictor variables to account for variation arising from linkages between the environment and seabird density.

Table 3: Distribution of seabird segments across months, years, and bands of distance to shore. Three contingency tables are shown: month by year, distance to shore by year, and distance to shore by month. Cell colors highlight areas of relatively high (red), medium (orange), and low (green) values within each contingency table. Colors vary by panel and stretch according to each panel's data range. The months between April and October, representing the modeled upwelling season, are shaded in grey.

Month by Year																
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	% of all observations	
Jan	0	0	0	0	0	0	0	0	0	0	0	168	0	0	1%	
Feb	0	0	0	0	0	0	0	0	0	0	0	0	195	0	2%	
Mar	0	0	0	0	0	0	0	0	161	254	156	131	150	0	7%	
Apr	0	0	0	0	0	0	0	0	0	175	0	0	0	0	2%	
May	41	96	29	148	380	77	144	187	211	162	93	178	168	115	18%	
Jun	73	144	511	153	182	542	258	231	361	281	397	513	308	176	36%	
Jul	79	122	316	186	212	227	172	168	286	254	162	158	450	209	26%	
Aug	0	0	0	0	0	0	56	51	30	0	0	0	46	0	2%	
Sep	0	96	0	0	0	0	17	52	53	50	0	0	189	0	4%	
Oct	0	0	0	0	0	45	0	0	0	0	51	201	0	0	3%	
Nov	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%	
Dec	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%	
% of all observations	2%	4%	7%	4%	7%	8%	6%	6%	10%	10%	7%	12%	13%	4%		

Distance to shore by Year																
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	% of all observations	
0 - 5 km	123	220	405	305	298	273	300	321	328	317	375	388	364	298	38%	
5 - 10 km	7	35	52	17	61	107	52	64	145	156	108	245	246	25	11%	
10 - 25 km	0	15	87	0	111	127	51	51	199	287	110	223	295	0	14%	
25 - 50 km	63	121	204	165	171	150	172	191	229	180	224	245	280	177	22%	
50 - 200 km	0	67	108	0	133	234	72	62	201	236	42	248	321	0	15%	
% of all observations	2%	4%	7%	4%	7%	8%	6%	6%	10%	10%	7%	12%	13%	4%		

Distance to shore by Month																
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec			% of all observations	
0 - 5 km	7	5	150	3	845	1668	1613	7	8	9	0	0			38%	
5 - 10 km	45	62	156	26	215	464	161	35	89	67	0	0			11%	
10 - 25 km	45	55	217	91	286	512	121	49	108	72	0	0			14%	
25 - 50 km	14	37	72	2	495	1003	887	11	31	20	0	0			22%	
50 - 200 km	57	36	257	53	188	483	219	81	221	129	0	0			15%	
% of all observations	1%	2%	7%	2%	18%	36%	26%	2%	4%	3%	0%	0%				



Geographic predictor variables were designed to account for variation in counts arising from spatial location *per se*. Isotropic x- and y-coordinates (based on projected longitude and latitude values) were included as predictor variables and their effects were modeled two ways. The first (x, y)-coordinate term allowed for smooth changes in numbers across the study area arising from spatial factors not captured by the other predictor variables. For example, none of the predictor variables capture colonization history or association with fishing vessels. The second term was formulated using radial basis functions to capture residual spatial autocorrelation in the data, after accounting for the effects of other predictor variables. Absolute distance to shelf break (200 m isobath), distance to colony (for Common Murre and Tufted Puffin), and distance to nesting habitat (for Marbled Murrelet) were also included as geographic predictor variables.

Topographic variables were designed to account for variation in counts arising from the direct and indirect effects of depth and seafloor features on bird distributions. Six different bathymetric datasets were used as topographic variables (Table 4). Depth, planform curvature, profile curvature, and slope were available through MARSPEC, a high-resolution GIS database of ocean climate and topographic layers (Sbrocco and Barber 2013; <http://www.marspec.org/>). A bathymetric position index was derived from the MARSPEC depth layer.

Physical oceanographic predictor variables were designed to account for variation in counts arising from the direct and indirect effects of the physical state and dynamics of the ocean. Five physical oceanographic predictor variables were developed from a range of data sources, including one from MARSPEC (Table 4). Remote sensing data were used to characterize sea surface salinity and temperature. Probabilities of cyclonic and anticyclonic eddy rings and probability of sea surface temperature fronts were derived from the remotely sensed variables.

Two biological predictor variables, chlorophyll *a* concentration and the frequency of chlorophyll peaks index (FCPI) developed by Suryan et al. (2012), were included to account for variation in counts arising from the direct and indirect effects of ocean productivity (Table 4). Although both predictors comprise measures of chlorophyll *a* concentration, they were not correlated (Table 5). This was expected since the FCPI is an indicator of anomalous conditions that differ from the chlorophyll *a* concentration average.

All of the physical oceanographic and biological variables that we considered are dynamic. To associate dynamic variables with seabird density and develop long-term predictions, dynamic variables, except for FCPI, were summarized into two seasonal climatologies. Data time series ranging from 11 to 22 years were used to estimate monthly mean climatologies. Monthly climatologies were then averaged within two non-overlapping seasons: April to October and November to March, which correspond to the major upwelling and downwelling oceanographic seasons off the coast of Washington, respectively (Mann and Lazier, 2006). The same temporal divisions were also used to partition seabird observations. FCPI was a non-seasonal climatology; however, Suryan et al. (2012) noted that the FCPI is in part related to peak seasonal chlorophyll *a* concentration values (i.e., during June-July).

Geographic, topographic, physical oceanographic, and biological predictor variables were spatially explicit. Each variable was calculated on a standard study grid projected onto zone 10N of the Universal Transverse Mercator coordinate system and with a spatial resolution of 3 km. When the native spatial resolution of a predictor variable was finer than that of the study grid, predictor values were averaged within study grid cells. When the native spatial resolution of a predictor variable was similar to or coarser than that of the study grid, bilinear interpolation was used to derive predictor values at the center of study grid cells. For many data sets the native spatial extent did not perfectly align with our land mask. To fill in missing values close to shore, we extrapolated values to the coastline using the “Springs” algorithm in the `inpaint_nans` MATLAB function (<http://www.mathworks.com/matlabcentral/fileexchange/4551-inpaint-nans>). Each survey transect segment was matched to the predictor variable values from the study grid cell that contained the midpoint of that segment.

Table 4: Predictor variables used in the models. Additional information is provided in Appendix B for predictors with an asterisk.

Predictor variable	Code	Native resolution	Source
<b>Survey variables</b>			
Survey platform	platform	Categorical variable	Seabird datasets
Survey ID	survey	Categorical variable	Seabird datasets
Transect ID	transect	Categorical variable	Seabird datasets
<b>Temporal variables</b>			
Julian day	jday	1 day	Seabird datasets
Year	year	1 year	Seabird datasets
Multivariate El Niño-Southern Oscillation Index (current and 12 month lag)	mei, mei.lag12	Monthly	<a href="#">NOAA ESRL</a>
North Pacific Gyre Oscillation Index (current and 12 month lag)	npgo, npgo.lag12	Monthly	<a href="#">Georgia Tech</a>
Pacific Decadal Oscillation Index (current and 12 month lag)	pdo, pdo.lag12	Monthly	<a href="#">NOAA ESRL</a>
Upwelling index (current and 12 month lag)	upi, upi.lag12	Monthly	<a href="#">Pacific Fisheries Environmental Laboratory</a>
<b>Geographic variables</b>			
X-coordinate	coords.x	n/a	Seabird datasets
Y-coordinate	coords.y	n/a	Seabird datasets
Distance to 200 m isobath*	dist2isobath	50 km	Derived from depth layer
Distance to Common Murre colony*	dist2comu	3 km	Derived from Washington Seabird Colony Catalog (source <a href="#">WDFW</a> )
Distance to Marbled Murrelet nesting habitat*	dist2mamu	3 km	Derived from Marbled Murrelet critical habitat (source: <a href="#">USFWS site</a> )
Distance to Tufted Puffin colony*	dist2tupu	3 km	Derived from Washington Seabird Colony Catalog (source <a href="#">WDFW</a> )
<b>Topographic variables</b>			
Depth*	bathy	30 seconds (~ 700 m)	<a href="#">MARSPEC</a>
Bathymetric position index (3 km)*	bpi.3km	30 seconds (~ 700 m)	Derived from depth layer
Bathymetric position index (20 km)*	bpi.20km	30 seconds (~ 700 m)	Derived from depth layer
Slope*	slope	30 seconds (~ 700 m)	<a href="#">MARSPEC</a>
Planform curvature*	plcurv	30 seconds (~ 700 m)	<a href="#">MARSPEC</a>
Profile curvature*	prcurv	30 seconds (~ 700 m)	<a href="#">MARSPEC</a>
<b>Physical variables (seasonal climatologies)</b>			
Probability of anticyclonic eddy ring*	anticyc	0.25 degrees (~ 25 km)	<a href="#">AVISO</a>
Probability of cyclonic eddy ring*	cyc	0.25 degrees (~ 25 km)	<a href="#">AVISO</a>
Sea surface salinity*	salinity	30 seconds (~ 700 m)	<a href="#">MARSPEC</a>
Sea surface temperature*	sst	0.05 degrees (~ 5.5 km)	<a href="#">Aqua MODIS</a>
Probability of sea surface temperature front*	front	0.05 degrees (~ 5.5 km)	<a href="#">GOES Imager</a>
<b>Biological variables (seasonal and non-seasonal climatologies)</b>			
Surface chlorophyll <i>a</i> *	chla	0.05 degrees (~ 5.5 km)	<a href="#">Aqua MODIS</a>
Frequency of chlorophyll peaks index*	fcpi	9 km	Provided by Rob Suryan

A small number of the spatially explicit predictor variables were correlated with each other (Table 5). Since some correlations remain relatively high (i.e., greater than 0.7), inferences regarding the association between relative variable importance and a functional ecological relationship with seabird density should be made with caution. The accuracy of predictions is less affected by collinearity among predictor variables.

Table 5: Pairwise Spearman’s rank correlation coefficients (rho) for spatial predictor variables (excluding x- and y-coordinates) during the months of April to October (above diagonal) and November to March (below diagonal). High correlations are highlighted ( $0.7 \leq |\rho| < 0.8$  in yellow,  $0.8 \leq |\rho| < 0.9$  in orange).

	bathy	bpi.20km	bpi.3km	dist2isobath	dist2comu	dist2mamu	dist2tupu	fcpi	plcurv	prcurv	slope	anticyc	chla	cyc	front	salinity	sst
bathy		0.12	0.26	0.66	-0.30	-0.76	-0.32	-0.29	-0.18	0.08	-0.51	-0.28	0.72	-0.01	-0.39	-0.59	0.01
bpi.20km	0.21		0.64	-0.31	-0.35	0.05	-0.37	-0.07	0.03	-0.13	0.28	-0.04	0.09	-0.38	0.16	0.30	-0.42
bpi.3km	0.15	0.50		-0.10	-0.40	-0.11	-0.42	-0.17	0.09	-0.21	0.34	-0.12	0.20	-0.36	0.00	0.17	-0.38
dist2isobath	0.78	-0.02	0.01		-0.14	-0.65	-0.15	-0.17	-0.24	0.17	-0.59	-0.08	0.59	0.17	-0.34	-0.63	0.31
dist2comu	-0.25	0.01	-0.08	-0.10				-0.01	0.07	-0.03	-0.08	0.28	-0.57	0.61	-0.02	-0.25	0.64
dist2mamu	-0.80	-0.01	-0.02	-0.59				0.25	0.25	-0.14	0.55	-0.07	-0.80	0.21	0.48	0.54	-0.10
dist2tupu	-0.26	0.00	-0.08	-0.11				0.00	0.11	-0.08	-0.06	0.27	-0.62	0.63	-0.03	-0.26	0.65
fcpi	-0.24	0.17	0.03	-0.21	-0.11	0.43	-0.11		0.06	-0.11	0.12	0.27	-0.23	-0.07	0.29	0.35	-0.14
plcurv	-0.09	-0.19	0.17	-0.06	-0.01	-0.02	0.00	-0.16		-0.45	0.13	-0.03	-0.26	0.03	-0.09	0.09	-0.05
prcurv	0.12	-0.16	-0.33	0.21	-0.11	-0.11	-0.12	0.02	-0.32		-0.11	-0.03	0.20	-0.09	0.08	0.02	0.02
slope	-0.75	-0.01	0.10	-0.77	0.17	0.53	0.17	0.08	0.07	-0.19		0.07	-0.41	-0.28	0.38	0.62	-0.36
anticyc	0.34	0.05	-0.03	0.49	-0.04	-0.30	-0.04	0.03	-0.11	0.15	-0.45		-0.18	-0.22	0.14	0.18	0.32
chla	0.83	-0.10	-0.04	0.68	-0.23	-0.84	-0.24	-0.49	0.02	0.14	-0.66	0.32		-0.30	-0.27	-0.33	-0.16
cyc	-0.46	0.09	-0.07	-0.13	0.49	0.47	0.49	0.32	-0.12	-0.03	0.13	0.30	-0.45		-0.13	-0.55	0.55
front	0.10	0.26	0.19	0.09	-0.09	-0.09	-0.09	0.06	-0.16	0.08	0.09	0.05	0.02	0.01		0.53	-0.15
salinity	-0.72	0.09	0.06	-0.68	-0.14	0.74	-0.14	0.51	0.01	-0.11	0.63	-0.41	-0.87	0.13	-0.02		-0.63
sst	-0.48	0.12	-0.02	-0.21	0.60	0.46	0.60	0.30	-0.11	-0.07	0.24	0.28	-0.53	0.85	0.07	0.17	

## Modeling process

A statistical modeling framework was used to relate bird sightings data from surveys to a range of temporal and spatial environmental predictor variables. The estimated relationships between relative occurrence and abundance of birds with predictor variables, after accounting for area surveyed, were then used to predict the distributions of birds across the entire study area. These predictions generated the maps of modeled relative density we present in this report. Separate models were developed for each combination of species and season for which there were sufficient data.

A boosted generalized additive modeling framework (Bühlmann and Hothorn 2007, Hofner et al. 2012) was used to estimate relationships between the numbers of birds counted per transect segment and the predictor variables, after accounting for area surveyed. Those relationships were then used to predict the relative density of each species throughout the study area in each season. Our main objective was to provide accurate predictions, so we chose a modeling framework that allowed for flexible relationships and multi-way interactions between predictor variables while accounting for sampling heterogeneity between and within datasets. This modeling framework was successfully applied to seabirds along the U.S. Atlantic coast (Kinlan et al. 2014).

## Likelihoods and model components

The number of individuals of a given species counted per transect segment was modeled using zero-inflated Poisson (Equation 1) and zero-inflated negative binomial likelihoods (Equation 2) to account for the overdispersed nature of the count data. Each component/parameter of the likelihood was modeled as a separate function of the predictor variables (Schmid et al. 2008, Mayr et al. 2012). For the zero-inflated Poisson likelihood, the two model components were the probability of an ‘extra’ zero ( $p$ ; also referred to as “zero-

inflation component”) and the mean of the Poisson distribution ( $\mu$ ; also referred to as “count component”). The same components were modeled for the zero-inflated negative binomial likelihood in addition to the dispersion parameter of the negative binomial distribution ( $\theta$ ). The probability of an extra zero was modeled on the logit scale while the mean of the Poisson/negative binomial distribution and the dispersion parameter of the negative binomial distribution were modeled on the log scale.

$$[1] \quad L(p, \mu; \mathbf{y}) = \prod_{i=1}^n \left\{ I_{y_i=0} [p + (1-p)e^{-\mu}] + I_{y_i>0} \left[ (1-p) \frac{\mu^{y_i} e^{-\mu}}{y_i!} \right] \right\}$$

$$[2] \quad L(p, \mu, \theta; \mathbf{y}) = \prod_{i=1}^n \left\{ I_{y_i=0} \left[ p + (1-p) \left( \frac{\theta}{\theta + \mu} \right)^\theta \right] + I_{y_i>0} \left[ (1-p) \frac{\Gamma(y_i + \theta)}{y_i! \Gamma(\theta)} \left( \frac{\theta}{\theta + \mu} \right)^\theta \left( \frac{\mu}{\theta + \mu} \right)^{y_i} \right] \right\}$$

In the above equations,  $n$  represents the total number of observations,  $I_{y_i=0}$  is an indicator for whether or not  $y_i = 0$  (i.e.,  $I_{y_i=0}$  equals 1 when  $y_i = 0$ , zero otherwise),  $I_{y_i>0}$  is an indicator for whether or not  $y_i > 0$ , and  $\Gamma()$  is the usual gamma function:  $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ .

### Base-learners

Within the boosting framework, each model component was constructed as a function of an ensemble of ‘base-learners.’ Each base-learner represented a specific functional relationship between a model component and one or more predictor variables. We utilized a suite of base-learners each representing different predictor variables, and different sets of base-learners were employed for different model components (Table 6).

Table 6: Base-learners employed in the boosted generalized additive modeling framework. Base-learner names are from the ‘mboost’ package for R (Hothorn et al. 2015, R Core Team 2015), and predictor variable names are defined in Table 4.

Name	Description	Predictor variables	Model component
bols	linear	intercept	$p, \mu, \theta$
bols	linear	platform	$p, \mu, \theta$
brandom	random effect	survey	$\theta$
brandom	random effect	transect	$p, \mu$
bbs	penalized regression spline <sup>1</sup>	year	$p, \mu$
bbs	penalized regression spline <sup>1</sup>	jday	$p, \mu$
btree	tree <sup>2</sup>	all climate indices (current and lagged)	$p, \mu$
bspacial	penalized tensor product <sup>1</sup>	coords.x coords.y	$p, \mu$
brad	penalized radial basis <sup>3</sup>	coords.x coords.y	$p, \mu$
btree	tree <sup>4</sup>	dist2isobath dist2comu, dist2mamu, dist2tupu all topographic, physical oceanographic, and biological variables	$p, \mu$

<sup>1</sup> P-spline basis

<sup>2</sup> Maximum depth = 1

<sup>3</sup> Matern correlation function

<sup>4</sup> Maximum depth = 4, 5

All spatially explicit predictor variables except geographic coordinates were included together in a single tree base-learner. The trees for that learner had a maximum depth of 4 or 5, which allowed for interacting effects among the spatially explicit predictor variables. Geographic coordinates appeared in two base learners, and those variables always entered the model as a pair. The remaining survey and temporal predictor variables entered the model individually, either through their own base-learners or, in the case of climate indices, one at a time through a tree base-learner with a maximum depth of 1. Thus, our model structure did not allow for interactions between temporal and spatial predictor variables.

#### *Effort offset*

The mean of the Poisson/negative binomial distribution model component ( $\mu$ ) was additionally modeled with an effort offset, corresponding to segment survey area in sq. km, that was log transformed prior to entering the model. Therefore, resulting model predictions corresponded directly to relative density values (birds per sq. km) rather than relative count values (birds per segment).

#### *Stochastic gradient boosting*

Stochastic gradient boosting was used to fit models whereby a sub-sample of the data was fitted in each boosting iteration (Friedman 2002). Rather than resampling the data for each boosting iteration, a set of 25 or 50 random samples was created before boosting, and one sample was randomly drawn from this set for each boosting iteration. Root mean square error (RMSE) was used to select the base-learner that gave the best fit to the gradient (all data) in each boosting iteration.

#### *Boosting offsets*

Model component estimates were initialized ('offset' in boosting terminology; Hofner et al. 2012) by conducting a preliminary generalized linear model (GLM) analysis. For that analysis, predictor variables were first reduced through principal component and cluster analyses to a smaller set of derived predictors. Those new predictors were then discretized into different numbers of classes. For each number of classes a GLM with a zero-inflated Poisson or zero-inflated negative binomial likelihood was fit, and the mean estimates for each model component were calculated. Model component estimates were then averaged across the fitted models with the different numbers of predictor classes, weighted by the Akaike Information Criterion (AIC) for those models.

#### *Tuning of shrinkage rate and number of boosting iterations*

A stratified (by transect ID)  $k$ -fold cross-validation approach was used to determine values for the shrinkage rate ( $nu$ ) and number of boosting iterations ( $mstop$ ) that resulted in the best predictive performance. The shrinkage rate was tuned first by fixing the number of boosting iterations and evaluating out-of-bag model performance in terms of the thresholded continuous rank probability score (CRPS) for different shrinkage rates. The number of boosting iterations was tuned second by fixing the shrinkage rate and evaluating out-of-bag model performance in terms of the negative log-likelihood. The number of boosting iterations at which performance was maximized was averaged across cross-validation samples (excluding the top and bottom 5%) and used as the number of boosting iterations for the final model fitting. If the number of boosting iterations was less than or greater than specified values, the shrinkage rate was decreased or increased, respectively, and the number of boosting iterations was tuned again. We allowed for a maximum of 20,000 boosting iterations, so models with boosting iterations above ~19,990 should be interpreted with caution as their performance may have improved with additional boosting iterations. A suite of cross-validation performance metrics were calculated during the tuning of  $mstop$ .

#### *Model performance and selection*

Four models were fit to each species-season combination: 1) zero-inflated Poisson with a maximum tree depth of 4 specified for the non-climate index predictors (Table 6), 2) zero-inflated Poisson with a maximum tree depth of 5, 3) zero-inflated negative binomial with a maximum tree depth of 4, and 4) zero-inflated negative

binomial with a maximum tree depth of 5. The performance of each of the four fitted models was evaluated from a suite of performance metrics. Cross-validation performance during the tuning of *mstop* in terms of the thresholded continuous rank probability score was used to select either the zero-inflated Poisson or the zero-inflated negative binomial model as the final best model for each species and season.

### *Spatial prediction*

The final fitted model for each species and season was used to predict relative density, defined as the mean number of individuals per square kilometer, throughout the study area. Relative density integrated both the zero-inflated and Poisson/negative binomial components of the likelihood.

Spatially explicit predicted values were calculated for each cell of the study grid from the values of the spatially explicit predictor variables for that cell. Thus, the predicted relative density in a given grid cell corresponds to predictions for a transect segment whose mid-point falls within that grid cell. All other predictor variables were set to their mean values.

### *Variable importance*

While our primary objective was not to determine the ecological drivers and mechanisms behind the spatial distributions of marine bird species in the study area, our model results do provide some indication of which variables were most useful for predicting those distributions. Those variables may provide useful starting points for future studies seeking ecological inference.

We calculated the relative importance of a given predictor variable in the final fitted models by summing the decrease in the negative log-likelihood in each boosting iteration attributable to that predictor variable. Thus, variable importance reflects the frequency with which a given predictor variable occurred in the selected base-learners across boosting iterations and that variable's ability to explain variation in the data when it was selected. When multiple predictor variables occurred in the selected base-learner for a given boosting iteration, the decrease in the negative log-likelihood was divided evenly among those predictor variables. Relative variable importance was re-scaled so that it summed to one across predictor variables.

### *Uncertainty*

Uncertainty in model predictions was estimated using a non-parametric bootstrapping framework. For each bootstrap iteration, the set of unique transect IDs was resampled with replacement, and the data for each transect ID were assigned weights proportional to the frequency of that ID in the sample. These data weights were then applied when fitting the model during that bootstrap iteration. Predictor variables that were not included in the final model were excluded from the bootstrap analysis. Two hundred bootstrap iterations were conducted producing a sample of predictions from which we calculated quantiles, confidence intervals, and coefficient of variation to characterize uncertainty in the predictions. This uncertainty information was mapped and is provided alongside seabird density maps.

### *Implementation*

The analysis was coded in R (R Core Team 2015) and relied on multiple existing contributed packages (e.g., *mboost*; Hothorn et al. 2015).

### *Review*

Contributors participated in several rounds of review in which model outputs were evaluated. This collaborative process led to significant improvements in model quality and the presentation of results. Over the course of model development, the following changes were made: excluding observations outside the study area and a narrow encircling buffer; improving the transect segmentation process using a routine developed by SWFSC; incorporating oceanographic climate indices as predictor variables; and adding diagnostic plots to address the impact of heterogeneous survey effort through space and time.

## Model Results

A range of performance metrics were chosen to assess model fit to observations, quantify model uncertainty, identify caveats, and describe the relationships between predictions and predictors. Table 7 and Appendix C show model performance metrics for final selected models and all fitted models, respectively (see *Model performance and selection* section for information on model selection).

We provide brief descriptions of model outputs and show corresponding maps for each modeled seabird and season combination in the remainder of this section. Models were developed only when there were greater than 50 segments with sightings for a given species; consequently, not all species were modeled during the downwelling season. This number was based on previous modeling experience and our assessment of sighting density requirements in the study area. The results intentionally focus on describing broad scale spatial patterns and how these relate to geographic places, predictors, and model diagnostics. The descriptions do not explore associations between model spatial patterns and seabird life history patterns, but this exploration would be a useful next step. Figure 1 can be used to identify the location of place names mentioned in the results.

### **Marbled Murrelet (*Brachyramphus marmoratus*)**

*April to October*

Predicted relative densities for Marbled Murrelet are high in waters less than 10 km from shore and along the Olympic Peninsula from Cape Flattery to the Copalis River (Figure 4). This stretch of nearshore water is within Olympic Coast National Marine Sanctuary and covers the three National Wildlife Refuges. The predictions of high density correspond well with observed patterns (Figure 4 inset). Relative densities decrease quickly with distance from shore. The segments of coast between Cape Flattery and Ozette Lake and between La Push and the Copalis River have the highest predicted relative densities for Marbled Murrelet within the study region. Model uncertainty is lowest near the shore and shows a clear longitudinal trend of increasing values from east to west (Figure 5b). This suggests low uncertainty and high confidence in coastal areas of high predicted relative density.

Diagnostics indicate the Marbled Murrelet model is very good (Table 7); however, since transect ID was selected more times than any other predictor for both the zero-inflation and count components of the model, there is room to improve the predictor set. When transect ID is chosen most often, it suggests that the environmental predictors are doing a poor job of explaining variability in density.

### **Tufted Puffin (*Fratercula cirrhata*)**

*April to October*

Areas of highest predicted relative density are nearshore (<10 km from the coast) along the Olympic Peninsula north of the Hoh River and in deep offshore waters south of the Juan de Fuca Canyon (Figure 6). A band of moderately-high density is predicted to connect these two regions of high density. The predicted relative density patterns are consistent with known breeding colonies near the coast and suggest a possibly important large feeding area offshore. Model uncertainties are relatively low throughout the study area, with CV values generally increasing from northeast to southwest (Figure 7b). Model fit diagnostics indicate the Tufted Puffin model is very good (Table 7). Similar to the Marbled Murrelet model, transect ID was selected more times than any other predictor for both components of the selected model, suggesting that the environmental predictors do a poor job of explaining variability in density.

### **Common Murre (*Uria aalge*)**

*April to October*

Predicted relative density is highest in nearshore waters along the coast and drops off sharply in waters further than 25 km from shore (Figure 8). The areas of highest relative density are predicted to occur in nearshore waters and adjacent to Cape Flattery, Flattery Rocks National Wildlife Refuge, and the mouths of the

Quillayute, Quinault, and Columbia rivers. CVs generally increase from east to west and are relatively small in nearshore waters, except in small patches adjacent to areas with the highest relative densities (Figure 9a-b). Additionally, there is a small region of high predicted relative density with high uncertainty (i.e., high CV) in the northernmost part of the study area, near Swiftsure Bank, where the Juan de Fuca Eddy creates a seasonal area of upwelling. This could potentially be important habitat for Common Murre; however this prediction cannot be confirmed due to lack of survey data in this region (Figure 8 inset). Model fit diagnostics are excellent (Table 7), and predicted and observed density spatial patterns correspond well where they overlap.

#### *November to March*

Relative densities of Common Murre are predicted to be high between the shelf-edge and 5 kilometers from shore, with a predicted latitudinal density gradient increasing towards the south of the study area (Figure 10). The highest predictions occur offshore from Grays Harbor and Willapa Bay. Unlike predictions for the months of April to October, Common Murres are predicted to be relatively rare in coastal waters less than 5 km from shore, especially in the southern half of the study area. This band of low density is clearly visible in observations. In general, uncertainty is low in areas of predicted high and moderate densities, and very low in waters deeper than 1000 m (Figure 11a-b).

Model fit diagnostics are excellent (Table 7). Over the majority of the study area, predicted and observed spatial patterns of density correspond well, but there is a small area of moderately high predicted relative density and fairly low uncertainty at the northwestern edge of the study region, which should be cautiously interpreted. There was no survey coverage in this area between November and March and the patterns are inconsistent with the rest of the predictions.

### **Black-footed Albatross (*Phoebastria nigripes*)**

#### *April to October*

Relative density of Black-footed Albatross is predicted to be high along the shelf-edge (Figure 12) with the highest densities near submarine canyons, such as Barkley, Nitinat, Juan De Fuca, and Grays Canyons. Predicted relative density is low on the continental shelf, except in shelf waters adjacent to submarine canyons. A broad area of moderately high density is predicted in the northwest of the study area over Barkley and Nitinat Canyons and extending eastward to Swiftsure Bank. Model uncertainty is low in coastal waters, but CV values increase and become highly variable offshore near areas of high predicted relative density (Figure 13a-b). This may be due to the fact that Black-footed Albatross are highly mobile throughout their range and there were too few sightings in the observations data set. Model fit diagnostics are generally good (Table 7).

#### *November to March*

Predictions of relative density are highest off the northern coast of Washington, directly west of the Olympic Peninsula in waters roughly 200 to 1000 m deep (Figure 14). This area includes an area of low uncertainty (i.e., low CV) above Quillayute and Juan de Fuca Canyons. The majority of offshore predictions are characterized by high uncertainty, most likely due to the small number of sightings (and surveys) between the months of November and March (Table 1, Figure 14 inset, Figure 15b). Densities of Black-footed Albatross are predicted to be relatively low in waters on the continental shelf (depths shallower than 200 m). The vast majority of sightings in the downwelling season were in March (Appendix D), suggesting the seasonal map may reflect individuals arriving from winter breeding sites (Naughton et al. 2007), and is not representative of Black-footed Albatross distribution in other months.

### **Northern Fulmar (*Fulmarus glacialis*)**

#### *April to October*

Predictions are highest between the Juan de Fuca Canyon, Swiftsure Bank and Nitinat Canyon (Figure 16). This area of high density includes in the general location of the Juan de Fuca Eddy, where upwelling increases nutrient concentrations (MacFadyen et al. 2008). A narrow band of moderately high density is predicted to



extend south of the Juan de Fuca Canyon, along the shelf edge (i.e. 200 m depth contour). Predictions are relatively certain over most of the continental shelf where densities are predicted to be low, and uncertain in deeper waters. There is an area of low uncertainty at the center of the area with highest relative density predictions (Figure 17a-b). Model fit diagnostics for the Northern Fulmar model are excellent (Table 7). The predicted distribution was lower than expected in the southern part of the study area, and this may be caused by the relatively low number of observations along the shelf edge between August and October, the latter portion of the upwelling season.

### **Pink-footed Shearwater (*Puffinus creatopus*)**

*April to October*

Areas of high predicted relative density occur between the 100 and 200 m depth contours on the continental shelf (Figure 18). The region west of the Olympic Peninsula and surrounding the Juan de Fuca Canyon contains the highest predicted relative densities and low uncertainty. Coastal waters (<5 km from shore) have low predicted densities and generally low uncertainty (i.e., low CV values) (Figure 19a-b). Low predictions and high uncertainty characterize waters west of the 1000 m depth contour.

Model performance metrics suggest a slightly poorer fit for Pink-footed Shearwater compared to models for other species, especially in areas where sightings occurred (Table 7). However, we feel the model performance is adequate and the resulting predictions of relative density are useful, especially for identifying regions of relatively low use. Transect ID was selected for the count component of the model more times than any other predictor, possibly suggesting that our suite of environmental predictor variables do a poor job of explaining variability in Pink-footed Shearwater densities during the months of April to October.

### **Sooty Shearwater (*Puffinus griseus*)**

*April to October*

A large contiguous zone of high predicted relative density runs south from nearshore waters off La Push, WA to the southern edge of the study area at the mouth of the Columbia River (Figure 20). This zone extends out to 25 km from shore and includes coastal waters off of Grays Harbor, Willapa Bay, and the Columbia River. Moderately high relative density is predicted on the continental shelf north of La Push, except for waters within 5 km from the coast. Observed densities are extremely variable across most of the study area, except in the south coastal region of the study area where the highest densities were observed. Given this variability and the nature of Sooty Shearwaters to form large aggregations, the extreme variability in CV values is expected. In general, model uncertainty is high in the south, southwest, and northeast regions of the study area, with smaller patches of high uncertainty near Willapa Bay and offshore between Grays Harbor and La Push, WA (Figure 21b). Model fit diagnostics are good (Table 7).

Table 7: Performance metrics for the final selected model of each species-season combination. Also shown are number of transect segments with sightings, number of individuals sighted, proportion of transect segments with sightings (prevalence), and mean number of individuals per transect segment with sightings. All model performance metrics were calculated on the full dataset, except columns labeled “Cross-val,” which denote statistics calculated separately on cross-validation data. AUC values were calculated as the area under the relevant ROC curve. Rank R refers to Spearman’s rank correlation coefficient (Spearman’s rho statistic) for the observed vs. predicted count. Median absolute [residual] error (non-zero) and median [residual] bias (non-zero) were standardized by dividing by mean no. individuals per segment with sightings and multiplying by 100, so that values shown represent percentages of the mean number of individuals observed per segment on segments with at least one sighting. For details on the Brier score and continuous ranked probability score (CRPS), please consult Brier (1950) and Gneiting and Raftery (2007).

Species code	Season	No. segments with		Mean no. individuals per segment with		Prevalence	Best model type	Max. tree depth	No. boosting iterations	AUC		Rank R	Gaussian rank correlation (non-zero)	Median absolute error (non-zero)		Median bias (non-zero)		Brier score (occupancy)		CRPS	
		sightings	individuals	sightings	Prevalence					Occupancy	Non-zero			Fit	Cross-val	Fit	Cross-val	Fit	Cross-val	Fit	Cross-val
tupu	summer	1,738	11,777	6.8	0.18	ZIP	4	19,995	0.92	0.75	0.60	0.63	21.6%	23.5%	-11.5%	-6.7%	0.08	0.09	0.08	0.09	
mamu	summer	1,625	5,604	3.4	0.16	ZINB	4	19,999	0.92	0.76	0.62	0.64	37.7%	32.4%	-33.2%	-24.7%	0.08	0.10	0.07	0.09	
comu	summer	6,533	293,713	45.0	0.64	ZINB	5	17,999	0.91	0.84	0.70	0.70	20.1%	27.1%	4.5%	12.0%	0.11	0.14	0.10	0.12	
comu	winter	405	6,516	16.1	0.33	ZIP	5	8,611	0.91	0.82	0.69	0.70	19.6%	23.9%	-2.9%	8.3%	0.12	0.11	0.10	0.10	
bfal	summer	421	3,008	7.1	0.04	ZINB	5	15,215	0.96	0.70	0.41	0.44	12.1%	11.9%	-11.1%	-11.6%	0.03	0.02	0.03	0.02	
bfal	winter	87	162	1.9	0.07	ZIP	4	17,986	0.97	0.78	0.46	0.43	37.0%	37.4%	-34.8%	-33.1%	0.03	0.03	0.03	0.03	
nofu	summer	463	2,916	6.3	0.05	ZINB	4	17,733	0.97	0.75	0.57	0.61	15.9%	15.8%	-11.3%	-11.1%	0.02	0.02	0.02	0.02	
pfsh	summer	611	3,977	6.5	0.06	ZIP	4	17,962	0.96	0.65	0.39	0.43	22.1%	19.5%	-12.5%	-11.1%	0.04	0.03	0.03	0.03	
sosh	summer	2,586	249,380	96.4	0.27	ZINB	4	17,999	0.91	0.72	0.49	0.52	8.3%	8.3%	1.5%	1.5%	0.11	0.09	0.10	0.08	

**Marbled murrelet (*Brachyramphus marmoratus*): April to October**

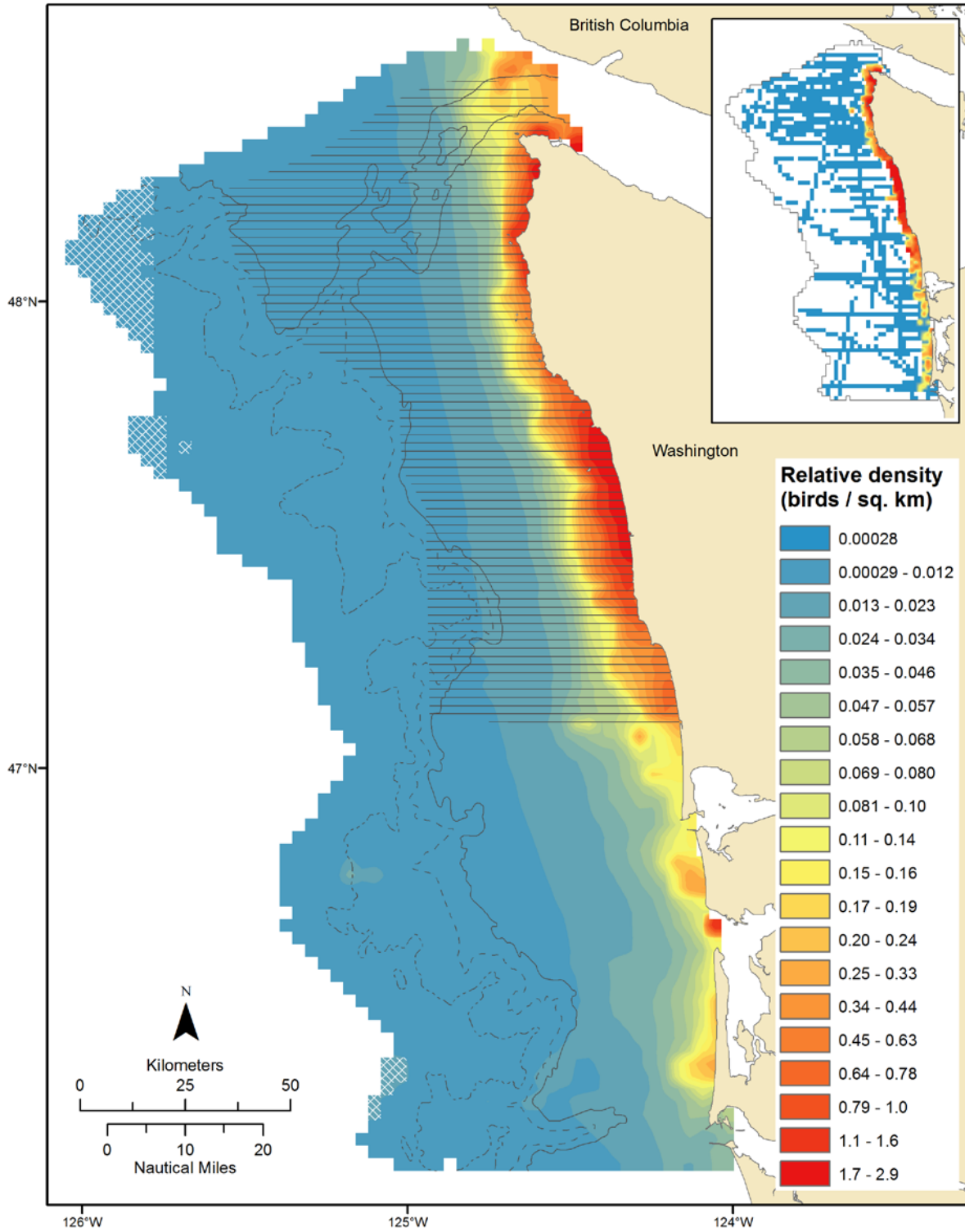


Figure 4: Long-term relative density (birds per sq. km) prediction map for Marbled Murrelet (*Brachyramphus marmoratus*) during the months of April to October. White cross-hatching represents areas of greater prediction uncertainty, where the coefficient of variation was greater than or equal to 0.8. Olympic Coast National Marine Sanctuary is designated by gray horizontal lines. 200, 500, and 1000 m isobath contours are shown as solid, large dashed, and small dashed gray lines, respectively. Observed density (birds per sq. km) is shown on the inset map. Prediction and observed density color gradient classes represent 5% quantile intervals calculated from predictions.

Marbled murrelet (*Brachyramphus marmoratus*): April to October

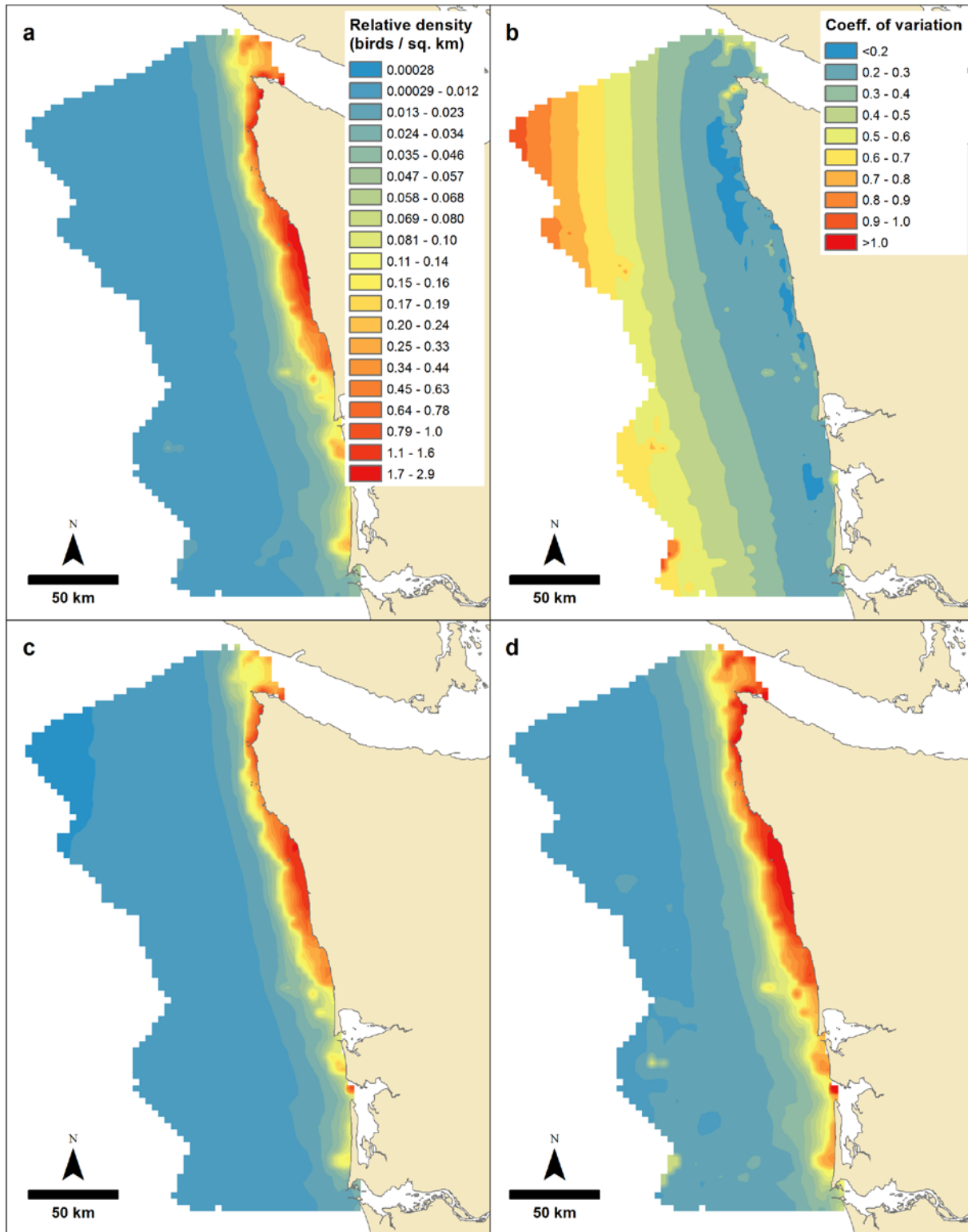


Figure 5: Long-term relative density (birds per sq. km) prediction maps for Marbled Murrelet (*Brachyramphus marmoratus*) during the months of April to October: a) 50% quantile of bootstrap (median), b) coefficient of variation, c) 5% quantile of bootstrap, d) 95% quantile of bootstrap. The legends for panels c and d are the same as for panel a.

**Tufted puffin (*Fratercula cirrhata*): April to October**

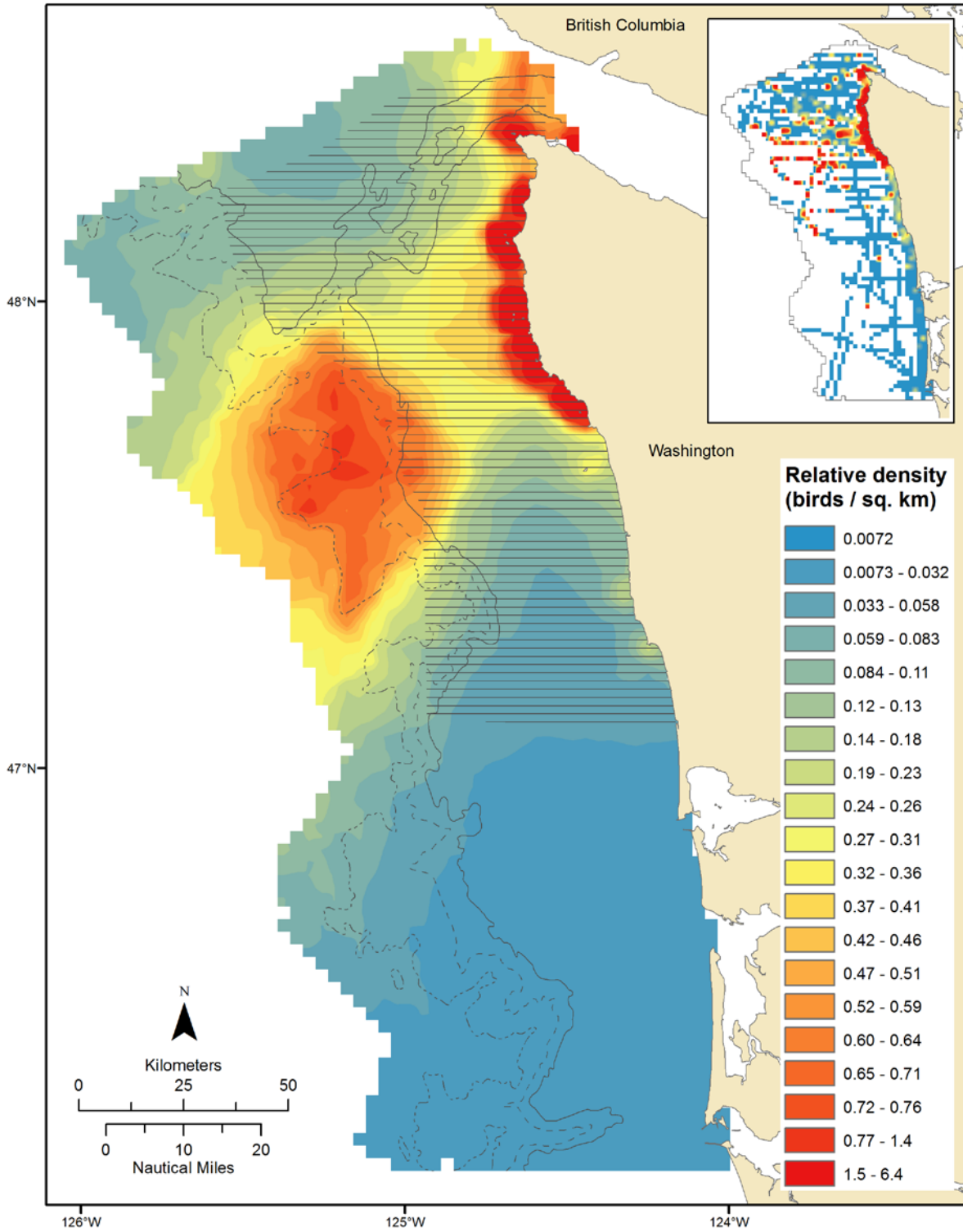


Figure 6: Long-term relative density (birds per sq. km) prediction map for Tufted Puffin (*Fratercula cirrhata*) during the months of April to October. Olympic Coast National Marine Sanctuary is designated by gray horizontal lines. 200, 500, and 1000 m isobath contours are shown as solid, large dashed, and small dashed gray lines, respectively. Observed density (birds per sq. km) is shown on the inset map. Prediction and observed density color gradient classes represent 5% quantile intervals calculated from predictions.

Tufted puffin (*Fratercula cirrhata*): April to October

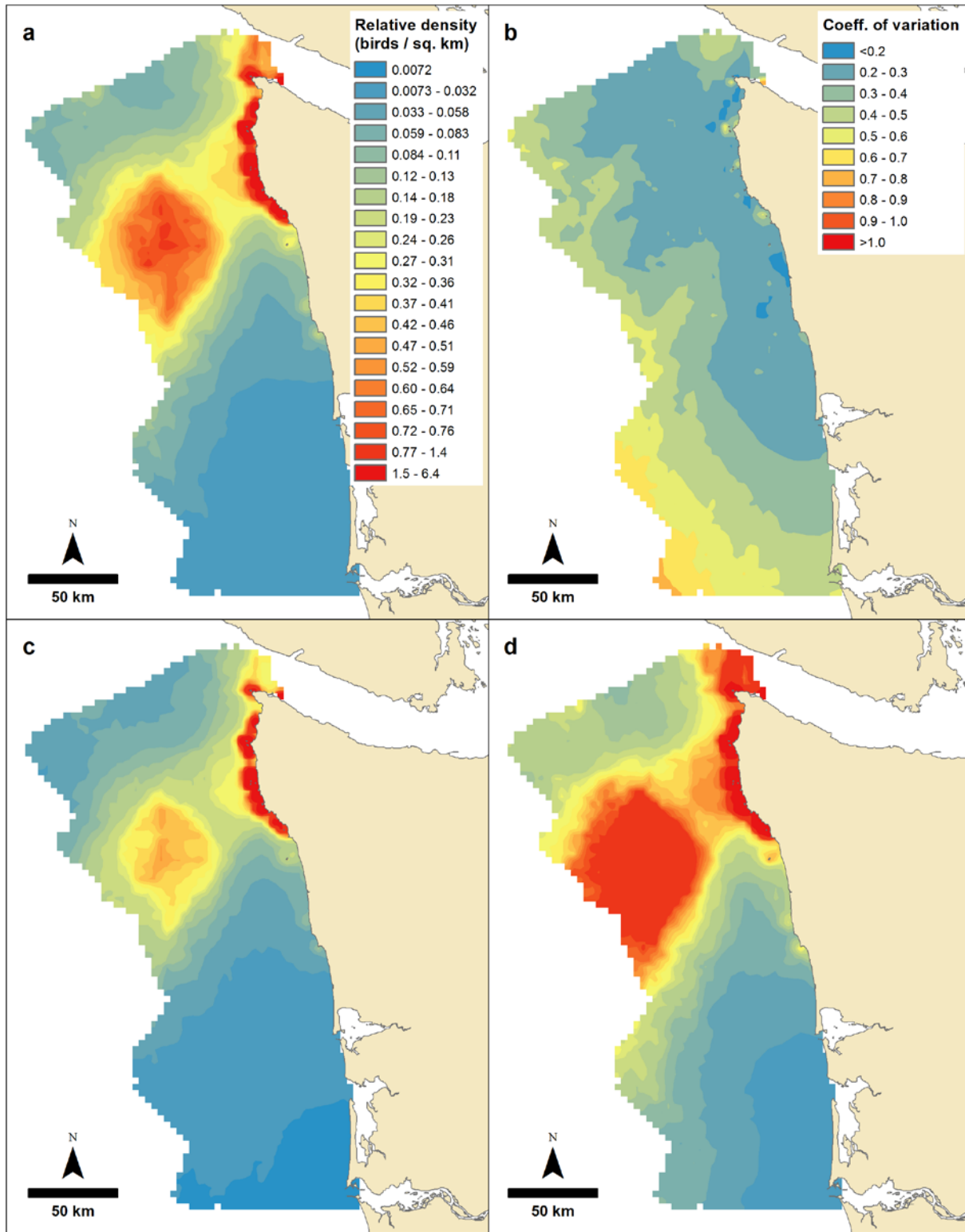


Figure 7: Long-term relative density (birds per sq. km) prediction maps for Tufted Puffin (*Fratercula cirrhata*) during the months of April to October: a) 50% quantile of bootstrap (median), b) coefficient of variation, c) 5% quantile of bootstrap, d) 95% quantile of bootstrap. The legends for panels c and d are the same as for panel a.

**Common murre (*Uria aalge*): April to October**

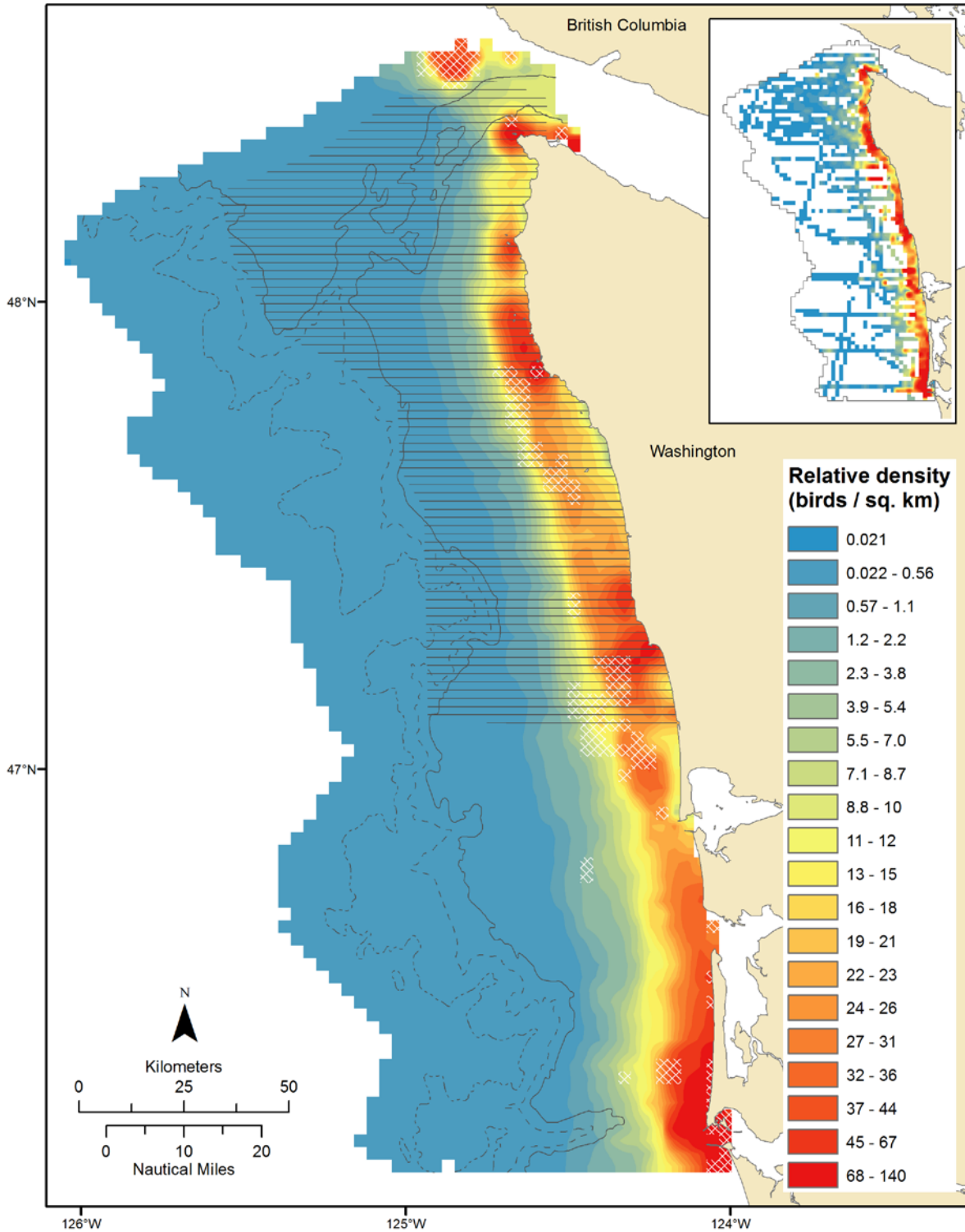


Figure 8: Long-term relative density (birds per sq. km) prediction map for Common Murre (*Uria aalge*) during the months of April to October. White cross-hatching represents areas of greater prediction uncertainty, where the coefficient of variation was greater than or equal to 0.8. Olympic Coast National Marine Sanctuary is designated by gray horizontal lines. 200, 500, and 1000 m isobath contours are shown as solid, large dashed, and small dashed gray lines, respectively. Observed density (birds per sq. km) is shown on the inset map. Prediction and observed density color gradient classes represent 5% quantile intervals calculated from predictions.



**Common murre (*Uria aalge*): April to October**

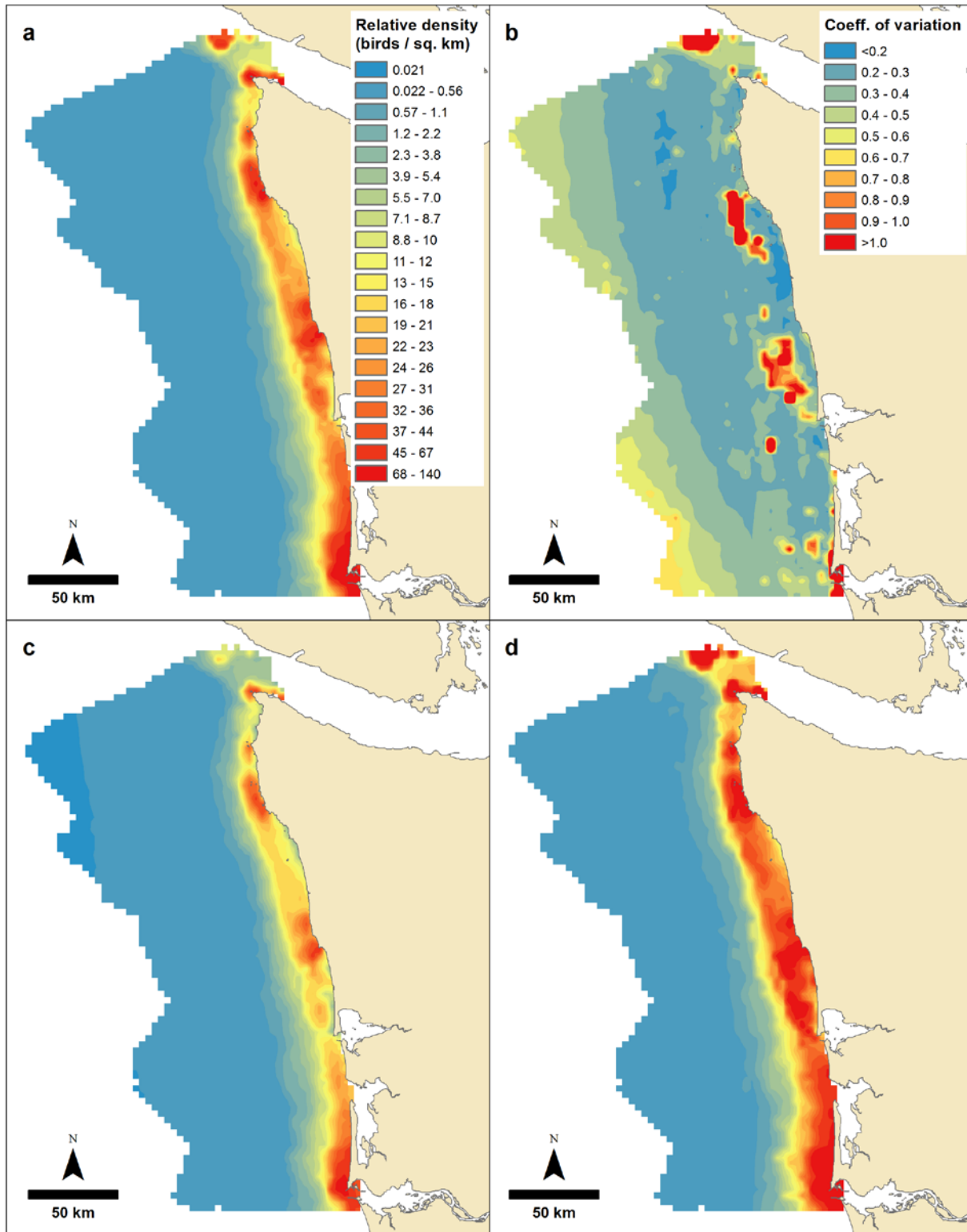


Figure 9: Long-term relative density (birds per sq. km) prediction maps for Common Murre (*Uria aalge*) during the months of April to October: a) 50% quantile of bootstrap (median), b) coefficient of variation, c) 5% quantile of bootstrap, d) 95% quantile of bootstrap. The legends for panels c and d are the same as for panel a.



**Common murre (*Uria aalge*): November to March**

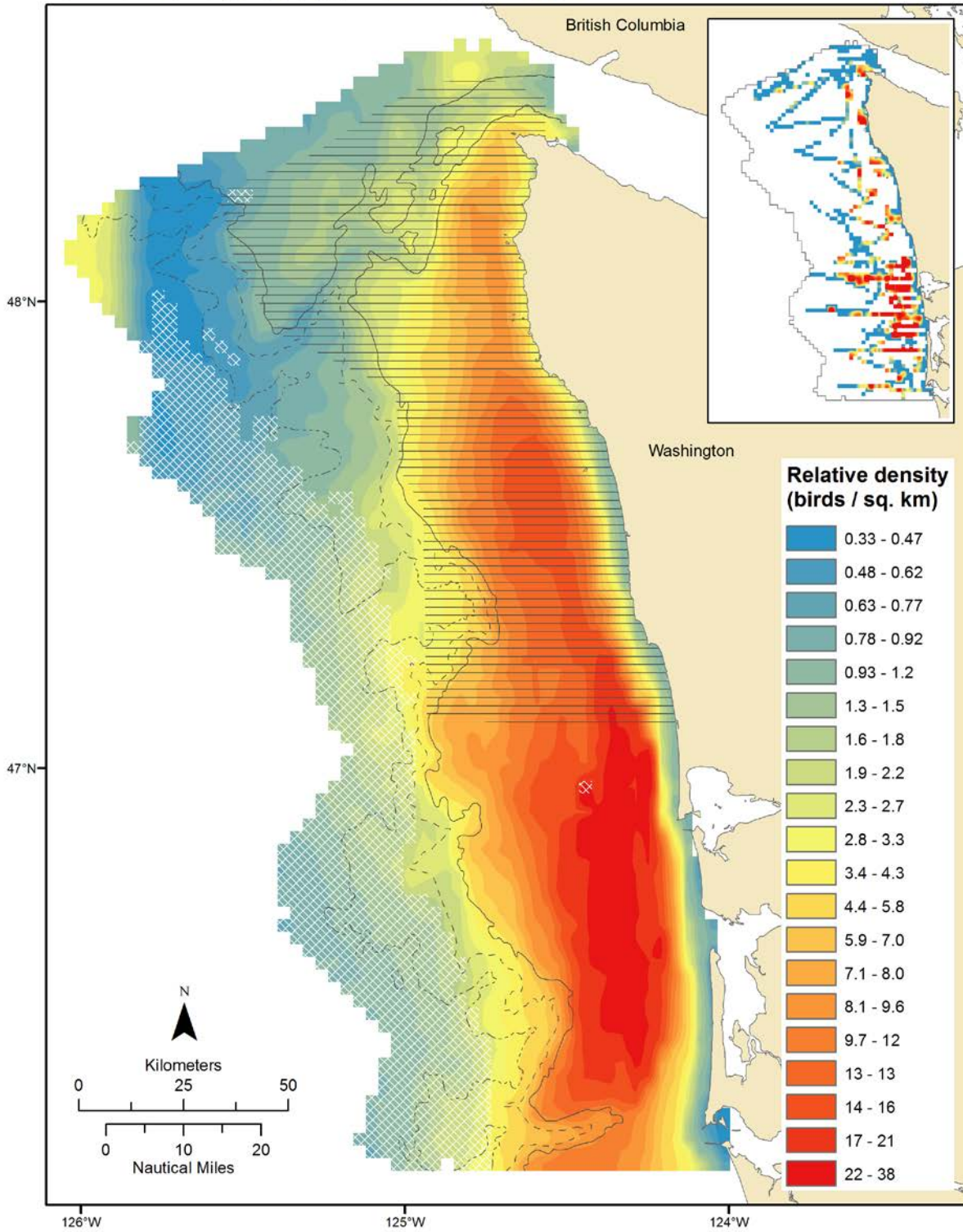


Figure 10: Long-term relative density (birds per sq. km) prediction map for Common Murre (*Uria aalge*) during the months of November to March. White cross-hatching represents areas of greater prediction uncertainty, where the coefficient of variation was greater than or equal to 0.8. Olympic Coast National Marine Sanctuary is designated by gray horizontal lines. 200, 500, and 1000 m isobath contours are shown as solid, large dashed, and small dashed gray lines, respectively. Observed density (birds per sq. km) is shown on the inset map. Prediction and observed density color gradient classes represent 5% quantile intervals calculated from predictions.

**Common murre (*Uria aalge*): November to March**

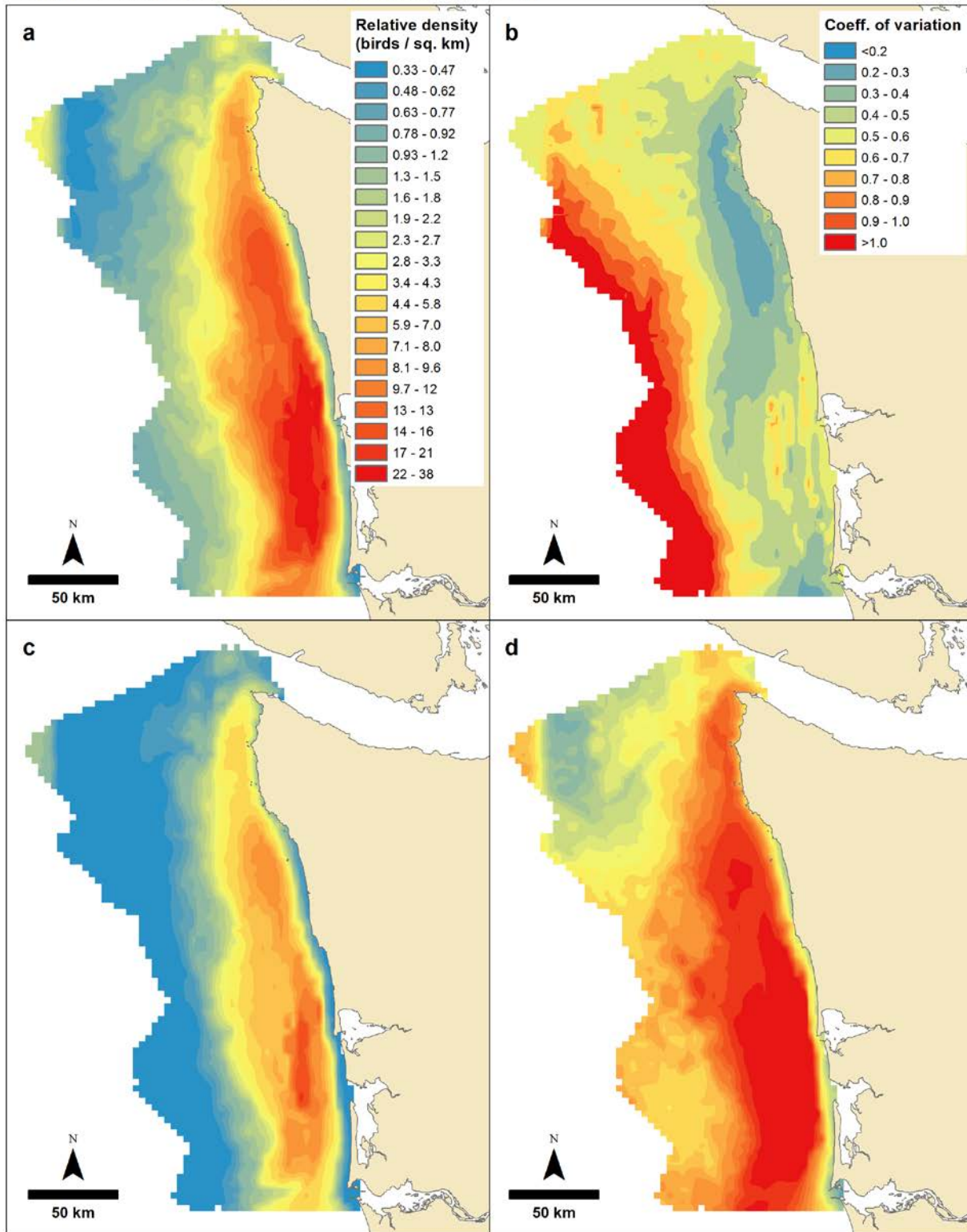


Figure 11: Long-term relative density (birds per sq. km) prediction maps for Common Murre (*Uria aalge*) during the months of November to March: a) 50% quantile of bootstrap (median), b) coefficient of variation, c) 5% quantile of bootstrap, d) 95% quantile of bootstrap. The legends for panels c and d are the same as for panel a.



**Black-footed albatross (*Phoebastria nigripes*): April to October**

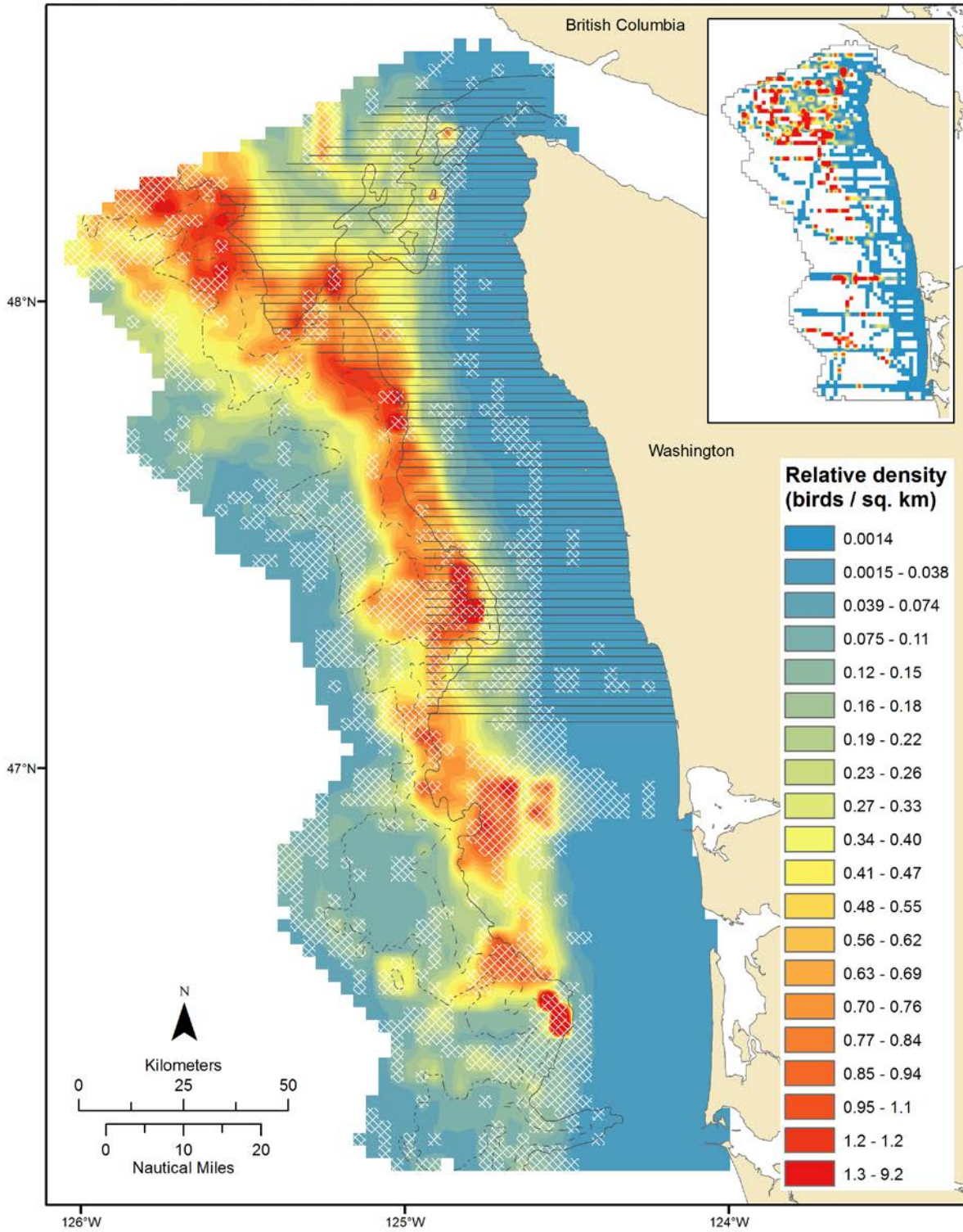


Figure 12: Long-term relative density (birds per sq. km) prediction map for Black-footed Albatross (*Phoebastria nigripes*) during the months of April to October. White cross-hatching represents areas of greater prediction uncertainty, where the coefficient of variation was greater than or equal to 0.8. Olympic Coast National Marine Sanctuary is designated by gray horizontal lines. 200, 500, and 1000 m isobath contours are shown as solid, large dashed, and small dashed gray lines, respectively. Observed density (birds per sq. km) is shown on the inset map. Prediction and observed density color gradient classes represent 5% quantile intervals calculated from predictions.

**Black-footed albatross (*Phoebastria nigripes*): April to October**

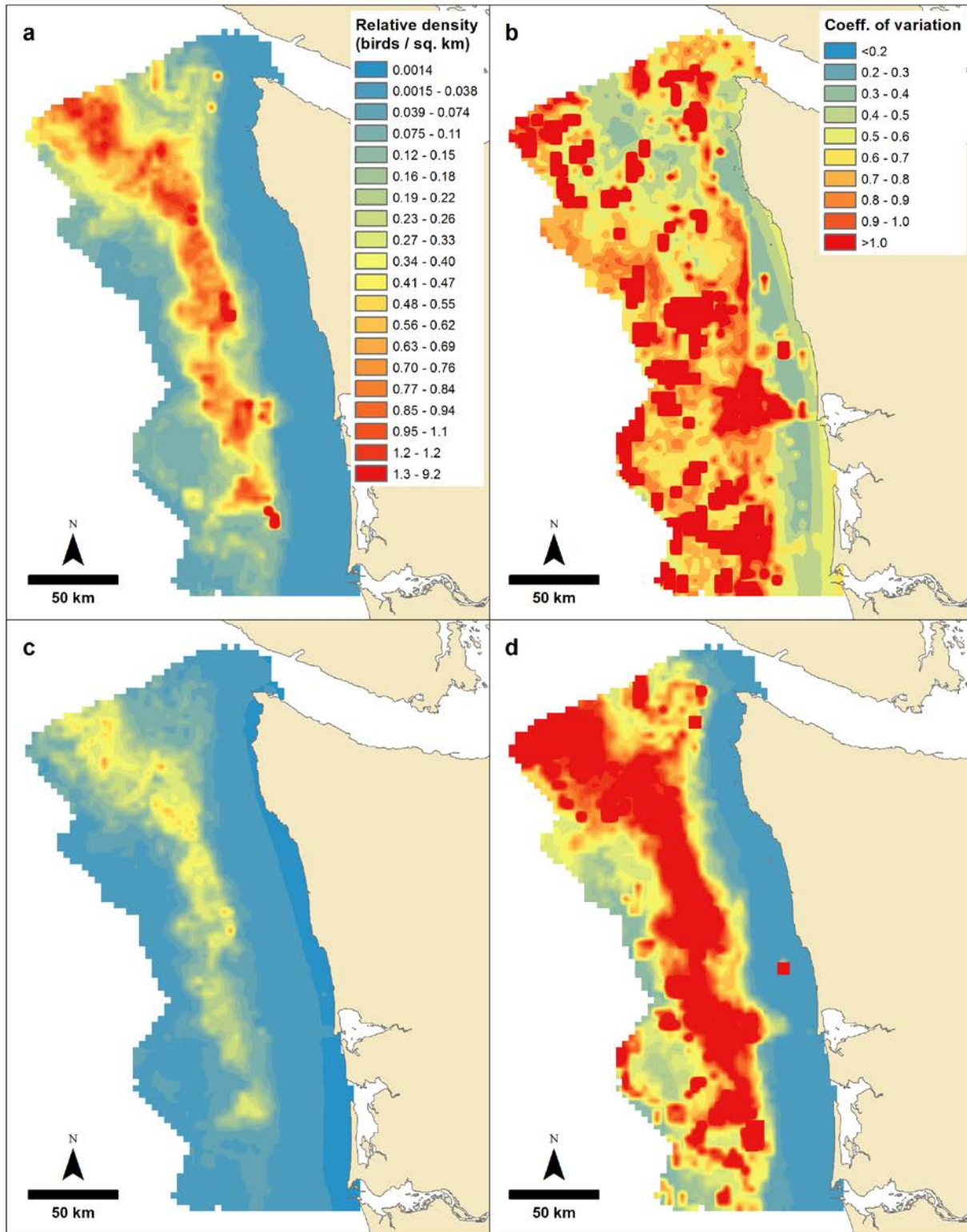


Figure 13: Long-term relative density (birds per sq. km) prediction maps for Black-footed Albatross (*Phoebastria nigripes*) during the months of April to October: a) 50% quantile of bootstrap (median), b) coefficient of variation, c) 5% quantile of bootstrap, d) 95% quantile of bootstrap. The legends for panels c and d are the same as for panel a.



**Black-footed albatross (*Phoebastria nigripes*): November to March**

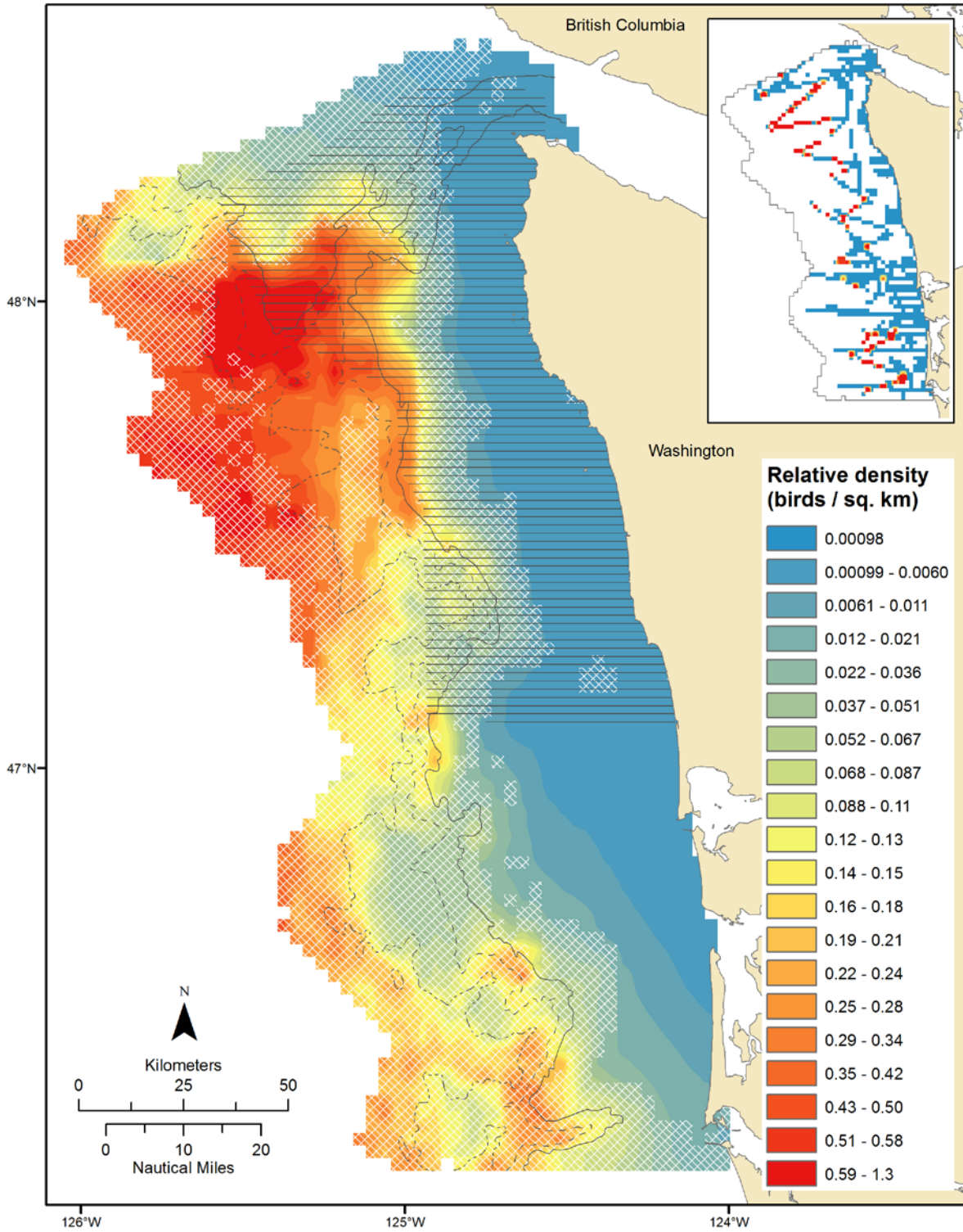


Figure 14: Long-term relative density (birds per sq. km) prediction map for Black-footed Albatross (*Phoebastria nigripes*) during the months of November to March. White cross-hatching represents areas of greater prediction uncertainty, where the coefficient of variation was greater than or equal to 0.8. Olympic Coast National Marine Sanctuary is designated by gray horizontal lines. 200, 500, and 1000 m isobath contours are shown as solid, large dashed, and small dashed gray lines, respectively. Observed density (birds per sq. km) is shown on the inset map. Prediction and observed density color gradient classes represent 5% quantile intervals calculated from predictions.

**Black-footed albatross (*Phoebastria nigripes*): November to March**

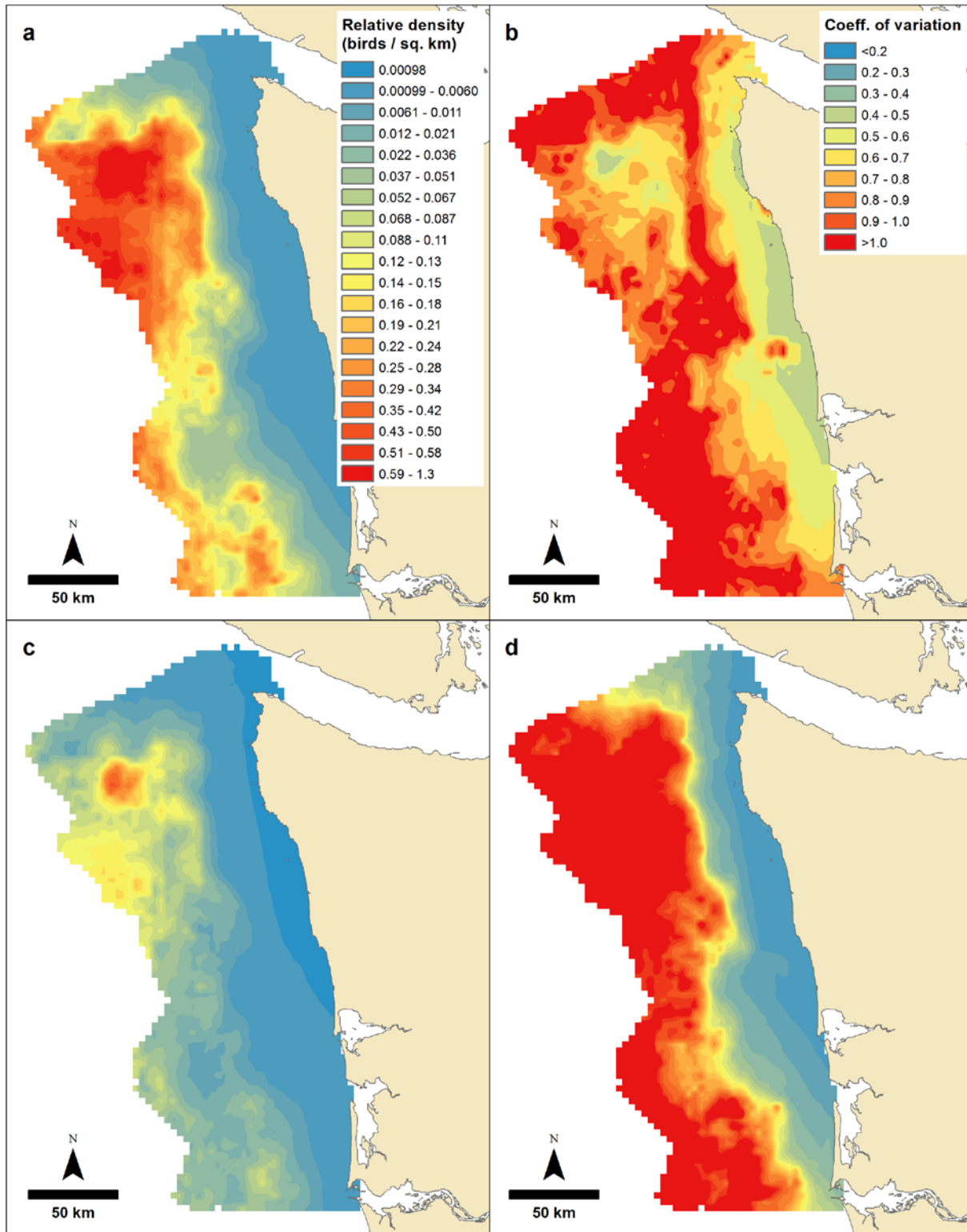


Figure 15: Long-term relative density (birds per sq. km) prediction maps for Black-footed Albatross (*Phoebastria nigripes*) during the months of November to March: a) 50% quantile of bootstrap (median), b) coefficient of variation, c) 5% quantile of bootstrap, d) 95% quantile of bootstrap. The legends for panels c and d are the same as for panel a.

**Northern fulmar (*Fulmarus glacialis*): April to October**

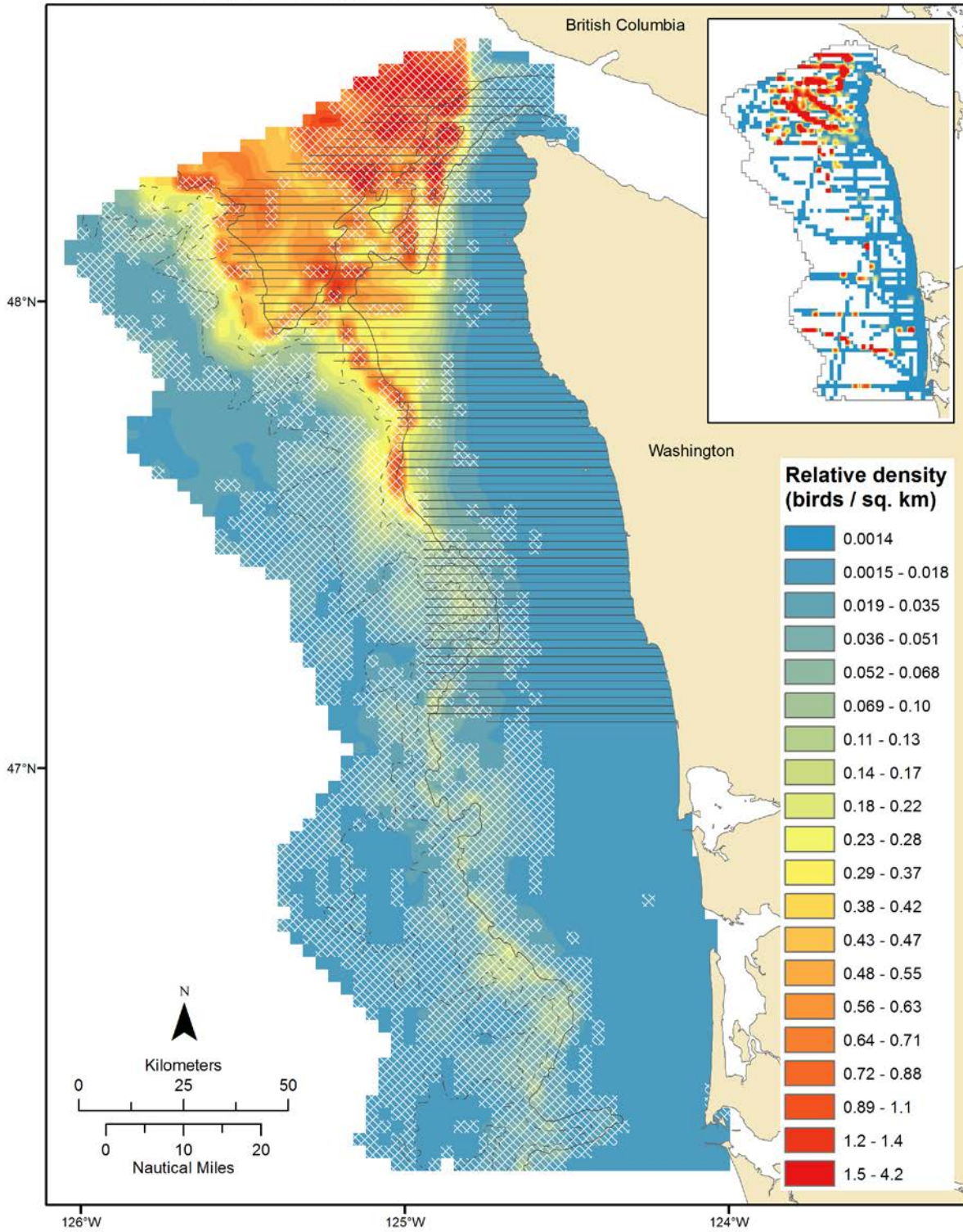


Figure 16: Long-term relative density (birds per sq. km) prediction map for Northern Fulmar (*Fulmarus glacialis*) during the months of April to October. White cross-hatching represents areas of greater prediction uncertainty, where the coefficient of variation was greater than or equal to 0.8. Olympic Coast National Marine Sanctuary is designated by gray horizontal lines. 200, 500, and 1000 m isobath contours are shown as solid, large dashed, and small dashed gray lines, respectively. Observed density (birds per sq. km) is shown on the inset map. Prediction and observed density color gradient classes represent 5% quantile intervals calculated from predictions.



Northern fulmar (*Fulmarus glacialis*): April to October

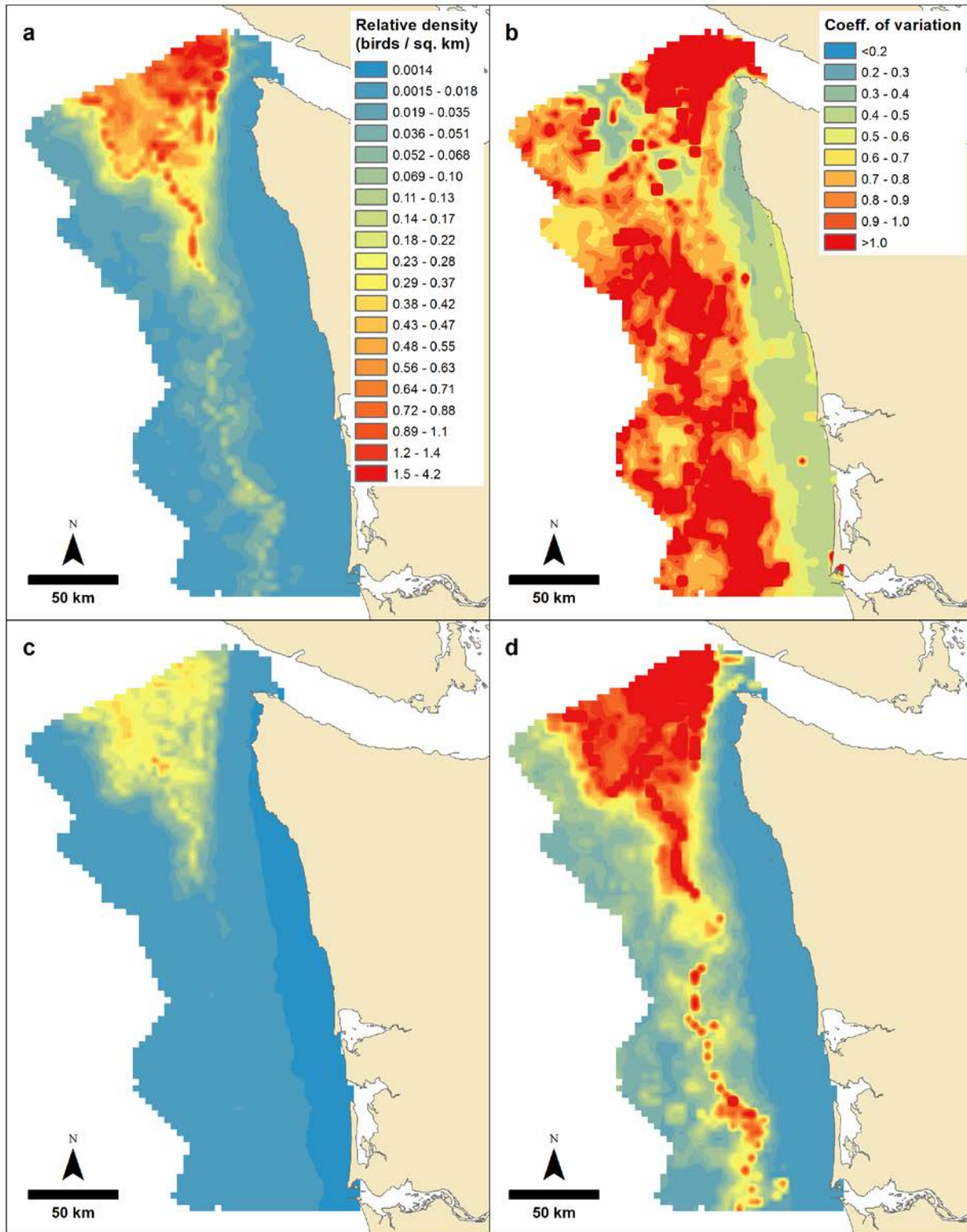


Figure 17: Long-term relative density (birds per sq. km) prediction maps for Northern Fulmar (*Fulmarus glacialis*) during the months of April to October: a) 50% quantile of bootstrap (median), b) coefficient of variation, c) 5% quantile of bootstrap, d) 95% quantile of bootstrap. The legends for panels c and d are the same as for panel a.



**Pink-footed shearwater (*Puffinus creatopus*): April to October**

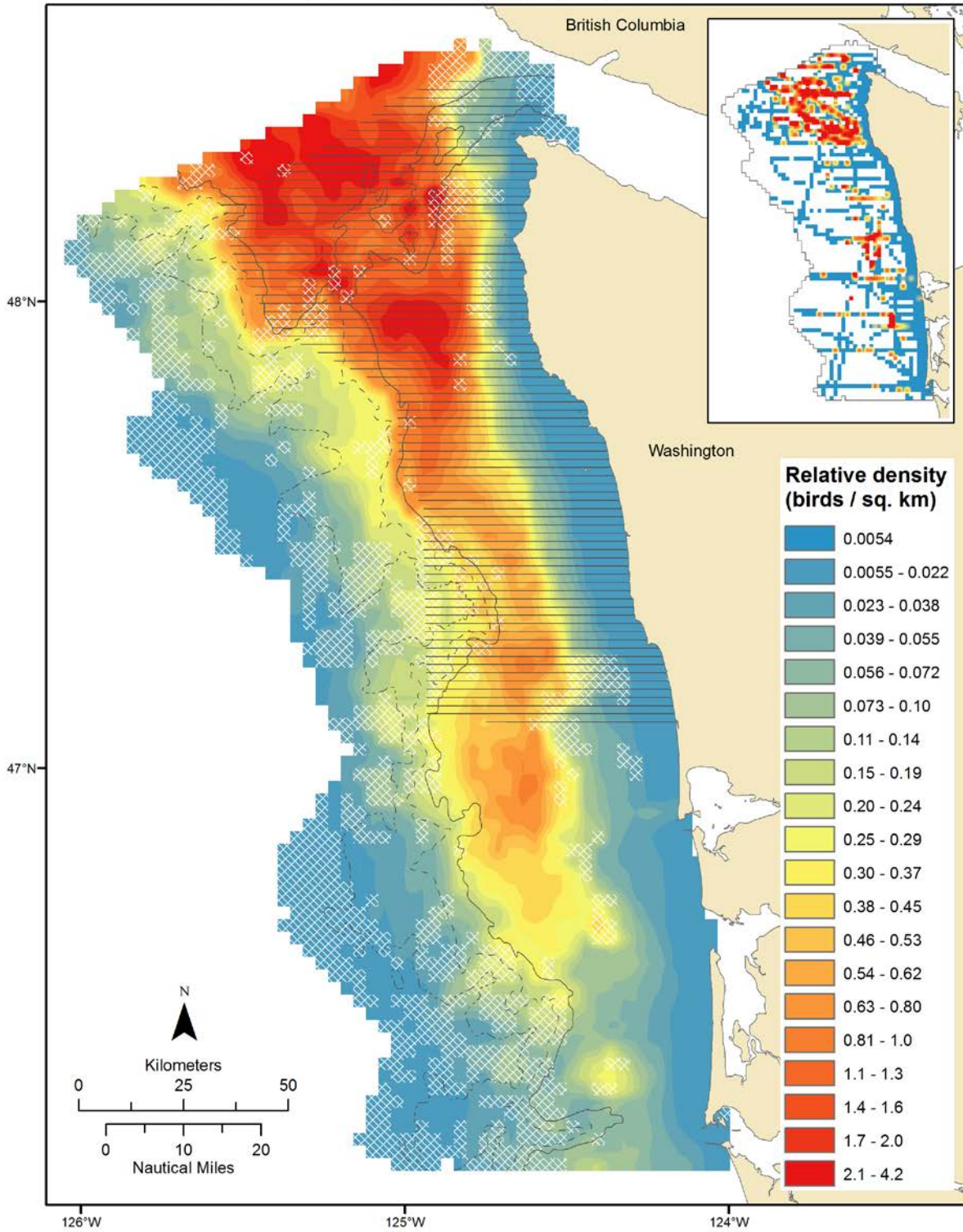


Figure 18: Long-term relative density (birds per sq. km) prediction map for Pink-footed Shearwater (*Puffinus creatopus*) during the months of April to October. White cross-hatching represents areas of greater prediction uncertainty, where the coefficient of variation was greater than or equal to 0.8. Olympic Coast National Marine Sanctuary is designated by gray horizontal lines. 200, 500, and 1000 m isobath contours are shown as solid, large dashed, and small dashed gray lines, respectively. Observed density (birds per sq. km) is shown on the inset map. Prediction and observed density color gradient classes represent 5% quantile intervals calculated from predictions.

**Pink-footed shearwater (*Puffinus creatopus*): April to October**

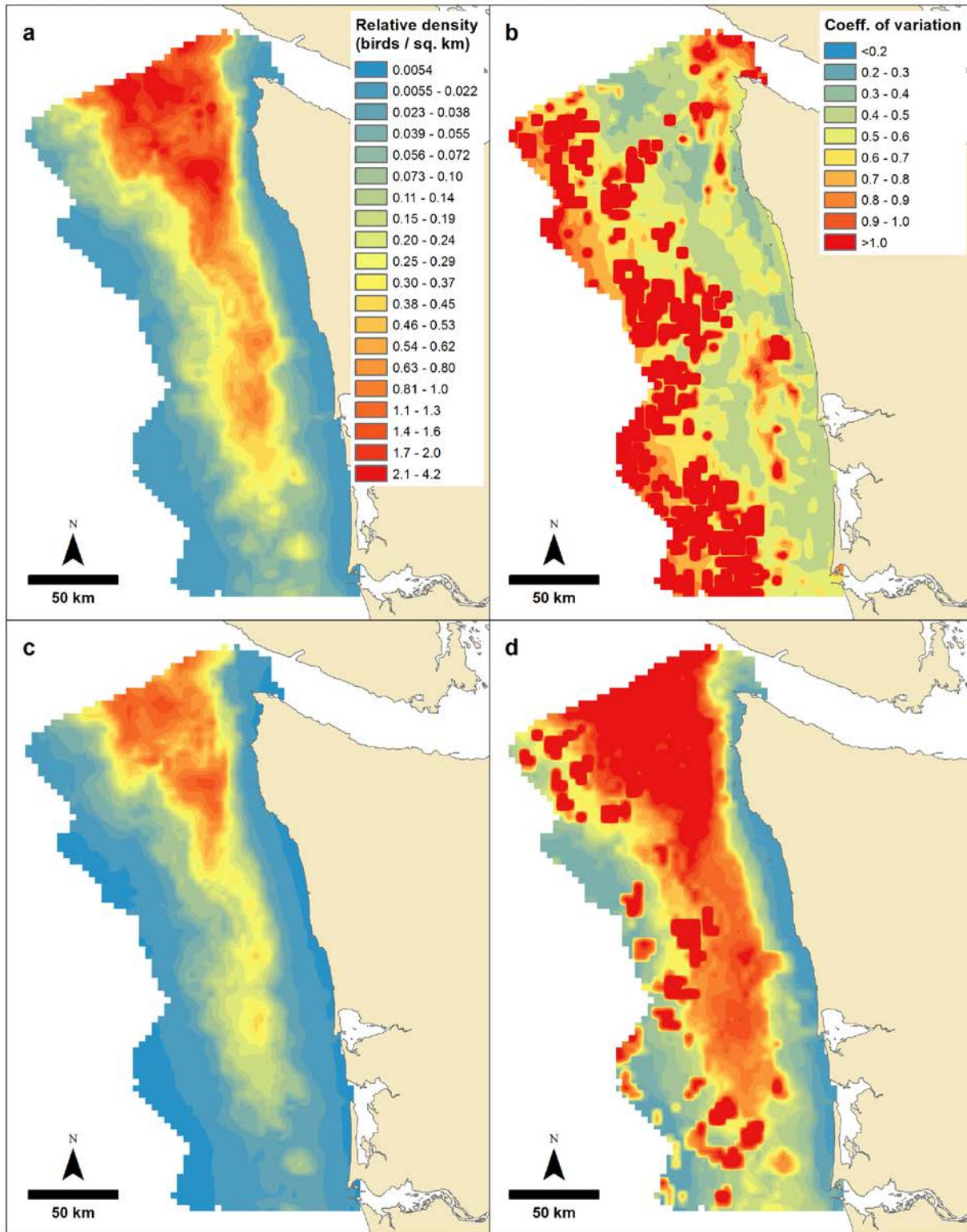


Figure 19: Long-term relative density (birds per sq. km) prediction maps for Pink-footed Shearwater (*Puffinus creatopus*) during the months of April to October: a) 50% quantile of bootstrap (median), b) coefficient of variation, c) 5% quantile of bootstrap, d) 95% quantile of bootstrap. The legends for panels c and d are the same as for panel a.



**Sooty shearwater (*Puffinus griseus*): April to October**

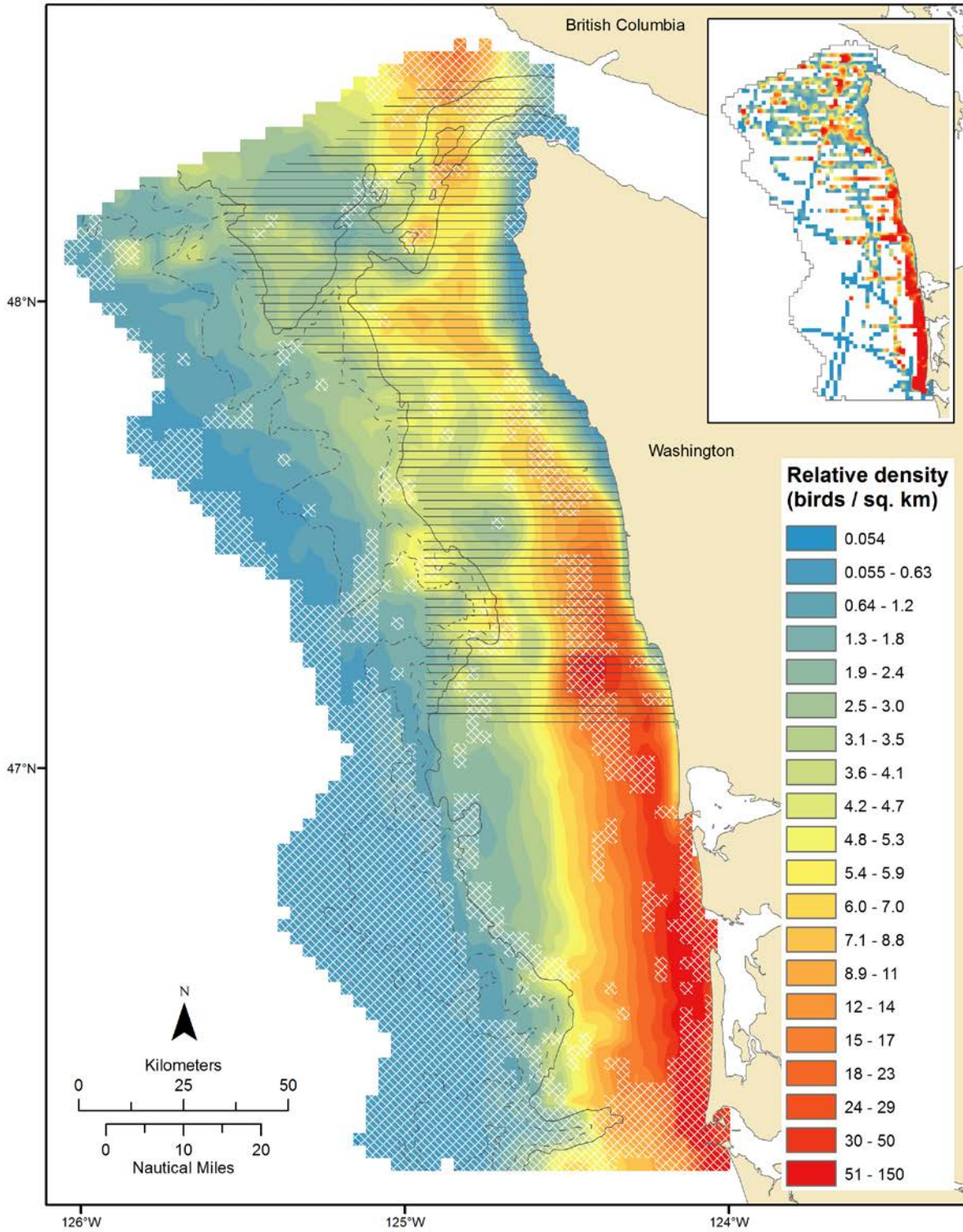


Figure 20: Long-term relative density (birds per sq. km) prediction map for Sooty Shearwater (*Puffinus griseus*) during the months of April to October. White cross-hatching represents areas of greater prediction uncertainty, where the coefficient of variation was greater than or equal to 0.8. Olympic Coast National Marine Sanctuary is designated by gray horizontal lines. 200, 500, and 1000 m isobath contours are shown as solid, large dashed, and small dashed gray lines, respectively. Observed density (birds per sq. km) is shown on the inset map. Prediction and observed density color gradient classes represent 5% quantile intervals calculated from predictions.

Sooty shearwater (*Puffinus griseus*): April to October

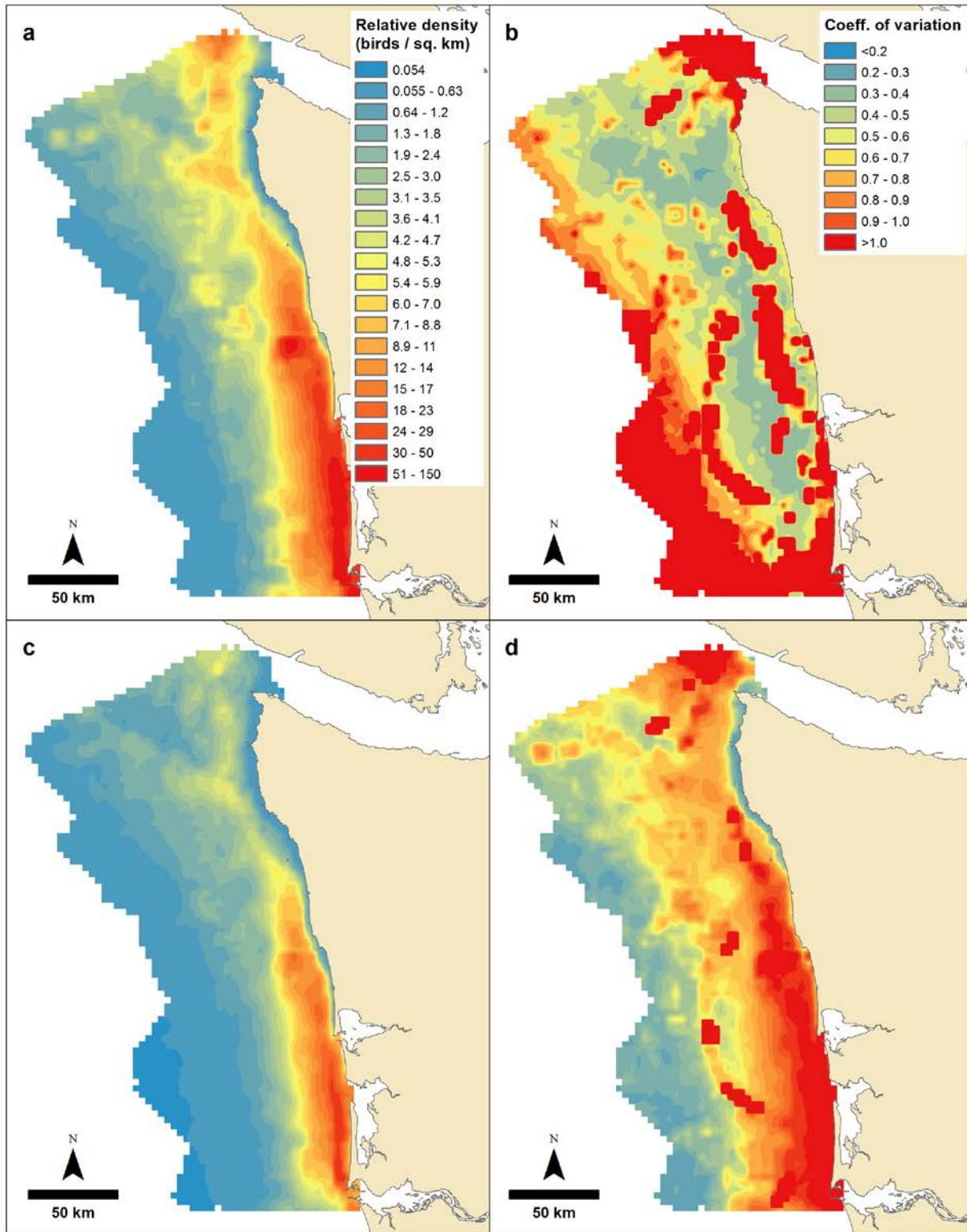


Figure 21: Long-term relative density (birds per sq. km) prediction maps for Sooty Shearwater (*Puffinus griseus*) during the months of April to October: a) 50% quantile of bootstrap (median), b) coefficient of variation, c) 5% quantile of bootstrap, d) 95% quantile of bootstrap. The legends for panels c and d are the same as for panel a.

## Discussion

The spatial models and associated maps and tables presented in this report provide information on the long-term spatial distribution of seven seabird species (Marbled Murrelet, Tufted Puffin, Common Murre, Black-footed Albatross, Northern Fulmar, Pink-footed Shearwater and Sooty Shearwater) from April to October, and two species (Common Murre and Black-footed Albatross) from November to March offshore of Washington. The models and maps are intended to distinguish persistent areas of high relative density from low relative density. It is important to contrast this approach with models and maps which address absolute seabird abundance, which require additional parameters such as probability of species detection. This work was completed to support marine spatial planning by the state of Washington; however, these data will benefit other organizations and other purposes including assessments of marine sanctuary condition, ecosystem health, coastal hazard impacts, and climate change.

All models show good to excellent performance based on model diagnostics and expert review; however users should not assume unqualified accuracy. A model, even a very good one, cannot be a perfect fit in all locations or times. In order to understand if specific points of deficiency exist and where they are located, we emphasize that density maps should be interpreted alongside additional supporting data. In particular, when using these data to make management decisions, we recommend:

- Evaluating model performance diagnostics to better understand the model fit and uncertainty,
- Evaluating spatial and temporal patterns of observations and residuals,
- Interpreting density maps alongside maps of spatially-explicit model uncertainty, represented in this report by the spatially-explicit coefficient of variation, and
- Confirming model findings using independent data, including expert opinion or other seabird observations.

This report provides information to support most of these recommendations. Model performance diagnostics are provided in Table 7, distributions of observations and residuals are in Table 3 and Appendix D, maps of spatially-explicit model uncertainty are provided in Figures 4 through 21, and this report includes expert reviews. We did not review model results against independent observation data and recommend these comparisons for follow up work. Menza et al. (2014) provide a list of seabird survey programs which offer independent data in the study area.

There are also several caveats for supporting data:

- Given the use of zero-inflated distributions, which are inherently complex, certain diagnostics (e.g., residual plots) may be different than when used with more common distributions.
- Any biases in species detection, observed habitats, or temporal periods that are inherent in observation data are propagated into the model results.
- Expert reviews were focused on coarse scale distributional patterns. Fine-scale expert review will be required for site-specific usage.

There are two modeling concepts that would benefit from further research. First, the development of predictive models raises many considerations concerning the appropriate spatial and temporal scales of assessment. These models used a climatological approach, where observations are linked to climatological covariates representing long-term environmental patterns (i.e., climatologies give coarse-scale temporal associations). An alternative approach is to link seabird observations to dynamic, and possibly fine-scale, contemporaneous covariates. Several contributors noted that there are important differences in these two approaches, which may affect model accuracy and predicted spatial patterns. There is an ongoing discussion among contributors, and other academic partners to compare results from the two approaches, but current research suggests that neither approach is

clearly superior to the other, and which approach is used should depend on the objectives of the work. Patrick Halpin at Duke University Marine Lab has forthcoming research on the topic (Brian Kinlan, pers. comm.). Second, the impact of heterogeneous survey effort is assumed to affect results. We present maps of survey effort and tables of effort categorized by distance to shore, year, and month, but do not address the precise impact of effort on predicted patterns of relative density. It is reasonable to assume that model diagnostic metrics are less accurate when and where there is less survey effort, but one should not assume model predictions are worse in the corresponding areas. The impact of heterogeneous survey effort is best evaluated using independent data.

An issue related to both of these modeling concepts is the impact of grouping temporally resolved sightings into seasons. Although the seasonal divisions were helpful to separate distinct life history patterns for most species, grouping sightings has the potential to mask important intra-seasonal changes in distribution, especially if they are out of sync with the oceanographic patterns used to partition seasons in this report. For instance, if a species is breeding from April to July and then feeds offshore from August to October, separate models created for the breeding and feeding periods will likely be better than a single model during the upwelling season. We flagged predictions of Black-footed Albatross in the downwelling season for a similar reason (see species commentary in model results). Evaluating plots of average density per month in Appendix D against known life history patterns can indicate if seasons are unlikely to represent expected species seasonal patterns.

The maps and models in this report are valuable for decision makers today and into the future, but they should be considered as part of an adaptive management strategy. New data sources may become available, new modeling approaches will improve fit of relationships between seabird observations and covariates, and new management objectives will dictate the need for new outputs. Over the next three years we know of two additional projects which will build off of this work. From 2015 to 2018, NCCOS and many other contributors of this modeling work will expand seabird models to the entire U.S. West Coast in support of the Bureau of Ocean Energy Management and in 2015 NCCOS will develop predictive models and maps for marine mammals in support of Washington and the Coastal Zone Management Program.

To encourage adaptive management, and integrate environmental change in a timely manner, there is tremendous value in at-sea observation field program investments. These field programs are immensely helpful to understand population-level distributions across many spatial scales.

Notwithstanding the caveats listed in this discussion, the maps presented here represent the first attempt to combine the eight seabird survey programs listed in Table 2, a substantial combination of nearshore and offshore survey effort. As far as we are aware, the compilation prepared for this report is the largest synthesis of recent seabird observations in the study area, in terms of both number of observations and number of programs combined. As such, these maps represent an important step towards understanding the long-term spatial distributions of the seven selected seabird species, design of effective spatial conservation and sustainable use of Washington's offshore resources.

## References

- Adams, J., J. Felis, J. W. Mason, and J. Y. Takekawa. 2014. Pacific Continental Shelf Environmental Assessment (PaCSEA): aerial seabird and marine mammal surveys off northern California, Oregon, and Washington, 2011-2012. U.S. Dept. of the Interior, Bureau of Ocean Energy Management, Pacific OCS Region, Camarillo, CA. OCS Study BOEM 2014-003. 266 pp.
- Barlow, J., A. Henry, J. V. Redfern, T. Yack, A. Jackson, C. Hall, E. Archer, and L. Ballance. 2010. Oregon, California and Washington line-transect and ecosystem (ORCAWALE) 2008 cruise report. NOAA-TM-NMFS-SWFSC-465.
- Brier, G. W. 1950. Verification of forecasts expressed in terms of probability. *Monthly Weather Review* 78: 1-3.
- Bühlmann, P. and T. Hothorn. 2007. Boosting algorithms: regularization, prediction and model fitting. *Statistical Science* 22(4): 477-505.
- Burger, A. E. 2003. Effects of the Juan de Fuca Eddy and upwelling on densities and distributions of seabirds off southwest Vancouver Island, British Columbia. *Marine Ornithology* 31: 113-122.
- Cotté, C., Y. H. Park, C. Guinet, and C. A. Bost. 2007. Movements of foraging king penguins through marine mesoscale eddies. *Proceedings of the Royal Society B: Biological Sciences* 274(1624): 2385-2391.
- Croxall, J. P., S. H. M. Butchart, B. Lascelles, A. J. Stattersfield, B. Sullivan, A. Symes, and P. Taylor. 2012. Seabird conservation status, threats and priority actions: a global assessment. *Bird Conservation International* 22: 1-34.
- Falxa, G., J. Baldwin, M. Lance, D. Lynch, S. K. Nelson, S. F. Pearson, M. G. Raphael, C. Strong, and R. Young. 2014. Marbled Murrelet effectiveness monitoring, Northwest Forest Plan: 2013 summary report. 20 pp.
- Ehler C. and F. Douvère. 2009. Marine spatial planning: a step-by-step approach toward ecosystem-based management. Intergovernmental Oceanographic Commission and Man and the Biosphere Programme. IOC Manual and Guides, No. 53, IOCAM Dossier No. 6, Paris, UNESCO.
- Friedman, J. H. 2002. Stochastic gradient boosting. *Computational Statistics & Data Analysis* 38: 367-378.
- Gneiting, T. and A. E. Raftery. 2007. Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association* 102: 359-378.
- Hanson, M. B., D. P. Noren, T. F. Norris, C. A. Emmons, M. M. Holt, E. Phillips, J. Zamon, and J. Menkel. 2010. Pacific Orca Distribution Survey (PODS) conducted aboard the NOAA ship McArthur II in March-April 2009. (STATE DEPT. CRUISE NO: 2009-002) Unpublished report, NWFSC, Seattle, WA.
- Hoefler, C. J. 2000. Marine bird attraction to thermal fronts in the California current system. *The Condor* 102(2): 423-427.
- Hofner, B., A. Mayr, N. Robinzonov, and M. Schmid. 2012. Model-based boosting in R: a hands-on tutorial using the R package mboost. Technical Report 120, Department of Statistics, University of Munich, Germany.



- Hothorn, T., P. Buehlmann, T. Kneib, M. Schmid, and B. Hofner. 2015. mboost: Model-Based Boosting, R package version 2.4-2. URL <http://CRAN.R-project.org/package=mboost>.
- Kinlan, B. P., R. Rankin, A. Winship, and C. Caldow. 2014. Modeling At-Sea Occurrence and Abundance of Marine Birds to Support Mid-Atlantic Marine Renewable Energy Planning. U.S. Department of the Interior, Bureau of Ocean Energy Management, Herndon, VA. OCS Study BOEM 2014. NOAA Technical Memorandum NOS NCCOS. In review.
- Kracker, L. and C. Menza. 2015. An Evaluation of Marine Mammal Surveys to Support Washington State's Marine Spatial Planning Process. NOAA Technical Memorandum NOS NCCOS 195. Silver Spring, MD. 37 pp.
- Mann, K. H. and J. R. N. Lazier. 2006. Dynamics of Marine Ecosystems: Biological-Physical Interactions in the Oceans. Third Edition. Blackwell Publishing. pp. 166–167, 194–204.
- MacFadyen A., B. M. Hickey, and M. G. G. Foreman. 2005. Transport of surface waters from the Juan De Fuca eddy region to the Washington coast. *Continental Shelf Research* 25: 2008–2021.
- Mayr, A., N. Fenske, B. Hofner, T. Kneib, and M. Schmid. 2012. Generalized additive models for location, scale and shape for high dimensional data---a flexible approach based on boosting. *Applied Statistics* 61(3): 403-427.
- McGillicuddy, D. J., A. R. Robinson, D. A. Siegel, H. W. Jannasch, R. Johnson, T. D. Dickey, J. McNeil, A. F. Michaels, and A. H. Knap. 1998. Influence of mesoscale eddies on new production in the Sargasso Sea. *Nature* 394: 263-266.
- Menza, C., T. Battista, and D. Dorfman. 2014. Technical and Mapping Support for Washington State Marine Spatial Planning. NOAA Technical Memorandum NOS NCCOS 181. Prepared by the NCCOS Center for Coastal Monitoring and Assessment Biogeography Branch. Silver Spring, MD. 37 pp.
- Naughton, M. B, M. D. Romano, T. S. Zimmerman. 2007. A Conservation Action Plan for Black-footed Albatross (*Phoebastria nigripes*) and Laysan Albatross (*P. immutabilis*), Ver. 1.0.
- Nur, N., J. Jahncke, M. P. Herzog, J. Howar, K. D. Hyrenbach, J. E. Zamon, D. G. Ainley, J. A. Wiens, K. Morgan, L. T. Balance, and D. Stralberg. 2011. Where the wild things are: predicting hotspots of seabird aggregations in the California Current System. *Ecological Applications* 21(6): 2241-2257.
- Office of National Marine Sanctuaries. 2008. Olympic Coast National Marine Sanctuary Condition Report 2008. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Office of National Marine Sanctuaries, Silver Spring, MD. 72 pp.
- Olson, D. B., G. L. Hitchcock, A. J. Mariano, C. J. Ashjian, G. Peng, R. W. Nero, and G. P. Podesta. 1994. Life on the edge: marine life and fronts. *Oceanography* 7(2): 52-60.
- Paleczny M, Hammill E, Karpouzi V, Pauly D. 2015. Population Trend of the World's Monitored Seabirds, 1950-2010. *PLoS ONE* 10(6): e0129342.
- R Core Team. 2015. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.



- Roberts, J. J., B. D. Best, D. C. Dunn, E. A. Treml, and P. N. Halpin. 2010. Marine Geospatial Ecology Tools: An integrated framework for ecological geoprocessing with ArcGIS, Python, R, MATLAB, and C++. *Environmental Modelling & Software* 25: 1197-1207.
- Sbrocco, E. J. and P. H. Barber. 2013. MARSPEC: ocean climate layers for marine spatial ecology. *Ecology* 94(4): 979-979.
- Schreiber, E. A. and J. Burger. 2001. *Biology of Marine Birds*. Boca Raton, FL. CRC Press. 744 pp.
- Schmid, M., S. Potapov, A. Pfahlberg, and T. Hothorn. 2008. Estimation and regularization techniques for regression models with multidimensional prediction functions. Technical Report 42, Department of Statistics, University of Munich, Germany.
- Suryan, R. M., J. A. Santora, and W. J. Sydeman. 2012. New approach for using remotely-sensed chlorophyll *a* to identify seabird hotspots. *Marine Ecology Progress Series* 451: 213-225
- U.S. Fish and Wildlife Service. 2005. *Regional Seabird Conservation Plan, Pacific Region*. U.S. Fish and Wildlife Service, Migratory Birds and Habitat Programs, Pacific Region, Portland, OR.
- U.S. Fish and Wildlife Service. 2013. *Birding in the United States: A Demographic and Economic Analysis*. Report 2011-1. U.S. Fish and Wildlife Service Division of Economics, Arlington, VA.
- Washington Department of Ecology. 2014. *Marine Spatial Plan for Washington's Pacific Coast: Summary of SEPA Scoping and Response to Comments*. Washington State Department of Ecology. 62 pp.
- Washington Department of Fish and Wildlife. 2015. *List of Species of Concern in Washington State*. Website <http://wdfw.wa.gov/conservation/endangered/All/> . Last accessed June 29, 2015.
- Wiebe, P. H., E. M. Hulbert, E. J. Carpenter, A. E. Jahn, G. P. Knapp III, S. H. Boyd, P. B. Ortner, and J. L. Cox. 1976. Gulf Stream cold core rings: large-scale interaction sites for open ocean plankton communities. *Deep Sea Research and Oceanographic Abstracts* 23(8): 695-710.
- Yen, P. P. W., W. J. Sydeman, and K. D. Hyrenbach. 2004. Marine bird and cetacean associations with bathymetric habitats and shallow-water topographies: implications for trophic transfer and conservation. *Journal of Marine Systems* 50: 79-99.
- Zamon, J. E., E. M. Phillips, and T. J. Guy. 2014. Marine bird aggregations associated with the tidally-driven plume and plume fronts of the Columbia River. *Deep-Sea Research Part II: Tropical Studies in Oceanography* 107: 85-95.

## Appendix A: Seabird survey program descriptions

Dataset Name	Data Collector(s)	Years	Geographic Coverage	Reference
Pacific Continental Shelf Environmental Assessment (PaCSEA)	USGS WERC, BOEM	2011-12	Grays Harbor to Columbia River; Oregon and northern California; shore to continental slope	Adams et al. 2014
ORegon, California, and WASHINGTON Line-transect Expedition (ORCAWALE) marine mammal survey	NOAA SWFSC	2001, 2008	US-Mexico to US-Canada border and seaward to 300 nmi	Barlow et al. 2010
Collaborative Survey of Cetacean Abundance and the Pelagic Ecosystem (CSCAPE)	NOAA SWFSC	2005	US-Mexico to US-Canada border and seaward to 300 nmi; but finer scale information was collected in National Marine Sanctuaries	Barlow et al. 2010
Seasonal Olympic Coast National Marine Sanctuary Surveys	NOAA OCNMS	2006-12	Pelagic OCNMS; La Push to Juan de Fuca Canyon	Online report: <a href="http://olympiccoast.noaa.gov/science/surveycruises/2012/marin seabirds.html">http://olympiccoast.noaa.gov/science/surveycruises/2012/marin seabirds.html</a>
Annual Olympic Coast National Marine Sanctuary Surveys	NOAA OCNMS	1995-2007	Grays Harbor to Cape Flattery; 20-500 fathoms	Online report: <a href="http://olympiccoast.noaa.gov/science/surveycruises/2011/seabird_density.html">http://olympiccoast.noaa.gov/science/surveycruises/2011/seabird_density.html</a>
NWFSC Northern California Current Seabird Surveys	NOAA NWFSC (Conservation Biology, Fish Ecology)	2008, 2009, 2012	Depends on location of Southern Resident Killer Whales; ranges from San Francisco, CA to Strait of Juan de Fuca, WA	Hanson et al. 2010
Northwest Forest Plan Marbled Murrelet Effectiveness Monitoring Program	USFS, USFWS, WDFW	2000-13	Outer Washington coast; shore to 8 km; also Salish Sea, Oregon, and California to San Francisco	Falxa et al. 2014
Pacific Coast Winter Seaduck Survey	Sea Duck Joint Venture, WDFW	2011	Cape Flattery to OR-CA border; shore to either 20 fathoms (36.58 m) or 3 nmi (5056 km), whichever was greater	Contact: Tim Bowman (USFWS)
Bonneville power plant seabird surveys	NOAA NWFSC (Fish Ecology)	2008-12	Coastal WA on the continental shelf edge	Zamon et al. 2014

## **Appendix B: Processing steps for environmental predictors**

### **Climate indices**

The North Pacific Gyre Oscillation Index, Pacific Decadal Oscillation Index, and upwelling index were processed by calculating a three-month moving average (using the values of the current month plus the previous two months) prior to analyses. A three-month moving average was not calculated for Multivariate El Niño-Southern Oscillation Index, as these data had already been ‘smoothed’ using a two-month moving average prior to obtaining them. For each climate index, two values were included as predictor variables: the value for the month and year of a given transect segment and the value for the same month one year previous (i.e., 12 month lag).

### **Seasonal mesoscale eddy probabilities**

Oceanic eddies are large circular currents with scales ranging from 10s to 100s of kilometers. Eddies can transfer nutrients across water masses and elevate primary production in upwelling cores (McGillicuddy et al. 1998), retain zooplankton (Weibe et al. 1976), and enhance top predator densities (Cotté et al. 2007, Burger 2006, Yen et al. 2006). Rotational patterns of mesoscale eddies are cyclonic or anticyclonic. Centers of anticyclonic eddies are concave down and are referred to as downwelling eddies and warm-core rings. Cyclonic eddies exhibit an opposite rotational pattern with concave-up (higher) centers and are referred to as upwelling eddies and cold-core rings.

Oceanic eddy climatologies were derived from a 21-year dataset (1993-2014) of daily AVISO sea surface height (SSH) imagery, specifically Mean Absolute Dynamic Topography (MADT). The AVISO website (<http://www.aviso.altimetry.fr/en/data/products/sea-surface-height-products/global/madt.html>) provides additional information for MADT and geospatial data layers in NetCDF format. The spatial resolution of AVISO SSH data is 0.25 degrees or approximately 25 km. The Okubo-Weiss Algorithm was applied to MADT using the Marine Geospatial Ecology Tools (MGET) geoprocessing ArcGIS toolbox (Roberts et al. 2010; <http://mgel.env.duke.edu/mget>) to identify anticyclonic and cyclonic eddies. The Okubo-Weiss algorithm parameter, “Minimum area-to-perimeter ratio of eddy cores” was modified to 0.45 in order to select for circular eddies; however, all other parameters remained set at default levels (Okubo-Weiss parameter threshold type = -0.2; Minimum area of eddy cores = 4; Minimum duration of eddy cores = 28 days).

Eddy climatologies are probability layers that estimate the probability of anticyclonic or cyclonic eddies in two seasons: April to October and November to March. Eddy probabilities were calculated at coincident pixels using the native resolution of MADT by summing the number of times each pixel was classified as an eddy, divided by the number of pixels with SSH data at the scale of the pixel frame.

### **Seasonal frontal probabilities**

Hydrographic fronts manifest across a wide variety of spatial and temporal scales and some may facilitate trophic energy transfer, where high concentrations of prey associated with fronts attract marine predators searching for food (Schneider 1990, Hoefler 2000, Olson et al. 2000).

A 13-year dataset of monthly sea surface temperature front probability composites (accessed from <http://coastwatch.pfeg.noaa.gov/erddap/griddap/erdGAtfntmday.html>) was used to create seasonal climatologies for two modeled time periods: April to October and November to March. The CoastWatch Oceanic Front Probability Index measures the probability of sea surface temperature front formation based on data from NOAA's GOES satellites. For each time period, front probabilities layers were averaged using the native resolution of the data (0.05 degrees; ~5.5 km) in order to calculate each seasonal climatological mean.

## **Bathymetry**

Seafloor topography has strong direct effects on marine ecosystems by steering the flow field and providing habitat for marine organisms. Discrete topographic features are important spatial predictors of seabird distribution and abundance, and influence foraging distributions of seabirds across a variety of spatial scales (e.g., continental shelf breaks / shelf slopes, submarine canyons, ledges, and shoals; Croll et al. 1998, Yen et al. 2004, Nur et al. 2011).

Depth, topographic slope, and planform and profile curvature data were taken from MARSPEC, a high-resolution GIS database of ocean climate layers intended for marine ecological niche modeling and other applications in marine spatial ecology (Sbrocco and Barber 2013; <http://www.marspec.org/>). MARSPEC uses the SRTM30 Plus Bathymetry version 6.0 data set for bathymetry (accessed from [http://topex.ucsd.edu/WWW\\_html/srtm30\\_plus.html](http://topex.ucsd.edu/WWW_html/srtm30_plus.html)). Bathymetric slope was measured in degrees ranging from 0° (flat surface) to 90° (vertical slope). Planform and profile curvature, derivatives of bathymetry, were generated from bathymetric data and the Surface tools within the Spatial Analyst toolbox in ArcGIS 9.3.1. Curvature layers are used to infer flow-field dynamics. Positive/negative values of planform curvature may indicate divergent/convergent flow, whereas, positive/negative values of profile curvature may indicate acceleration/deceleration of the flow field (Sbrocco and Barber 2013).

The bathymetric position index (BPI) is a measure of depth relative to a surrounding neighborhood. The bathymetric position index was calculated for 2 spatial scales (3 km and 20 km) to capture topographic features and complexity of the seafloor (e.g., flat bottom, trough, and steep wall) at moderate and coarse spatial scales. BPI for both scales was calculated with the Spatial Analyst toolbox in ArcGIS 10.2 using an annulus neighborhood with an inner radius of 1 cell and outer radius approximately half the corresponding scale.

## **Seasonal salinity**

Monthly sea surface salinity climatologies were extracted from the MARSPEC monthly climatological dataset (Sbrocco and Barber 2013). The average of monthly climatologies from April to October and November to March were used to produce two seasonal climatologies.

## **Seasonal sea surface temperature**

Sea surface temperature data was derived from Aqua MODIS, processed to a level 3 monthly composite for years 2002-12. Monthly composites were averaged into two seasonal climatologies: April to October and November to March.

## **Seasonal chlorophyll *a* concentration**

Chlorophyll *a* data was derived from Aqua MODIS, processed to a level 3 monthly composite for years 2002-12. Monthly composites were averaged to create two seasonal climatologies: April to October and November to March.

## **Frequency of chlorophyll peaks index (FCPI)**

Chlorophyll *a* data for this climatology was derived from the SeaWiFS satellite, unlike the mean climatologies derived from the MODIS sensor. FCPI is a 9-year index that represents chlorophyll *a* intensity above a modeled mean of chlorophyll *a* concentration across all months (January 1998 through December 2006). The index is not seasonal, so the same climatological values are used for both upwelling and downwelling seasons. See Suryan et al. 2012 for details.

## **Distance to 200 m isobath**

The 200 m isobath was derived from the MARSPEC depth layer and distance to the 200 m isobath was calculated using the Spatial Analyst toolbox in ArcGIS 10.2 across the entire study area at a resolution of 100 m.

**Distance to colonies and critical habitat**

Locations of Common Murre and Tufted Puffin colonies were extracted from the Washington Seabird Catalog and converted to shapefiles as points. The locations of Marbled Murrelet nesting habitat were taken from the Final Revised Marbled Murrelet Critical Habitat Designation which can be downloaded from the USFWS Washington Fish and Wildlife Office (<http://www.fws.gov/wafwo/mamu.html>). Marbled Murrelets nest inland in forests that are generally characterized by large trees with large branches or deformities for use as nest platforms. Distance to colonies and critical habitat were calculated using the Spatial Analyst toolbox in ArcGIS 10.

## Appendix C: Model performance metrics (full model assemblage)

Performance metrics for all models of each species-season combination. Models are sorted in descending order of performance in terms of the thresholded continuous rank probability score (CRPS) from cross-validation tuning of the number of boosting iterations (column marked with \*). The selected “best” model is in bold font. All model performance metrics were calculated on the full dataset, except columns labeled “Cross-val,” which denote statistics calculated separately on cross-validation data. AUC values were calculated as the area under the relevant ROC curve. Rank R refers to Spearman’s rank correlation coefficient (Spearman’s rho statistic) for the observed vs. predicted count. Median absolute [residual] error (non-zero) and median [residual] bias (non-zero) were standardized by dividing by mean no. individuals per segment with sightings and multiplying by 100, so that values shown represent percentages of the mean number of individuals observed per segment on segments with at least one sighting. For details on the Brier score and continuous ranked probability score (CRPS), please consult Brier (1950) and Gneiting and Raftery (2007). Note that risk is not directly comparable across model types (e.g., ZIP vs. ZINB).

Species code	Season	Model type	Max. tree depth	No. boosting iterations	AUC		Rank R (non-zero)	Gaussian rank correlation (non-zero)		Median absolute error (non-zero)		Median bias (non-zero)		Brier score (occupancy)		CRPS		Log-likelihood	Risk (cross-val)
					Occupancy	Non-zero		correlation	Fit	Cross-val	Fit	Cross-val	Fit	Cross-val	Fit	Cross-val*			
mamu	summer	ZINB	4	19,999	0.92	0.76	0.62	0.64	37.7%	32.4%	-33.2%	-24.7%	0.080	0.098	0.073	0.089	-5287	2717	
mamu	summer	ZINB	5	19,999	0.92	0.76	0.62	0.64	38.0%	32.2%	-33.2%	-24.9%	0.081	0.098	0.074	0.089	-5317	2715	
mamu	summer	ZIP	5	20,000	0.92	0.67	0.46	0.50	34.7%	32.2%	-29.9%	-23.0%	0.082	0.097	0.075	0.089	-5910	2933	
mamu	summer	ZIP	4	20,000	0.92	0.66	0.45	0.49	34.5%	32.0%	-29.5%	-23.2%	0.082	0.098	0.075	0.090	-5956	2934	
tupu	summer	ZIP	4	19,995	0.92	0.75	0.60	0.63	21.6%	23.5%	-11.5%	-6.7%	0.084	0.094	0.077	0.086	-7982	4241	
tupu	summer	ZIP	5	18,998	0.91	0.74	0.57	0.61	22.8%	25.2%	-11.0%	-6.9%	0.089	0.102	0.081	0.093	-8480	4426	
tupu	summer	ZINB	5	20,000	0.91	0.77	0.62	0.63	22.9%	23.4%	-12.7%	-9.5%	0.091	0.106	0.081	0.095	-6404	3184	
tupu	summer	ZINB	4	19,999	0.91	0.77	0.62	0.64	22.0%	23.4%	-12.6%	-9.4%	0.090	0.106	0.081	0.095	-6403	3200	
comu	summer	ZINB	5	17,999	0.91	0.84	0.70	0.70	20.1%	27.1%	4.5%	12.0%	0.113	0.137	0.102	0.124	-29549	12673	
comu	summer	ZINB	4	17,998	0.91	0.84	0.70	0.70	20.3%	27.0%	4.9%	11.9%	0.113	0.137	0.102	0.124	-29552	12684	
comu	summer	ZIP	5	17,953	0.92	0.86	0.74	0.74	16.5%	19.3%	3.0%	8.6%	0.107	0.124	0.107	0.130	-116272	59537	
comu	summer	ZIP	4	17,961	0.92	0.86	0.74	0.75	16.0%	19.5%	2.6%	8.4%	0.107	0.125	0.106	0.131	-115756	62410	
comu	winter	ZIP	5	8,611	0.91	0.82	0.69	0.70	19.6%	23.9%	-2.9%	8.3%	0.118	0.110	0.103	0.098	-3479	1560	
comu	winter	ZIP	4	8,954	0.91	0.82	0.69	0.70	19.2%	23.4%	-2.2%	6.6%	0.117	0.110	0.102	0.100	-3376	1629	
comu	winter	ZINB	5	17,668	0.88	0.84	0.71	0.72	18.4%	17.9%	-4.5%	0.4%	0.136	0.123	0.114	0.101	-1730	534	
comu	winter	ZINB	4	17,971	0.88	0.84	0.71	0.72	19.7%	19.0%	-4.7%	0.4%	0.135	0.126	0.113	0.104	-1727	536	
bfal	summer	ZINB	5	15,215	0.96	0.70	0.41	0.44	12.1%	11.9%	-11.1%	-11.6%	0.027	0.023	0.026	0.022	-1497	546	
bfal	summer	ZINB	4	15,379	0.96	0.71	0.43	0.45	12.2%	12.0%	-11.6%	-11.6%	0.027	0.023	0.026	0.022	-1506	540	
bfal	summer	ZIP	4	12,615	0.91	0.59	0.16	0.18	13.1%	12.3%	-12.4%	-11.9%	0.029	0.024	0.028	0.023	-2677	889	
bfal	summer	ZIP	5	3,089	0.90	0.65	0.31	0.34	12.6%	12.8%	-11.8%	-11.6%	0.028	0.025	0.028	0.024	-4731	2798	
bfal	winter	ZIP	4	17,986	0.97	0.78	0.46	0.43	37.0%	37.4%	-34.8%	-33.1%	0.032	0.028	0.031	0.027	-224	69	
bfal	winter	ZIP	5	17,952	0.97	0.76	0.43	0.42	35.2%	43.6%	-31.8%	-37.1%	0.032	0.029	0.031	0.028	-224	71	
bfal	winter	ZINB	4	16,743	0.97	0.83	0.57	0.54	39.7%	47.7%	-35.6%	-38.0%	0.033	0.029	0.032	0.028	-232	70	
bfal	winter	ZINB	5	16,181	0.97	0.83	0.55	0.51	41.5%	48.4%	-38.4%	-37.7%	0.034	0.030	0.033	0.029	-238	74	
nofu	summer	ZINB	4	17,733	0.97	0.75	0.57	0.61	15.9%	15.8%	-11.3%	-11.1%	0.022	0.019	0.022	0.018	-1645	522	
nofu	summer	ZINB	5	17,969	0.97	0.75	0.56	0.59	16.3%	16.0%	-11.7%	-11.5%	0.022	0.019	0.022	0.018	-1654	525	
nofu	summer	ZIP	5	17,970	0.96	0.72	0.50	0.56	19.5%	17.6%	-13.3%	-13.9%	0.022	0.019	0.022	0.018	-2259	1103	
nofu	summer	ZIP	4	16,333	0.96	0.71	0.48	0.54	19.1%	16.5%	-13.2%	-13.6%	0.022	0.019	0.022	0.018	-2408	1089	
pfsh	summer	ZIP	4	17,962	0.96	0.65	0.39	0.43	22.1%	19.5%	-12.5%	-11.1%	0.036	0.026	0.035	0.025	-4084	1365	
pfsh	summer	ZIP	5	16,509	0.96	0.66	0.40	0.44	21.4%	20.1%	-12.1%	-10.5%	0.036	0.027	0.035	0.026	-4046	1416	
pfsh	summer	ZINB	5	17,997	0.96	0.68	0.45	0.47	23.7%	20.2%	-13.2%	-12.3%	0.036	0.027	0.035	0.026	-2558	757	
pfsh	summer	ZINB	4	17,991	0.96	0.67	0.43	0.45	22.0%	20.9%	-13.3%	-12.8%	0.036	0.027	0.035	0.026	-2567	758	
sosh	summer	ZINB	4	17,999	0.91	0.72	0.49	0.52	8.3%	8.3%	1.5%	1.5%	0.110	0.092	0.098	0.081	-12882	4304	
sosh	summer	ZINB	5	17,999	0.91	0.72	0.51	0.53	8.0%	8.2%	1.3%	1.5%	0.110	0.092	0.098	0.081	-12876	4293	
sosh	summer	ZIP	5	14,643	0.88	0.72	0.51	0.56	6.2%	6.1%	1.2%	1.2%	0.124	0.103	0.113	0.093	-110921	59066	
sosh	summer	ZIP	4	14,035	0.88	0.71	0.51	0.55	6.5%	6.3%	1.3%	1.2%	0.125	0.103	0.114	0.093	-109763	52523	

## Appendix D: Select marginal and residual plots

Figures D1-9: Select marginal and residual plots for the final selected model of each species-season combination. Panels from left to right, top to bottom are: 1) observed monthly average density (birds per sq. km); 2) marginal plot of  $\text{logit}(1 - p)$  vs. Julian day; 3) marginal plot of  $\log(\mu)$  vs. Julian day; 4) blank; 5) marginal plot of  $\text{logit}(1 - p)$  vs. year; 6) marginal plot of  $\log(\mu)$  vs. year; 7) scatterplot of Pearson residuals vs. latitude; 8) scatterplot of Pearson residuals vs. longitude; 9) scatterplot of Pearson residuals vs. date. Marginal plots show patterns in the functional relationship between transformed versions of the zero-inflation and count model components and Julian day and year, after accounting for the effects of all other predictors in the model. Gray shading represents  $\pm 1$  standard deviation from the mean. Absent marginal plots indicate either Julian day or year was not selected as an important predictor in the final model.

Panel schema

1	2	3
4	5	6
7	8	9

Marbled murrelet (*Brachyramphus marmoratus*): April to October

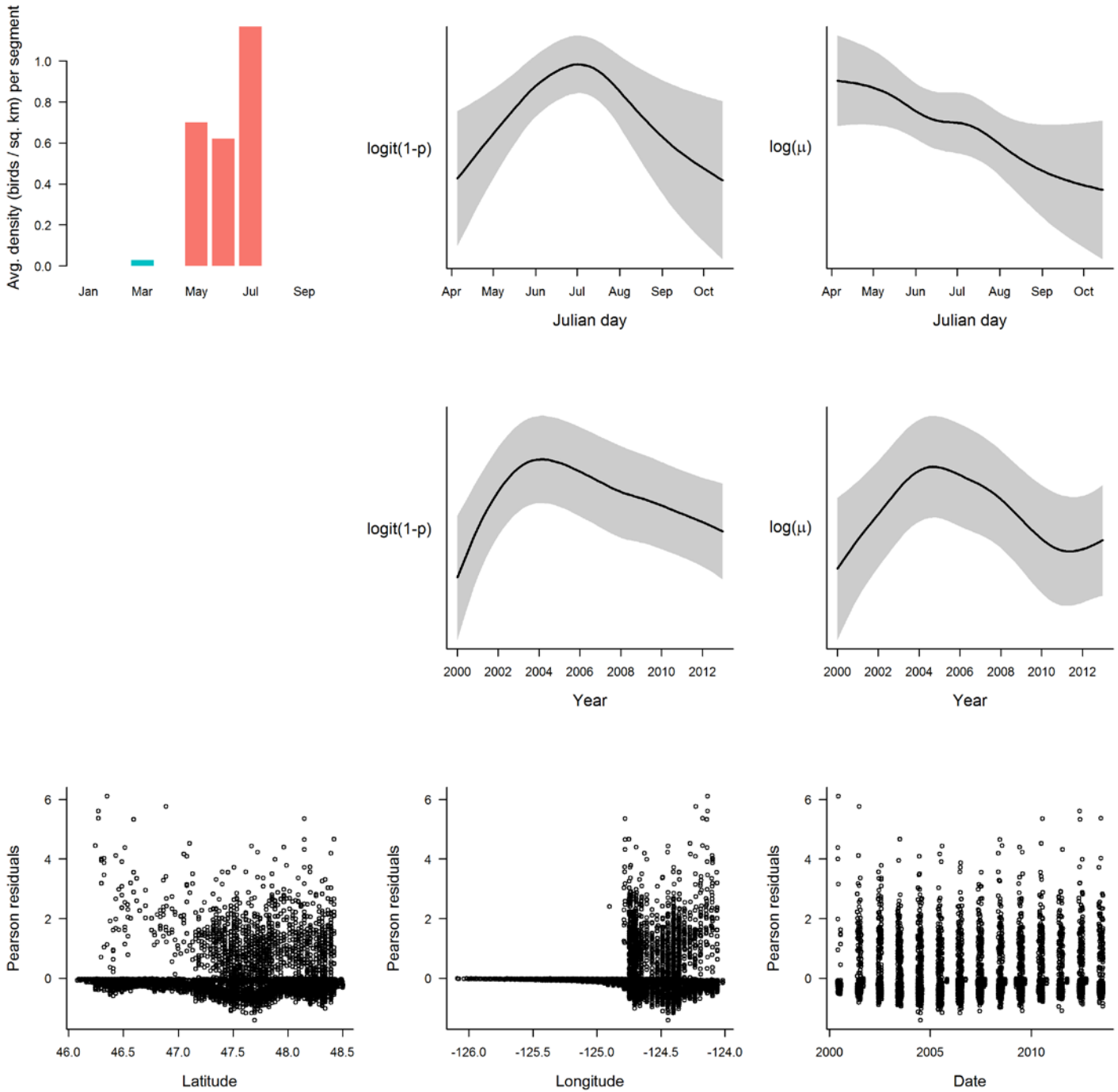


Figure D1: Marginal and residual plots for Marbled Murrelet (*Brachyramphus marmoratus*) during the months of April to October.



Tufted puffin (*Fratercula cirrhata*): April to October

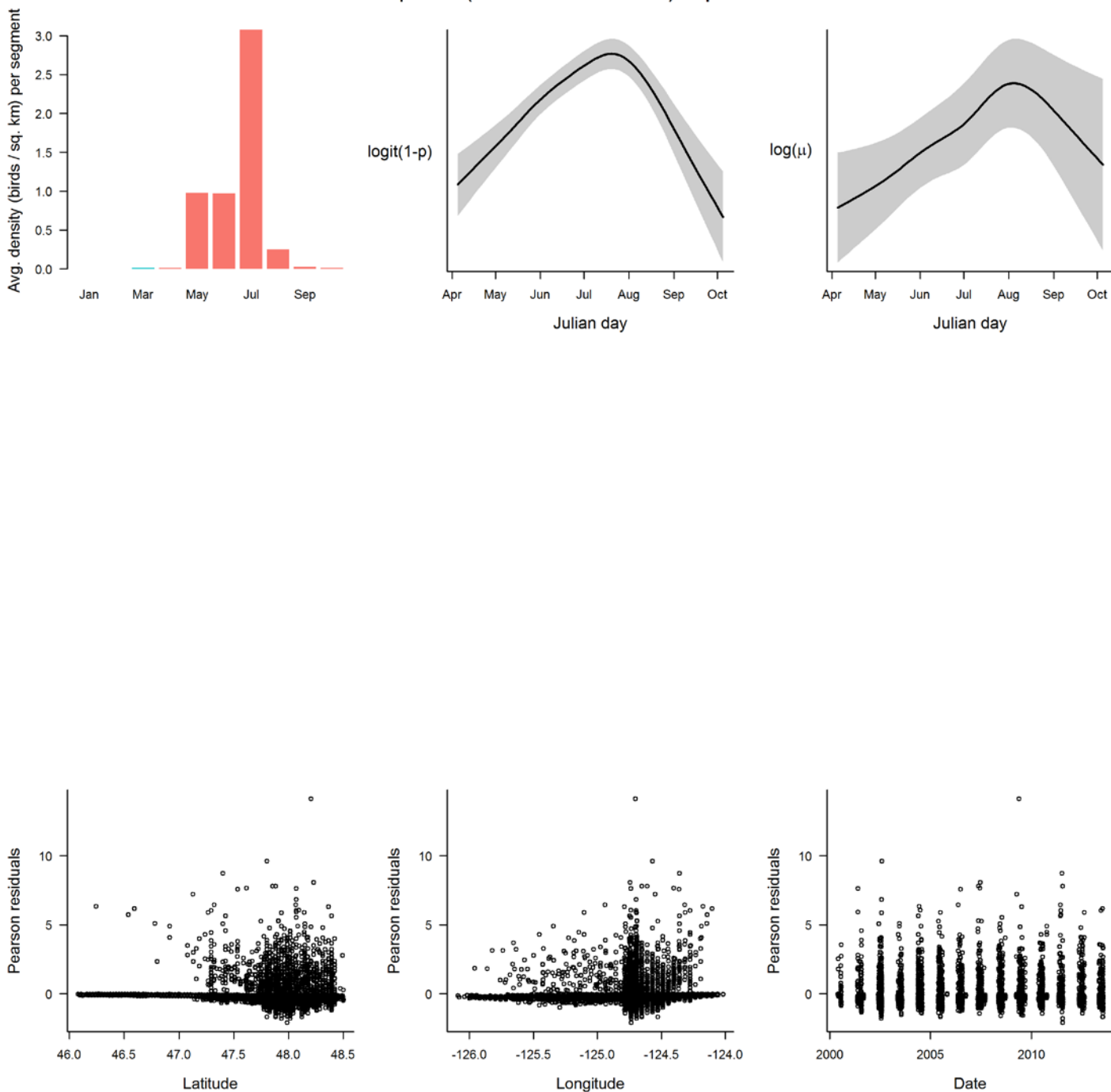


Figure D2: Marginal and residual plots for Tufted Puffin (*Fratercula cirrhata*) during the months of April to October.

Common murre (*Uria aalge*): April to October

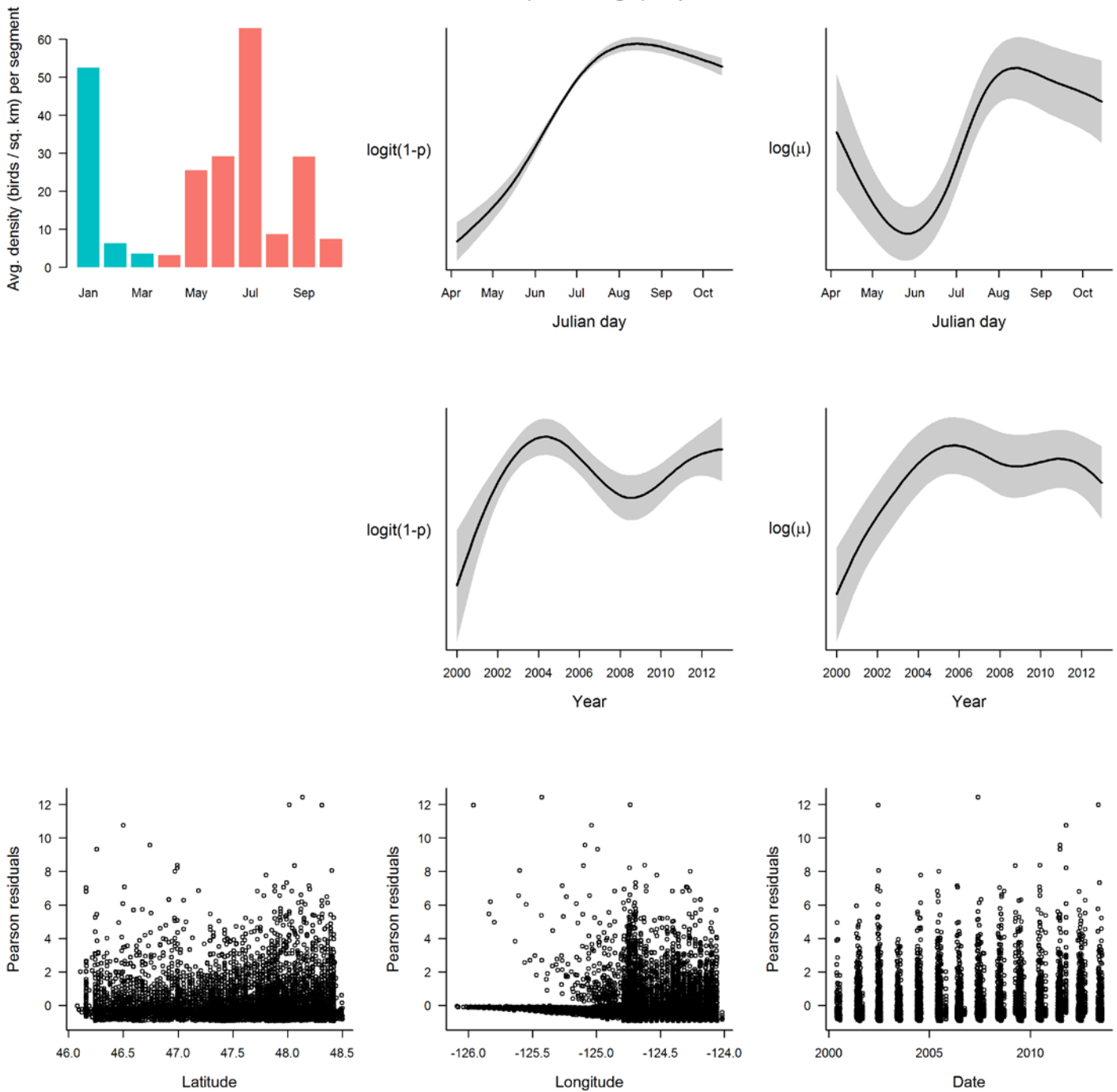


Figure D3: Marginal and residual plots for Common Murre (*Uria aalge*) during the months of April to October.

Common murre (*Uria aalge*): November to March

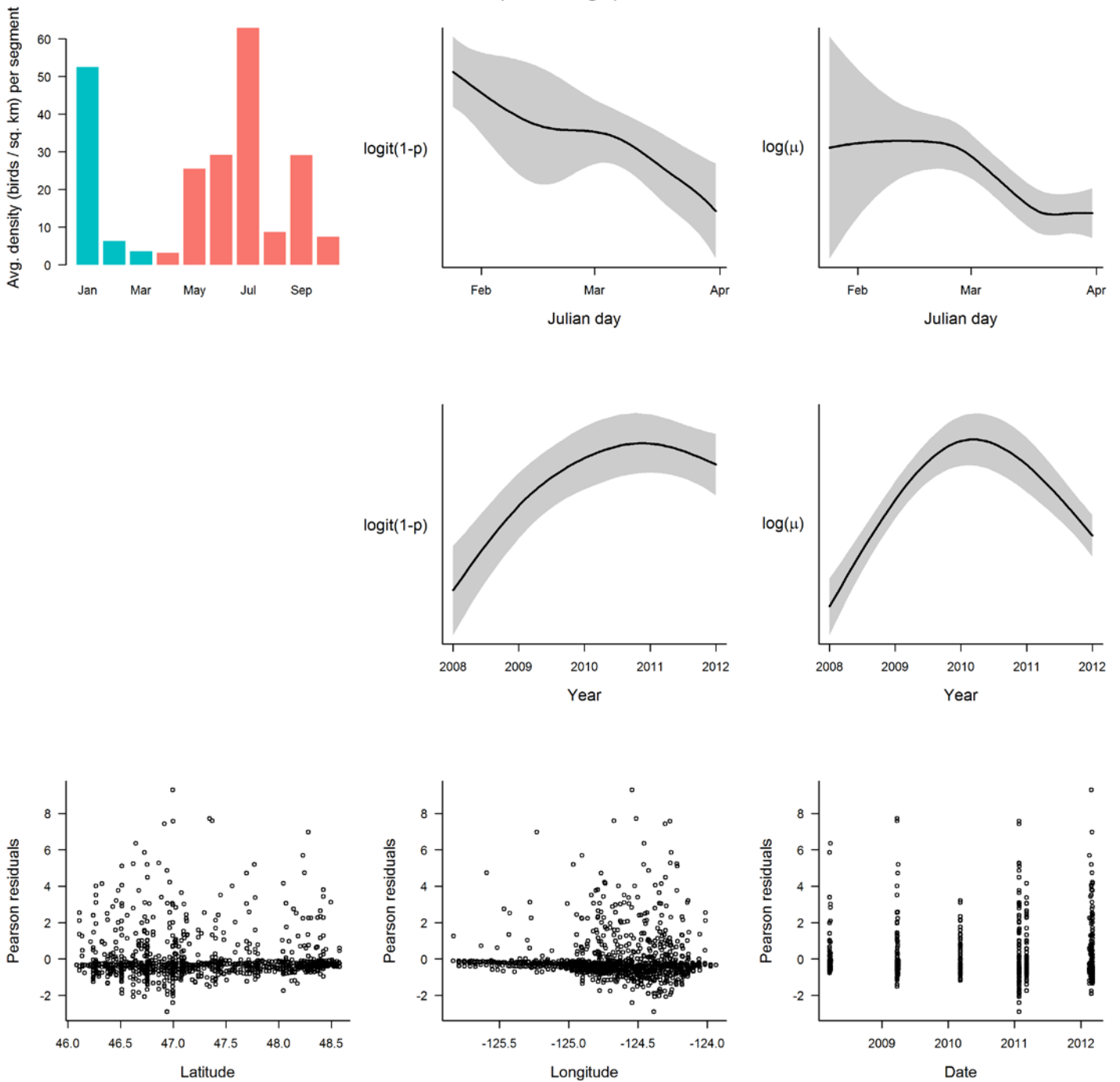


Figure D4: Marginal and residual plots for Common Murre (*Uria aalge*) during the months of November to March.

Black-footed albatross (*Phoebastria nigripes*): April to October

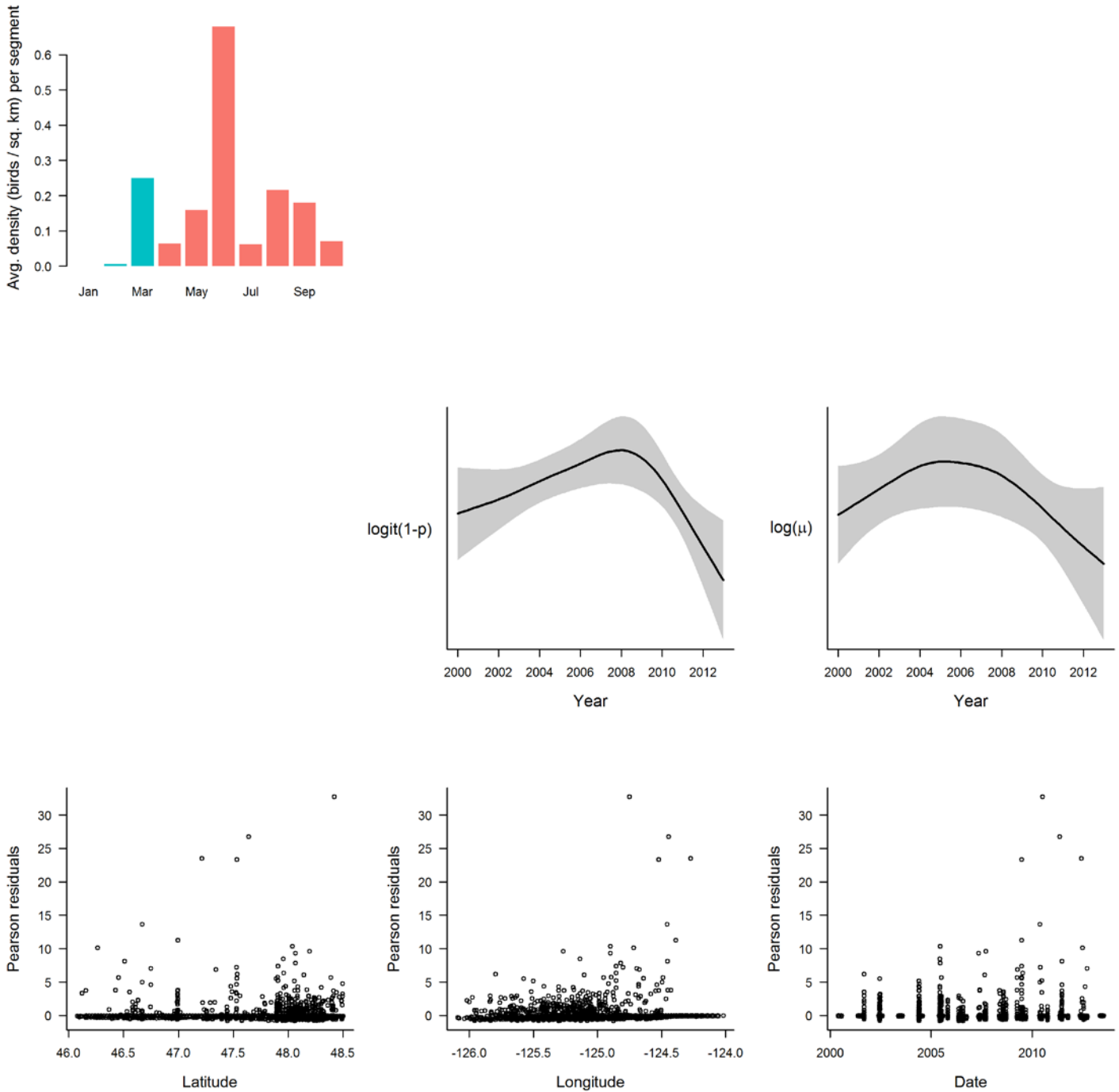


Figure D5: Marginal and residual plots for Black-footed Albatross (*Phoebastria nigripes*) during the months of April to October.

Black-footed albatross (*Phoebastria nigripes*): November to March

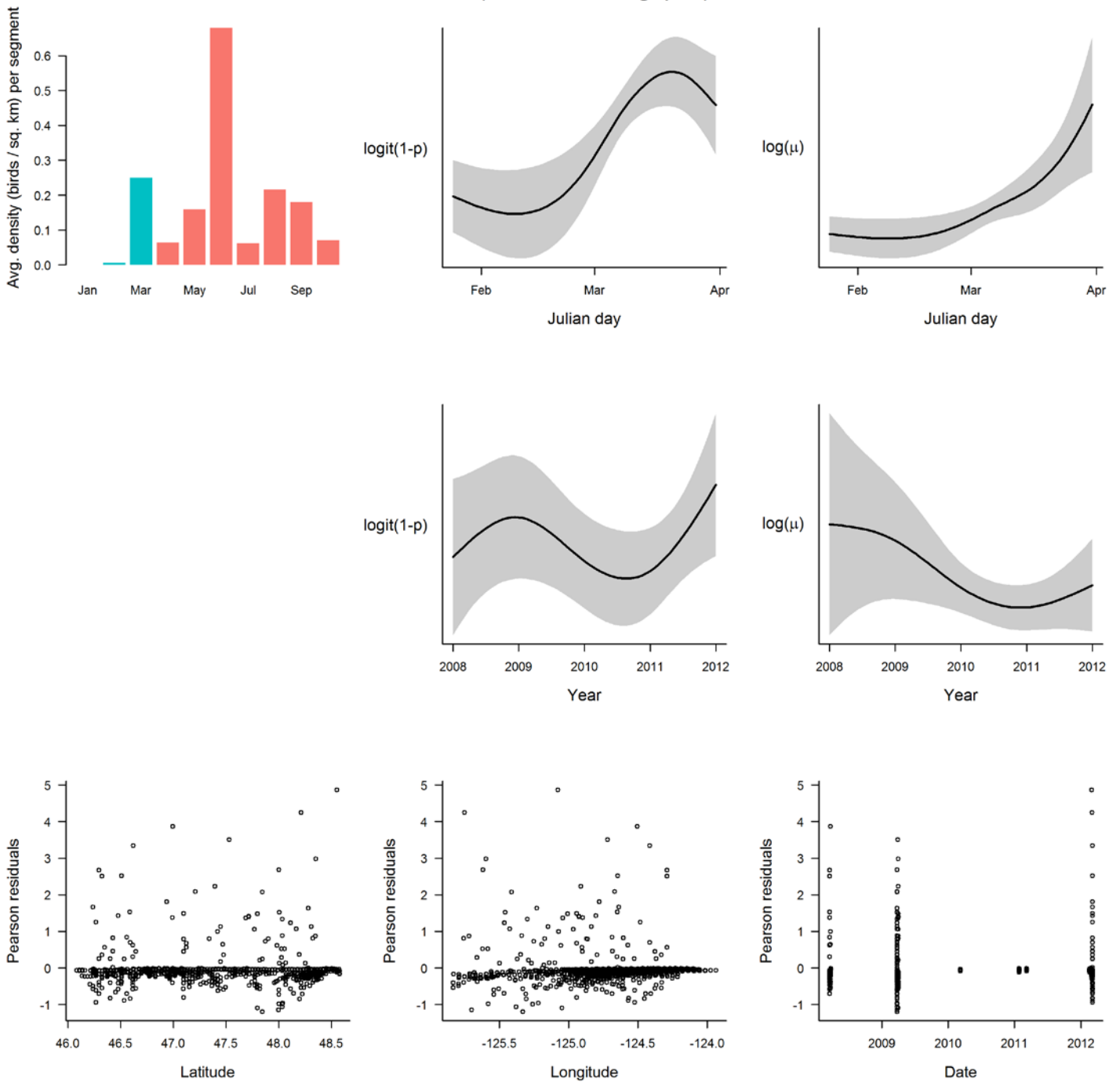


Figure D6: Marginal and residual plots for Black-footed Albatross (*Phoebastria nigripes*) during the months of November to March.

Northern fulmar (*Fulmarus glacialis*): April to October

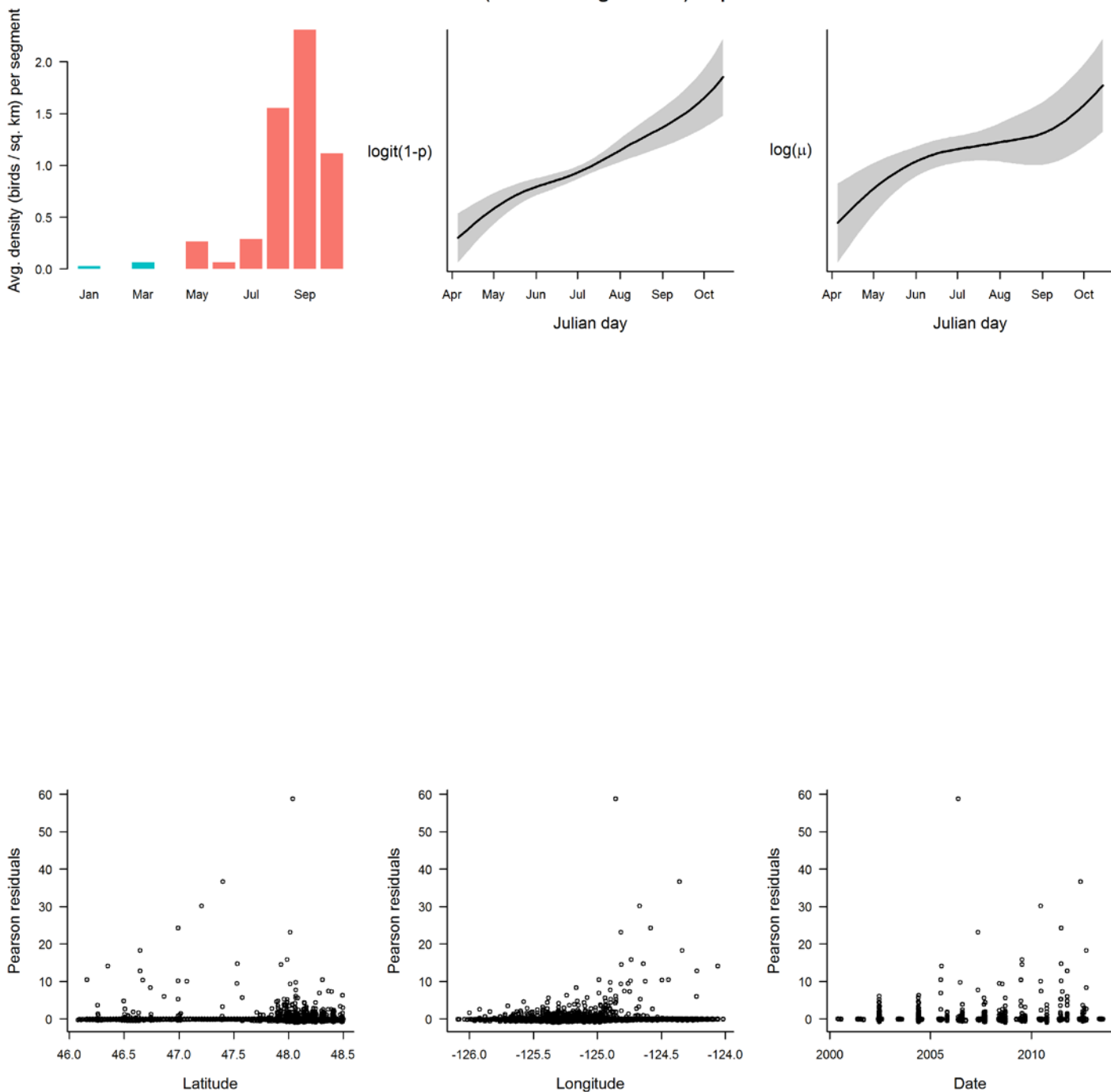


Figure D7: Marginal and residual plots for Northern Fulmar (*Fulmarus glacialis*) during the months of April to October.

Pink-footed shearwater (*Puffinus creatopus*): April to October

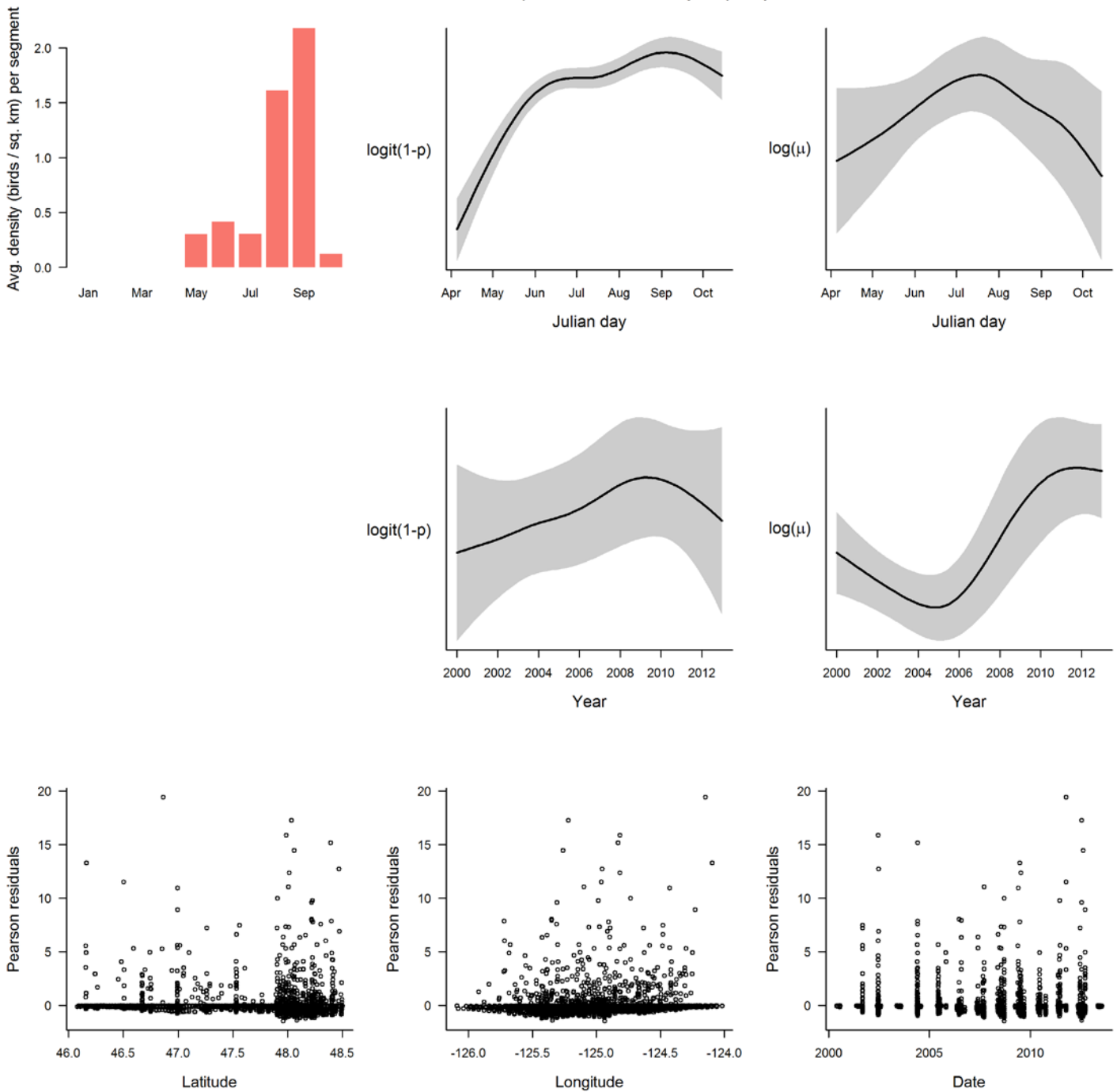


Figure D8: Marginal and residual plots for Pink-footed Shearwater (*Puffinus creatopus*) during the months of April to October.



Sooty shearwater (*Puffinus griseus*): April to October

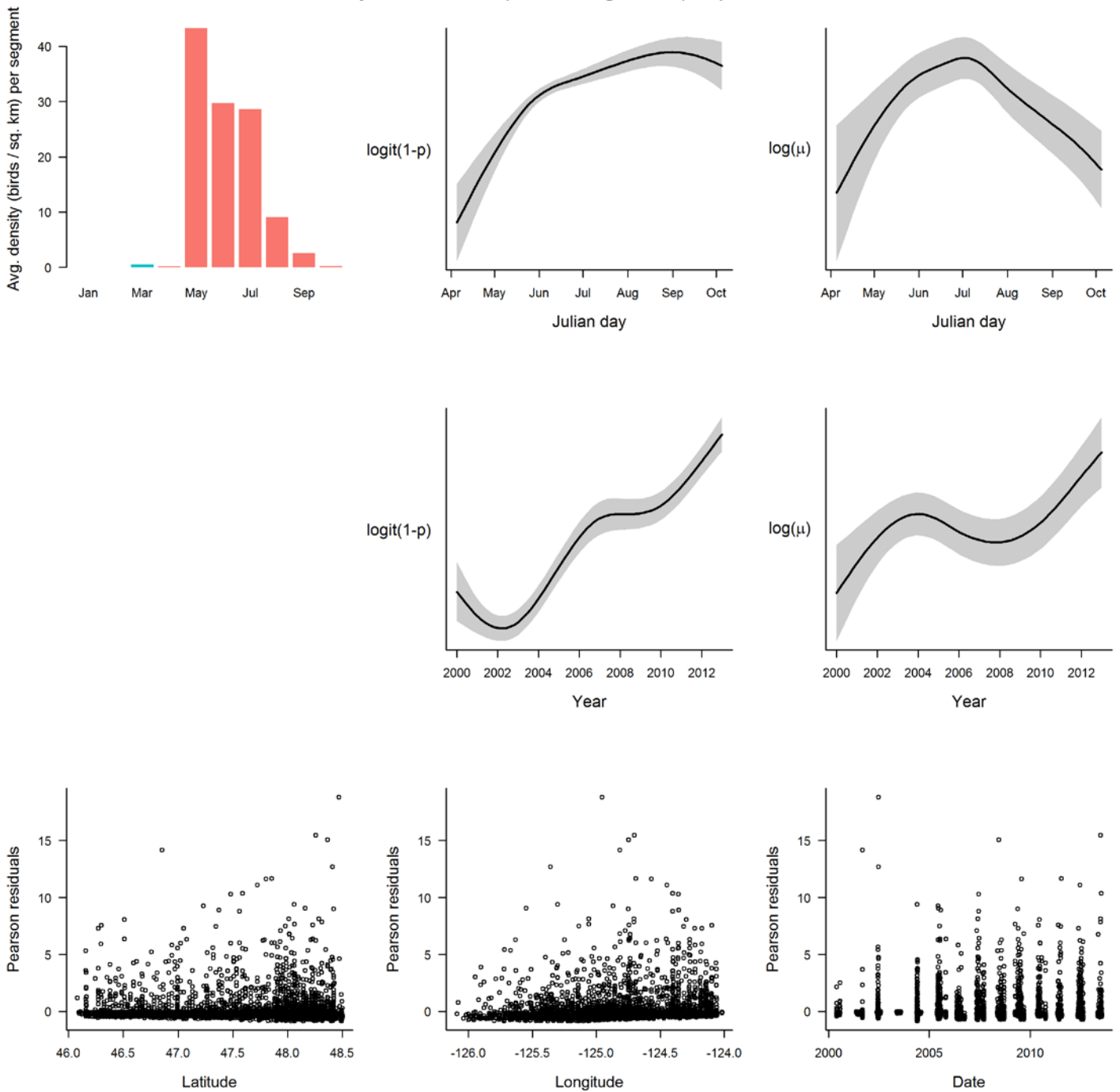


Figure D9: Marginal and residual plots for Sooty Shearwater (*Puffinus griseus*) during the months of April to October.

## Appendix E: Variable importance figures

Figure E1: Mean relative importance of predictor variables for the months of April to October, calculated by averaging across species within the zero-inflation ( $p$ ) and count ( $\mu$ ) components of the selected models.

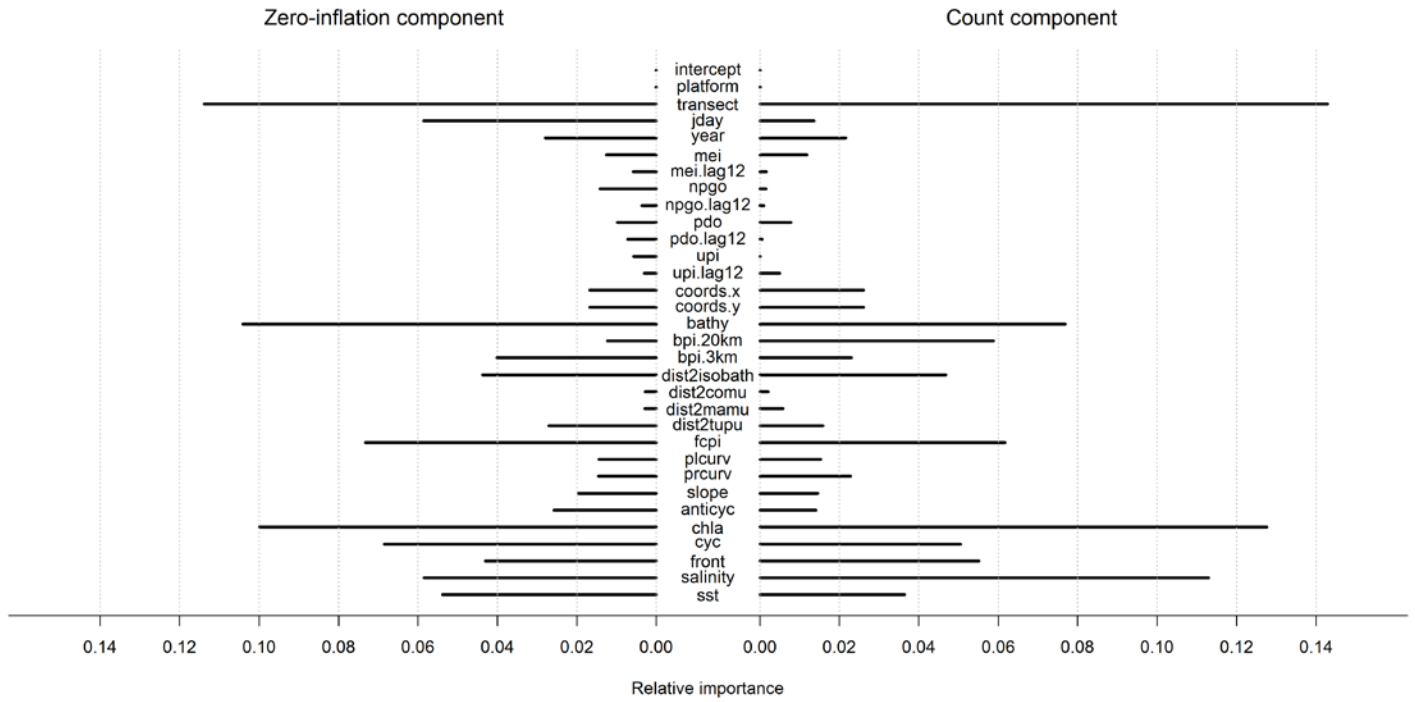
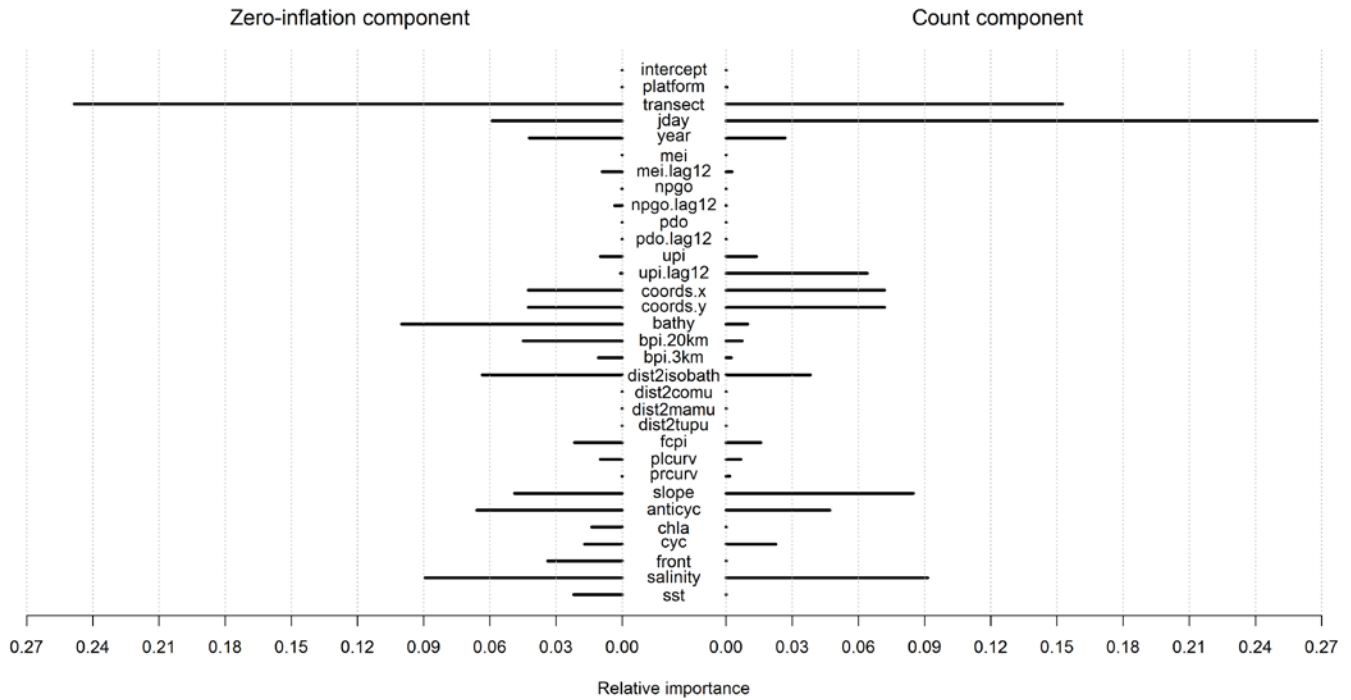


Figure E2: Mean relative importance of predictor variables for the months of November to March, calculated by averaging across species within the zero-inflation ( $p$ ) and count ( $\mu$ ) components of the selected models.



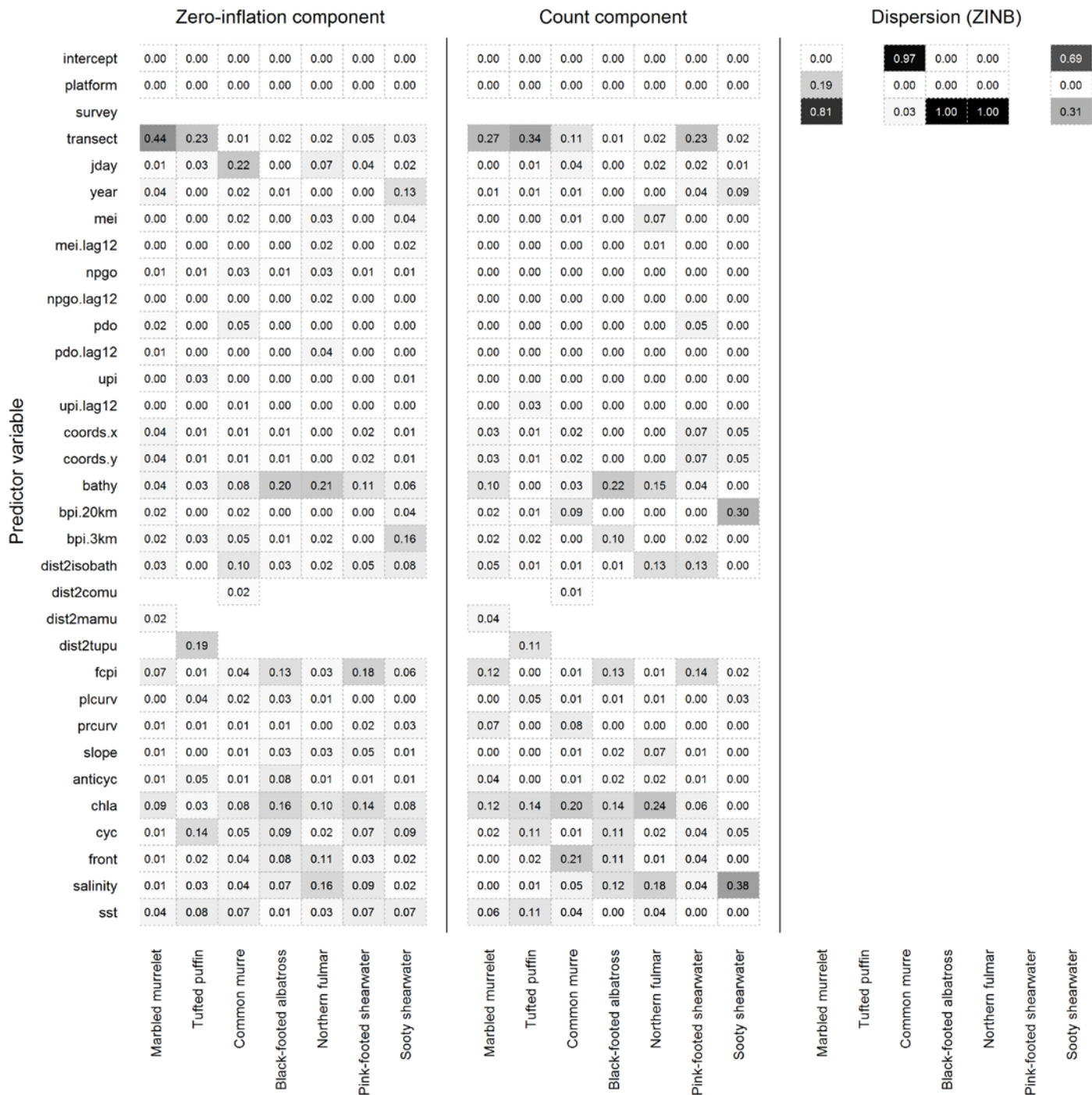


Figure E3: Relative importance of predictor variables for the selected model of each species during the months of April to October. Empty cells represent predictors that were not modeled for a given species. When all cells of the dispersion (ZINB) model component are empty for a given species, the final fitted model assumed a zero-inflated Poisson distribution.

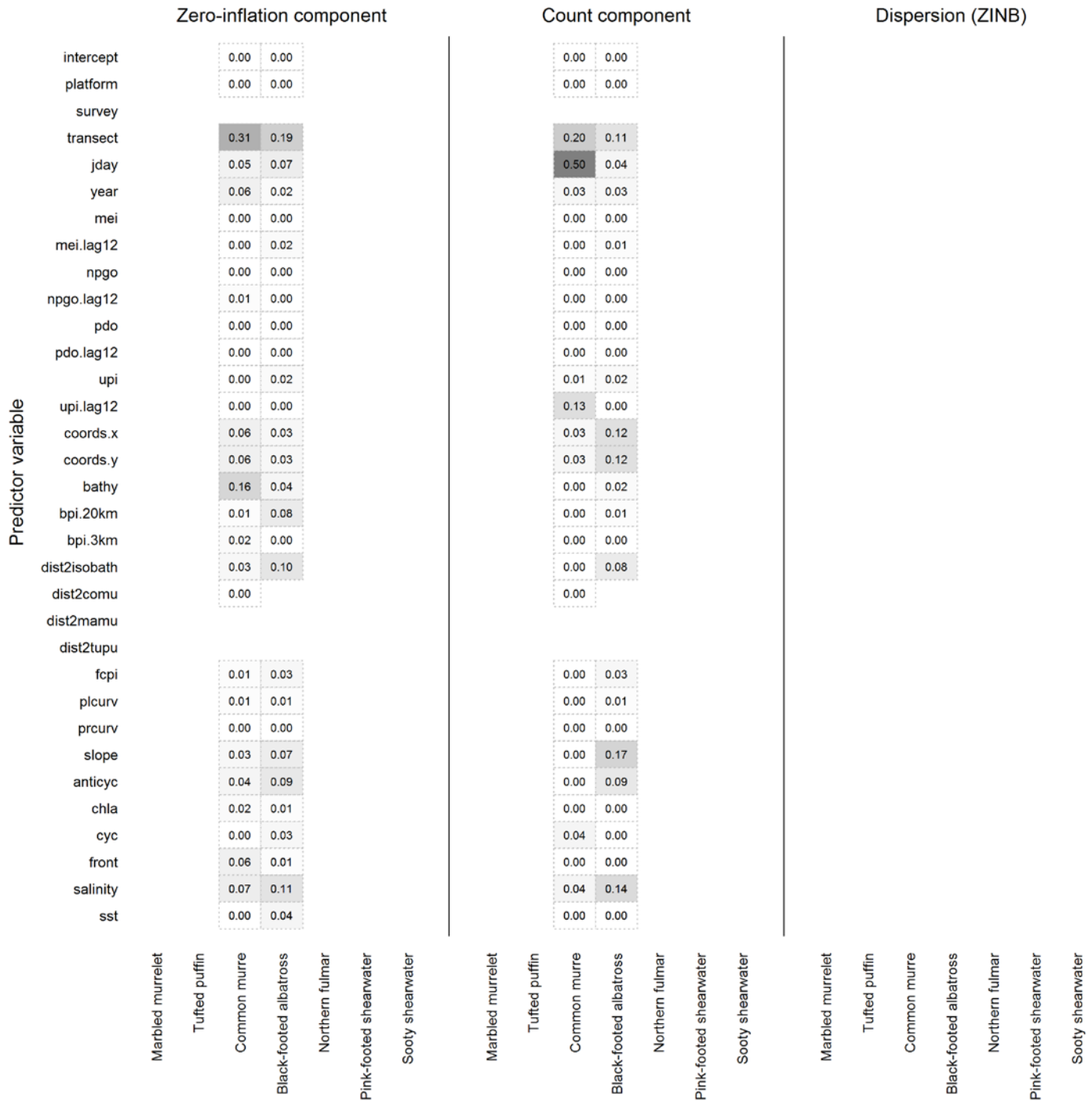


Figure E4: Relative importance of predictor variables for the selected model of each species during the months of November to March. Empty cells represent predictors that were not modeled for a given species. When all cells of the dispersion (ZINB) model component are empty for a given species, the final fitted model assumed a zero-inflated Poisson distribution.