Code: UKCAL-CWF-CON-EIA-RPT-00007-7B30

Volume 7B Proposed Development (Offshore) Appendices

Appendix 6-1 Offshore Ornithology Baseline Characterisation Report Annex 16 MRSea Modelling Report

Caledonia Offshore Wind Farm Ltd

5th Floor Atria One, 144 Morrison Street, Edinburgh, EH38EX

Volume 7B Appendix 6-1 Annex 16 MRSea Modelling Report

Code	UKCAL-CWF-CON-EIA-RPT-00007-7B30
Revision	Issued
Date	18 October 2024

This document contains the following report: 'Caledonia Offshore Windfarm. MRSea modelling of key seabird species' as prepared by Black Bawks Data Science Ltd in February 2024. For the purpose of Consent Application, the document has been retitled to: 'Volume 7B, Appendix 6-1, Annex 16: MRSea Modelling Report', alongside the addition of a new front cover.

OCEAN WINDS

Caledonia Offshore Windfarm

MRSea modelling of key seabird species

Revision History				
Rev.	Description	Prepared	Reviewed	Date
1	Original generated	[redacted]	[redacted]	23/09/2024
2	Draft to client	[redacted] / [redacted]	[redacted]	23/11/2024
3	V3 to client	[redacted] / [redacted]	[redacted]	23/02/2024

Contents

Tables

Figures

List of Abbreviations

ANOVA – Analysis of Variance

BIC – Bayesian Information Criterion

CReSS – Complex Regional Spatial Smoother

CV – Coefficient of Variation

GAM – Generalised Additive Model

GEBCO – General Bathymetric Chart of the Oceans

GIS – Geographic Information System

GLM – Generalised Linear Model

JNCC – Joint Nature Conservation Committee

MRSea – Marine Renewables Strategic Environmental Assessment

OWF – Offshore Wind Farm

QL – Quasi-likelihood

RSE – Relational Standard Error

SALSA – Spatially Adaptive Local Smoothing Algorithm

SST – Sea Surface Temperature

UTM – Universal Transverse Mercator

VIF – Variance Inflation Factor

1 Introduction

As part of the environmental impact assessment framework for the proposed Caledonia offshore wind farm, an assessment of the distribution and abundance of key species was contracted to be undertaken using the Marine Renewables Strategic Environmental Assessment (MRSea Windows package) in R (Scott-Hayward, 2013). The goal of this work was to first quantify the abundance and distribution of key seabird species while considering environmental covariates that could impact their ecology. Further to this, we developed and utilized a spatial apportioning approach for observations unable to be identified to species level which allowed for modelling and thus appropriate accounting for variability in the data. We also applied the random forests algorithm to create predictions of the distribution of birds based on several wind farm scenarios. The use of random forests also allowed for us to make some assessment of displacement based on the partial dependence plots. Finally, using the MRSea outputs, we performed a hot-spot analysis to determine if there were any parts of the proposed development area that had persistent hot-spots for the key species of the analysis.

2 Methods statement

Seabird observation and effort data from 24 digital aerial surveys conducted by APEM between May 2021 – April 2023 were used for the spatial modelling of species' monthly (or yearly, if too few observations) distribution and abundance. Aerial surveys were flown over the proposed Caledonia Offshore Wind Farm (OWF) development, wherein transects included the site and a 4km buffer. Data for eight species were processed: kittiwake (*Rissa tridactyla*), razorbill (*Alca torda*), guillemot (*Cepphus grylle*), puffin (*Fratercula arctica*), gannet (*Morus bassanus*), fulmar (*Fulmarus glacialis*), great black-backed gull (*Larus marinus*), and herring gull (*Larus argentatus*). All species were modeled with data for "all birds" (both flying and sitting behaviours). For kittiwake, gannet, and fulmar, there were sufficient observations to create additional models specifically for birds observed while flying. For all species except auks (puffin, razorbill, guillemot), only detections identified to species level were included in the analysis. For the auks, observations identified to species level were supplemented with counts apportioned using a proportional technique based on species groupings provided by APEM. Proportional no-id apportioning was made at the transect level to ensure a more accurate proportional representation of non-identified species (see section 2.5.1 for more detail).

To model survey-specific bird distribution and abundance, we used the Complex Regional Spatial Smoother (CReSS) spatial modelling method using Spatially Adaptive Local Smoothing Algorithm (SALSA) based model selection (Scott-Hayward, 2013). The models effectively fit the relationship between the observations (count response variable) and the environment (candidate covariates, listed in table 1) at each location which was then be used to estimate the animal density over the area of interest. For each survey, individual counts for each species were assigned to the midpoint of the respective aerial image footprint to produce the input data for each species-specific model. This generated a count variable (the dependent variable) for each footprint, and the footprint area thus became the offset for the model to ensure predicted outputs represented density. Covariate values were then assigned to the midpoint of each segment such that the resulting model input data frame included survey-specific species counts and covariate values for each transect segment.

Count data from aerial transect surveys were correlated to consecutive measurements through space and time. Furthermore, due to environmental and prey conditions, the number of animals in any given area is more likely to be similar for points closer temporally, than those more distant in time. Models fitted to (relative) abundance data attempt to explain animal abundance at any location, but the information (covariate data) that describes why animals are found in high/low numbers at specific locations is frequently missing from the model, leaving patterns in the model's noise component (model residuals). These patterns are also expected to be similar along the track lines. This (positive) correlation in model residuals along the track lines violates a critical assumption for standard statistical models that require an independent set of residuals (such as Generalised Linear Models (GLMs) / Generalised Additive Models (GAMs)). Ignoring this violation can invalidate all model-based precision estimates (e.g., standard errors, confidence intervals, and p-values), resulting in overly complex models that can suggest irrelevant environmental covariates are statistically significant. Transect data are frequently prone to such spatiotemporal autocorrelation, which violates a core assumption of GLMs/GAMs. Thus, transect ID was included as a blocking factor in the analysis to control for autocorrelation in the model. This informed the model that correlation within a transect is accepted and that transect independence is assumed.

2.1 Model Inference

For all models we assumed that the data followed a quasi-poisson distribution, which is typical when working with overdispersed count data (overdispersion occurs when the variance of the data is greater than the mean). This distribution assumes: 1) observations can be counted; 2) the rate at which observations occur can be calculated; 3) No two observations can occur at the same time and/or in the same place; 4) the variance is a linear function of the mean; and 5) that all observations are independent (Ver Hoef and Boveng, 2007). However, count data from aerial transect surveys are likely to be correlated (i.e., dependent) to consecutive measurements through space and time. Furthermore, due to environmental and prey conditions, the number of animals in any given area is more likely to be similar for points closer temporally, than those more distant in time. Models fitted to (relative) abundance data attempt to explain animal abundance at any location, but the variables that determine why animals are found in high/low numbers at specific locations are frequently missing from the model, leaving patterns in the model's noise component (model residuals). These patterns are also expected to be similar along the track lines. This, temporal and spatial autocorrelation in model residuals along the track lines violates a critical assumption for standard statistical models which require an independent set of residuals (such as Generalised Linear Models (GLMs) / Generalised Additive Models (GAMs)). Ignoring this violation can invalidate all model-based precision estimates (e.g., standard errors, confidence intervals, and p-values), resulting in overly complex models that can result in a Type 1 statistical error whereby irrelevant environmental covariates are statistically significant. Transect data are frequently prone to such spatiotemporal autocorrelation, which violates a core assumption of GLMs/GAMs. To address this issue, transect ID was included as a blocking factor in the analysis to control for autocorrelation in the model. Including transect ID mitigates autocorrelation issues by including it as a factorial variable that describes the distribution of the individuals, accounting for the correlation within the transect. This is standard practice when accounting for autocorrelation in transect data. A one-way Analysis of Variance (ANOVA) was performed to determine the statistical significance of covariates in the predictive model. Partial dependence plots were used to investigate covariates that have significant relationships with the data in the model. Further model inference was made by examining the cumulative residual plots output by the models.

2.2 Selection of Model Covariates

First, a full model with all appropriate terms (e.g., as identified from Table 1) was fit for each species without a smooth term for the spatial component. This allowed the potential relationships between covariates and species observations to be initially unhindered by spatial information. Variance Inflation Factors (VIF) were then used to select covariate terms from the initial model fitting process that should be removed due to collinearity based on a VIF threshold value of 2. There is not a standardised approach to choosing this value, but in general, lower threshold values constitute a more conservative approach to eliminating issues with multicollinearity. The flexibility of the smoother-related term for each model term was then chosen, followed by the model selection for the two-dimensional smoother term for the spatial component. Segment area was then incorporated into the model as an offset term as the transects' division may have resulted in slightly different dimensions. Each model was permitted to retain the covariates as a smooth or linear term (or omitted completely). SALSA was used to fit a smooth function for each covariate. Model selection for both the covariates and spatially based smoothers was conducted by using an objective fit measure (Bayesian Information Criterion (BIC) for quasi-likelihood (QL) models). Models that permit over-dispersion for Poisson-style counts are QL based, necessitating QL-based fit scores.

Table 1. Candidate covariates for model selection in MRSea analysis

Model Covariate	Definition	Source
Survey ID	Unique ID for each survey	APEM Aerial Surveying
Bathymetry	Depth below sea surface (m)	GEBCO Gridded Bathymetry Data 2019
Bathymetric Slope	Change in bathymetry between pixels	GEBCO Gridded Bathymetry Data 2019
Bathymetric aspect	Direction bathymetric slope faces	GEBCO Gridded Bathymetry Data 2019
SeaSurface Temperature (SST)	Interpolated sea surface temperature on daily 0.02 degree grid	Copernicus (European North West Shelf/Iberia Biscay Irish Seas – High Resolution ODYSSEA L4 Sea Surface Temperature Analysis)
SST gradient	Change in SST between pixels/ slope of SST	Copernicus (European North West Shelf/Iberia Biscay Irish Seas – High Resolution ODYSSEA L4 Sea Surface Temperature Analysis
Sandeel predicted density	Probability of presence of buried sandeel in the North Sea study region.	Marine Scotland (Langton et al., 2021)
Sandeel probability of presence	Predicted density of buried sandeel in the North Sea study region (number per m^2)	Marine Scotland, (Langton et al., 2021)
Distance to coast	Distance to coast (m)	GIS (UTMs)
Distance to colony	Distance to nearest colony (m)	JNCC Seabird Monitoring database
Segment area	Area of each segment within a transect (m^2)	APEM Aerial Surveying
Spatial component	Northing and Easting	GIS (UTMs)
Distance to wind turbine	Distance to the nearest operational turbine (m)	Moray East Wind Farm, Beatrice Wind Farm

2.3 Knot Placement and Basis Function Details

The number of "knots" used for the model and the effective range of each knot (the spatial extent to which each knot influences the fitted surface) are both key factors in determining the model flexibility for the spatial surfaces in this setting. The candidate models were chosen from a range of models that varied in the number of knots provided and the effective range (R-value) of each knot because the optimal choices for both values are always unknown. The starting knot positions on the spatial surface were chosen to maximize coverage across the spatial area (via a space filling algorithm; (John et al., 1995)), and these positions were allowed to move according to the SALSA (Walker et al., 2011) model selection technique. We used the local exponential basis function, defined as:

$$((exp(-d/r2))$$

where d is the Euclidian distance, allowing for varied R-values over the surface. A variable number of knots (2-40 depending on data sparsity; the number is denoted by the degrees of freedom in the model) were used for the candidate models, and an objective fit criterion was employed to select the best model(s). In effect, the position of the knot placement, and to a lesser extent the number of knots, reflect the complexities of the spatial relationship between bird abundance and the covariates chosen for the study. Knot locations were identified separately for each survey to accommodate differences in survey effort and bird distributions across surveys.

2.4 Geo-Referenced Results

The species-specific fitted surfaces were generated by making predictions within a grid using the final model at a 1km x 1km resolution. The grid is a series of regular points spaced at 1km resolution across the surface of the area of interest. These regular points are associated with the same environmental covariates as those used in the modelling process. This allows the trained MRSea model to make predictions of animal density on each of those points. Those data can then be visualized or interpolated to create surfaces. These grids were projected as the Universal Transverse Mercator (UTM; Zone 30) projection. To measure uncertainty spatially and in the population estimates, the model was bootstrapped 1,000 times (wherein random subsets of the modelled coefficients are drawn from a multivariate normal distribution and predictions are made for each grid cell, 1,000 times). From this we calculated the mean predicted density, the upper and lower 95% confidence limits, and the coefficient of variation (CV; as defined by the ratio of the standard deviation to the mean). These measures of uncertainty were visualized and are presented in the results section.

2.5 Abundance Estimates from MRSea Density Surfaces

Abundance estimates were calculated by summing the grid cells across the prediction surface at the temporal scale specified by the exploratory analysis. To calculate abundance estimates within the survey area, we summed grid cells that fall within the boundary or touch the edges, ensuring that grid cells at the boundary are clipped to the boundary footprint and adjusted for the new area. The upper and lower confidence limits of the population estimate were calculated by determining the 95% confidence limits of the 1000 bootstrapped surfaces. The bootstrap outputs were not normally distributed, and the mean was more sensitive to outliers therefore the median was chosen as a more representative measure of central tendency when presenting abundance estimates. We note that in previous work in the offshore wind industry in the UK, the term “coefficient of variation” is often applied to population estimates derived from bootstrap exercises. However, this is a misnomer because the standard deviation of a bootstrapped mean is the standard error. The calculation of the ratio of the standard error to the mean is called the relational standard error (RSE), and not the CV as has been previously used. For the sake of accuracy in the vocabulary used and to distinguish this measure of uncertainty from the spatial CV, we present the uncertainty as the RSE, but note that this is equivalent to what has been referred to as the CV in past work.

2.5.1 Species Apportioning

To account for auks that have not been identified to species level, we applied a spatial proportional apportioning technique based on the species groupings defined by the digital aerial survey provider. For example, if there is a species grouping called “large auks” which is made up of possible guillemots or razorbill, then the proportional species composition will be applied within spatial bins that will be used for modelling. In other words, if 80 guillemot and 20 razorbill were identified in a survey, and a spatial bin/cluster had 10 unidentified large auks, then 8 guillemot and 2 razorbills would be apportioned into that bin. Spatial proportional no-id apportioning was first attempted at a transect level (which ensures a more accurate proportional representation of non-identified species). In cases when that was unable to be performed, then it was performed at the survey level. In the case that a proportional approach leads to fractional values (for example, if the split was 80/20 for guillemot to razorbill, and there were 7 unidentified large auks in a cluster, this would equate to 5.6 guillemot and 1.4 razorbill. This would be rounded to 6 guillemot and 1 razorbill). The new observations apportioned from the non-identified birds were then used for MRSea modelling.

2.6 Hotspot Analysis

A hotspot analysis was performed upon completion of all species models. The analysis was informed by the “all birds” survey-level MRSea models. Model outputs from all species and survey-level outputs were normalized on a scale of 0 –1 and then averaged. The cell-by-cell coefficient of variation was calculated by dividing the standard deviation by the mean. This provided a single magnitude (I.e., mean normalized prediction layer) and persistence (I.e., variability as defined by the CV) layer. The upper and lower 95th percentiles of all values in the magnitude and persistence layers were computed to use as thresholds for determining if a grid cell is persistent hot or cold spot. This is categorized as per table 2.

Table 2. Classifications of hot spots as defined by percentiles from magnitude and persistence layers

Persistence	Magnitude	Classification
CV > 95th percentile	Mean > 95th percentile	Persistent hot spot
CV > 95th percentile	Mean < 5th percentile	Persistent cold spot
CV < 5th percentile	Mean > 95th percentile	Volatile hot spot
CV < 5th percentile	Mean < 5th percentile	Volatile cold spot
CV > 5th percentile, < 95th percentile	Mean > 95th percentile	Transient hot spot
CV > 5th percentile, < 95th percentile	Mean < 5th percentile	Transient cold spot

CV > 5th percentile, < 95th percentile	Mean > 5th percentile, < 95th percentile	Transient ambient spot
CV > 95th percentile	Mean > 5th percentile, < 95th percentile	Persistent ambient spot
CV < 5th percentile	> 5th percentile, < 95th percentile	Volatile ambient spot

2.7 Assessment of Displacement from MRSea

We made use of the precise Moray East offshore wind farm operational turbine locations and timing of installation to build a series of "distance to turbine" predictor layers that were used in the modelling process. We constructed monthly "distance to turbine" layers that were representative of the active turbines in the water at the time of each survey (akin to monthly SST layers). We then examined the partial dependence plots of the distance to turbine layers to make assessments of displacement because inflection points in those figures could indicate if displacement is occurring, and at what spatial scales.

2.8 Random forests

An additional analysis was conducted which performed spatial displacement scenarios based on proposed turbine locations. The distance to turbine layers from the operational Moray East offshore wind farm were used to train a spatial model that employs a machine learning algorithm called random forests. The random forest algorithm is a supervised machine learning algorithm used widely for regression and classification problems in machine learning (Breiman, 2001). A random forest is a classifier that includes many decision trees on various subsets of a given dataset. The classifier takes the average decision of that subset to improve the predictive performance. It is built on the idea of ensemble learning, where multiple classifiers are integrated to solve complex problems and improve model performance. In the same way that a forest with many trees is more robust, a random forest algorithm with more decision trees will have greater accuracy and higher predictive ability.

Random Forests for regression represent a powerful ensemble learning approach, particularly adept at handling complex relationships in data (Hastie et al., 2008). The foundation of Random Forests lies in the construction of decision trees. In the context of regression, each tree is essentially a sequence of binary decisions or splits that recursively partition the input space. What sets Random Forests apart is the introduction of randomness in the process. During the creation of each tree, a random subset of the training data, known as a bootstrap sample, is drawn with replacement. Moreover, at each node of the tree, only a random subset of predictors is considered for determining the best split. This deliberate injection of variability helps prevent individual trees from fitting noise in the data and encourages diversity among the constituent trees.

The aggregation of these independently grown and varied trees defines the strength of Random Forests. For regression, the final prediction is an average (or weighted average) of the predictions from all the trees, providing a robust and accurate estimate of the target variable. Importantly, Random Forests are equipped to handle both numerical and categorical predictors, making them versatile across a range of real-world datasets. The algorithm's adaptability, resilience to overfitting, and ability to capture non-linear relationships make it particularly well-suited for regression tasks where the underlying data structure is complex and multifaceted.

The calculation of regression trees within the Random Forest framework involves recursive binary splits based on predictor variables. At each node, the algorithm identifies the predictor and split point that minimizes the sum of squared differences between the observed and predicted values. The splitting process continues until a predefined stopping criterion is met, such as a minimum node size or a maximum tree depth. This recursive partitioning results in a tree structure where each terminal node, or leaf, represents a distinct subset of the data. During prediction, an input observation traverses the tree, and its target value is determined by the average of the training observations in the terminal node to which it belongs. This process is repeated for each tree in the ensemble, and the final prediction for regression is the aggregated result of all individual tree predictions. The collective strength of Random Forests lies in their ability to capture intricate relationships in the data while minimizing the risk of overfitting (Breiman 2001; Cutler et al., 2012; Hastie et al., 2008). Furthermore, its non-parametric nature means that many of the typical frequentist statistical assumptions do not have to be adhered to, thus increasing the likelihood of generating statistically appropriate estimations.

In a spatial context, this is applied in an almost identical fashion to MRSea, where observations or counts are associated with environmental covariates which allows for a model to be trained. This trained model is then applied to a grid of regular points to generate predictions in space. We used the same environmental covariates as MRSea, however, the spatial component was replaced by spatial autocorrelation terms, which helped capture the nature of flocking behaviour by marine birds. To generate the predicted scenarios of displacement, Ocean Winds was consulted to create three plausible wind turbine configurations, and each of those scenarios was used to create distance to wind turbine layers, which was used to predict displacement for Guillemot, Razorbill, Kittiwake, Fulmar, Puffin and Gannet. This algorithm was used as opposed to MRSea because MRSea requires the spatial knots as calculated using the raw observations which have been captured prior to any turbines being installed. The aim of this process was to create predictive scenarios of turbines being installed in various spatial configurations without being constrained by the locations of birds as they existed prior to the installation.

2.9 Displacement simulation

To assess the influence of the distance to turbine predictor layer in the model, we built a random forests model using the caret (Kuhn 2022) and ranger (Wright and Ziegler 2017) packages in R version 4.3.1. Using a grid-based tuning approach, we tested values of mtry (i.e., the number of randomly selected variables at each split), from 3 to 10, and the number of minimum observations in each terminal node between 80 and 100. As a metric for model selection, we used root mean squared error (RMSE) with five-fold cross validation. The model with the lowest RMSE was selected as the best model.

To assess the effect of turbines on the distribution of key species, the partial dependence plots of the distance to turbine covariate were generated and examined. A partial dependence plot is a graphical method used in machine learning to visualize the marginal effect of a feature on the predicted outcome of a model while marginalizing over the values of all other features. It helps to understand the relationship between a feature and the target variable in isolation, holding other features constant or averaging over their values. However, in terms of understanding displacement, we note that we are incorporating observations of birds throughout the entire survey area (i.e., outside of the likely distance of immediate impact of turbines on distribution as they would be outside of visual range). Furthermore, relationships noted (e.g., higher densities of birds predicted further away from turbines) could simply be due to other factors (e.g., presence of more favourable foraging conditions at a site that happens to fall at some distance away from the turbine blades), Thus these interpretations must be taken cautiously. Visual acuity varies between species depending on their foraging ecology, but based on data collected for Northern fulmar, the smallest low-contrast object that can be seen near sea-level would have to be 13m in diameter at 11 km distance (Mitkus et al., 2016). Thus, a conservative estimate of 20 km was applied as the maximum distance along the x-axis for interpretation of the partial dependence plots.

Next, we used plausible scenarios of turbine distributions to generate spatial predictions, under the assumption that the distribution of birds in the region was representative of the post-construction phase. For the purposes of this analysis, we use all flying and sitting birds combined under the assumption that the impact will be the same for both behaviour types.

Four scenarios were modeled: Scenario 1; where no new turbines were constructed - this represents the baseline scenario as per the digital aerial surveys. Scenario 2; where turbines were constructed only in the northern part of the proposed wind farm. Scenario 3; where turbines were constructed only in the southern part of the proposed wind farm. And Scenario 4; a combination of scenarios 2 and 3 (Figure 1).

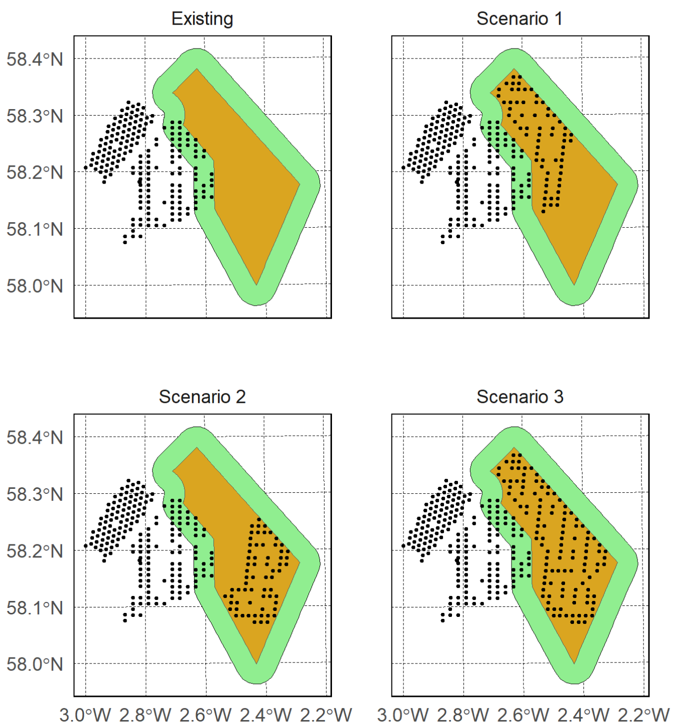


Figure 1: Turbine scenarios for examining distributional responses post-construction

Grid cells in the predicted outputs were adjusted for unequal areas around the borders of the survey region and predicted counts were summed to generate population estimates for each survey and turbine scenario.

2.10 Interpretation of Random Forests outputs

Although random forests offers a number of advantages as a predictive algorithm, interpretation of the outputs must be done carefully, particularly when attempting to understand mechanistic relationships. Firstly, Random forests are comprised of multiple decision trees, where each tree is trained on a subset of the data and with a subset of features. This ensemble approach often leads to improved performance compared to individual decision trees. However, interpreting the collective decision-making process of many trees can be complex, especially when there are many trees in the forest. Also, the building blocks of random forests are capable of learning complex non-linear relationships in the data. As a result, random forests can model intricate decision boundaries that may not be easily understandable or interpretable, especially in high-dimensional spaces. Further to this, even though decision trees are simpler models compared to some other machine learning algorithms, individual trees in a random forest can still be quite complex, especially if the dataset contains many features or if the trees are allowed to grow deep. Interpreting the decisions made by these trees can be challenging, particularly when trying to understand how they collectively contribute to the final prediction.

3 Results

Models were generated for each species using observations of that species for all behaviours, these will be referred to as the "all-birds" models. Additional models were generated for kittiwake, gannet, fulmar, and great black-backed gulls using only observations of flying birds, these models will be referred to as "flying-birds" models. Mean density surfaces for each survey from MRSea outputs mapped to the Caledonia OWF site are provided in figures 2 - 17 (all kittiwake), figures 21-24 (flying kittiwake), figures 40-41 (all gannet), figures 51-52 (flying gannet), figures 62-66 (all fulmar), figures 85-89 (flying fulmar), figures 108-112 (guillemot), figures 130-133 (razorbill), figures 148-150 (puffin), figure 162 (all great black-backed gull), figure 169 (flying great black-backed gull), figure 176 (herring gull).

3.1 Kittiwake

3.1.1 All Birds Model

Table 3. Candidate and final covariates for all kittiwakes model.

Starting model covariates after VIF-based collinearity removal	Final model covariates after removal by SALSA
Survey ID	Sandeel Density
Bathymetry	Distance to colony
Sandeel Density	
Distance to colony	
Distance to turbine	
Standard deviation of Sea Surface Temperature	

Distribution maps generated using MRSea (Figure 2-17) suggest that kittiwake are widely distributed throughout the area during the breeding season (15 April – 31 August), with higher concentrations in the south. In the non-breeding season, distribution declines, but is again generally higher in the south. The highest densities were observed in the southern third of the study area during July 2021.

Model fit was poor with a marginal R squared value of 0.04 and root mean squared error of 13.13. Cumulative residuals in the model showed that there was a moderate relationship between predicted and observed values across the lower range of predicted values, though the model tends to under-predict for values between ~1.8 and ~5 birds/ km^2 (bottom row, Figure 19).

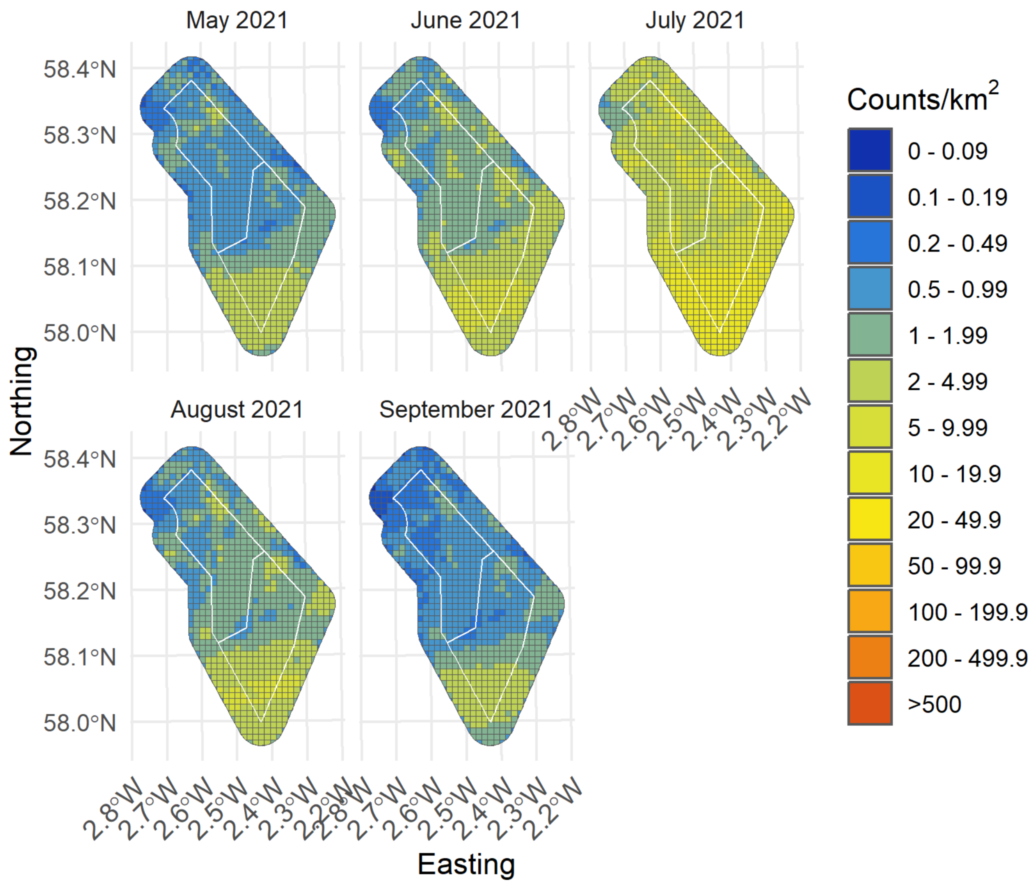


Figure 2. Median density of all kittiwake in the survey area for months with sufficient observations between May and September 2021

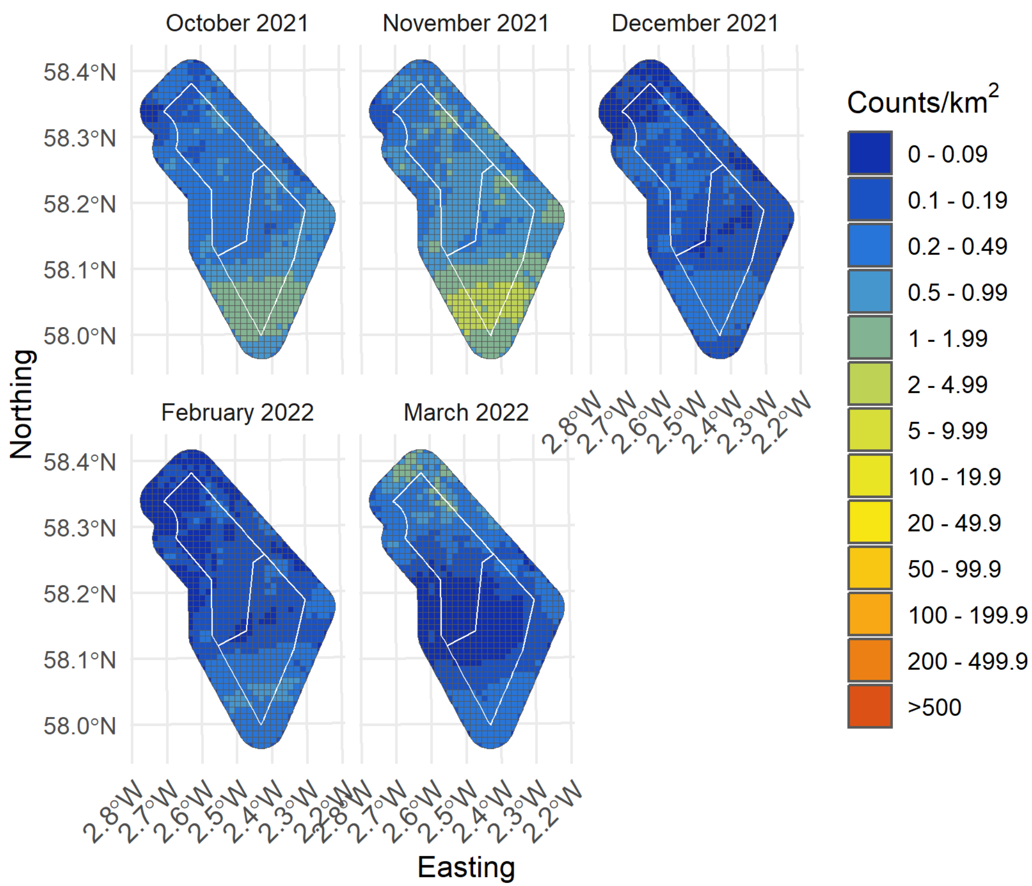


Figure 3. Median density of all kittiwake in the survey area for months with sufficient observations between October 2021 and March 2022

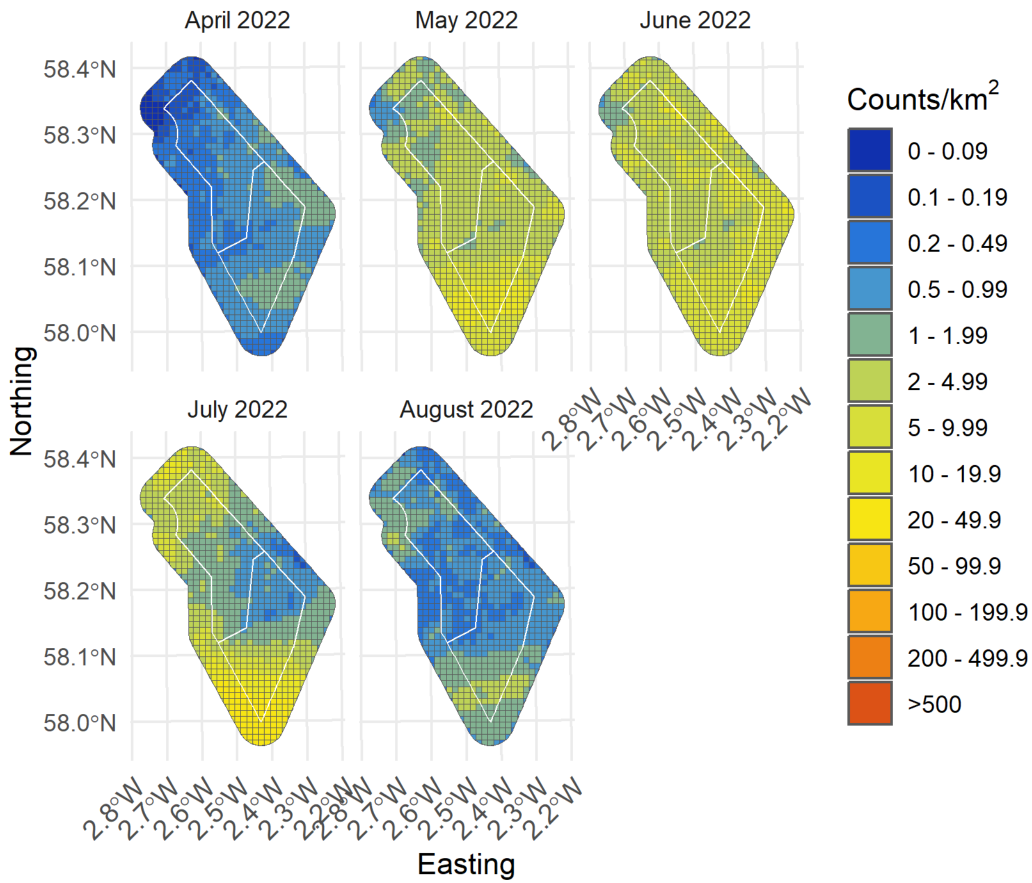


Figure 4. Median density of all kittiwake in the survey area for months with sufficient observations between April and August 2022

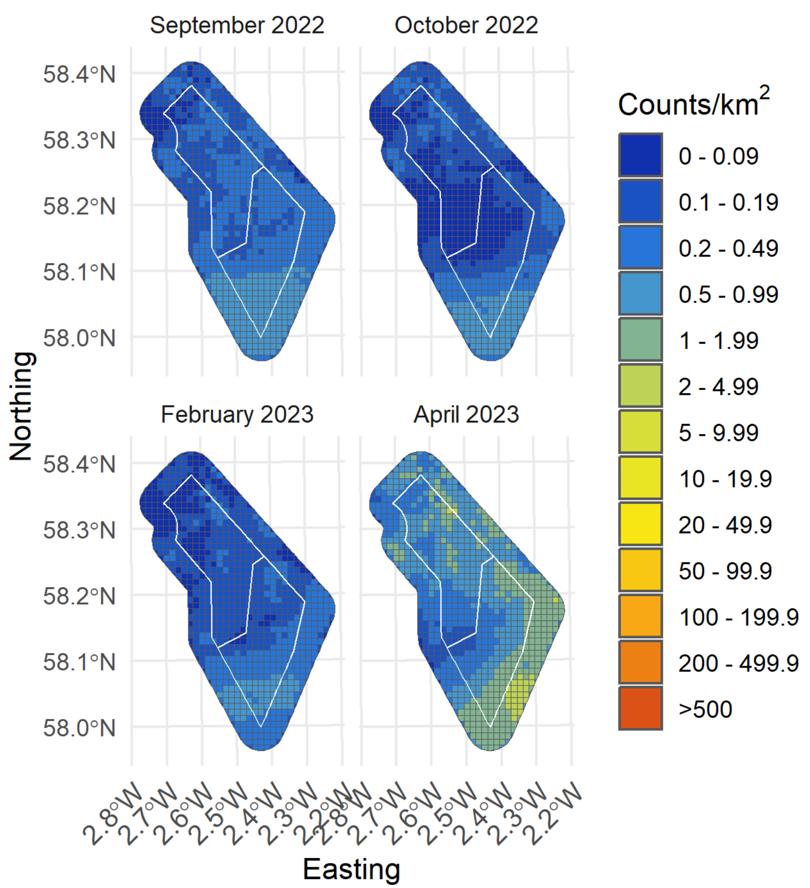


Figure 5. Median density of all kittiwake in the survey area for months with sufficient observations between September 2022 and April 2023

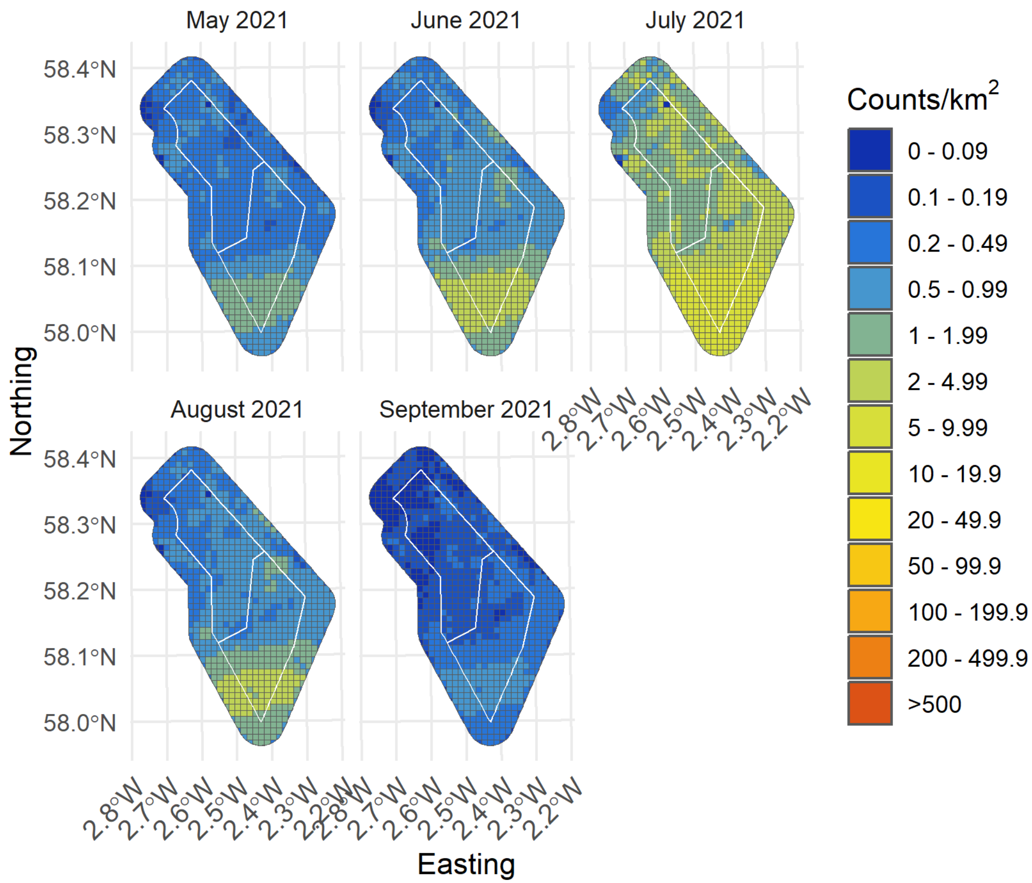


Figure 6. Lower confidence limit of density of all kittiwake in the survey area for months with sufficient observations between May and September 2021

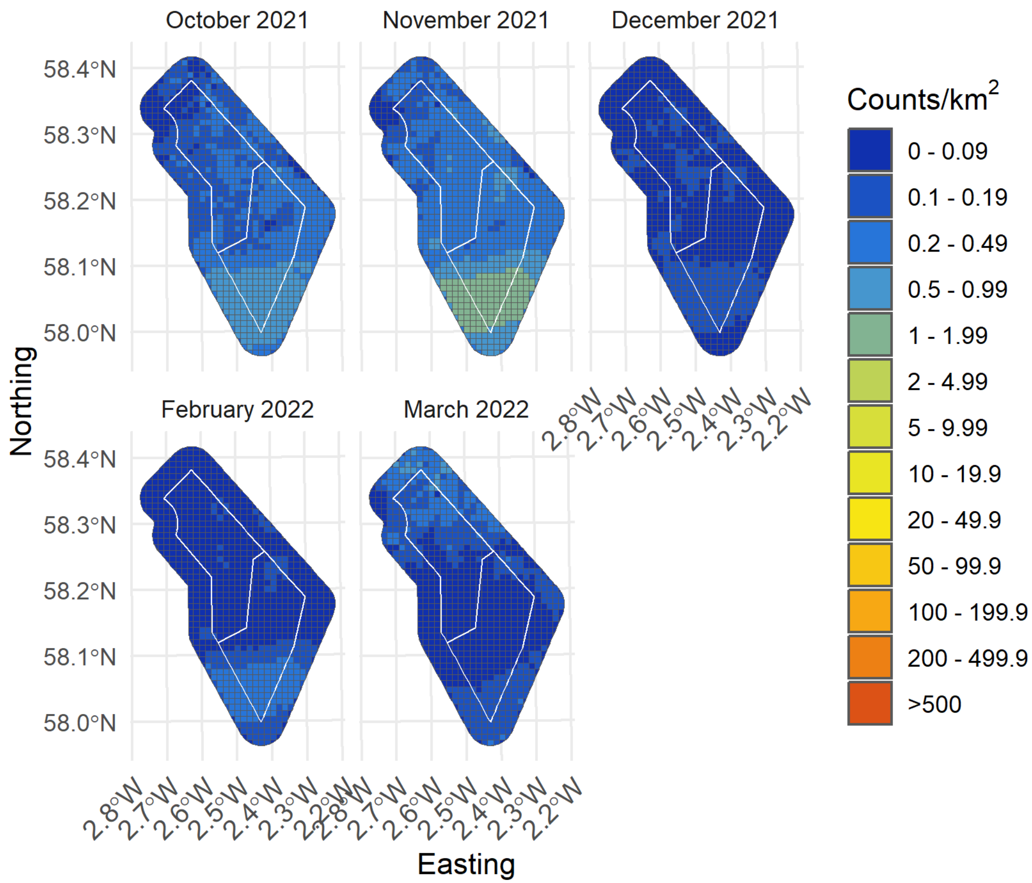


Figure 7. Lower confidence limit of density of all kittiwake in the survey area for months with sufficient observations between October 2021 and March 2022

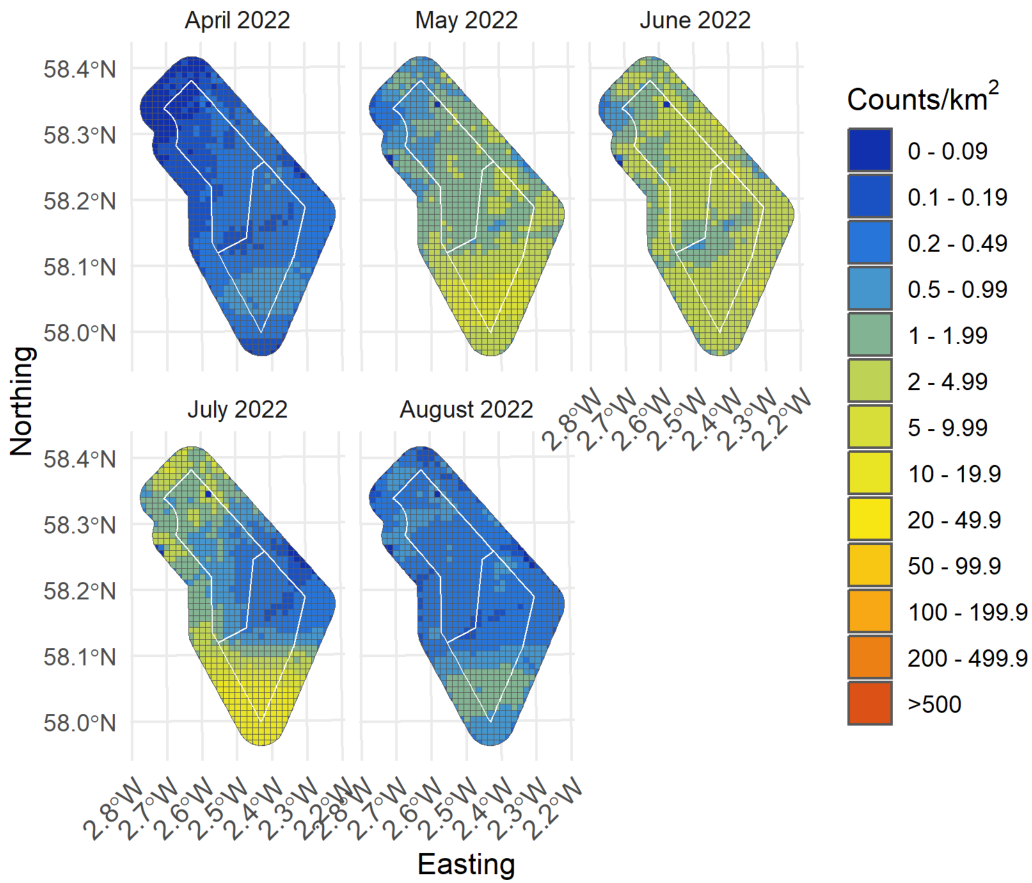


Figure 8. Lower confidence limit of density of all kittiwake in the survey area for months with sufficient observations between April and August 2022

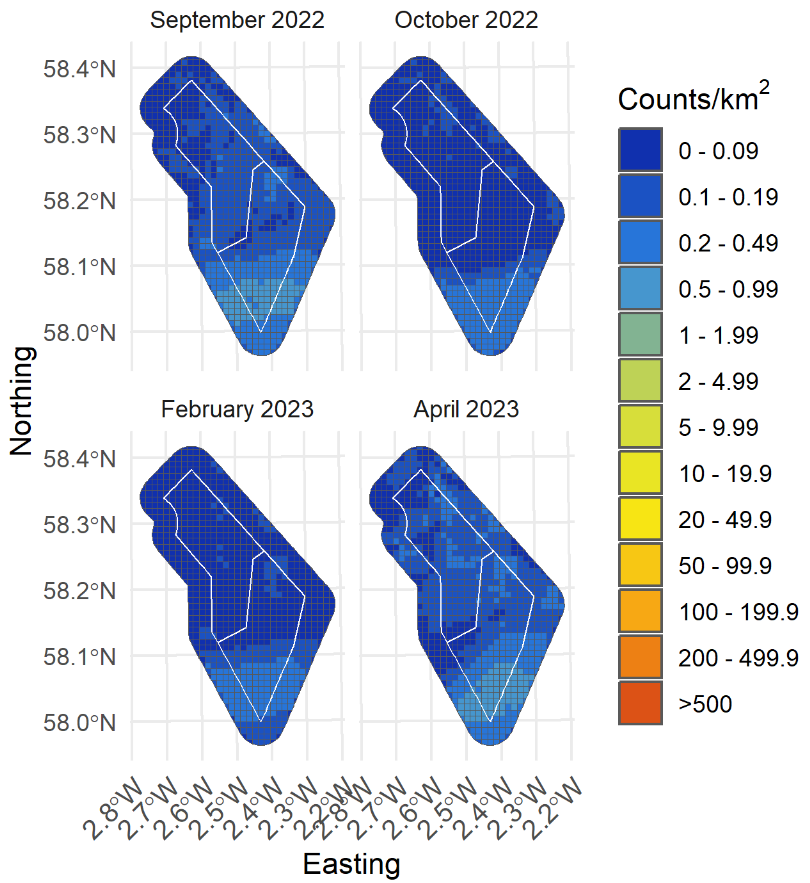


Figure 9. Lower confidence limit of density of all kittiwake in the survey area for months with sufficient observations between September 2022 and April 2023

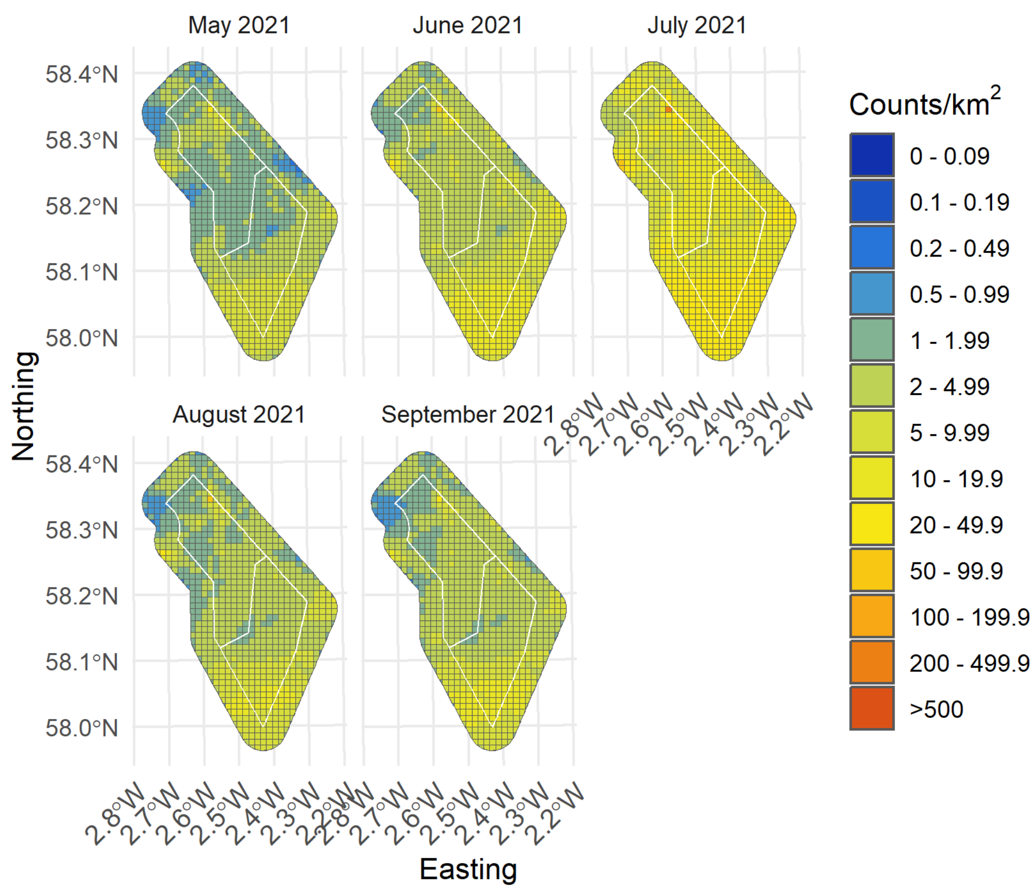


Figure 10. Upper confidence limit of density of all kittiwake in the survey area for months with sufficient observations between May and September 2021

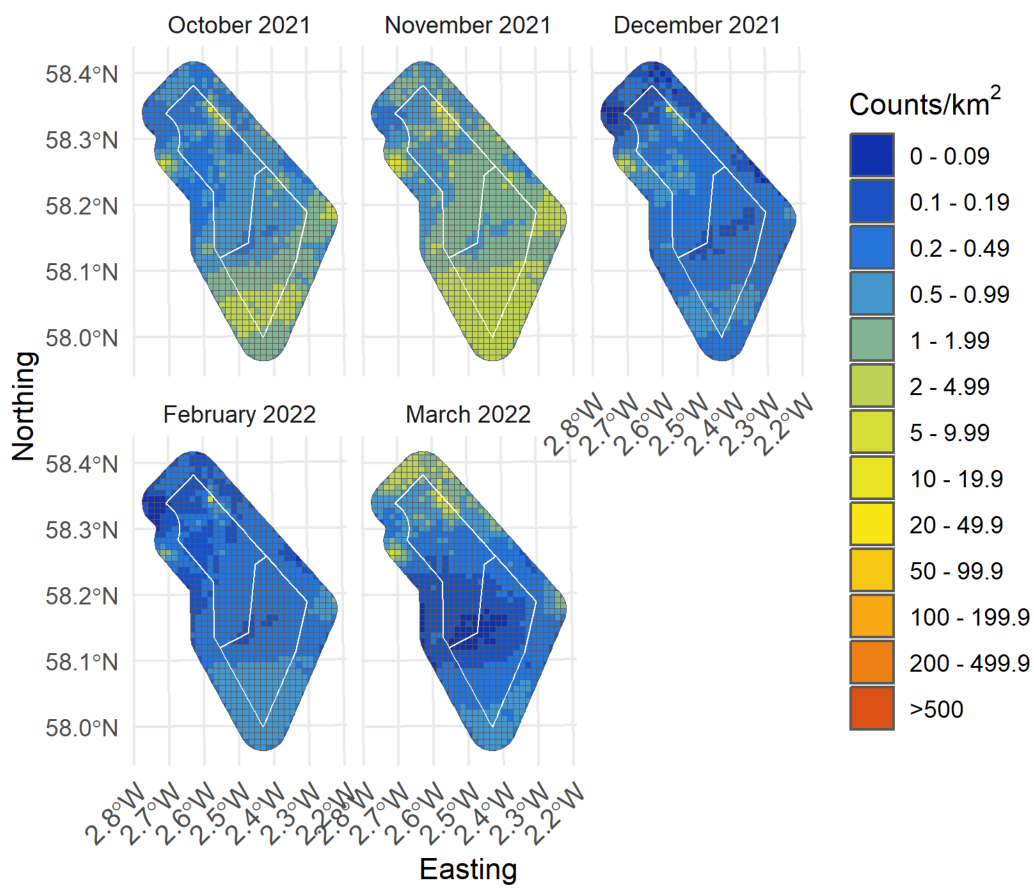


Figure 11. Upper confidence limit of density of all kittiwake in the survey area for months with sufficient observations between October 2021 and March 2022

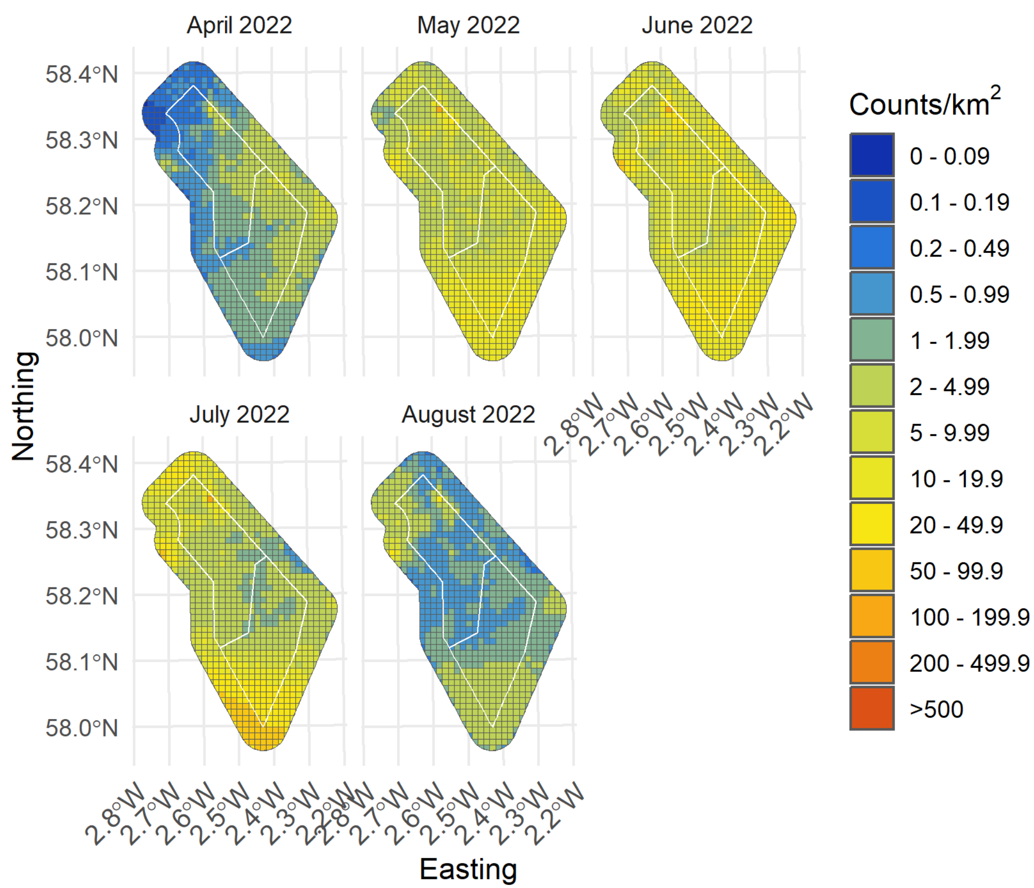


Figure 12. Upper confidence limit of density of all kittiwake in the survey area for months with sufficient observations between April and August 2022

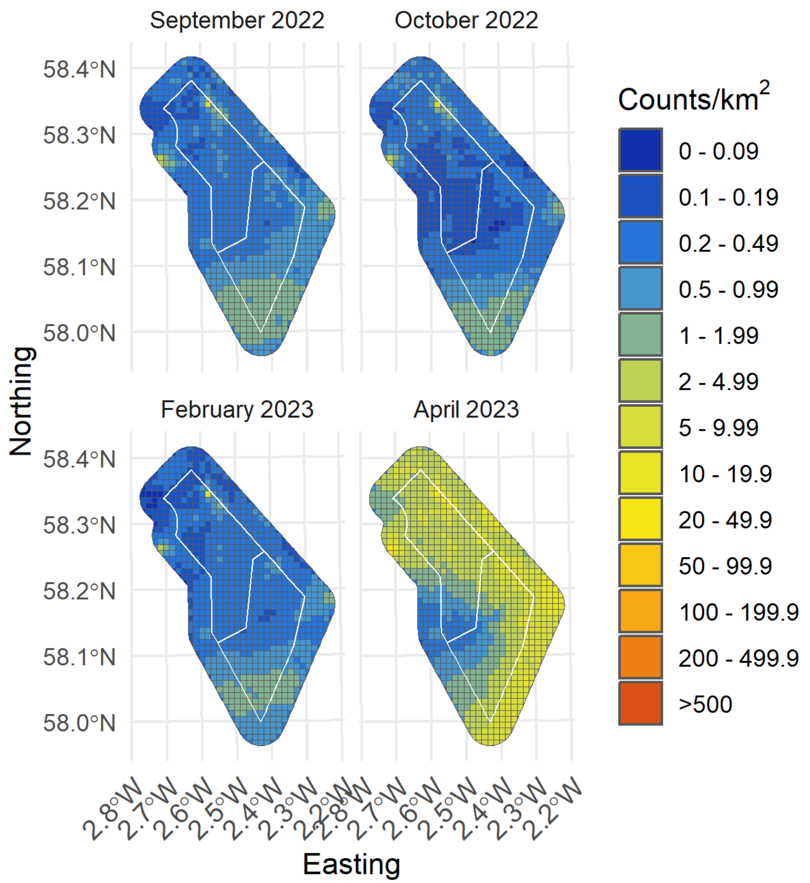


Figure 13. Upper confidence limit of density of all kittiwake in the survey area for months with sufficient observations between September 2022 and April 2023

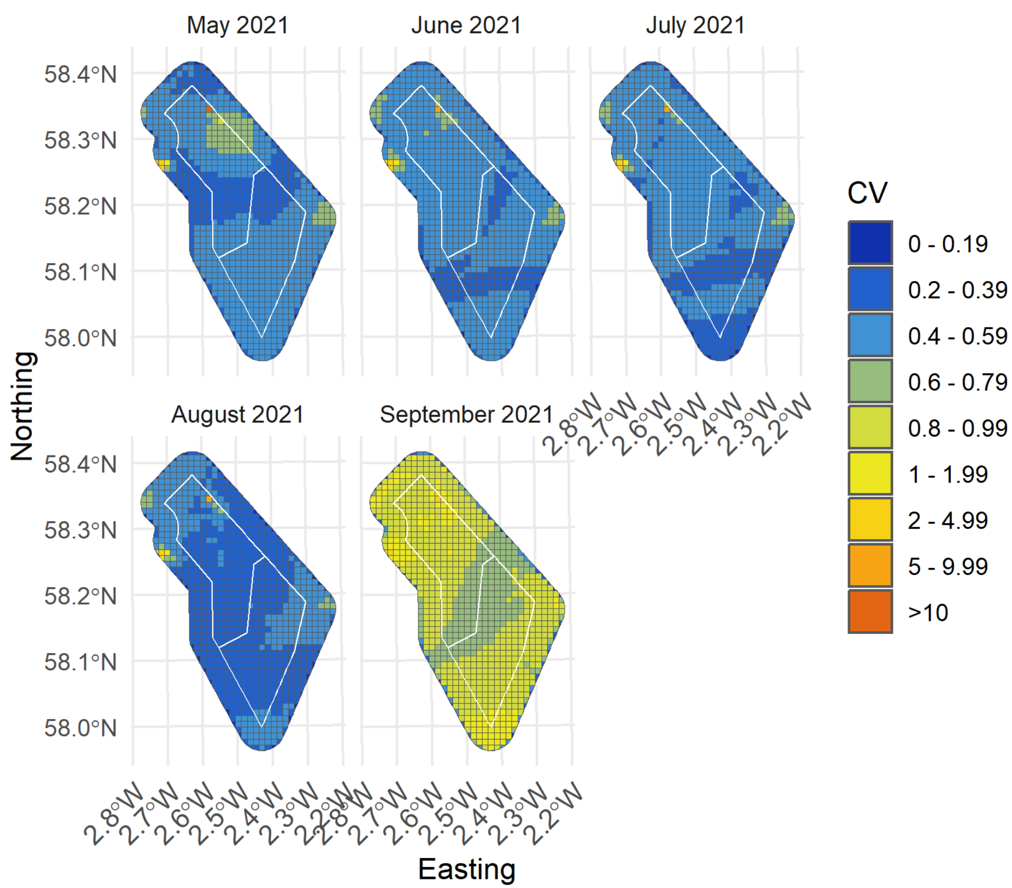


Figure 14. Spatial coefficient of variation of predicted densities of all kittiwake from MRSea across the survey area for months with sufficient observations between May and September 2021

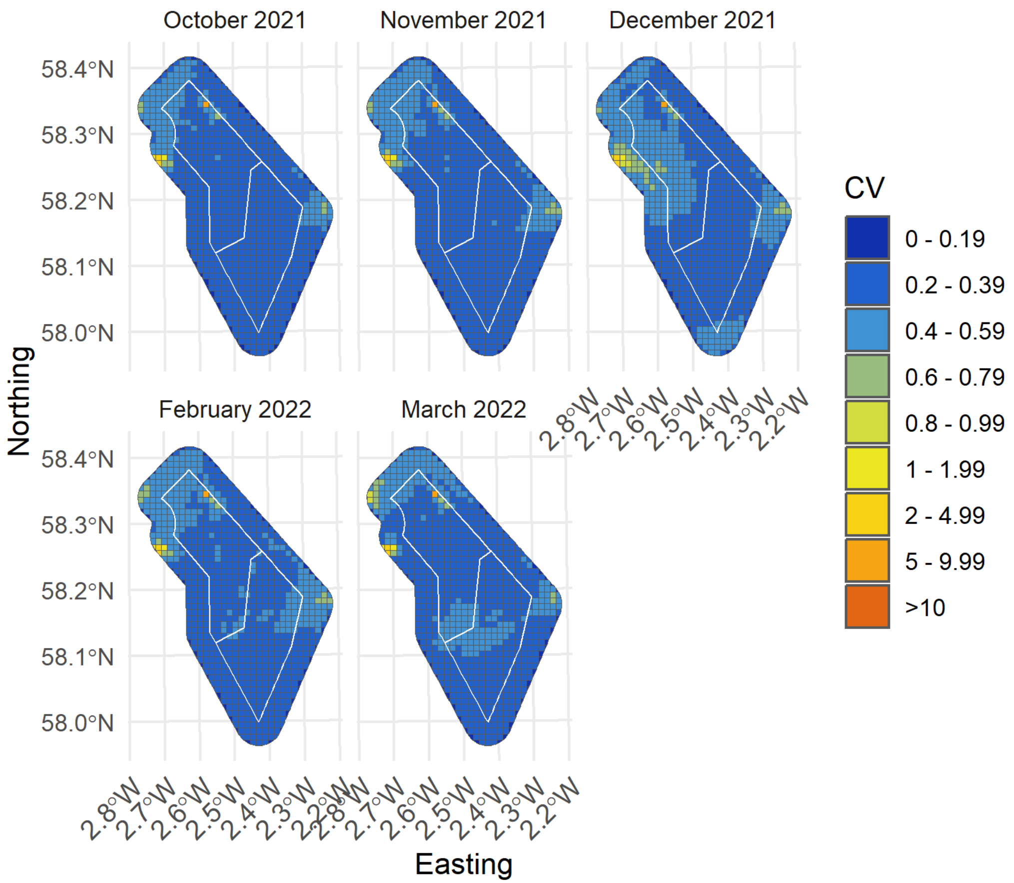


Figure 15. Spatial coefficient of variation of predicted densities of all kittiwake from MRSea across the survey area for months with sufficient observations between October 2021 and March 2022

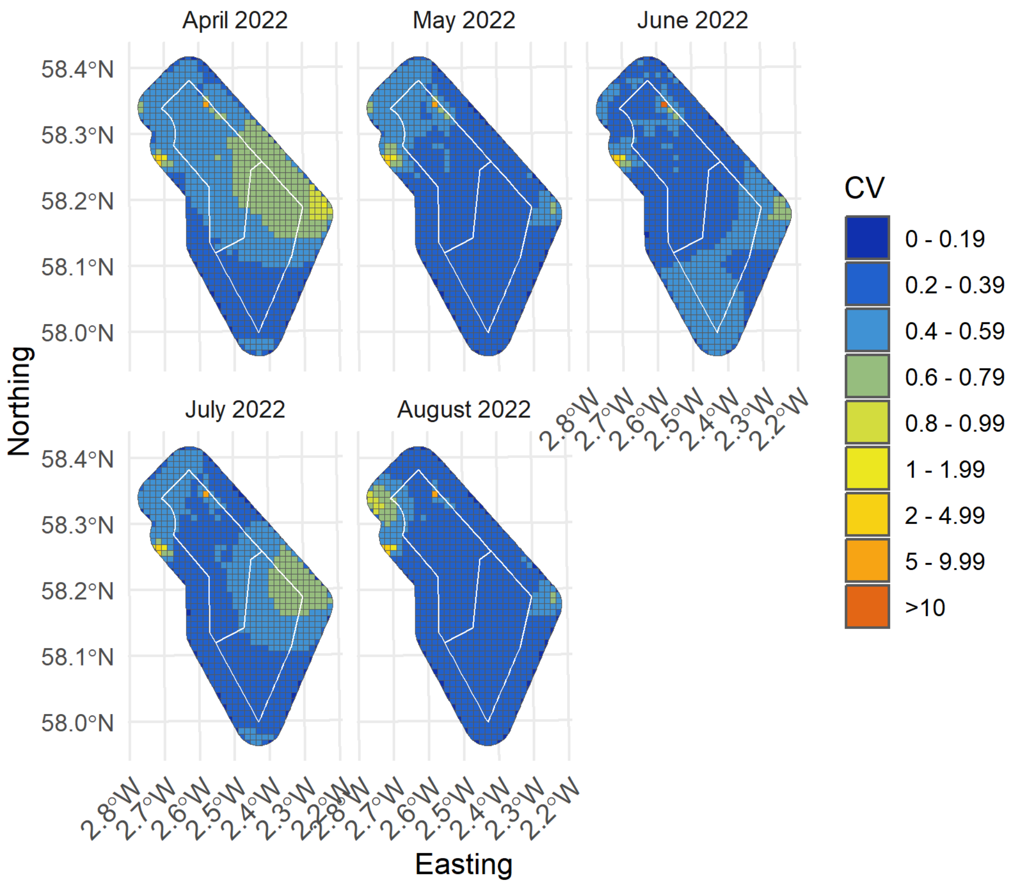


Figure 16. Spatial coefficient of variation of predicted densities of all kittiwake from MRSea across the survey area for months with sufficient observations between April and August 2022

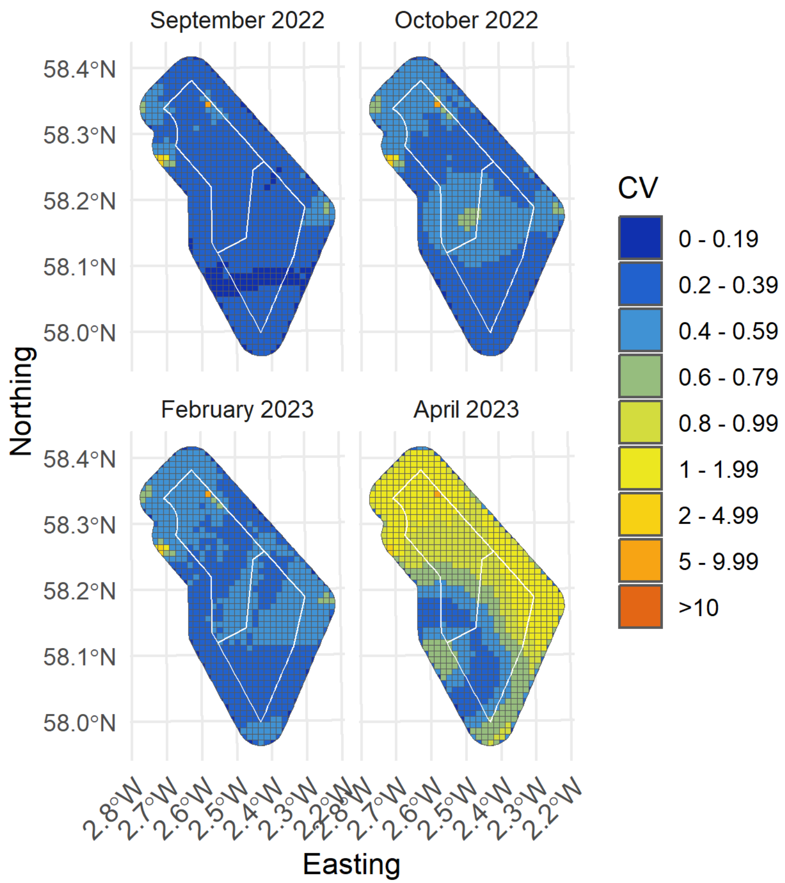


Figure 17. Spatial coefficient of variation of predicted densities of all kittiwake from MRSea across the survey area for months with sufficient observations between September 2022 and April 2023

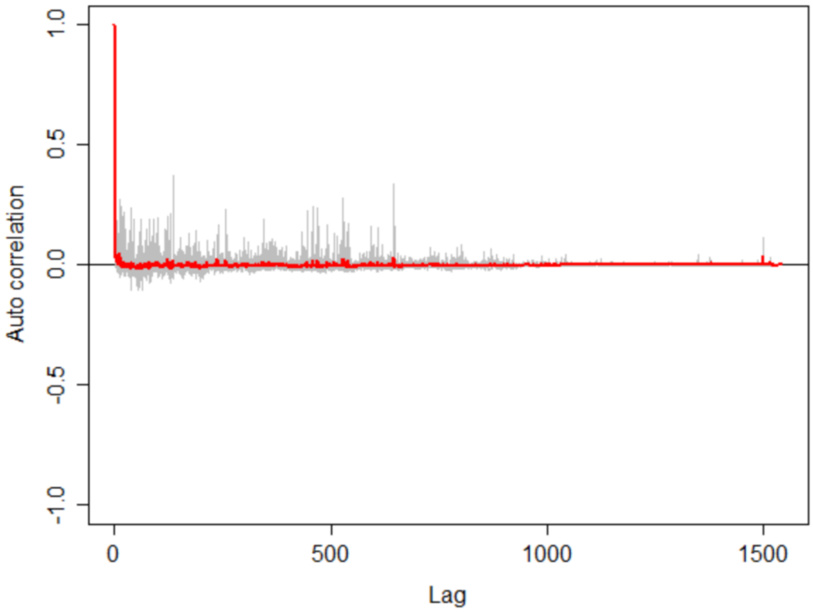


Figure 18. Autocorrelation test for kittiwake density surface models when using transect as a blocking feature in MRSea showing no significant correlation. A Runs test on the data prior to using transect as a blocking feature gave a p-value of << 0.0001 (i.e., that the data were significantly autocorrelated when not using a blocking feature)

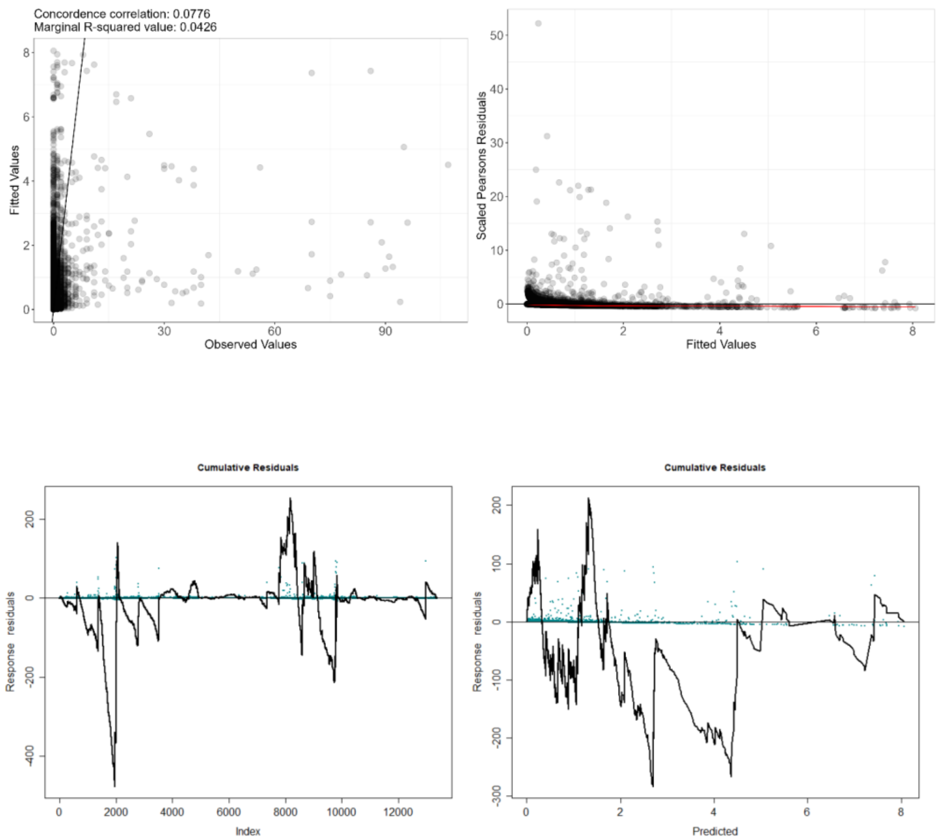


Figure 19. Fitted (MRSea predictions) versus observed counts of all kittiwake

Table 4. ANOVA results from the best MRSea model for all kittiwake as selected by cross-validation

Variable	Degrees of Freedom	Chi-square	P-value
Sandeel Density	3	31.23	<<0.001
Distance to Colony	5	17.63	0.003
X/Y (location)	10	41.85	<<0.001

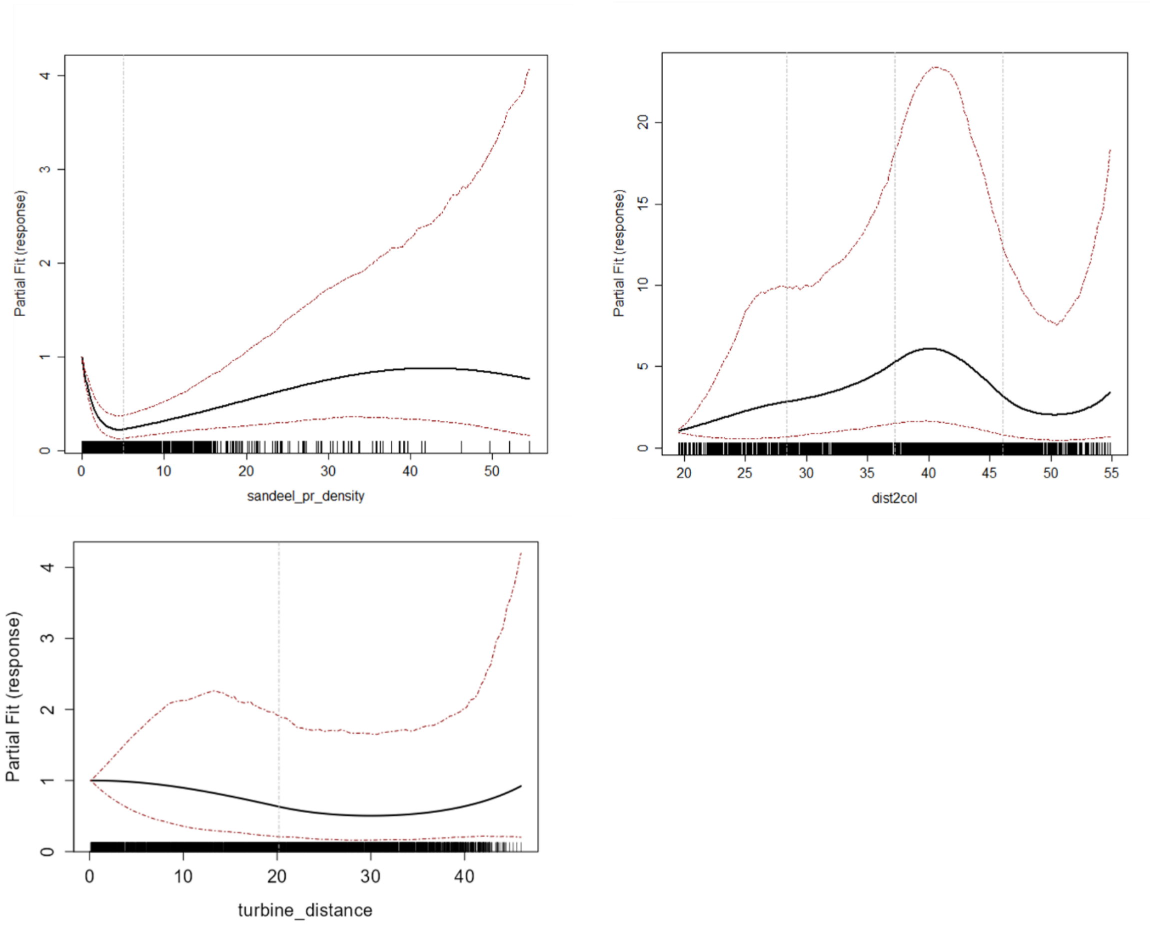


Figure 20. Partial dependence plots for significant variables for all kittiwake from MRSea. Note that distance to turbine was not a significant variable but was included to demonstrate the relationship. (Clockwise from top left: sandeel density, distance to colony, distance to turbine)

3.1.2 Flying Birds Model

Table 5. Candidate and final covariates for flying kittiwake model

Starting model covariates after VIF-based collinearity removal	Final model covariates after removal by SALSA
Survey ID	Sandeel Density
Bathymetry	Distance to colony
Sandeel Density	
Distance to colony	
Distance to turbine	
Standard deviation of Sea Surface Temperature	

Distribution maps generated using MRSea (Figure 21-36) suggest elevated densities of flying kittiwake during the breeding season (15 April – 31 August) of both years, but were higher in

2022. In the non-breeding season, there is a notable drop in kittiwake density. The highest densities were observed in the south of the study area during July 2022.

Model fit was poor with a marginal R squared value of 0.06 and root mean squared error of 3.94. Cumulative residuals in the model showed that there was a poor relationship between predicted and observed values particularly when predicted counts were above ~2.5 birds/km^2 (Figure 38).

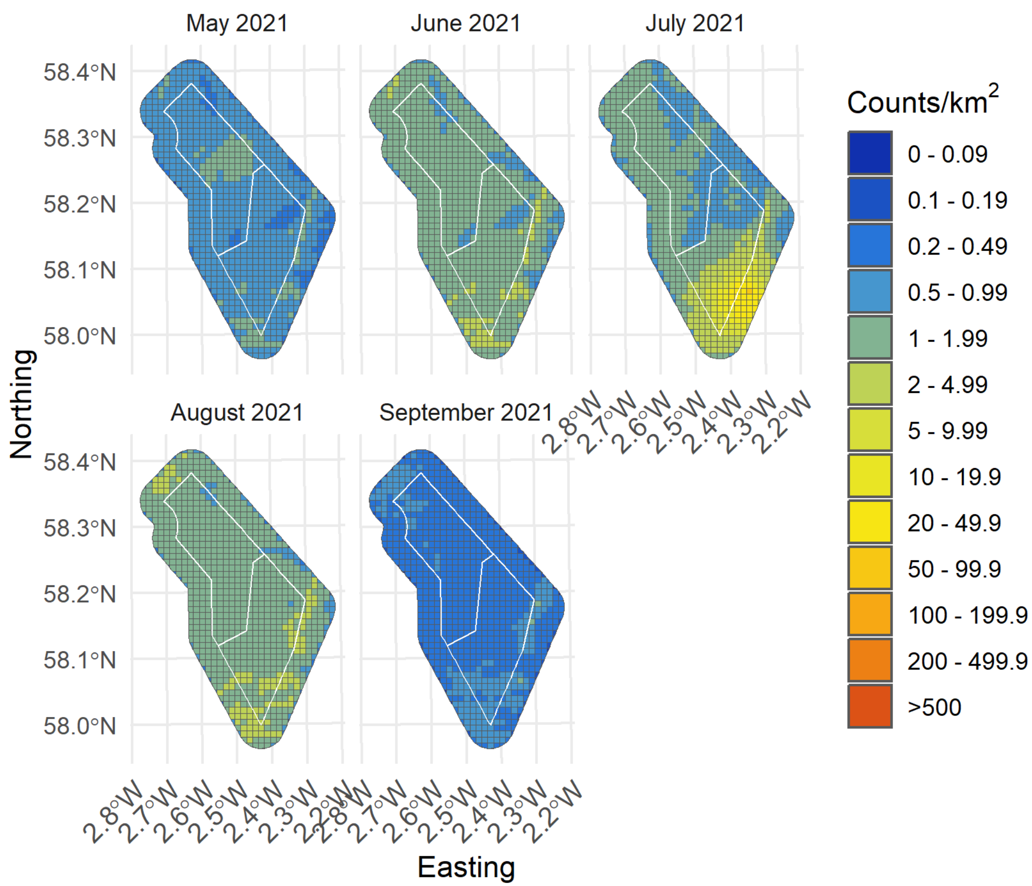


Figure 21. Median density of flying kittiwakes in the survey area for months with sufficient observations between May and September 2021

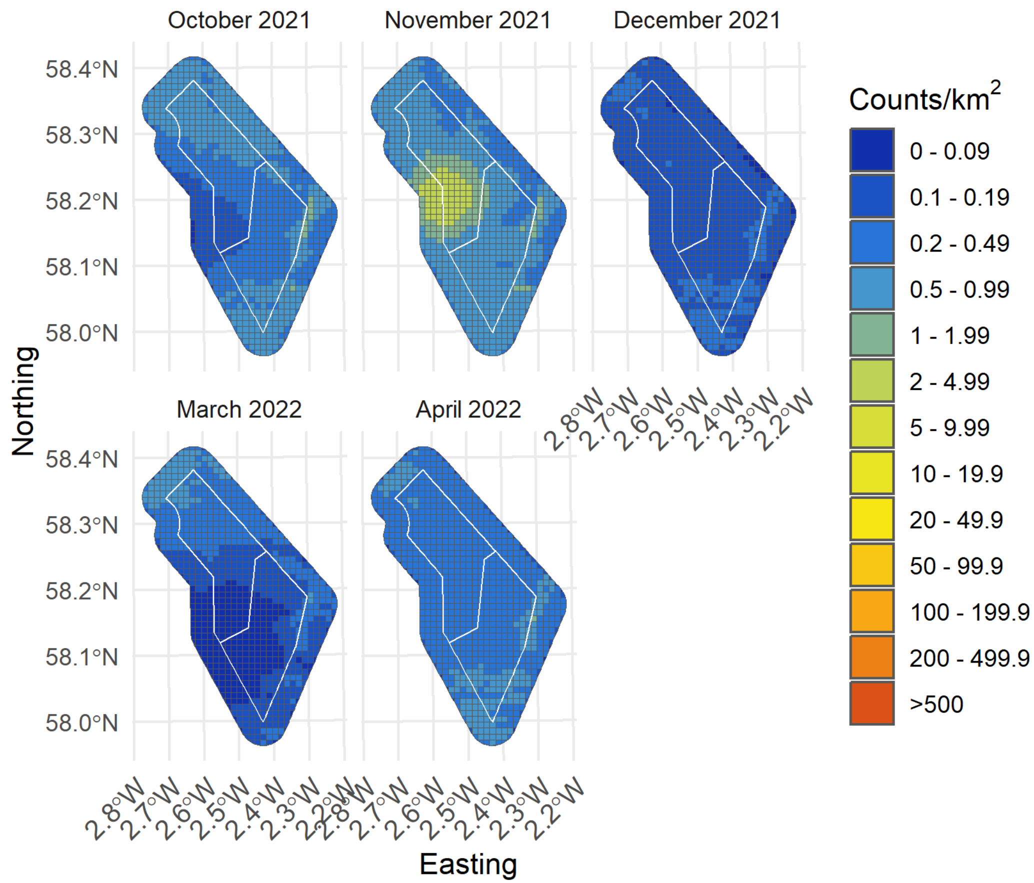


Figure 22. Median density of flying kittiwakes in the survey area for months with sufficient observations between October 2021 and April 2022

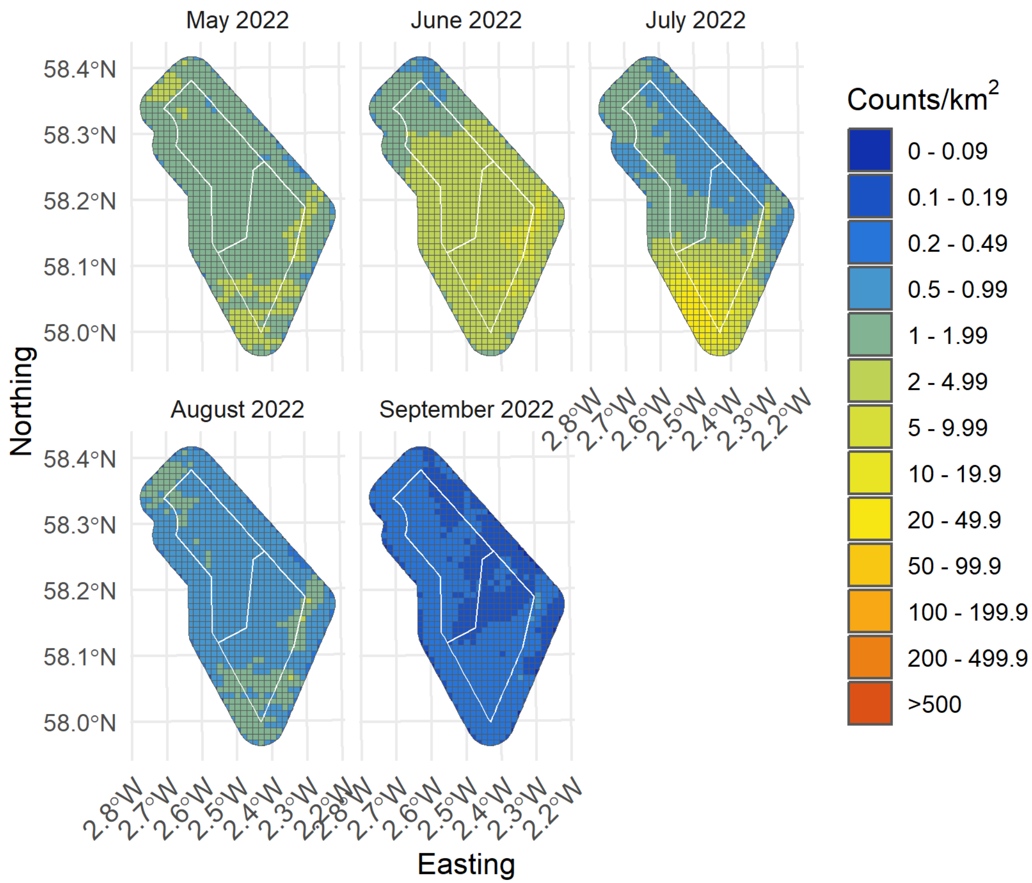


Figure 23. Median density of flying kittiwakes in the survey area for months with sufficient observations between May and September 2022

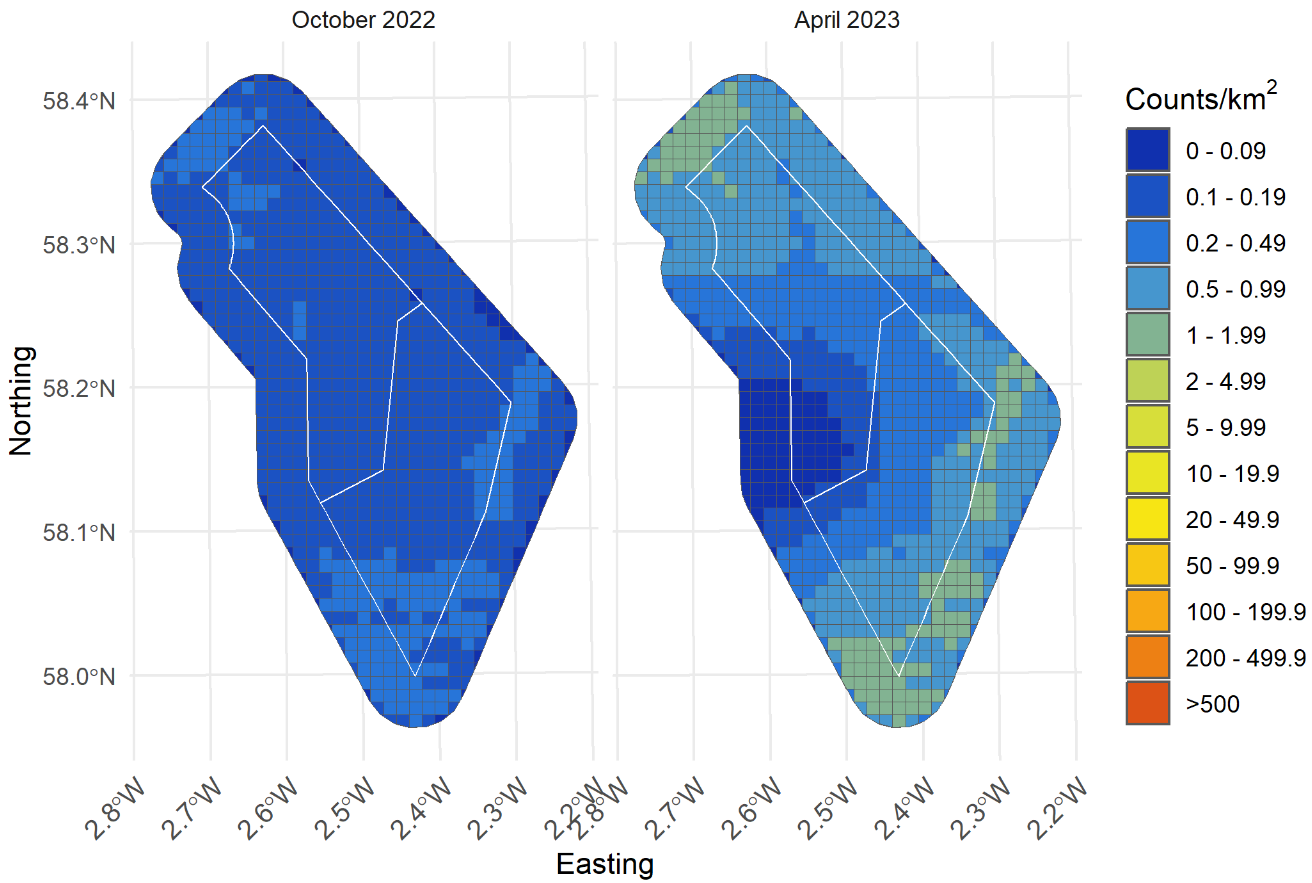


Figure 24. Median density of flying kittiwakes in the survey area for months with sufficient observations between October 2022 and April 2023

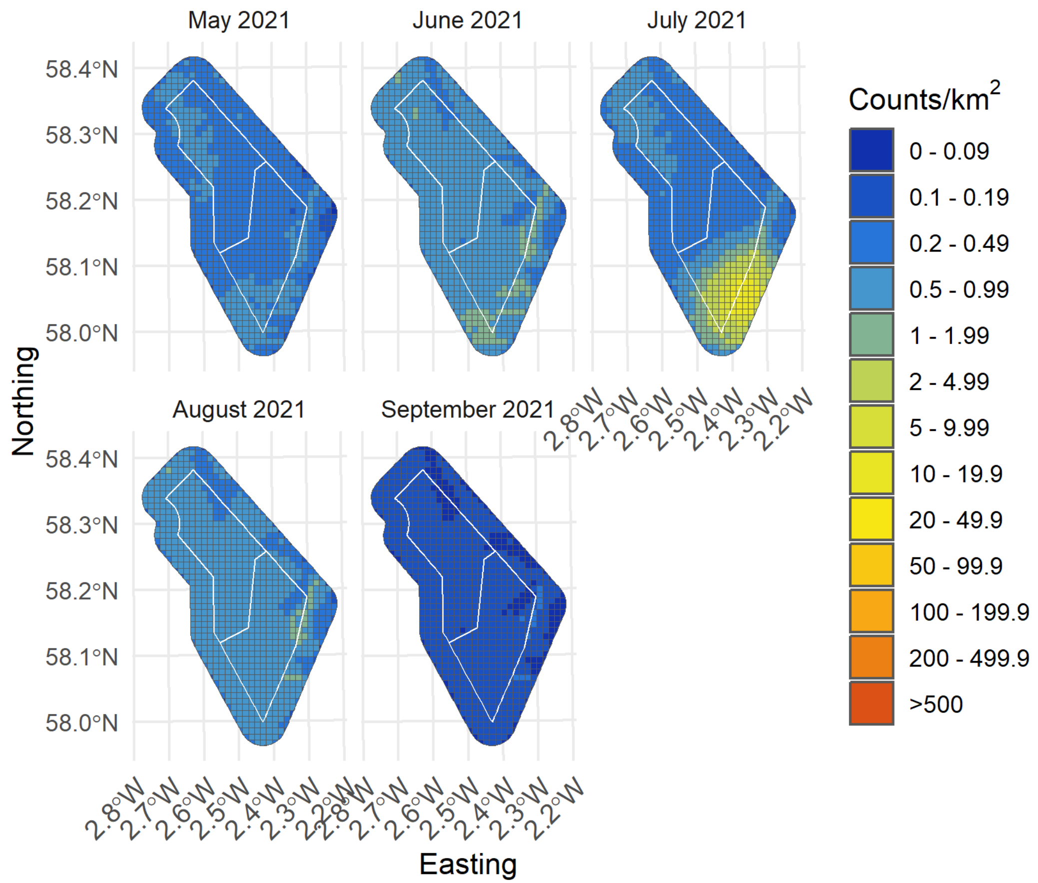


Figure 25. Lower confidence limit of density of flying kittiwakes in the survey area for months with sufficient observations between May and September 2021

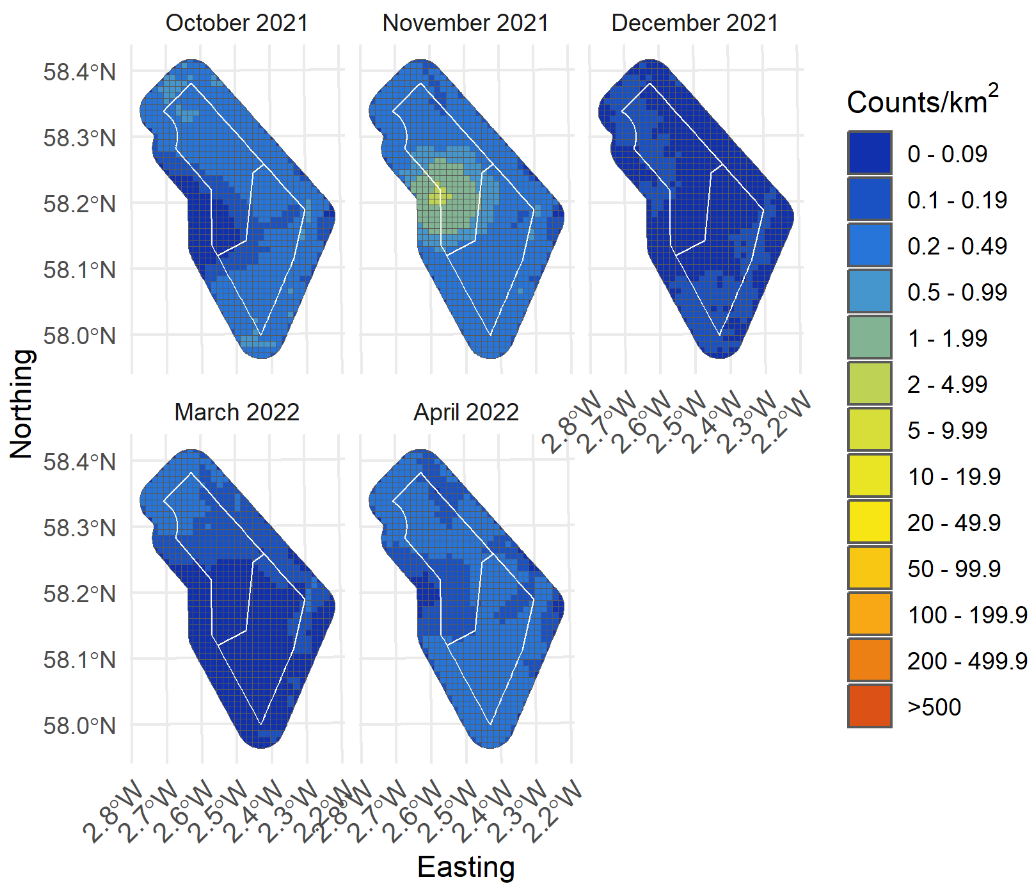


Figure 26. Lower confidence limit of density of flying kittiwakes in the survey area for months with sufficient observations between October 2021 and April 2022

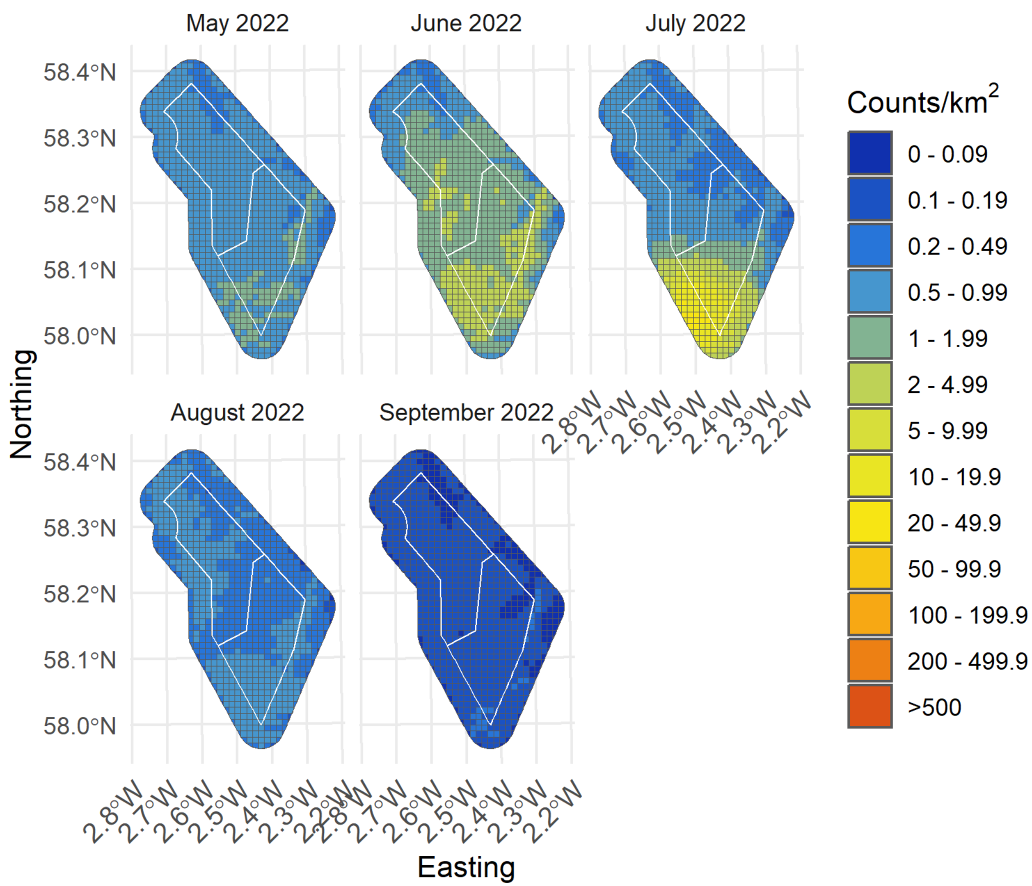


Figure 27. Lower confidence limit of density of flying kittiwakes in the survey area for months with sufficient observations between May and September 2022

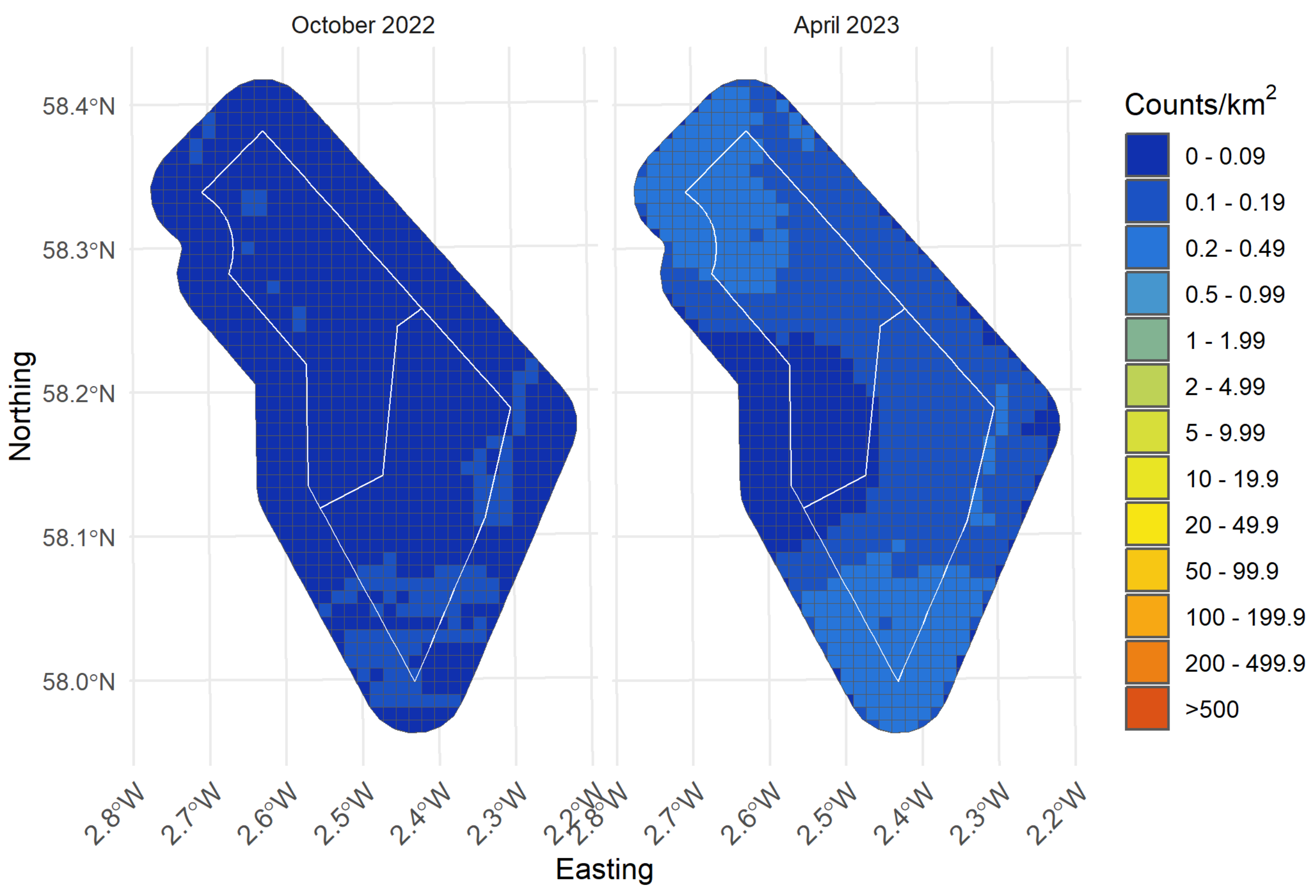


Figure 28. Lower confidence limit of density of flying kittiwakes in the survey area for months with sufficient observations between October 2022 and April 2023

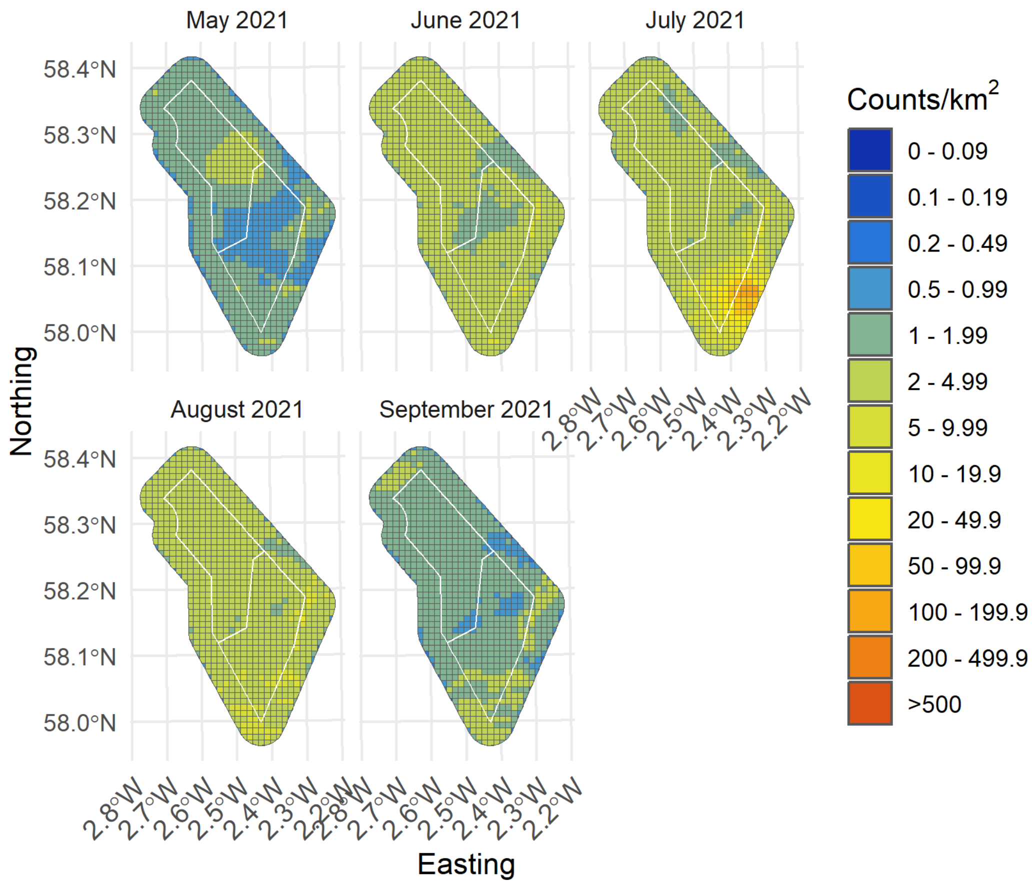


Figure 29. Upper confidence limit of density of flying kittiwakes in the survey area for months with sufficient observations between May and September 2021

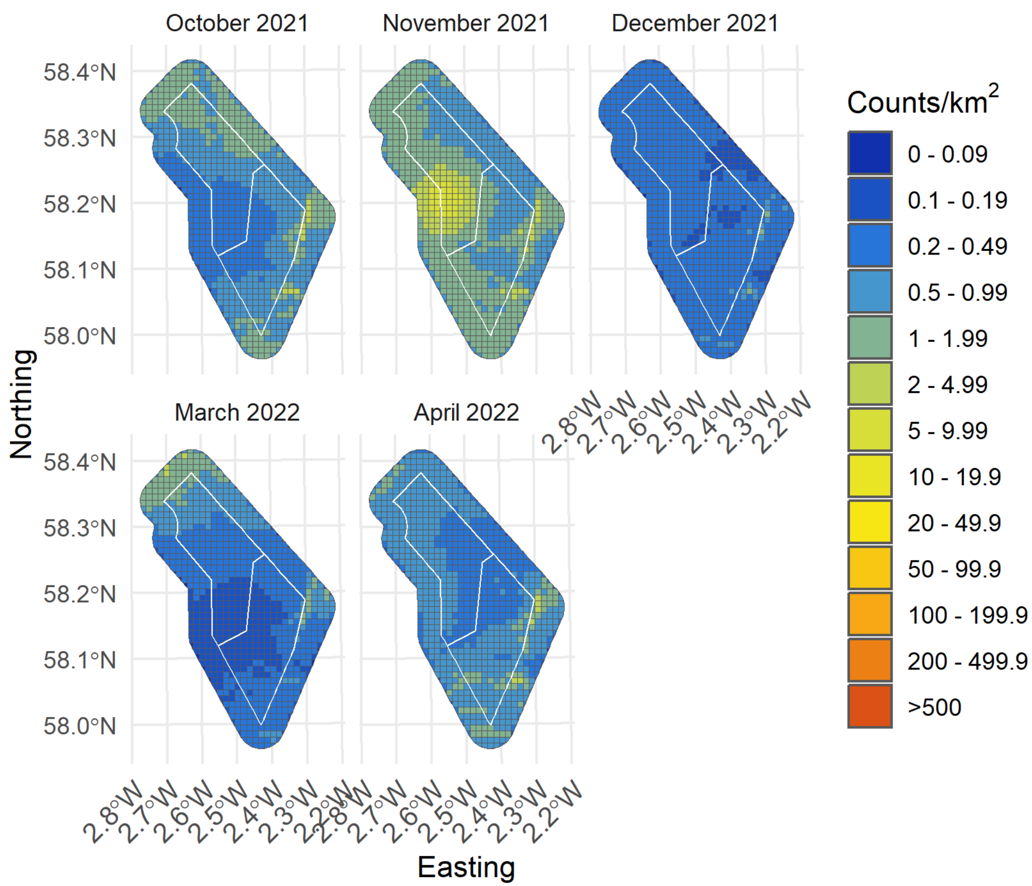


Figure 30. Upper confidence limit of density of flying kittiwakes in the survey area for months with sufficient observations between October 2021 and April 2022

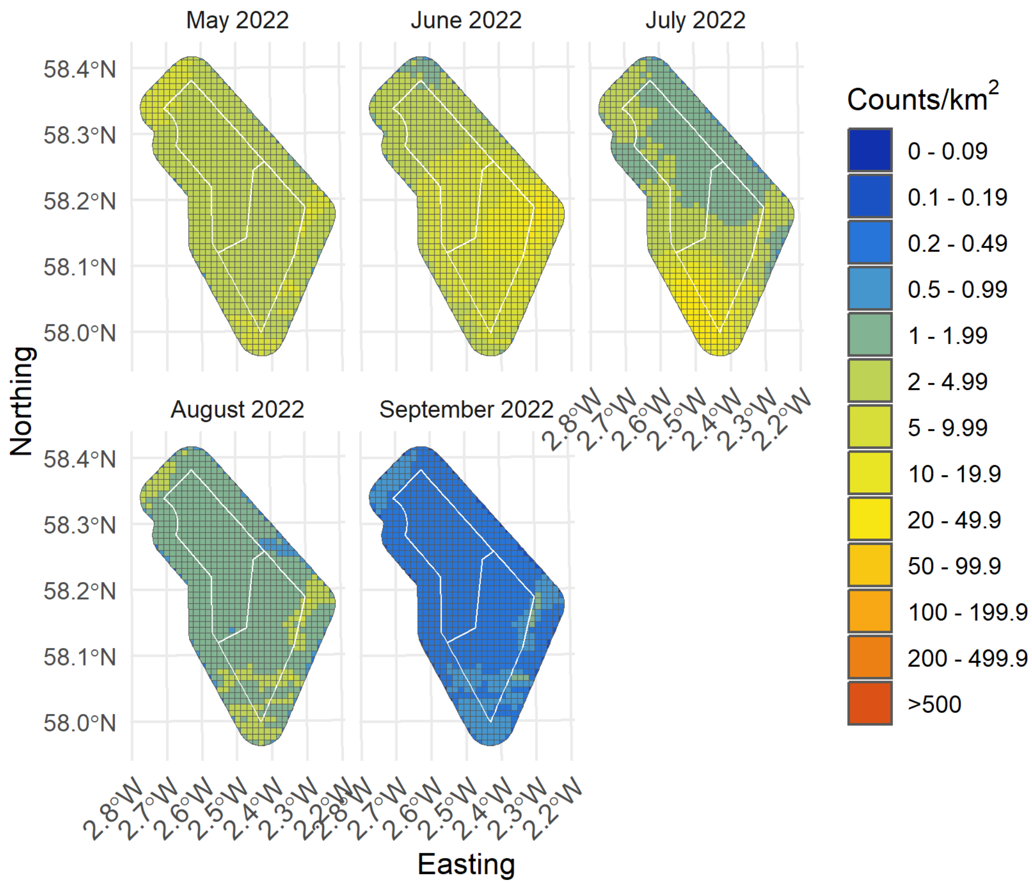


Figure 31. Upper confidence limit of density of flying kittiwakes in the survey area for months with sufficient observations between October 2021 and April 2022

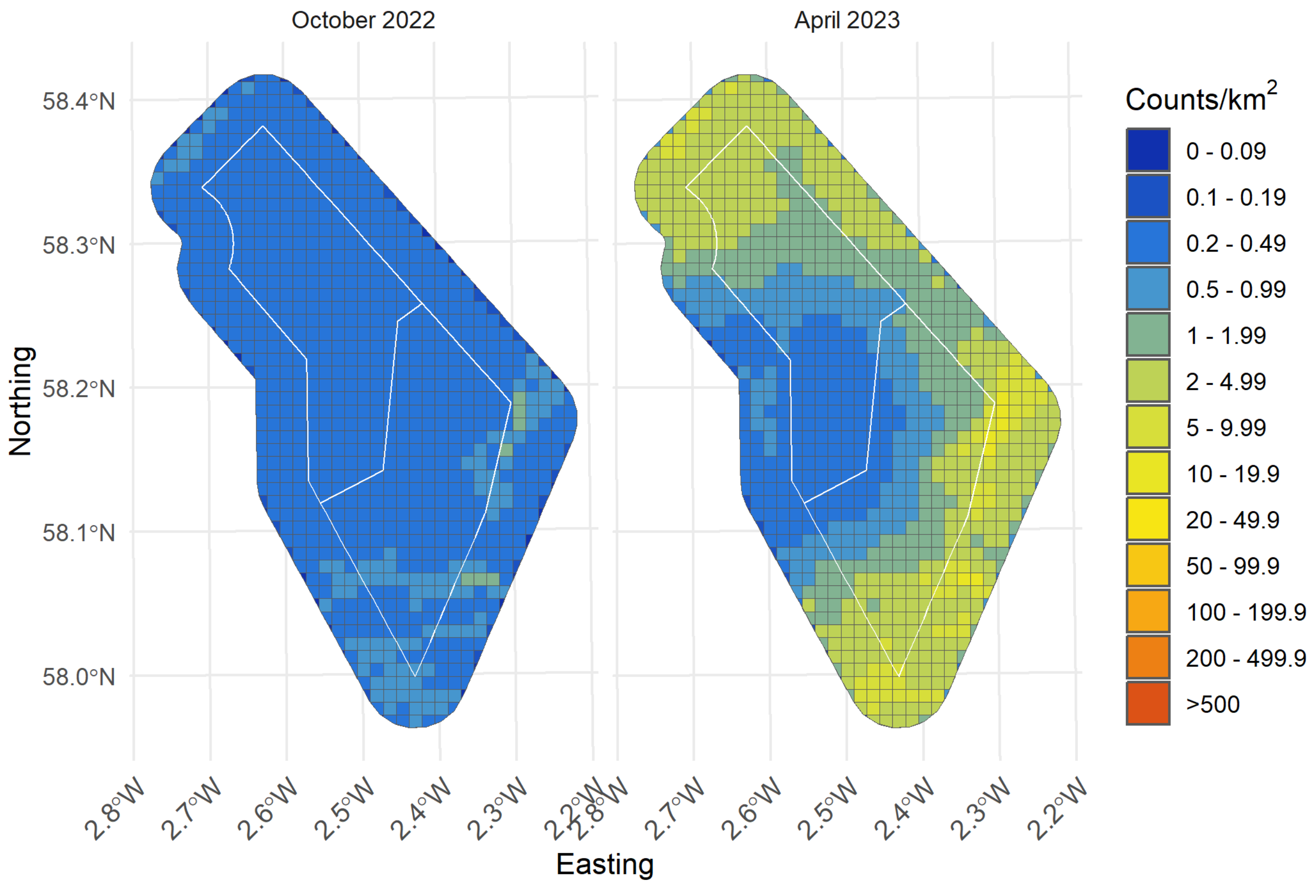


Figure 32. Upper confidence limit of density of flying kittiwakes in the survey area for months with sufficient observations between October 2022 and April 2023

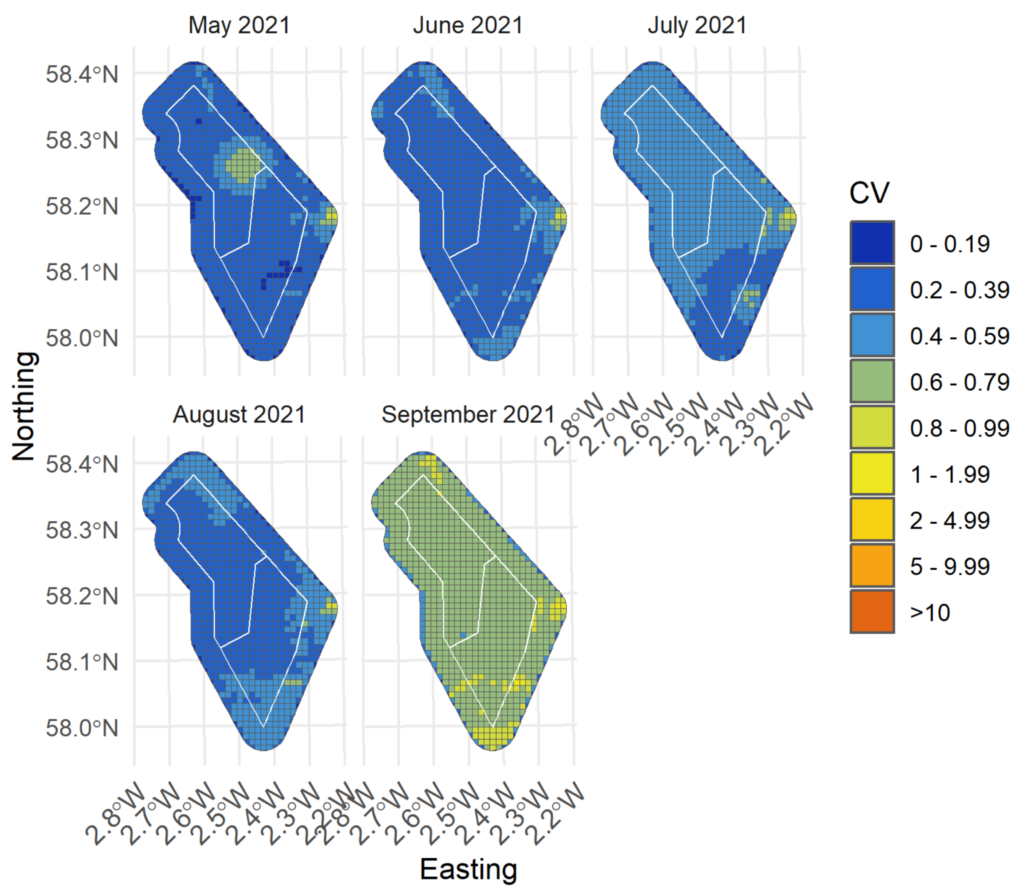


Figure 33. Spatial coefficient of variation of predicted densities of flying kittiwakes from MRSea across the survey area for months with sufficient observations between May and September 2021

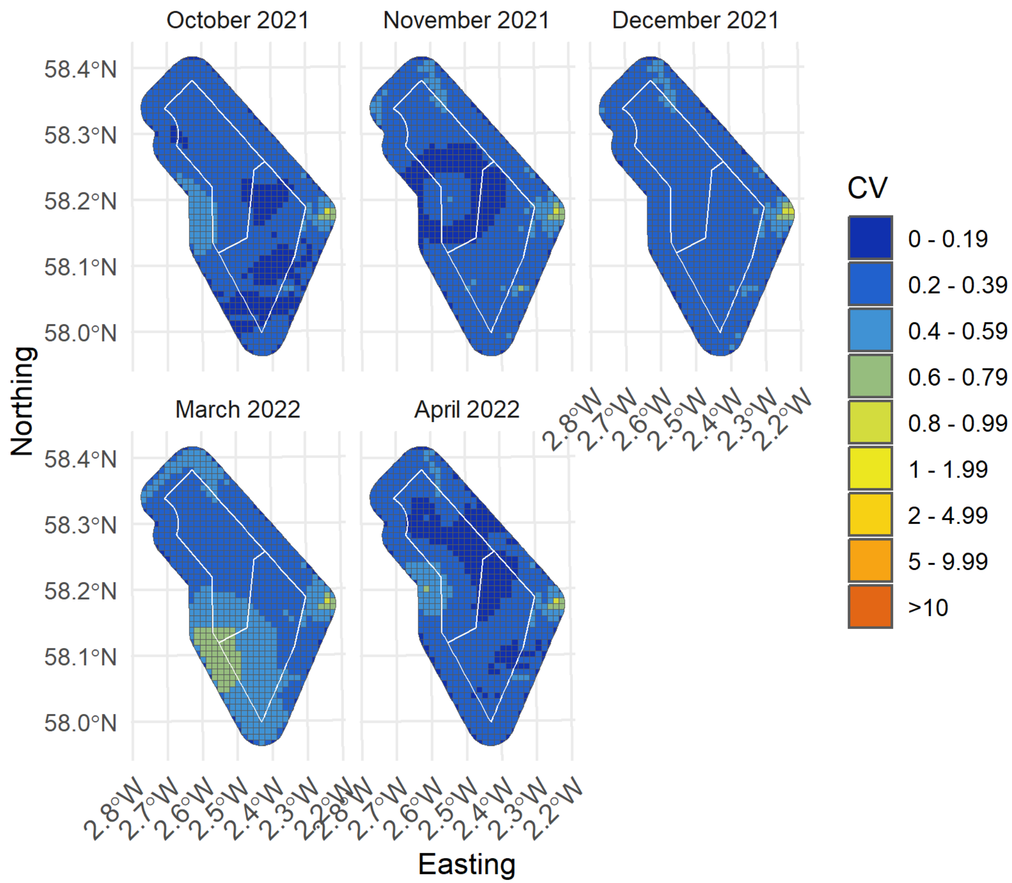


Figure 34. Spatial coefficient of variation of predicted densities of flying kittiwakes from MRSea across the survey area for months with sufficient observations between October 2021 and April 2022

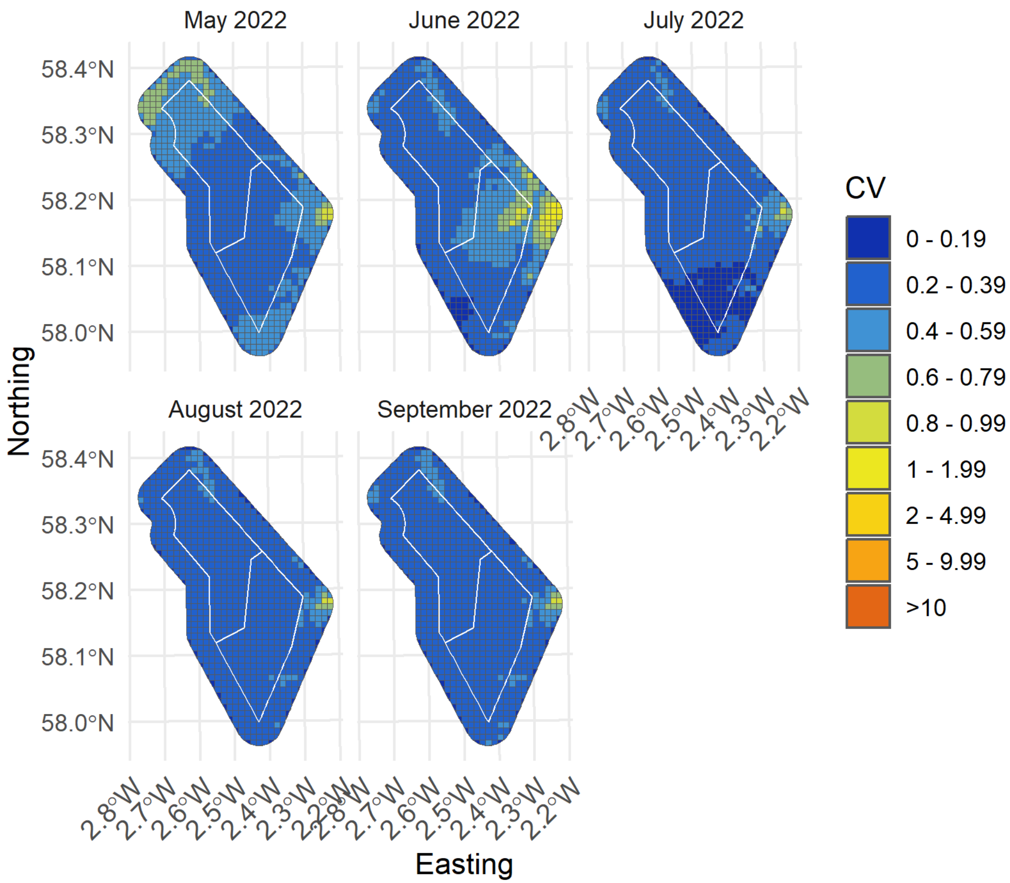


Figure 35. Spatial coefficient of variation of predicted densities of flying kittiwakes from MRSea across the survey area for months with sufficient observations between May and September 2022

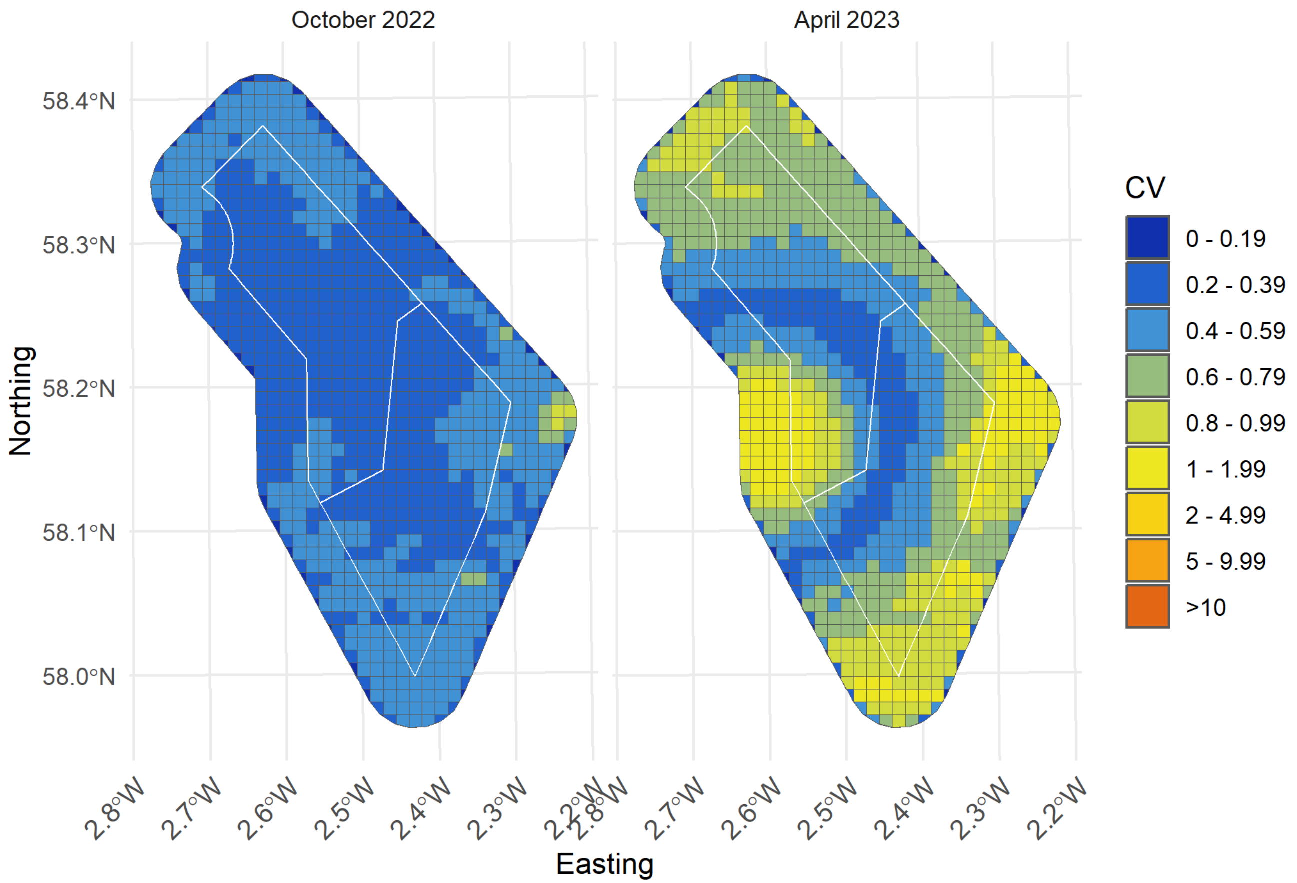


Figure 36. Spatial coefficient of variation of predicted densities of flying kittiwakes from MRSea across the survey area for months with sufficient observations between October and April 2023

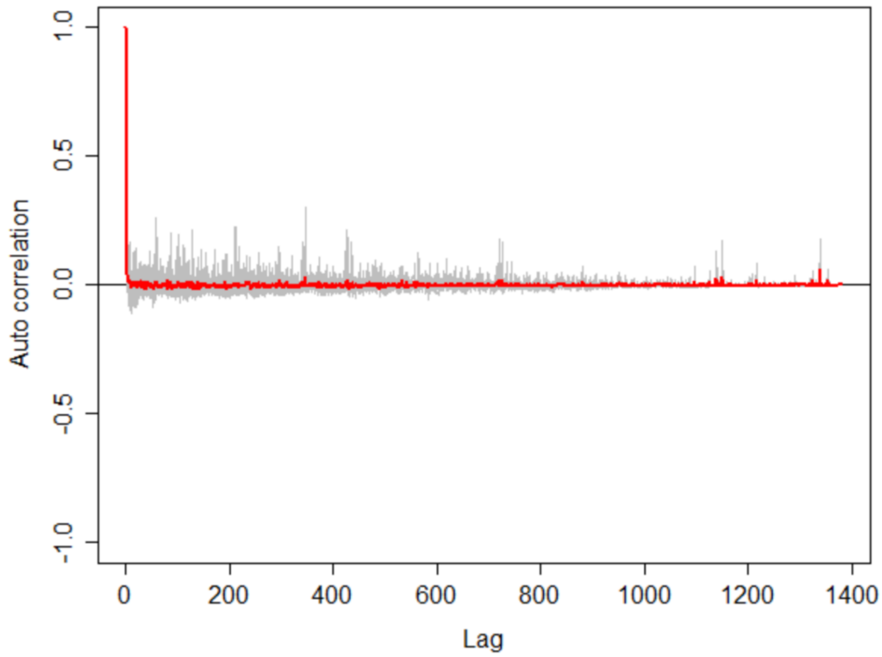


Figure 37. Autocorrelation test for kittiwake density surface models when using transect as a blocking feature in MRSea showing no significant correlation. A Runs test on the data prior to using transect as a blocking feature gave a p-value of << 0.0001 (i.e., that the data were significantly autocorrelated when not using a blocking feature)

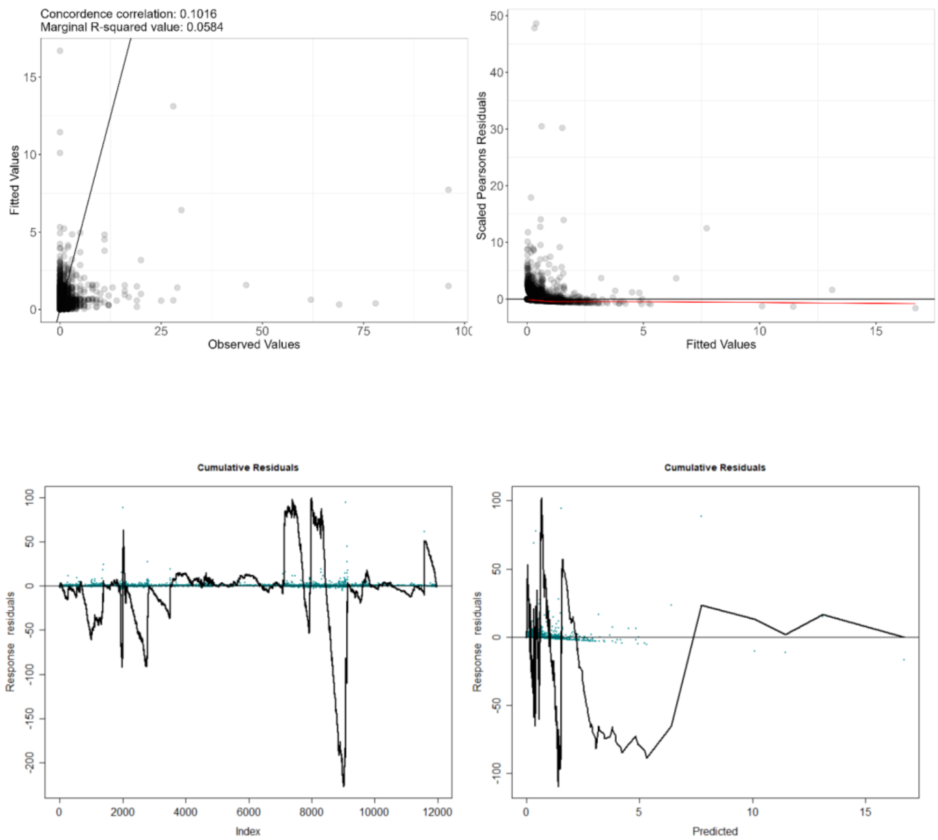


Figure 38. Fitted (MRSea predictions) versus observed counts of flying kittiwake

Table 6. ANOVA results from the best MRSea model for flying kittiwakes as selected by cross-validation

Variable	Degrees of Freedom	Chi-square	P-value
Bathymetry	1	4.32	0.038
Distance to Colony	3	2.01	0.569
Distance to Turbine	1	0.98	0.323
X/Y (location)	10	127.07	<<0.001

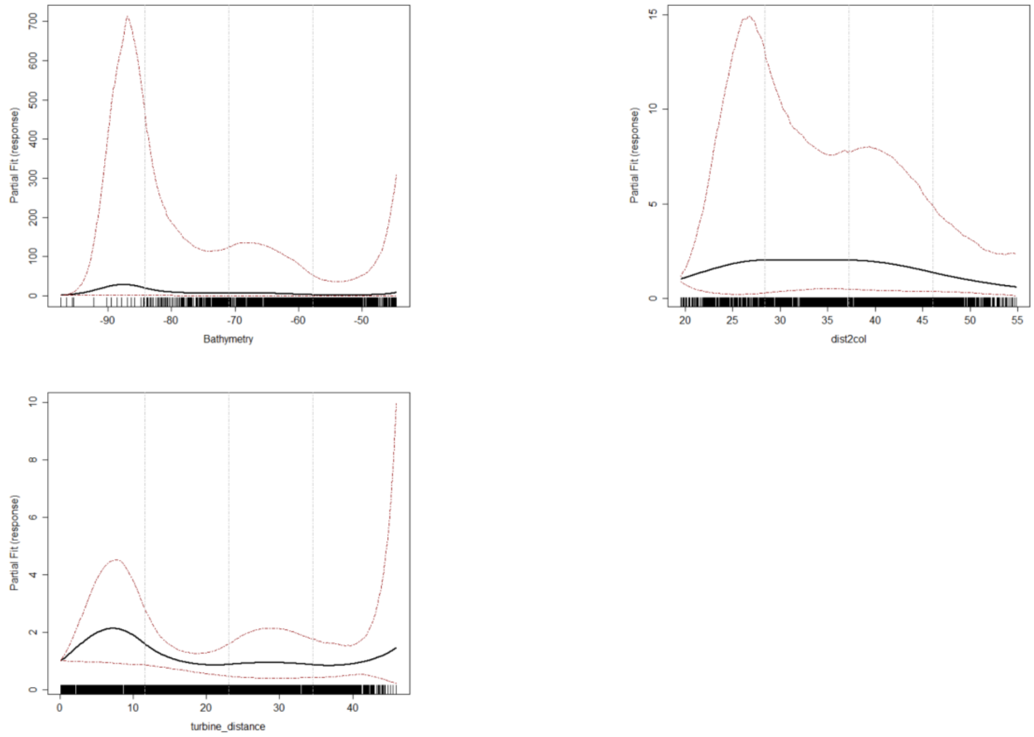


Figure 39. Partial dependence plots for significant variables for flying kittiwakes from MRSea. (Clockwise from top left: bathymetry, distance to colony, distance to turbine)

3.2 Gannet

3.2.1 All Birds Model

Table 7. Candidate and final covariates for all gannets model

Starting model covariates after VIF-based collinearity removal	Final model covariates after removal by SALSA
Aspect	Aspect
Slope	Slope
Bathymetry	Bathymetry
Sandeel Presence	Sandeel Presence
Distance to colony	Distance to colony
Distance to turbine	Distance to turbine
Standard deviation of Sea Surface Temperature	Standard deviation of Sea Surface Temperature

Distribution maps generated using MRSea (Figure 39-40) suggest that gannets are distributed mostly across the southern half of the study area during the breeding season (15 March – 30 September). In the non-breeding season, gannets were largely absent from the study area. The highest densities were observed along the south east border of the study area in July 2022.

Model fit was poor with a marginal R squared value of 0.04 and root mean squared error of 0.87. Cumulative residuals in the model showed that there was a poor relationship between predicted and observed values across most of the range of predicted values, with the model generally over predicting for values under 0.4 birds/km^2, and over predicting for values above. (Figure 48).

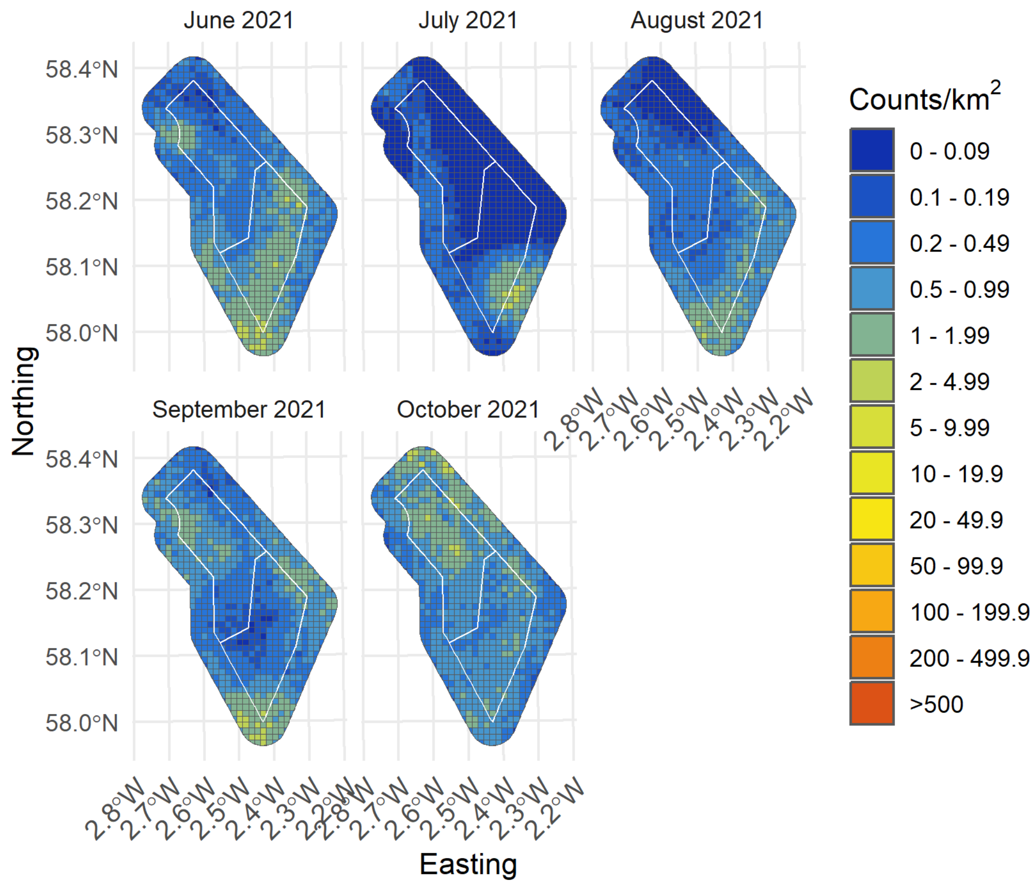


Figure 40. Median density of all gannets in the survey area for months with sufficient observations between June and October 2021

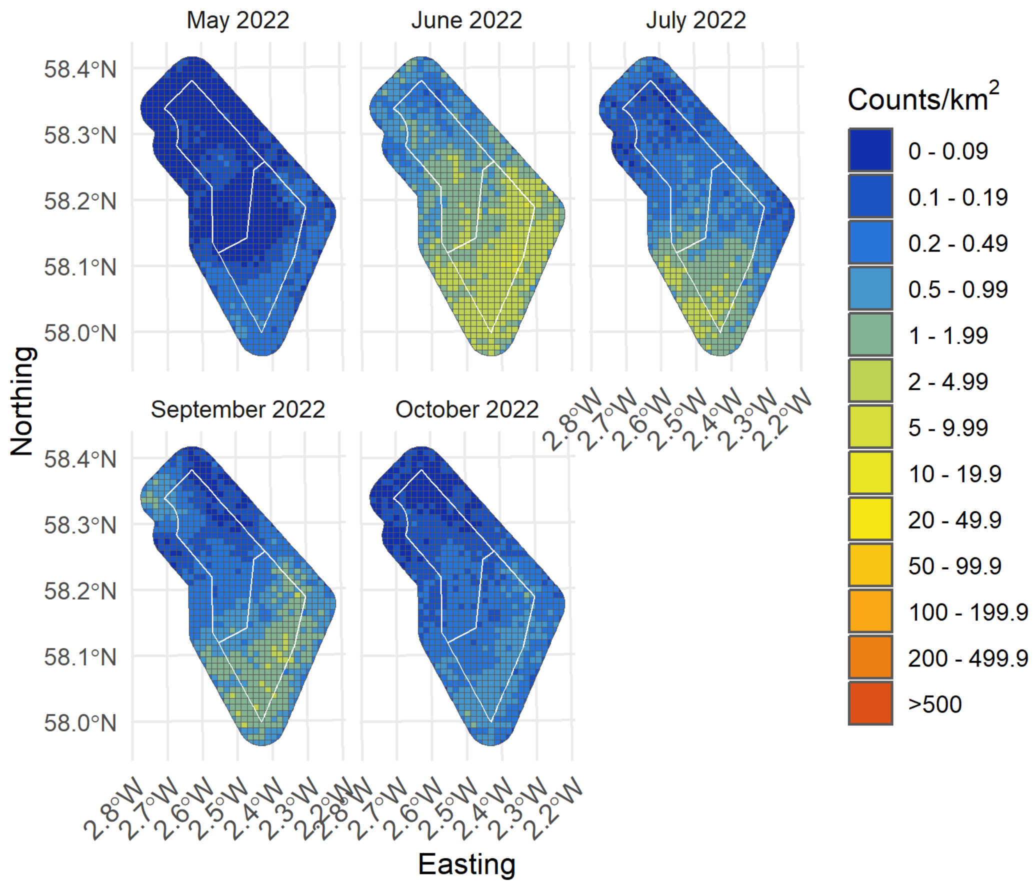


Figure 41. Median density of all gannets in the survey area for months with sufficient observations between May and October 2022

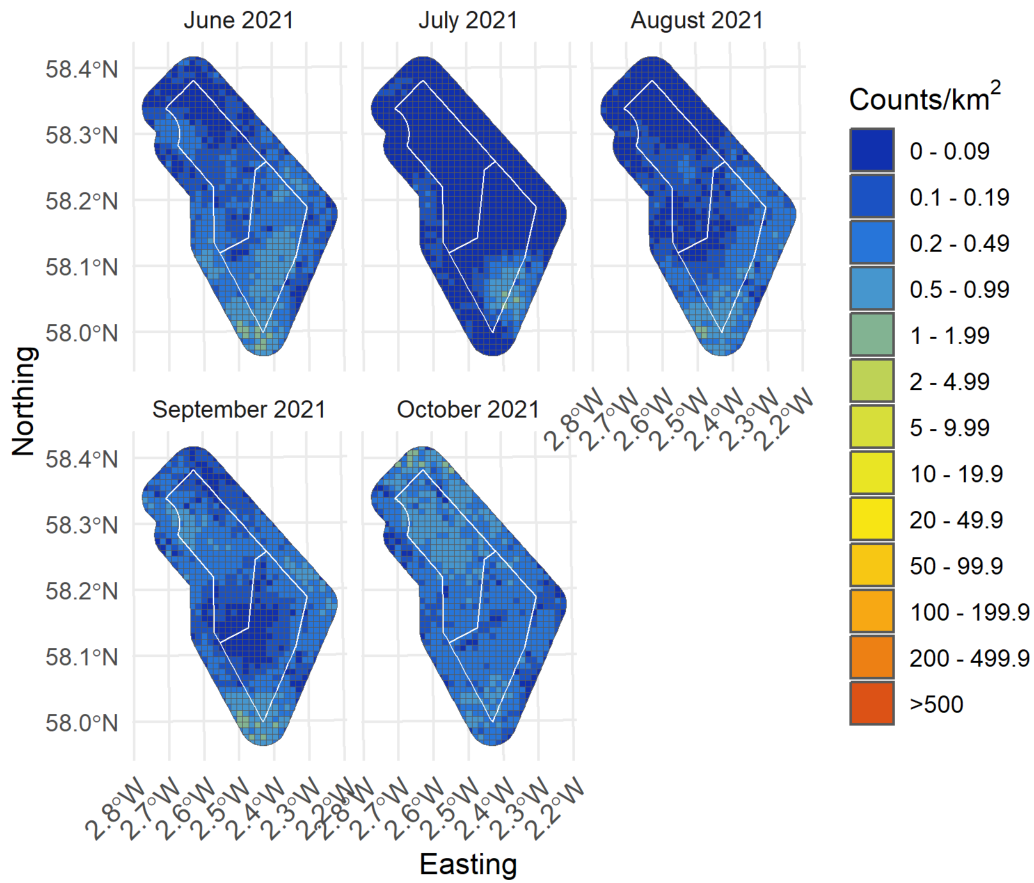


Figure 42. Lower confidence limit of density of all gannets in the survey area for months with sufficient observations between June and October 2021

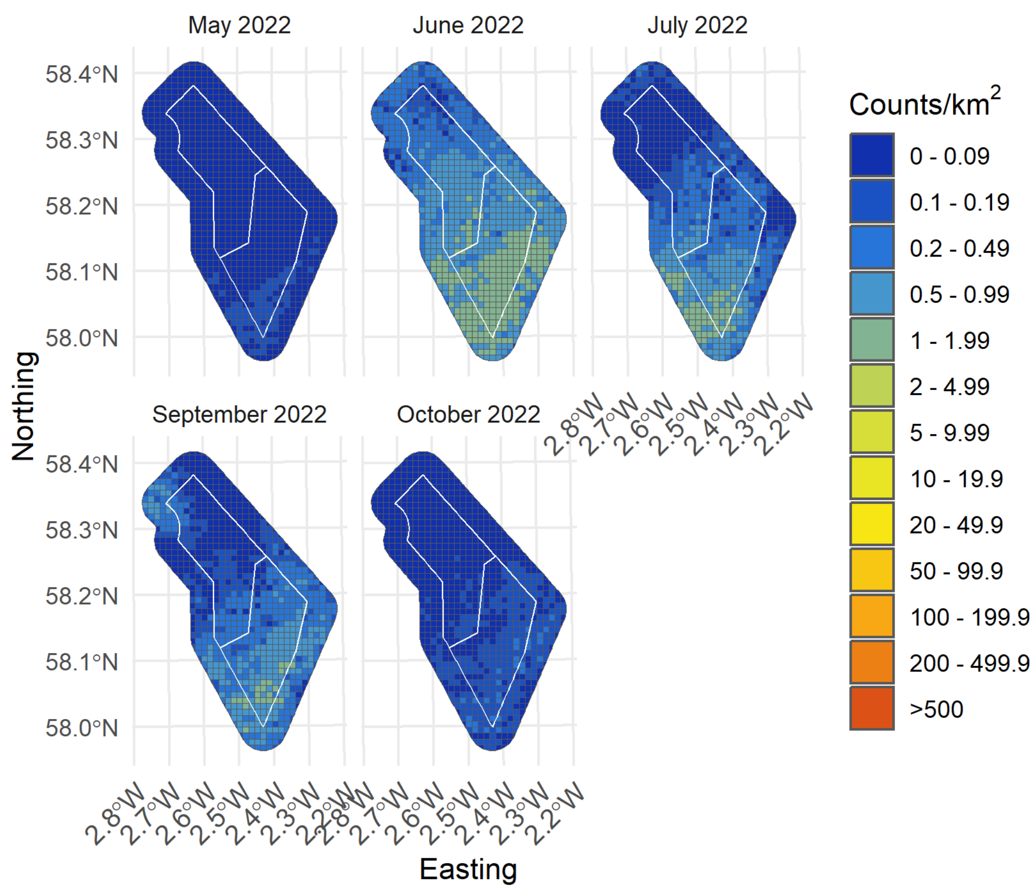


Figure 43. Lower confidence limit of density of all gannets in the survey area for months with sufficient observations between may and October 2022

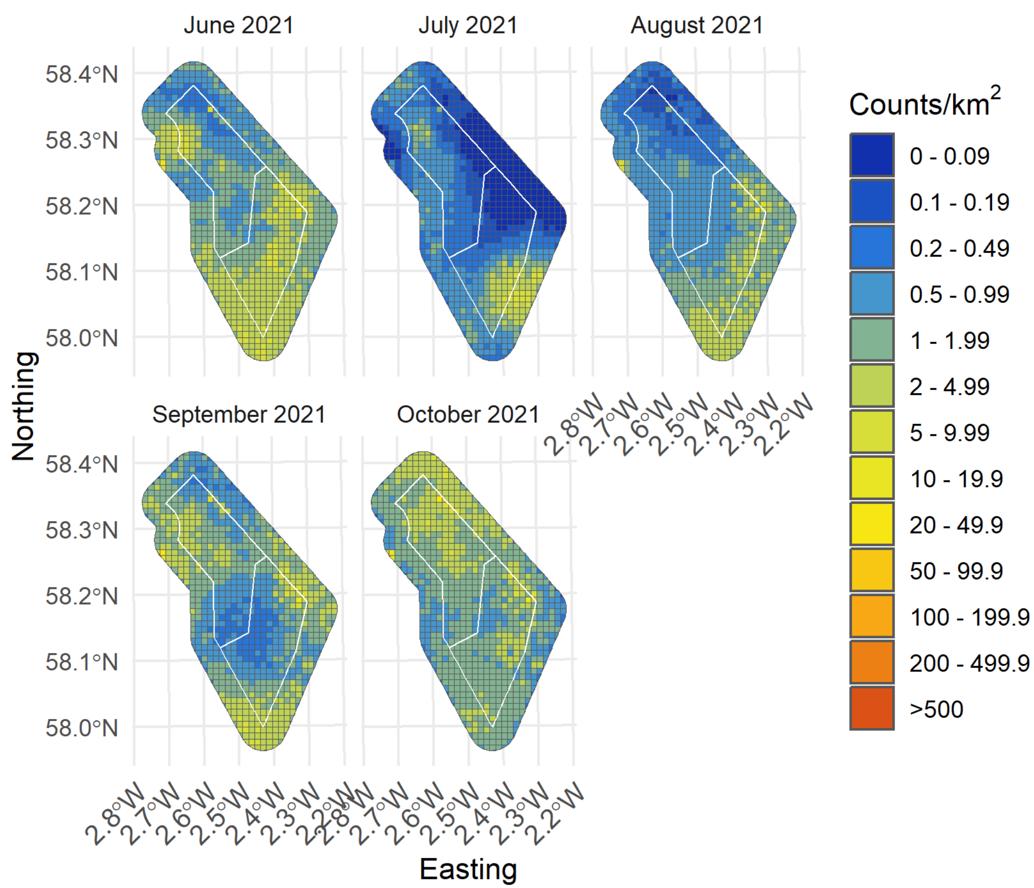


Figure 44. Upper confidence limit of density of all gannets in the survey area for months with sufficient observations between June and October 2021

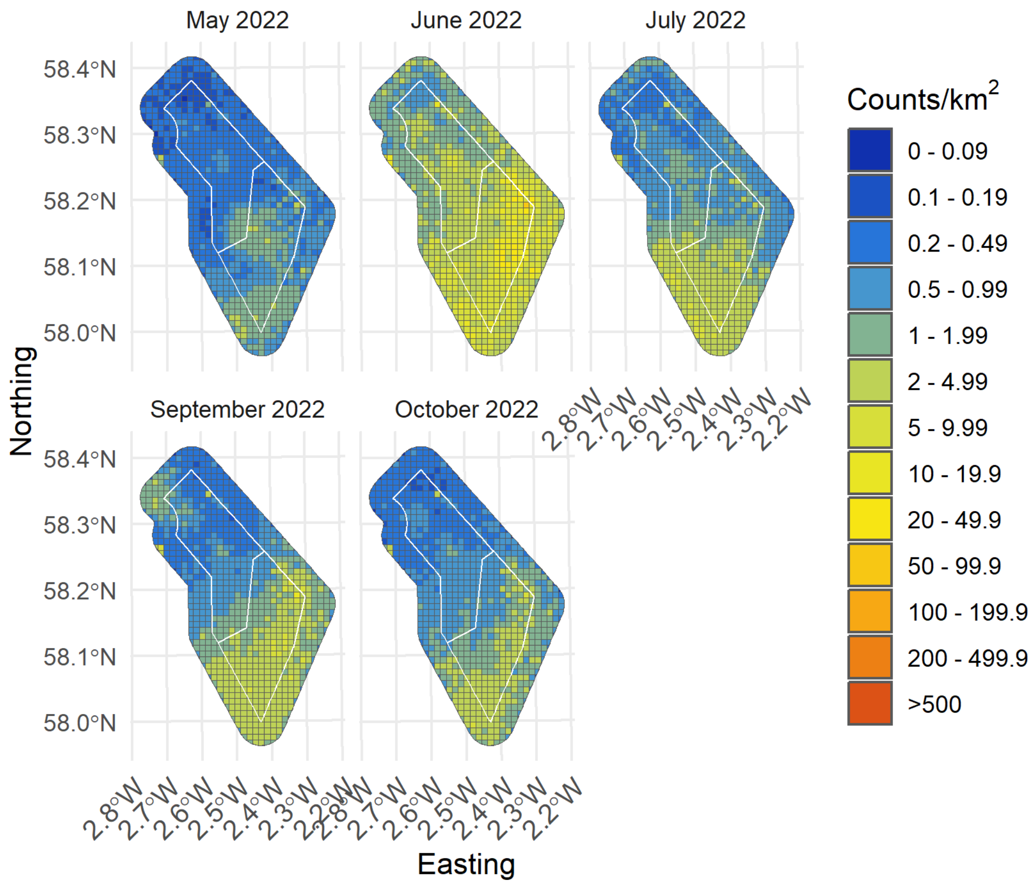


Figure 45. Upper confidence limit of density of all gannets in the survey area for months with sufficient observations between May and October 2022

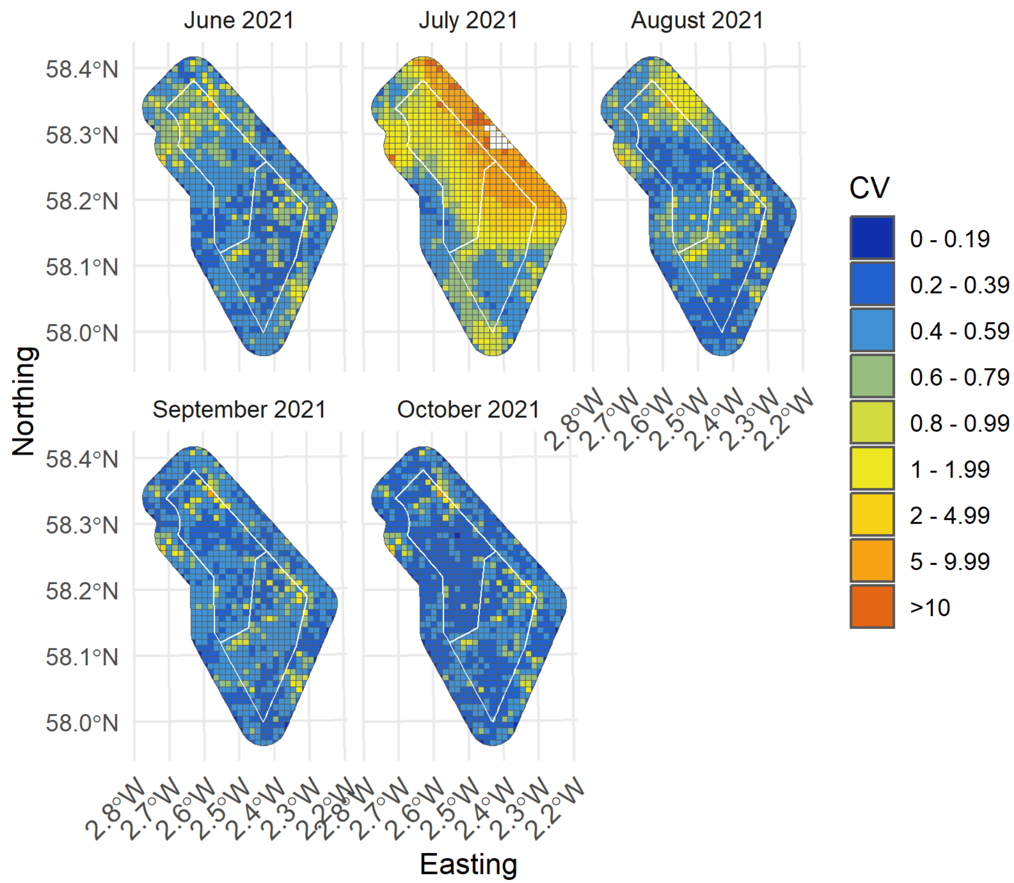


Figure 46. Spatial coefficient of variation of predicted densities of all gannets from MRSea across the survey area for months with sufficient observations between June and October 2021

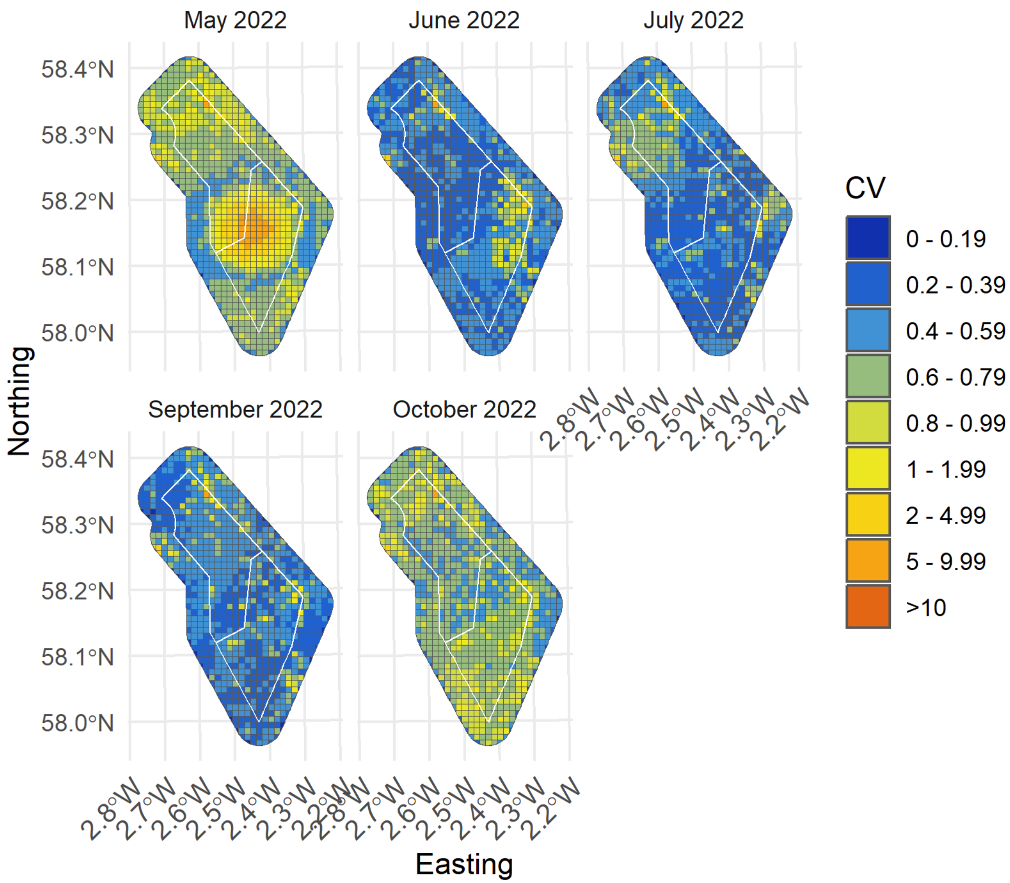


Figure 47. Spatial coefficient of variation of predicted densities of all gannets from MRSea across the survey area for months with sufficient observations between May and October 2022

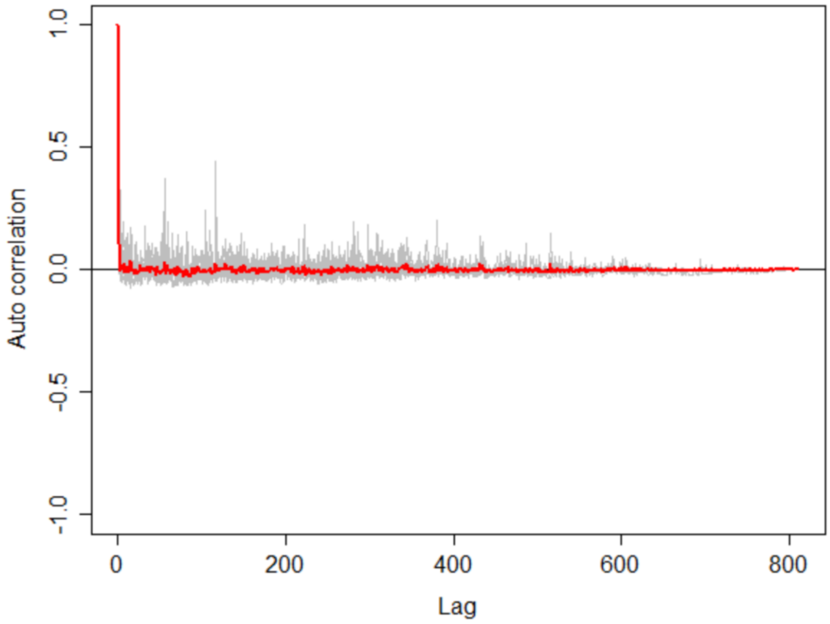


Figure 48. Autocorrelation test for gannet density surface models when using transect as a blocking feature in MRSea showing no significant correlation. A Runs test on the data prior to using transect as a blocking feature gave a p-value of << 0.0001 (i.e., that the data were significantly autocorrelated when not using a blocking feature)

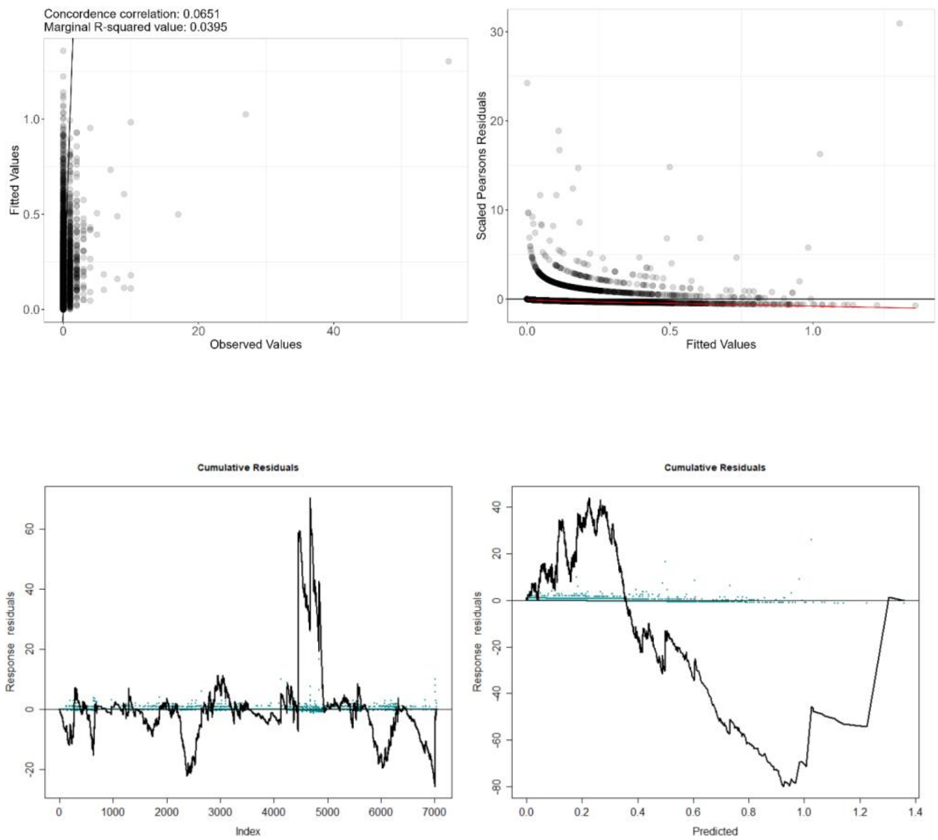


Figure 49. Fitted (MRSea predictions) versus observed counts of all gannets

Table 8. ANOVA results from the best MRSea model for all gannets as selected by cross-validation

Variable	Degrees of Freedom	Chi-square	P-value
Aspect	5	2.60	0.762
Slope	5	4.07	0.539
Bathymetry	5	2.83	0.726
Sandeel Presence	5	3.79	0.58
Distance to Colony	5	17.66	0.003
Turbine Distance	5	6.03	0.303
Sea Surface Temperature (SD)	5	23.91	<<0.001
X/Y (location)	10	84.62	<<0.001

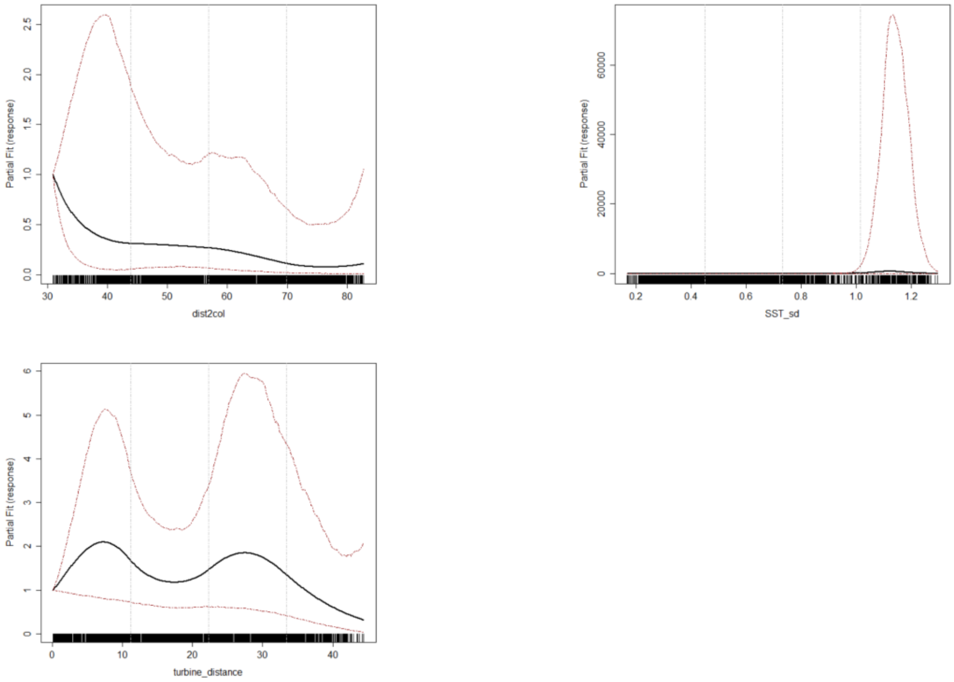


Figure 50. Partial dependence plots for significant variables for all gannets from MRSea. Note that distance to turbine was not a significant variable but was included to demonstrate the relationship. (Clockwise from top left: distance to colony, daily standard deviation of sea surface temperature, distance to turbine)

3.2.2 Flying Birds Model

Table 9. Candidate and final covariates for flying gannets model.

Starting model covariates after VIF-based collinearity removal	Final model covariates after removal by SALSA
Aspect	Distance to turbine
Slope	Daily mean of Sea Surface Temperature
Bathymetry	
Sandeel Presence	
Distance to colony	
Distance to turbine	
Daily mean of Sea Surface Temperature	

Distribution maps generated using MRSea (Figure 50-51) suggest that flying gannets are distributed mostly across the southern half of the study area during the breeding season (15 March – 30 September). In the non-breeding season, gannets were largely absent from the study area. The highest densities were observed in the southern third of the study area in June 2022.

Model fit was poor with a marginal R squared value of 0.01 and root mean squared error of 0.60. Cumulative residuals in the model showed that there was a moderate relationship between predicted and observed values across most of the range of predicted values, except between ~1.3 and 0.2 birds/km^2, where the model tended to underpredict density (Figure 59).

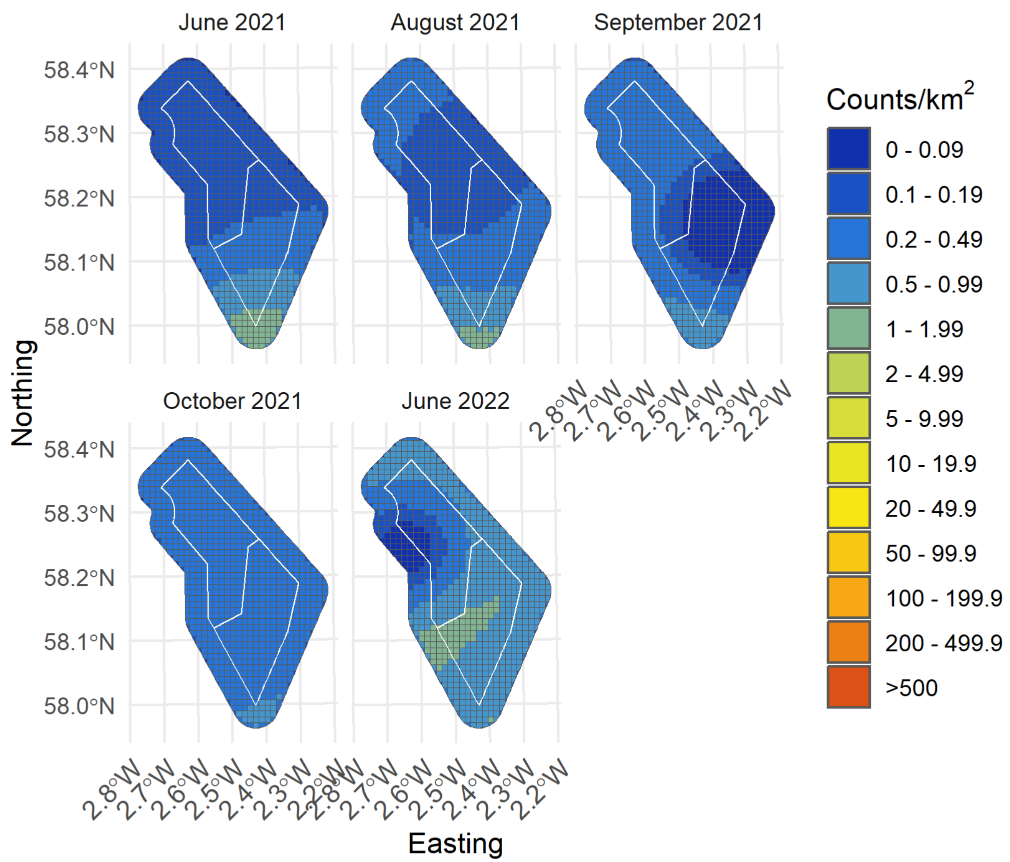


Figure 51. Median density of flying gannets in the survey area for months with sufficient observations between June 2021 and June 2022

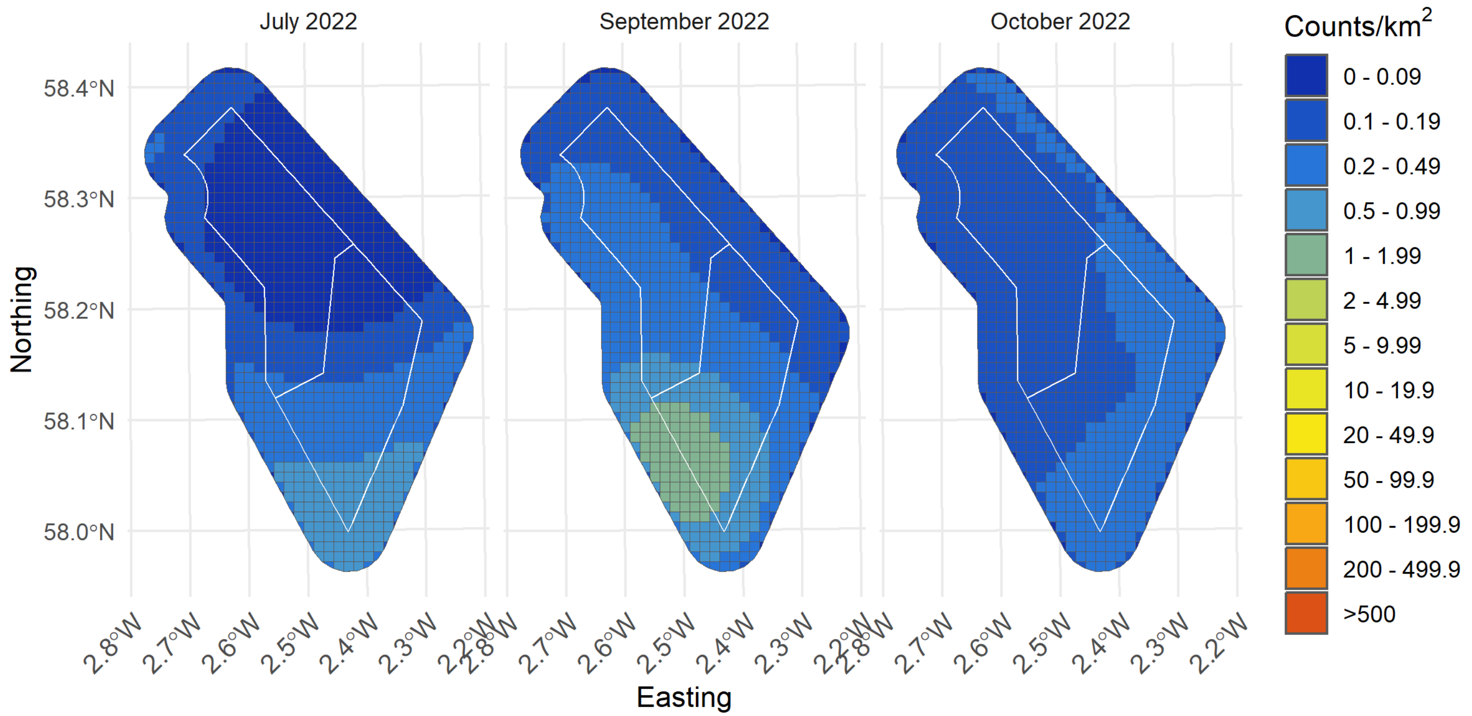


Figure 52. Median density of flying gannets in the survey area for months with sufficient observations between July and October 2022

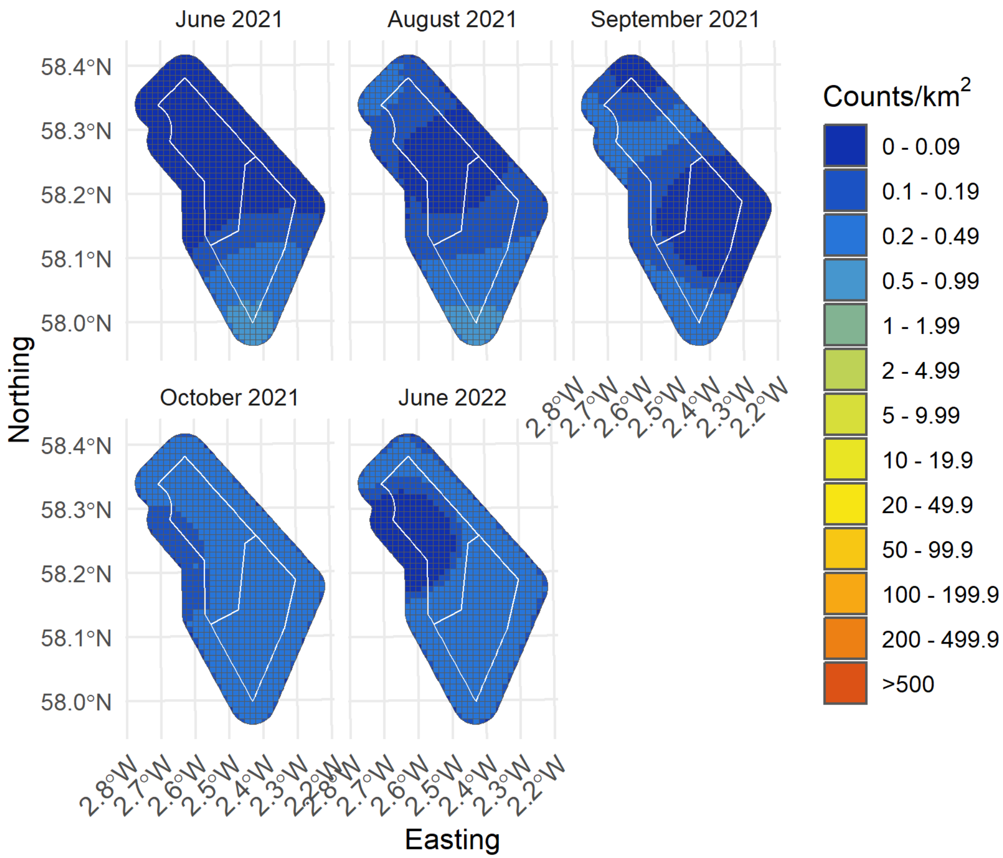


Figure 53. Lower confidence limit of density of flying gannets in the survey area for months with sufficient observations between June 2021 and June 2022

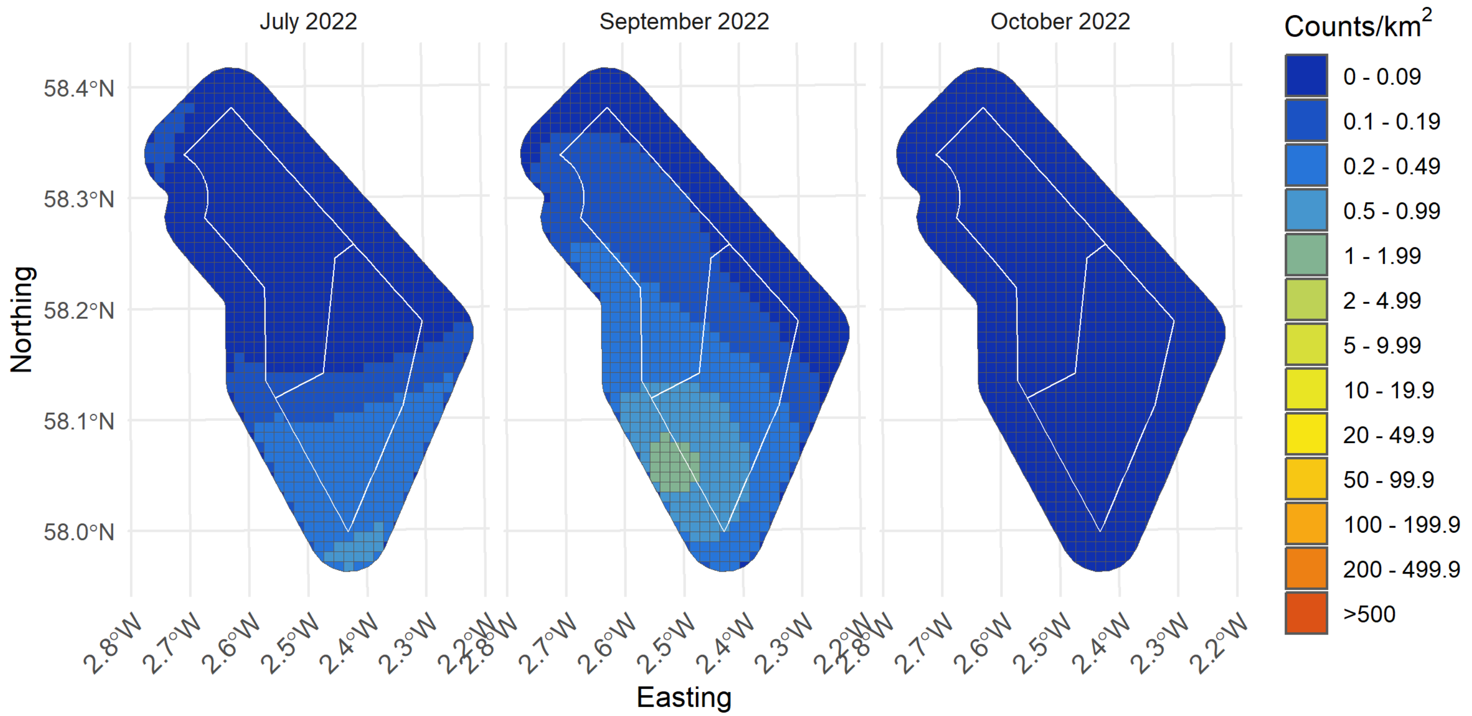


Figure 54. Lower confidence limit of density of flying gannets in the survey area for months with sufficient observations between July and October 2022

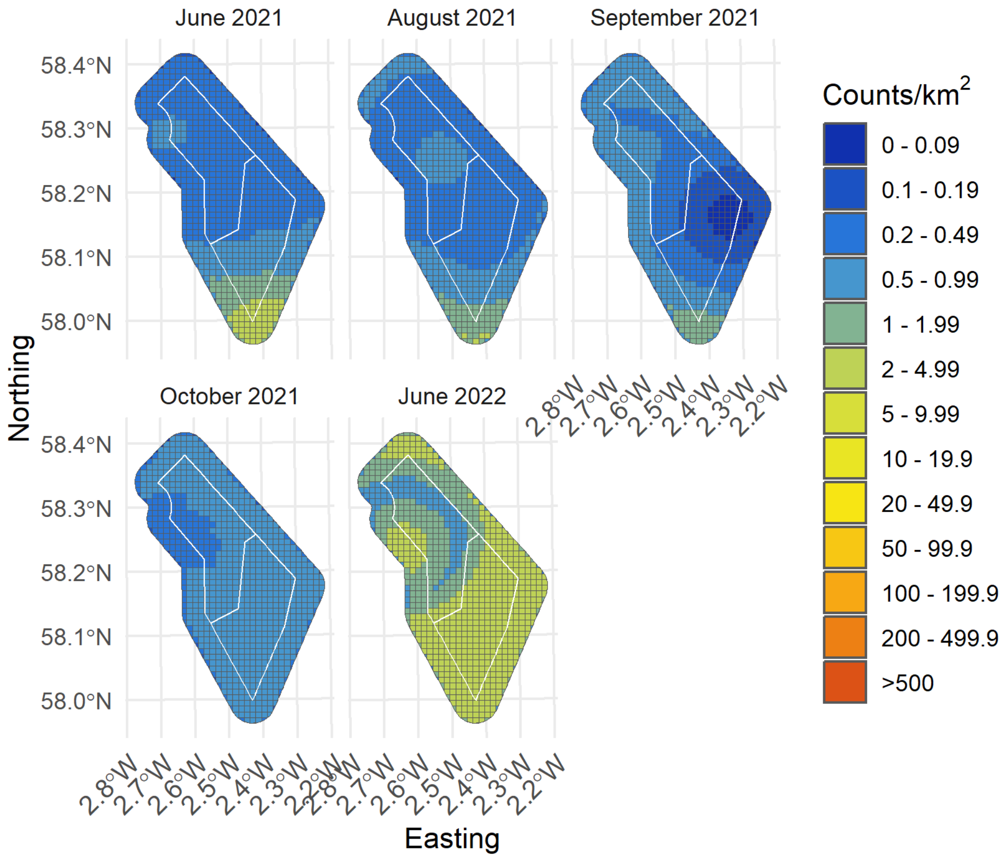


Figure 55. Upper confidence limit of density of flying gannets in the survey area for months with sufficient observations between June 2021 and June 2022

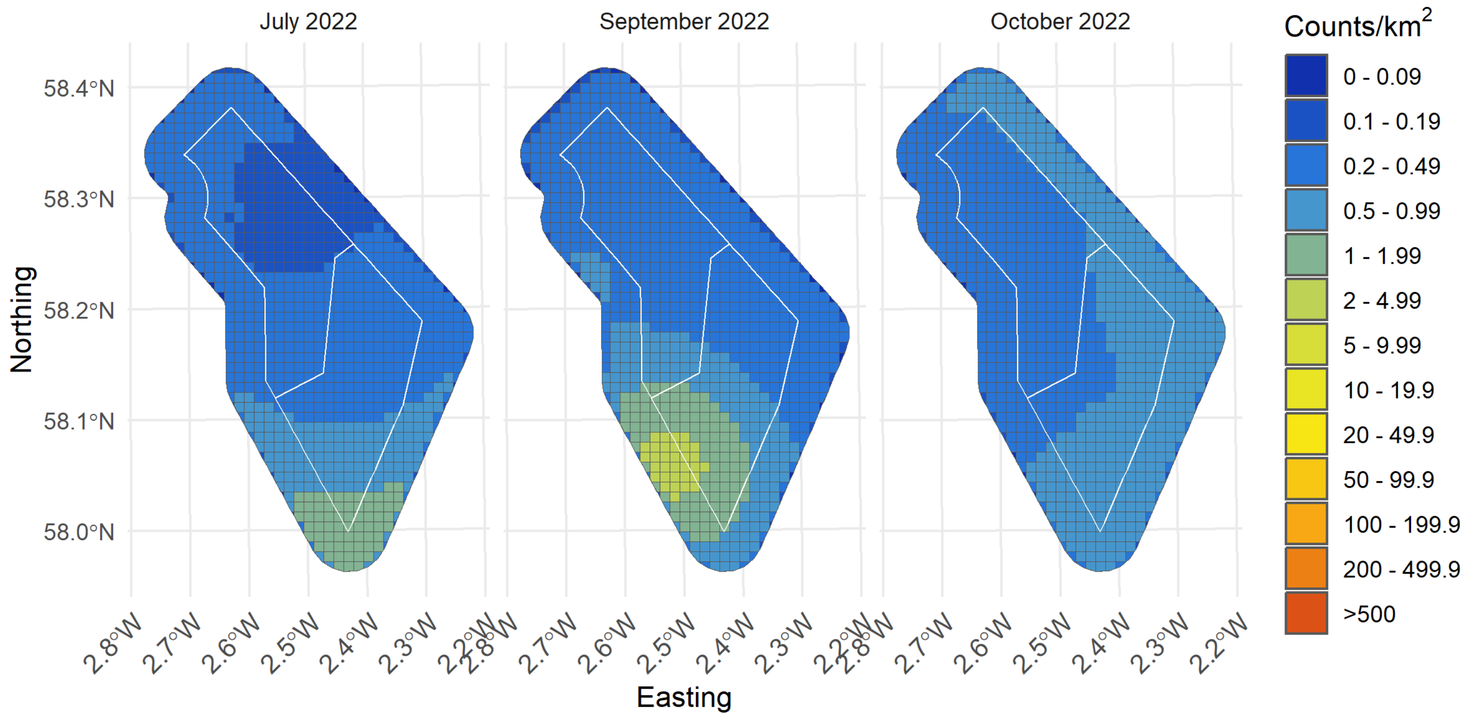


Figure 56. Upper confidence limit of density of flying gannets in the survey area for months with sufficient observations between July and October 2022

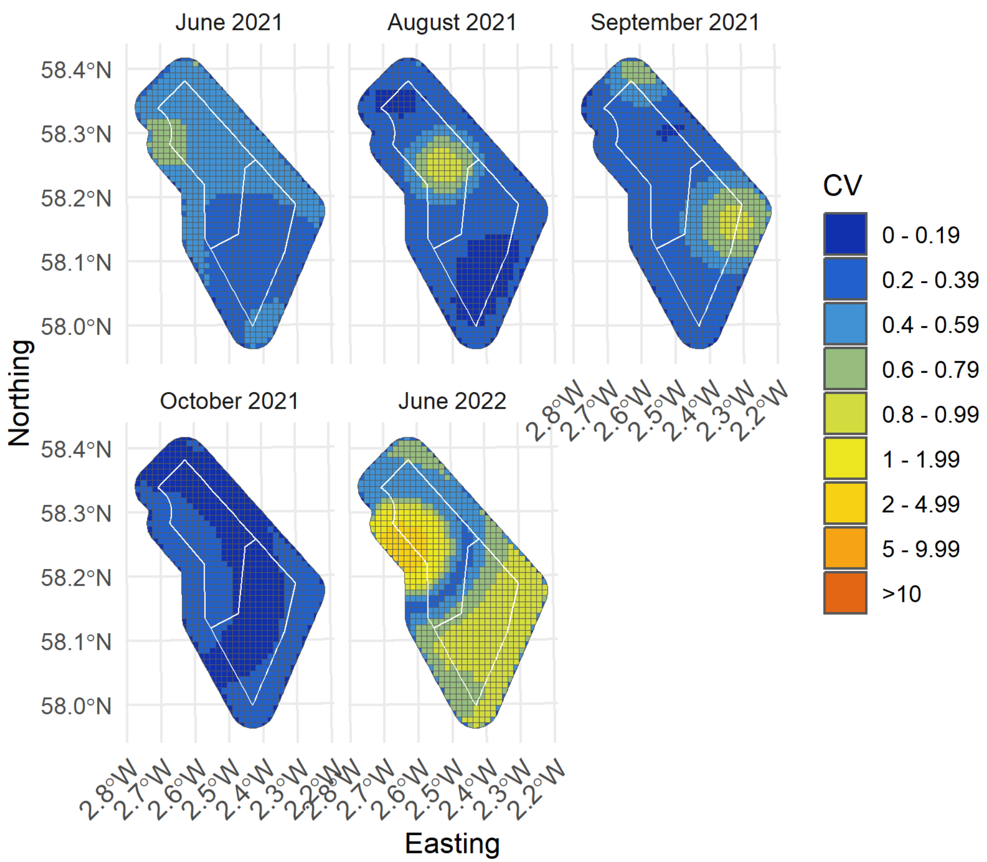


Figure 57. Spatial coefficient of variation of predicted densities of flying gannets from MRSea across the survey area for months with sufficient observations between June 2021 and June 2022

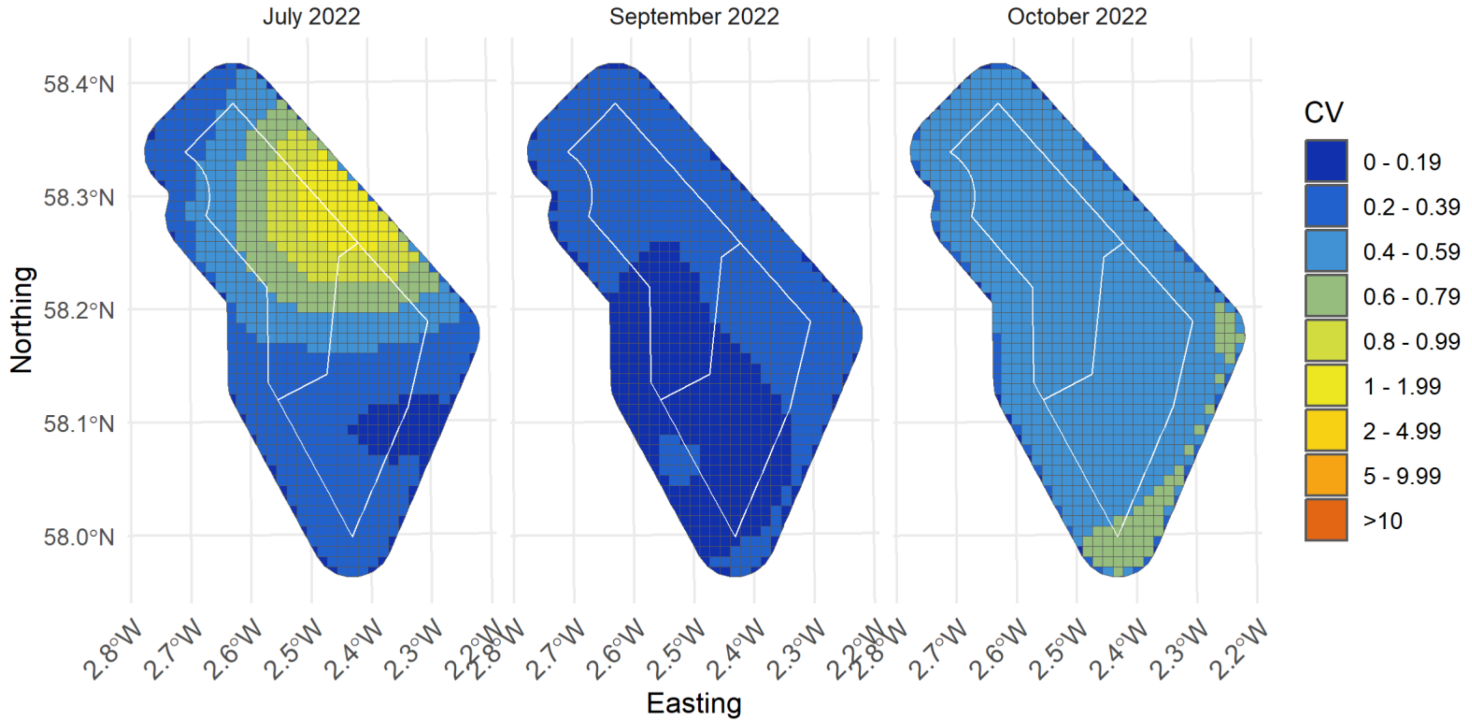


Figure 58. Spatial coefficient of variation of predicted densities of flying gannets from MRSea across the survey area for months with sufficient observations between July and October 2022

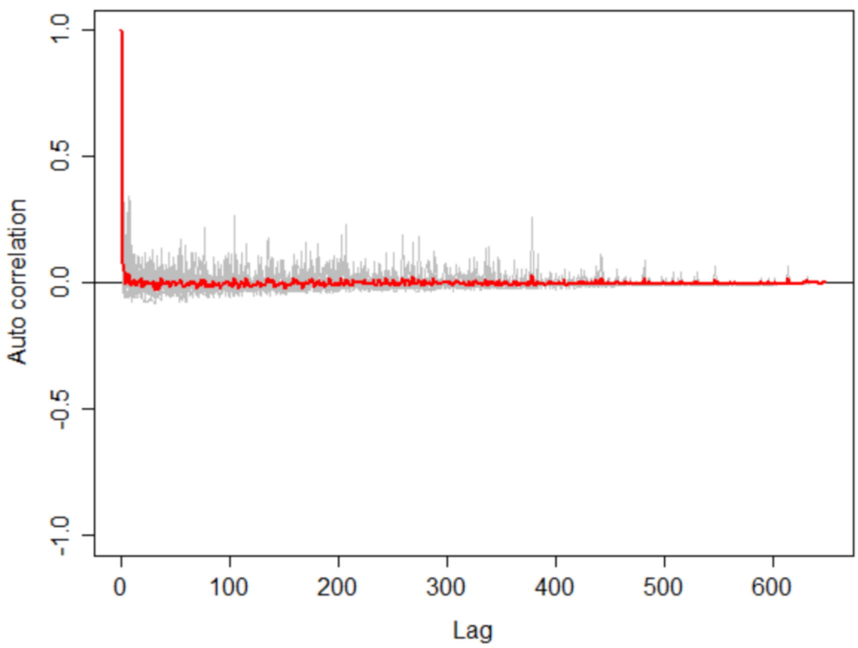


Figure 59. Autocorrelation test for gannet density surface models when using transect as a blocking feature in MRSea showing no significant correlation. A Runs test on the data prior to using transect as a blocking feature gave a p-value of << 0.0001 (i.e., that the data were significantly autocorrelated when not using a blocking feature)

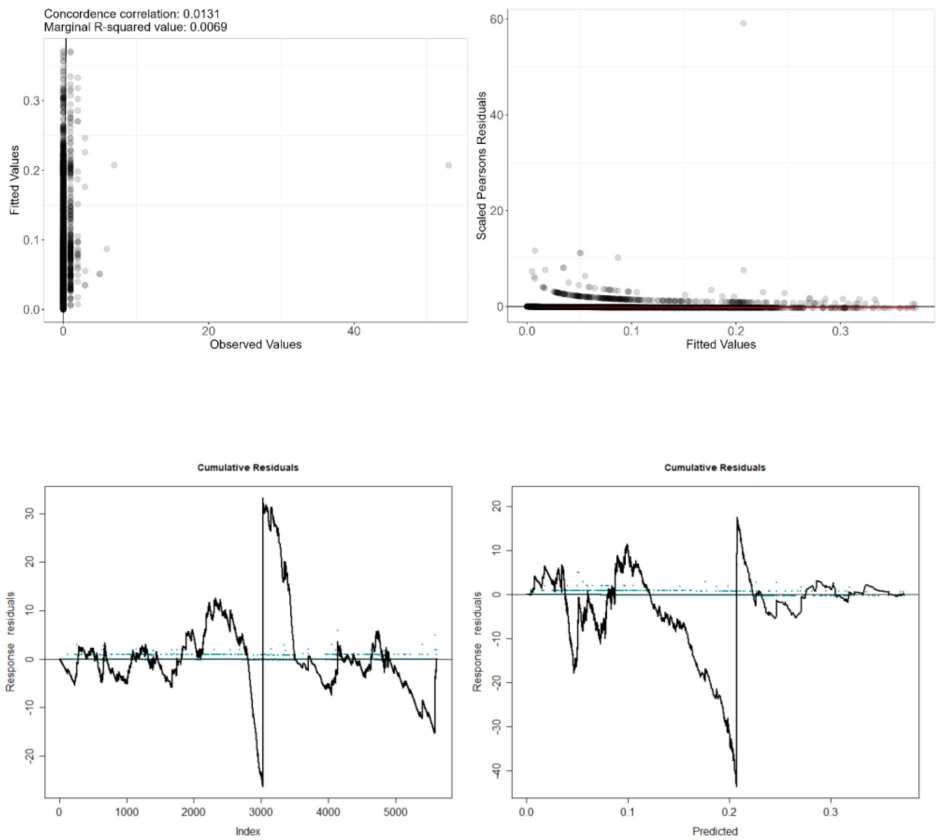


Figure 60. Fitted (MRSea predictions) versus observed counts of flying gannets

Table 10. ANOVA results from the best MRSea model for flying gannets as selected by cross-validation

Variable	Degrees of Freedom	Chi-square	P-value
Turbine Distance	1	1.86	0.173
Sea Surface Temperature	1	0.66	0.416
X/Y (location)	10	52.64	<<0.001

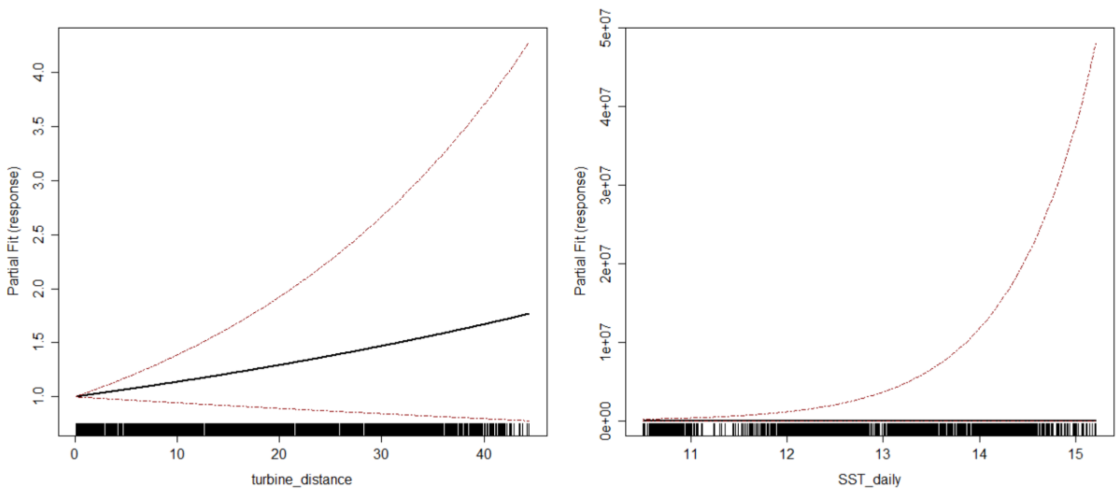


Figure 61. Partial dependence plots for significant variables for flying gannets from MRSea. Note that distance to turbine was not a significant variable, but was included to demonstrate the relationship. (left: distance to turbine; right: daily mean of sea surface temperature)

3.3 Fulmar

3.3.1 All Birds Model

Table 11. Candidate and final covariate terms for all fulmar model.

Starting model covariates after VIF-based collinearity removal	Final model covariates after removal by SALSA
Bathymetry	Bathymetry
Sandeel Density	Distance to colony
Distance to colony	Distance to turbine
Distance to turbine	

Distribution maps generated using MRSea (Figure 61-65) suggest that all fulmar are distributed in higher densities in the southern and western areas of the study area during the breeding season (01 April - 15 September). In the non-breeding season, fulmar tend to be distributed in higher densities along the south edge of the study area. The highest densities were observed around the southern tip of the study area during July 2022.

Model fit was moderate with a marginal R squared value of 0.14 and root mean squared error of 5.20. Cumulative residuals in the model showed that there was a poor relationship between predicted and observed values across most of the range of predicted values, and tended to underpredict bird density (Figure 82).

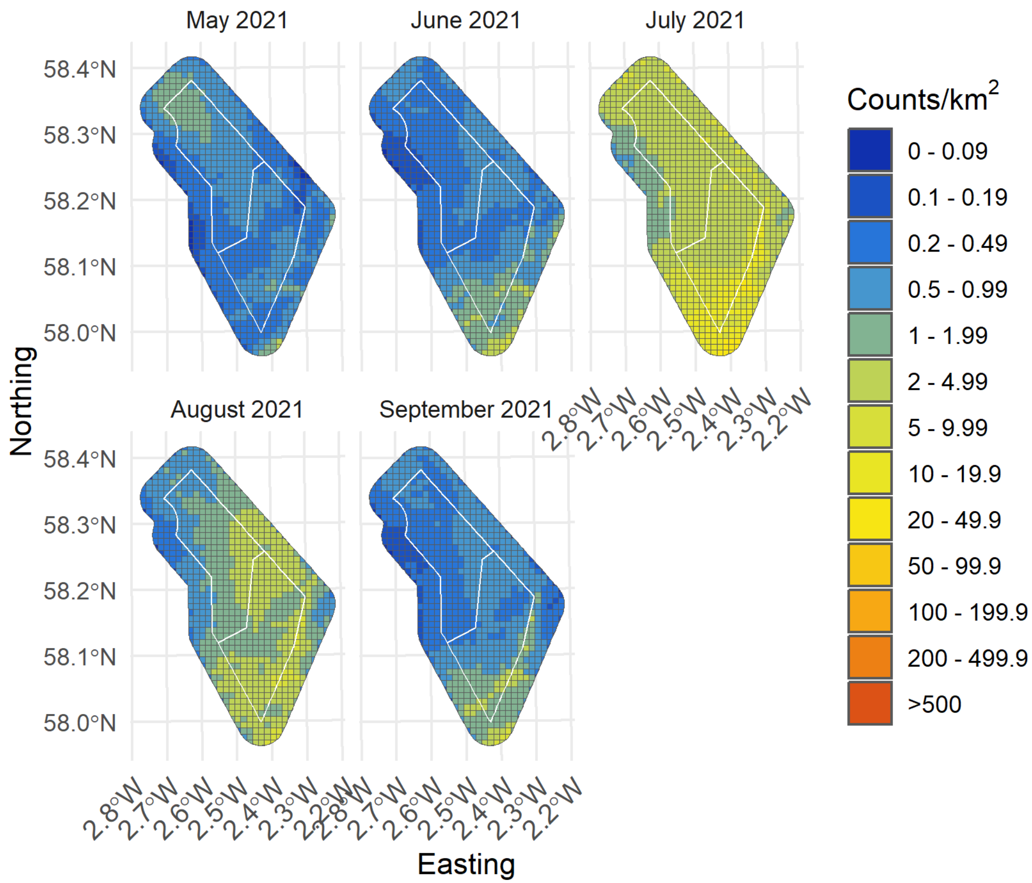


Figure 62. Median density of all fulmar in the survey area for months with sufficient observations between May and September 2021

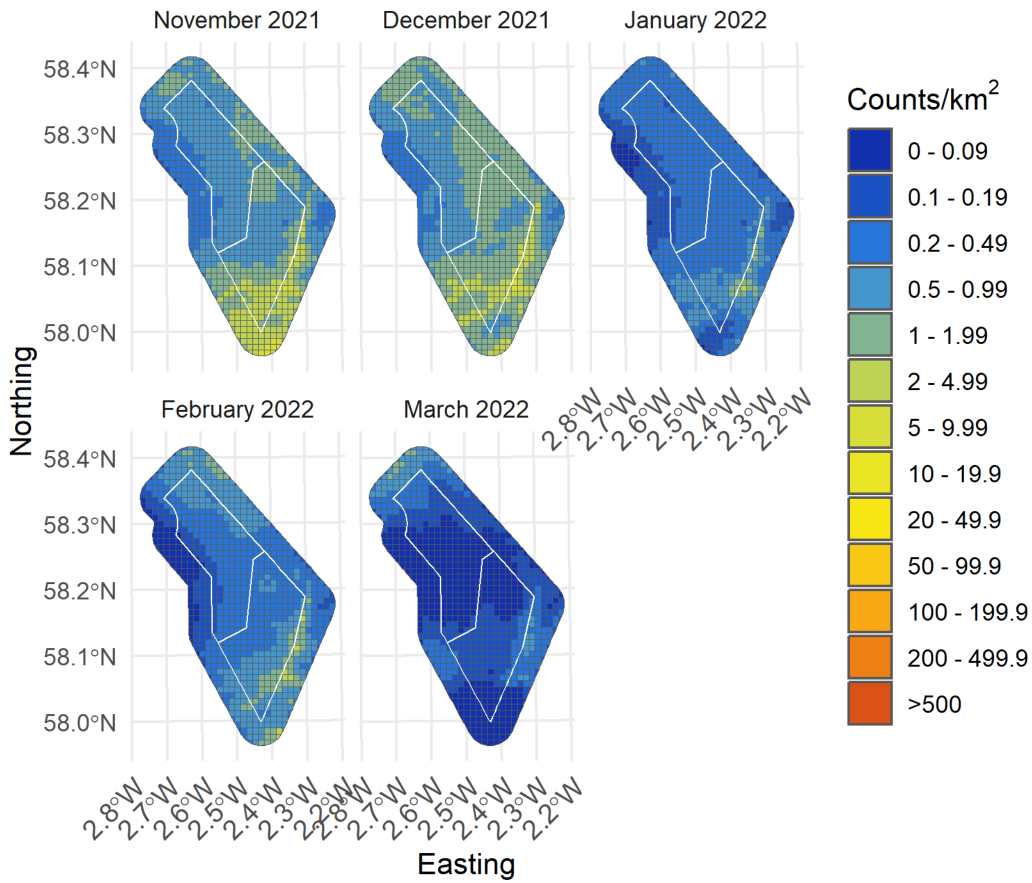


Figure 63. Median density of all fulmar in the survey area for months with sufficient observations between November 2021 and March 2022

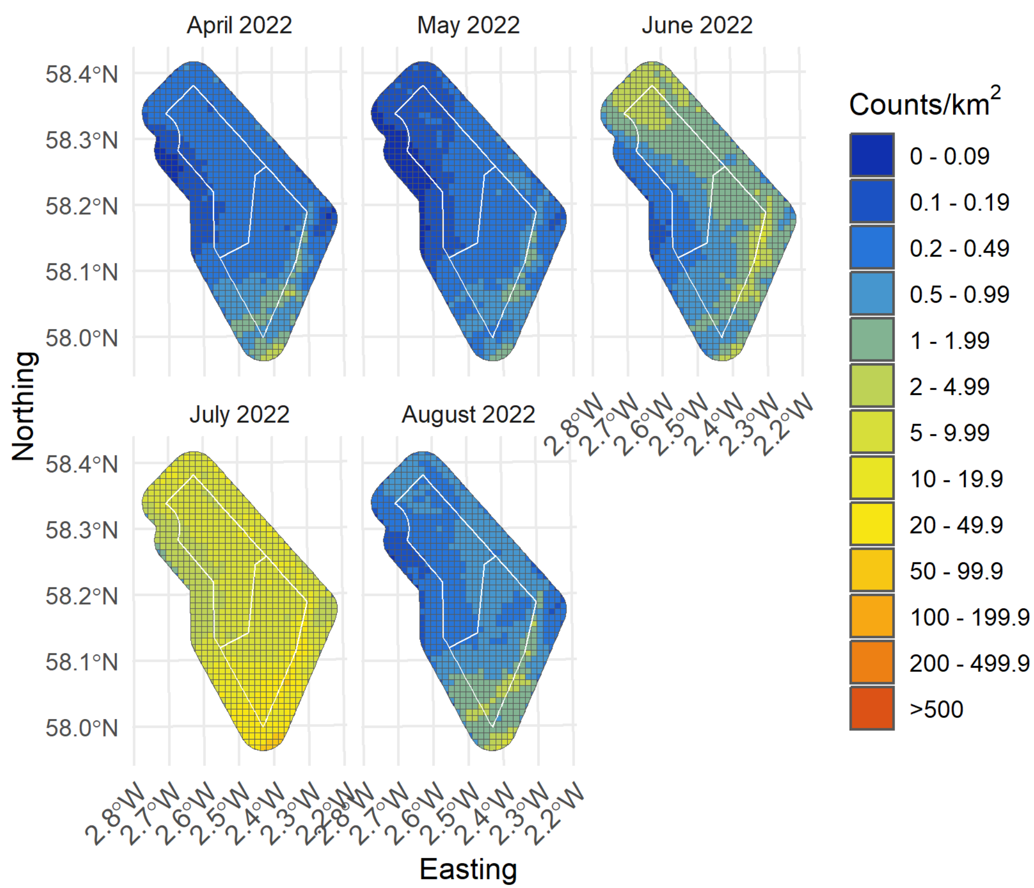


Figure 64. Median density of all fulmar in the survey area for months with sufficient observations between April and August 2022

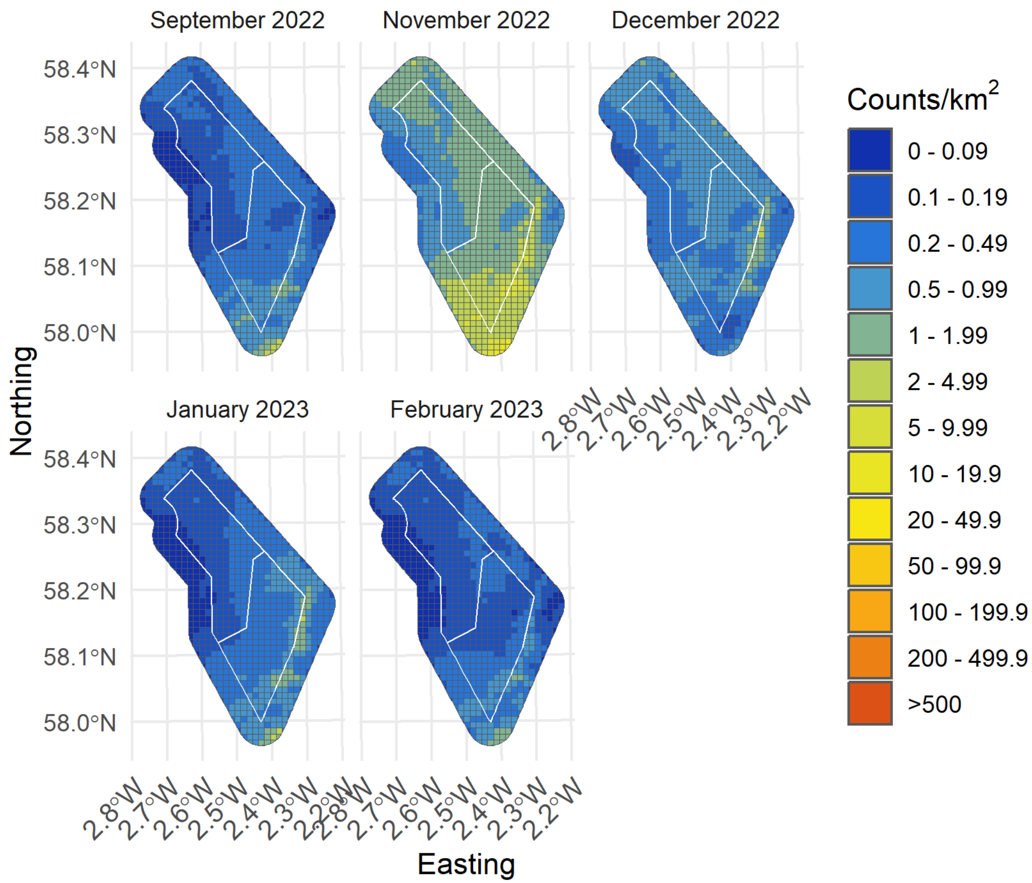


Figure 65. Median density of all fulmar in the survey area for months with sufficient observations between September 2022 and February 2023

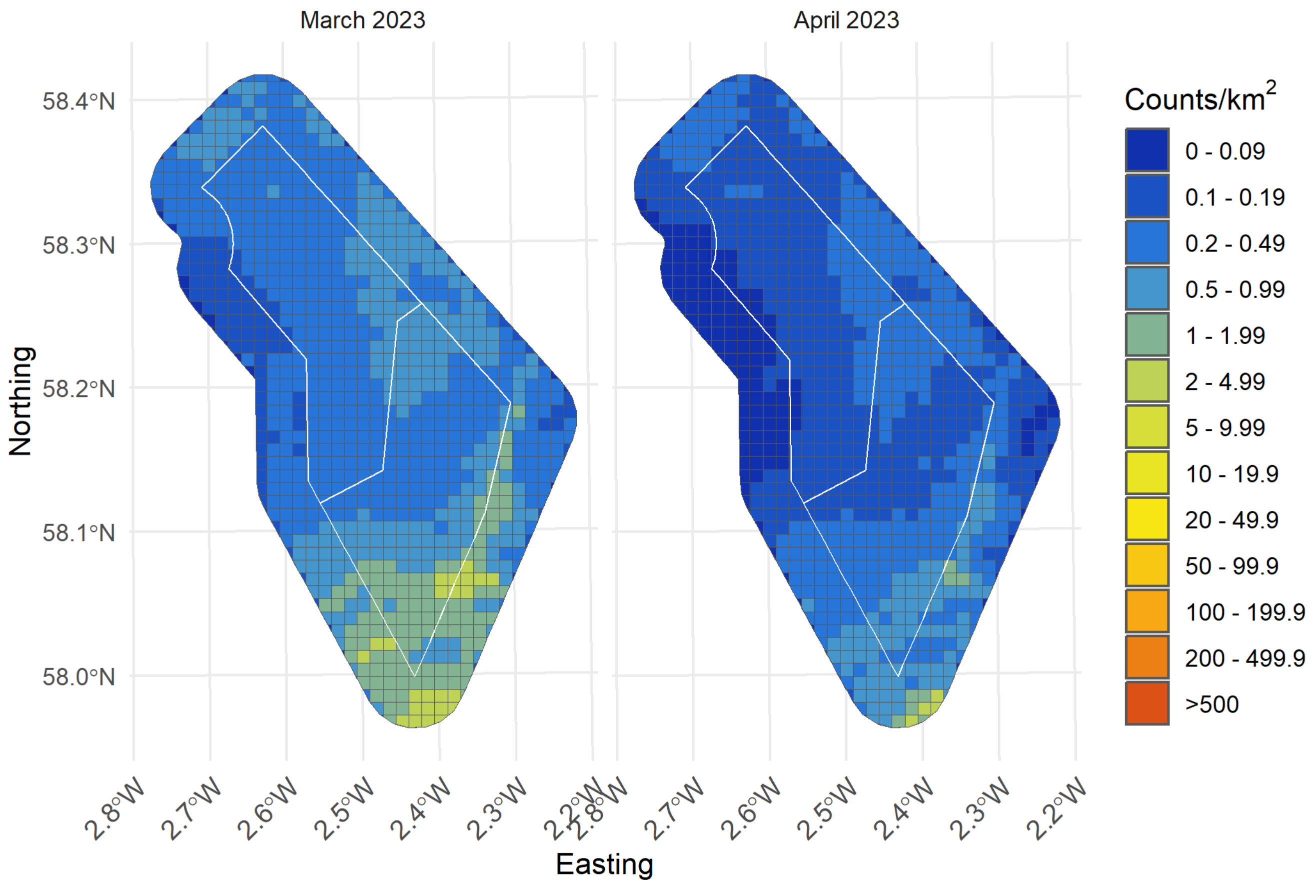


Figure 66. Median density of all fulmar in the survey area for months with sufficient observations between March and April 2023

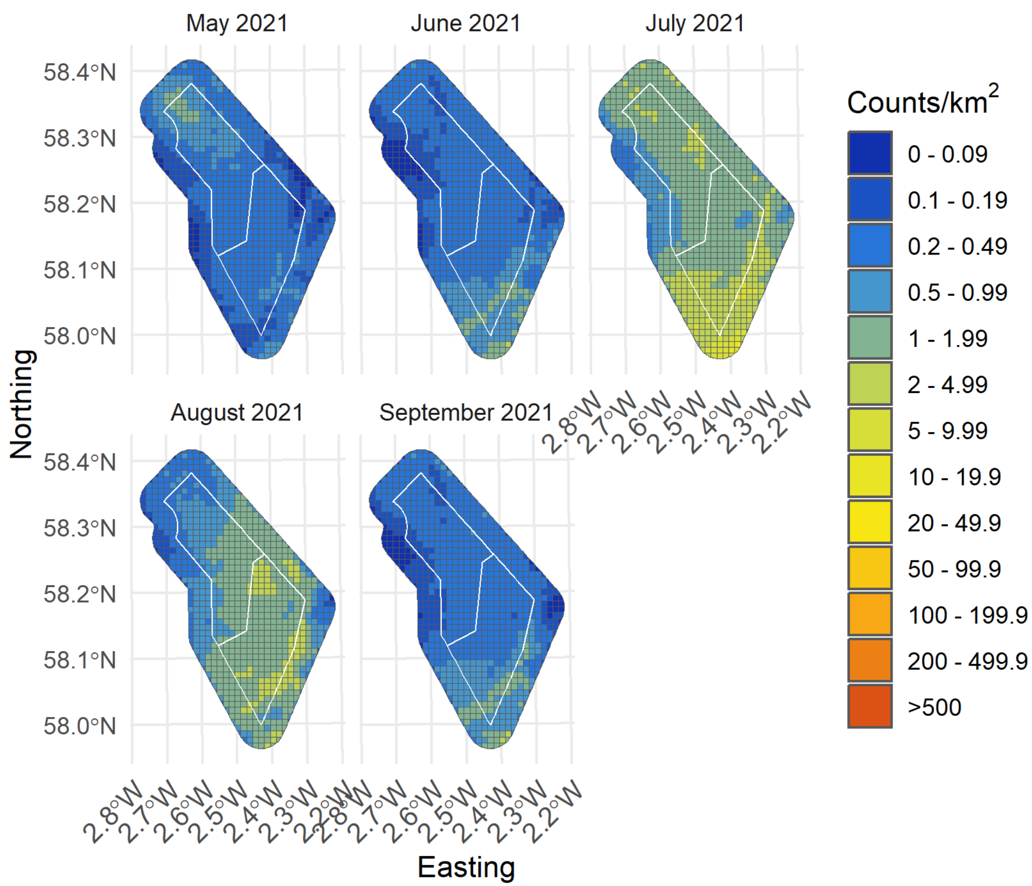


Figure 67. Lower confidence limit of density of all fulmar in the survey area for months with sufficient observations between May and September 2021

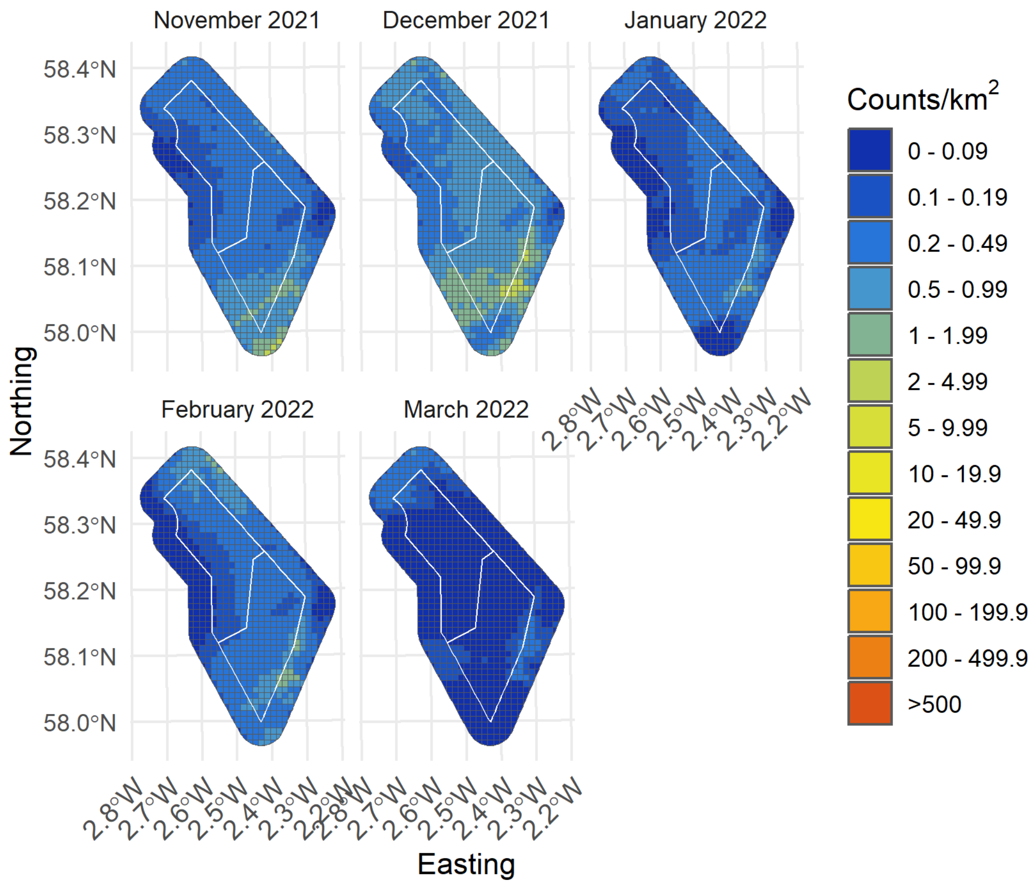


Figure 68. Lower confidence limit of density of all fulmar in the survey area for months with sufficient observations between November 2021 and March 2022

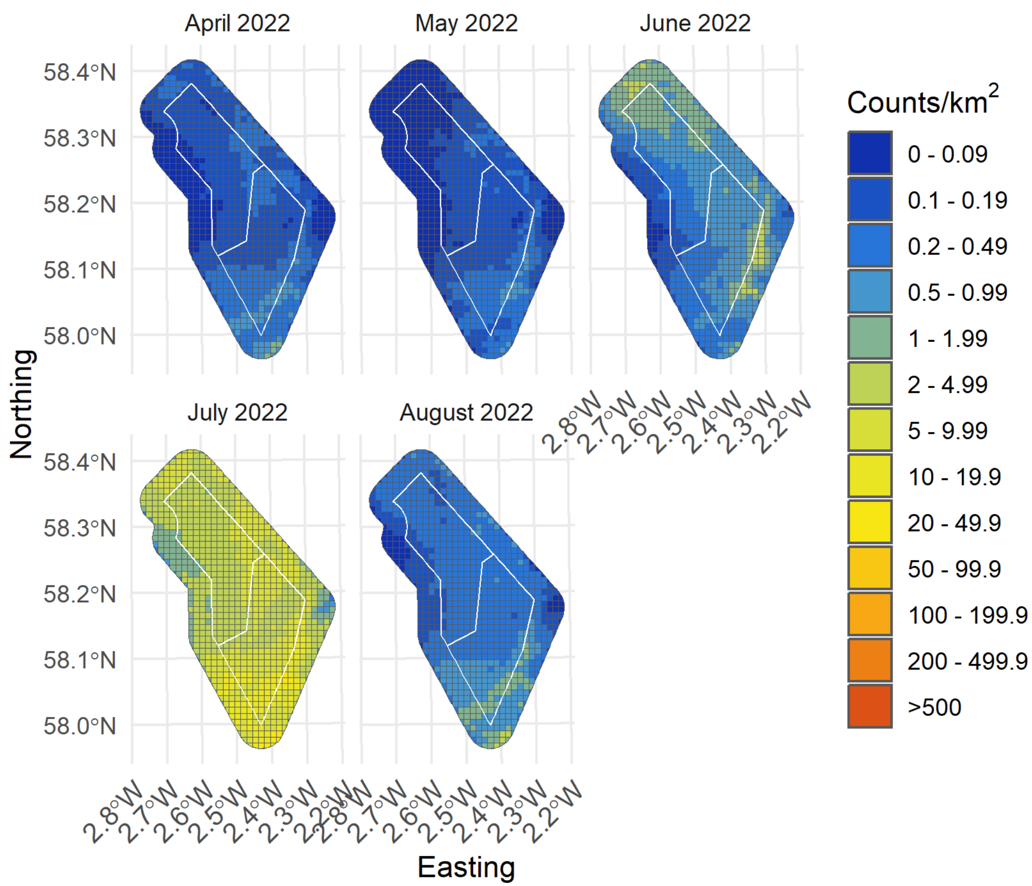


Figure 69. Lower confidence limit of density of all fulmar in the survey area for months with sufficient observations between April and August 2022

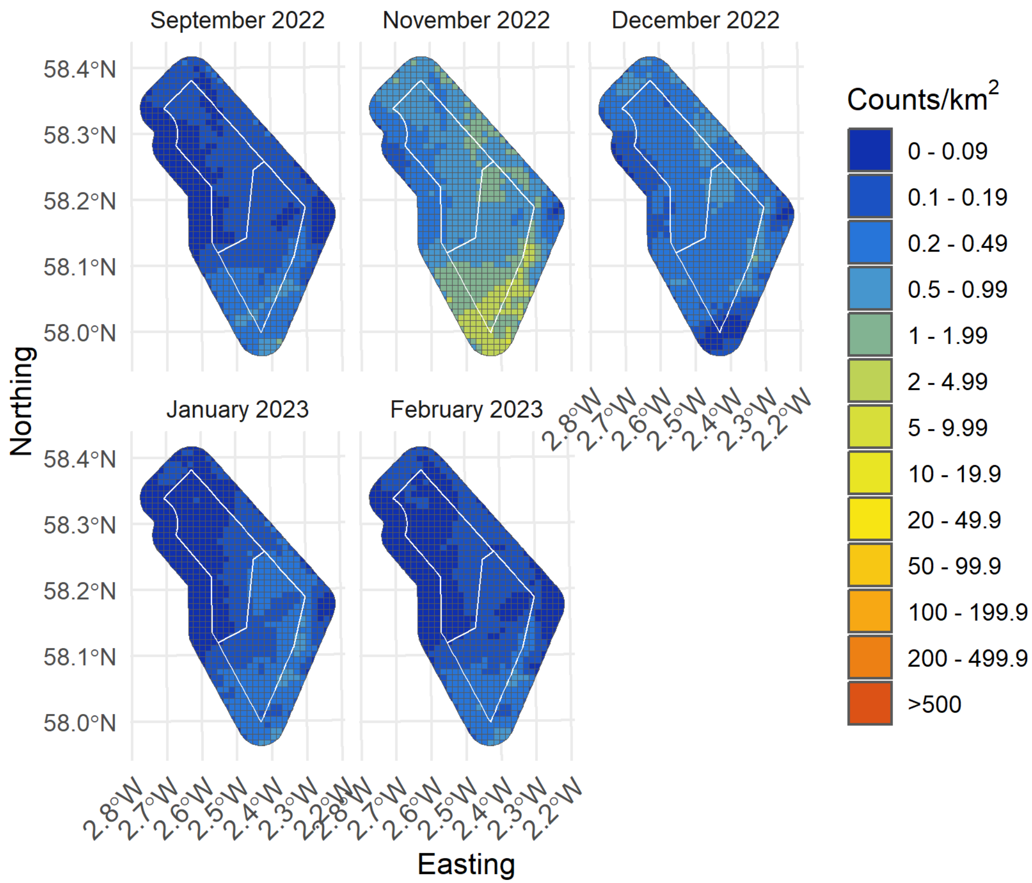


Figure 70. Lower confidence limit of density of all fulmar in the survey area for months with sufficient observations between September 2022 and February 2023

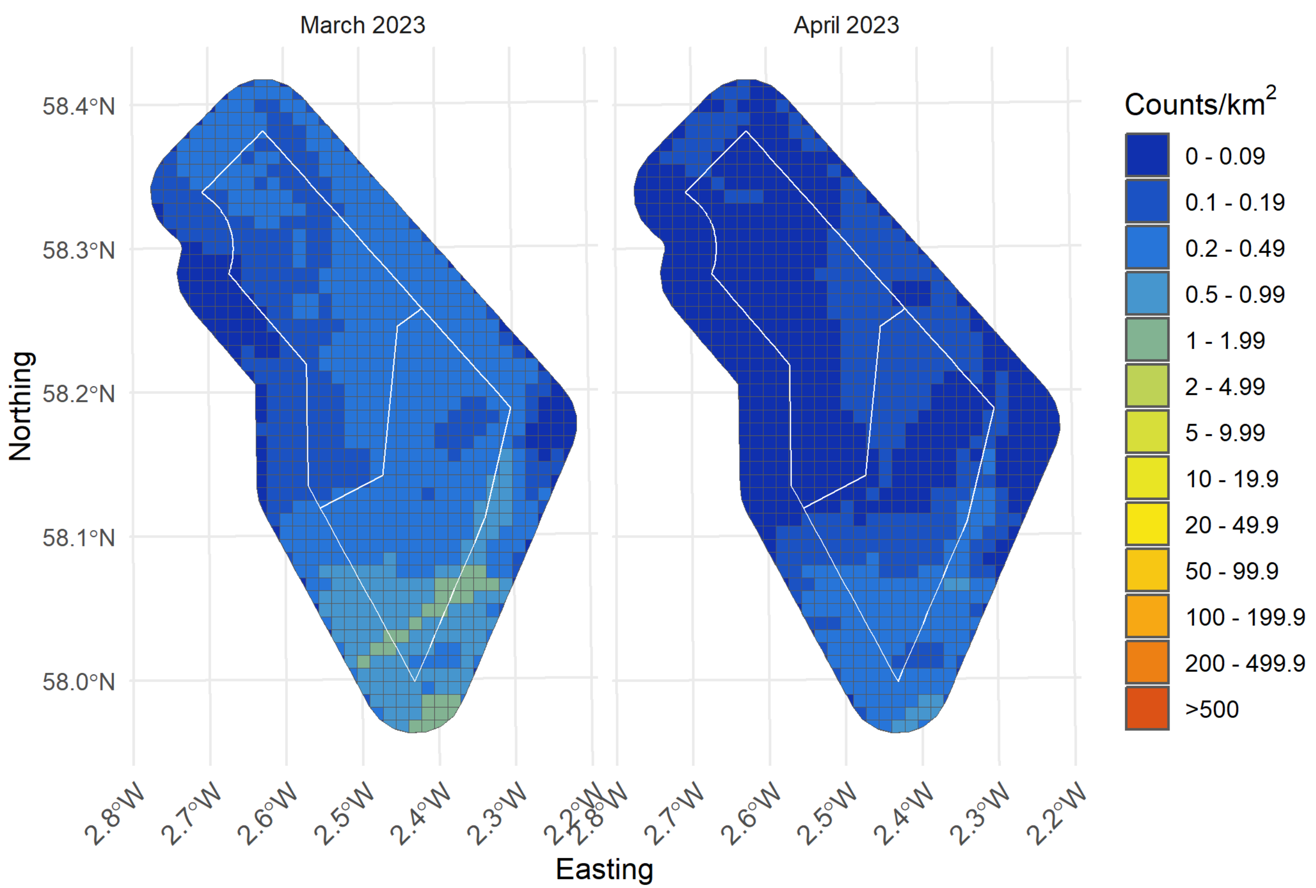


Figure 71. Lower confidence limit of density of all fulmar in the survey area for months with sufficient observations between March and April 2023

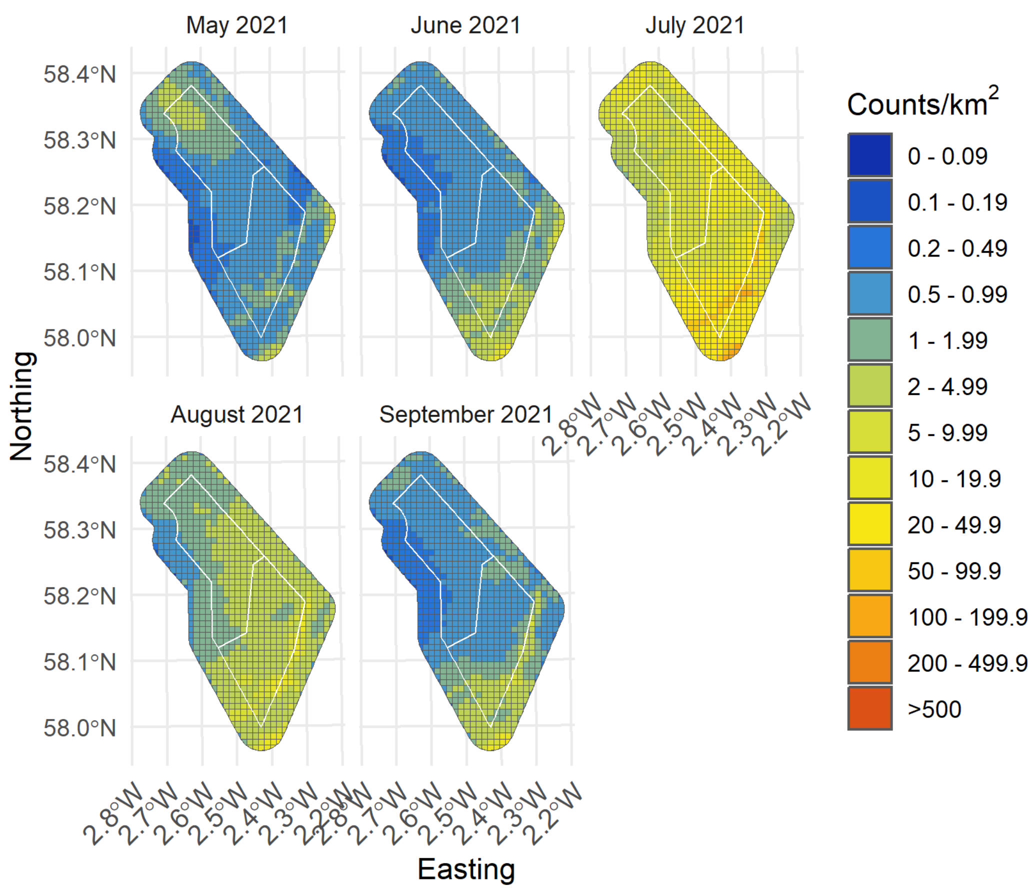


Figure 72. Upper confidence limit of density of all fulmar in the survey area for months with sufficient observations between May and September 2021

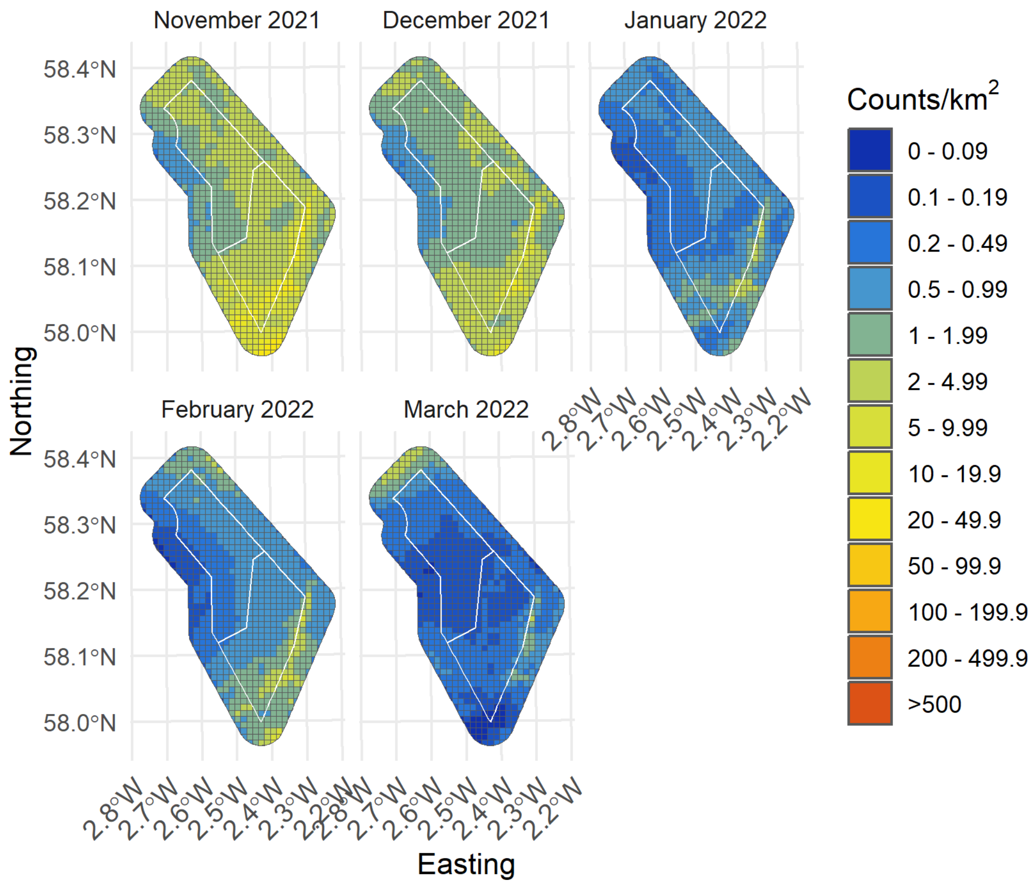


Figure 73. Upper confidence limit of density of all fulmar in the survey area for months with sufficient observations between November 2021 and March 2022

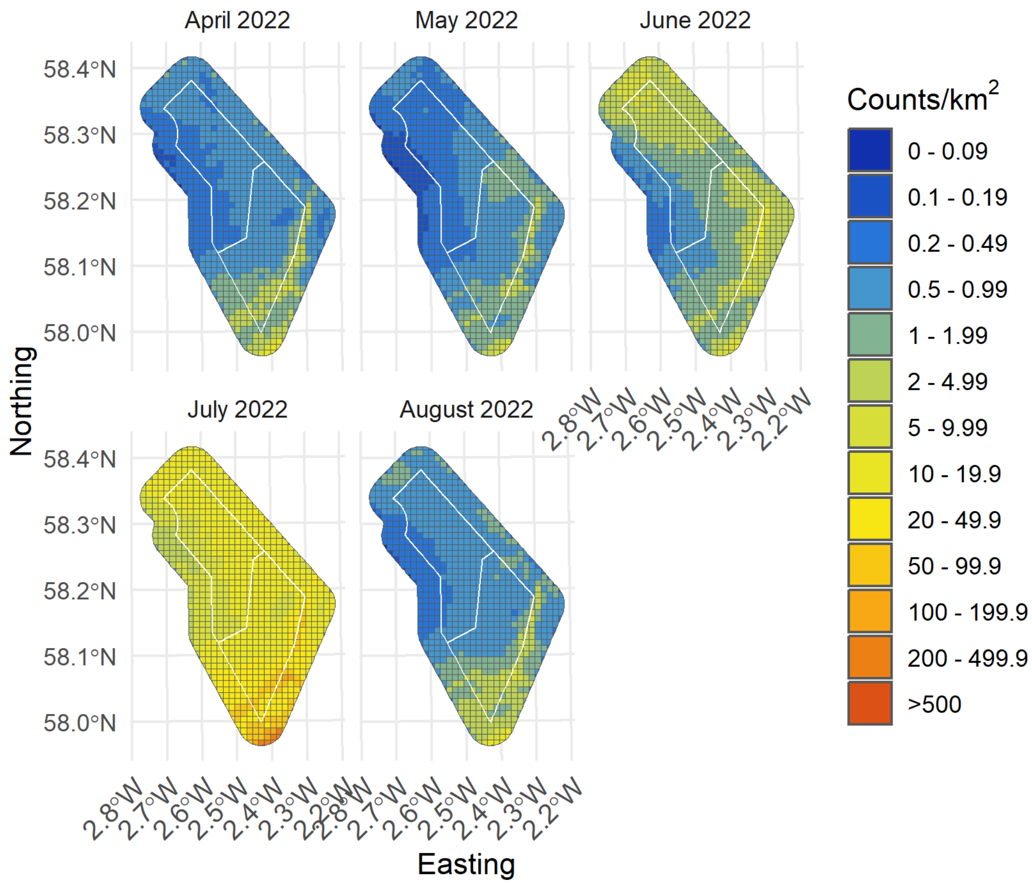


Figure 74. Upper confidence limit of density of all fulmar in the survey area for months with sufficient observations between April and August 2022

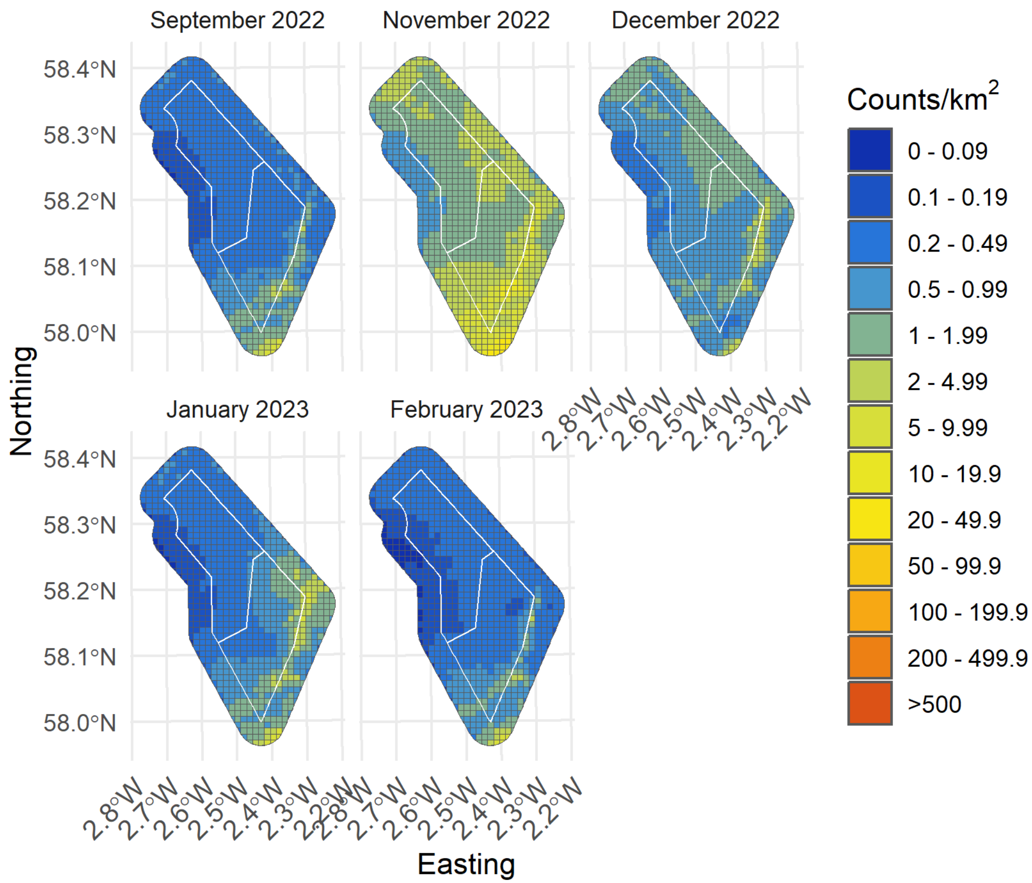


Figure 75. Upper confidence limit of density of all fulmar in the survey area for months with sufficient observations between September 2022 and February 2023

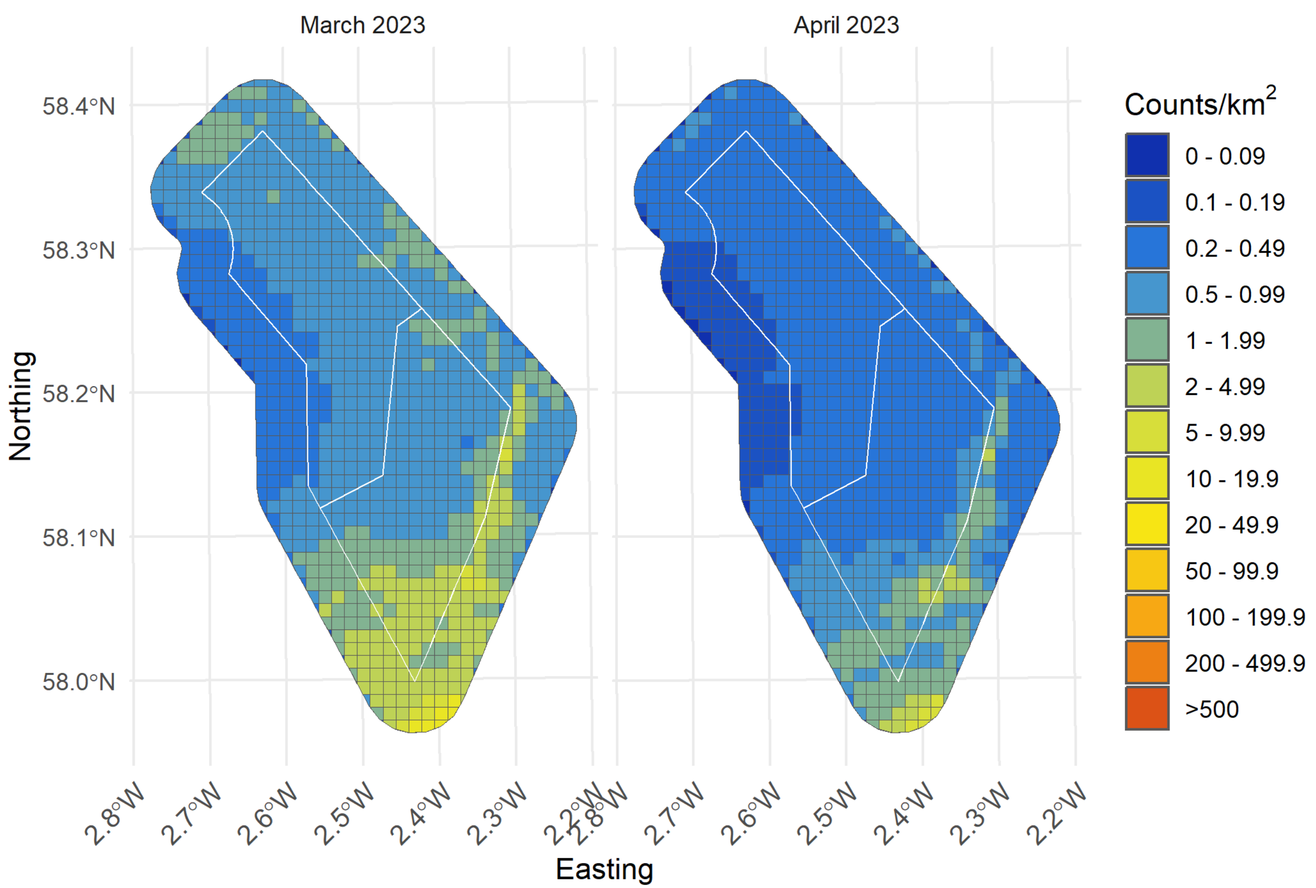


Figure 76. Upper confidence limit of density of all fulmar in the survey area for months with sufficient observations between March and April 2023

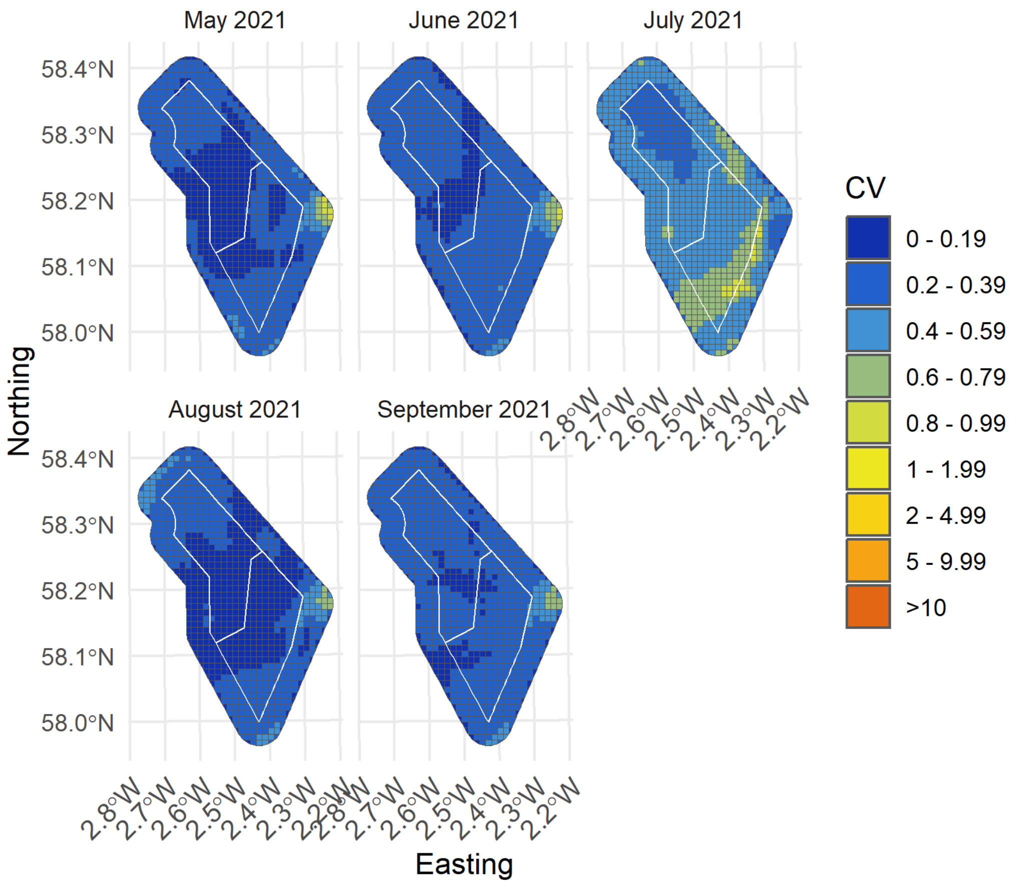


Figure 77. Spatial coefficient of variation of predicted densities of all fulmar from MRSea across the survey area for months with sufficient observations between May and September 2021

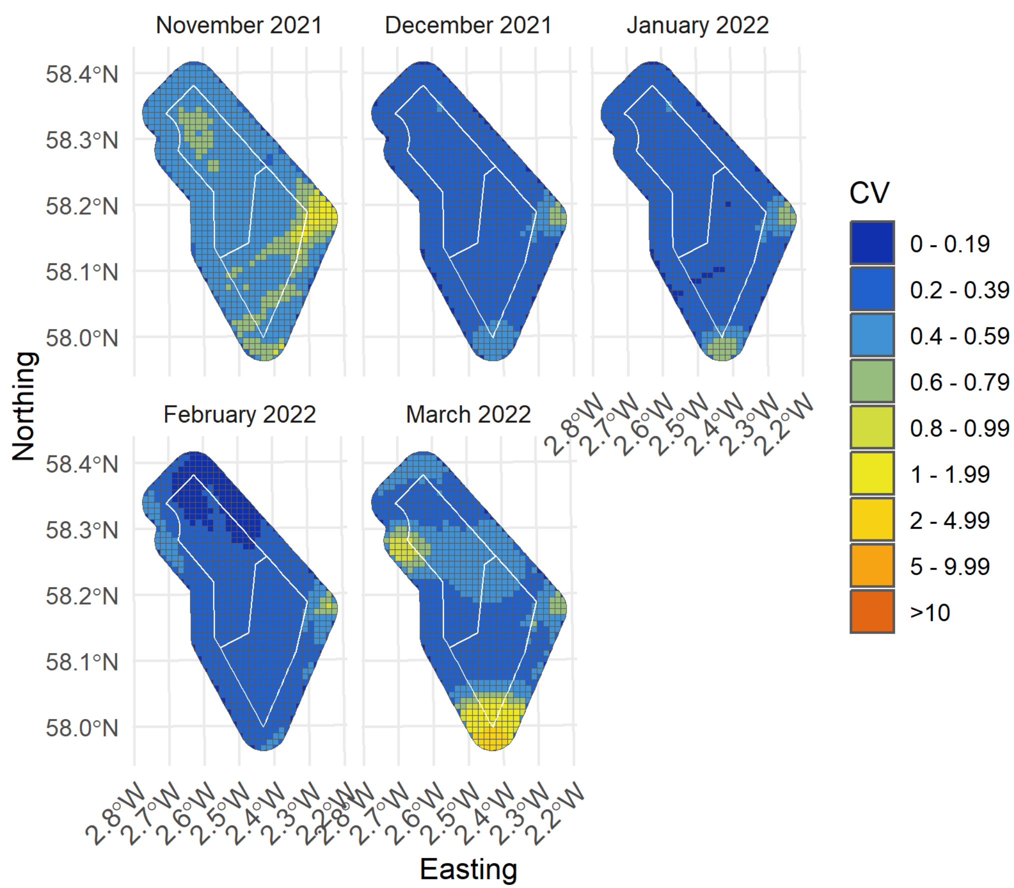


Figure 78. Spatial coefficient of variation of predicted densities of all fulmar from MRSea across the survey area for months with sufficient observations between November 2021 and March 2022

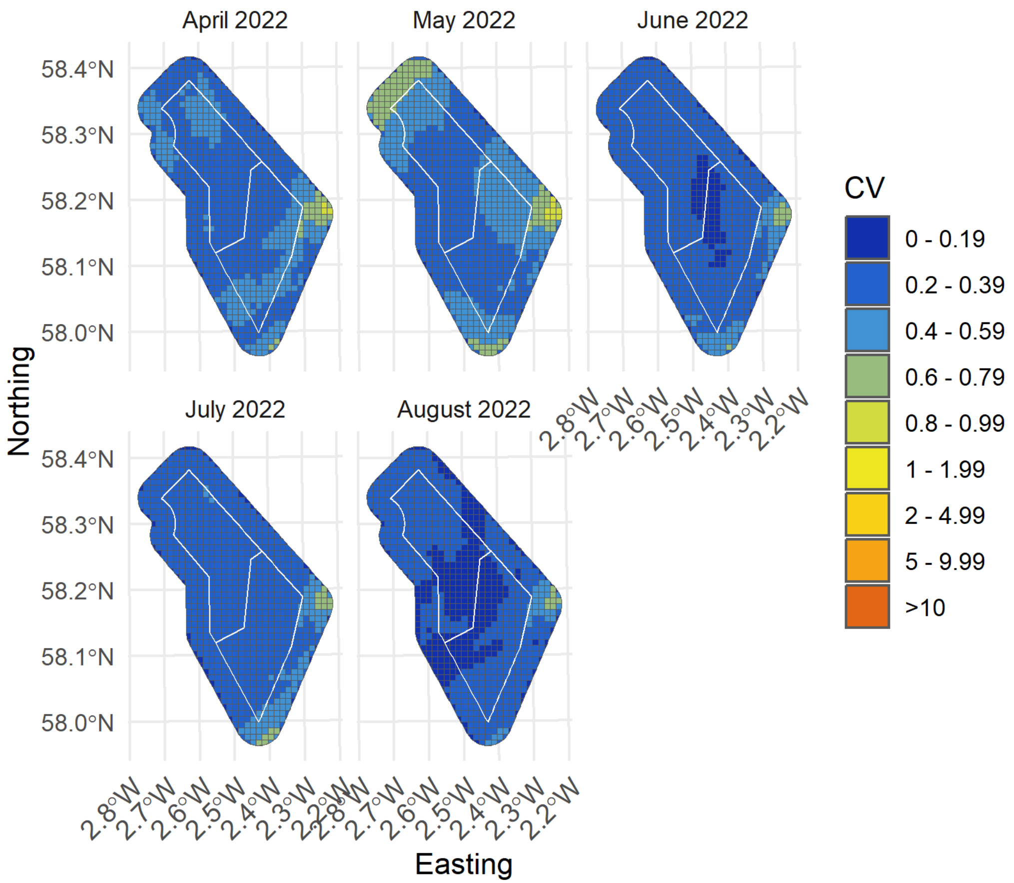


Figure 79. Spatial coefficient of variation of predicted densities of all fulmar from MRSea across the survey area for months with sufficient observations between April and August 2022

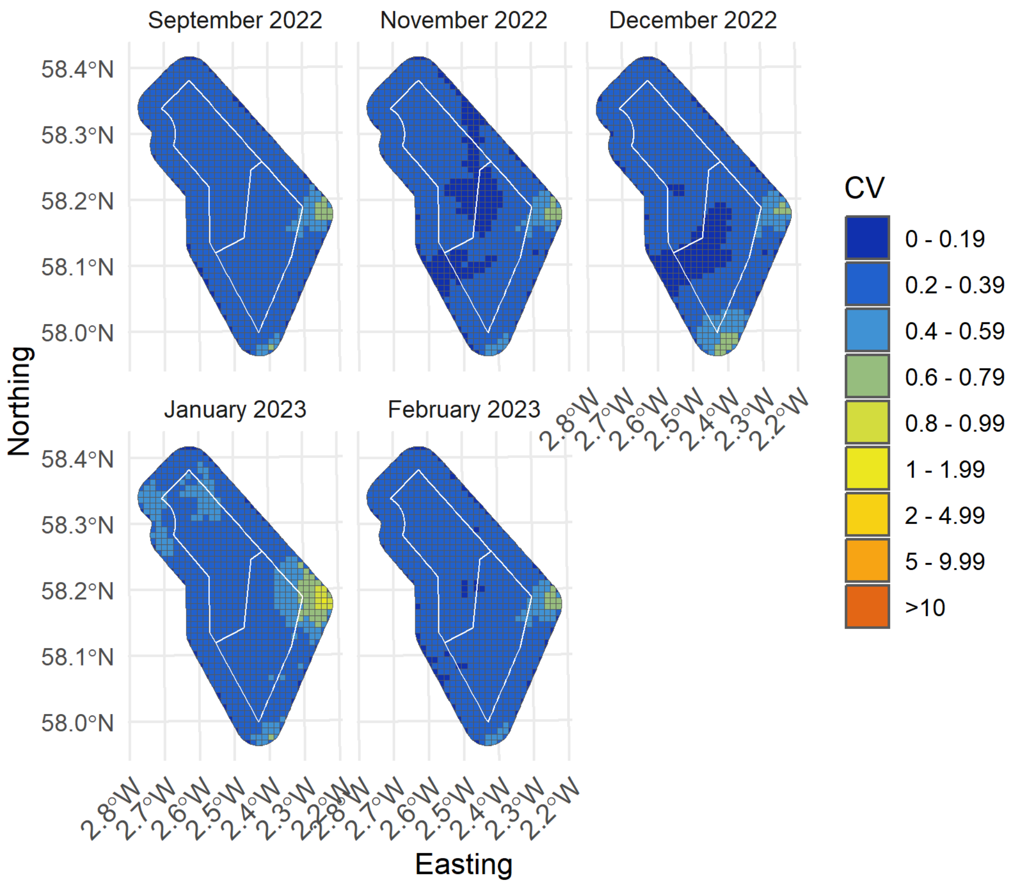


Figure 80. Spatial coefficient of variation of predicted densities of all fulmar from MRSea across the survey area for months with sufficient observations between September 2022 and February 2023

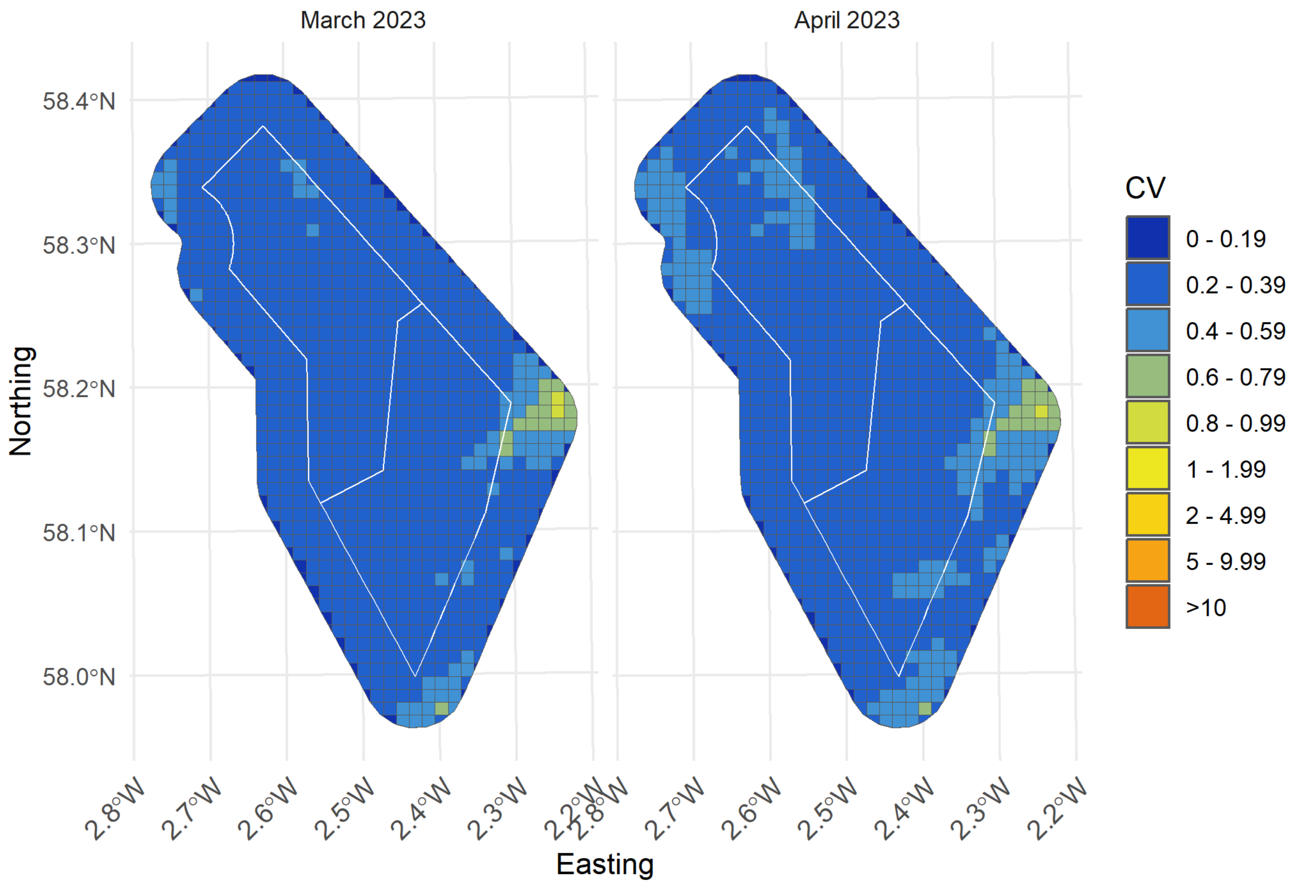


Figure 81. Spatial coefficient of variation of predicted densities of all fulmar from MRSea across the survey area for months with sufficient observations between March and April 2023

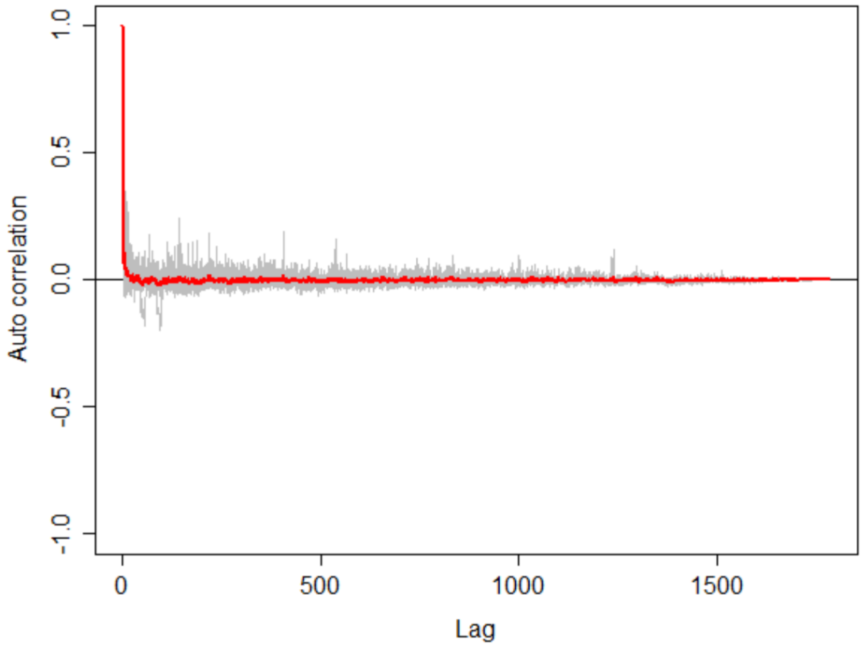


Figure 82. Autocorrelation test for fulmar density surface models when using transect as a blocking feature in MRSea showing no significant correlation. A Runs test on the data prior to using transect as a blocking feature gave a p-value of << 0.0001 (i.e., that the data were significantly autocorrelated when not using a blocking feature)

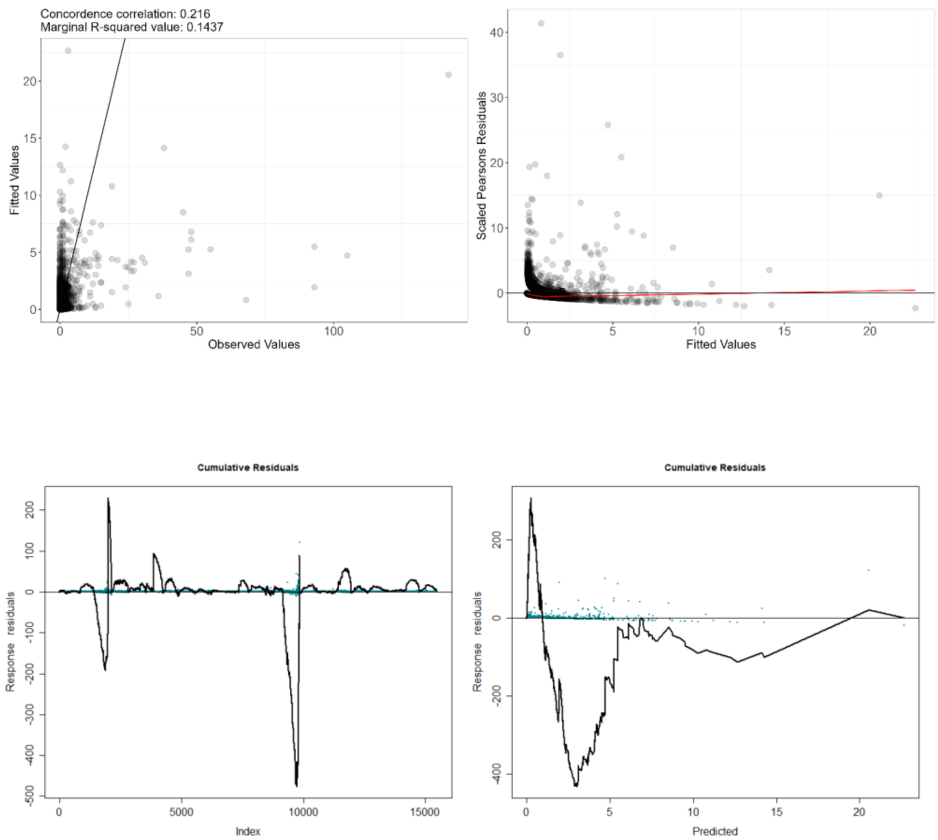


Figure 83. Fitted (MRSea predictions) versus observed counts of all fulmar

Table 12. ANOVA results from the best MRSea model for All fulmar as selected by cross-validation

Variable	Degrees of Freedom	Chi-square	P-value
Bathymetry	1	13.73	<<0.001
Distance to Colony	3	14.84	0.002
Distance to Turbine	5	62.19	<<0.001
X/Y (location)	15	57.36	<<0.001

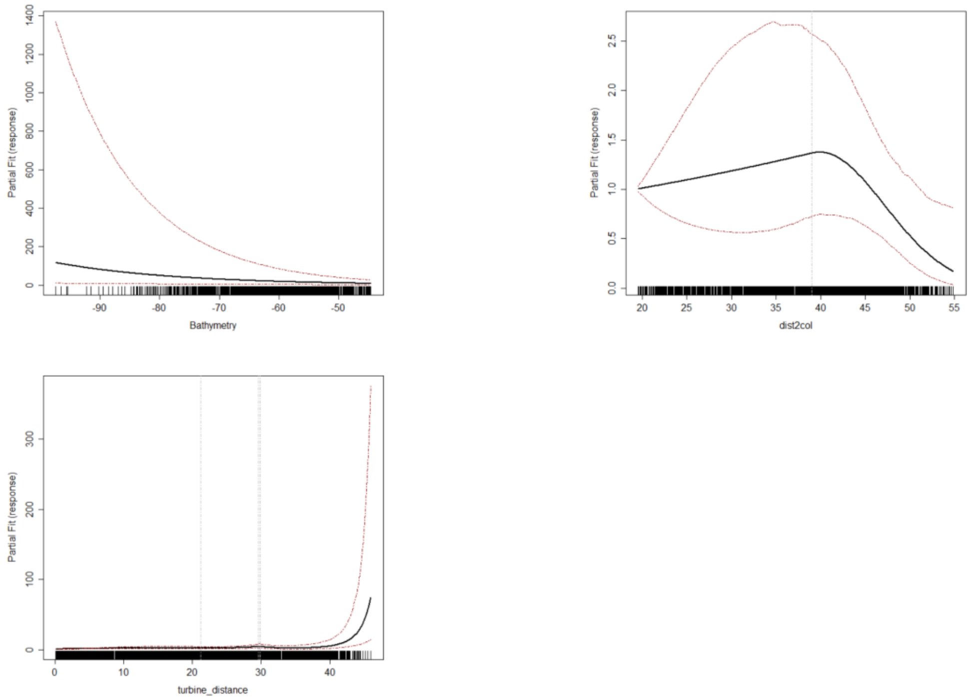


Figure 84. Partial dependence plots for significant variables for all fulmar from MRSea. (Clockwise from top left, bathymetry, distance to colony, distance to turbine)

3.3.2 Flying Birds Model

Table 13. Candidate and final covariates for flying fulmar model.

Starting model covariates after VIF-based collinearity removal	Final model covariates after removal by SALSA
Aspect	Aspect
Slope	Slope
Bathymetry	Distance to turbine
Sandeel Density	Daily Mean of Sea Surface Temperature
Distance to colony	
Distance to turbine	
Daily Mean of Sea Surface Temperature	

Distribution maps generated using MRSea (Figure 84-88) suggest that flying fulmar are distributed in higher densities during the breeding season (01 April - 15 September), in particular in July and August. In the non-breeding season, density declines considerably, however there are some interesting hotspots in the north and east corners in November 2021. The highest densities were observed consistently throughout the majority of the study area in July 2022, with slightly lower densities occurring along the western edge.

Model fit was moderate with a marginal R squared value of 0.12 and root mean squared error of 0.21. Cumulative residuals in the model showed that there was a relatively poor relationship between predicted and observed values across most of the range of predicted values, but were typically bounded around 0 across the whole (Figure 105).

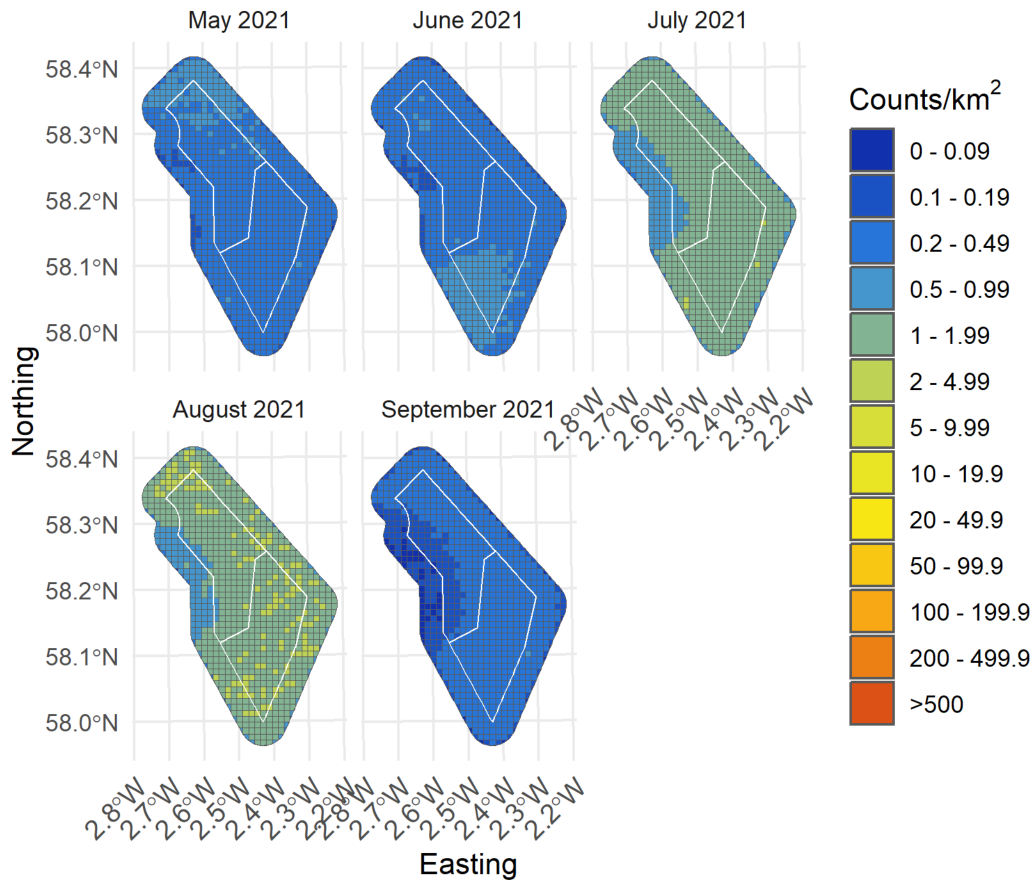


Figure 85. Median density of flying fulmar in the survey area for months with sufficient observations between May and September 2021

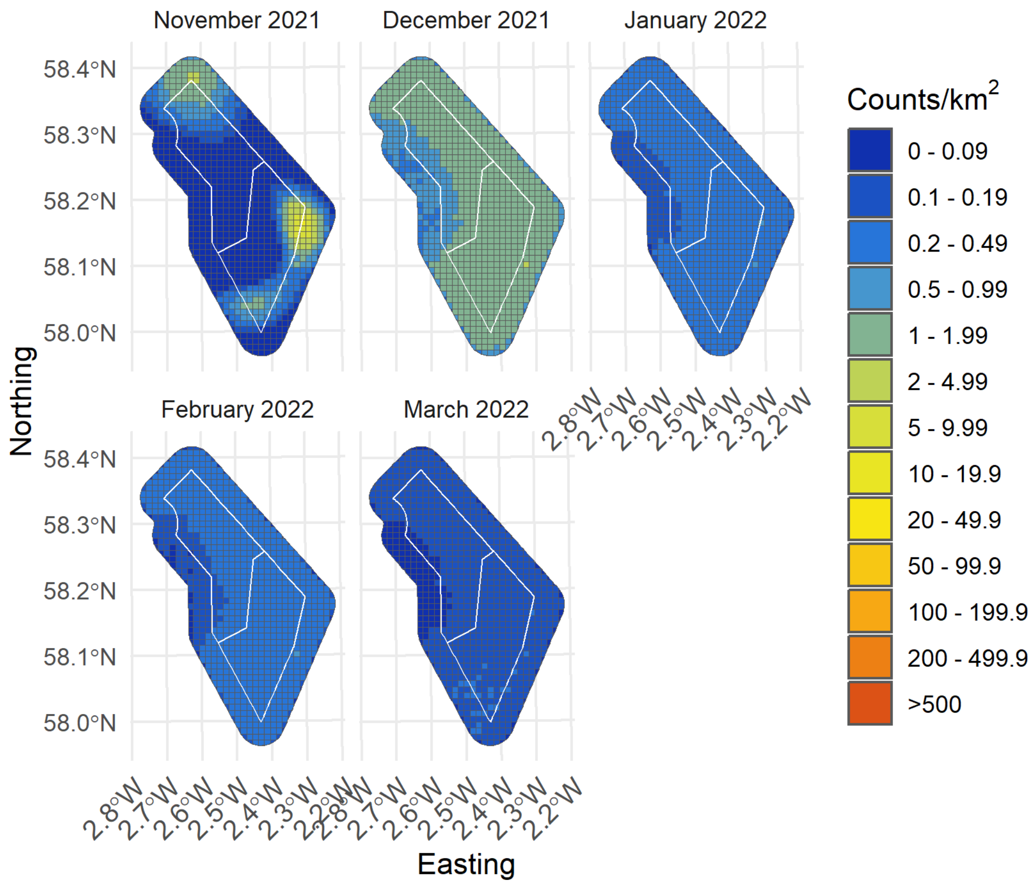


Figure 86. Median density of flying fulmar in the survey area for months with sufficient observations between November 2021 and March 2022

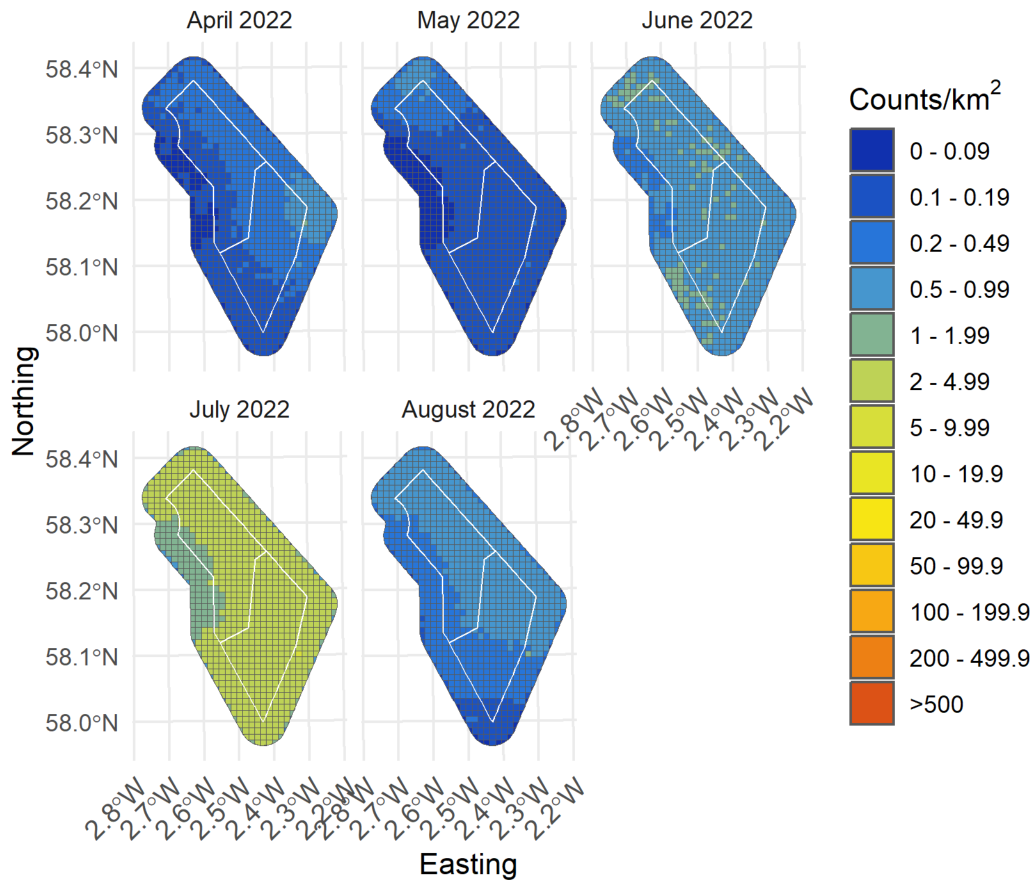


Figure 87. Median density of flying fulmar in the survey area for months with sufficient observations between April and August 2022

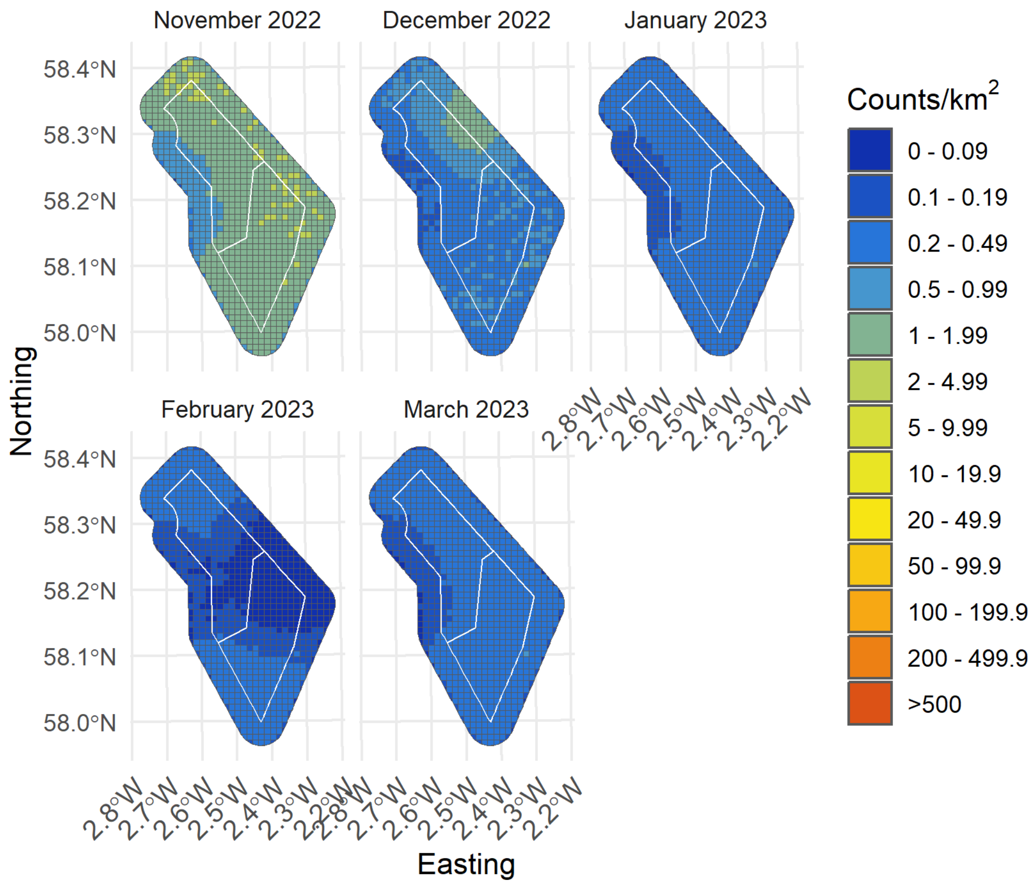


Figure 88. Median density of flying fulmar in the survey area for months with sufficient observations between November 2022 and March 2023

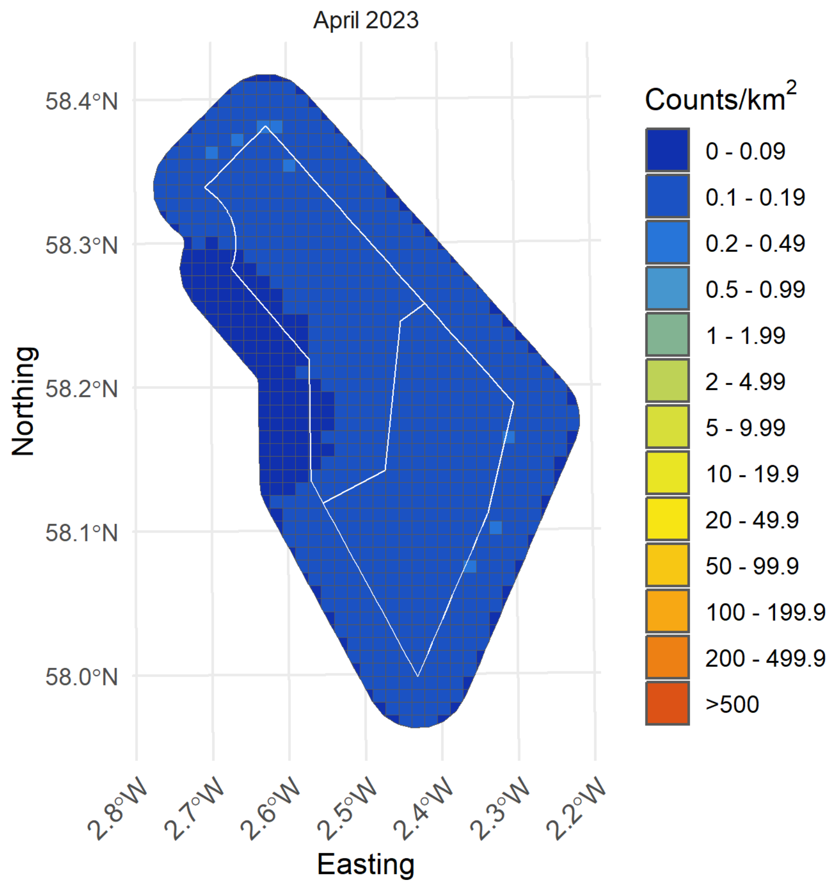


Figure 89. Median density of flying fulmar in the survey area for months with sufficient observations in April 2023

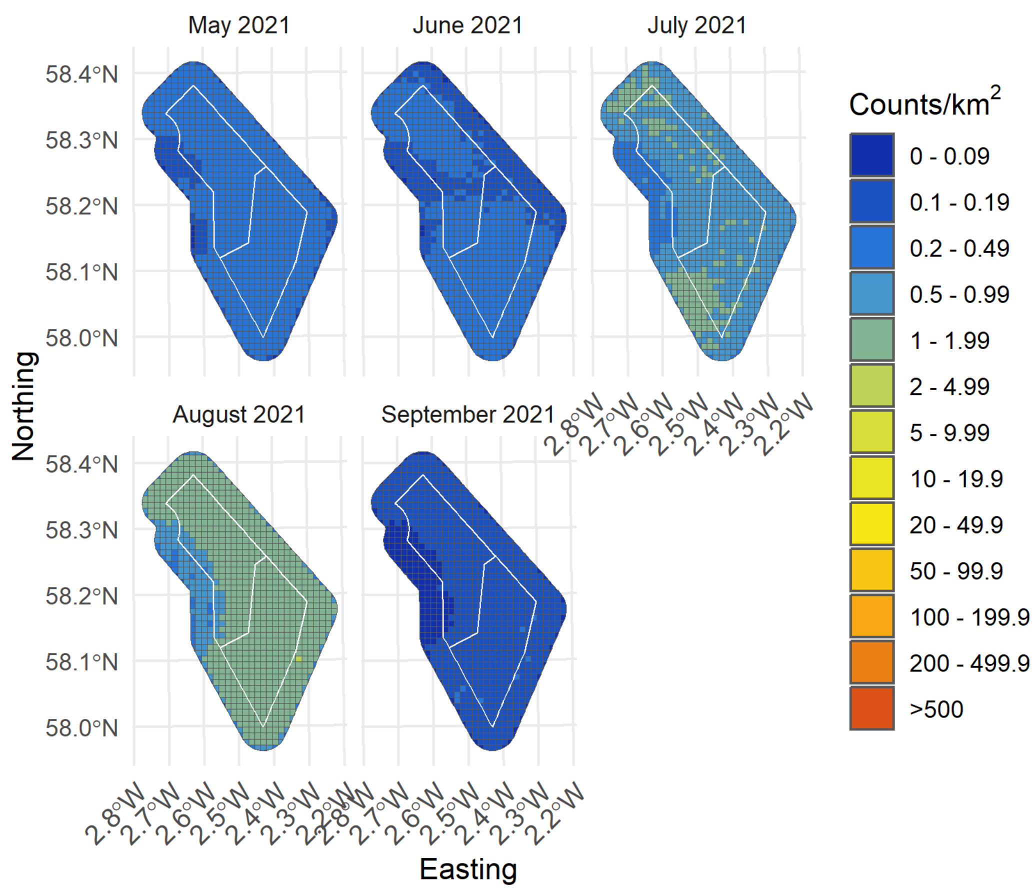


Figure 90. Lower confidence limit of density of flying fulmar in the survey area for months with sufficient observations between May and September 2021

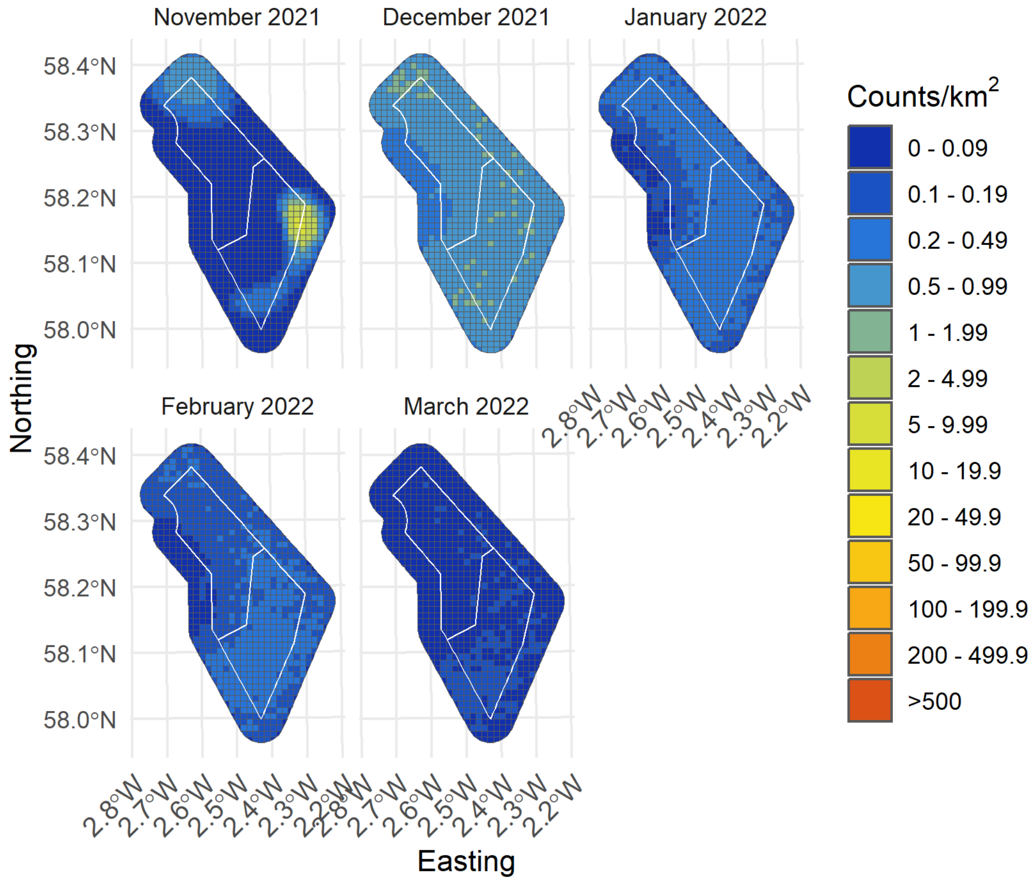


Figure 91. Lower confidence limit of density of flying fulmar in the survey area for months with sufficient observations between November 2021 and March 2022

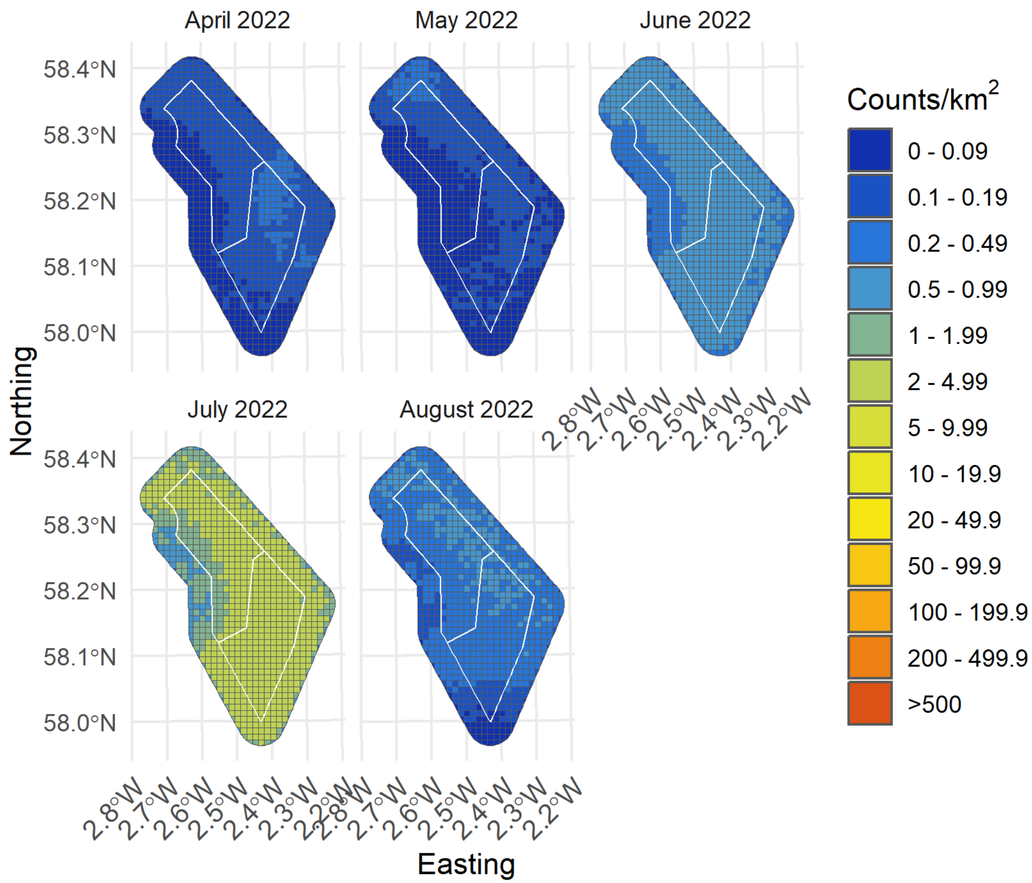


Figure 92 .Lower confidence limit of density of flying fulmar in the survey area for months with sufficient observations between April and August 2022

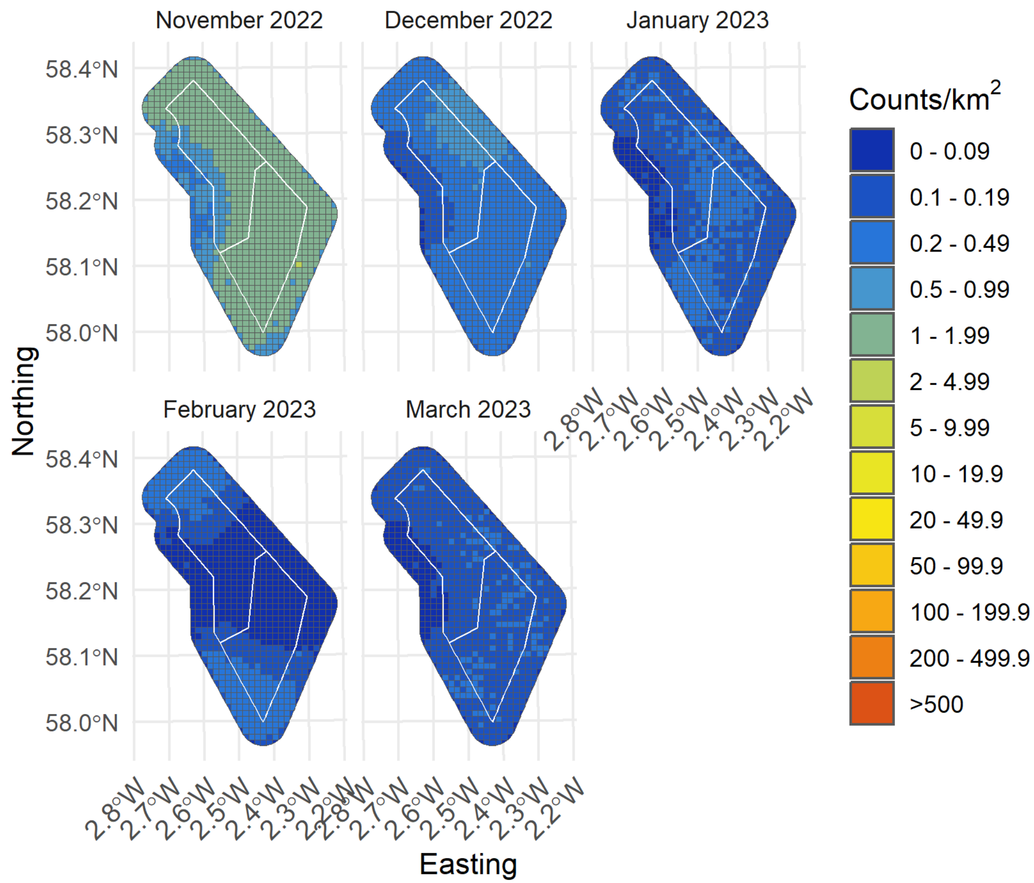


Figure 93. Lower confidence limit of density of flying fulmar in the survey area for months with sufficient observations between November 2022 and March 2023

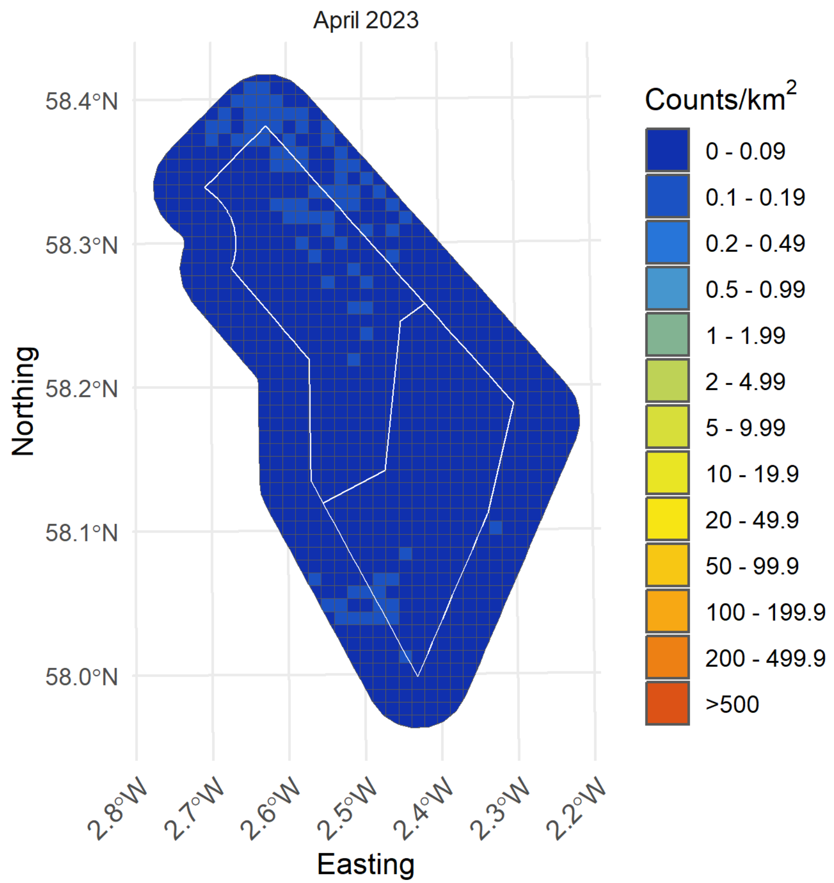


Figure 94. Lower confidence limit of density of flying fulmar in the survey area in April 2023

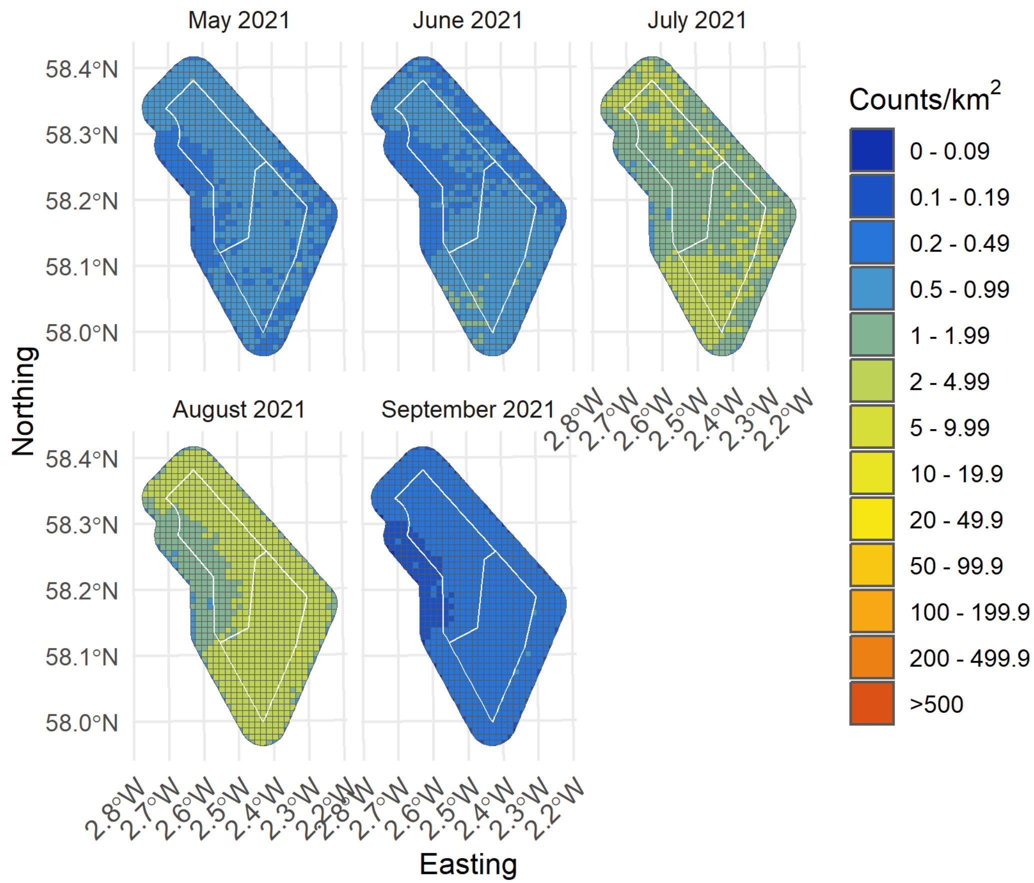


Figure 95. Upper confidence limit of density of flying fulmar in the survey area for months with sufficient observations between May and September 2021

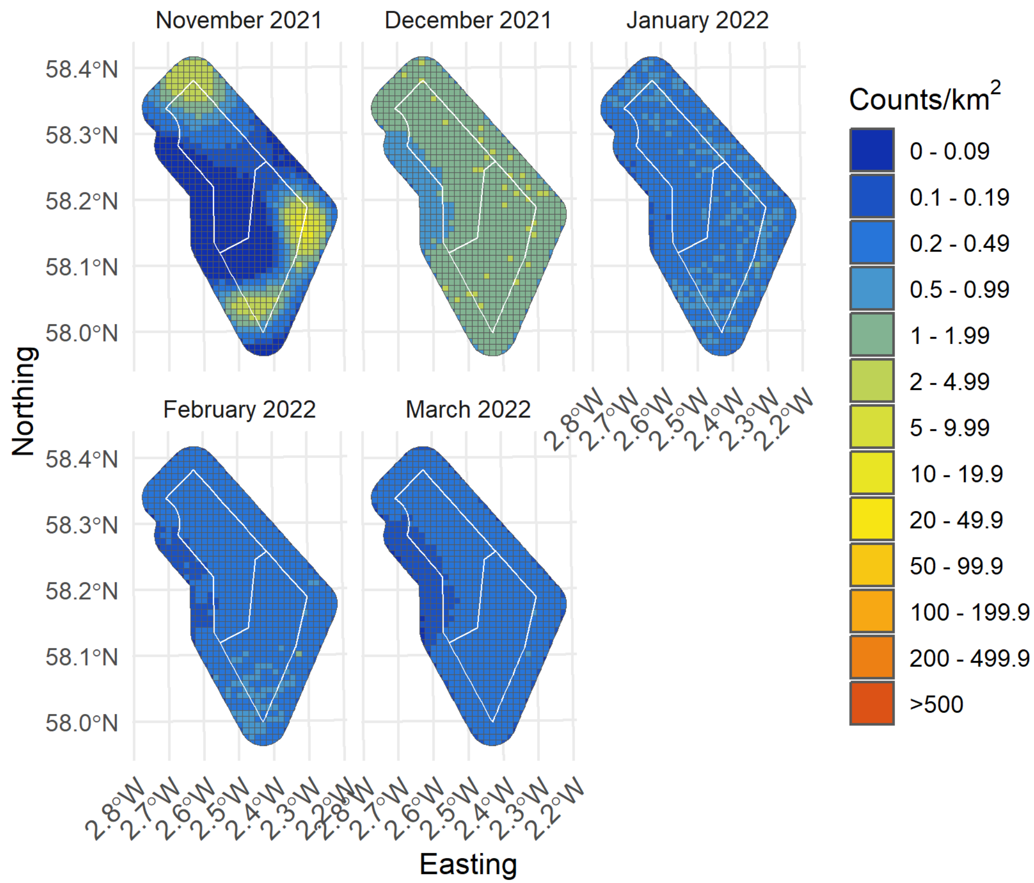


Figure 96. Upper confidence limit of density of flying fulmar in the survey area for months with sufficient observations between November 2021 and March 2022

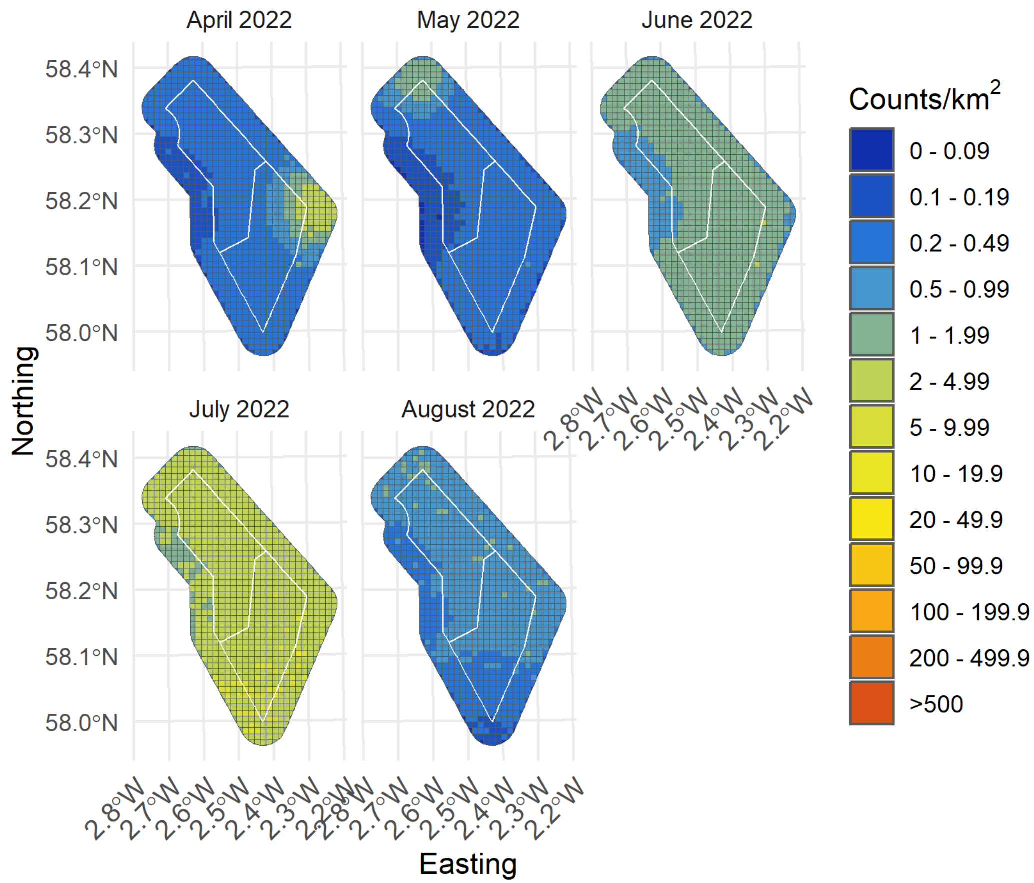


Figure 97. Upper confidence limit of density of flying fulmar in the survey area for months with sufficient observations between April and August 2022

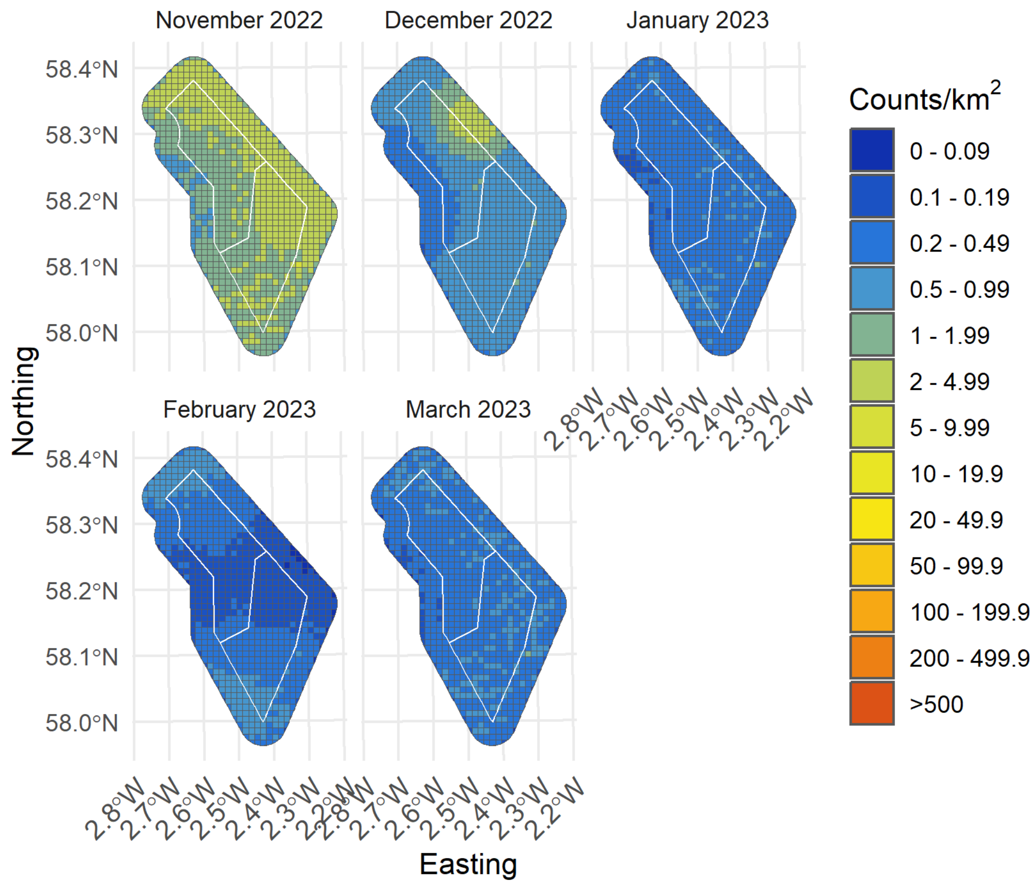


Figure 98. Upper confidence limit of density of flying fulmar in the survey area for months with sufficient observations between November 2022 and March 2023

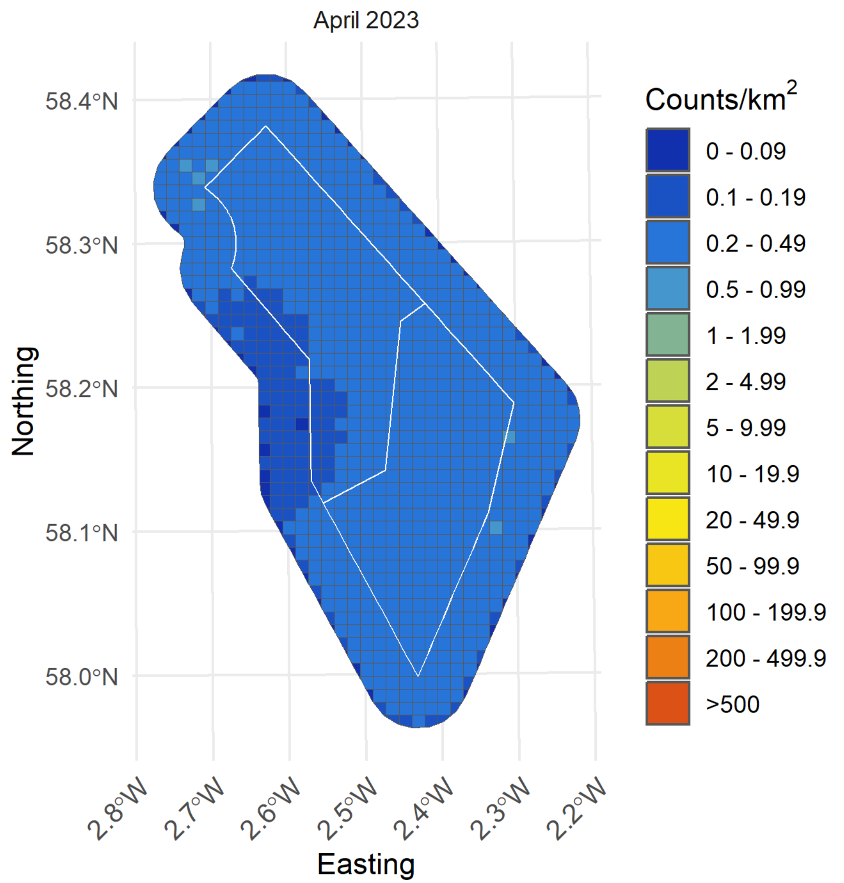


Figure 99. Upper confidence limit of density of flying fulmar in the survey area in April 2023

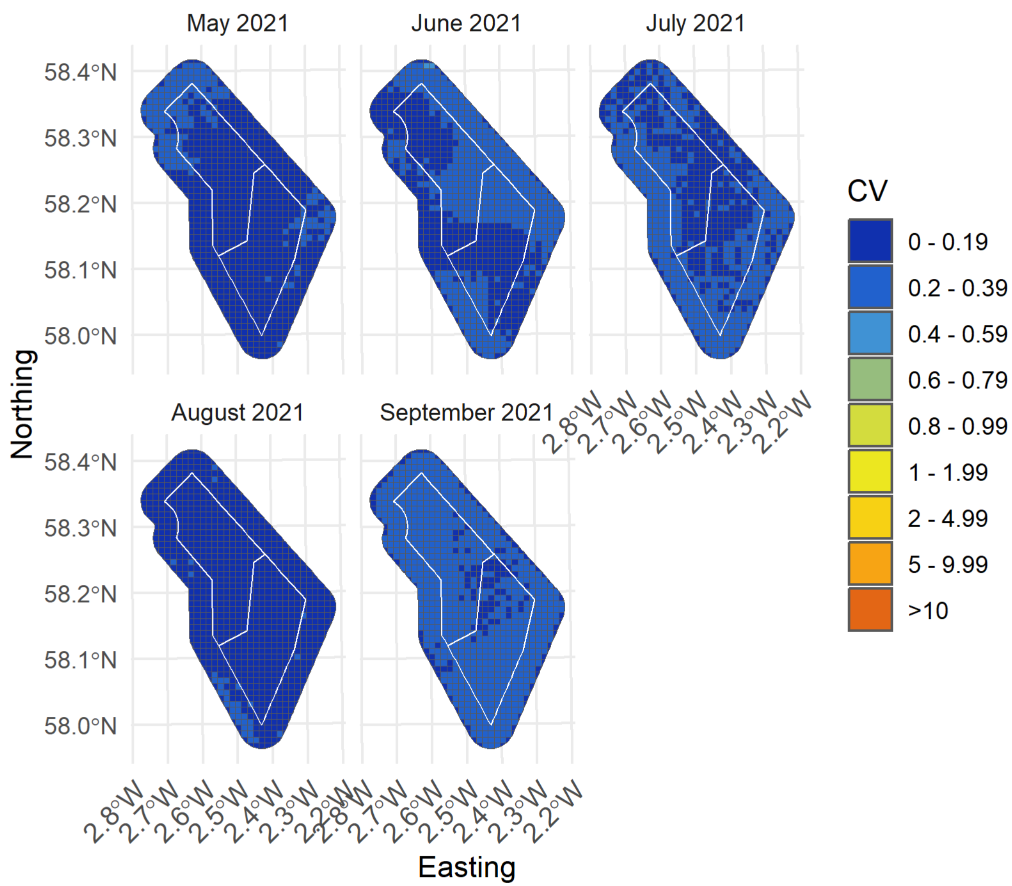


Figure 100. Spatial coefficient of variation of predicted densities of flying fulmar from MRSea across the survey area for months with sufficient observations between May and September 2021

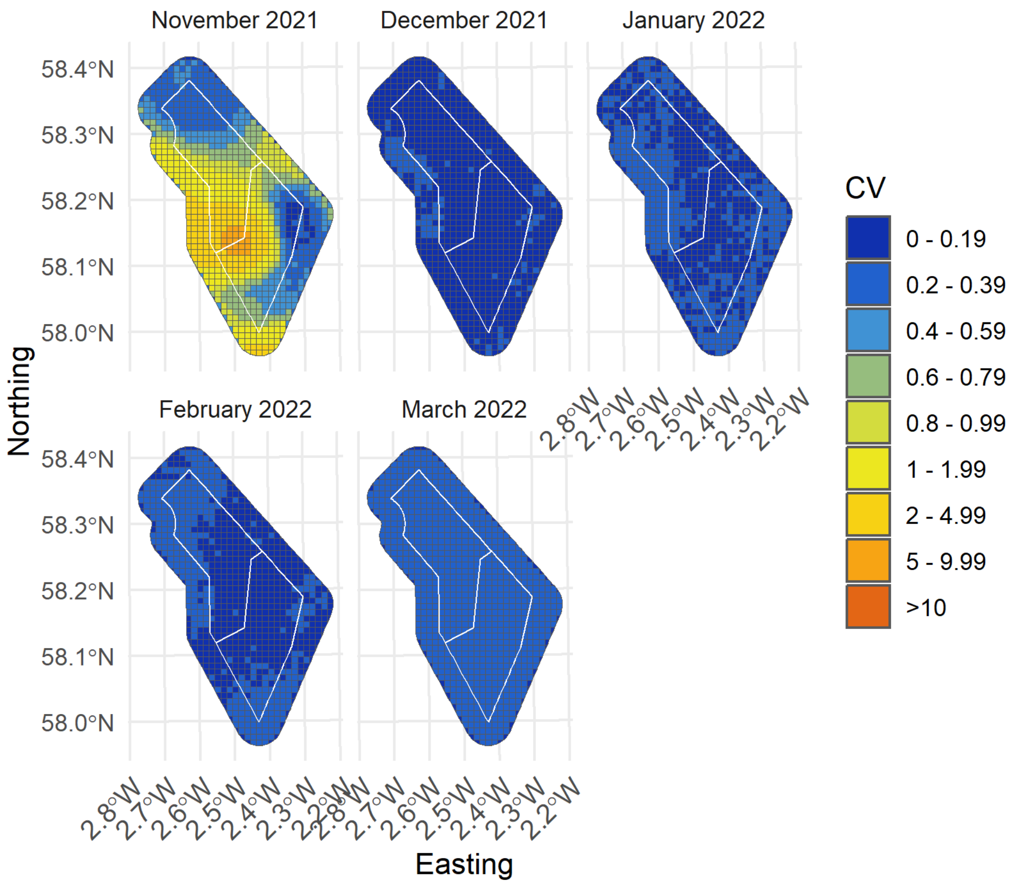


Figure 101. Spatial coefficient of variation of predicted densities of flying fulmar from MRSea across the survey area for months with sufficient observations between November 2021 and March 2022

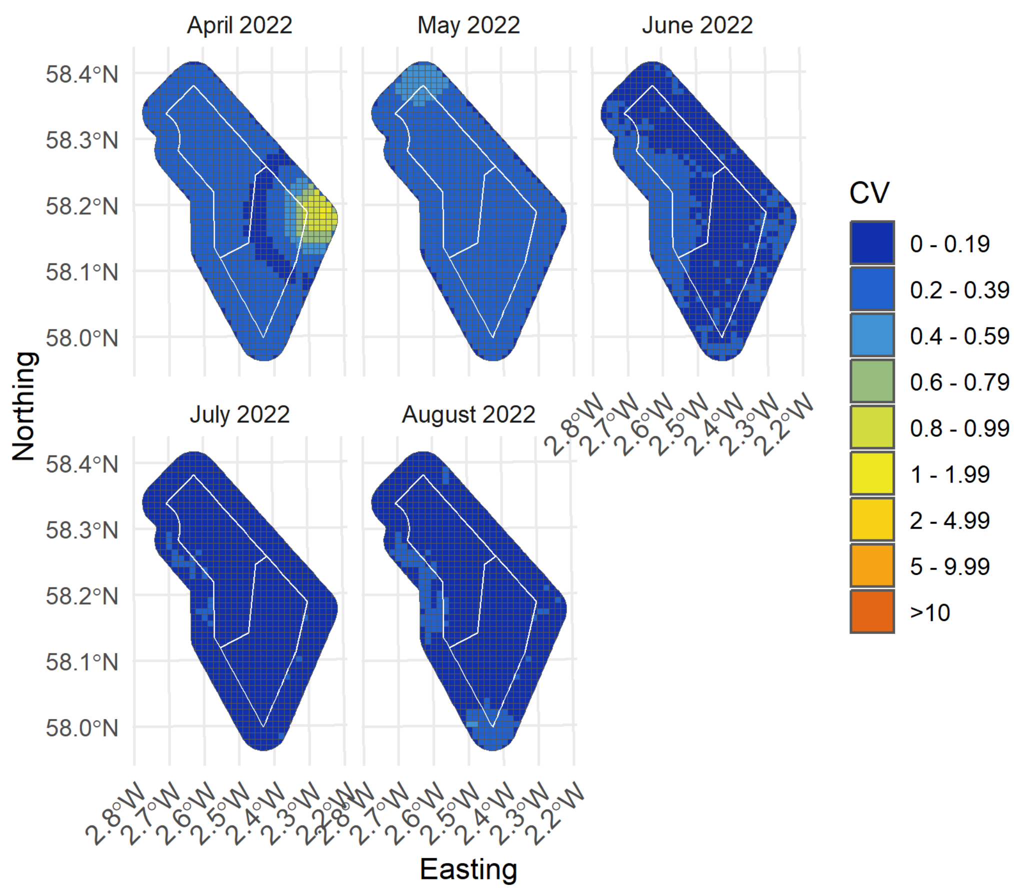


Figure 102. Spatial coefficient of variation of predicted densities of flying fulmar from MRSea across the survey area for months with sufficient observations between April and August 2022

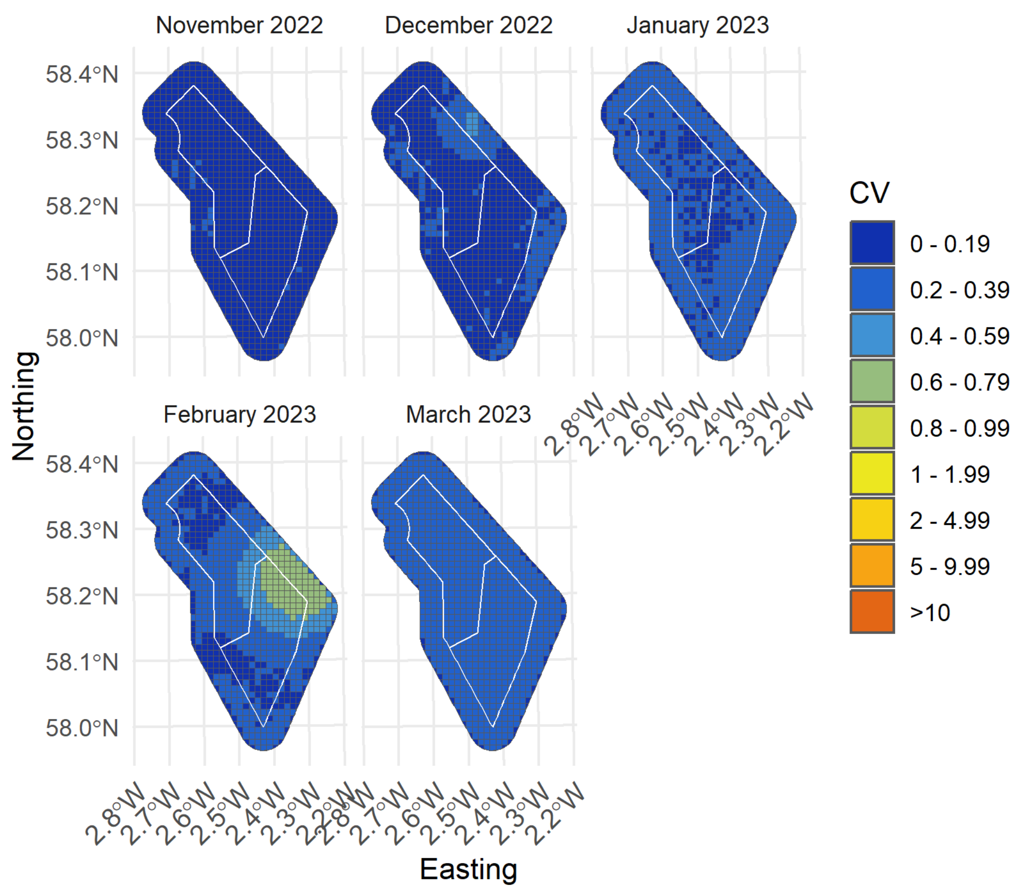


Figure 103. Spatial coefficient of variation of predicted densities of flying fulmar from MRSea across the survey area for months with sufficient observations between November 2022 and March 2023

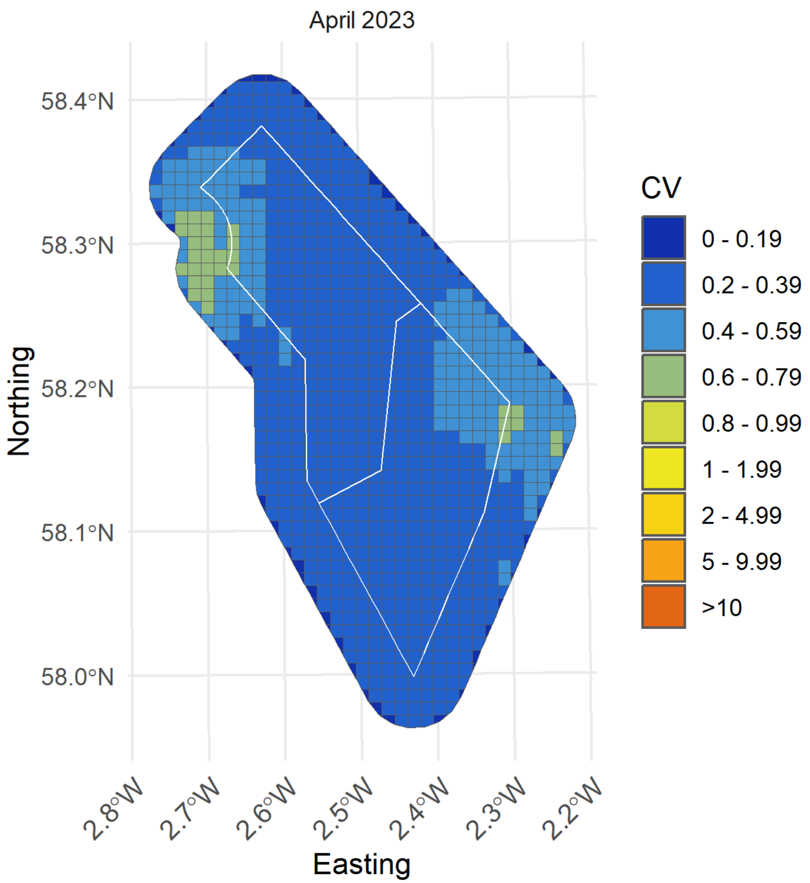


Figure 104. Spatial coefficient of variation of predicted densities of flying fulmar from MRSea across the survey area in April 2023

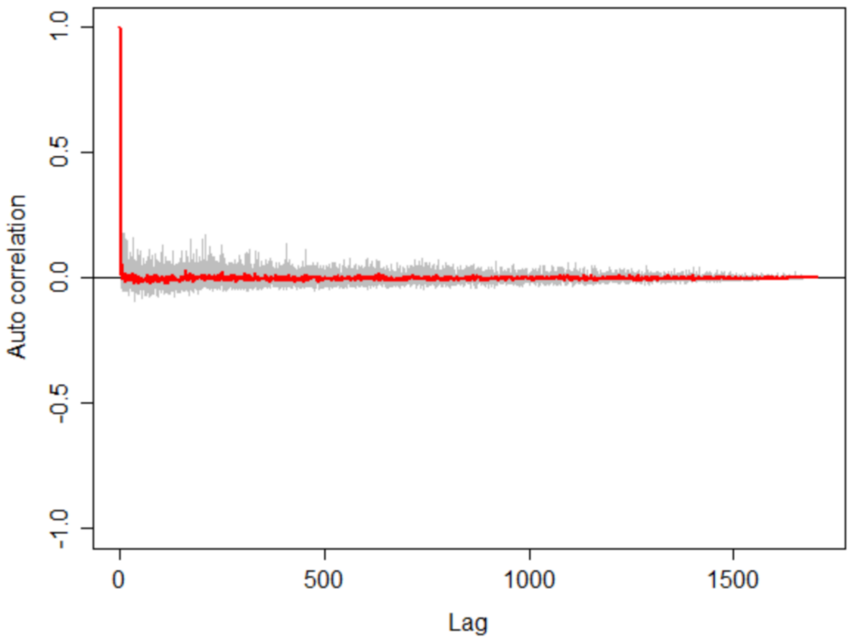


Figure 105. Autocorrelation test for fulmar density surface models when using transect as a blocking feature in MRSea showing no significant correlation. A Runs test on the data prior to using transect as a blocking feature gave a p-value of << 0.0001 (i.e., that the data were significantly autocorrelated when not using a blocking feature)

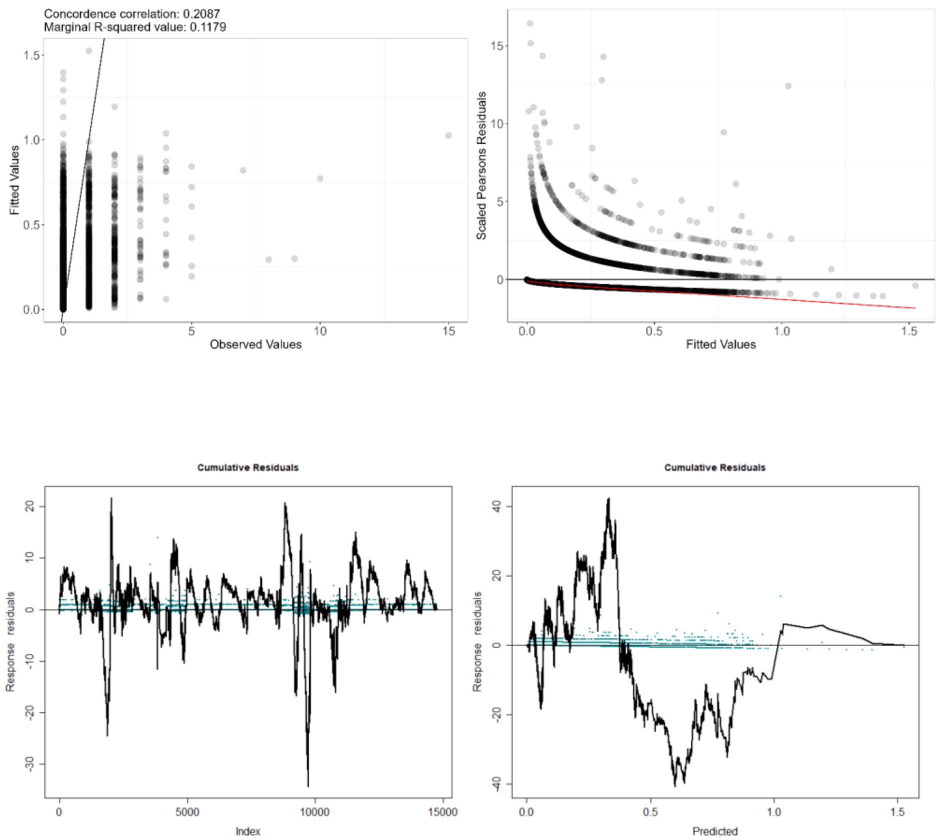


Figure 106. Fitted (MRSea predictions) versus observed counts of flying fulmar

Table 14. ANOVA results from the best MRSea model for flying fulmar as selected by cross-validation

Variable	Degrees of Freedom	Chi-square	P-value
Aspect	5	9.14	0.104
Slope	5	6.51	0.259
Distance to Turbine	3	31.41	<<0.001
Sea Surface Temperature (daily)	5	6.78	0.238
X/Y Location	10	130.91	<<0.001

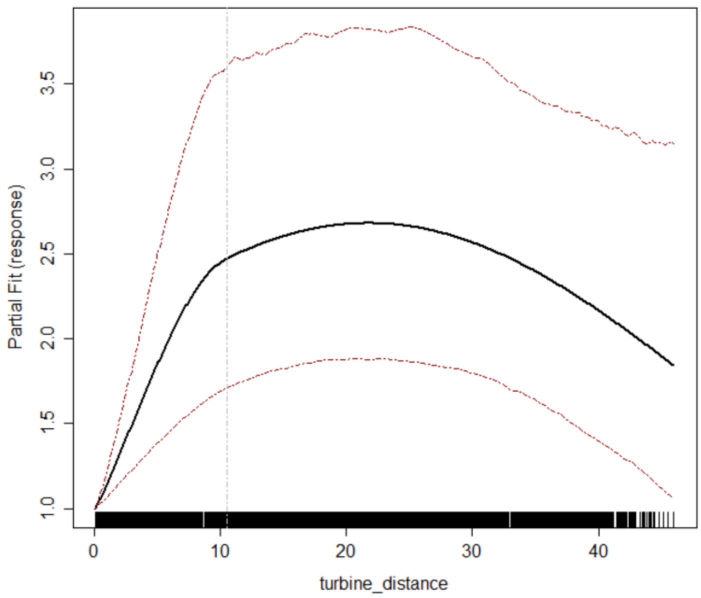


Figure 107. Partial dependence plots for significant variables (distance to turbine) for flying fulmar from MRSea

3.4 Guillemot

Table 15. Candidate and final covatiates for guillemot MRSea model

Starting model covariates after VIF-based collinearity removal	Final model covariates after removal by SALSA
Bathymetry	Bathymetry
Sandeel Density	Sandeel Density
Distance to colony	Distance to colony
Distance to turbine	Distance to turbine
Daily mean of Sea Surface Temperature	Daily mean of Sea Surface Temperature

Distribution maps generated using MRSea (Figure 107-111) suggest that guillemot are widely distributed throughout the study area during the breeding season (01 April – 15 August). In the non-breeding season, distribution is much patchier and density declines. The highest densities were observed along the eastern corner of the study area in May 2022.

Broadly, model fit was better for guillemot than for most species, with a marginal R squared value of 0.18 and root mean squared error of 37.33. Cumulative residuals in the model showed that there was a poor relationship between predicted and observed values across most of the range of predicted values, but typically bounded around 0 across the whole (Figure 128).

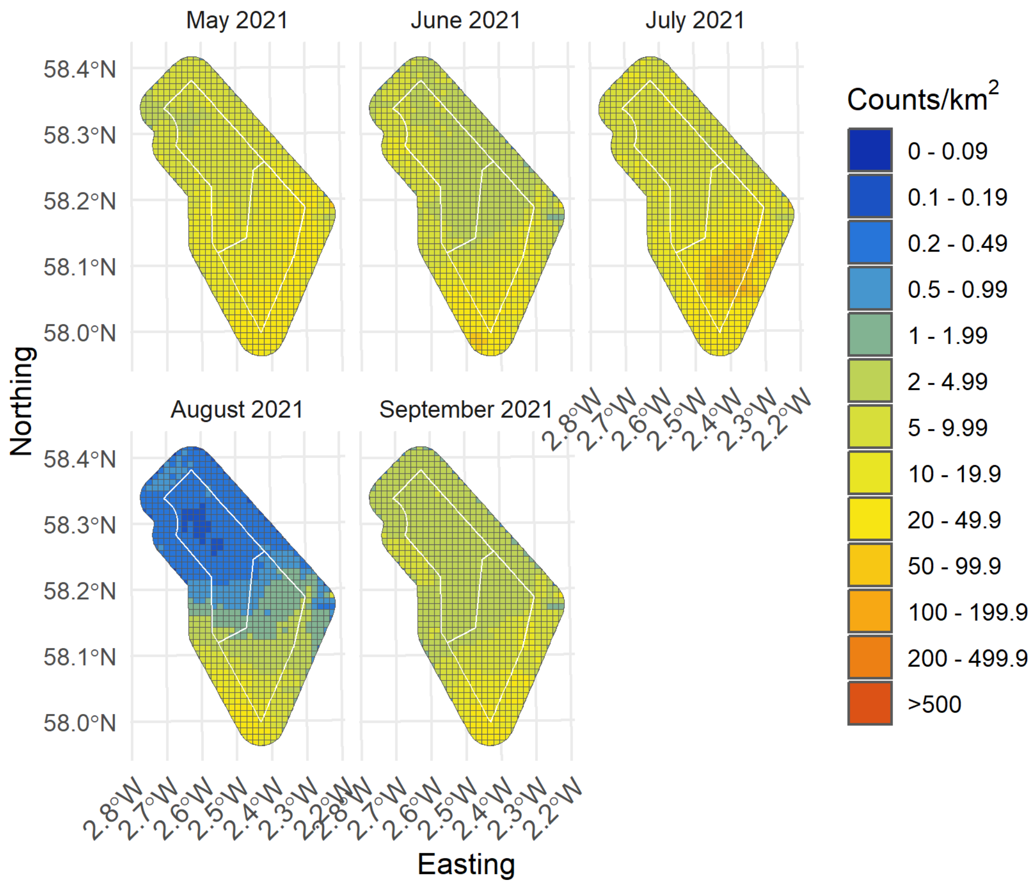


Figure 108. Median density of all guillemot in the survey area for months with sufficient observations between May and September 2021

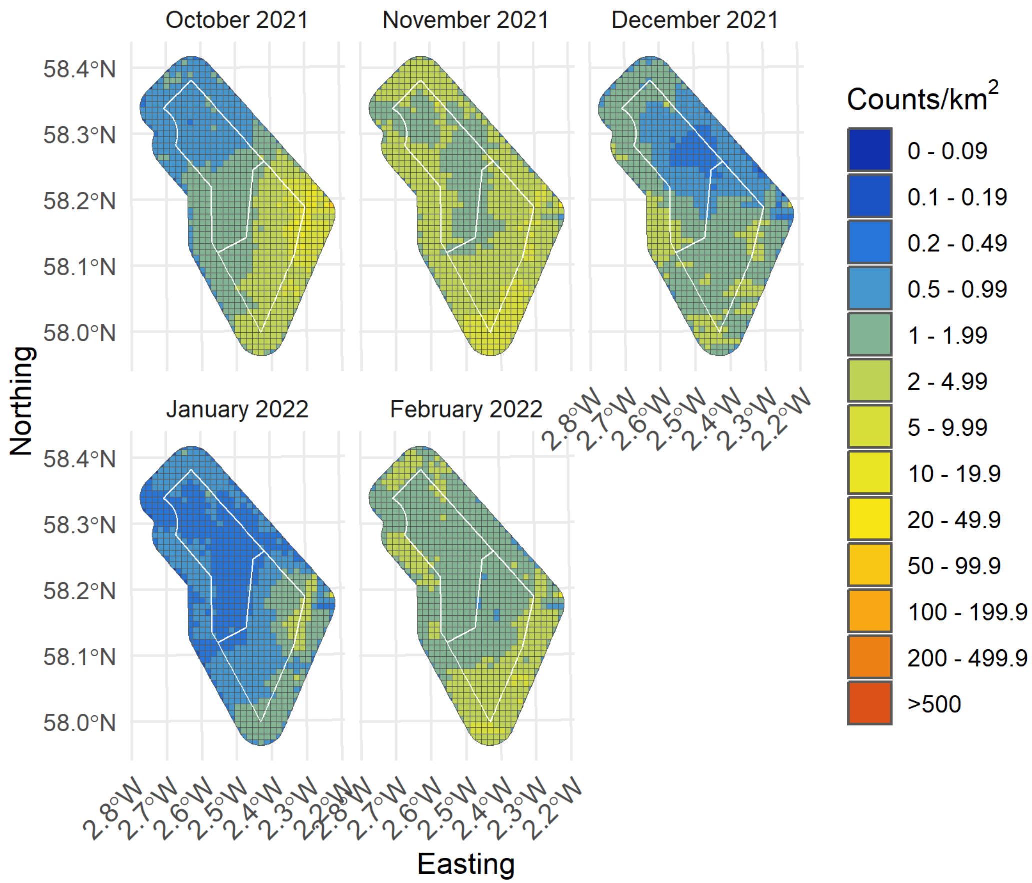


Figure 109. Median density of all guillemot in the survey area for months with sufficient observations between October 2021 and February 2022

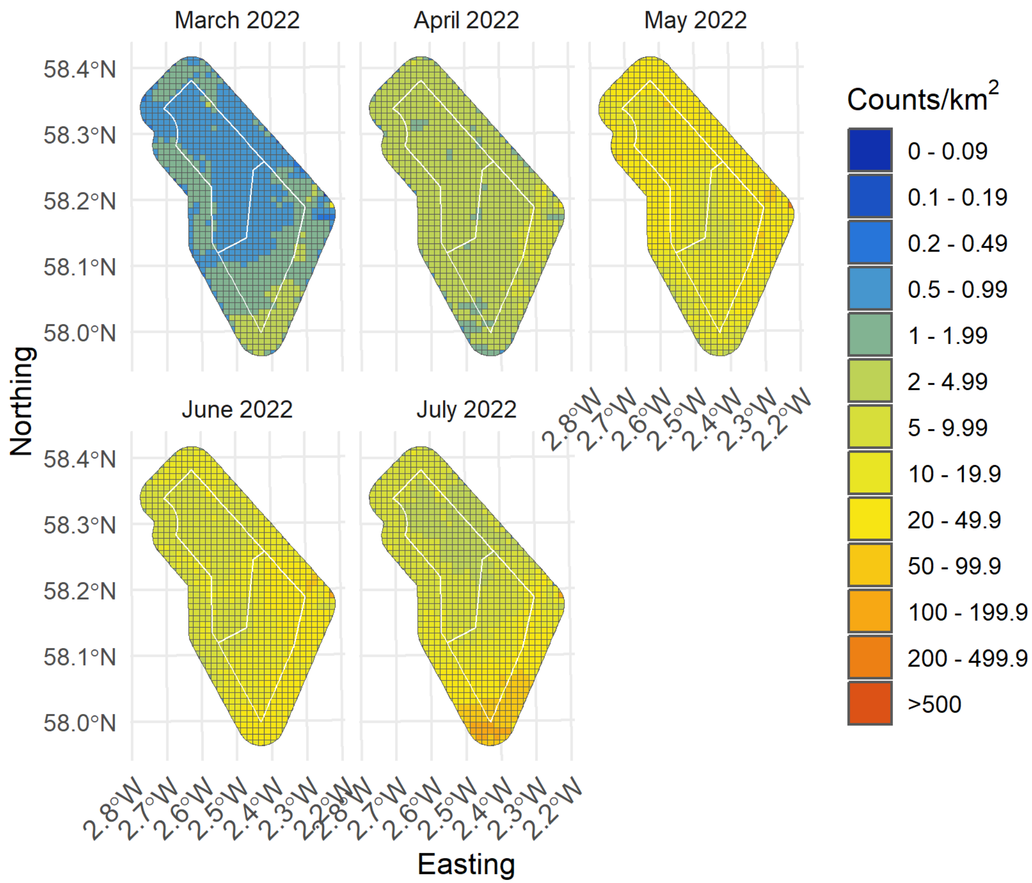


Figure 110. Median density of all guillemot in the survey area for months with sufficient observations between March and July 2022

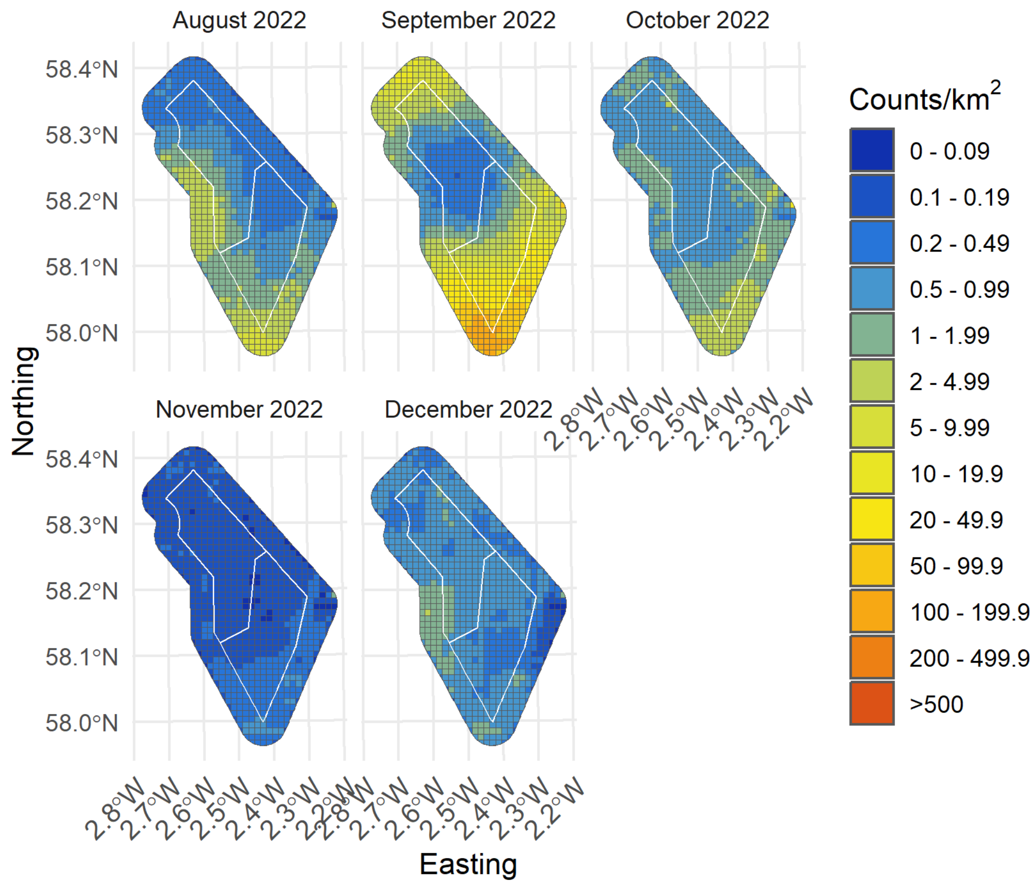


Figure 111. Median density of all guillemot in the survey area for months with sufficient observations between August and December 2022

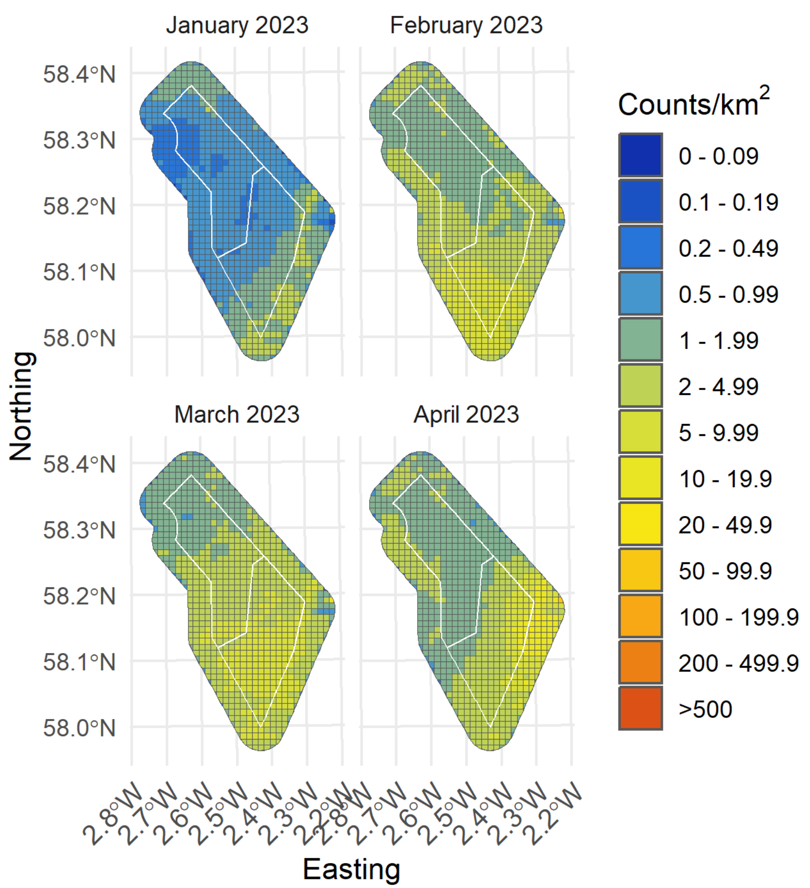


Figure 112. Median density of all guillemot in the survey area for months with sufficient observations between January and April 2023

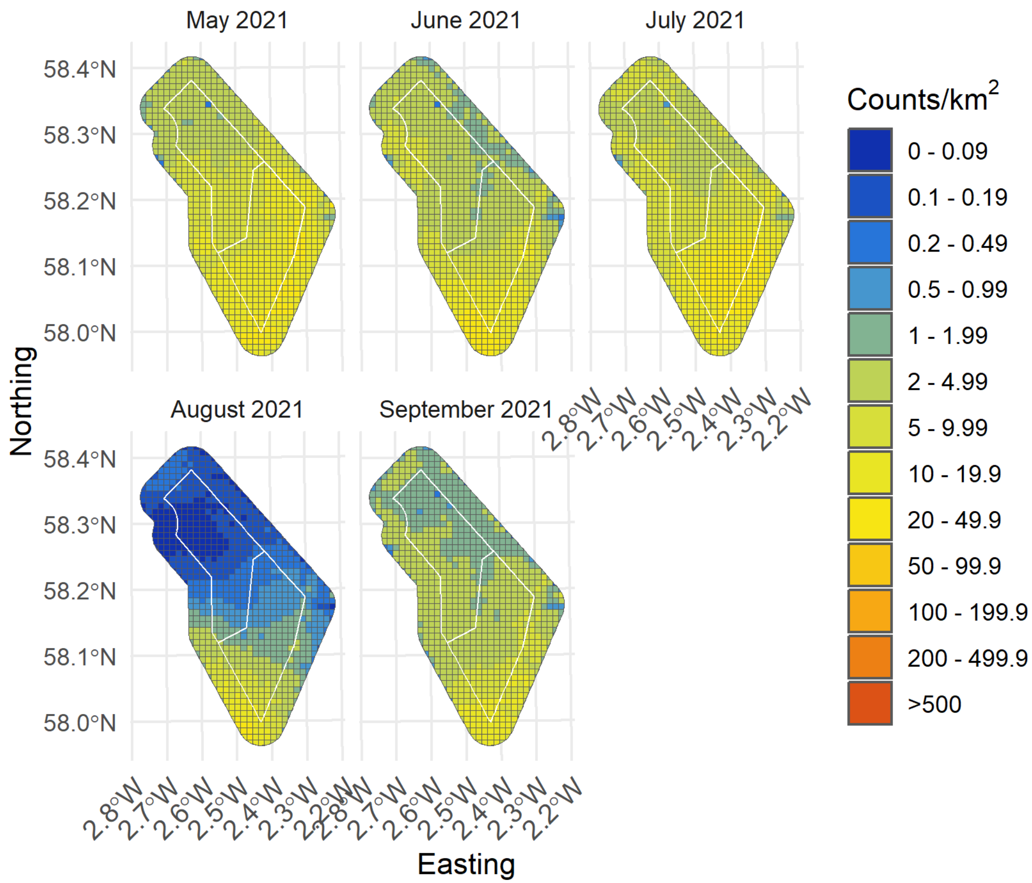


Figure 113. Lower confidence limit of density of all guillemot in the survey area for months with sufficient observations between May and September 2021

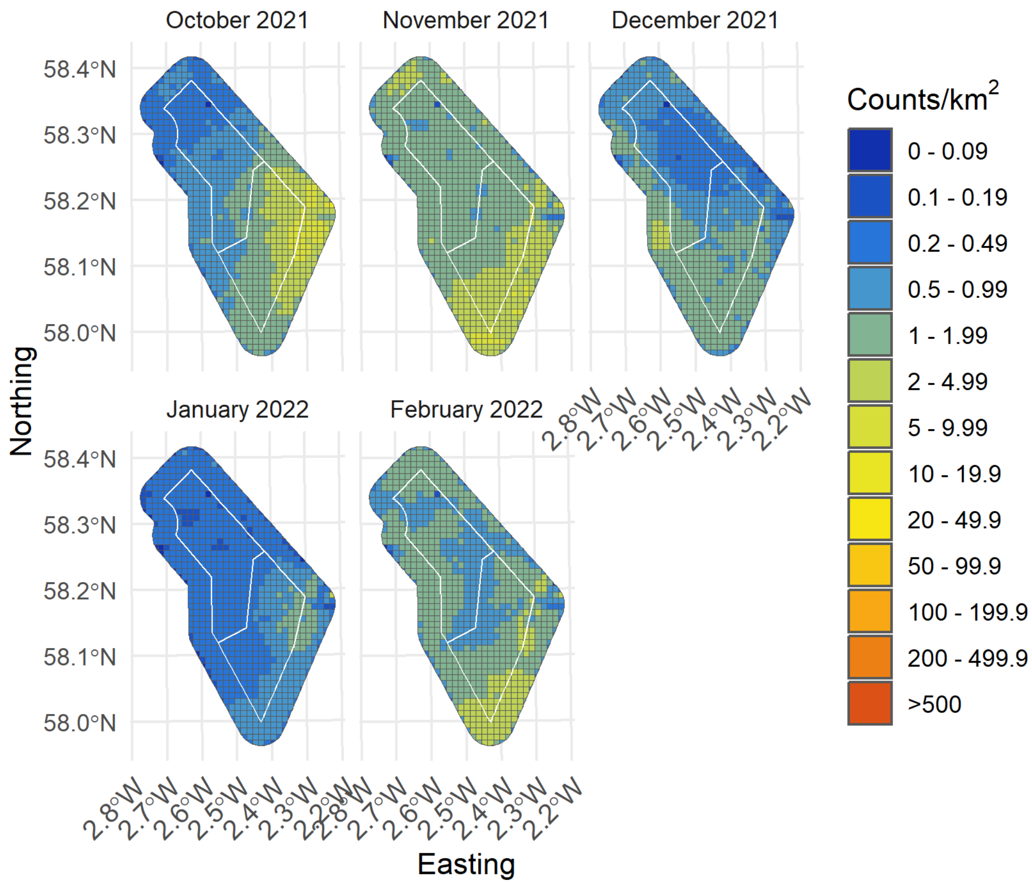


Figure 114. Lower confidence limit of density of all guillemot in the survey area for months with sufficient observations between October 2021 and February 2022

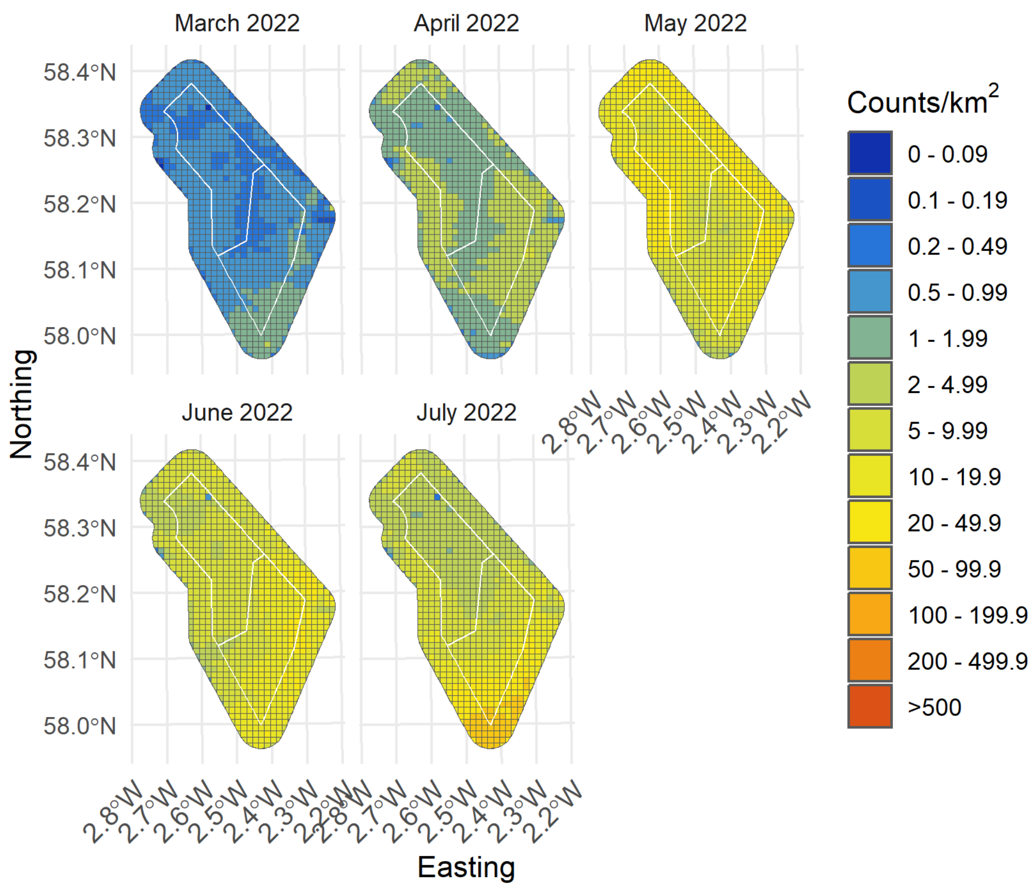


Figure 115. Lower confidence limit of density of all guillemot in the survey area for months with sufficient observations between March and July 2022

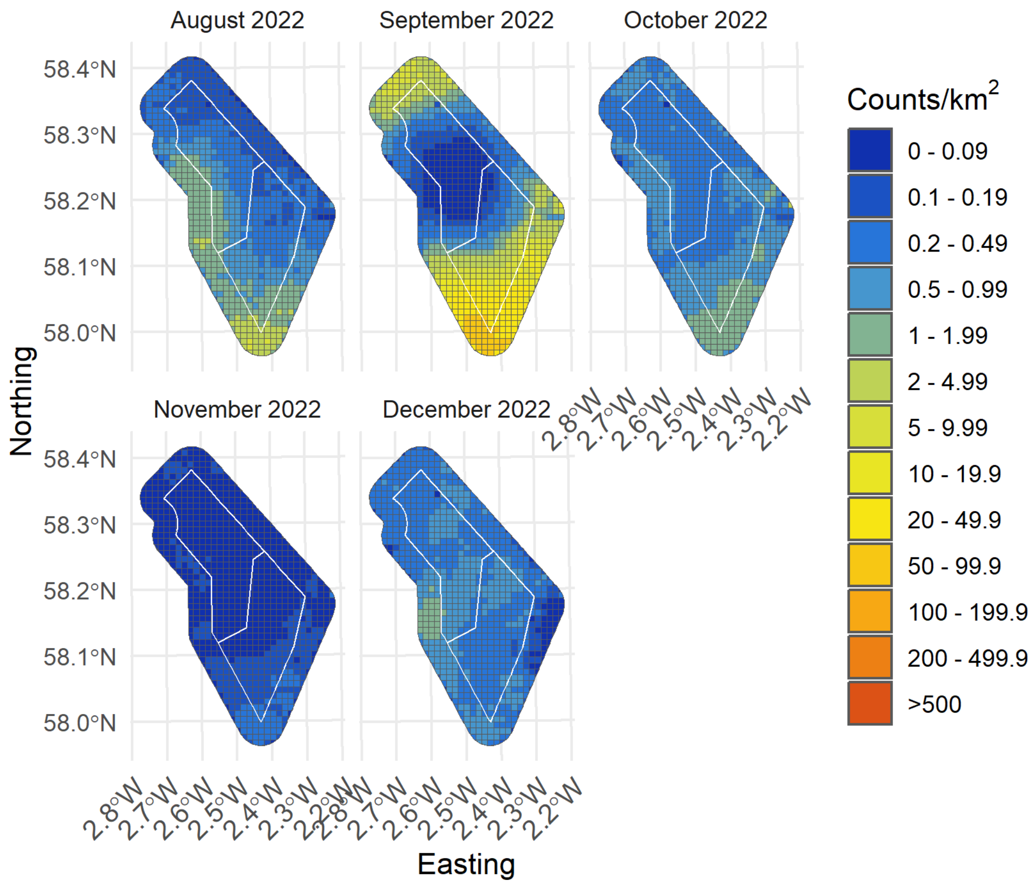


Figure 116. Lower confidence limit of density of all guillemot in the survey area for months with sufficient observations between August and December 2022

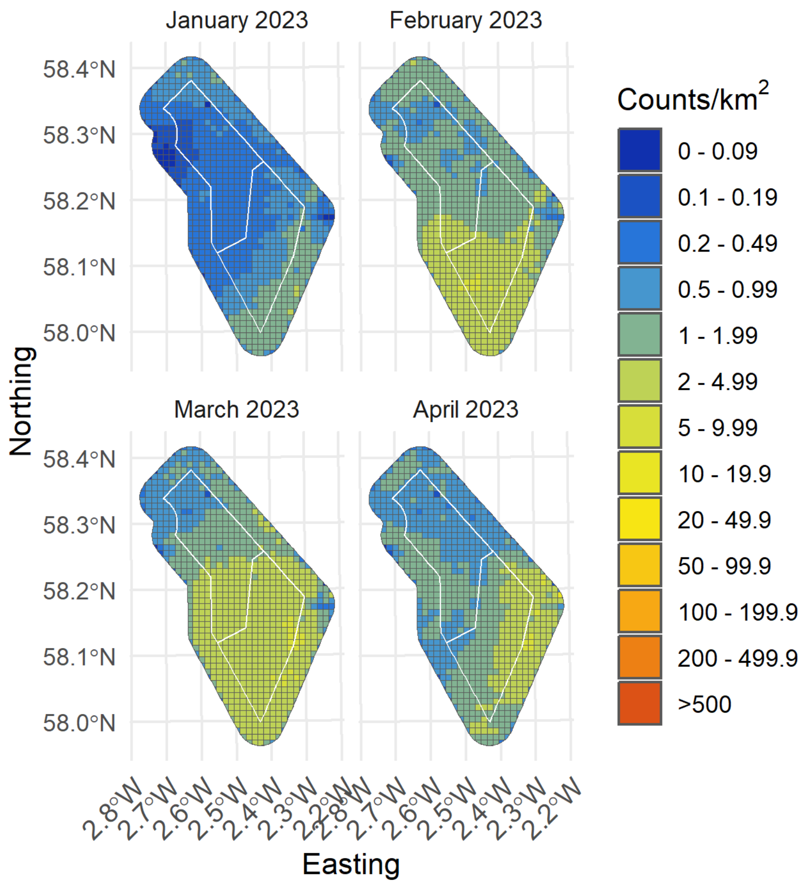


Figure 117. Lower confidence limit of density of all guillemot in the survey area for months with sufficient observations between January and April 2023

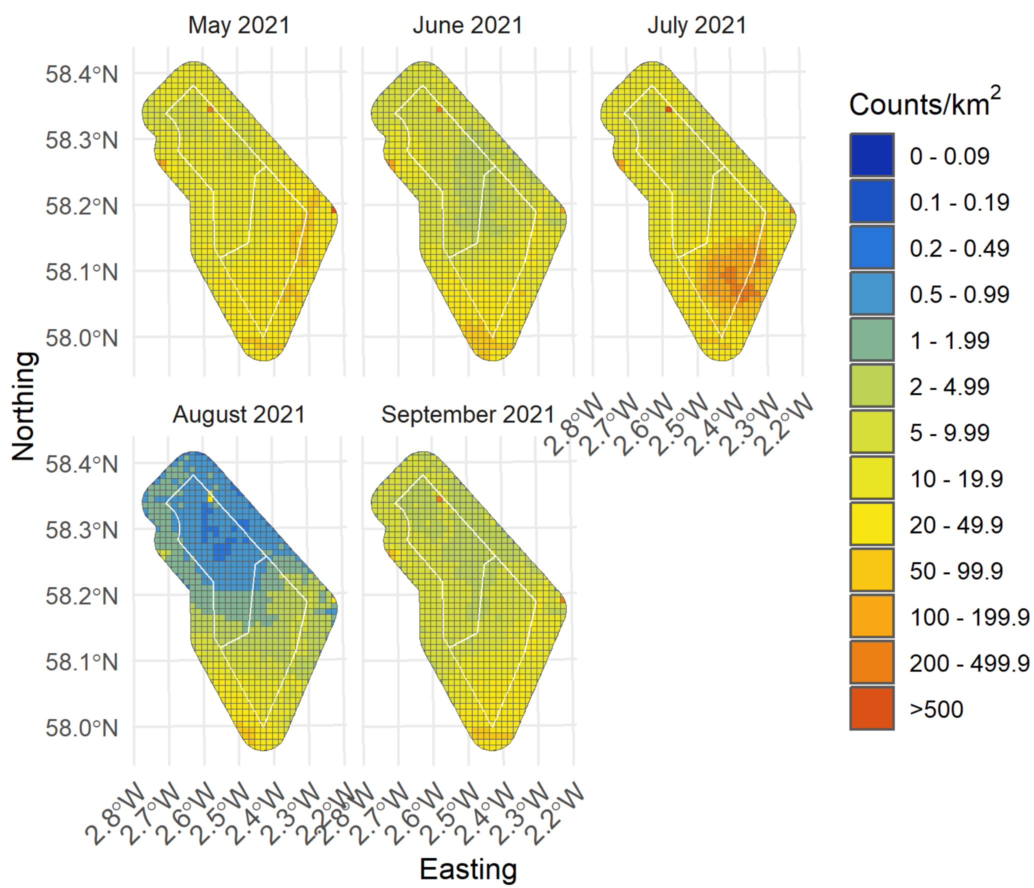


Figure 118. Upper confidence limit of density of all guillemot in the survey area for months with sufficient observations between May and September 2021

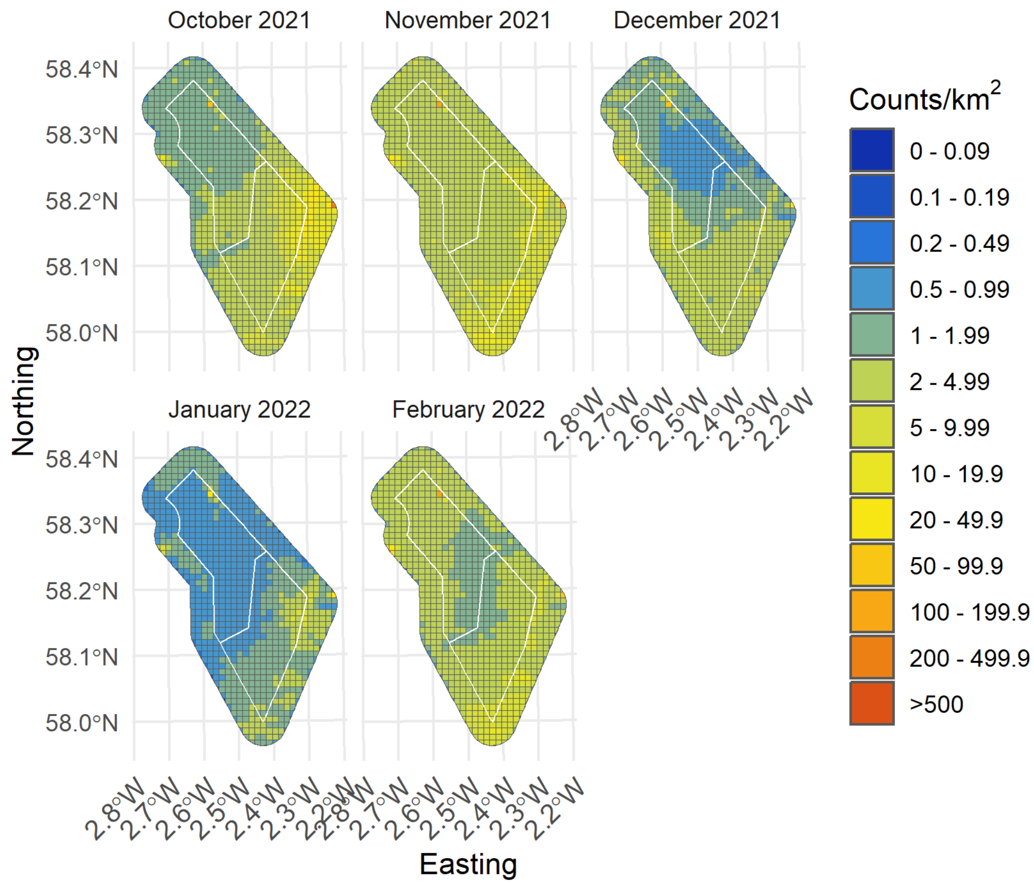


Figure 119. Upper confidence limit of density of all guillemot in the survey area for months with sufficient observations between October 2021 and February 2022

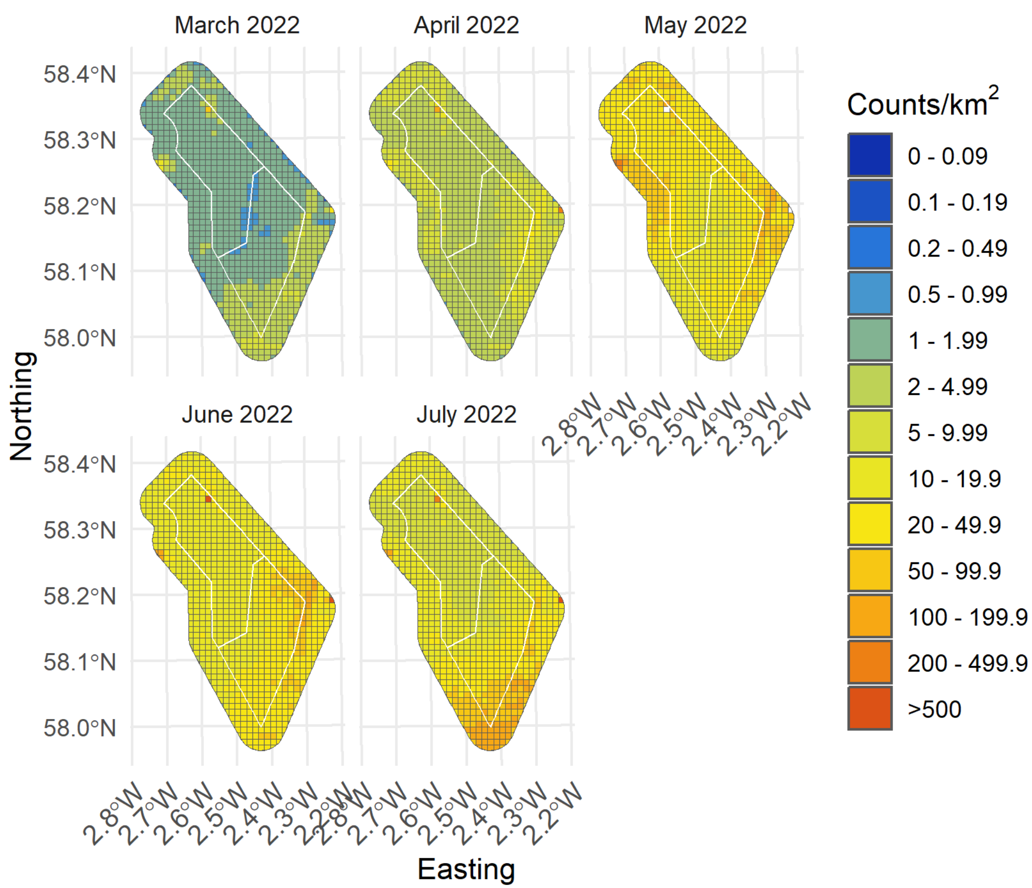


Figure 120. Upper confidence limit of density of all guillemot in the survey area for months with sufficient observations between March and July 2022

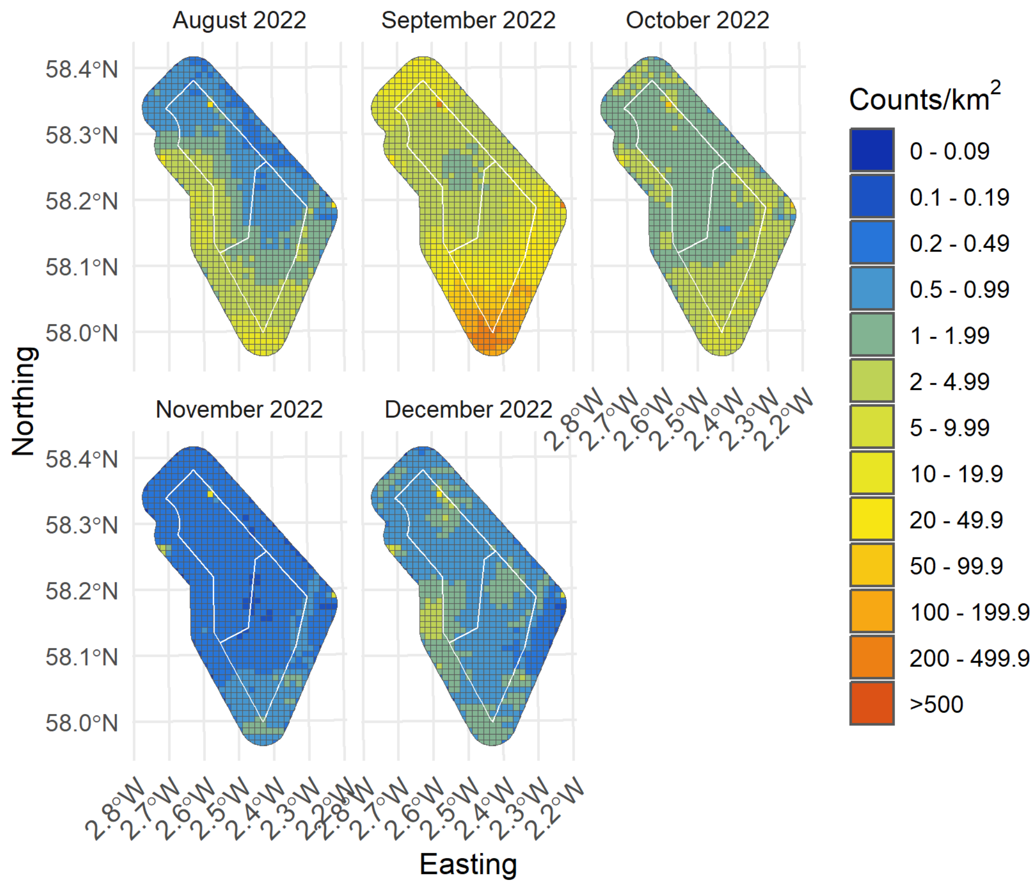


Figure 121. Upper confidence limit of density of all guillemot in the survey area for months with sufficient observations between August and December 2022

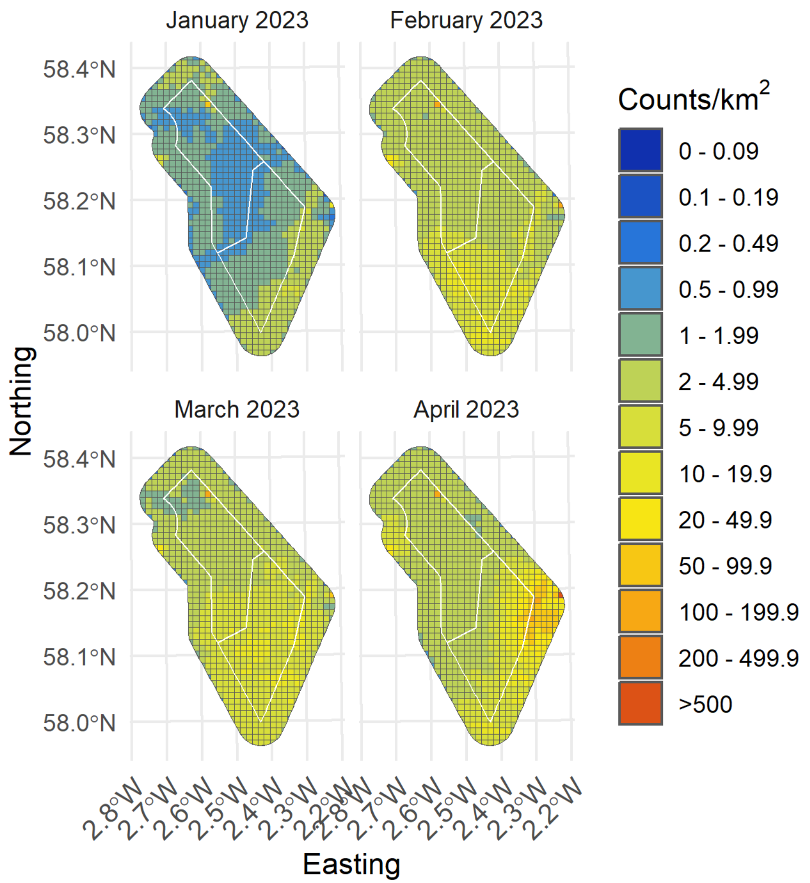


Figure 122. Upper confidence limit of density of all guillemot in the survey area for months with sufficient observations between January and April 2023

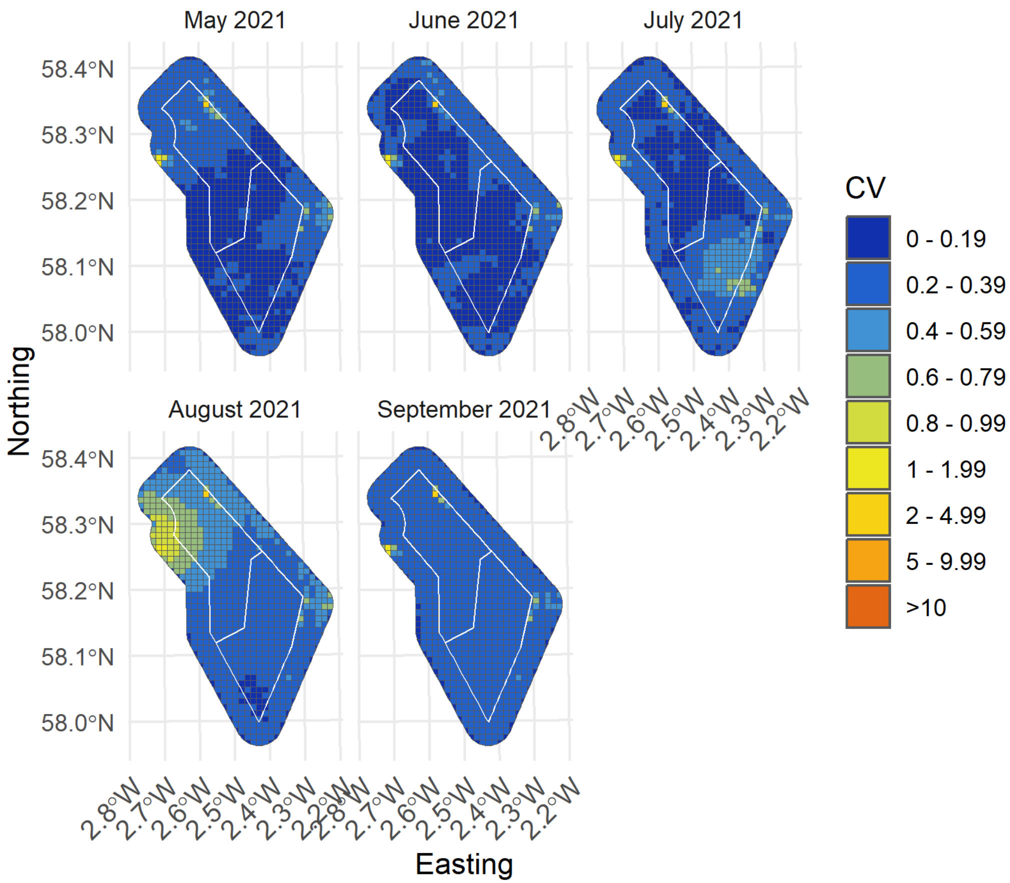


Figure 123. Spatial coefficient of variation of predicted densities of all guillemot from MRSea across the survey area for months with sufficient observations between May and September 2021

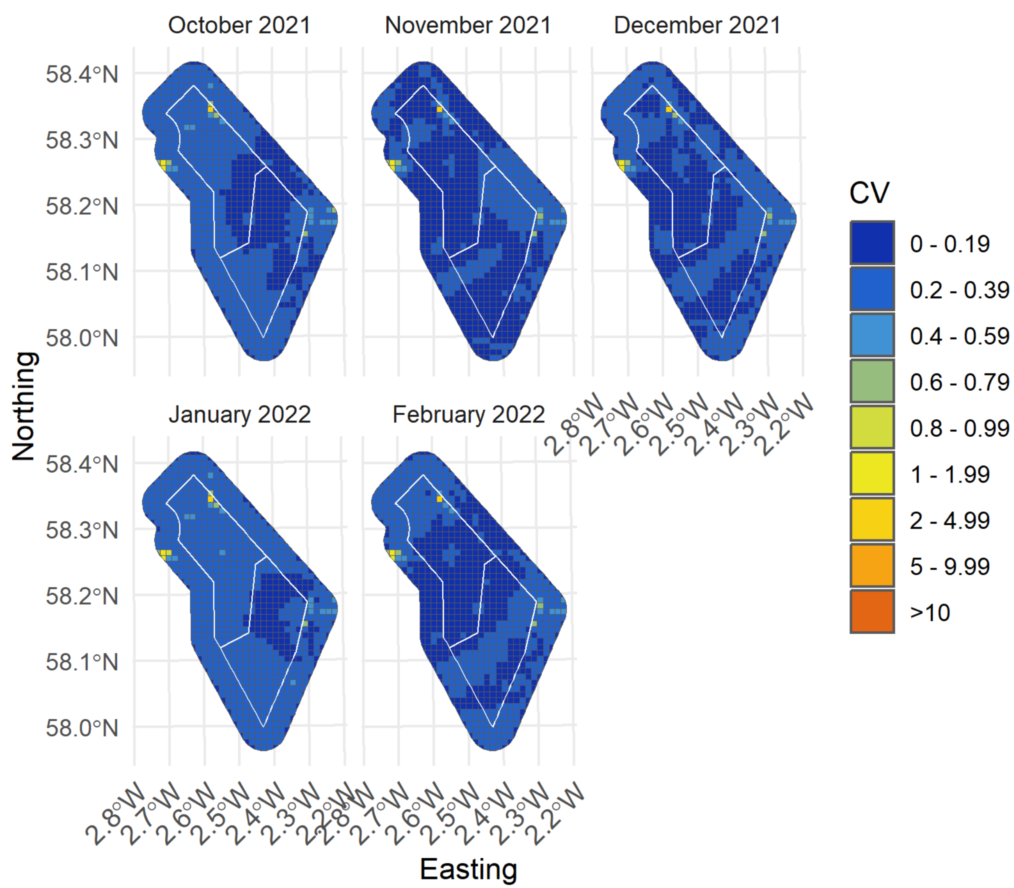


Figure 124. Spatial coefficient of variation of predicted densities of all guillemot from MRSea across the survey area for months with sufficient observations between October 2021 and February 2022

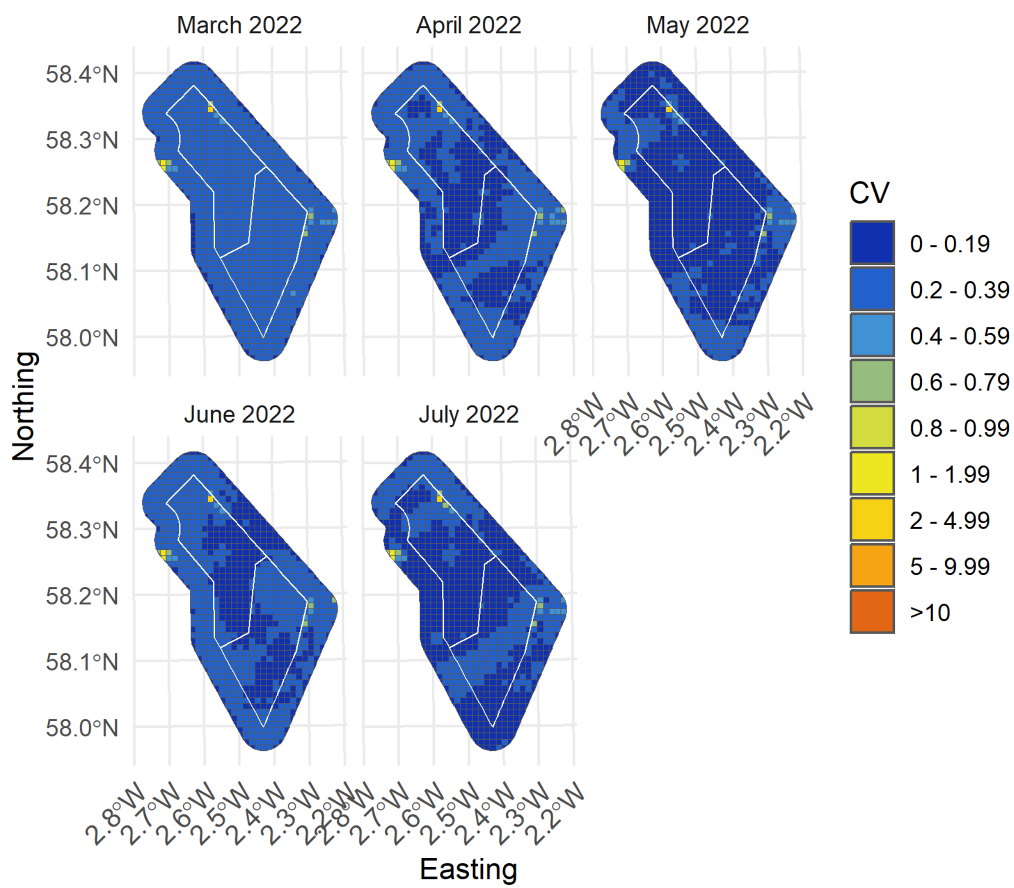


Figure 125. Spatial coefficient of variation of predicted densities of all guillemot from MRSea across the survey area for months with sufficient observations between March and July 2022

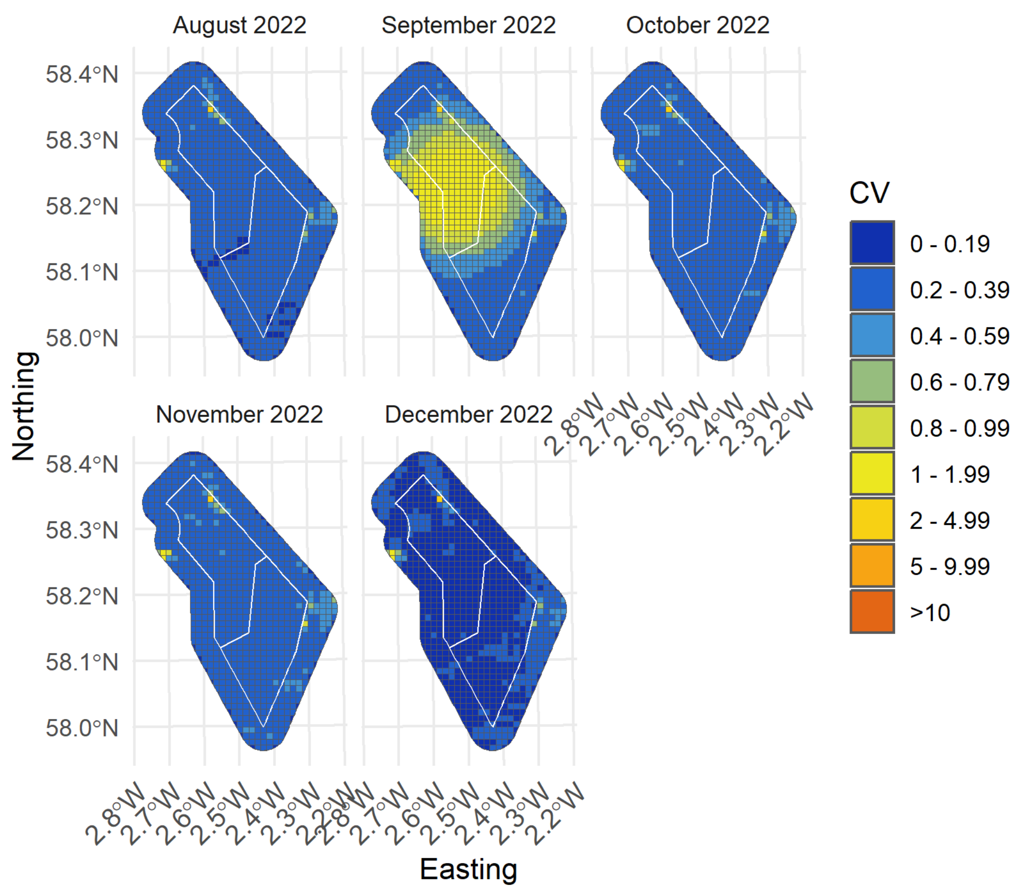


Figure 126. Spatial coefficient of variation of predicted densities of all guillemot from MRSea across the survey area for months with sufficient observations between August and December 2022

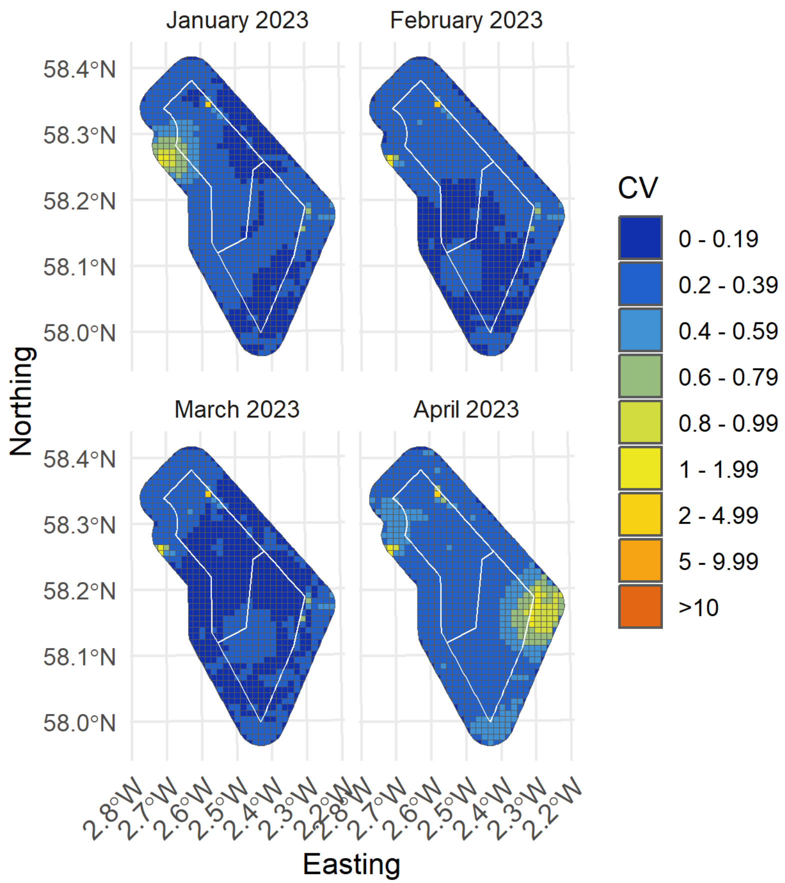


Figure 127. Spatial coefficient of variation of predicted densities of all guillemot from MRSea across the survey area for months with sufficient observations between January and April 2023

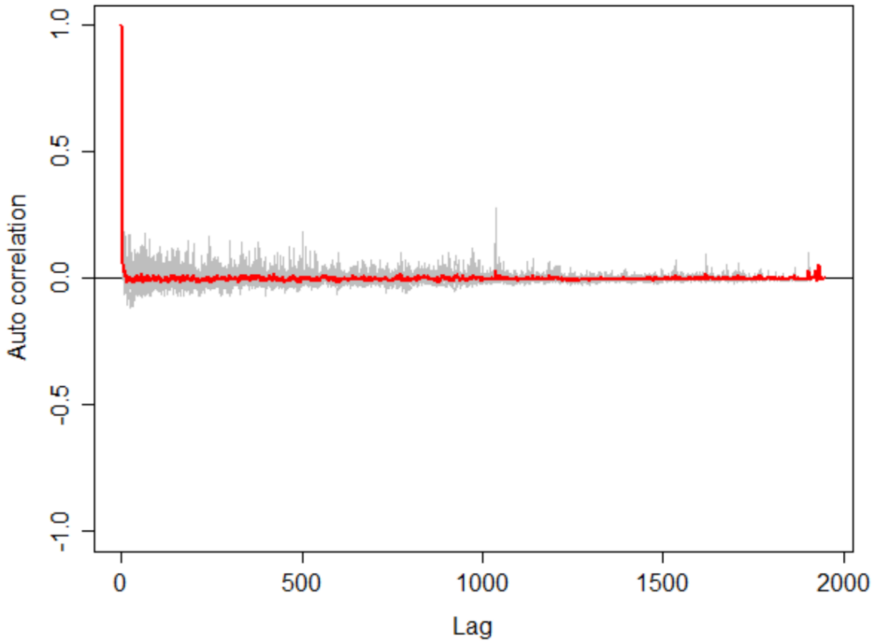


Figure 128. Autocorrelation test for guillemot density surface models when using transect as a blocking feature in MRSea showing no significant correlation. A Runs test on the data prior to using transect as a blocking feature gave a p-value of << 0.0001 (i.e., that the data were significantly autocorrelated when not using a blocking feature)

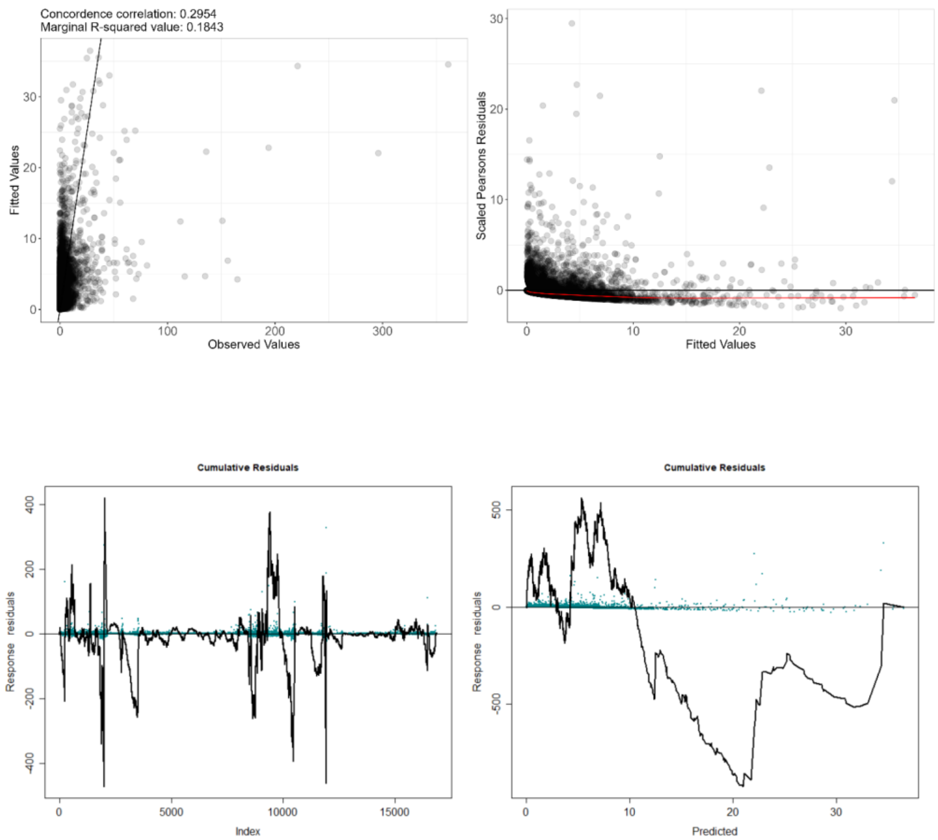


Figure 129. Fitted (MRSea predictions) versus observed counts of all guillemot

Table 16. ANOVA results from the best MRSea model for all guillemot as selected by cross-validation

Variable	Degrees of Freedom	Chi-square	P-value
Bathymetry	3	5.73	0.125
Sandeel Density	3	10.53	0.015
Distance to Colony	4	21.47	<<0.001
Distance to Turbine	3	4.59	0.205
SST (daily)	5	120.24	<<0.001
X/Y (location)	15	183.63	<<0.001

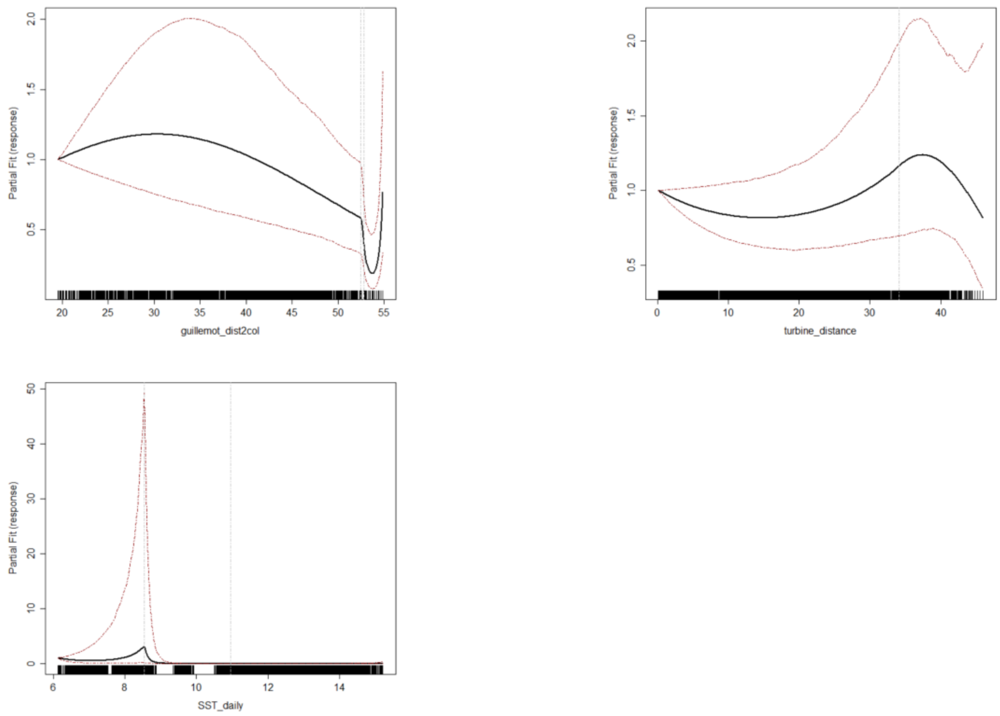


Figure 130. Partial dependence plots for significant variables for all guillemot from MRSea, note that distance to turbine was not a significant covariate in the model but was included for reference. (Clockwise from top left: distance to colony, distance to turbine, daily mean of sea surface temperature.)

3.5 Razorbill

Table 17. Candidate and final covariates for razorbill MRSea model.

Starting model covariates after VIF-based collinearity removal	Final model covariates after removal by SALSA
Bathymetry	Bathymetry
Sandeel Density	Sandeel Density
Distance to colony	Distance to colony
Distance to turbine	Distance to turbine
Daily mean of Sea Surface Temperature	Daily mean of Sea Surface Temperature

Distribution maps generated using MRSea (Figure 130 - 133) suggest that the highest densities of all razorbill tend to be distributed throughout the southern region of the study area during the breeding season (01 April - 15 August). The highest densities occur during the non-breeding season, in September 2022, where birds are widely distributed across the entire region.

Model fit was moderate with a marginal R squared value of 0.07 and root mean squared error of 0.85. Cumulative residuals in the model showed that there was a moderate relationship between predicted and observed values across most of the range of predicted values, but bounded around 0 across the whole (Figure 147).

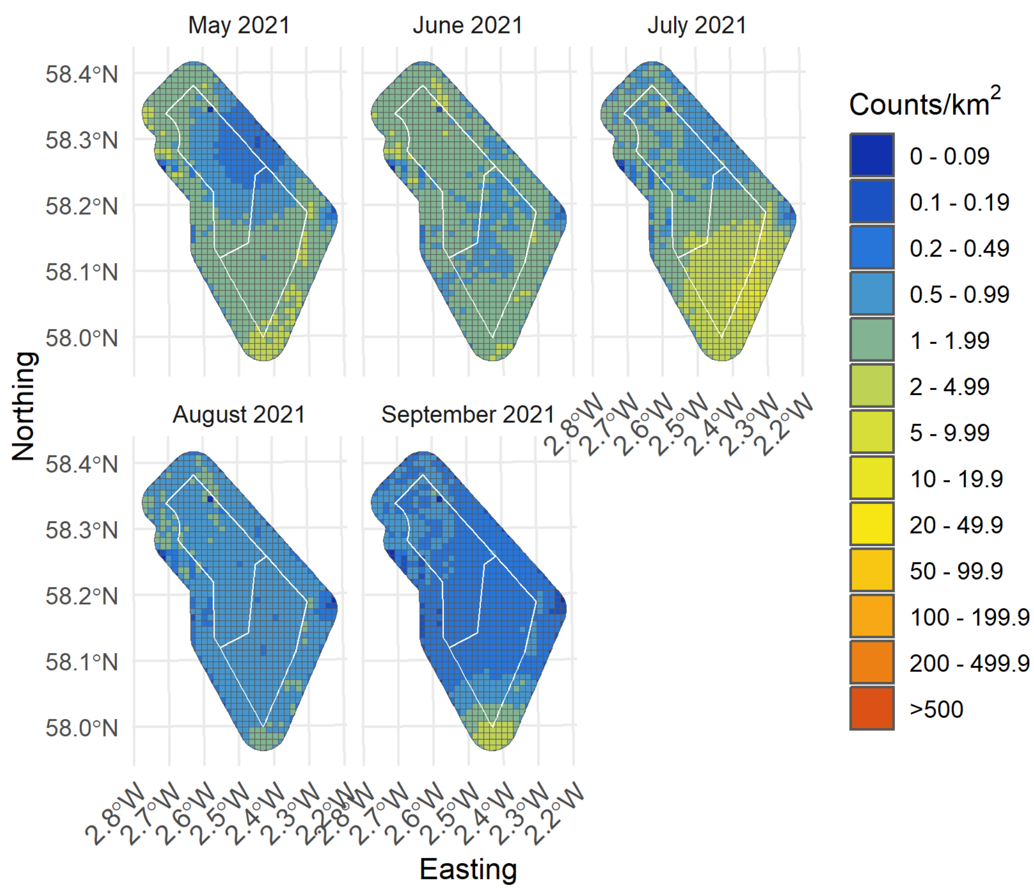


Figure 131. Median density of all razorbill in the survey area for months with sufficient observations between May and September 2021

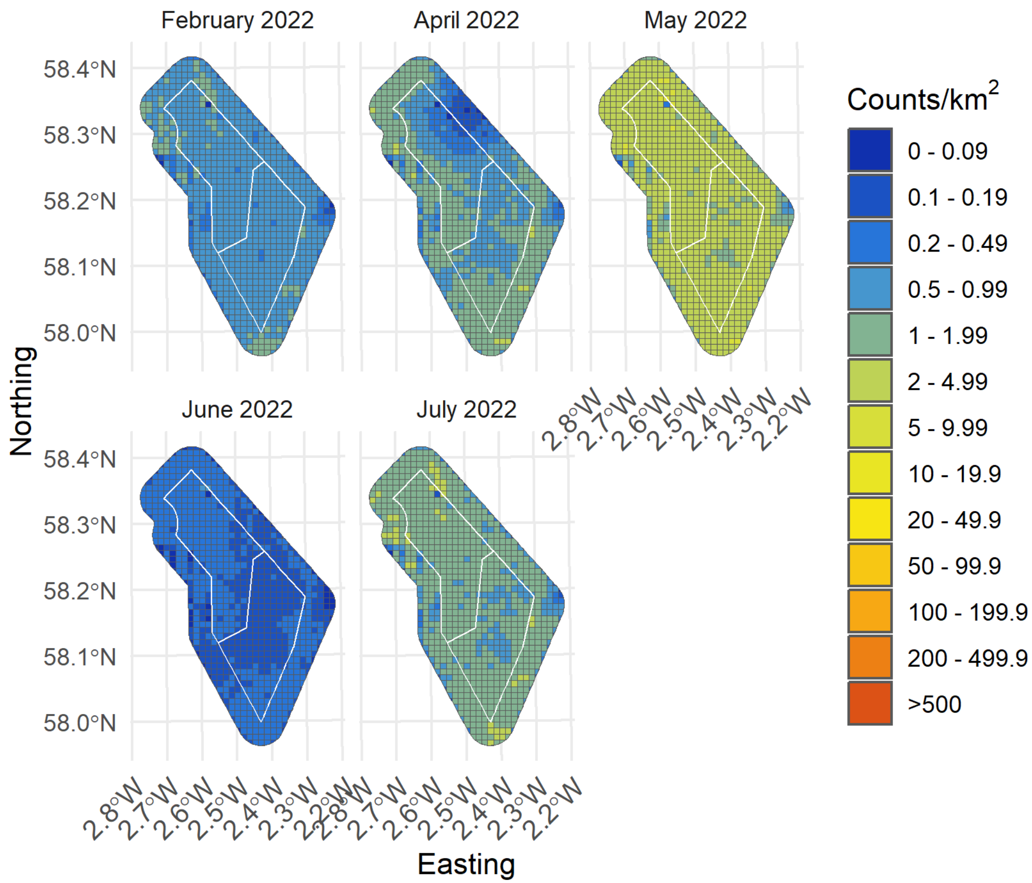


Figure 132. Median density of all razorbill in the survey area for months with sufficient observations between February and July 2022

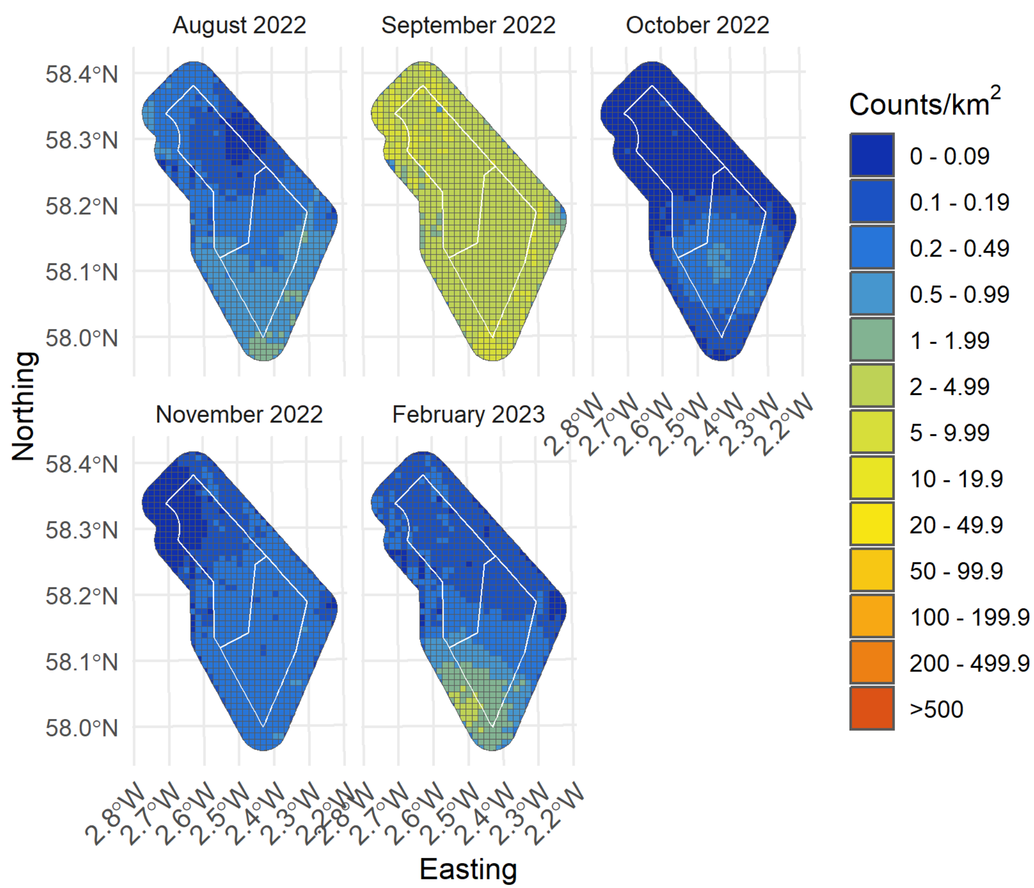


Figure 133. Median density of all razorbill in the survey area for months with sufficient observations between August 2022 and February 2023

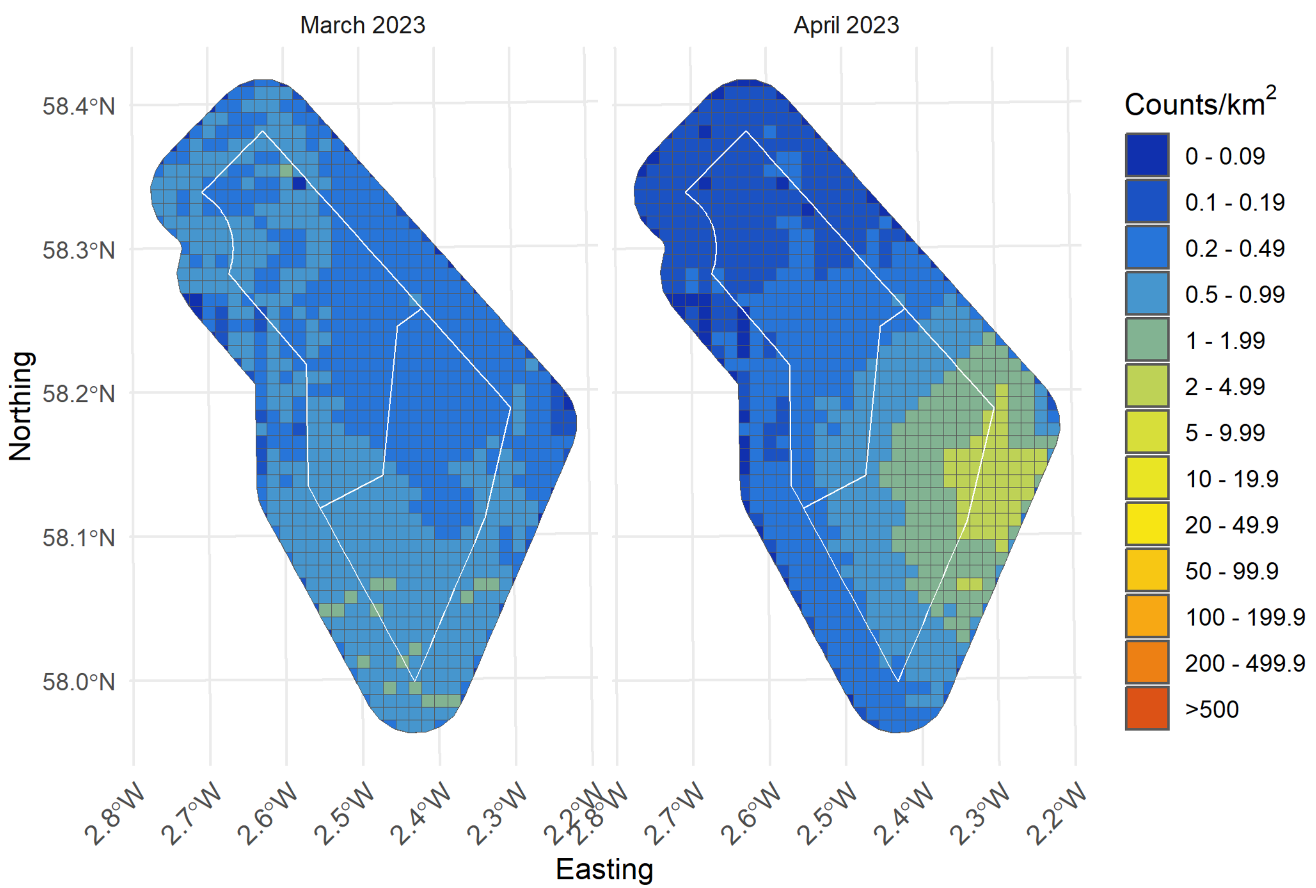


Figure 134. Median density of all razorbill in the survey area for months with sufficient observations between March and April 2023

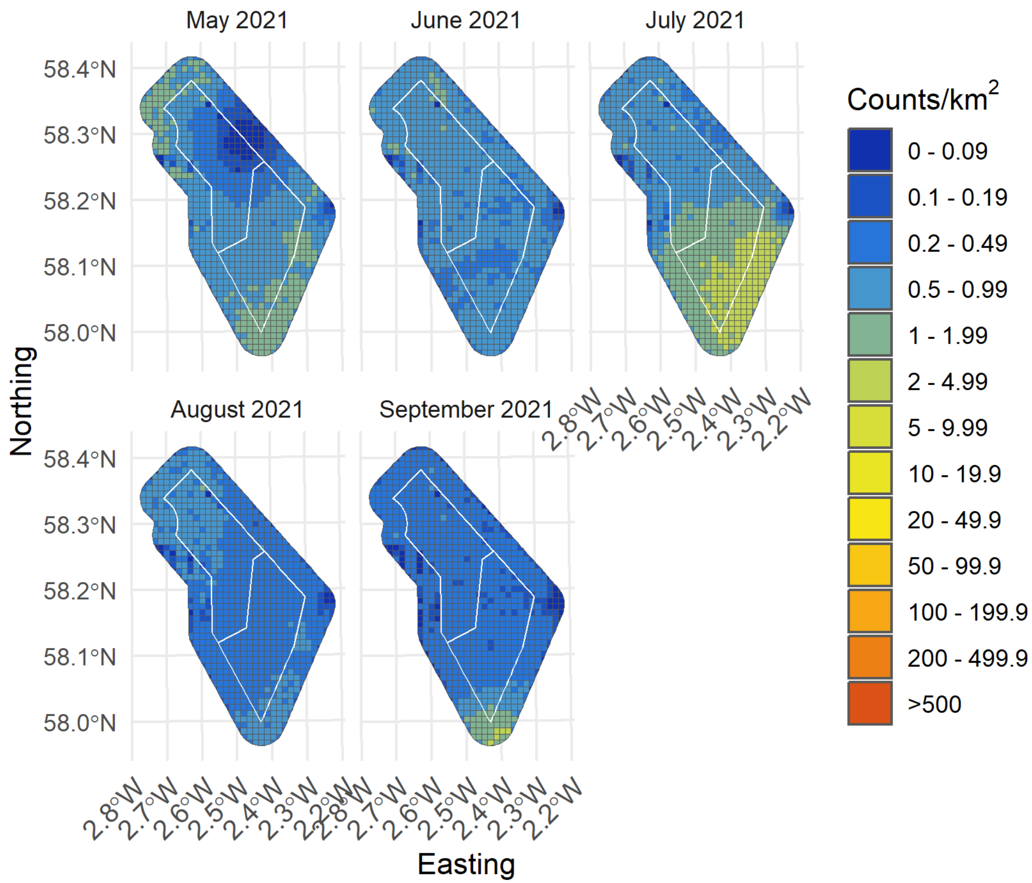


Figure 135. Lower confidence limit of density of all razorbill in the survey area for months with sufficient observations between May and September 2021

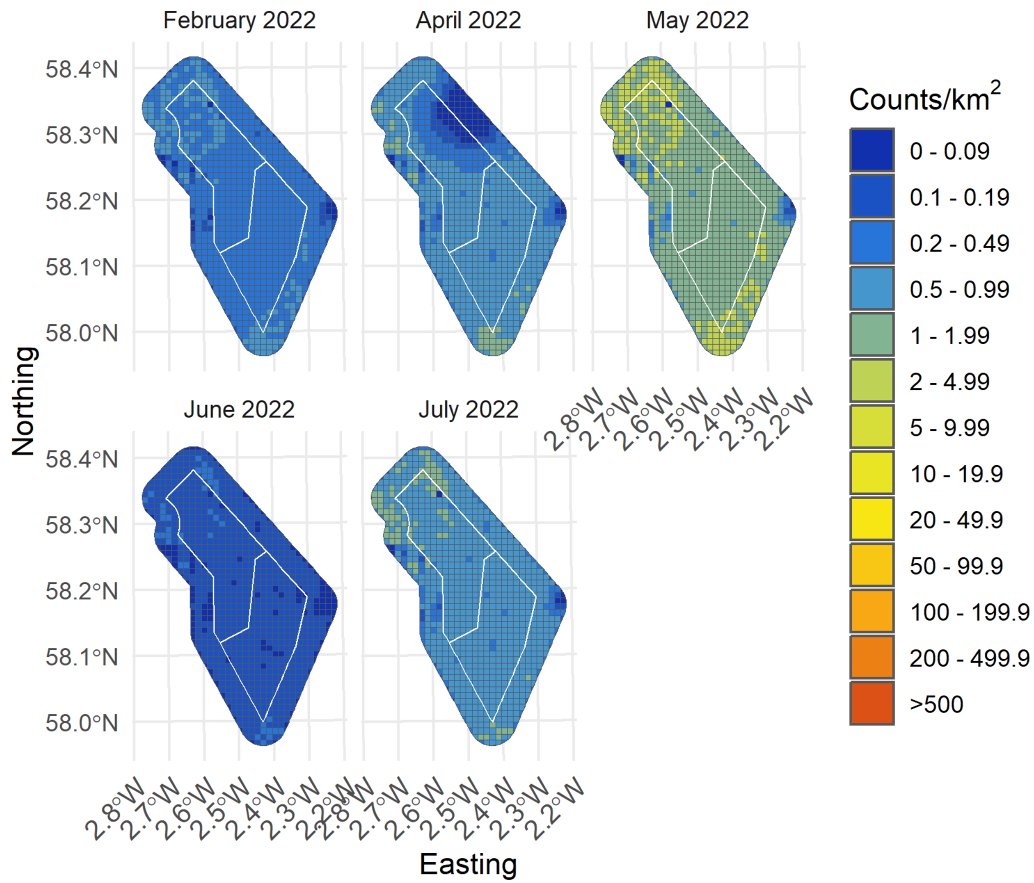


Figure 136. Lower confidence limit of density of all razorbill in the survey area for months with sufficient observations between February and July 2022

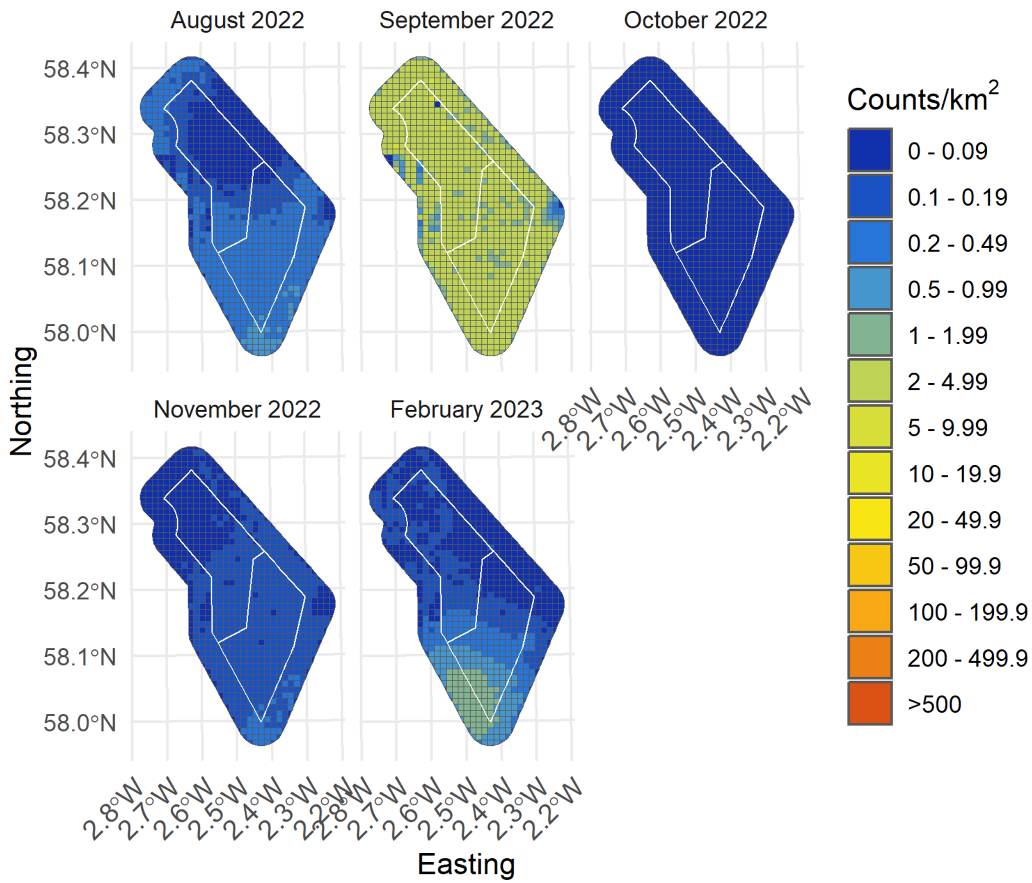


Figure 137. Lower confidence limit of density of all razorbill in the survey area for months with sufficient observations between August 2022 and February 2023

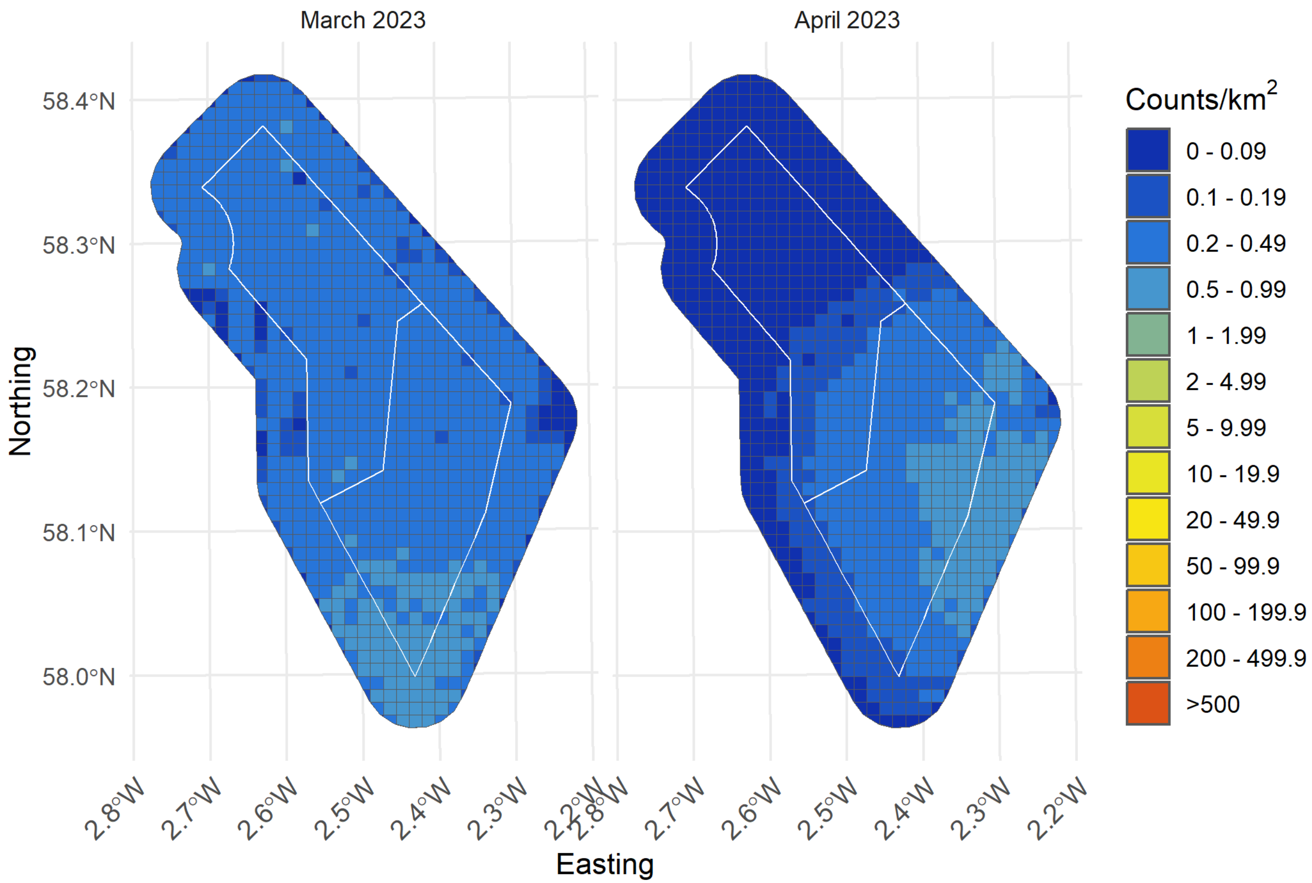


Figure 138. Lower confidence limit of density of all razorbill in the survey area for months with sufficient observations between March and April 2023

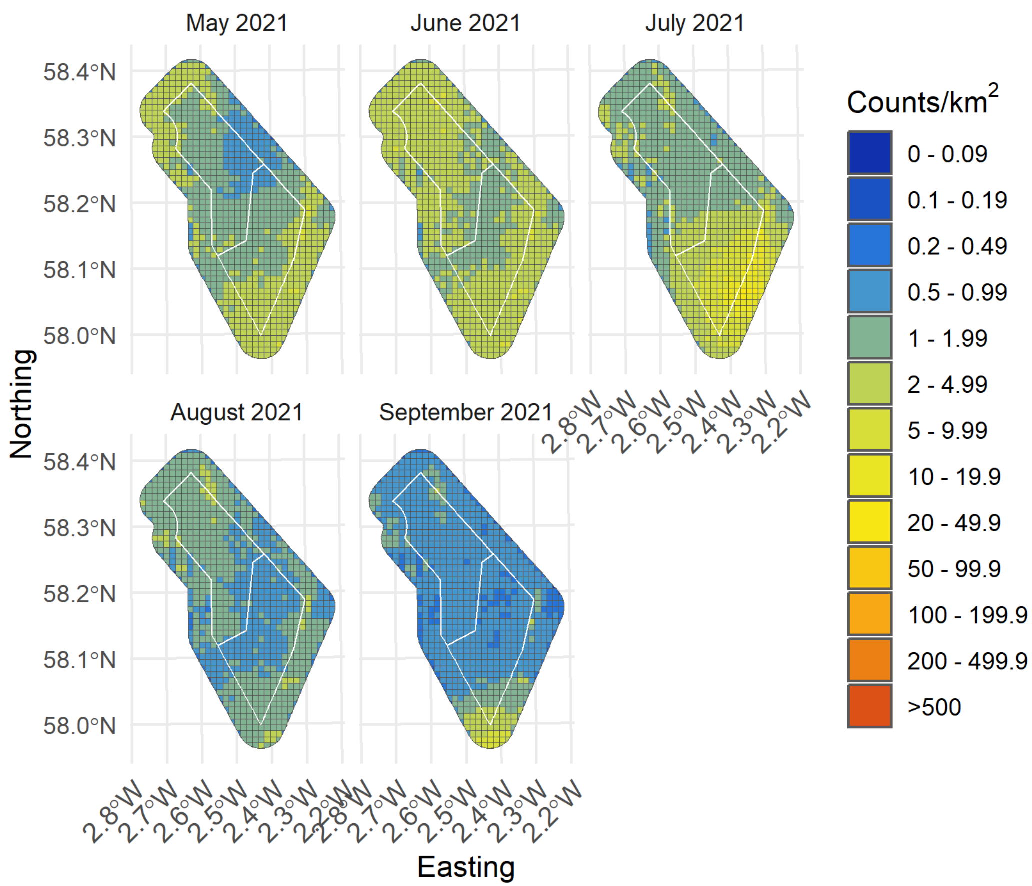


Figure 139. Upper confidence limit of density of all razorbill in the survey area for months with sufficient observations between May and September 2021

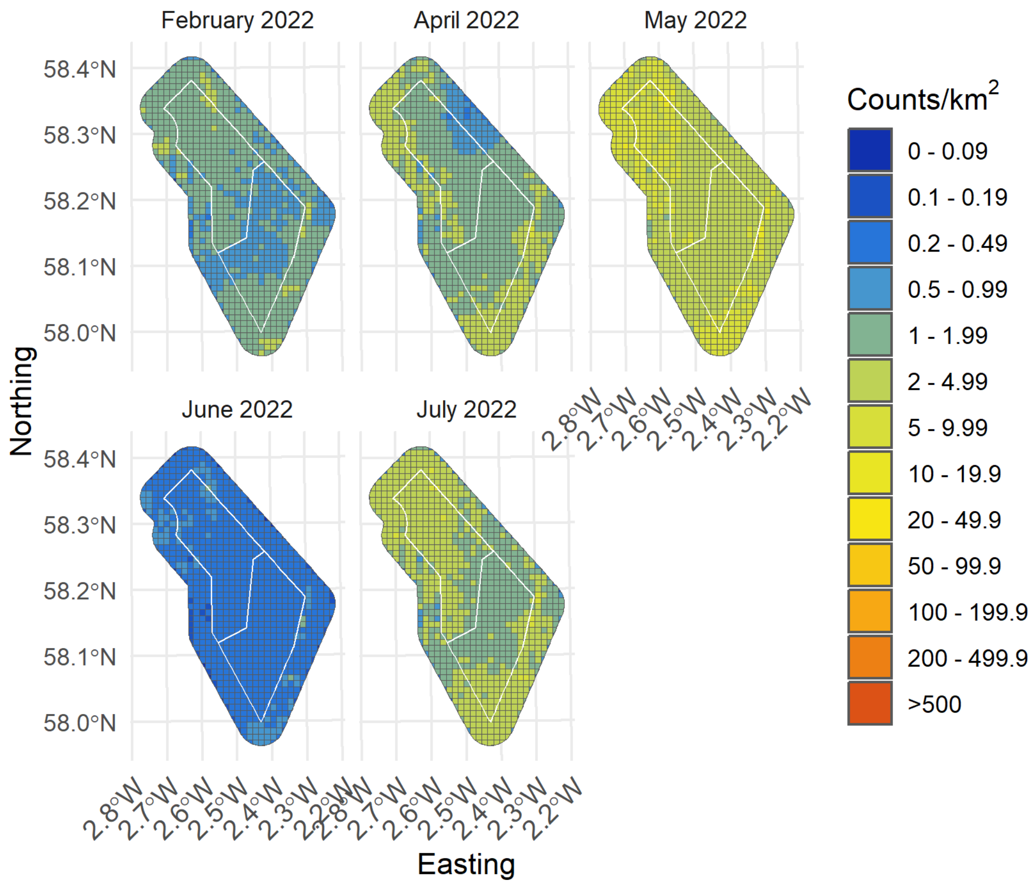


Figure 140. Upper confidence limit of density of all razorbill in the survey area for months with sufficient observations between February and July 2022

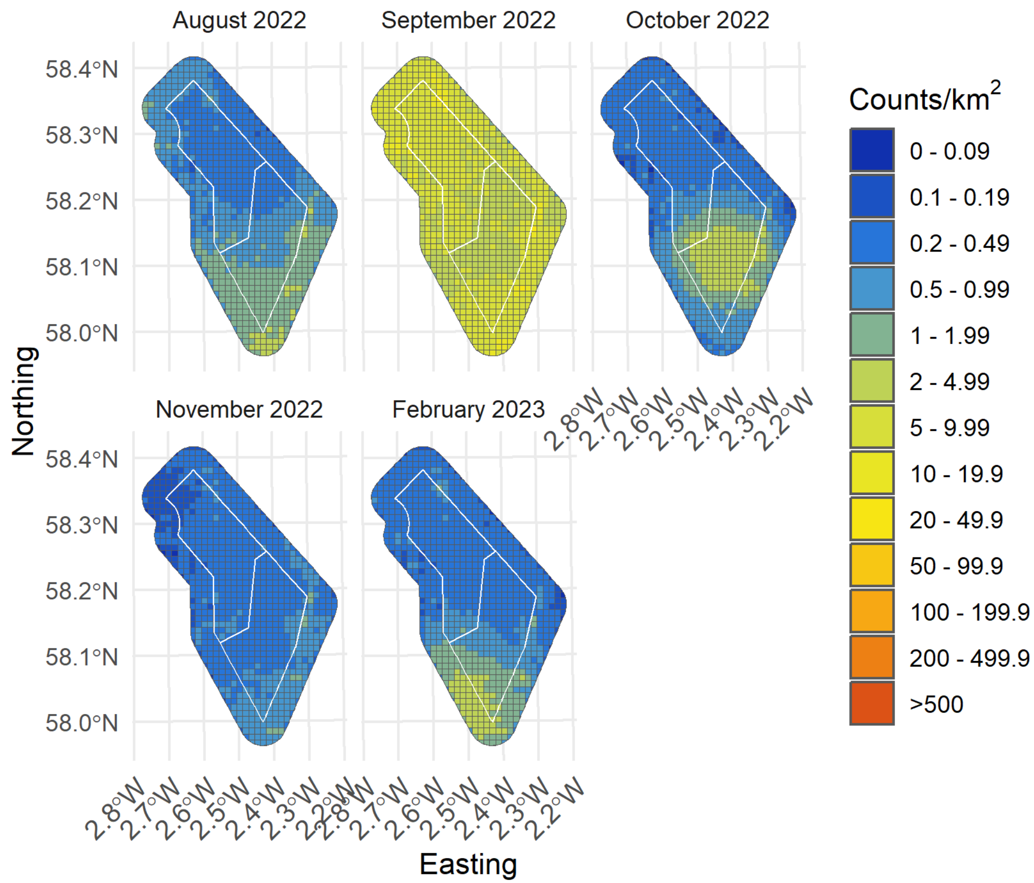


Figure 141. Upper confidence limit of density of all razorbill in the survey area for months with sufficient observations between August 2022 and February 2023

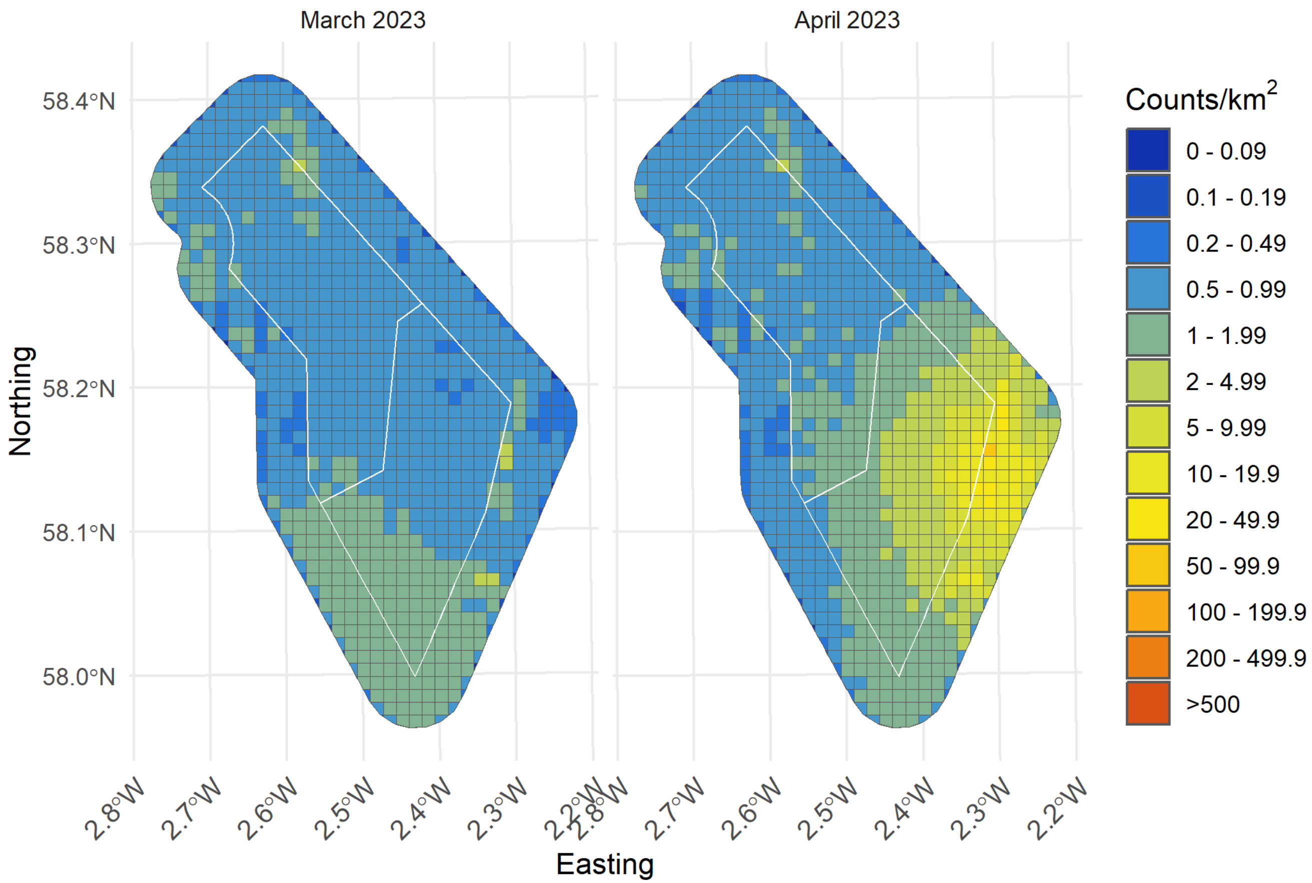


Figure 142. Upper confidence limit of density of all razorbill in the survey area for months with sufficient observations between March and April 2023

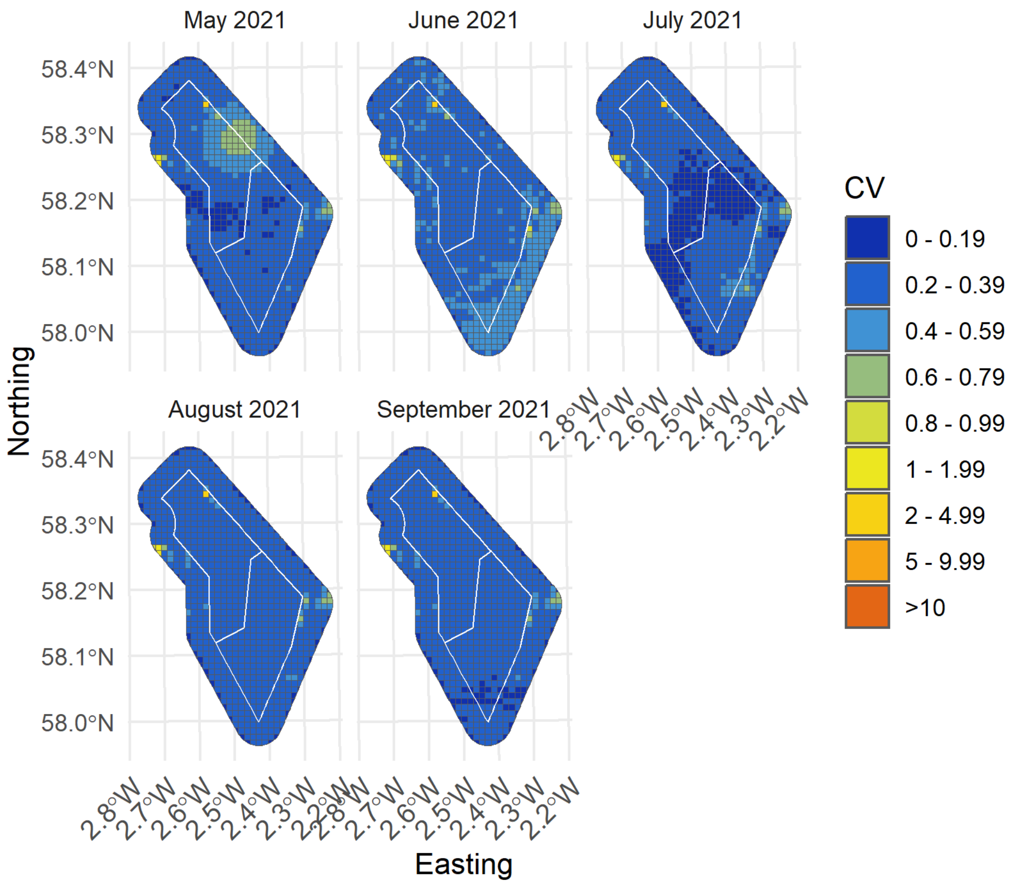


Figure 143. Spatial coefficient of variation of predicted densities of all razorbill from MRSea across the survey area for months with sufficient observations between May and September 2021

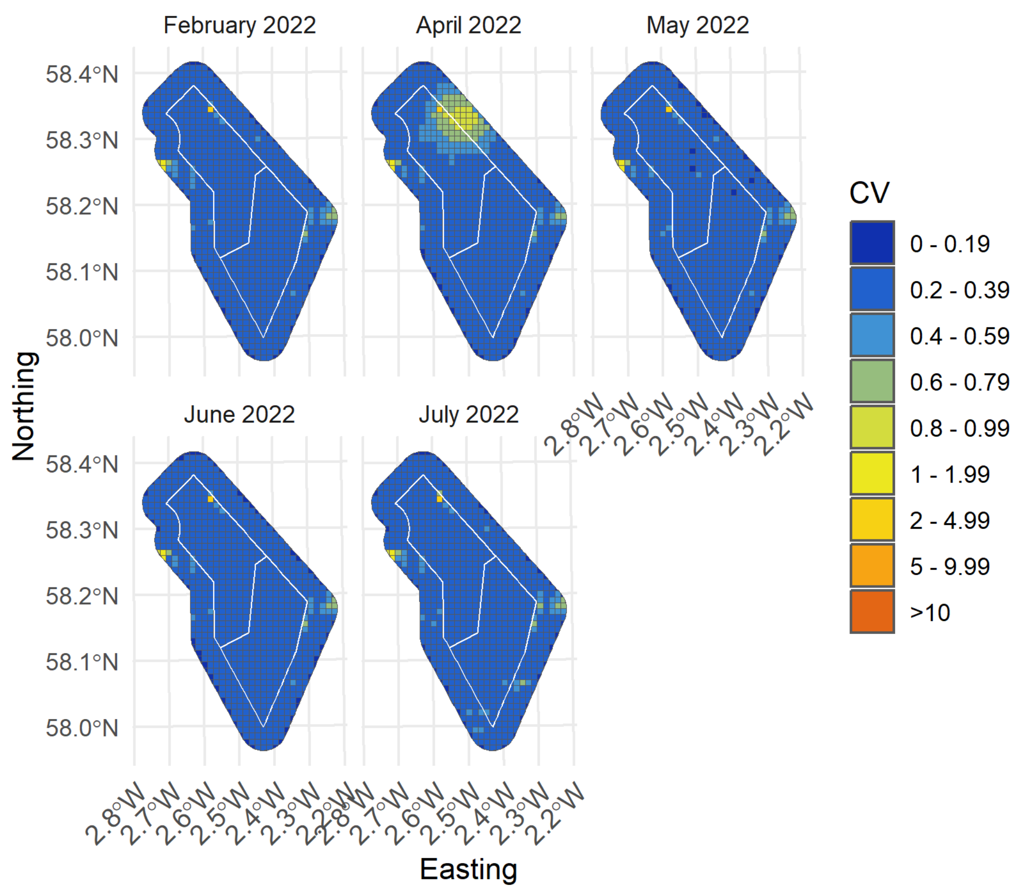


Figure 144. Spatial coefficient of variation of predicted densities of all razorbill from MRSea across the survey area for months with sufficient observations between February and July 2022

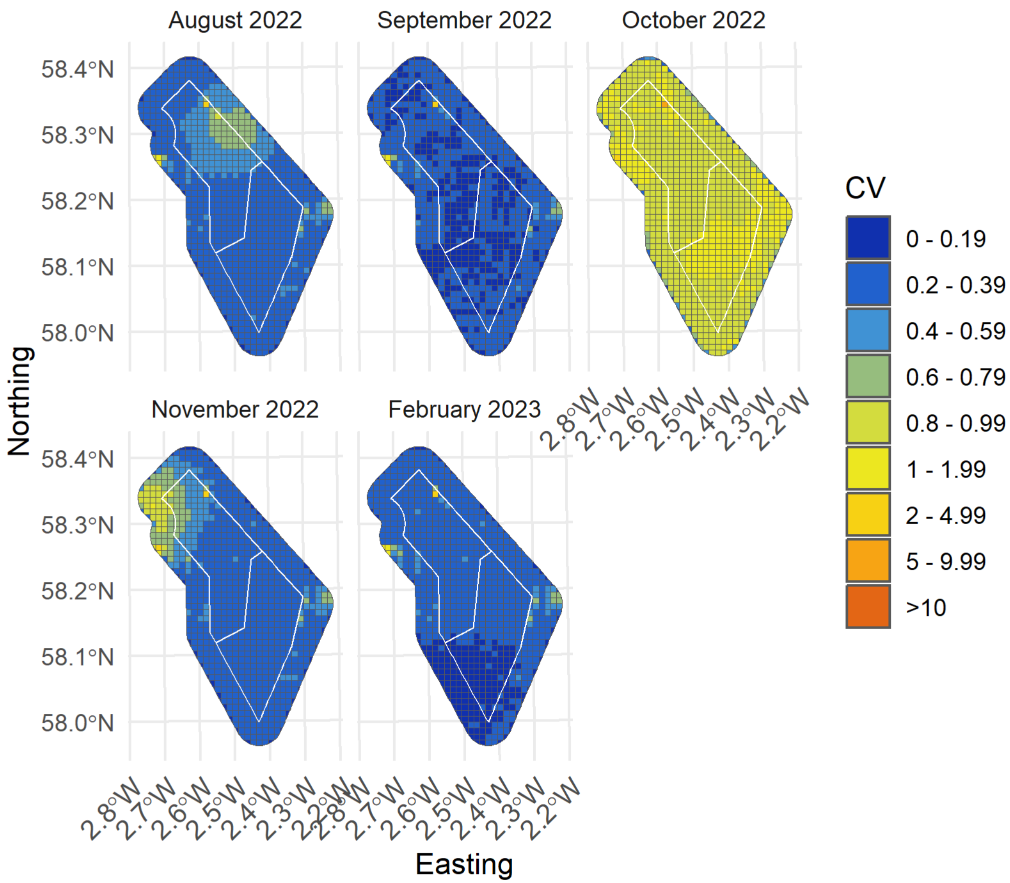


Figure 145. Spatial coefficient of variation of predicted densities of all razorbill from MRSea across the survey area for months with sufficient observations between August 2022 and February 2023

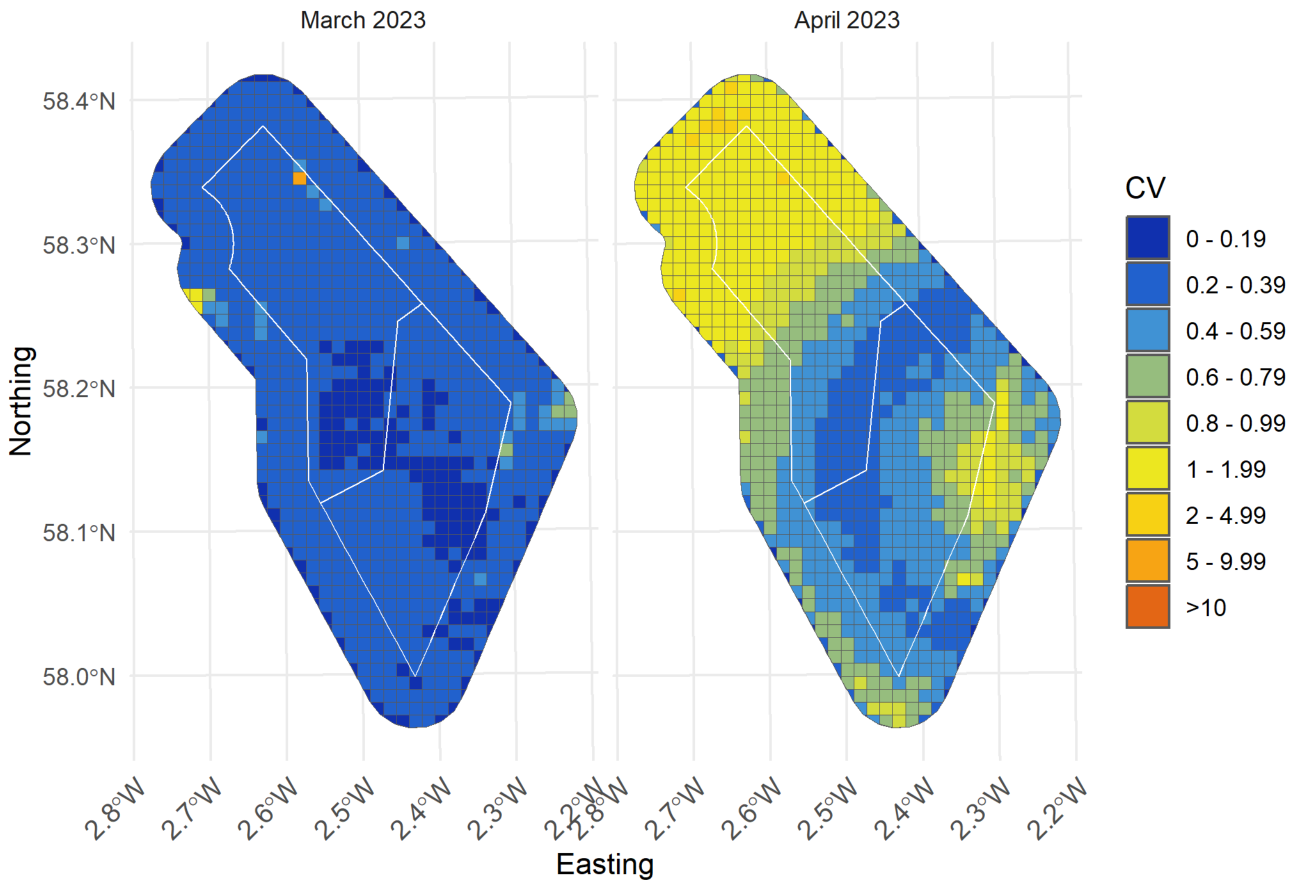


Figure 146. Spatial coefficient of variation of predicted densities of all razorbill from MRSea across the survey area for months with sufficient observations between March and April 2023

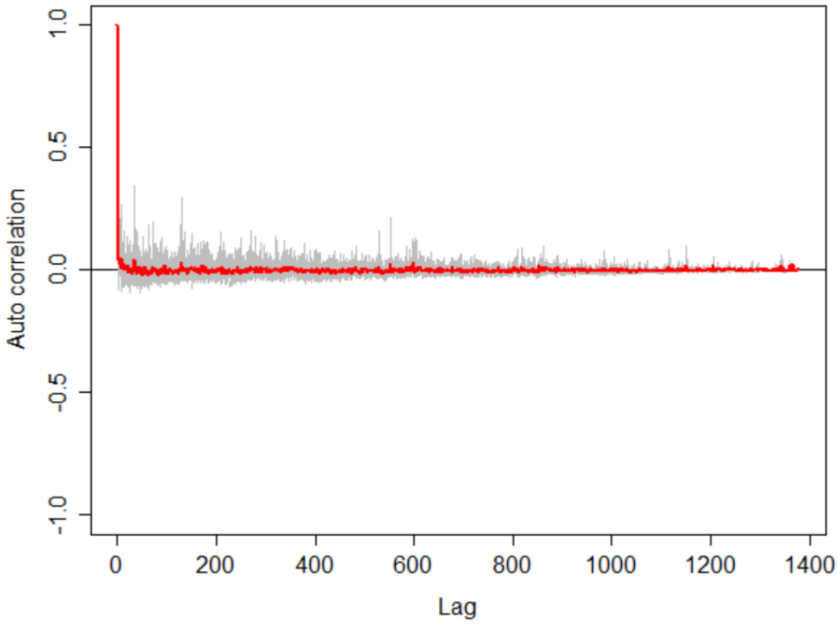


Figure 147. Autocorrelation test for razorbill density surface models when using transect as a blocking feature in MRSea showing no significant correlation. A Runs test on the data prior to using transect as a blocking feature gave a p-value of << 0.0001 (i.e., that the data were significantly autocorrelated when not using a blocking feature)

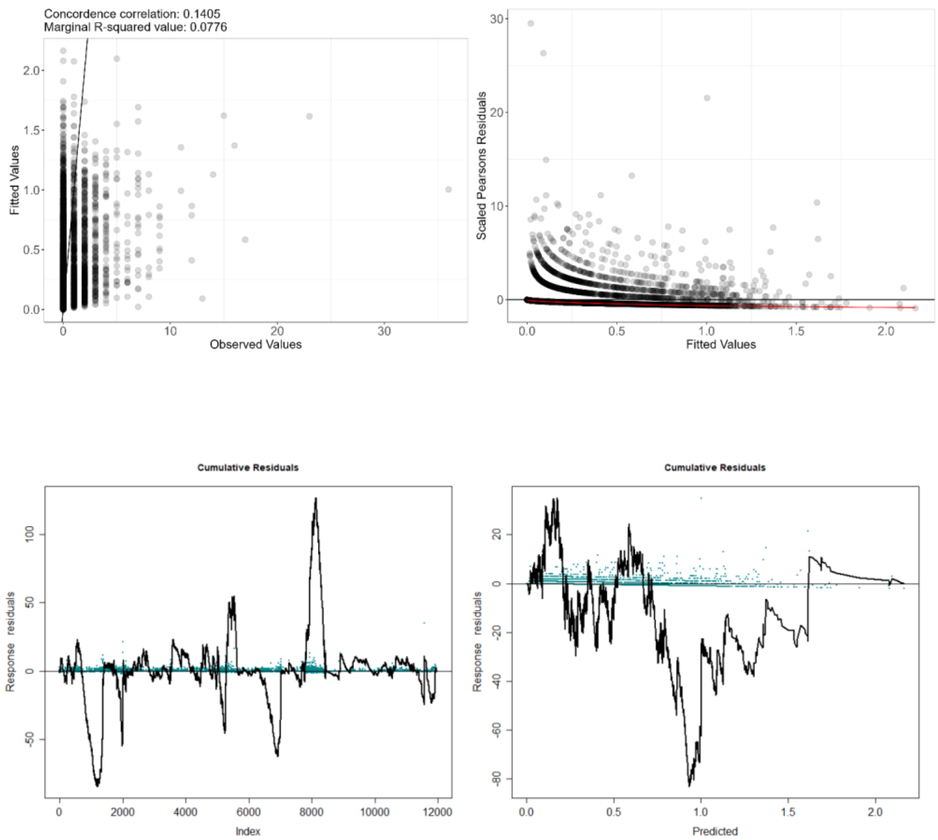


Figure 148. Fitted (MRSea predictions) versus observed counts of all razorbill

Table 18. ANOVA results from the best MRSea model for all razorbill as selected by cross-validation

Variable	Degrees of Freedom	Chi-square	P-value
Aspect	3	2.26	0.52
Slope	3	3.66	0.3
Bathymetry	3	5.12	0.163
Sandeel Density	3	3.52	0.318
Distance to Colony	3	16.79	<<0.001
Distance to Turbine	3	14.68	0.002
X/Y (location)	10	97.17	<<0.001

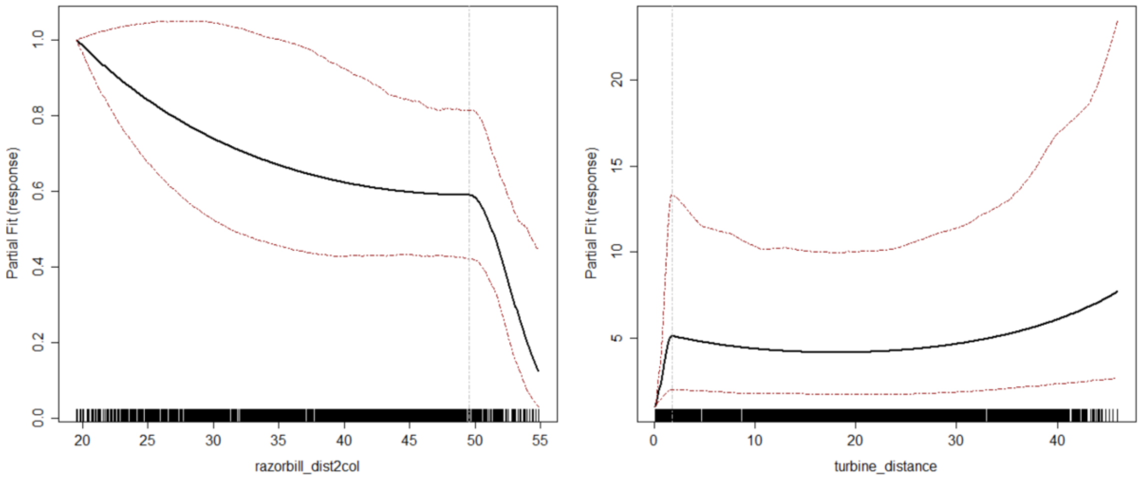


Figure 149. Partial dependence plots for significant variables for all razorbill from MRSea

3.6 Puffin

Table 19. Candidate and final covariates for puffin MRSea model

Starting model covariates after VIF-based collinearity removal	Final model covariates after removal by SALSA
Aspect	Aspect
Slope	Slope
Bathymetry	Bathymetry
Sandeel Density	Sandeel Density
Distance to colony	Distance to colony
Distance to turbine	Distance to turbine
Standard deviation of Sea Surface Temperature	Standard deviation of Sea Surface Temperature

Distribution maps generated using MRSea (Figure 149-151) suggest that puffin are distributed more generally towards the eastern half of the study area during the breeding season (01 April - 15 August). In the non-breeding season, with the exception of September - October 2021 and September 2022, puffins are mostly absent from the study area (less than 20 observations). The highest densities were observed in the center and southern regions of the study area in August 2021.

Model fit was better for Puffin than for most species, with a marginal R squared value of 0.17 and root mean squared error of 0.77. Cumulative residuals in the model showed that there was a moderate relationship between predicted and observed values across most of the range of predicted value, but the model tended to underpredict for densities above 1.75 birds/km2 (Figure 160).

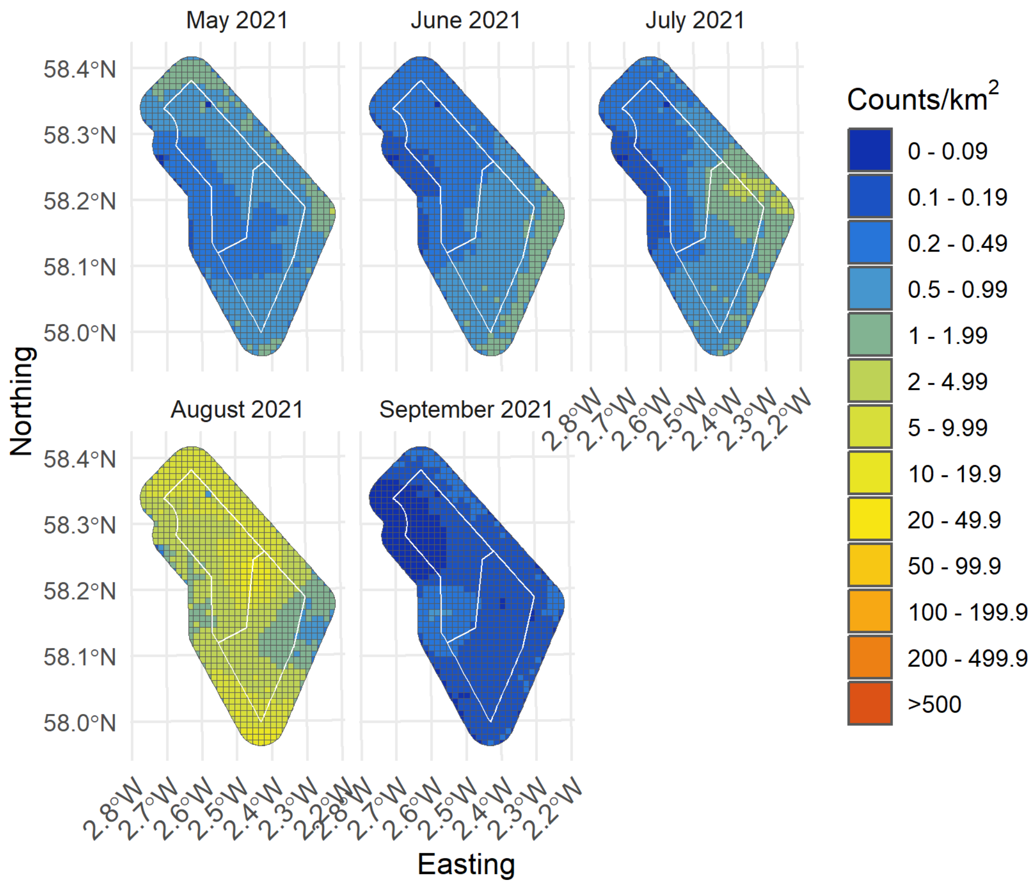


Figure 150. Median density of all puffin in the survey area for months with sufficient observations between May and September 2021

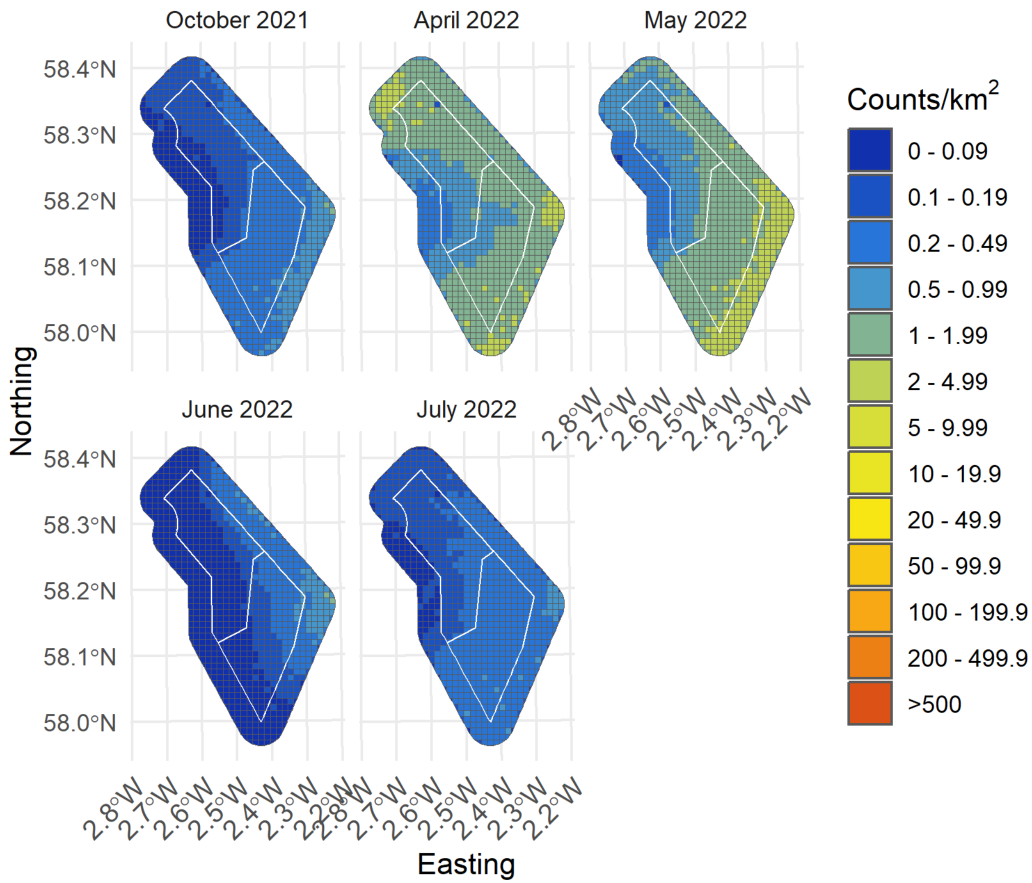


Figure 151. Median density of all puffin in the survey area for months with sufficient observations between October 2021 and July 2022

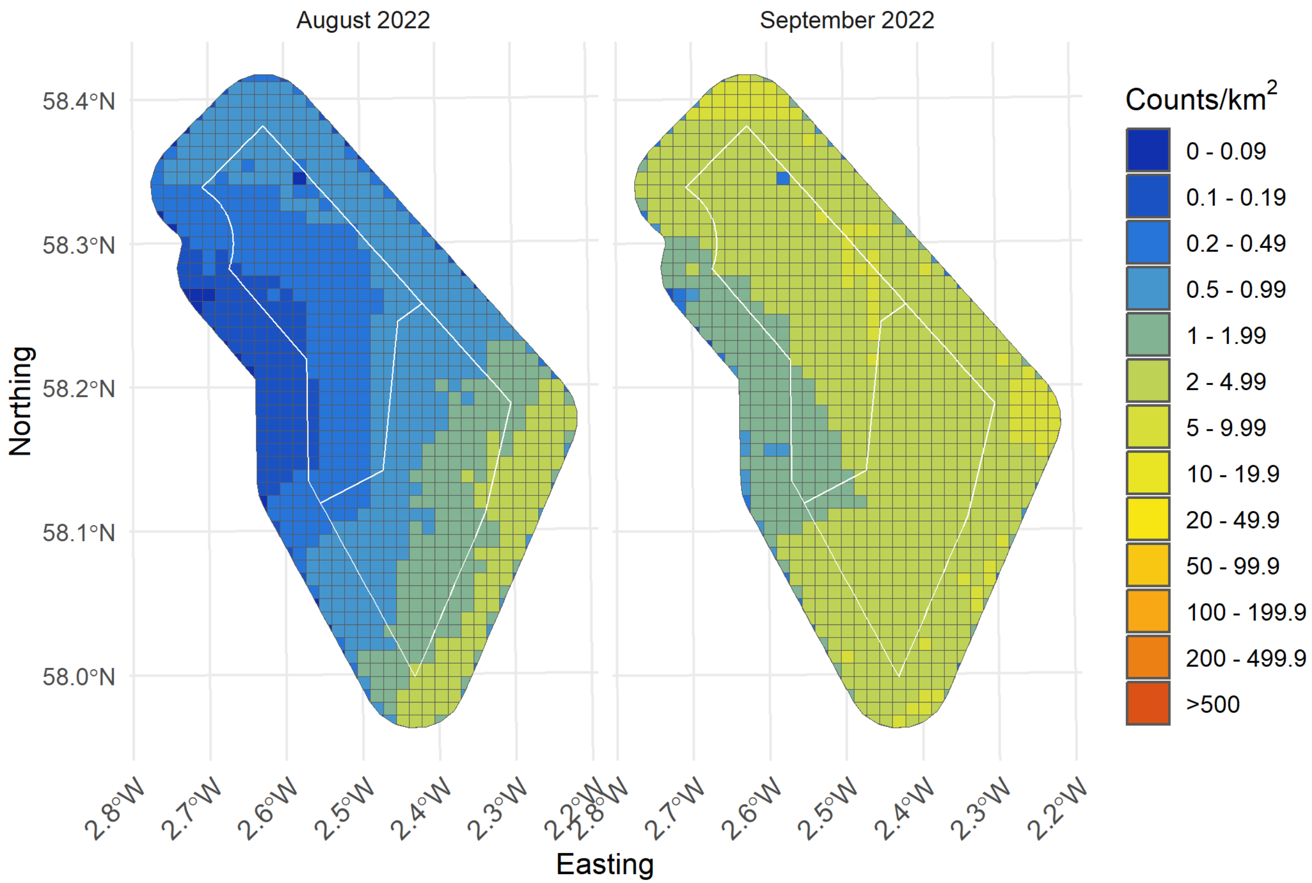


Figure 152. Median density of all puffin in the survey area for months with sufficient observations between August and September 2022

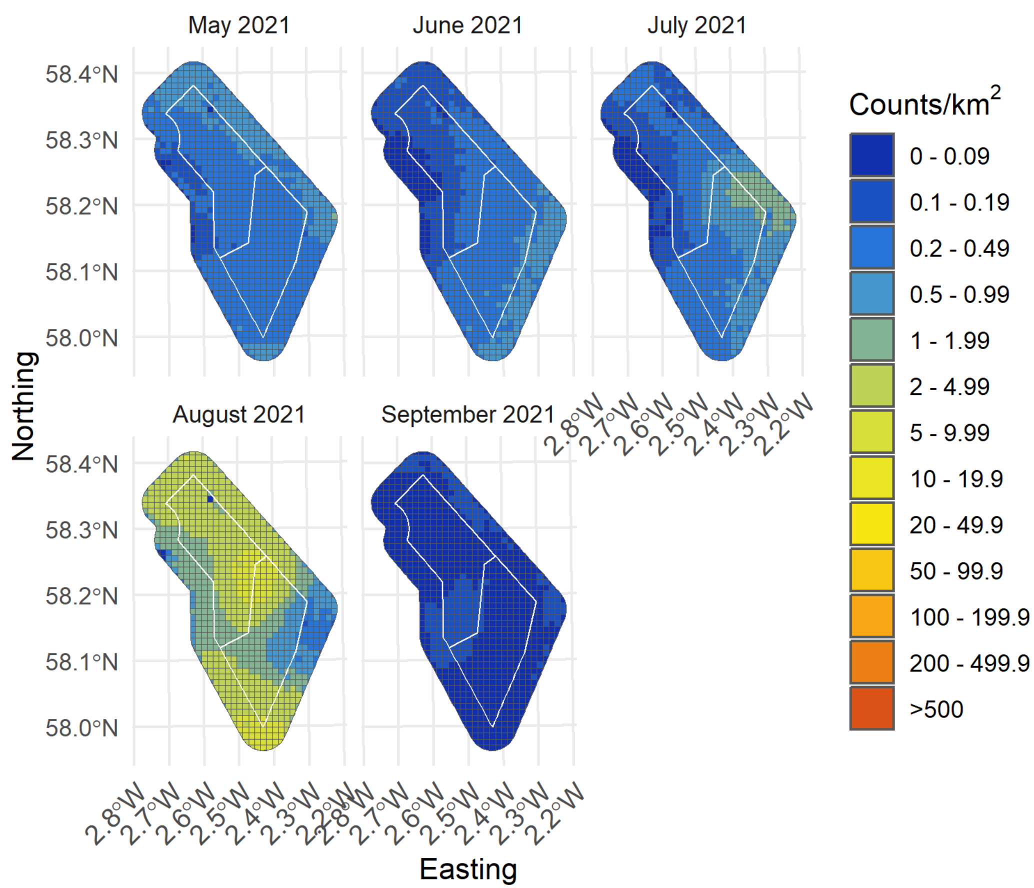


Figure 153. Lower confidence limit of density of all puffin in the survey area for months with sufficient observations between May and September 2021

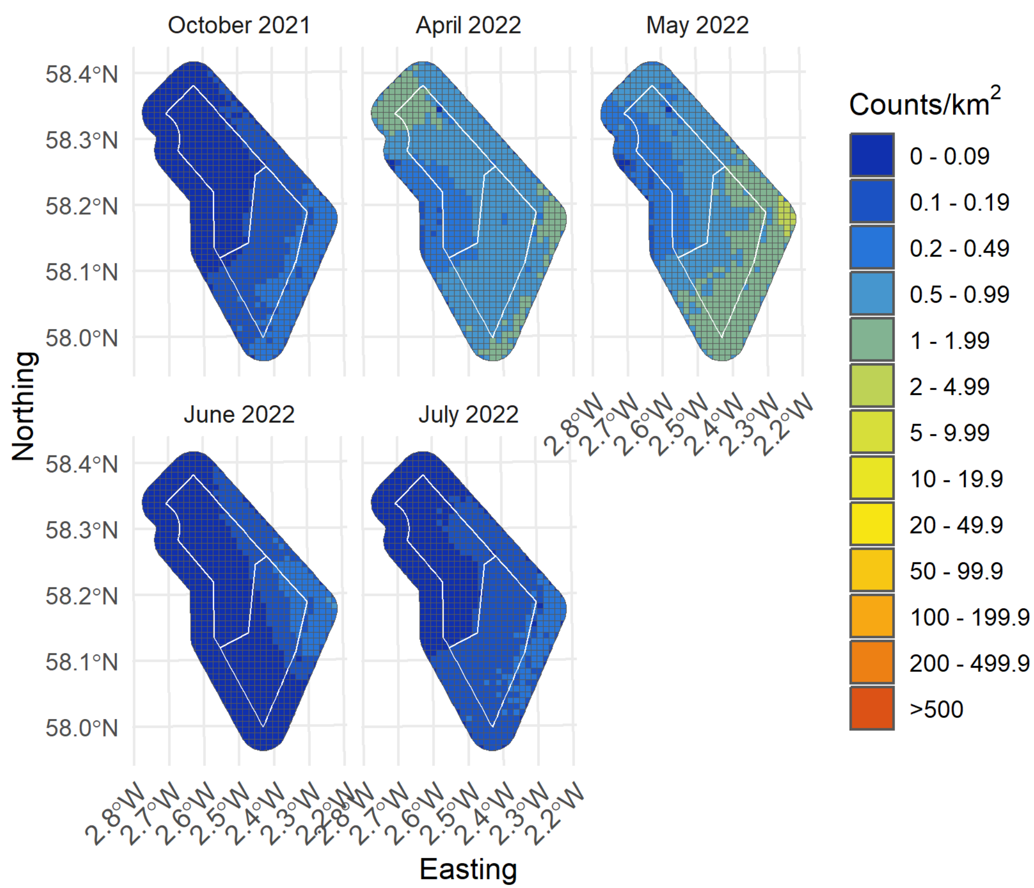


Figure 154. Lower confidence limit of density of all puffin in the survey area for months with sufficient observations between October 2021 and July 2022

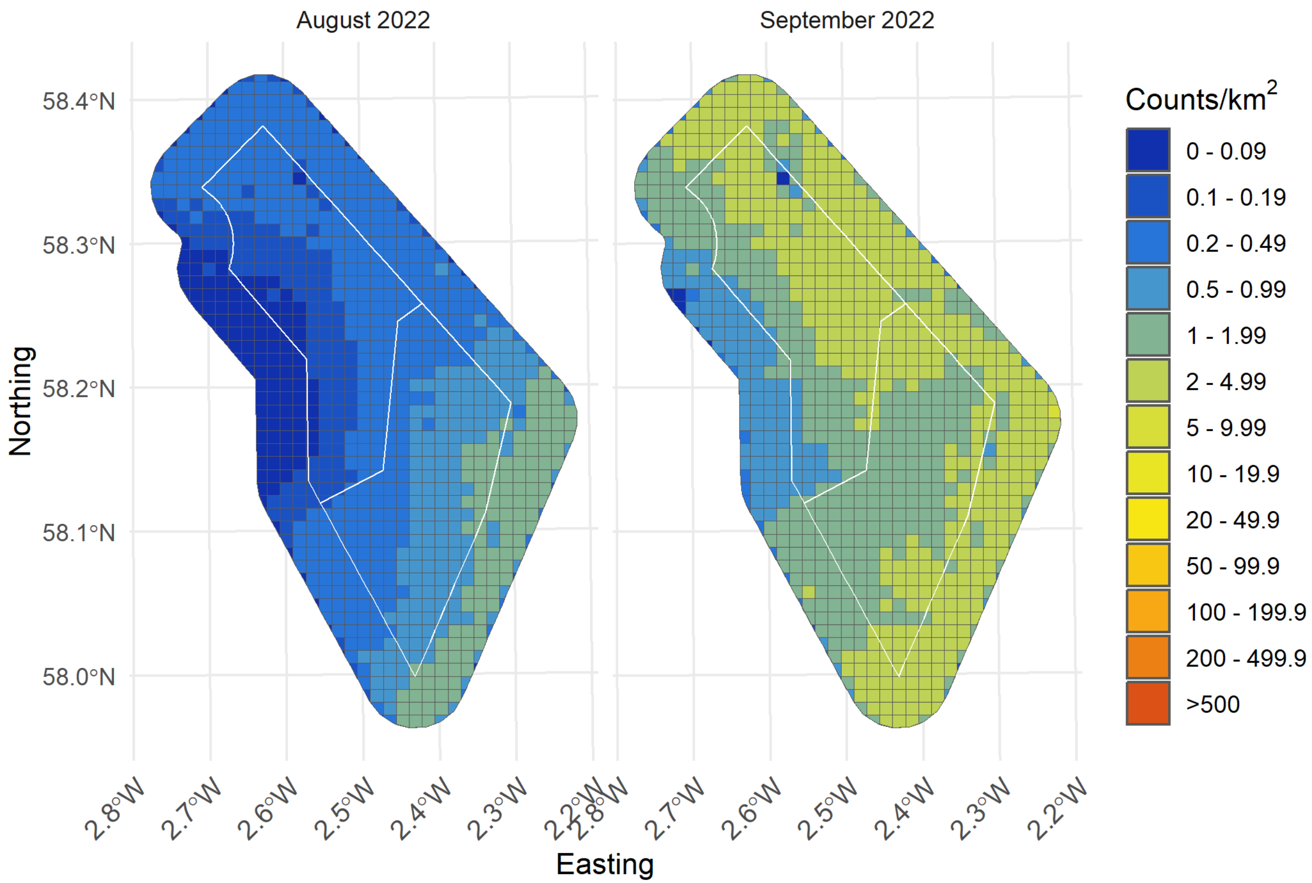


Figure 155. Lower confidence limit of density of all puffin in the survey area for months with sufficient observations between August and September 2022

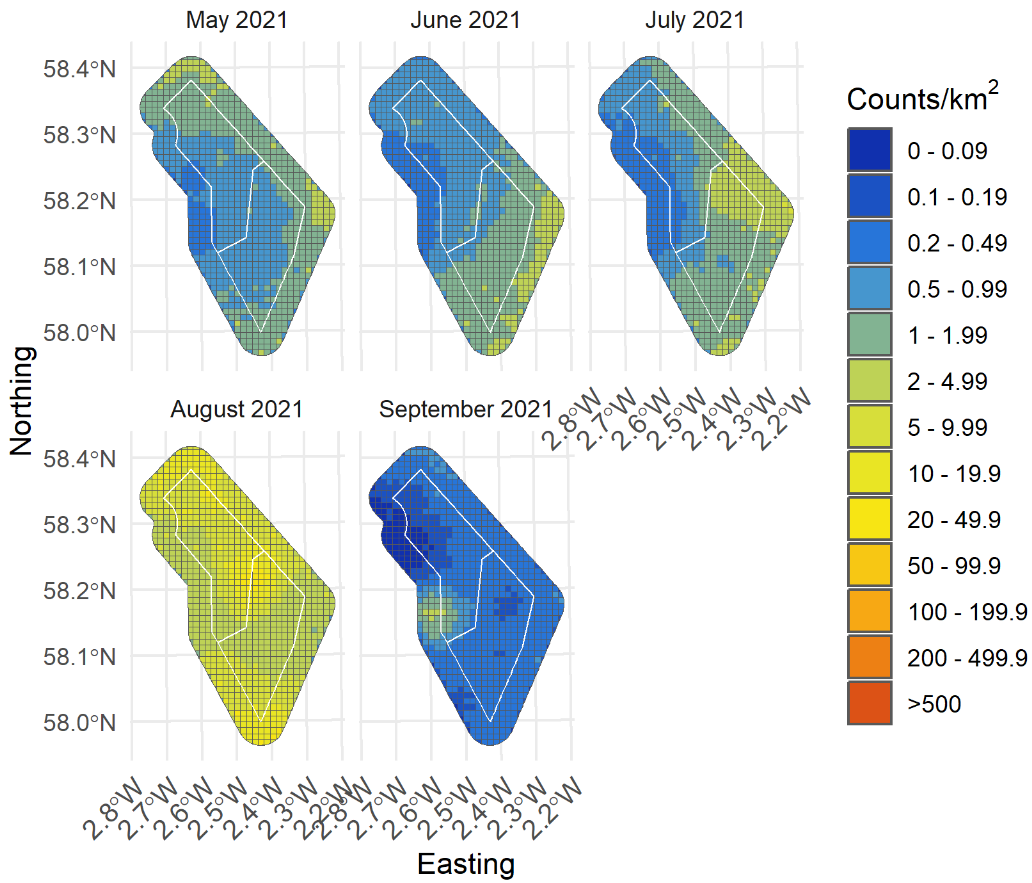


Figure 156. Upper confidence limit of density of all puffin in the survey area for months with sufficient observations between May and September 2021

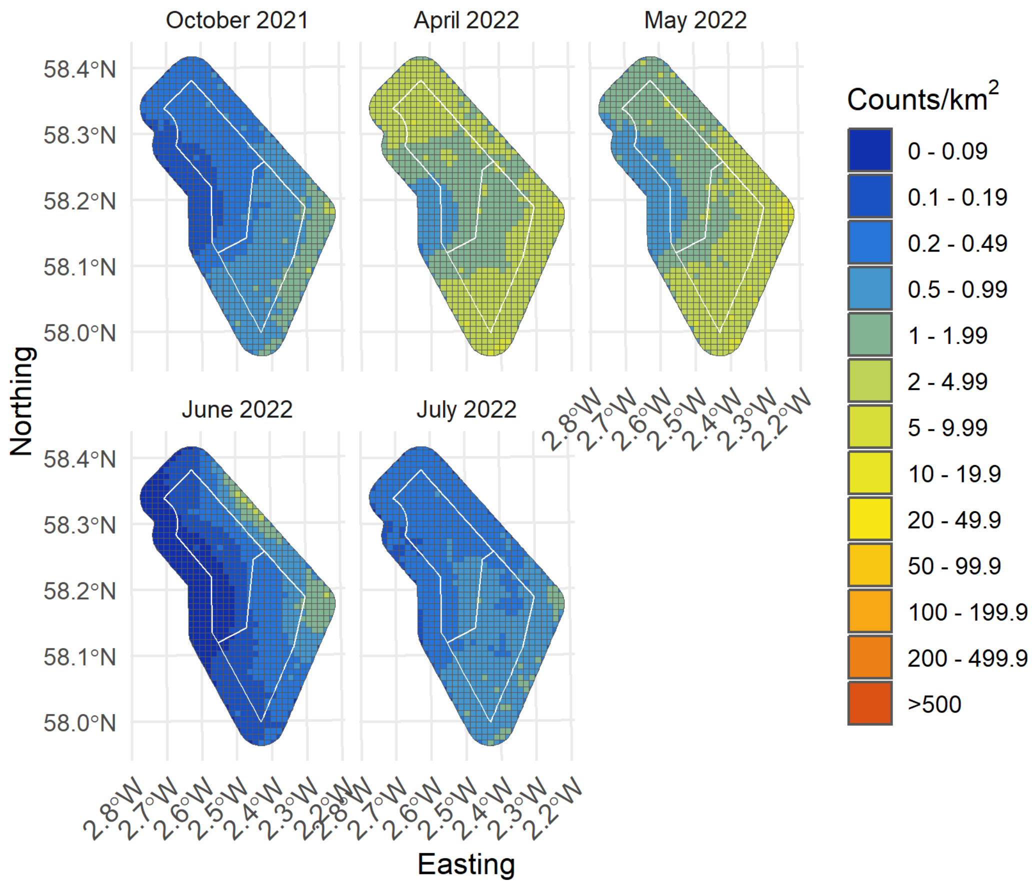


Figure 157. Upper confidence limit of density of all puffin in the survey area for months with sufficient observations between October 2021 and July 2022

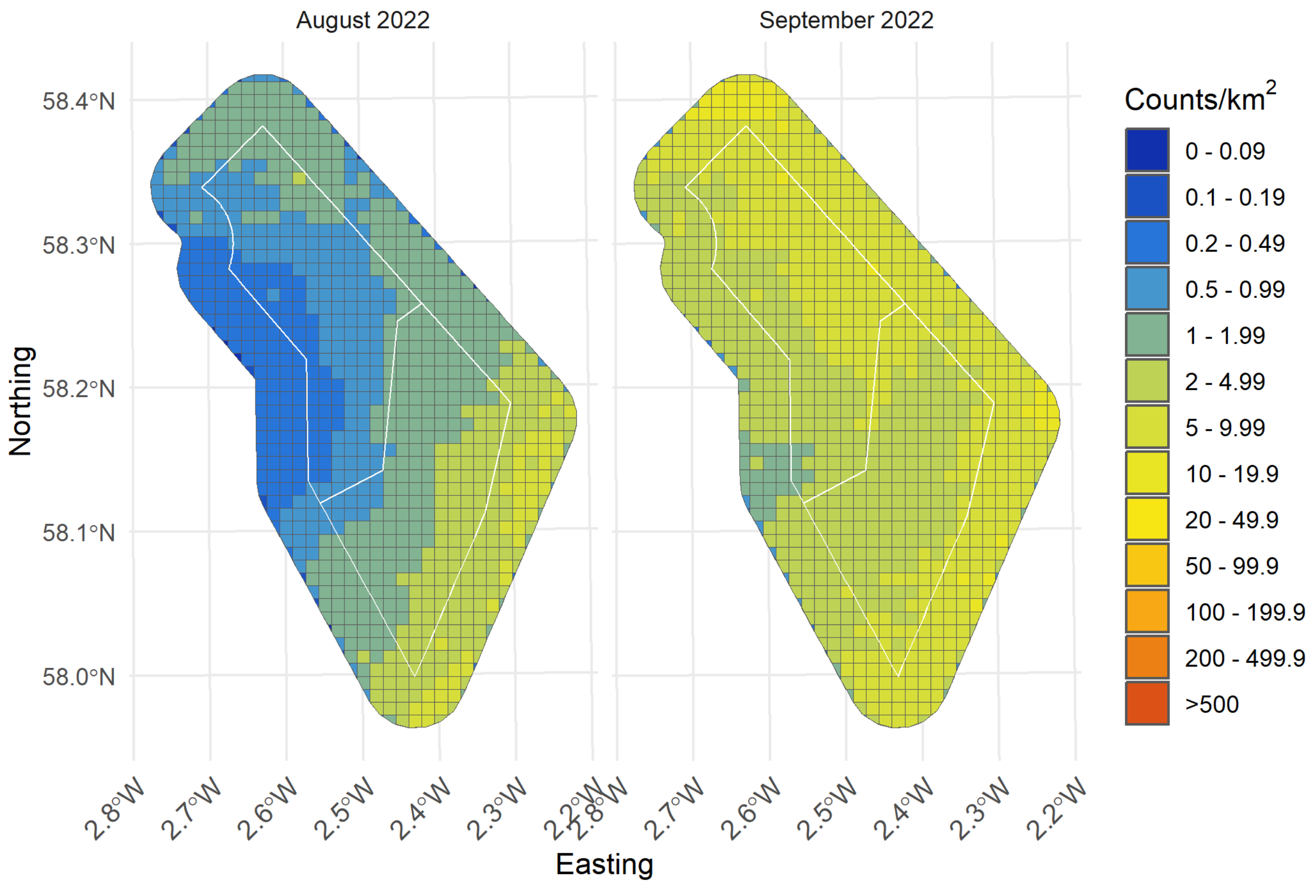


Figure 158. Upper confidence limit of density of all puffin in the survey area for months with sufficient observations between August and September 2022

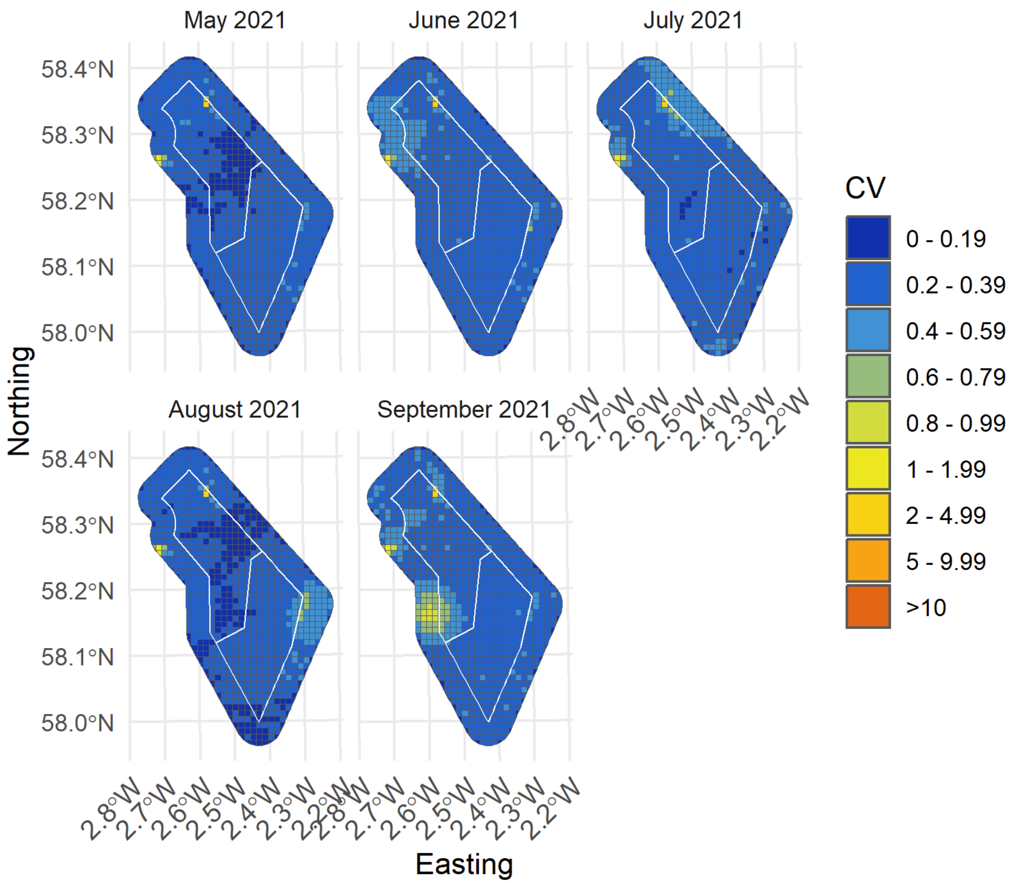


Figure 159. Spatial coefficient of variation of predicted densities of all puffin from MRSea across the survey area for months with sufficient observations between May and September 2021

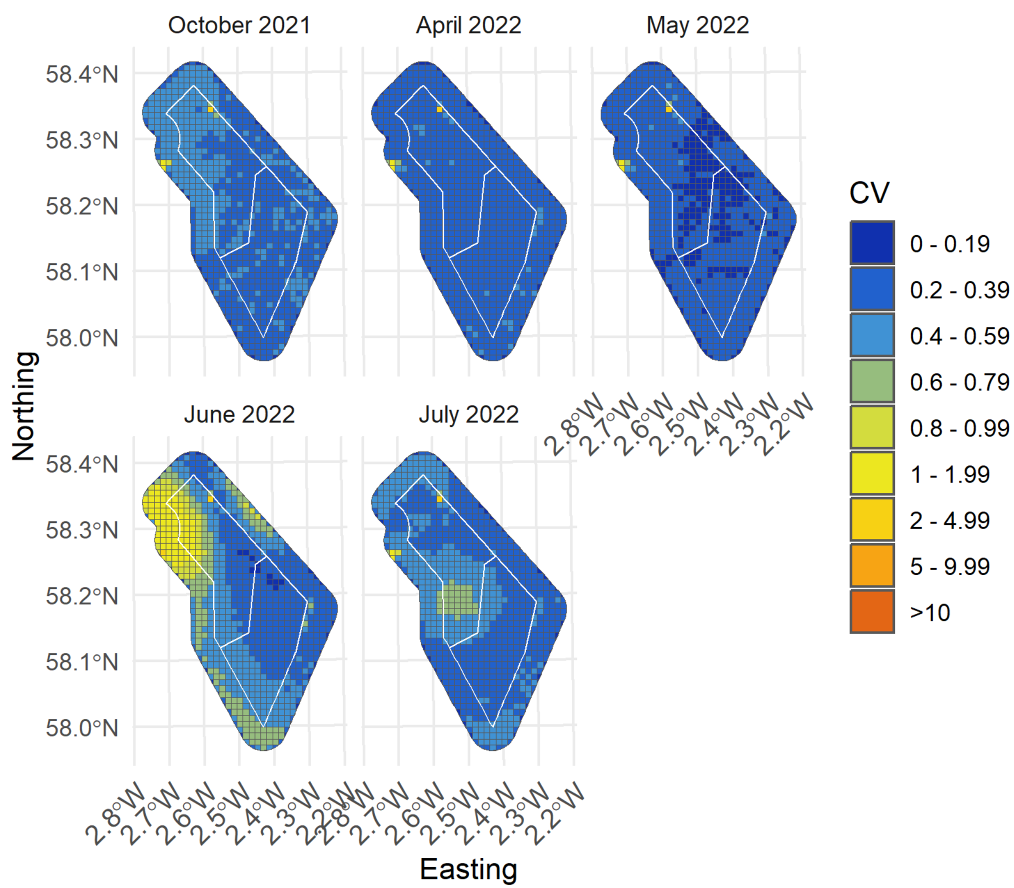


Figure 160. Spatial coefficient of variation of predicted densities of all puffin from MRSea across the survey area for months with sufficient observations between October 2021 and July 2022

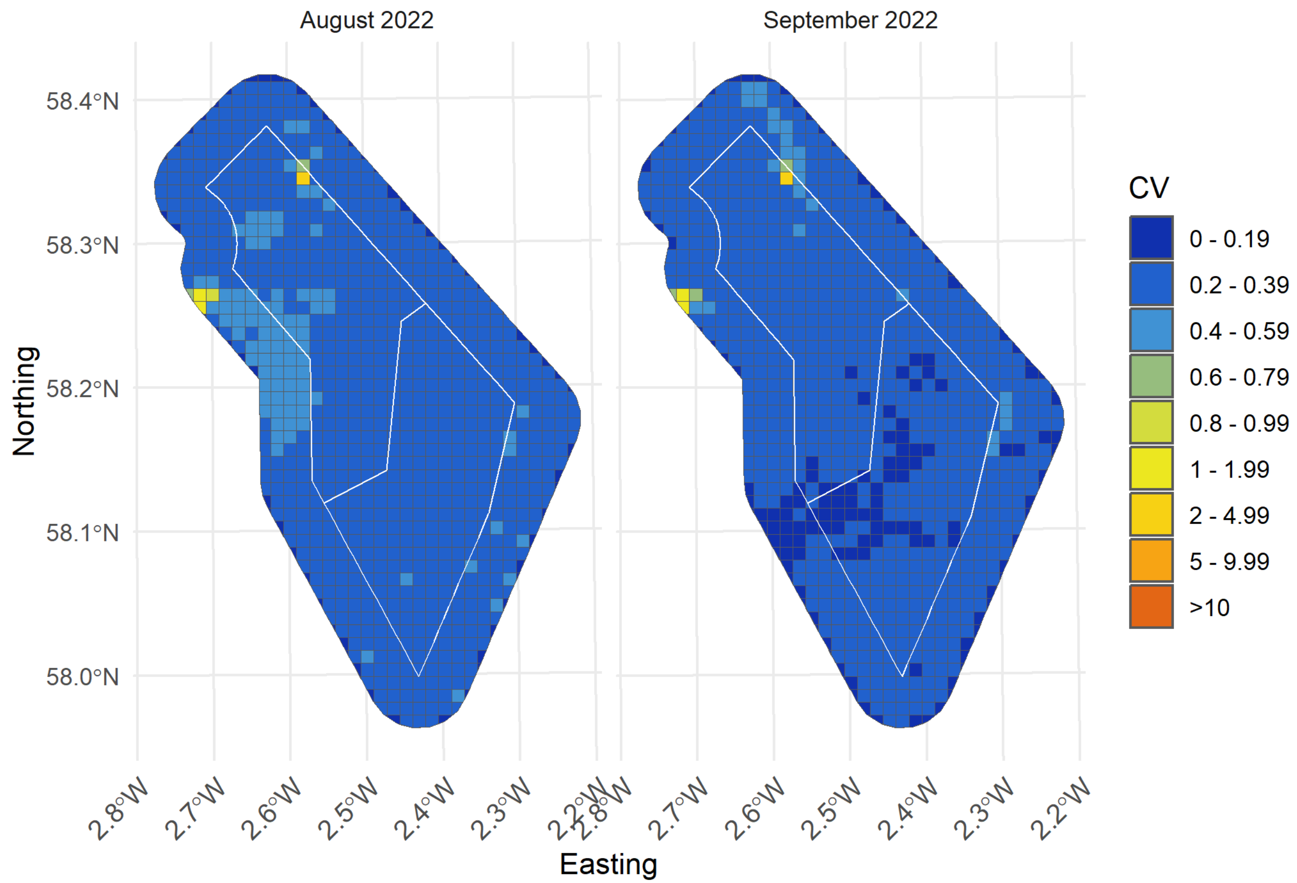


Figure 161. Spatial coefficient of variation of predicted densities of all puffin from MRSea across the survey area for months with sufficient observations between August and September 2022

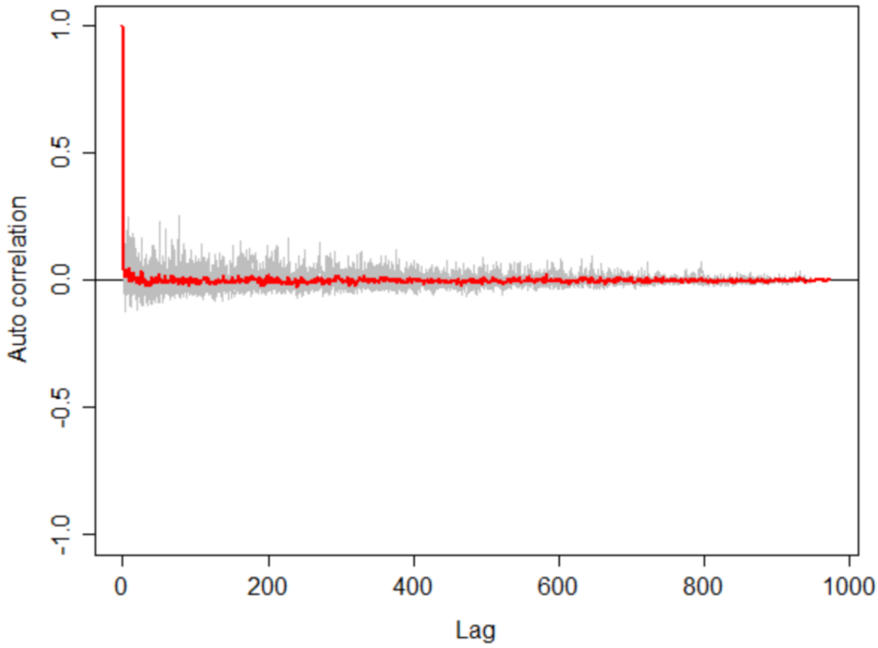


Figure 162. Autocorrelation test for puffin density surface models when using transect as a blocking feature in MRSea showing no significant correlation. A Runs test on the data prior to using transect as a blocking feature gave a p-value of << 0.0001 (i.e., that the data were significantly autocorrelated when not using a blocking feature)

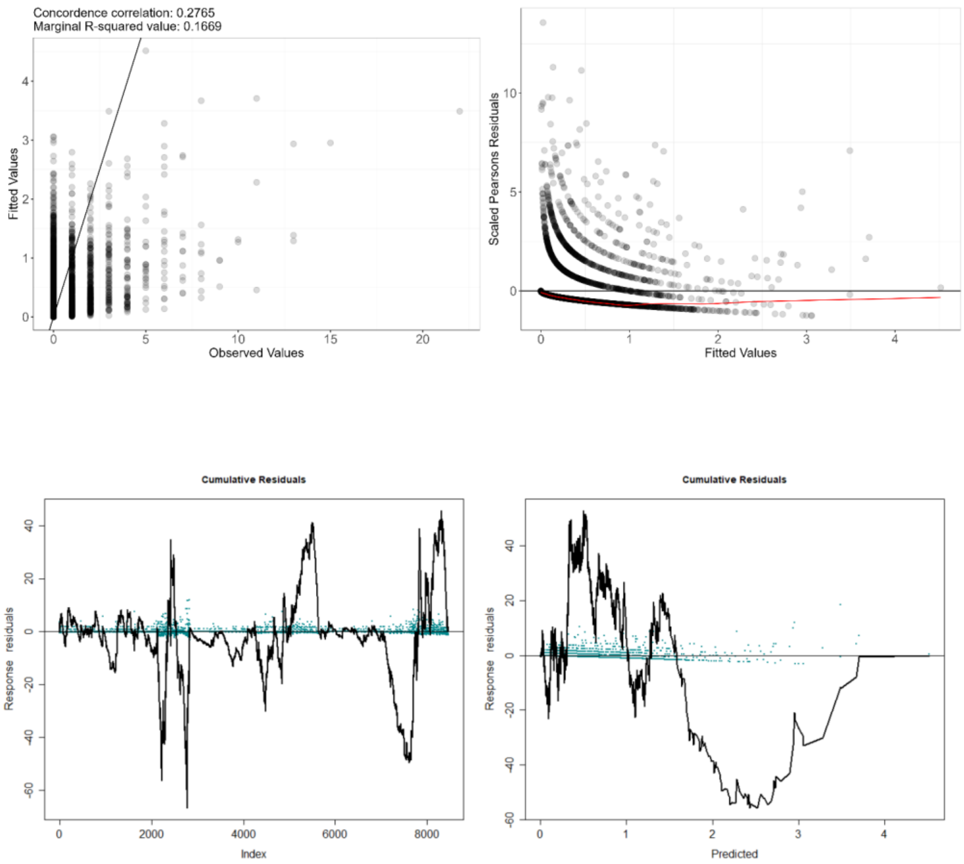


Figure 163. Fitted (MRSea predictions) versus observed counts of all puffin

Table 20. ANOVA results from the best MRSea model for all puffin as selected by cross-validation

Variable	Degrees of Freedom	Chi-square	P-value
Aspect	3	4.35	0.226
Slope	3	0.97	0.808
Bathymetry	3	1.64	0.651
Sandeel Density	3	2.03	0.566
Distance to Colony	4	6.98	0.137
Distance to Turbine	3	21.63	<<0.001
SST (SD)	3	21.70	<<0.001
X/Y (location)	10	85.76	<<0.001

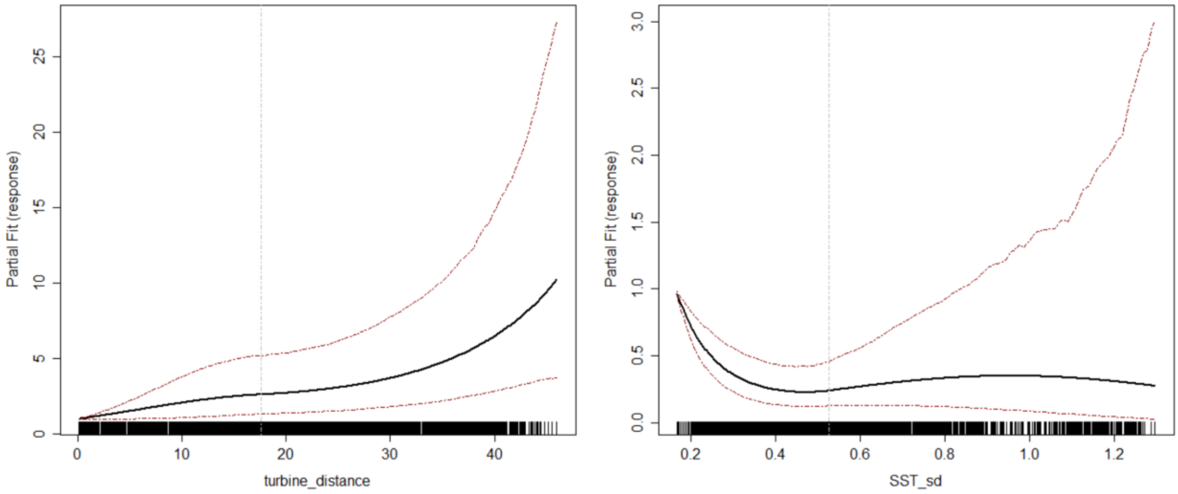


Figure 164. Partial dependence plots for significant variables for all puffin from MRSea. (Left: Distance to turbine; right: daily standard deviation of sea surface temperature.)

3.7 Great Black-backed Gull

3.7.1 All birds model

Table 21. Candidate and final covariates for all great black-backed gull MRSea model

Starting model covariates after VIF-based collinearity removal	Final model covariates after removal by SALSA
Aspect	Distance to colony
Distance to colony	Distance to turbine
Distance to turbine	Daily mean of Sea Surface Temperature
Daily mean of Sea Surface Temperature	

Due to a limited number of observations, it was more appropriate to model great black-backed gulls by survey year rather than month. Distribution maps generated using MRSea (Figure 164) suggest that all great black-backed gulls are minimally distributed throughout the study area for both years of surveys. Higher densities are estimated along the central eastern boundary of the study area.

Model fit was moderate with a marginal R squared value of 0.10 and root mean squared error of 0.01. Cumulative residuals in the model showed that there was poor relationship between predicted and observed values where the model tended to under predict above values of ~0.25 birds/km^2 (Figure 169).

Figure 165. Median density of all great black-backed gull in the survey area for each year of surveys.

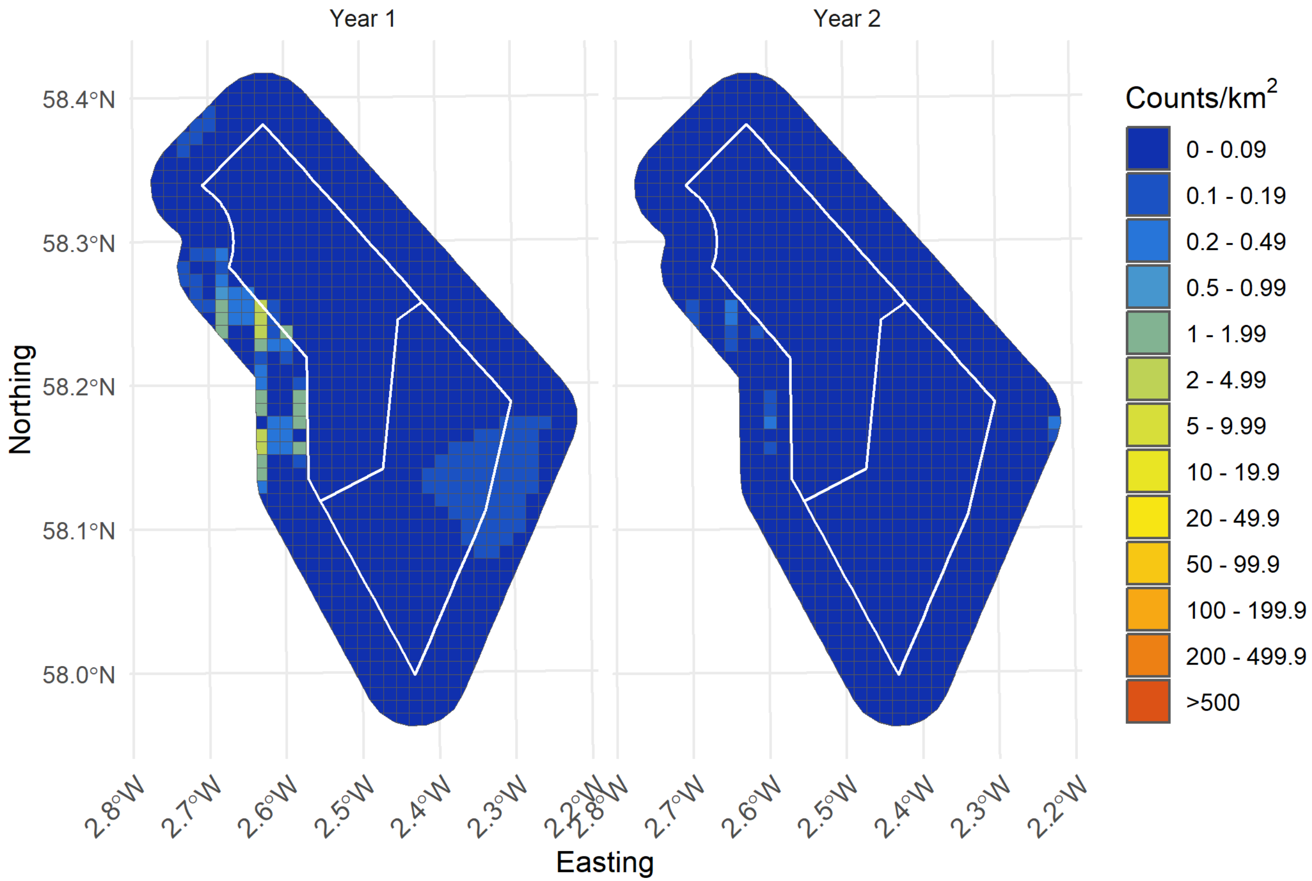


Figure 166. Lower confidence limit of density of all great black-backed gull in the survey area for each year of surveys

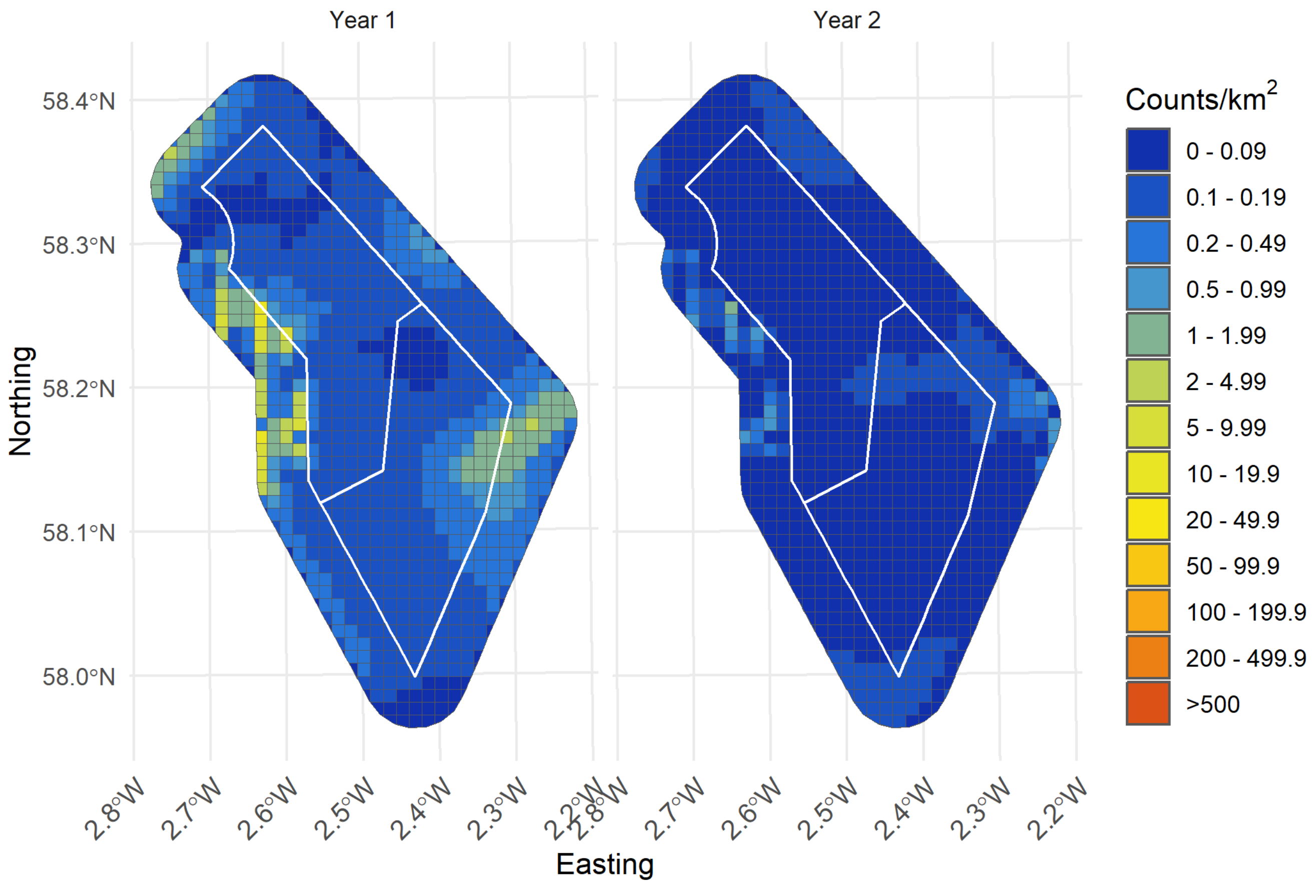


Figure 167. Upper confidence limit of density of all great black-backed gulls in the survey area for each year of surveys

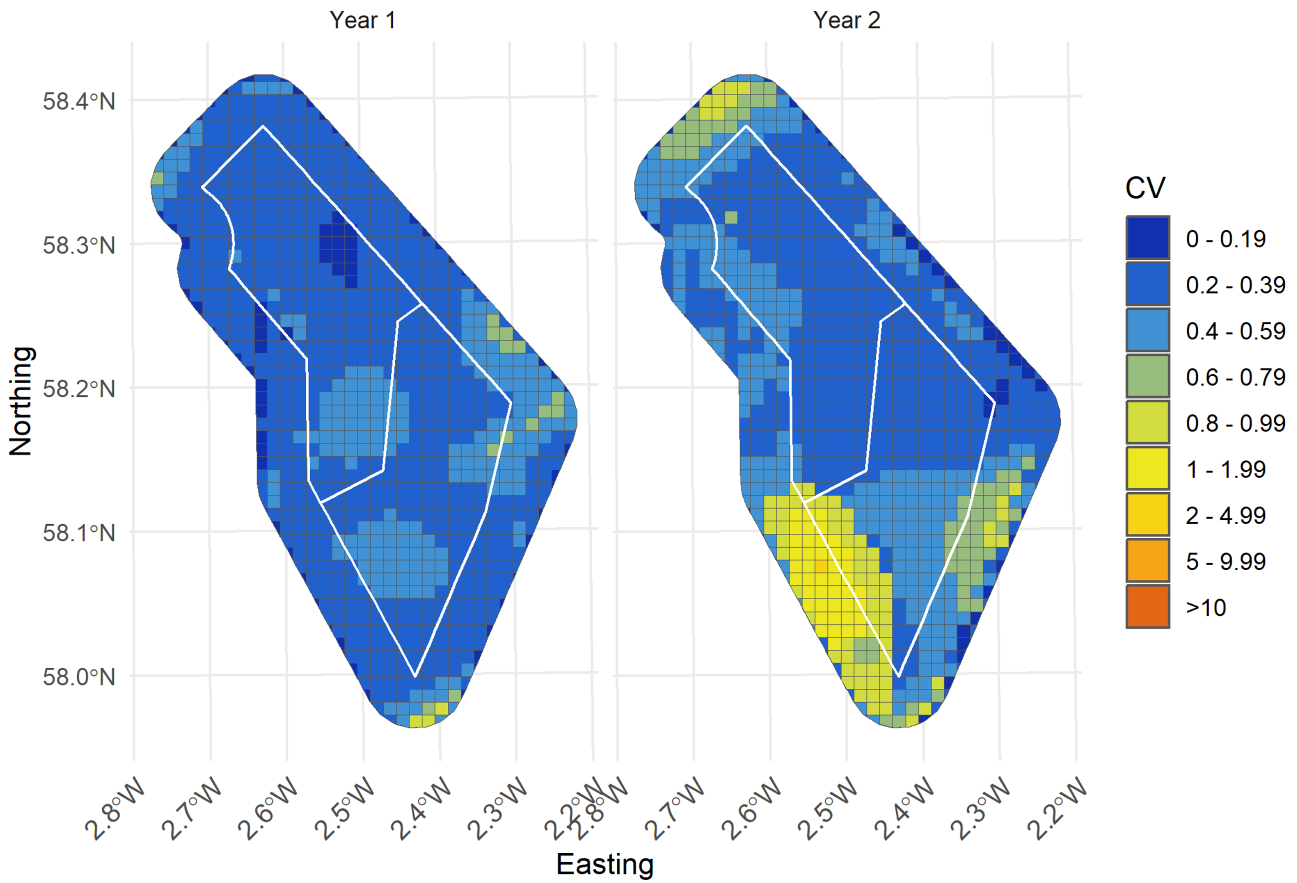


Figure 168. Spatial coefficient of variation of predicted densities of all great black-backed gull from MRSea across the survey area for each year of surveys

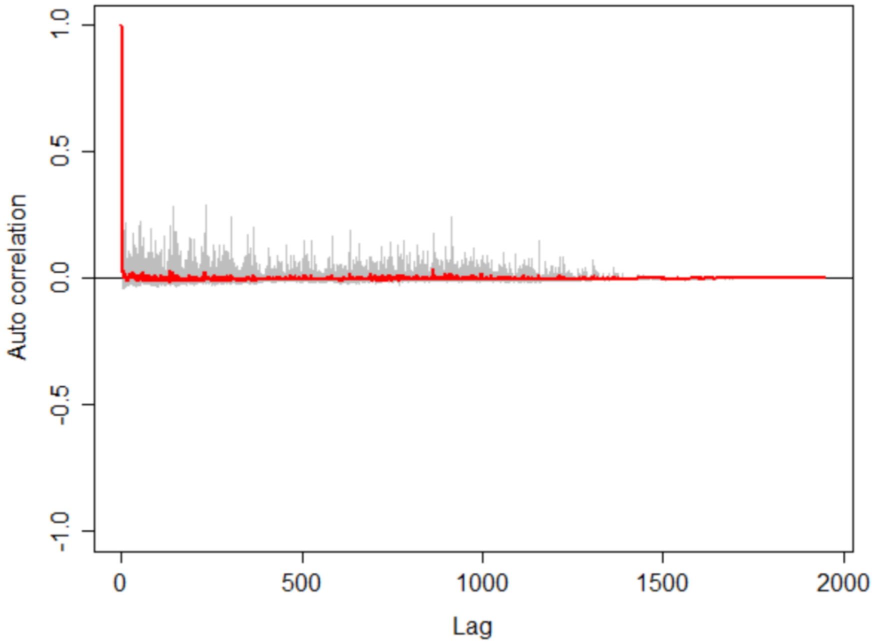


Figure 169. Autocorrelation test for great black-backed gull density surface models when using transect as a blocking feature in MRSea showing no significant correlation. A Runs test on the data prior to using transect as a blocking feature gave a p-value of << 0.0001 (i.e., that the data were significantly autocorrelated when not using a blocking feature)

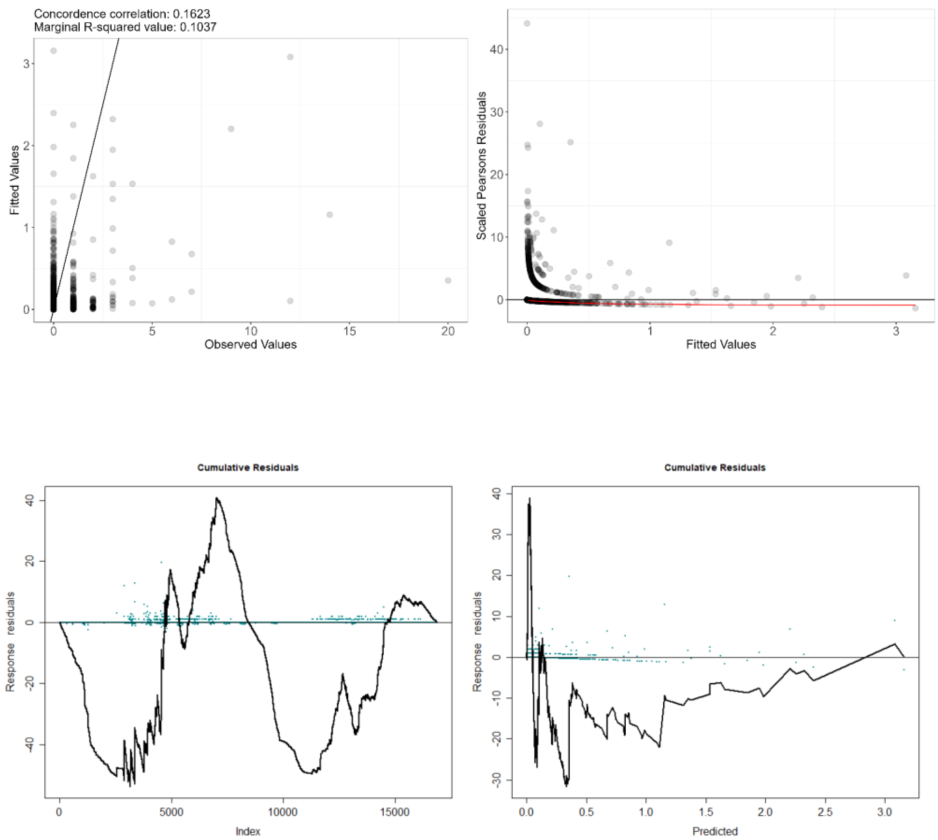


Figure 170. Fitted (MRSea predictions) versus observed counts of all great black-backed gull

Table 22. ANOVA results from the best MRSea model for all great black-backed gull as selected by cross-validation

Variable	Degrees of Freedom	Chi-square	P-value
Distance to Colony	5	18.53	0.002
Distance to Turbine	3	234.67	<<0.001
SST (daily)	5	65.72	<<0.001
X/Y (location)	15	120.48	<<0.001

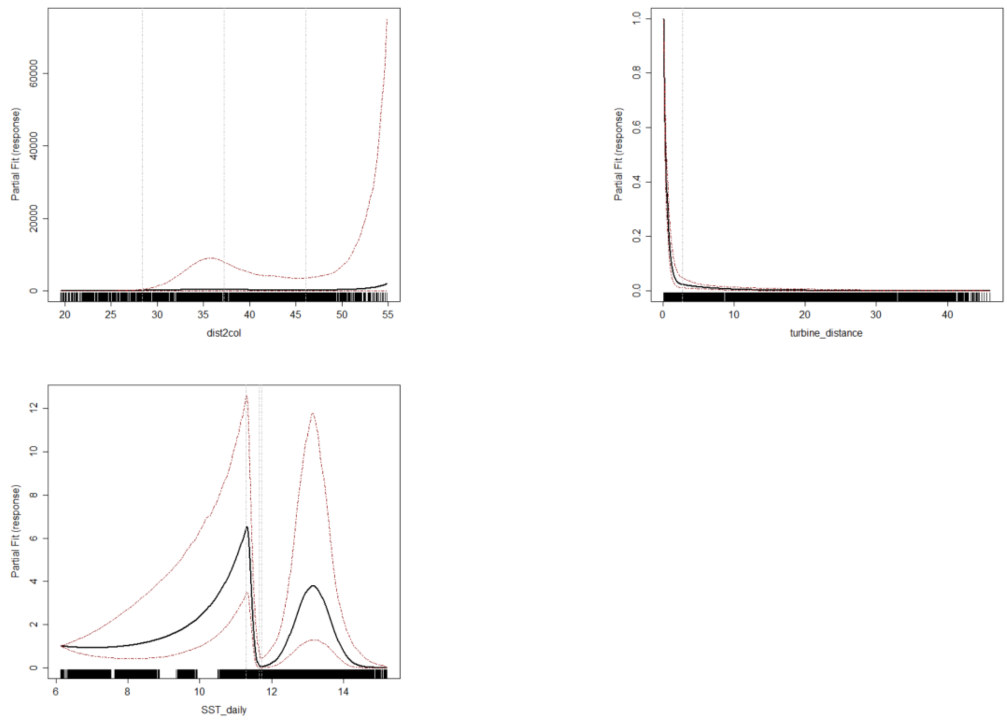


Figure 171. Partial dependence plots for significant variables for all great black-backed gull from MRSea (Clockwise from top left: distance to colony, distance to turbine, daily mean of sea surface temperature)

3.7.2 Flying birds model

Table 23. Candidate and final covariates for flying great black-backed gull MRSea model

Starting model covariates after VIF-based collinearity removal	Final model covariates after removal by SALSA
Bathymetry	Bathymetry
Distance to colony	Distance to colony
Distance to turbine	Distance to turbine

Due to a limited number of observations, it was more appropriate to model great black-backed gulls by survey year rather than month. Distribution maps generated using MRSea (Figure 168) suggest that flying great black-backed gulls are very minimally distributed throughout the study area for both years. Year one has slightly higher densities along the central western boundary of the study area. Year two has lower densities in that region, but higher densities in the northern corner of the study area.

Model fit was poor with a marginal R squared value of 0.004 and root mean squared error of 0.01. Cumulative residuals in the model showed that there was moderate relationship between predicted and observed values across most of the range of predicted values, but residuals were bounded around 0 across the whole (Figure 173).

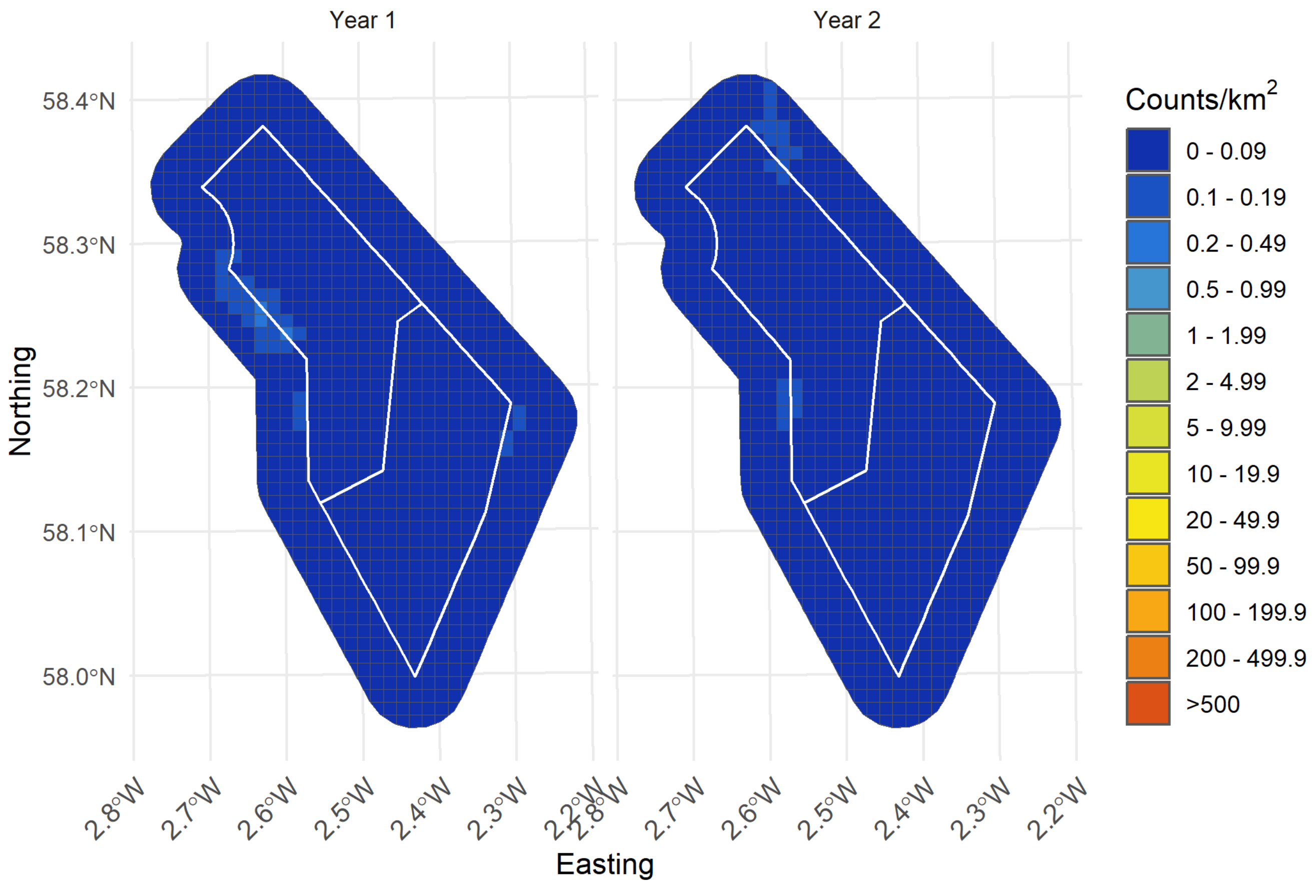


Figure 172. Median density of flying great black-backed gull in the survey area for each year of surveys.

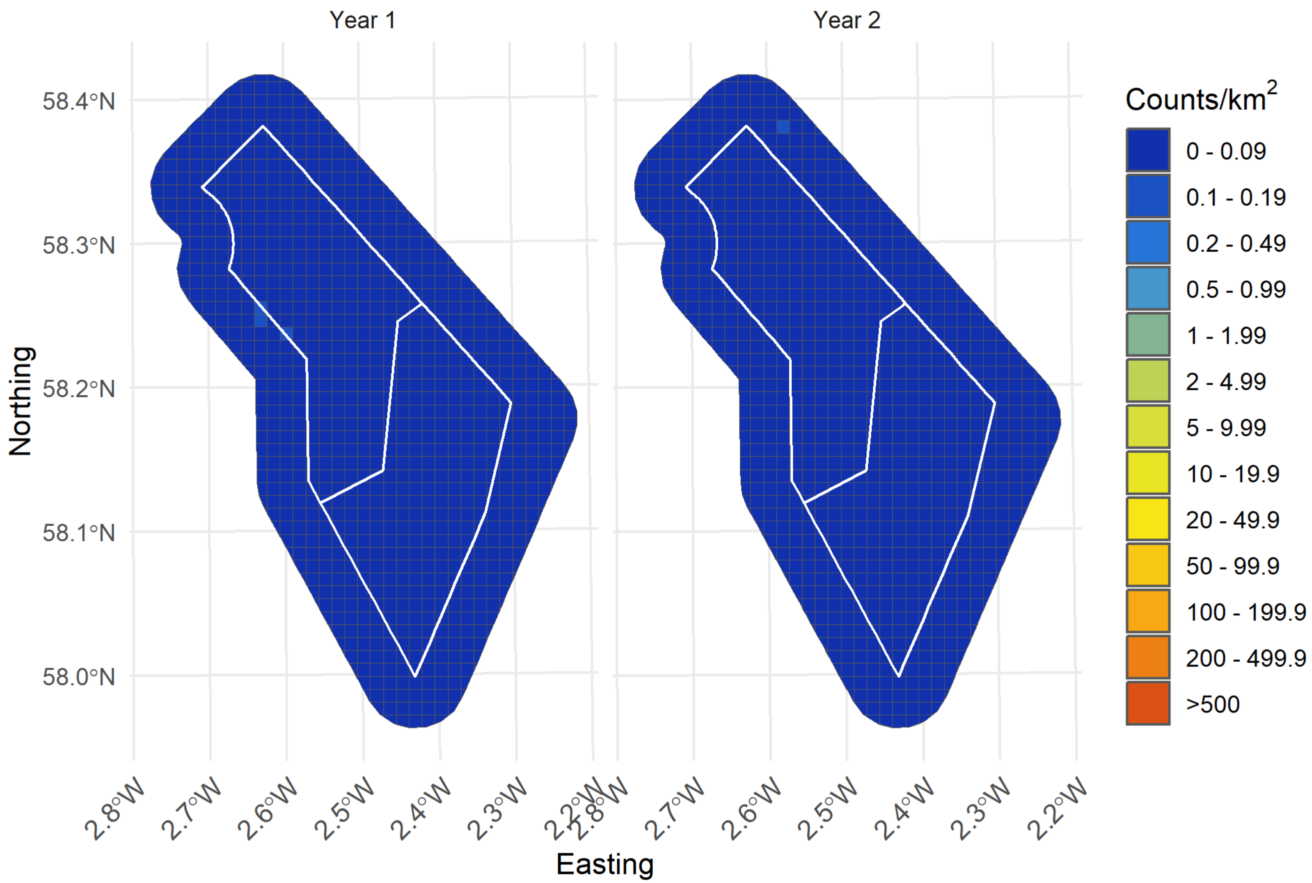


Figure 173. Lower confidence limit of density of flying great black-backed gull in the survey area for each year of surveys

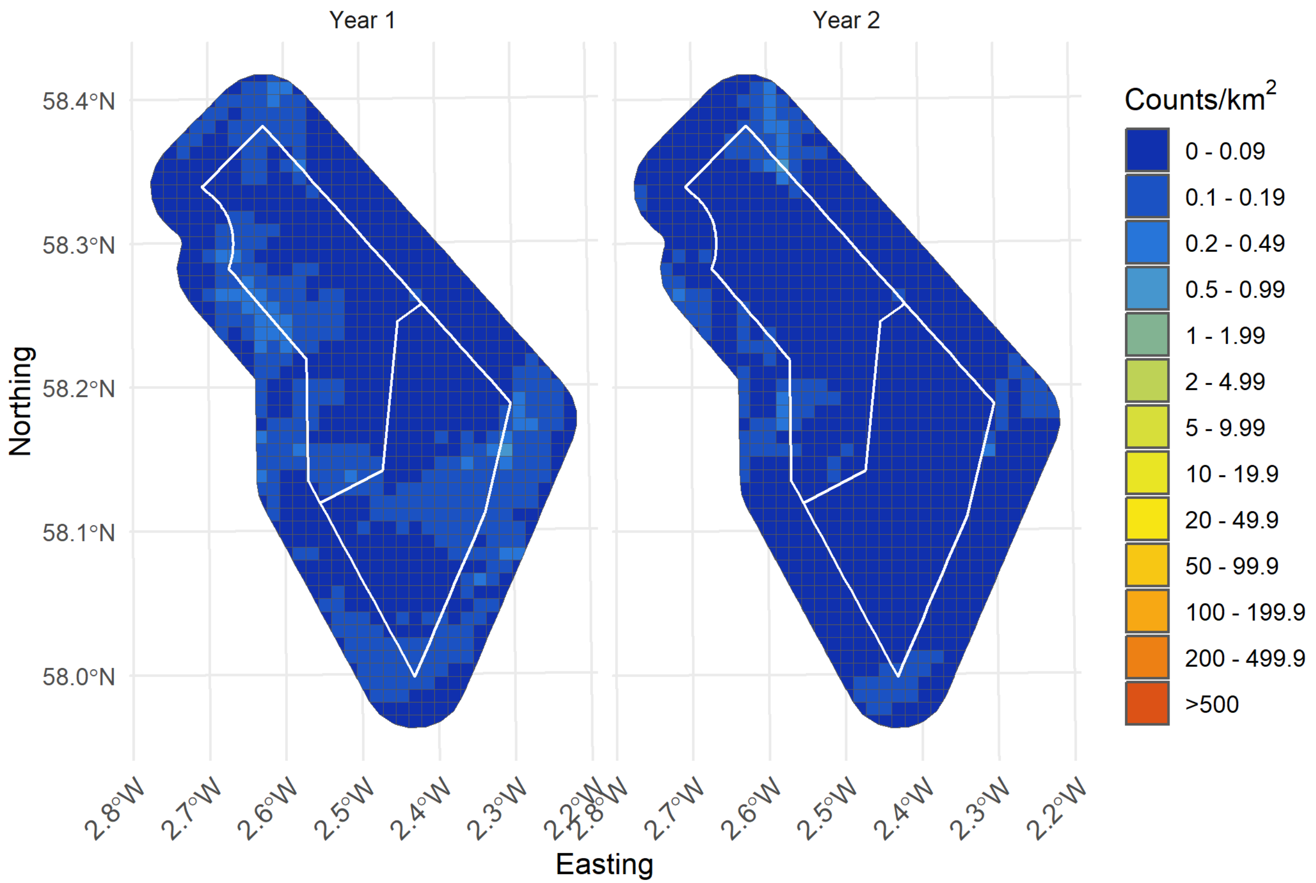


Figure 174. Upper confidence limit of density of flying great black-backed gulls in the survey area for each year of surveys

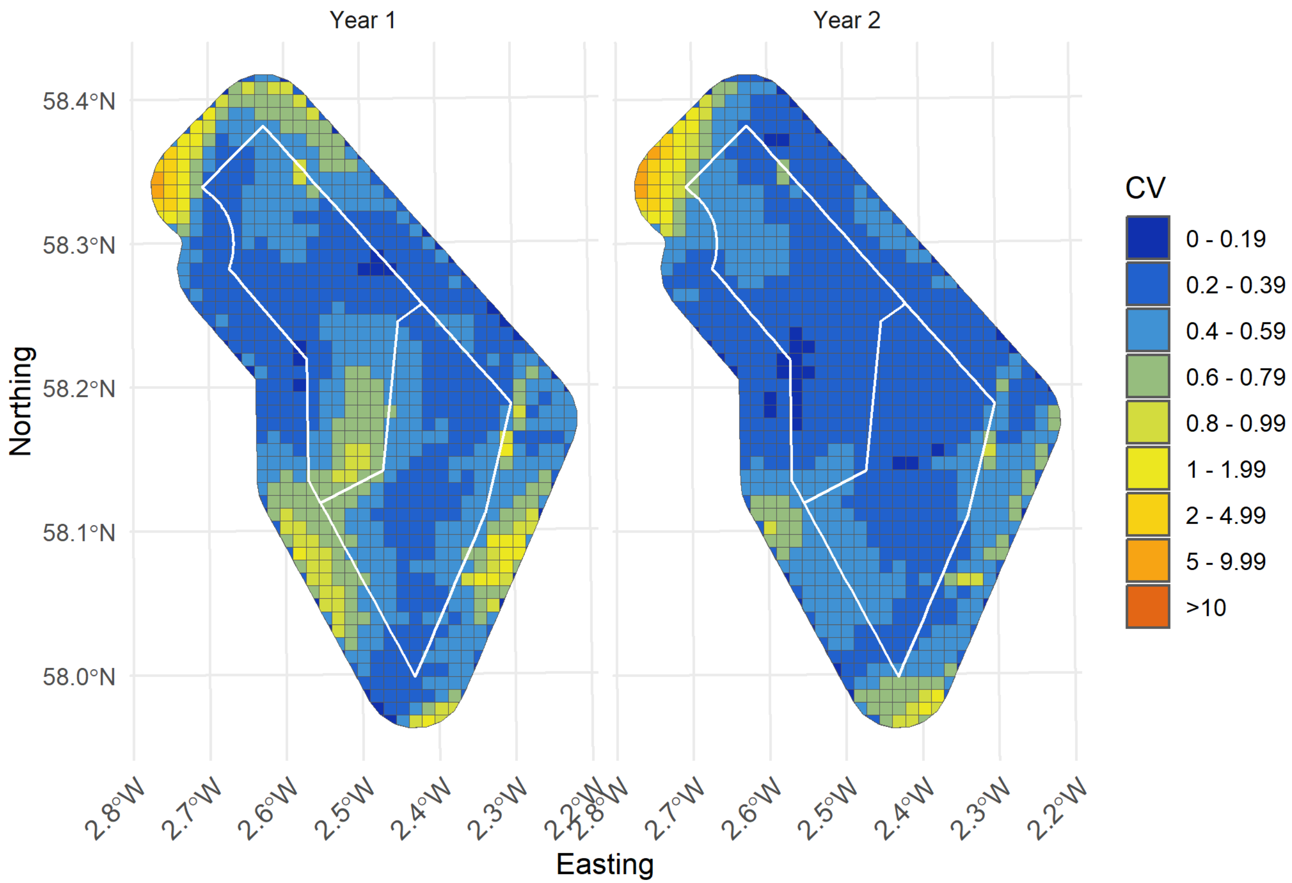


Figure 175. Spatial coefficient of variation of predicted densities of flying great black-backed gull from MRSea across the survey area for each year of surveys

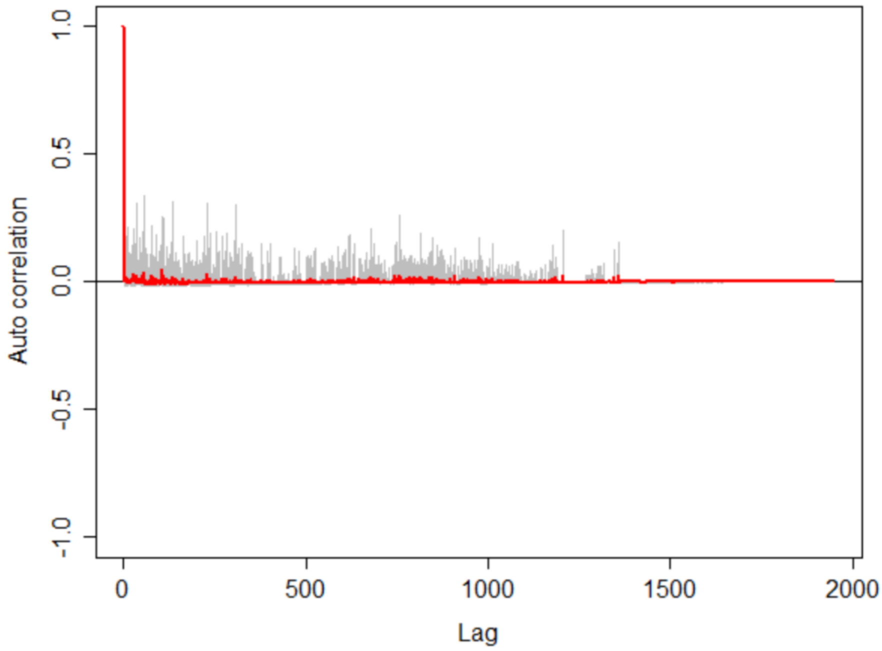


Figure 176. Autocorrelation test for great black-backed gull density surface models when using transect as a blocking feature in MRSea showing no significant correlation. A Runs test on the data prior to using transect as a blocking feature gave a p-value of << 0. 0001 (i.e., that the data were significantly autocorrelated when not using a blocking feature)

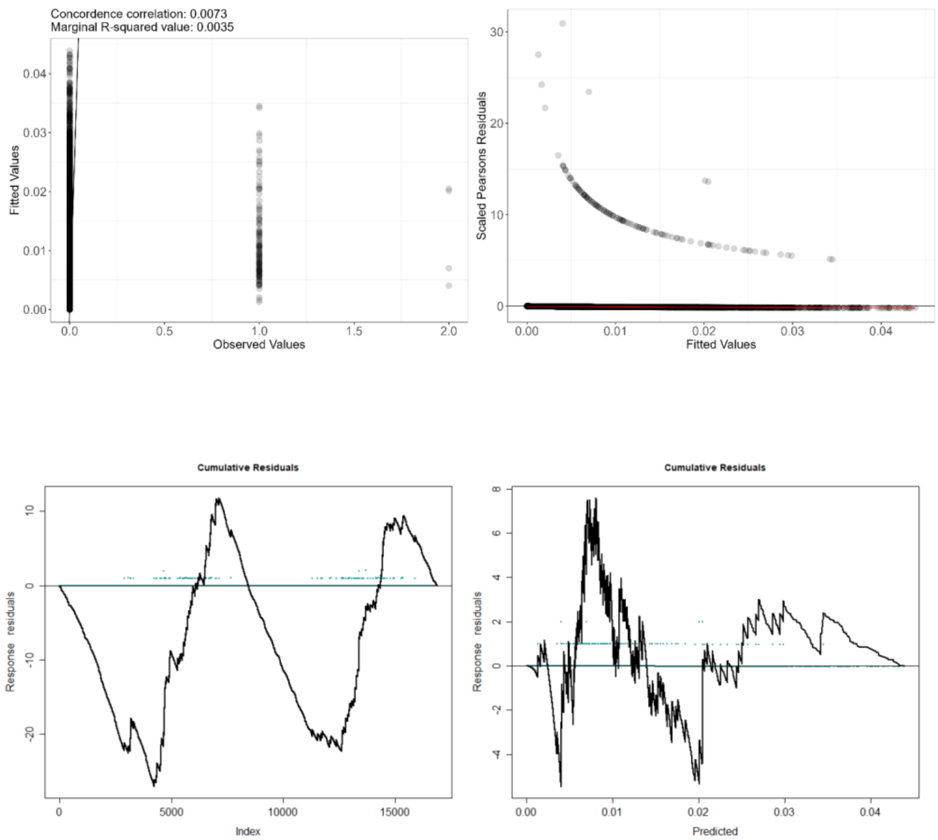


Figure 177. Fitted (MRSea predictions) versus observed counts of flying great black-backed gull

Table 24. ANOVA results from the best MRSea model for flying Great Black-backed Gull as selected by cross-validation

Variable	Degrees of Freedom	Chi-square	P-value
Bathymetry	5	4.89	0.429
Distance to Colony	5	36.23	<<0.001
Distance to Turbine	5	51.62	<<0.001
X/Y (location)	15	79.72	<<0.001

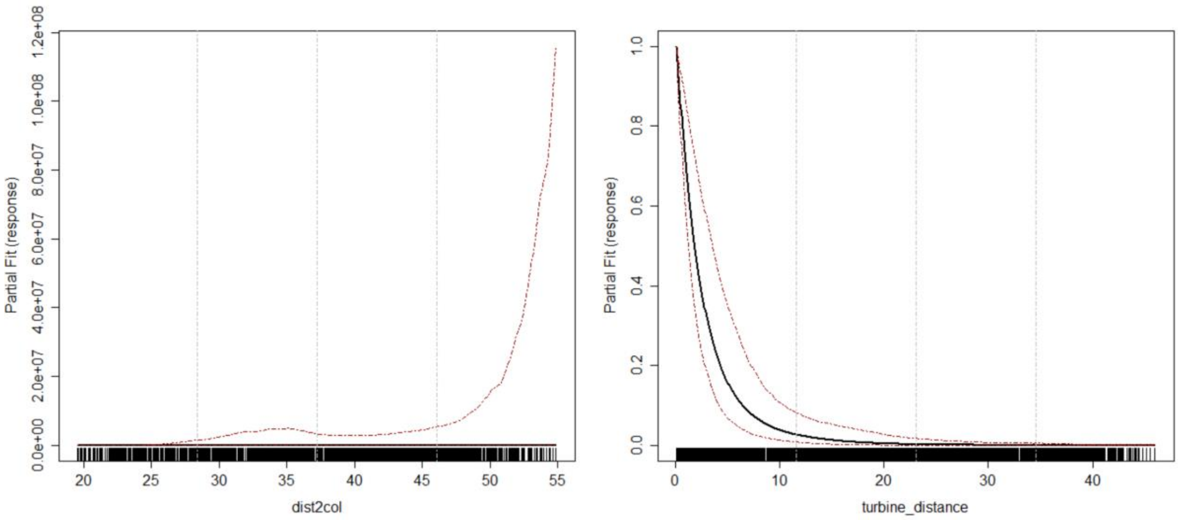


Figure 178. Partial dependence plots for significant variables for flying Great Black-backed Gull from MRSea (left: distance to colony, right: distance to tubine)

3.8 Herring Gull

Table 25. Candidate and final covariates for herring gull MRSea model

Starting model covariates after VIF-based collinearity removal	Final model covariates after removal by SALSA
Bathymetry	Bathymetry
Distance to colony	Distance to colony
Distance to turbine	Distance to turbine
Daily mean of Sea Surface Temperature	Daily mean of Sea Surface Temperature

Due to a limited number of observations, we decided to model herring gulls by survey year rather than month. Distribution maps generated using MRSea (Figure 178) suggest that herring gull occur in minimal numbers throughout the study area. The highest densities were modeled to be along the southern edge during the first year.

Model fit was relatively poor with a marginal R squared value of 0.03 and root mean squared error of 0.03. Cumulative residuals in the model showed that there was a poor relationship between predicted and observed values across most of the range of predicted values, especially above densities of 0.1 (Figure 183). Due to there being so few observations, we had to aggregate across the all the months of surveys and therefore the SD was very high, and our confidence in this model is low.

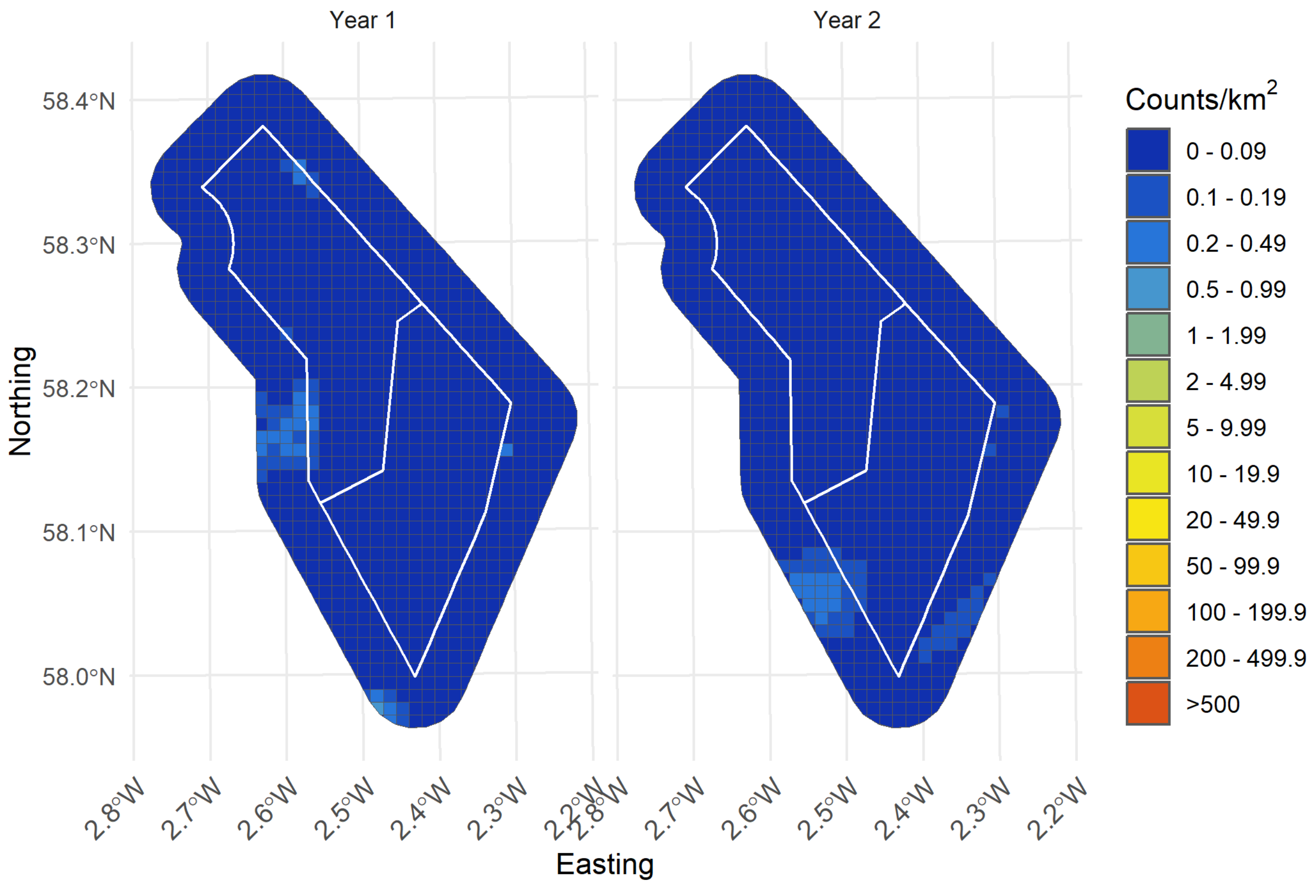


Figure 179. Median density of all herring gull in the survey area for each year of surveys

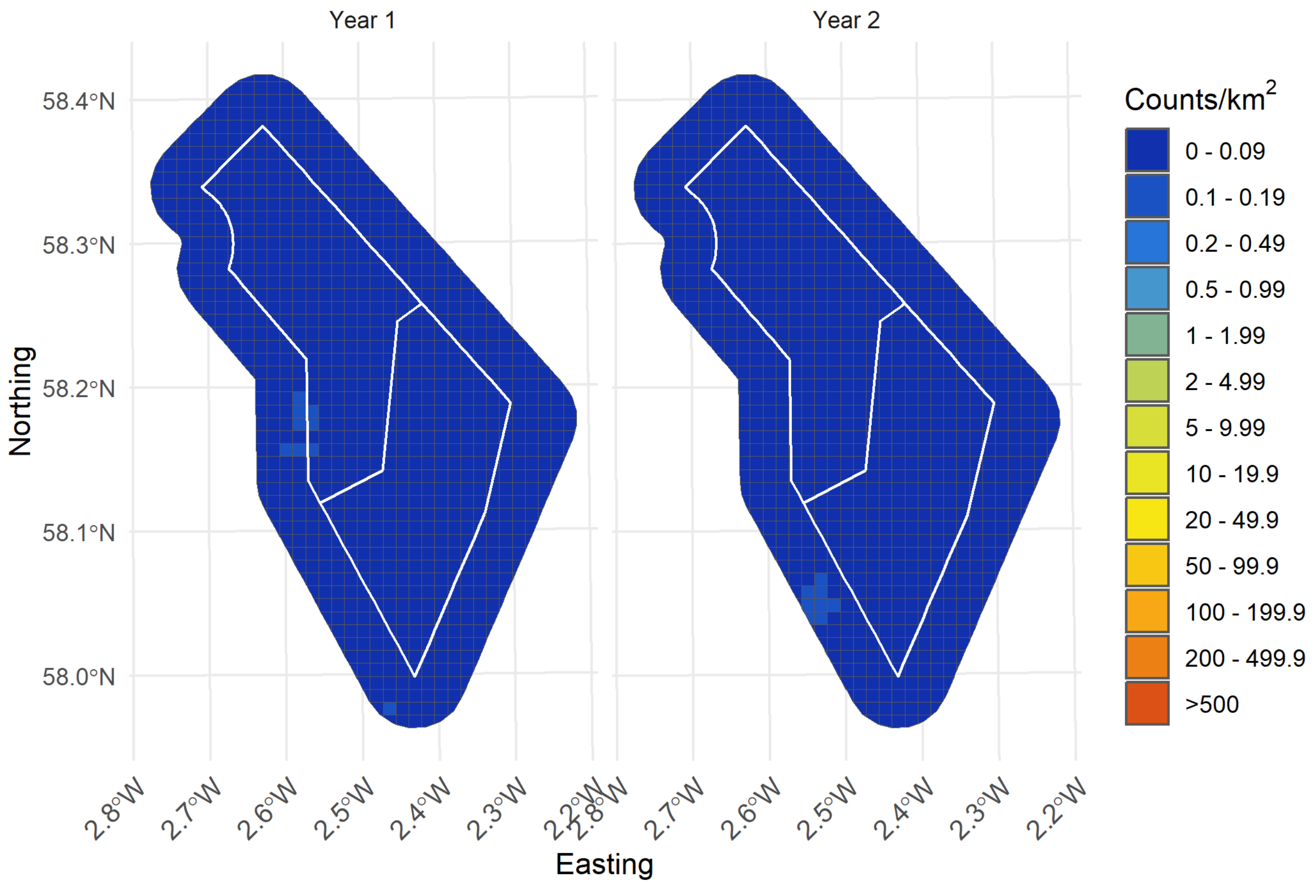


Figure 180. Lower confidence limit of density of all herring gull in the survey area for each year of surveys

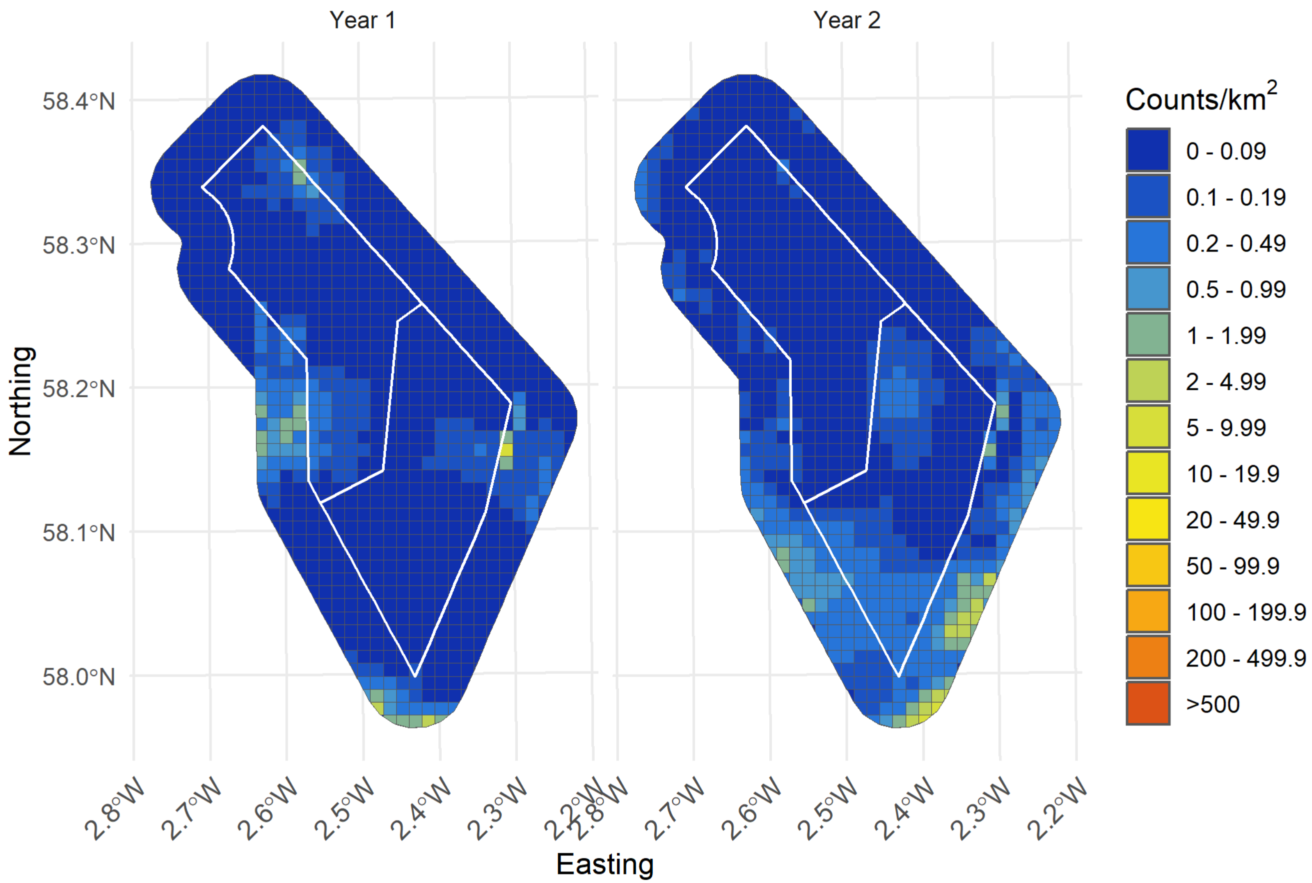


Figure 181. Upper confidence limit of density of all herring gulls in the survey area for each year of surveys

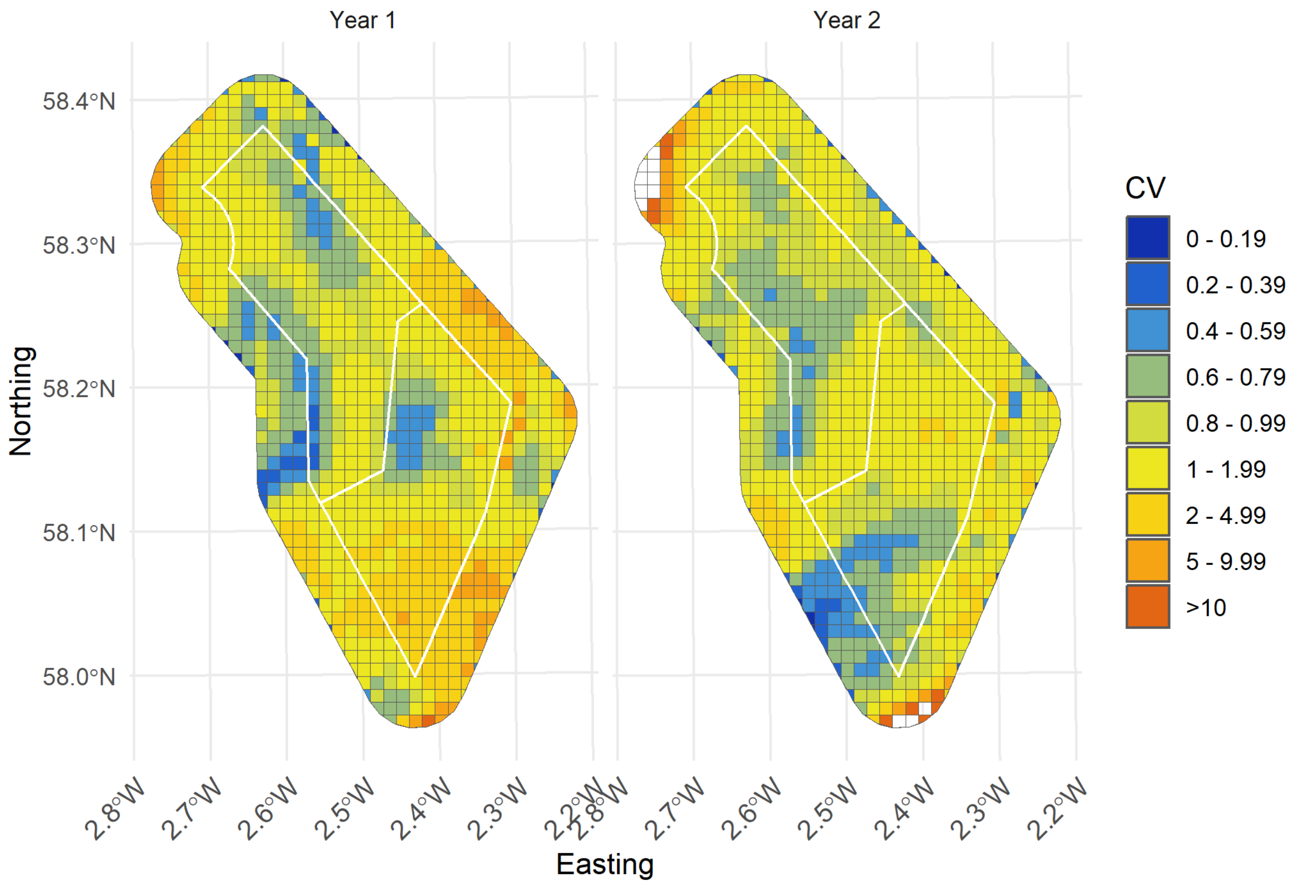


Figure 182. Spatial coefficient of variation of predicted densities of all herring gull from MRSea across the survey area for each year of surveys

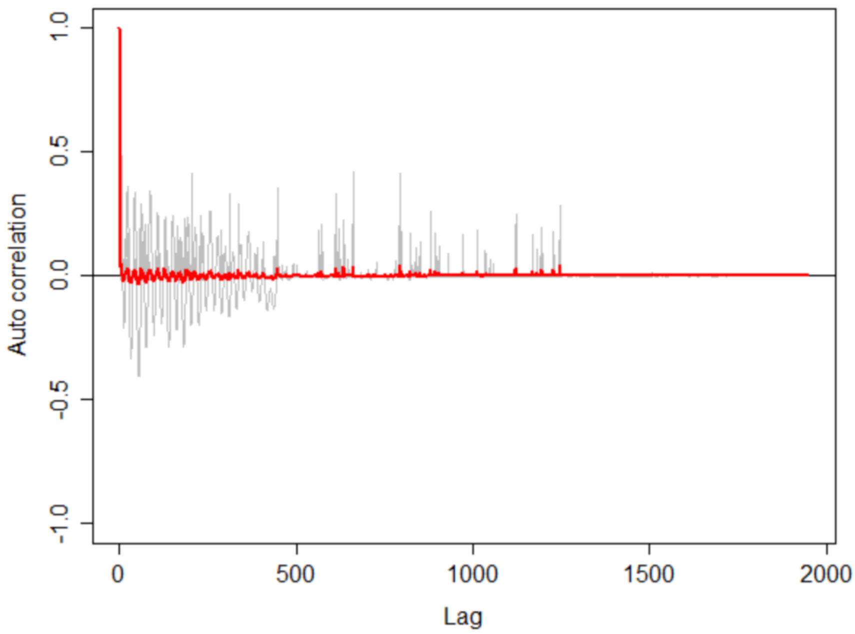


Figure 183. Autocorrelation test for herring gull density surface models when using transect as a blocking feature in MRSea showing no significant correlation. A Runs test on the data prior to using transect as a blocking feature gave a p-value of << 0.0001 (i.e., that the data were significantly autocorrelated when not using a blocking feature)

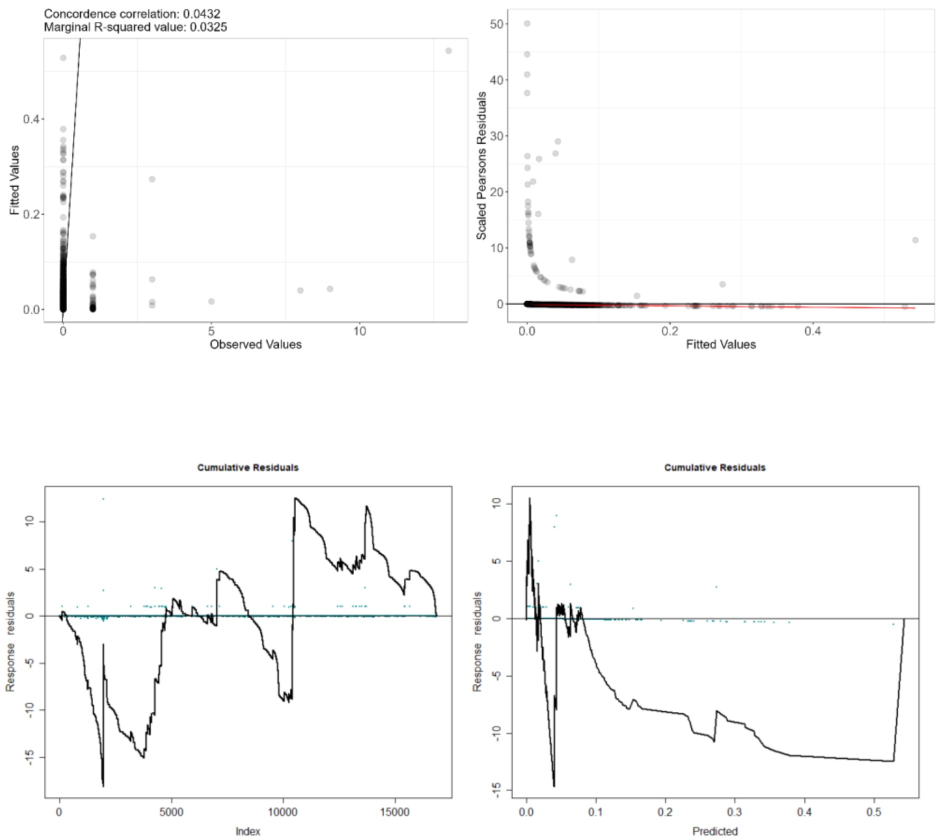


Figure 184. Fitted (MRSea predictions) versus observed counts of all herring gull

Table 26. ANOVA results from the best MRSea model for all herring gull as selected by cross-validation

Variable	Degrees of Freedom	Chi-square	P-value
Bathymetry	5	23.77	<<0.001
Distance to Colony	5	2.16	0.826
Distance to Turbine	5	29.70	<<0.001
SST (daily)	5	17.94	0.003
X/Y (location)	15	134.92	<<0.001

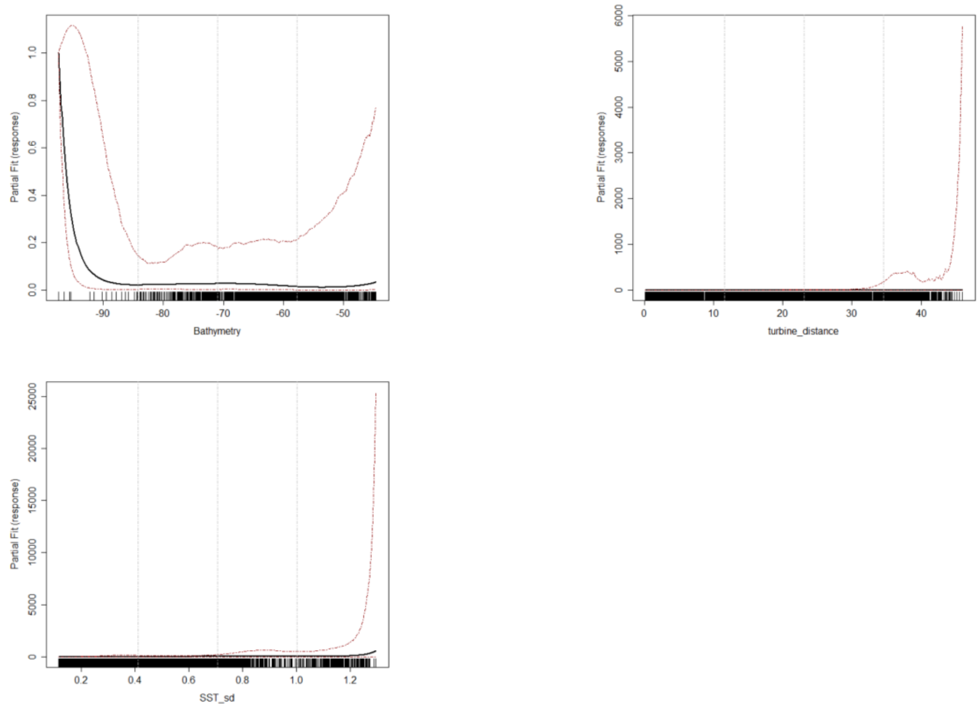


Figure 185. Partial dependence plots for significant variables for all herring gull from MRSea (Clock wise from top left: Bathymetry, distance to turbine, daily standard deviation of sea surface temperature)

4 References

Breiman, L., 2001. Random forests. *Machine learning*, *45*, pp.5-32.

Cutler, Adele, D. Richard Cutler, and John R. Stevens. 2012. "Random Forests." In, 157–75. Springer New York. https://doi.org/10.1007/978-1-4419-9326-7_5.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2008. "Random Forests." In, 587–604. Springer New York. https://doi.org/10.1007/978-0-387-84858-7_15.

John, P. W. M., M. E. Johnson, L. M. Moore, and D. Ylvisaker. 1995. "Minimax Distance Designs in Two-Level Factorial Experiments." *Journal of Statistical Planning and Inference* 44 (2): 249–63. https://doi.org/10.1016/0378-3758(94)00047-y.

Kuhn, Max. 2022. "Caret: Classification and Regression Training." https://CRAN.R-project.org/package=caret.

Langton, R., Boulcott, P. and Wright P.J. 2021. "A verified distribution model for the lesser sandeel Ammodytes marinus." *Marine Ecology Progress Series*. https://doi.org/10.3354/meps13693.

Mitkus, M., Nevitt, G.A., Danielsen, J. and Kelber, A., 2016. Vision on the high seas: spatial resolution and optical sensitivity in two procellariiform seabirds with different foraging strategies. Journal of Experimental Biology, 219(21), pp.3329-3338.

Scott-Hayward, L., Oedekoven,C.. 2013. "User Guide for the MRSea Package: Statistical Modelling of Bird and Cetacean Distributions in Offshore Renewables Development Areas." *University of St. Andrews Contract for Marine Scotland*.

Ver Hoef, J. M., & Boveng, P. L. 2007. "Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data?" *Ecology*, *88*(11): 2766-2772.

Walker, C. G., M. L. Mackenzie, C. R. Donovan, and M. J. O'Sullivan. 2011. "SALSA: a Spatially Adaptive Local Smoothing Algorithm." *Journal of Statistical Computation and Simulation* 81 (2): 179–91. https://doi.org/10.1080/00949650903229041.

Wright, Marvin N., and Andreas Ziegler. 2017. "Ranger: A Fast Implementation of Random Forests for High Dimensional Data in c++ and r" 77. https://doi.org/10.18637/jss.v077.i01.

I Exploratory Data Analysis

I.1 Introduction

In order to properly model the data we first performed an exploratory data analysis to understand how many of each species were in each survey, and how they were spatially distributed throughout the study area. We did this by examining the observation data by species, then mapping species distribution in the study area for the smallest temporal scale that the number of observations would allow. The target temporal scale was by survey, but if the number of observations was too low, we try breeding season, then survey year, and finally for all surveys. MRSea models can fail with a minimal number of observations, therefore we used a threshold of 20 observations per species, Where possible, we mapped the observations plotting behaviours as sitting (red) or flying (blue). We found 624 observations of animals that were not identified to species level, which could supplement some data-poor surveys for some species through apportioning. We were confident we have sufficient data to model seven out of thirteen of the key species (See table I-1). For guillemot, puffin, and razorbill, we did not suggest modelling unique behaviours (sitting/flying) as these species do not fly high enough to be considered at risk for collision with a turbine. We summarise our findings in table I-1, then provide more detail and plots below.

I.2 Summary Table

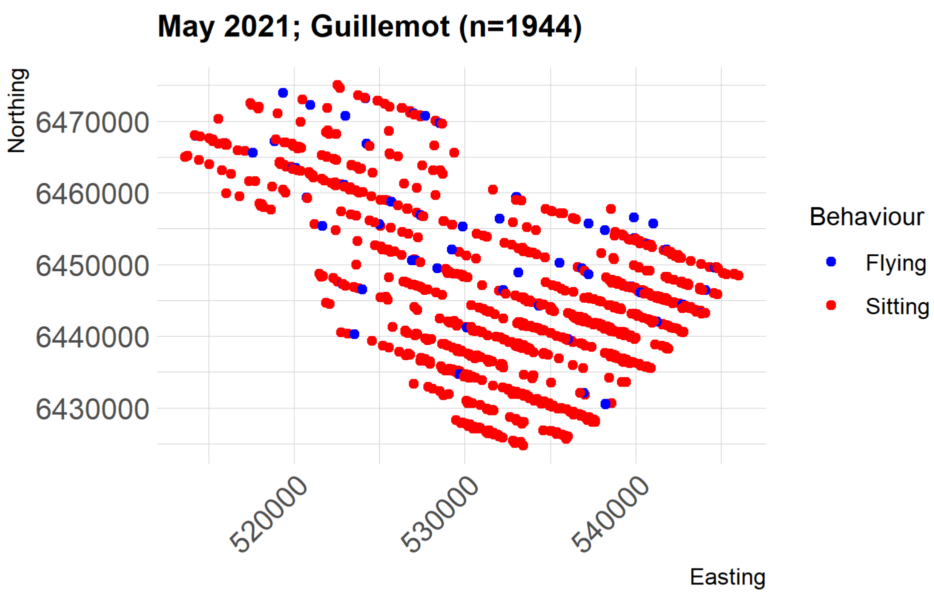
Table I-1. Summary of all key species, including whether they have sufficient observations for modelling, at which temporal scale (survey, breeding season, survey year, or all surveys) we suggest that they should be modelled (# surveys that can be modelled / # total surveys are in brackets), and which behaviours we can model. Number of available surveys are in brackets.

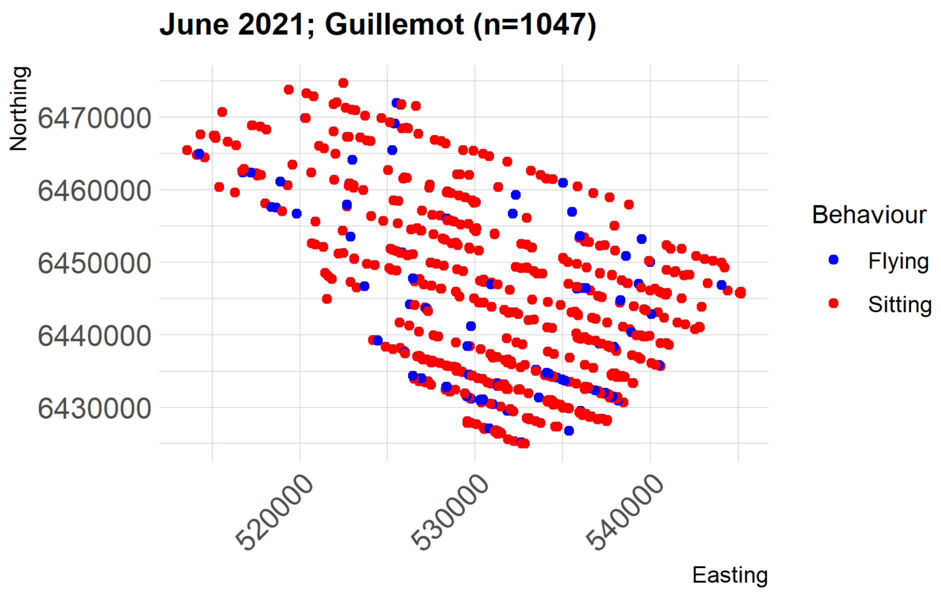
Species	Can be modelled?	Temporal Scale?	Behaviour
Guillemot	Yes	Survey (22/24)	All
Razorbill	Yes	Survey (15/24)	All
Kittiwake	Yes	Survey (19/24)	All, Fly
Fulmar	Yes	Survey (22/24)	All, Fly
Puffin	Yes	Survey (13/24)	All
Gannet	Yes	Survey (10/24)	All, Fly
Great Black-backed Gull	Yes	Survey Year	All, Fly
Herring Gull	Yes	All Surveys	All
Great Skua	No		
Manx Shearwater	No		
Common Gull	No		
Arctic Tern	No		
Common Tern	No		

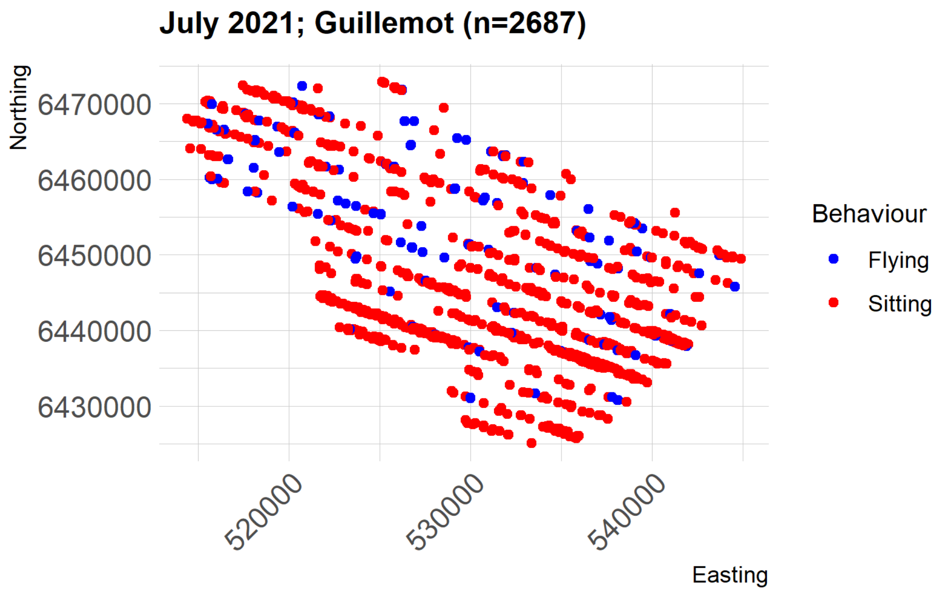
I.3 Results by Species

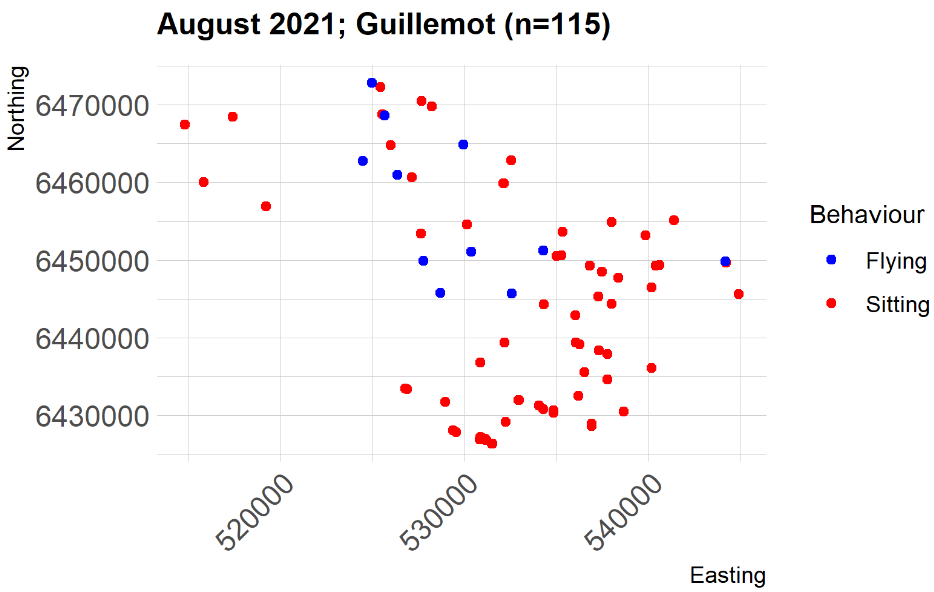
I.3.1 Guillemot

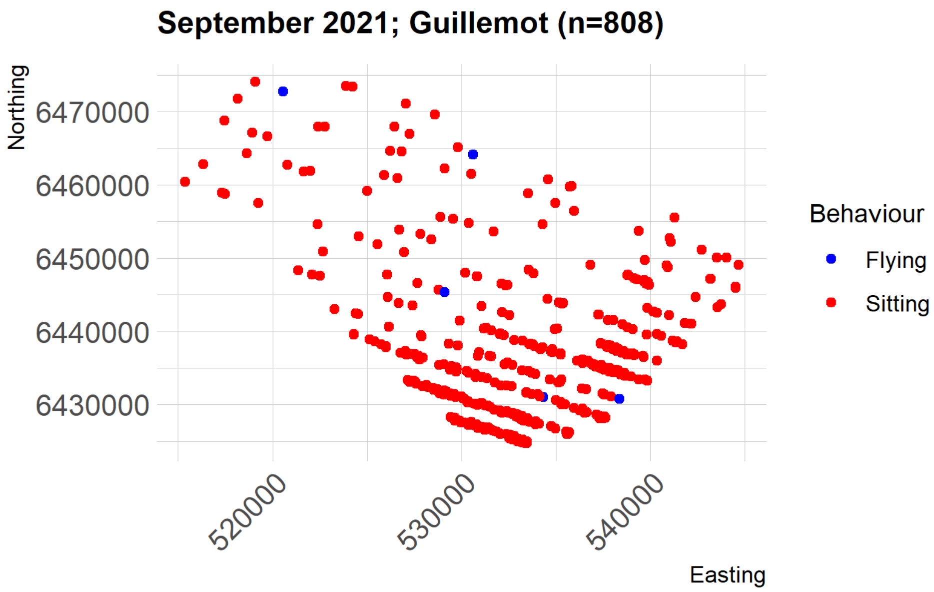
Guillemot are the most common species throughout all surveys consisting of almost half of all observations (N=20135). Sufficient observations were recorded for 22/24 surveys. Surveys 8 and 19 (Dec 2021, Nov 2022) do not have enough observations (more than 20), and will not be included in modelling. We could also model guillemot by breeding season (April - 15 August). The August surveys from both years fall within the breeding season (01 Aug 2021, 10 Aug 2022). We suggest modelling all behaviours (flying & sitting) together as guillemot do not fly high enough to be at risk for collision with turbines.

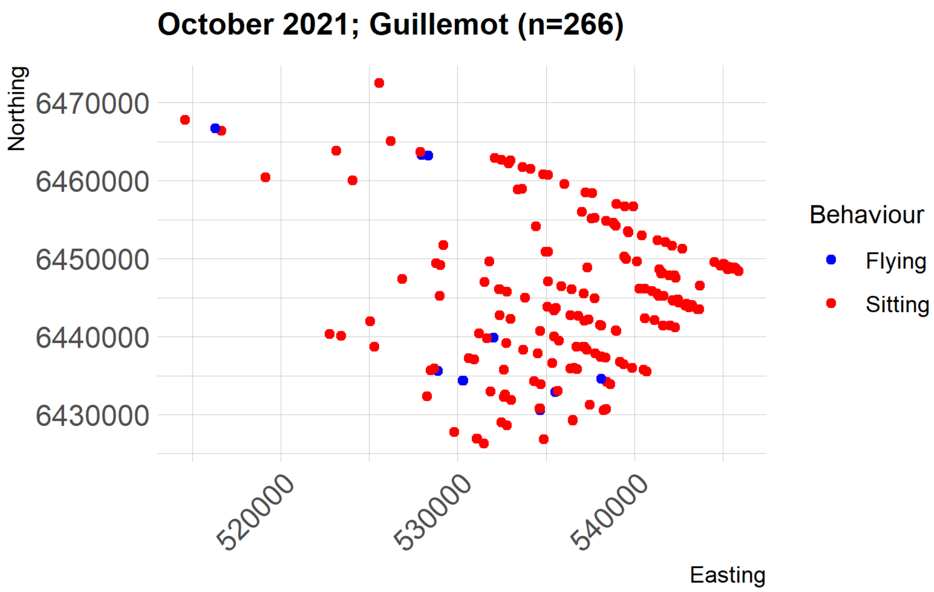


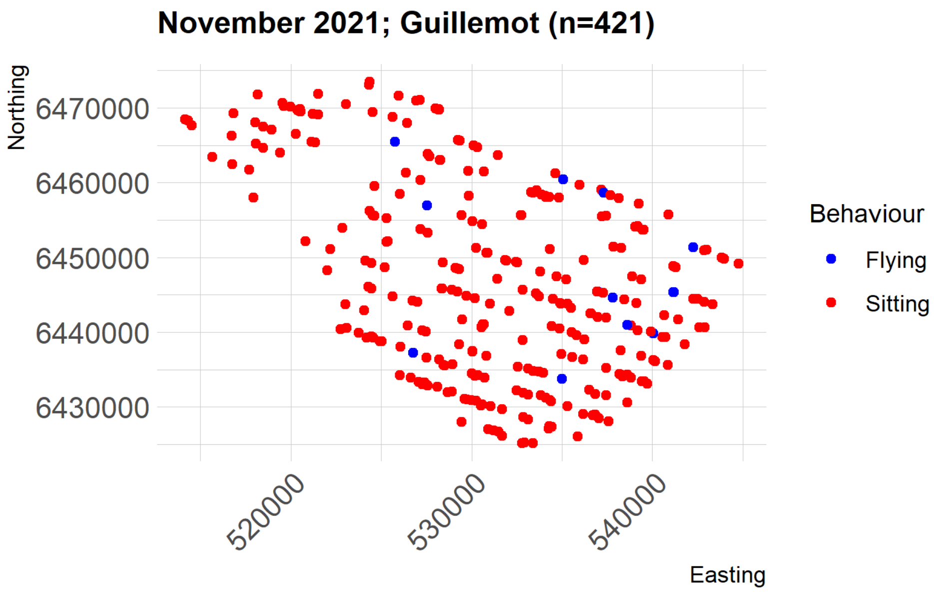

June 2021; Guillemot (n=1047)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

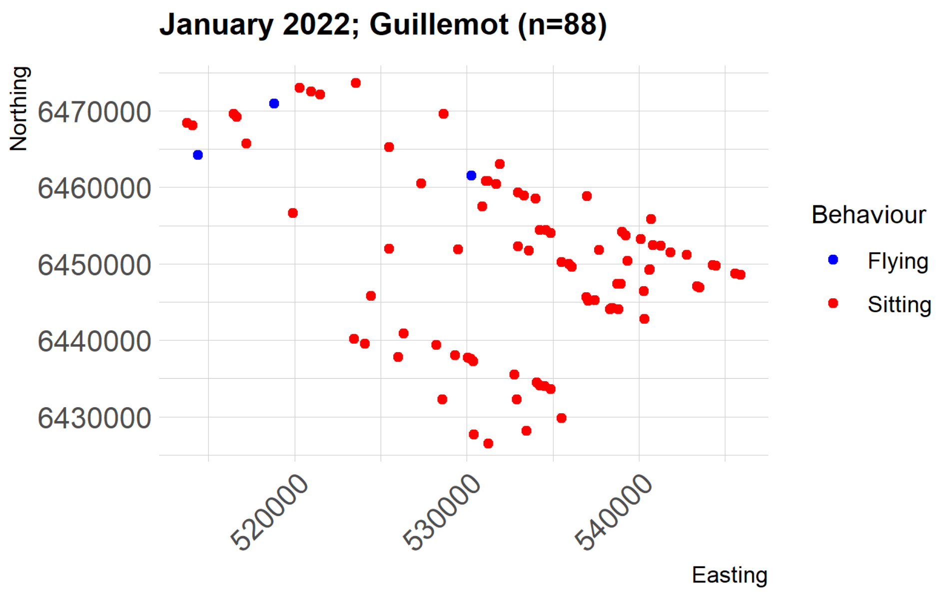

July 2021; Guillemot (n=2687)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

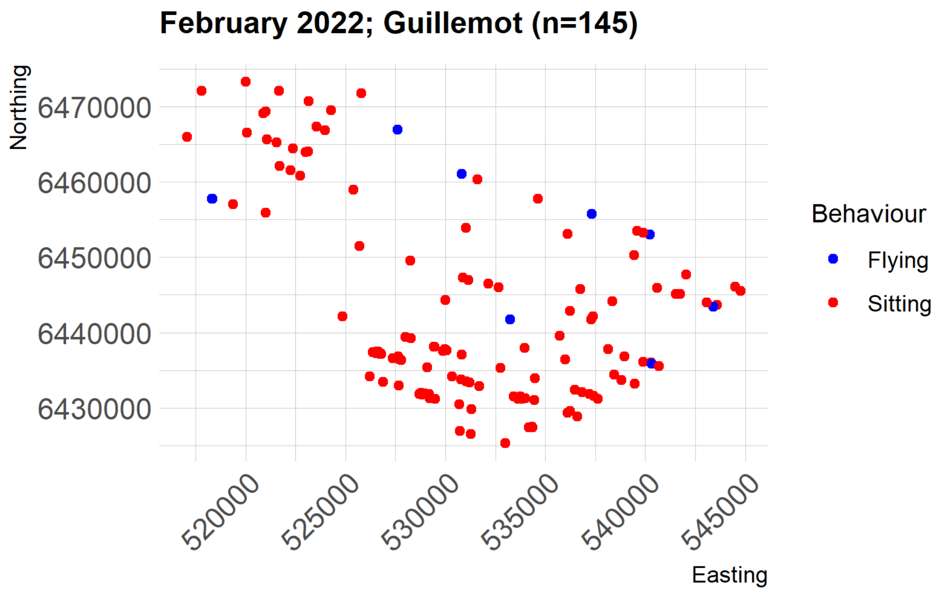

August 2021; Guillemot (n=115)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

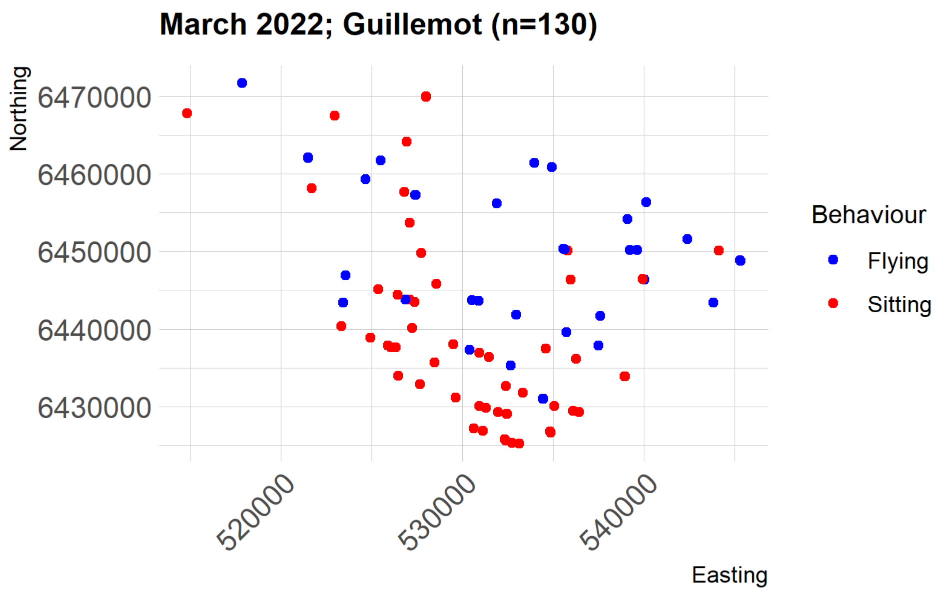

September 2021; Guillemot (n=808)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

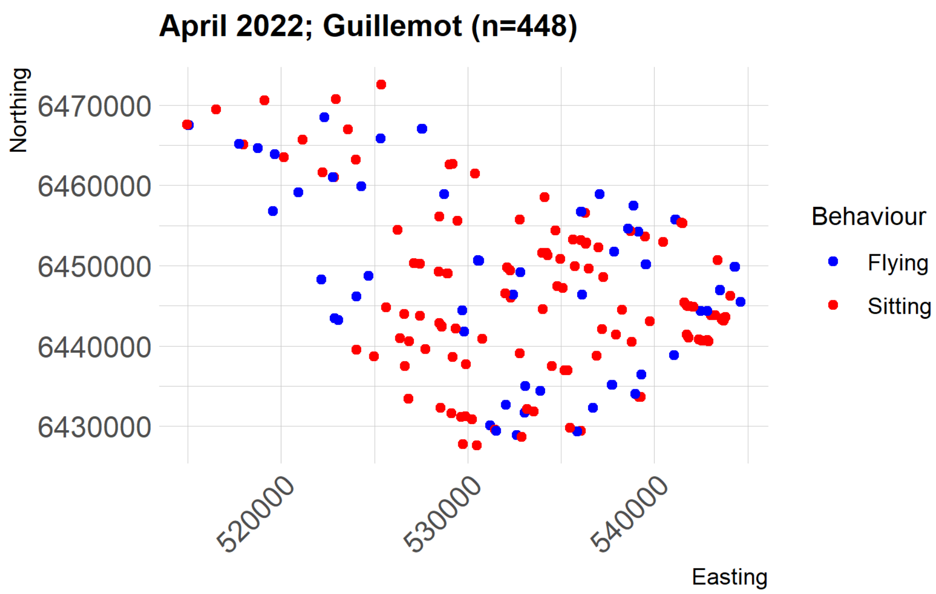

October 2021; Guillemot (n=266)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

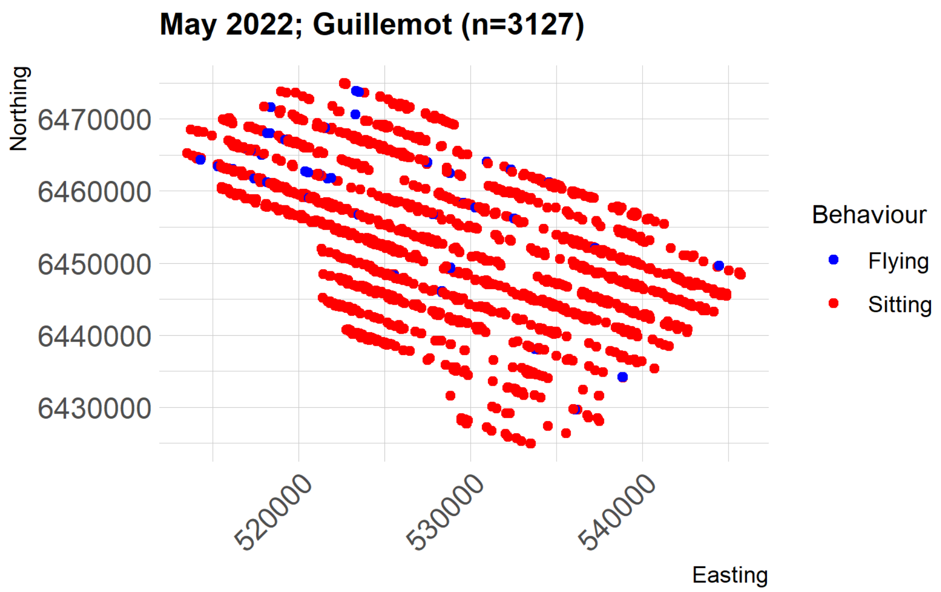

November 2021; Guillemot (n=421)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

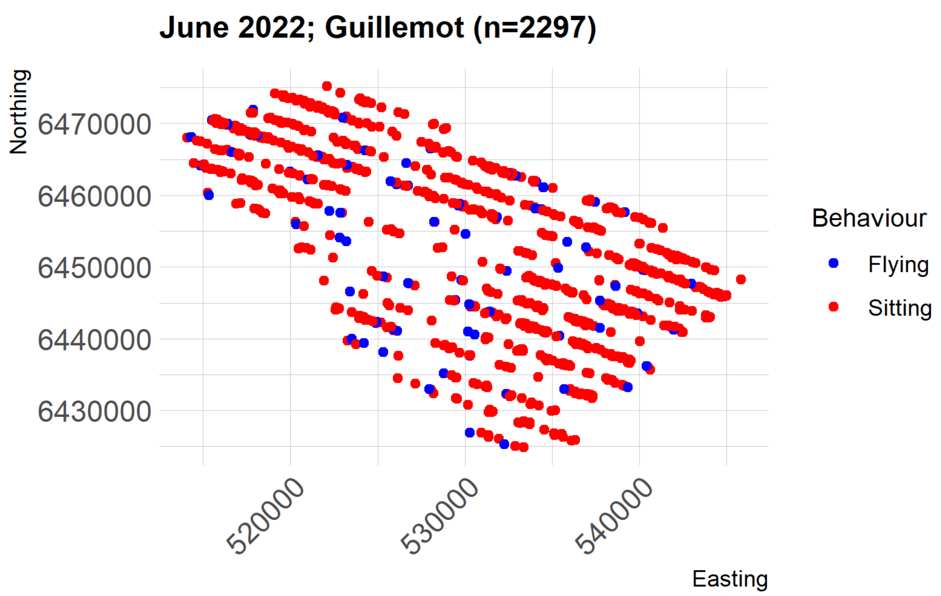

January 2022; Guillemot (n=88)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

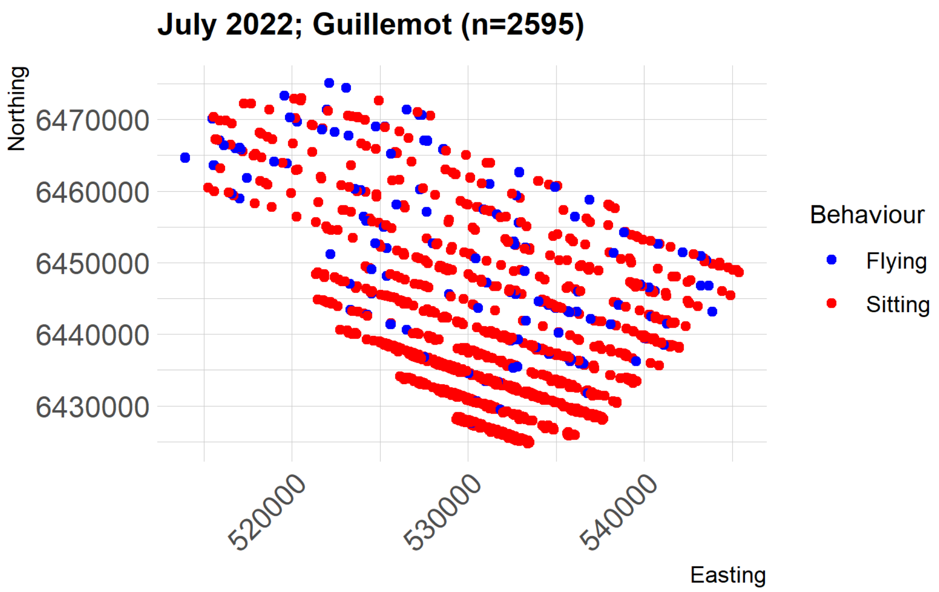

February 2022; Guillemot (n=145)
Northing
6470000
6460000
6450000
6440000
6430000
520000
525000
530000
535000
540000
545000
Easting
Behaviour
Flying
Sitting

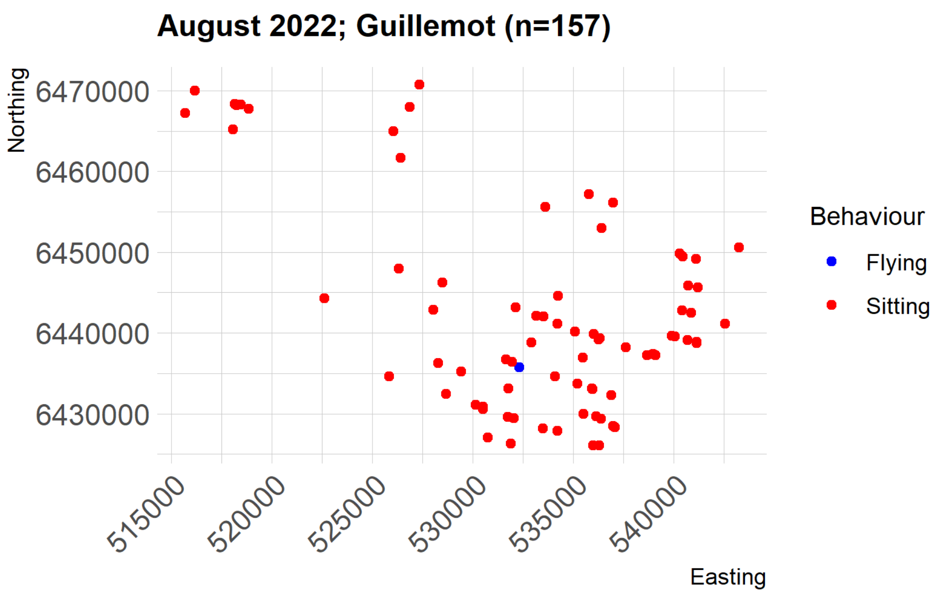

March 2022; Guillemot (n=130)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

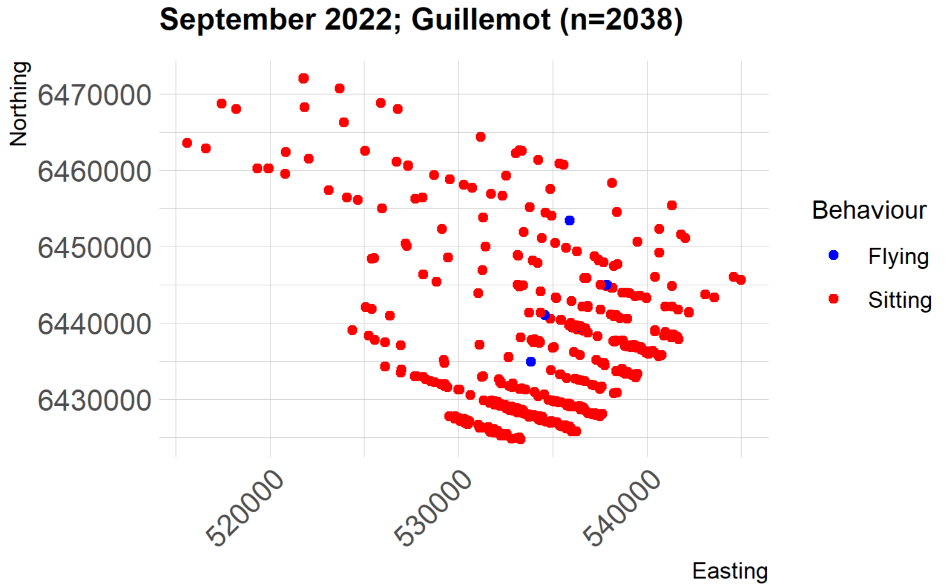

April 2022; Guillemot (n=448)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

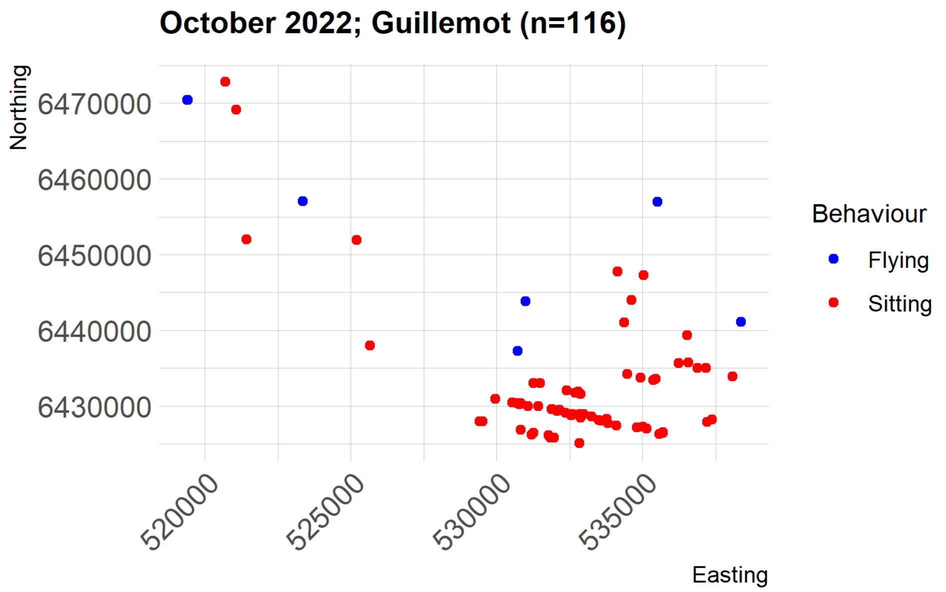

May 2022; Guillemot (n=3127)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

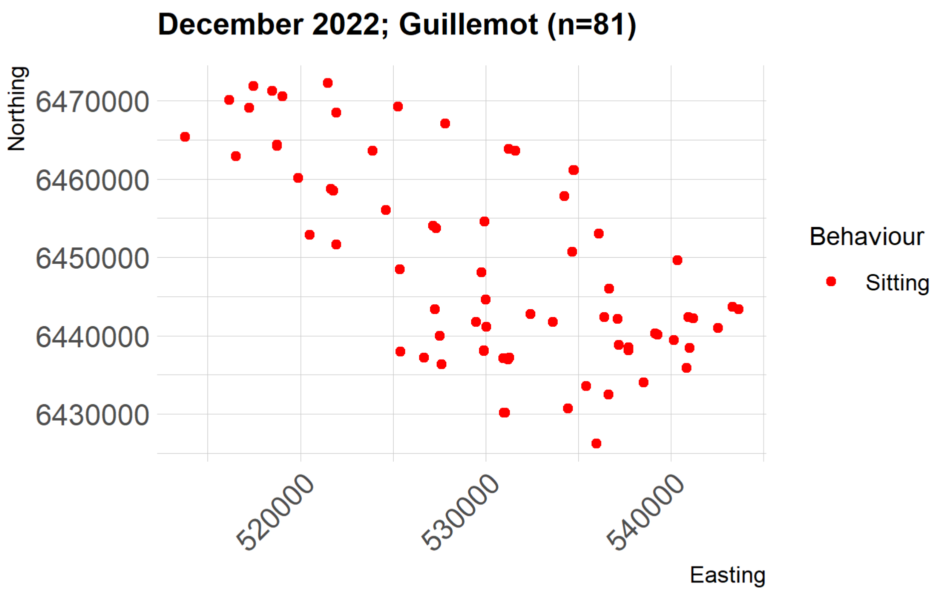

June 2022; Guillemot (n=2297)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

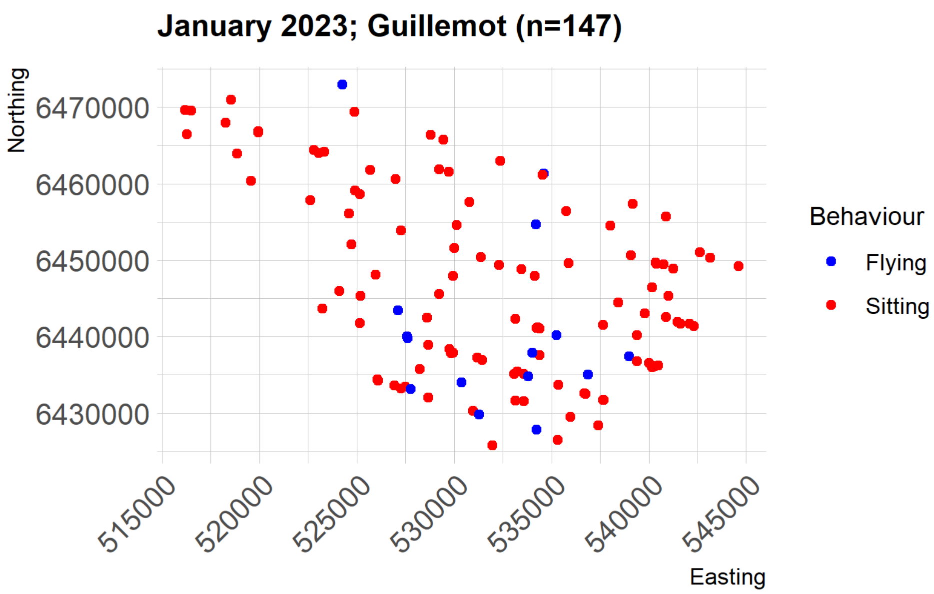

July 2022; Guillemot (n=2595)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

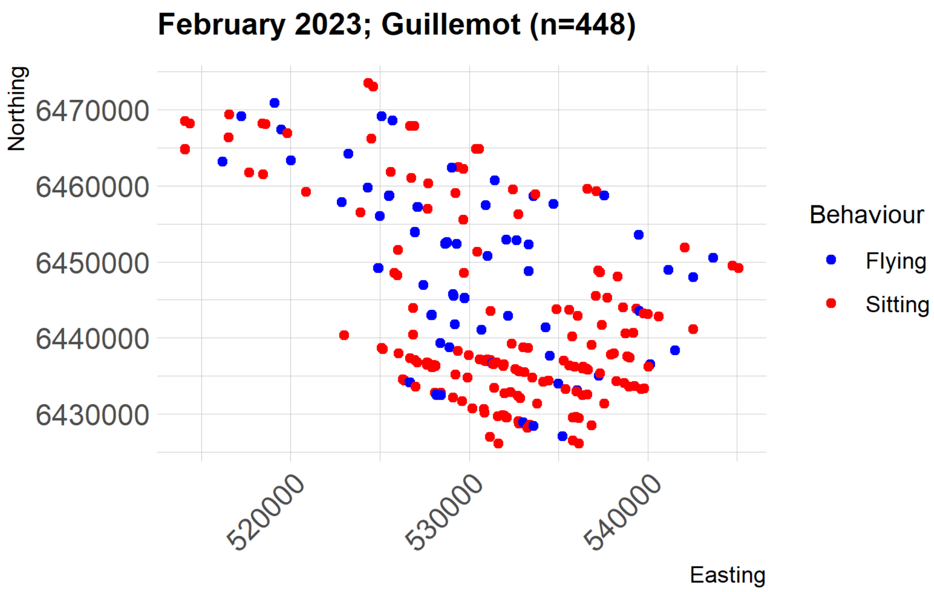

August 2022; Guillemot (n=157)
Northing
6470000
6460000
6450000
6440000
6430000
515000
520000
525000
530000
535000
540000
Easting
Behaviour
Flying
Sitting

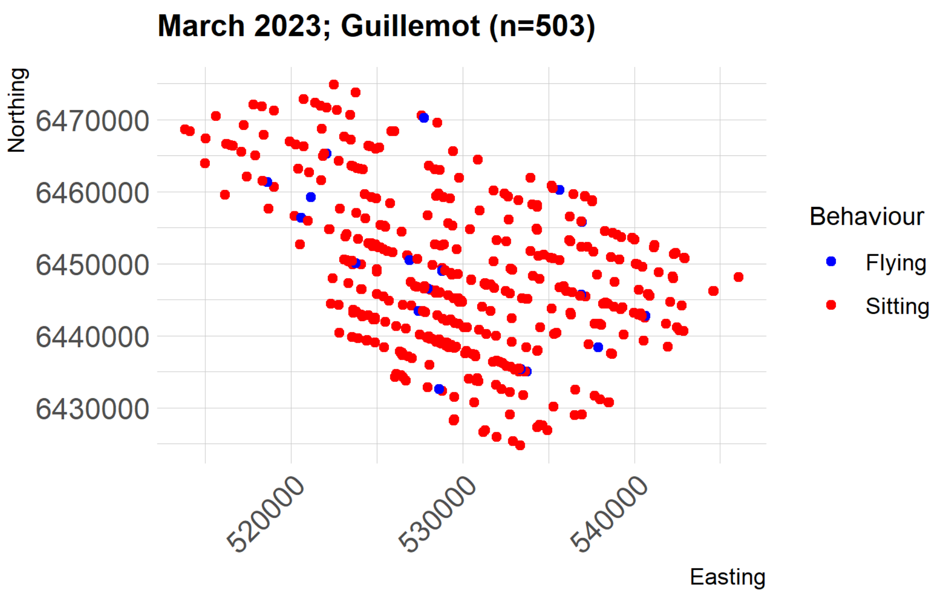

September 2022; Guillemot (n=2038)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

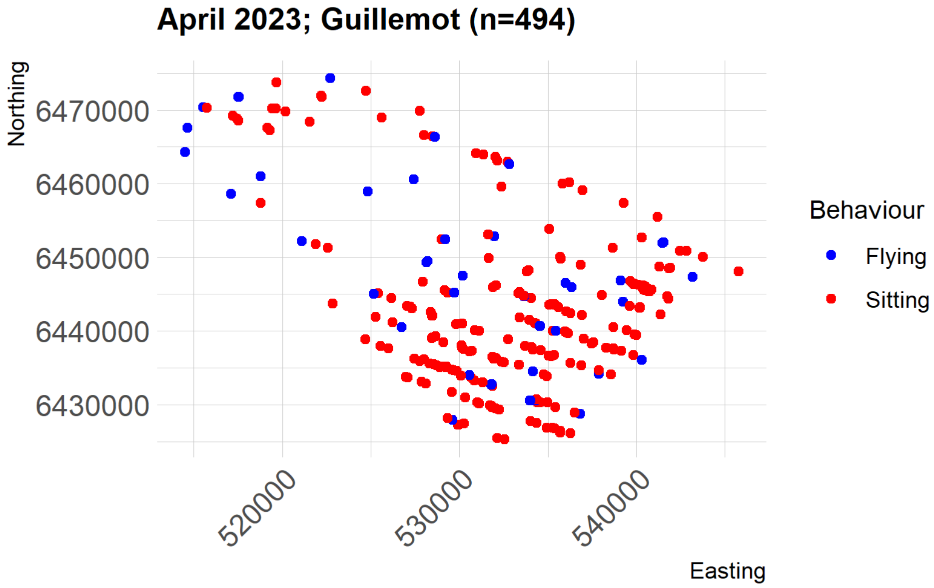

October 2022; Guillemot (n=116)
Northing
6470000
6460000
6450000
6440000
6430000
520000
525000
530000
535000
Easting
Behaviour
Flying
Sitting

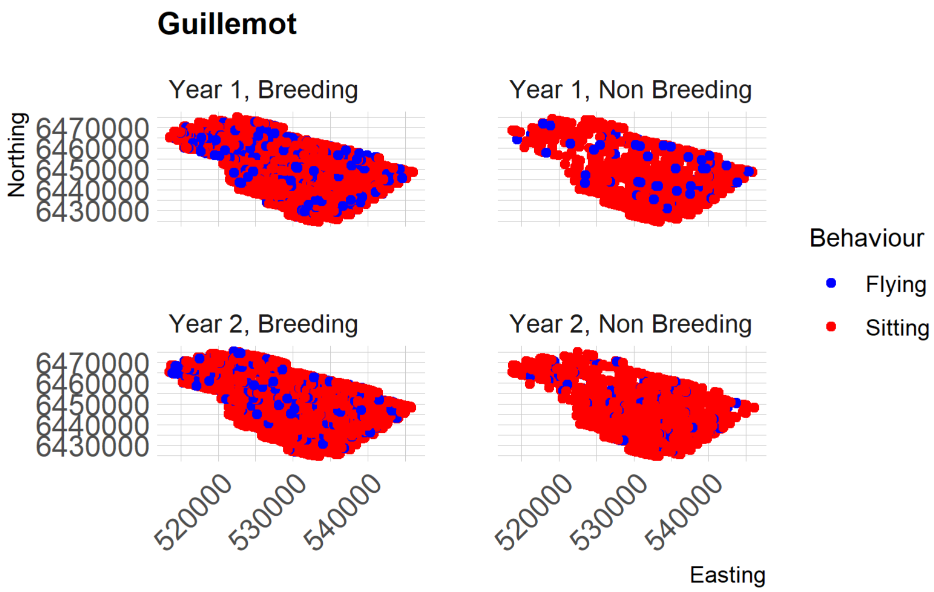

December 2022; Guillemot (n=81)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Sitting


January 2023; Guillemot (n=147)
Northing
6470000
6460000
6450000
6440000
6430000
515000
520000
525000
530000
535000
540000
545000
Easting
Behaviour
Flying
Sitting


February 2023; Guillemot (n=448)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

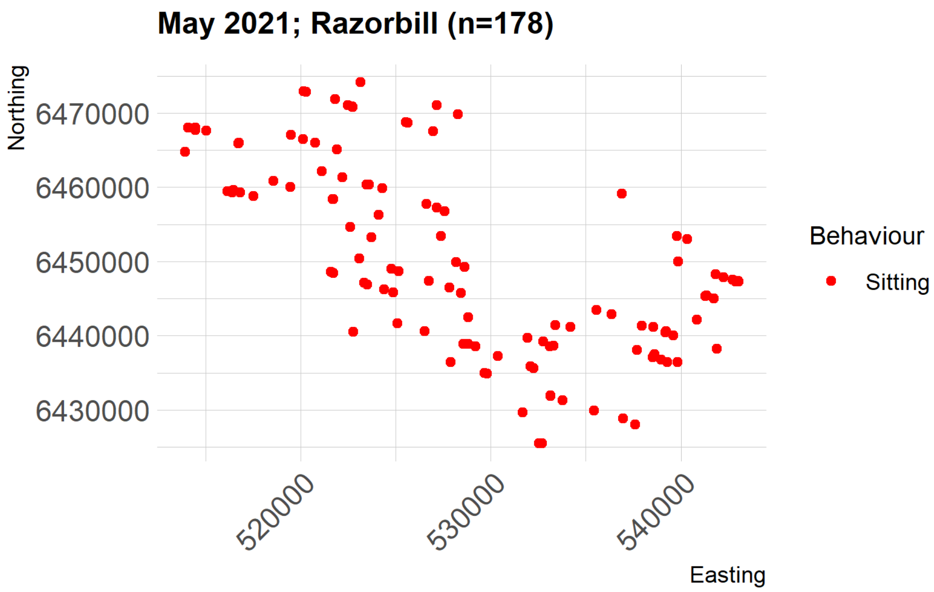

March 2023; Guillemot (n=503)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

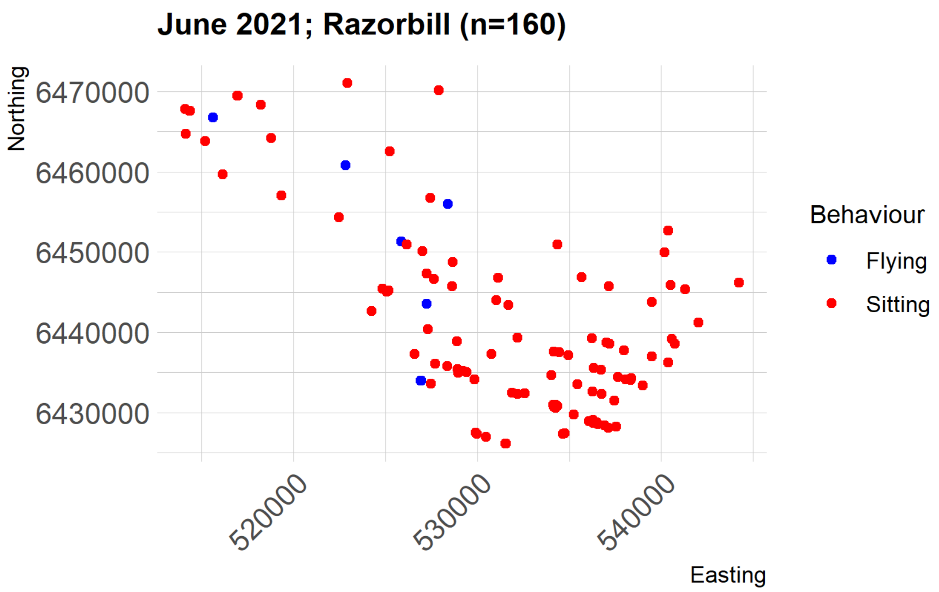

April 2023; Guillemot (n=494)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

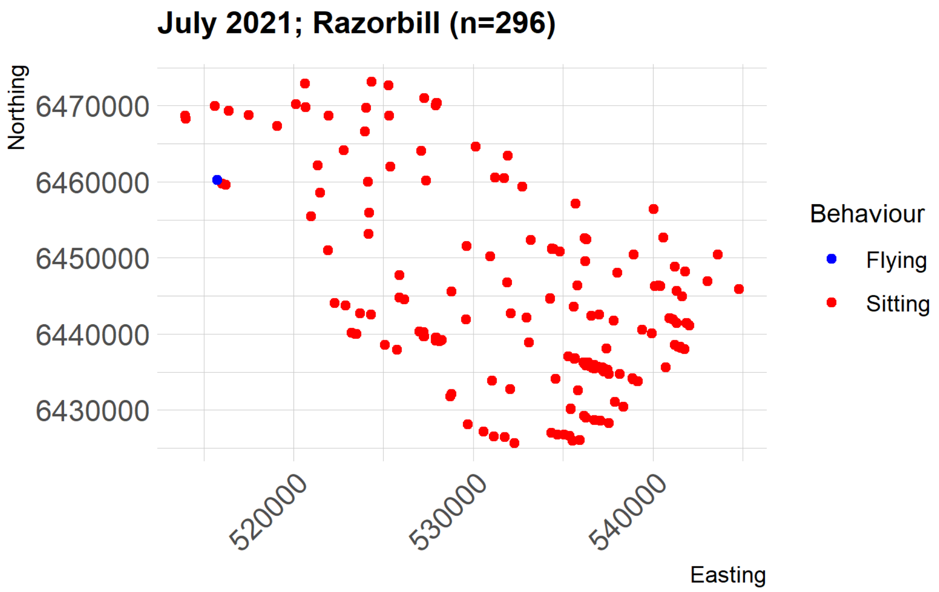

Guillemot
Year 1, Breeding
Year 1, Non Breeding
Year 2, Breeding
Year 2, Non Breeding
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

I.3.2 Razorbill

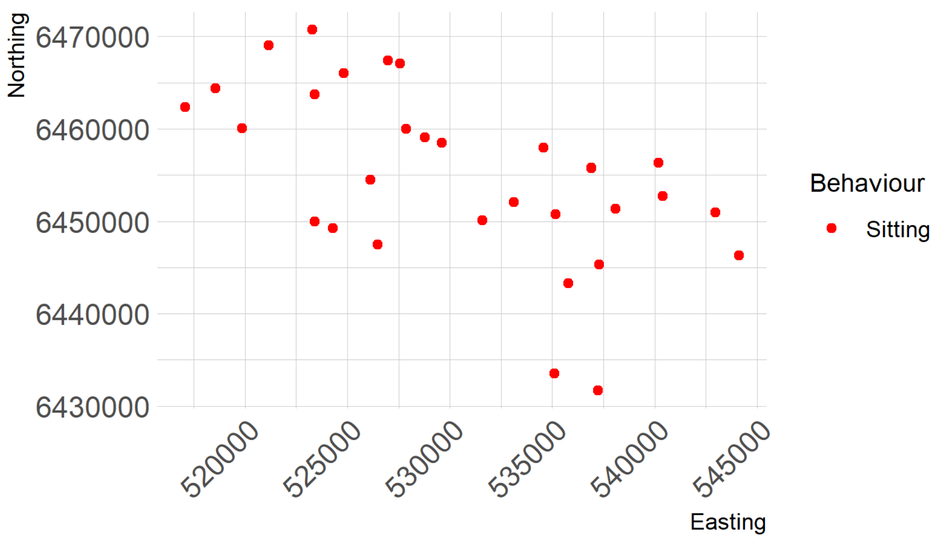
We suggest modelling razorbill (N=2542) by survey for the surveys that had sufficient observations (15/24). We could also model razorbill by breeding season (April - 15 August). The August surveys from both years fall within the breeding season (01 Aug 2021, 10 Aug 2022). We suggest modelling all behaviours (flying & sitting) together as razorbill do not fly high enough to be at risk for collision with turbines.



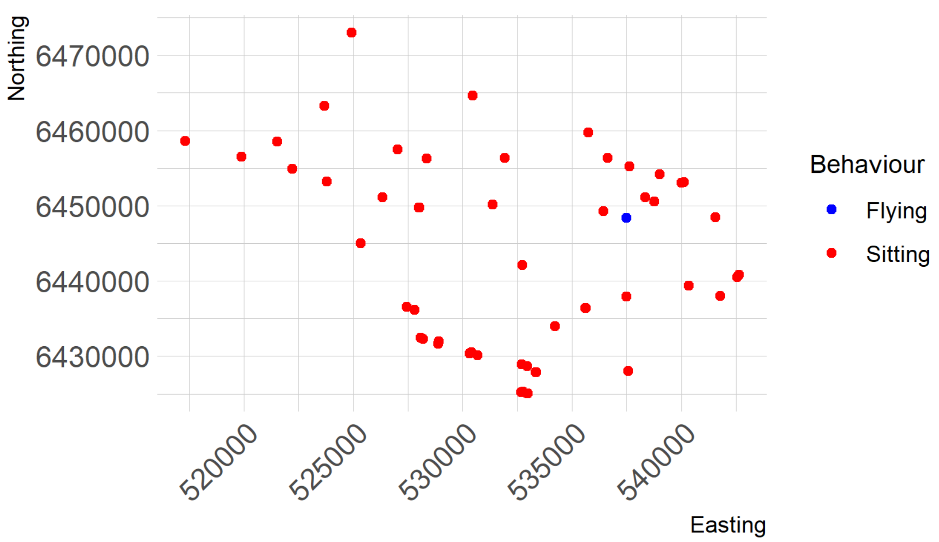

June 2021; Razorbill (n=160)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

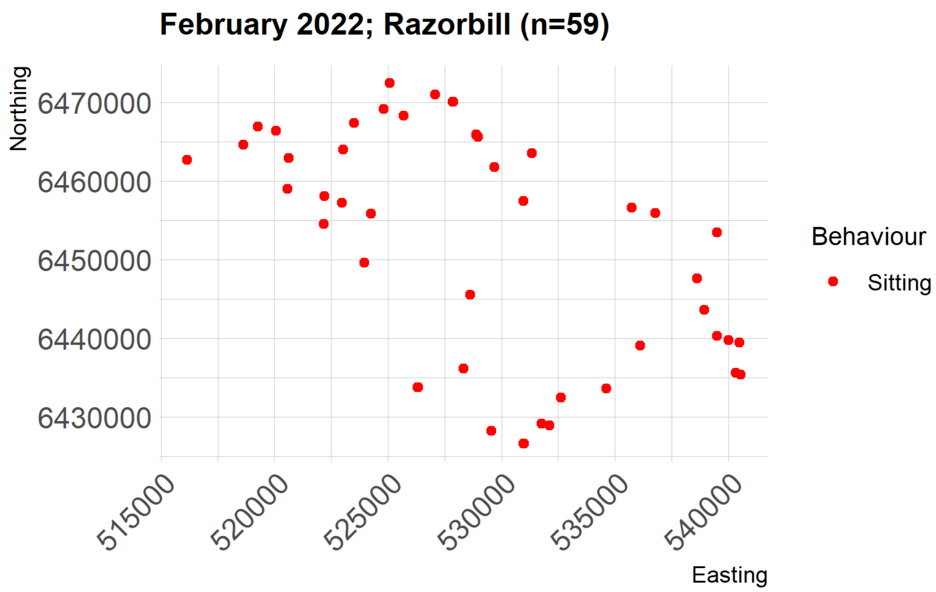

July 2021; Razorbill (n=296)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

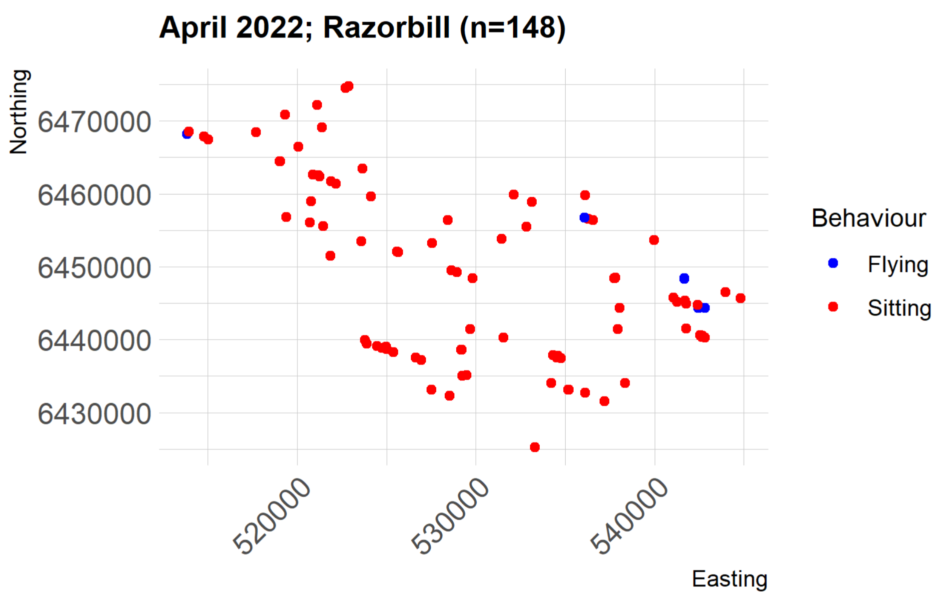
August 2021; Razorbill (n=61)

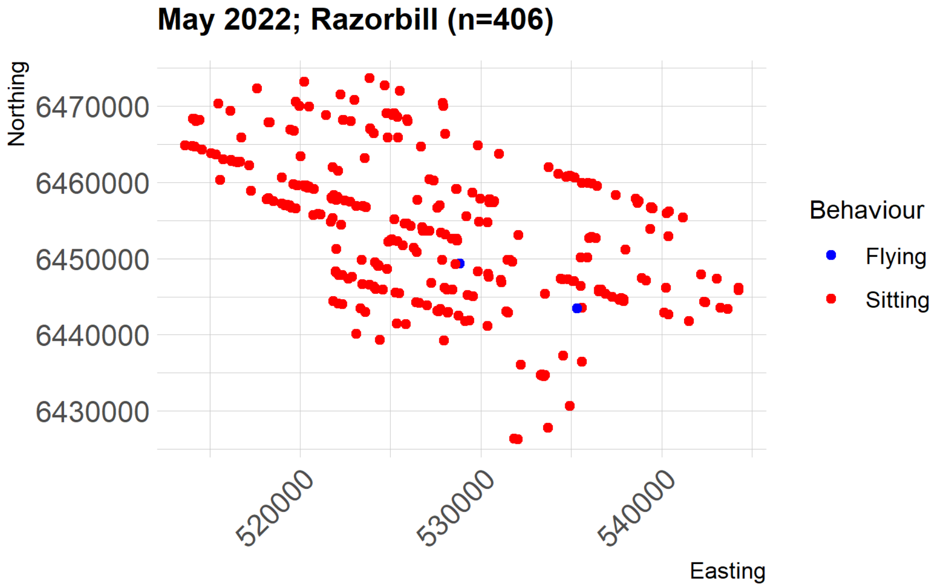


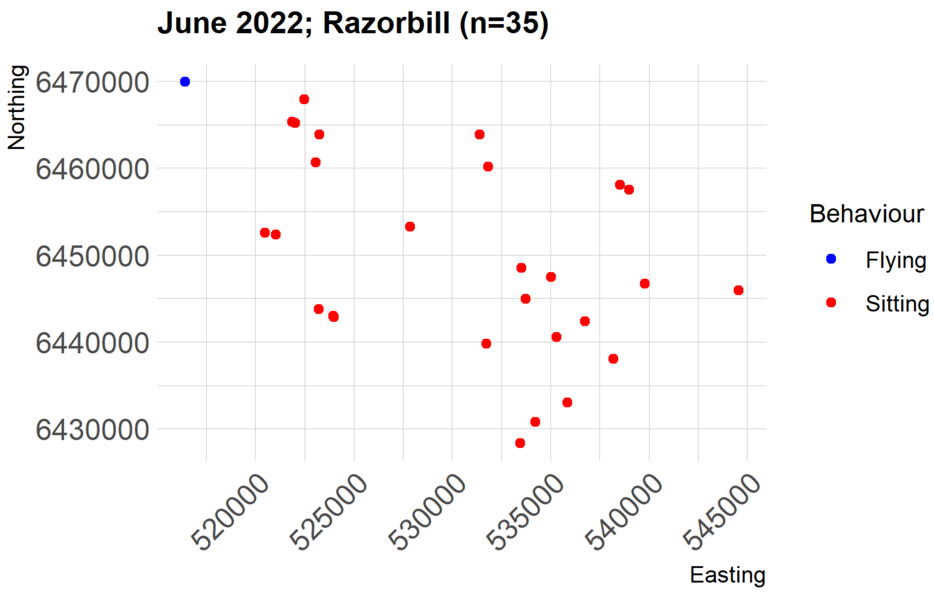
September 2021; Razorbill (n=80)

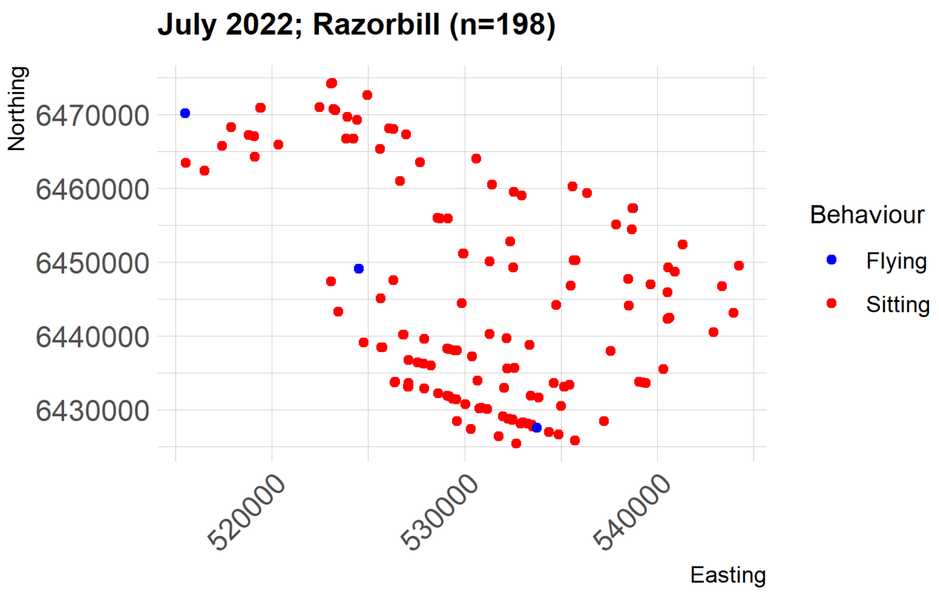


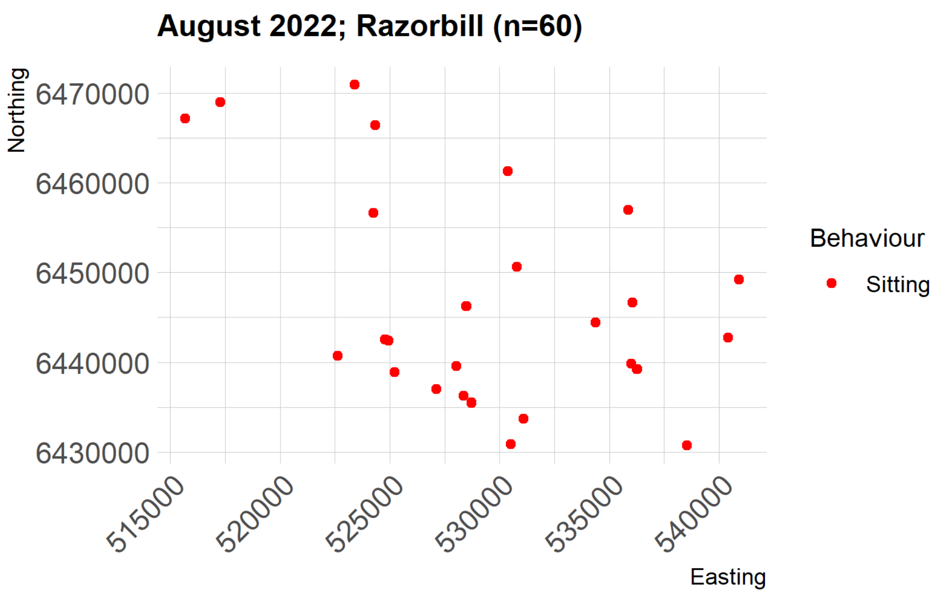

February 2022; Razorbill (n=59)
Northing
6470000
6460000
6450000
6440000
6430000
515000
520000
525000
530000
535000
540000
Easting
Behaviour
Sitting

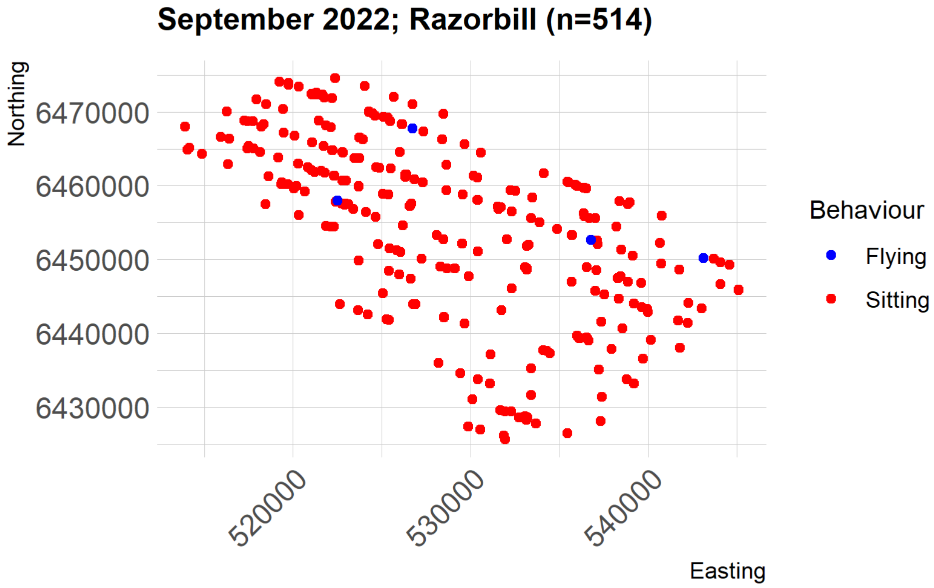

April 2022; Razorbill (n=148)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

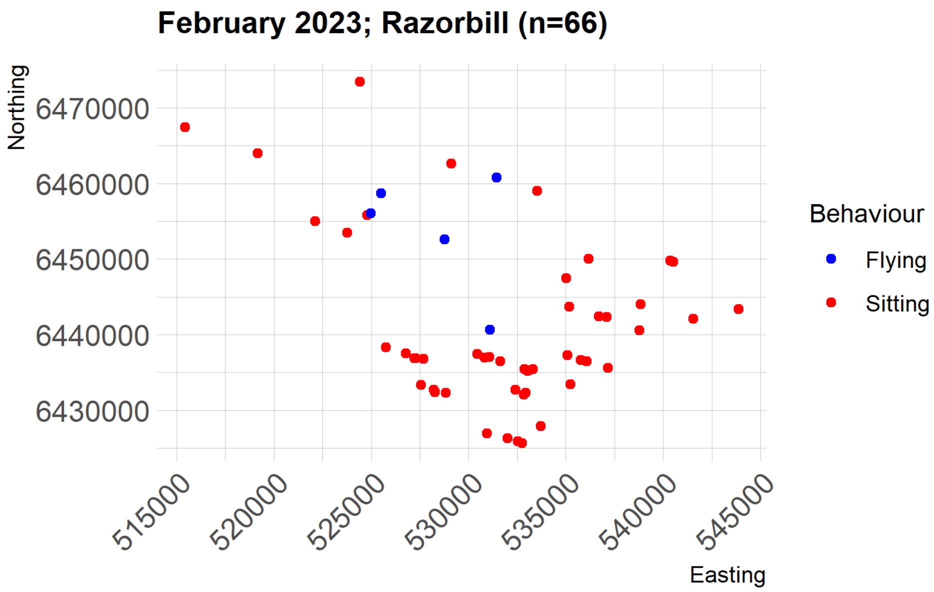

May 2022; Razorbill (n=406)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

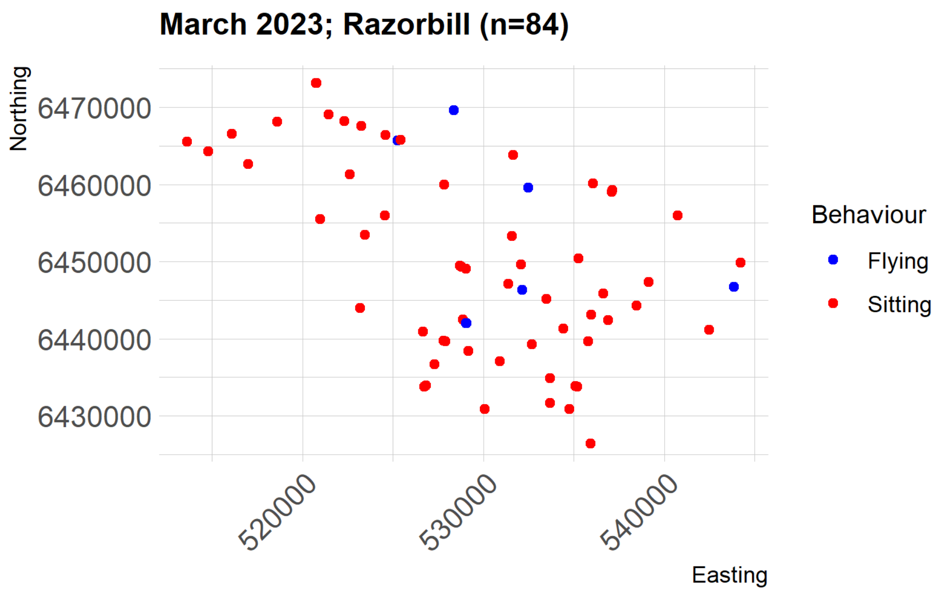

June 2022; Razorbill (n=35)
Northing
6470000
6460000
6450000
6440000
6430000
520000
525000
530000
535000
540000
545000
Easting
Behaviour
Flying
Sitting

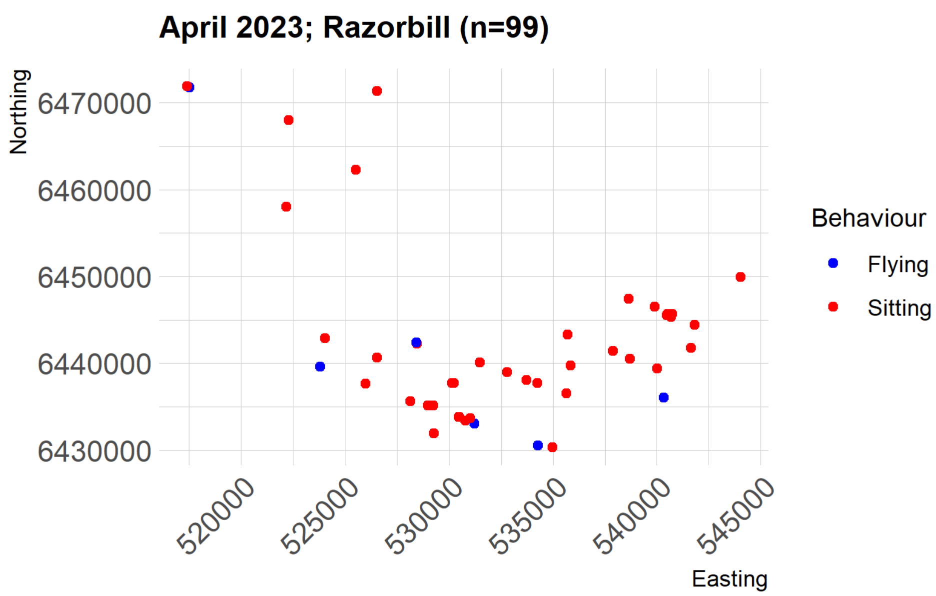

July 2022; Razorbill (n=198)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

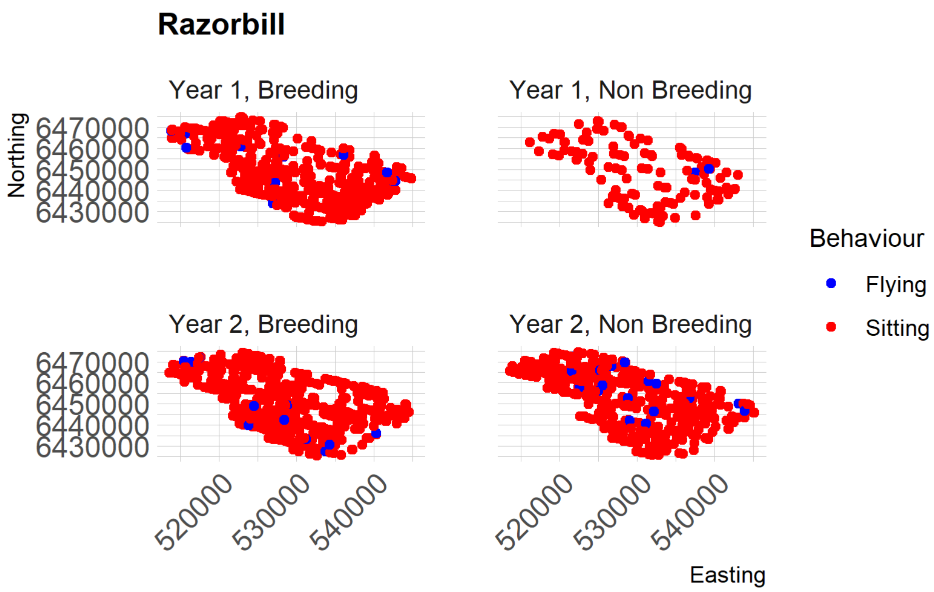

August 2022; Razorbill (n=60)
Northing
6470000
6460000
6450000
6440000
6430000
515000
520000
525000
530000
535000
540000
Easting
Behaviour
Sitting


September 2022; Razorbill (n=514)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting


February 2023; Razorbill (n=66)
Northing
6470000
6460000
6450000
6440000
6430000
515000
520000
525000
530000
535000
540000
545000
Easting
Behaviour
Flying
Sitting

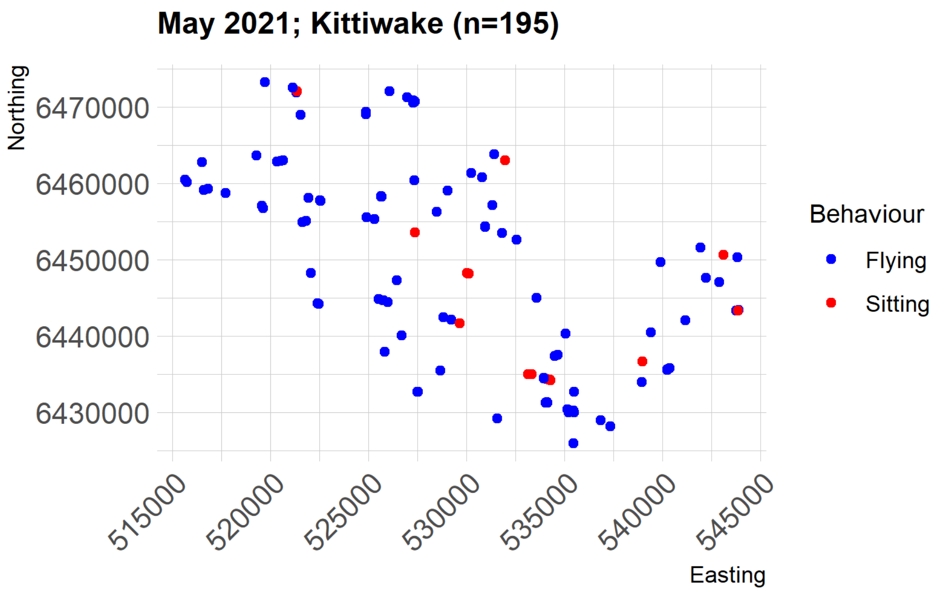

March 2023; Razorbill (n=84)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

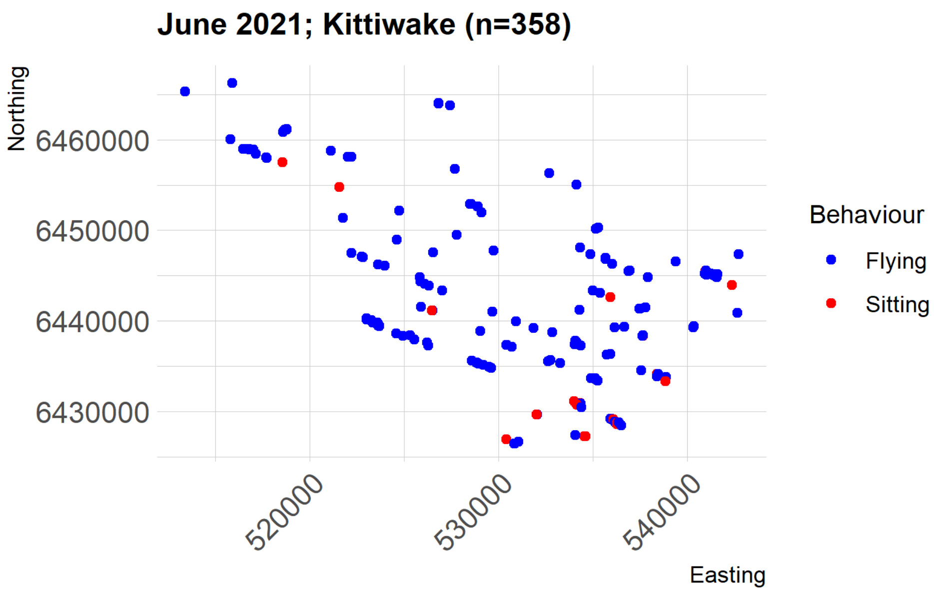

April 2023; Razorbill (n=99)
Northing
6470000
6460000
6450000
6440000
6430000
520000
525000
530000
535000
540000
545000
Easting
Behaviour
Flying
Sitting

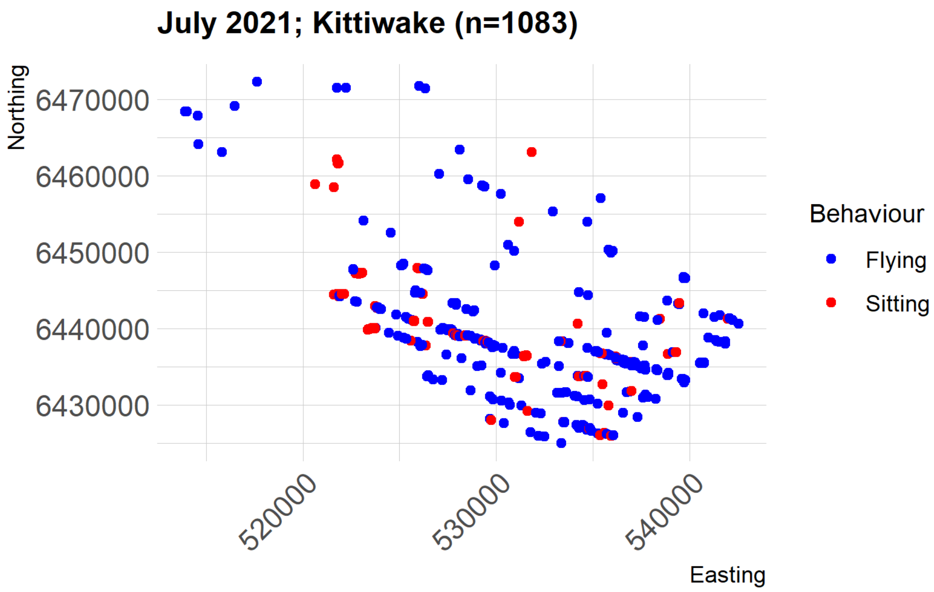

Razorbill
Year 1, Breeding
Year 1, Non Breeding
Year 2, Breeding
Year 2, Non Breeding
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

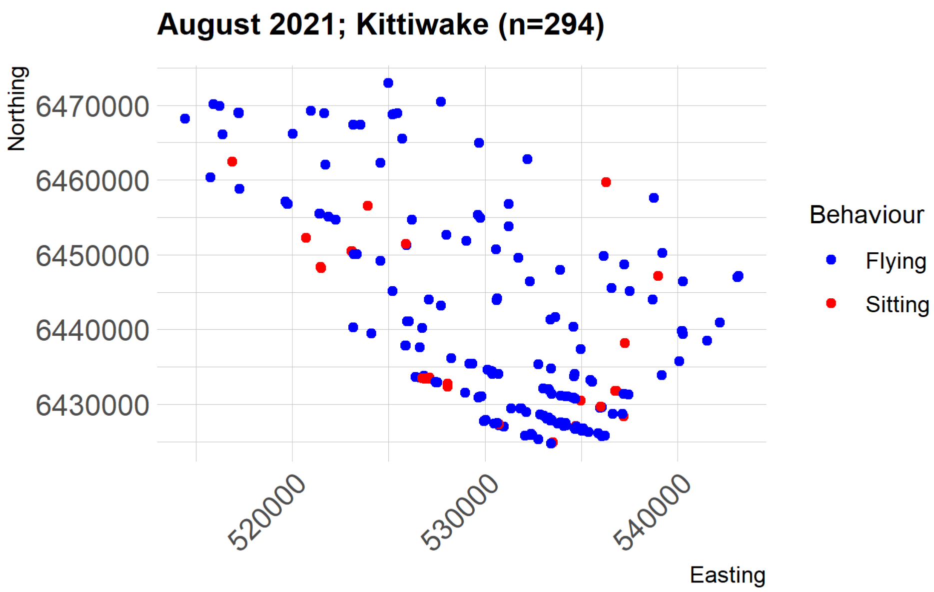
I.3.3 Kittiwake

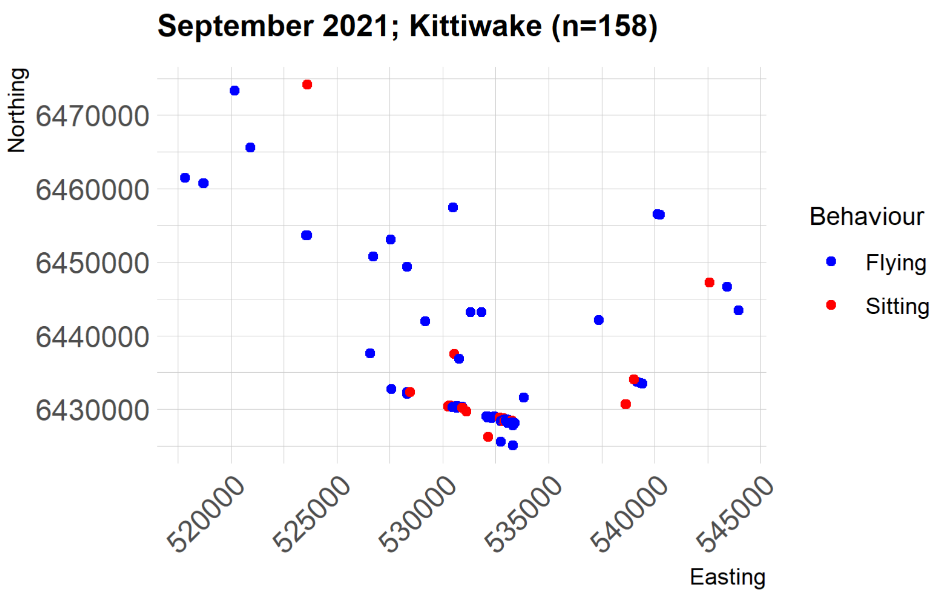
We suggest modelling kittiwake (N=5366) by survey for the surveys that had sufficient observations (19/24). We have sufficient observations of kittiwake to include behaviour in the model. We could also model by breeding season (15 April - August). The April 2023 survey does fall outside of the breeding season (4 Apr 2023) and is therefore accounted for as "Non-Breeding." The April 2022 survey (26 Apr 2022), however does fall within the breeding season.

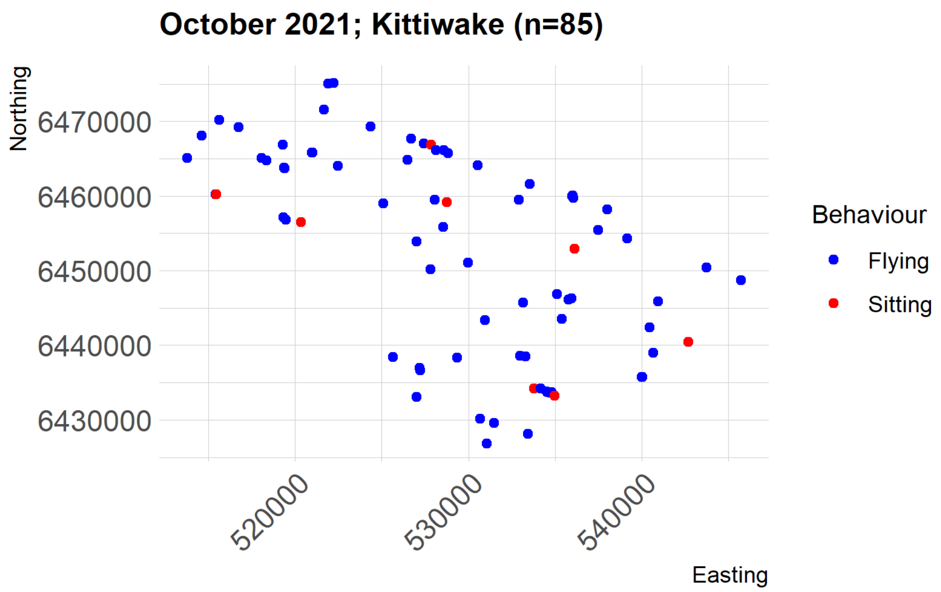


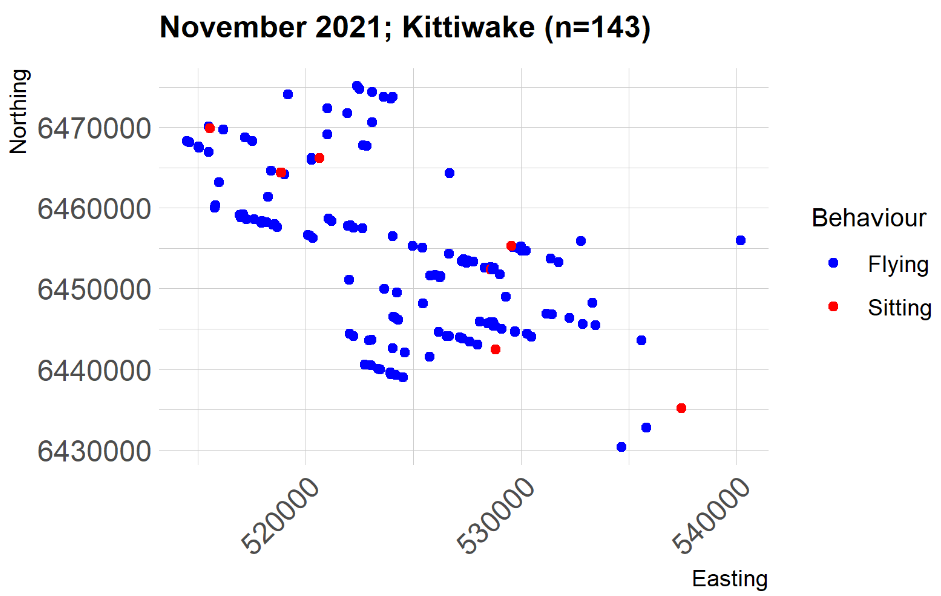

June 2021; Kittiwake (n=358)
Northing
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

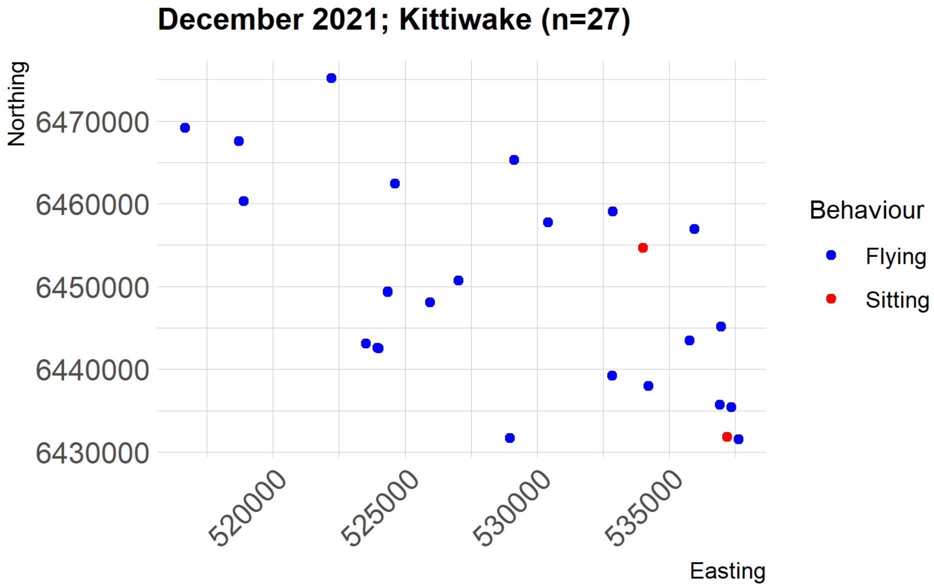

July 2021; Kittiwake (n=1083)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

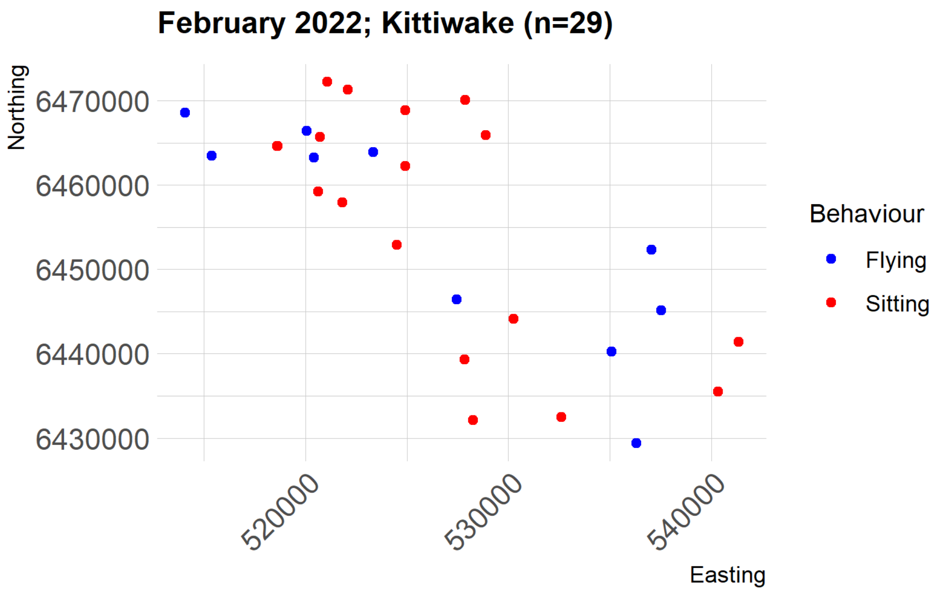

August 2021; Kittiwake (n=294)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

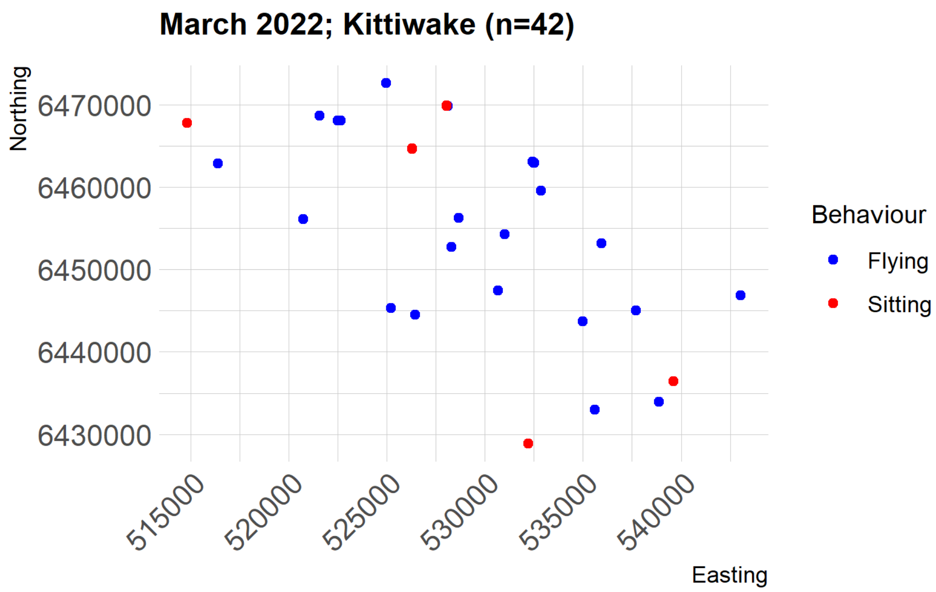

September 2021; Kittiwake (n=158)
Northing
6470000
6460000
6450000
6440000
6430000
520000
525000
530000
535000
540000
545000
Easting
Behaviour
Flying
Sitting

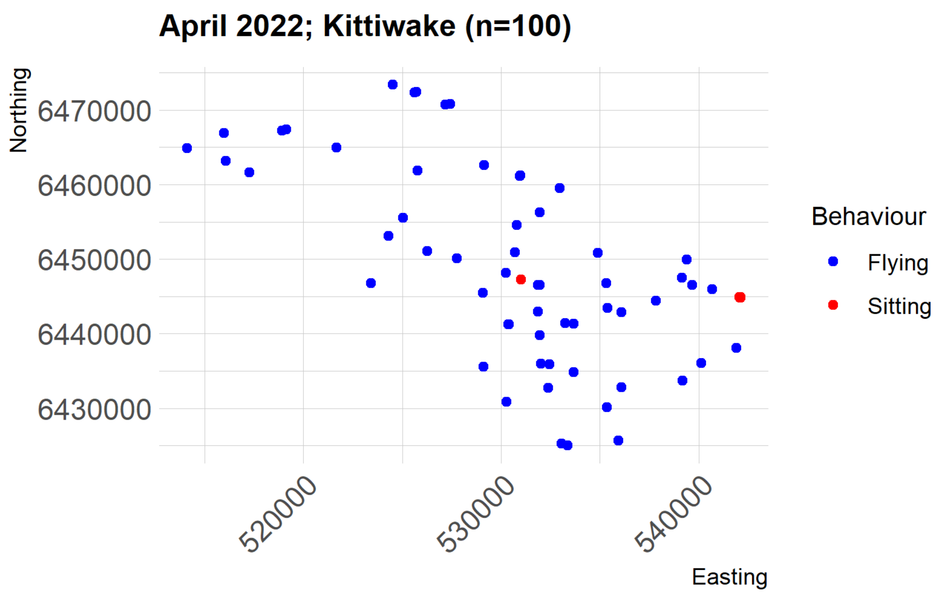

October 2021; Kittiwake (n=85)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

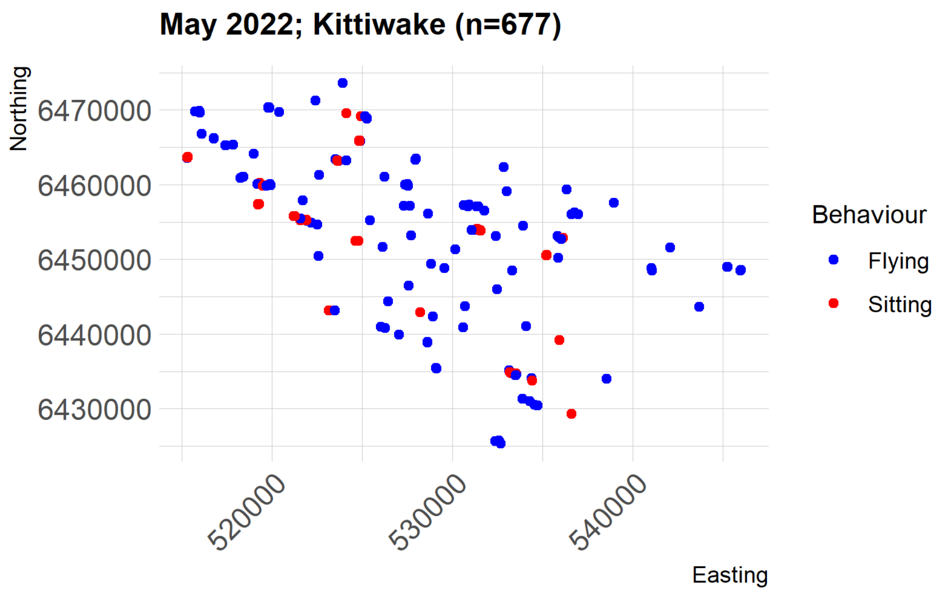

November 2021; Kittiwake (n=143)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

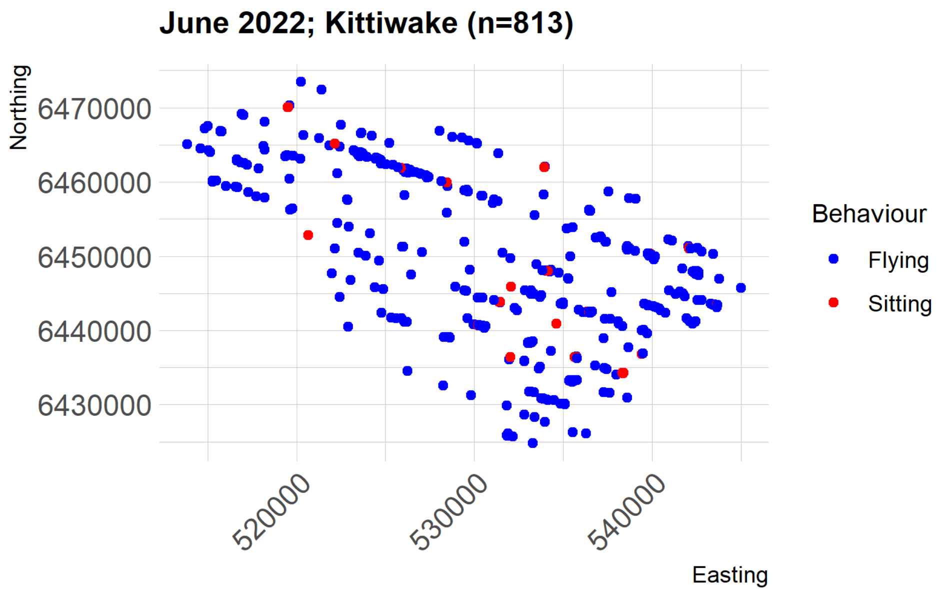

December 2021; Kittiwake (n=27)
Northing
6470000
6460000
6450000
6440000
6430000
520000
525000
530000
535000
Easting
Behaviour
Flying
Sitting

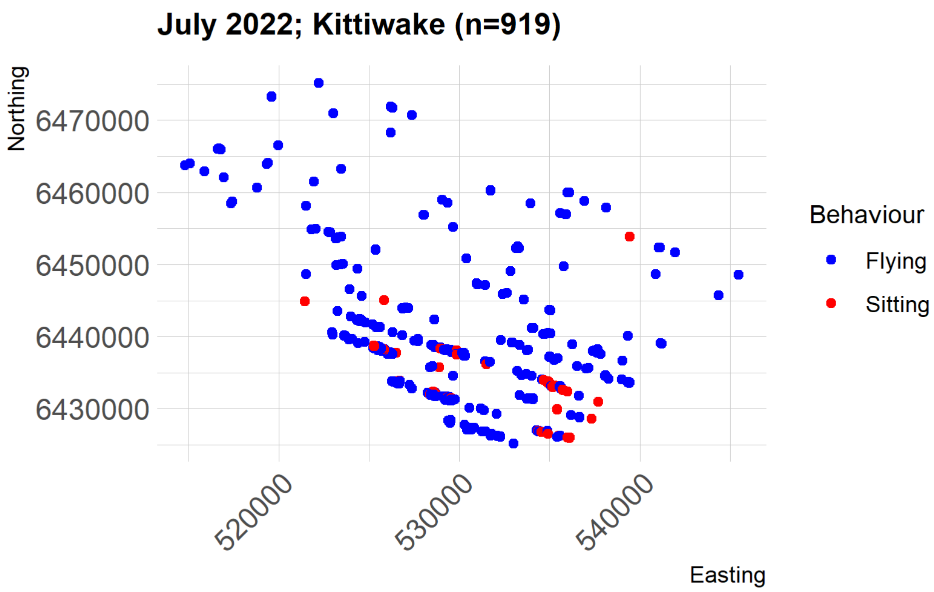

February 2022; Kittiwake (n=29)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

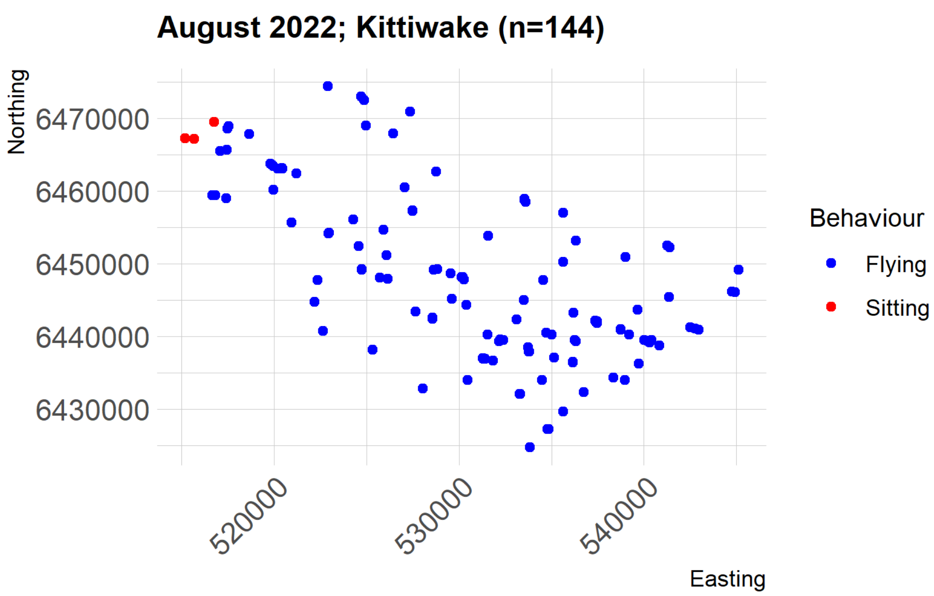

March 2022; Kittiwake (n=42)
Northing
6470000
6460000
6450000
6440000
6430000
515000
520000
525000
530000
535000
540000
Easting
Behaviour
Flying
Sitting

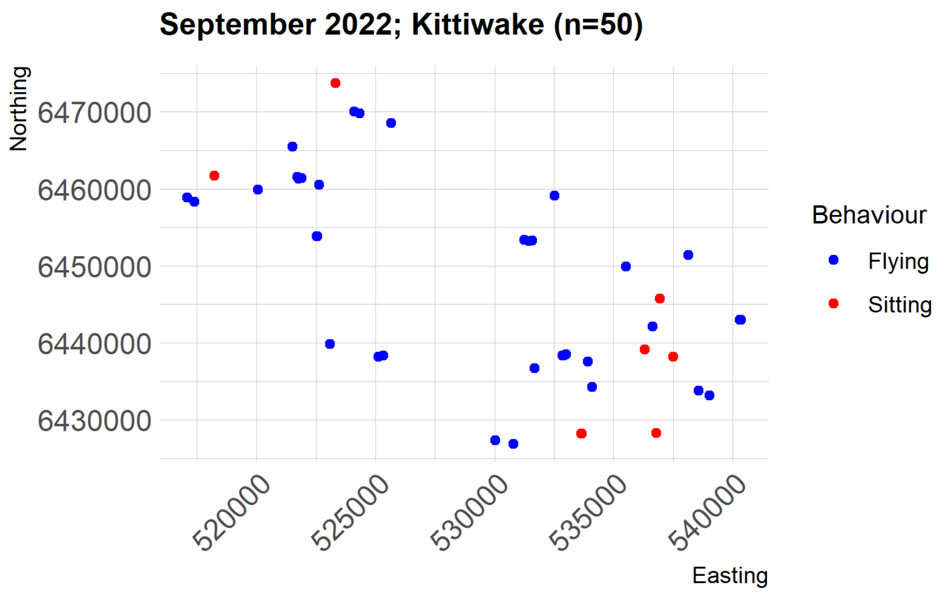

April 2022; Kittiwake (n=100)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

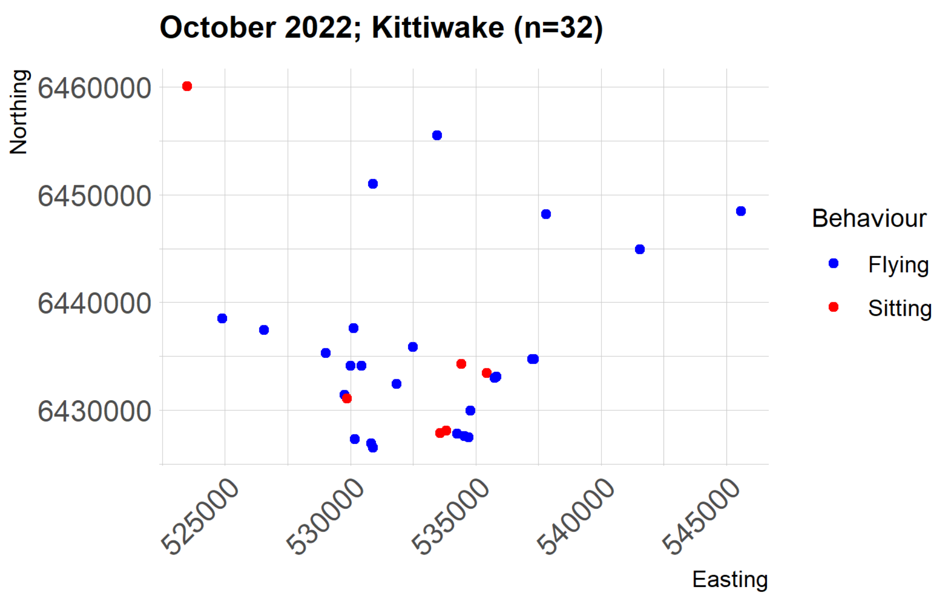

May 2022; Kittiwake (n=677)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

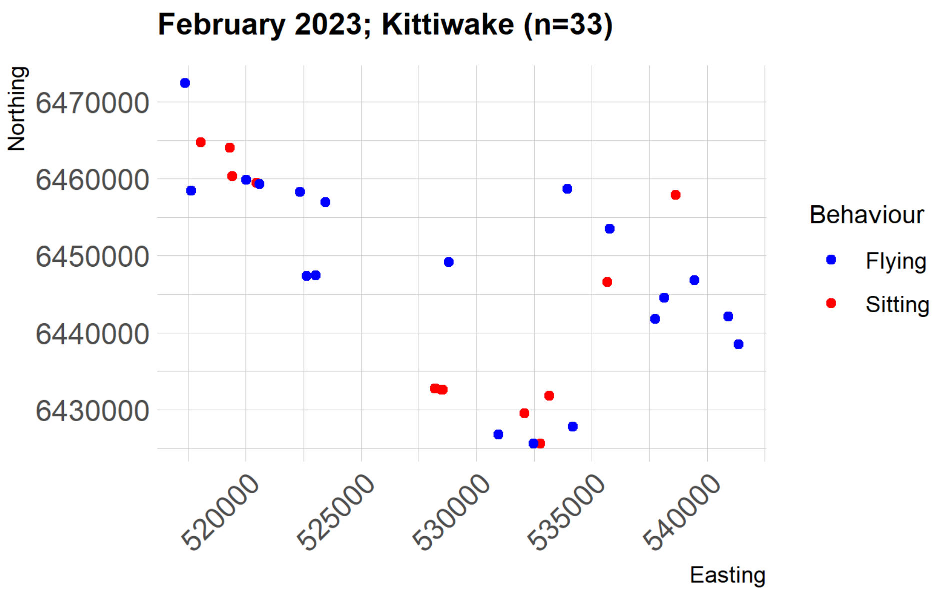

June 2022; Kittiwake (n=813)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

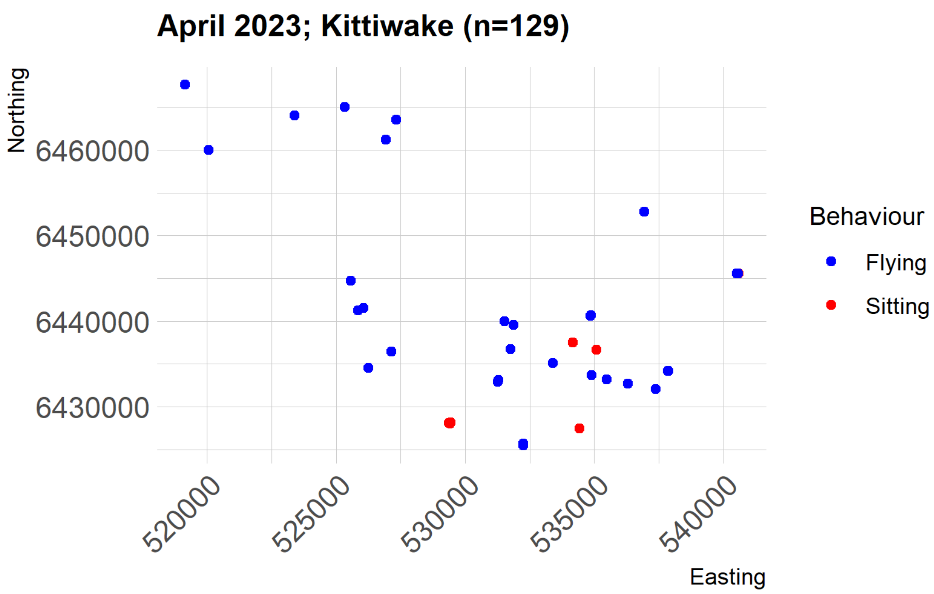

July 2022; Kittiwake (n=919)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

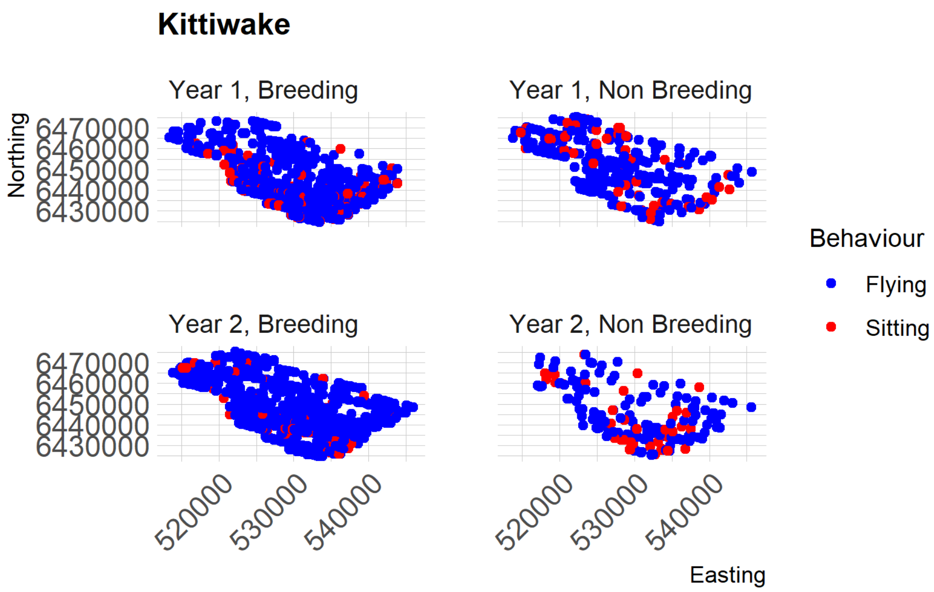

August 2022; Kittiwake (n=144)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting


September 2022; Kittiwake (n=50)
Northing
6470000
6460000
6450000
6440000
6430000
520000
525000
530000
535000
540000
Easting
Behaviour
Flying
Sitting


October 2022; Kittiwake (n=32)
Northing
6460000
6450000
6440000
6430000
525000
530000
535000
540000
545000
Easting
Behaviour
Flying
Sitting

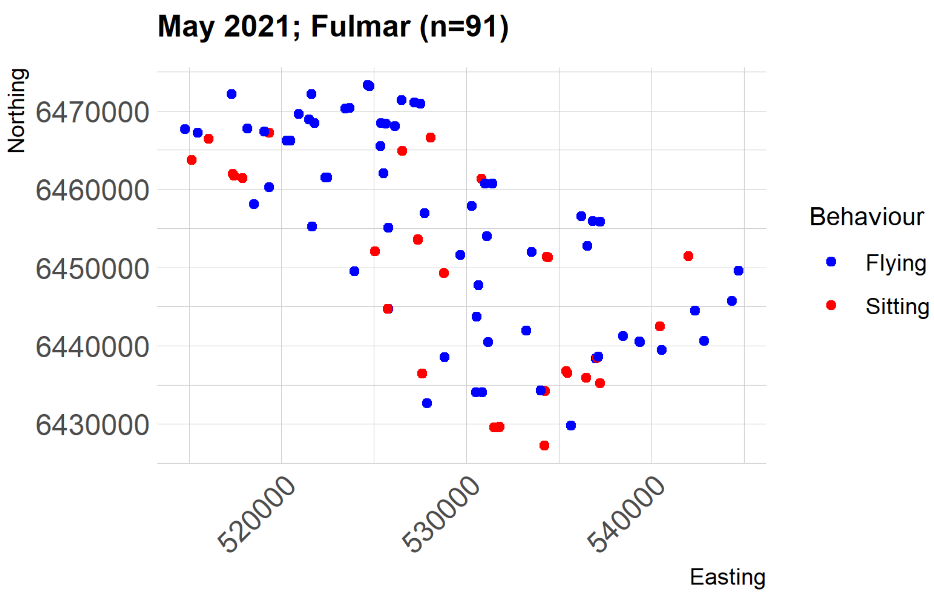

February 2023; Kittiwake (n=33)
Northing
6470000
6460000
6450000
6440000
6430000
520000
525000
530000
535000
540000
Easting
Behaviour
Flying
Sitting

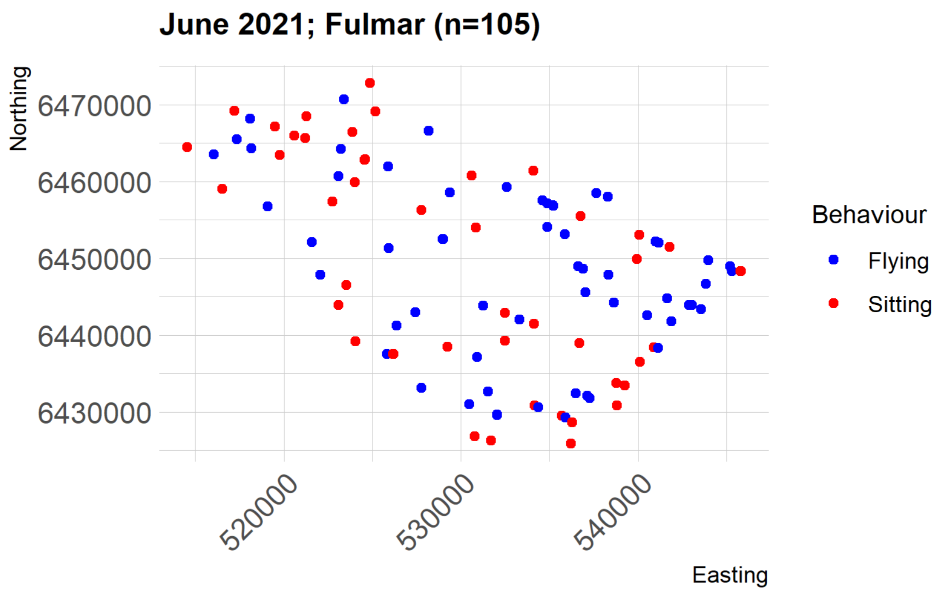

April 2023; Kittiwake (n=129)
Northing
6460000
6450000
6440000
6430000
520000
525000
530000
535000
540000
Easting
Behaviour
Flying
Sitting

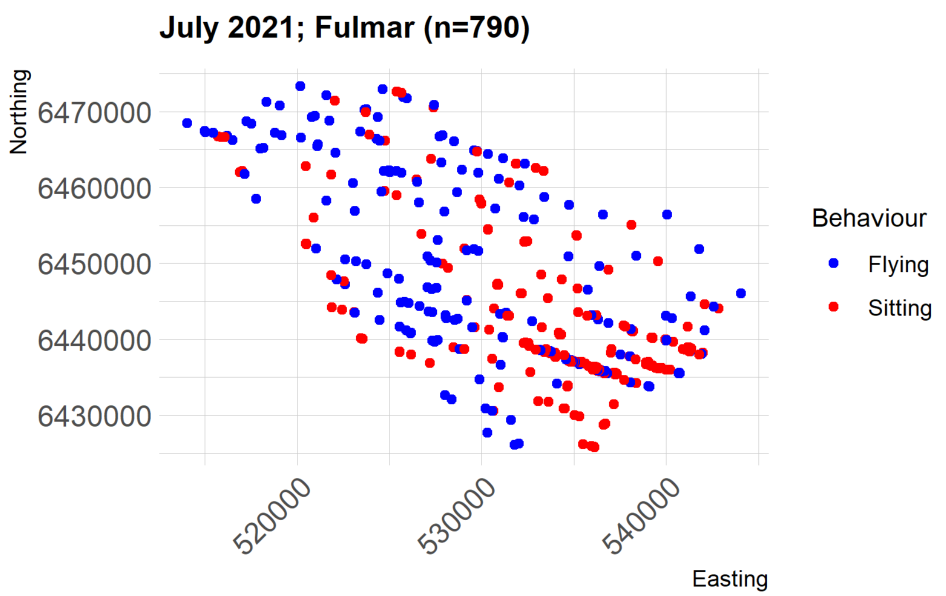

Kittiwake
Year 1, Breeding
Year 1, Non Breeding
Year 2, Breeding
Year 2, Non Breeding
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

I.3.4 Fulmar

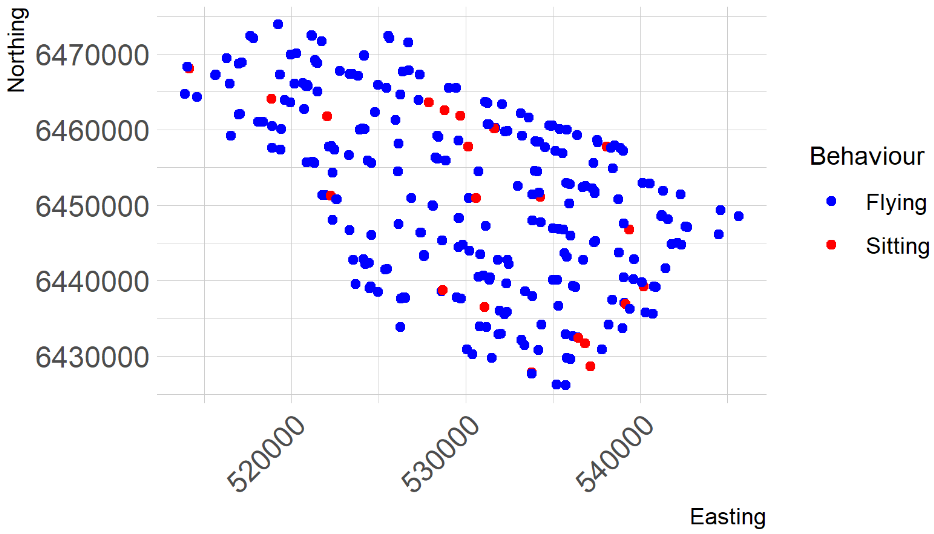
We suggest modelling fulmar (N=4611) by survey because the majority of surveys had sufficient observations (22/24). We have sufficient observations of fulmar to include behaviour in the model. We could also model by breeding season (April - 15 September). The September surveys from both years fall within the breeding season (14 Sep 2021, 11 Sep 2022).



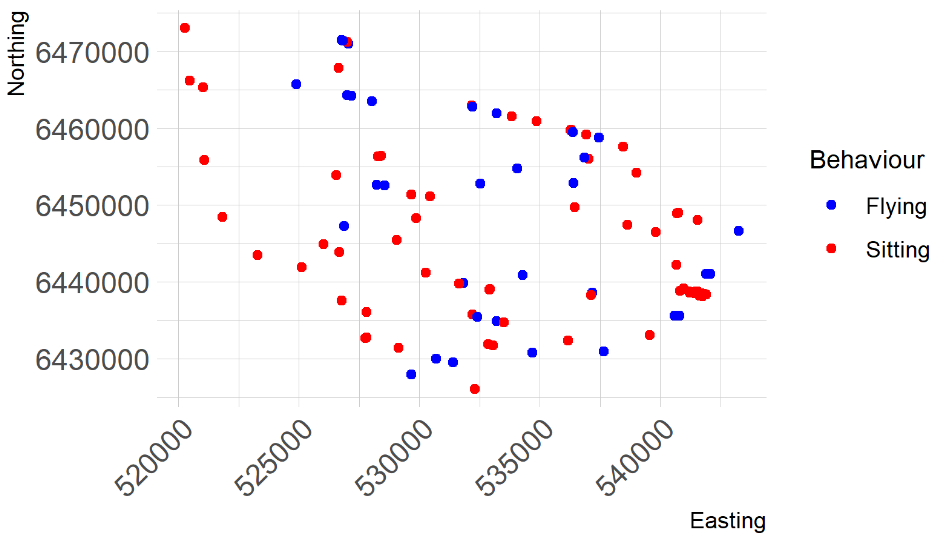

June 2021; Fulmar (n=105)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting


July 2021; Fulmar (n=790)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

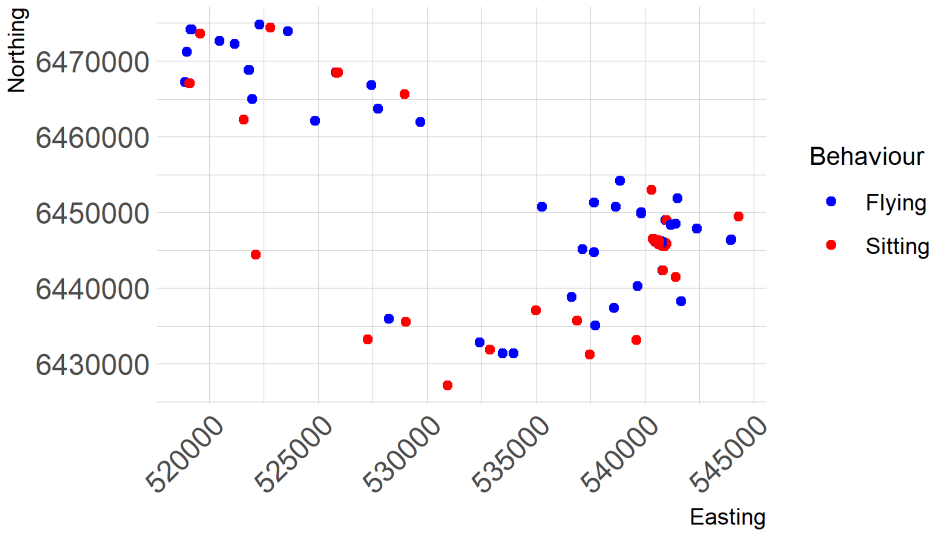
August 2021; Fulmar (n=285)



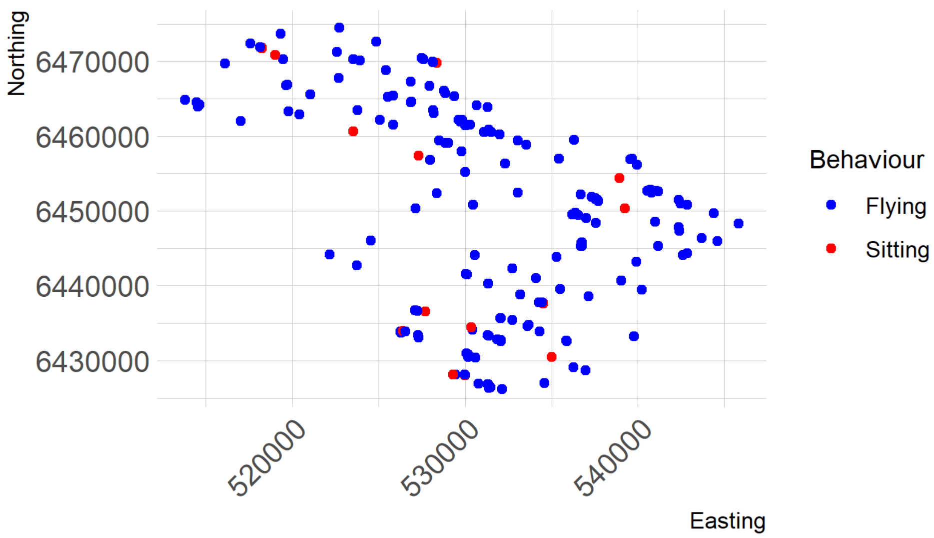
September 2021; Fulmar (n=113)

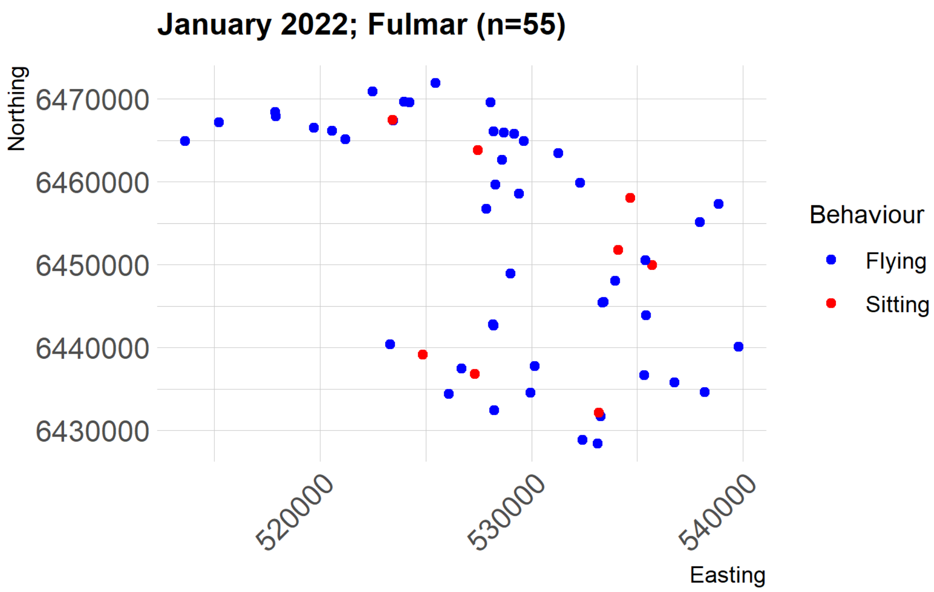


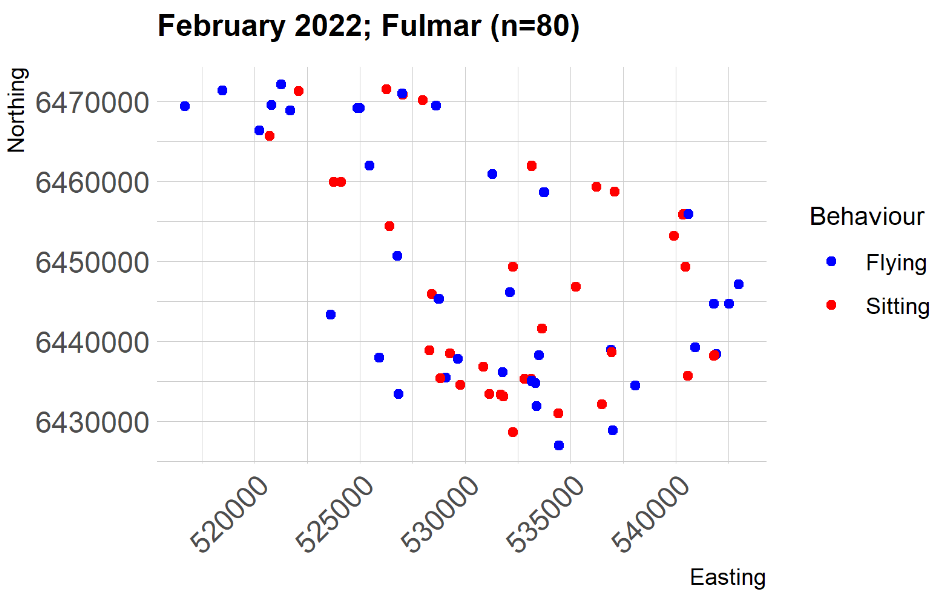
November 2021; Fulmar (n=180)

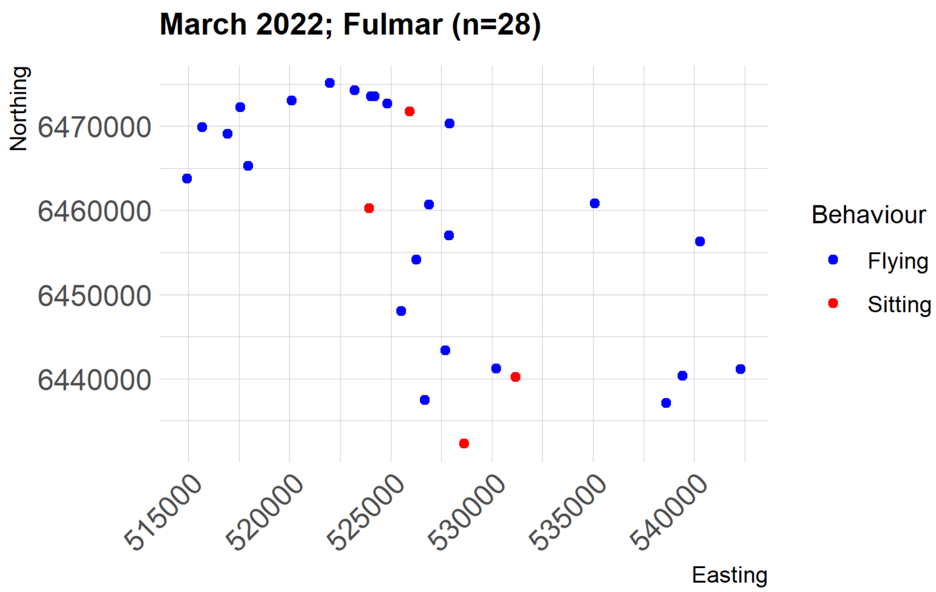


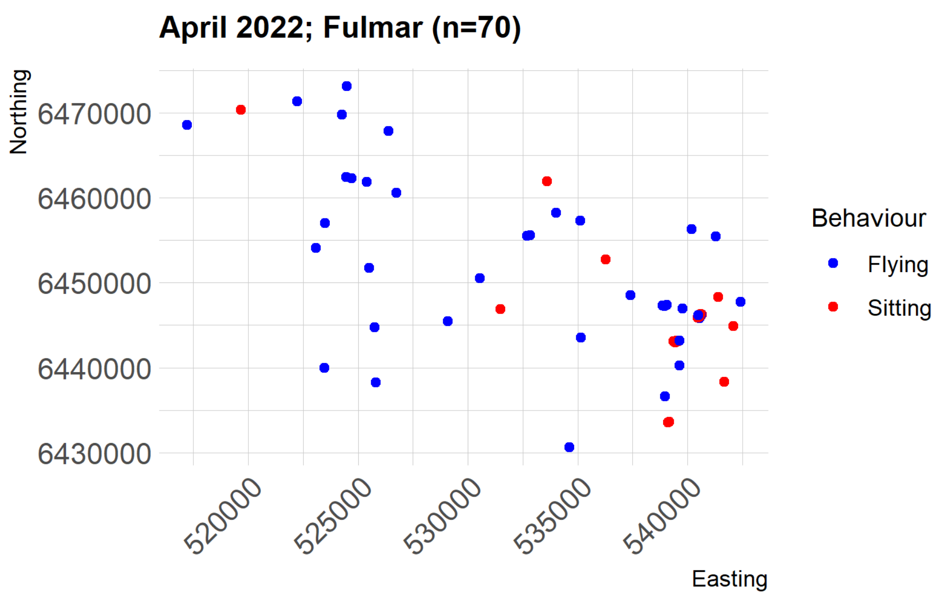
December 2021; Fulmar (n=187)

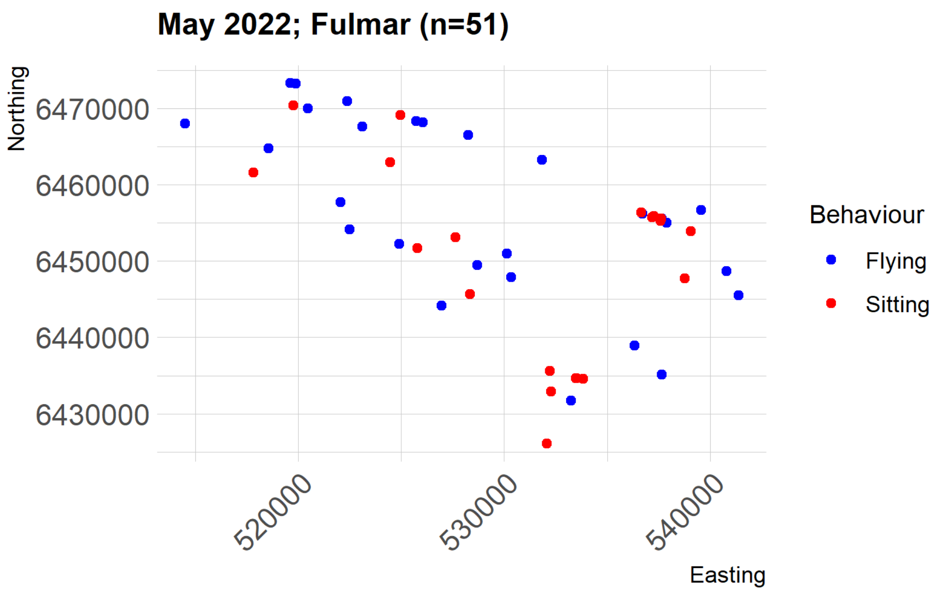


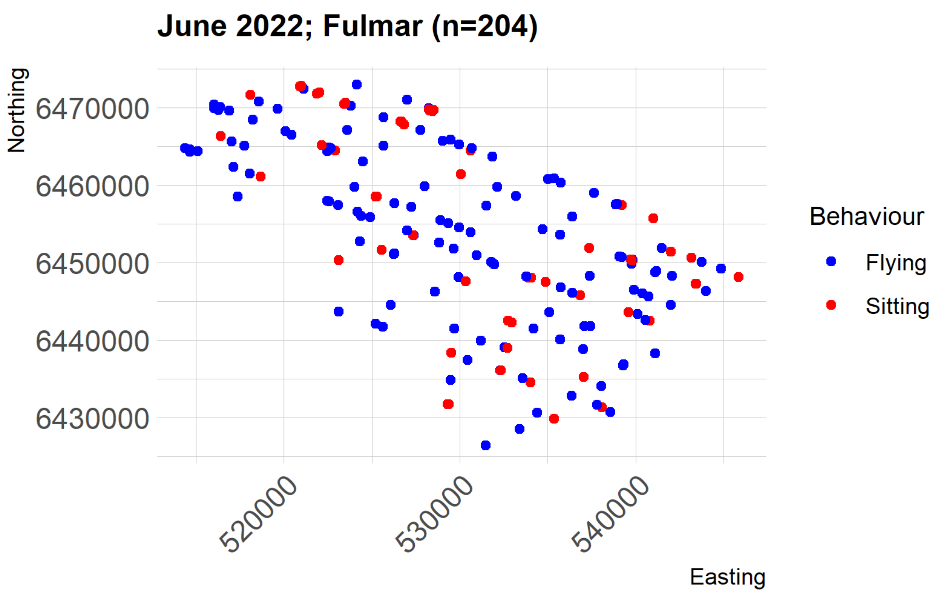

January 2022; Fulmar (n=55)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

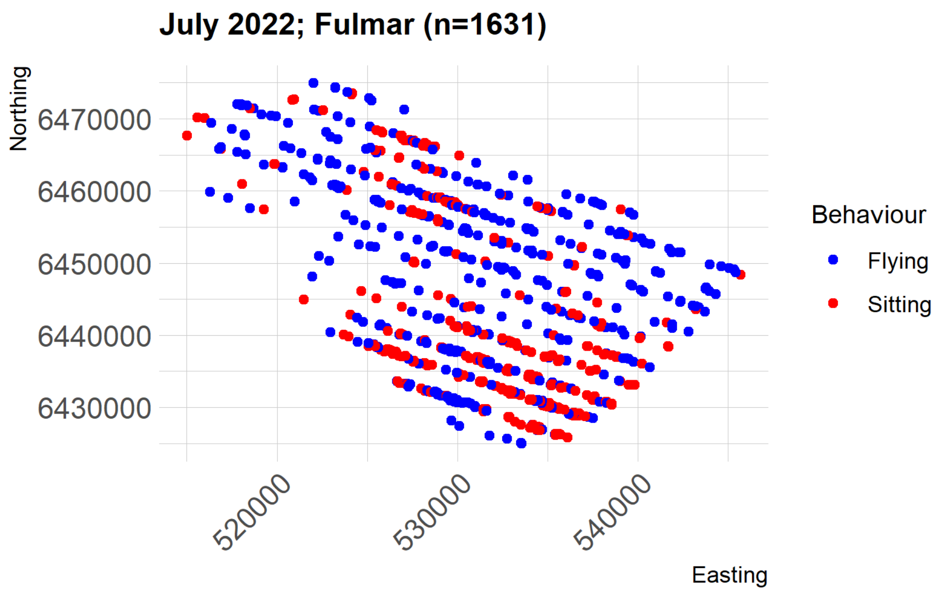

February 2022; Fulmar (n=80)
Northing
6470000
6460000
6450000
6440000
6430000
520000
525000
530000
535000
540000
Easting
Behaviour
Flying
Sitting

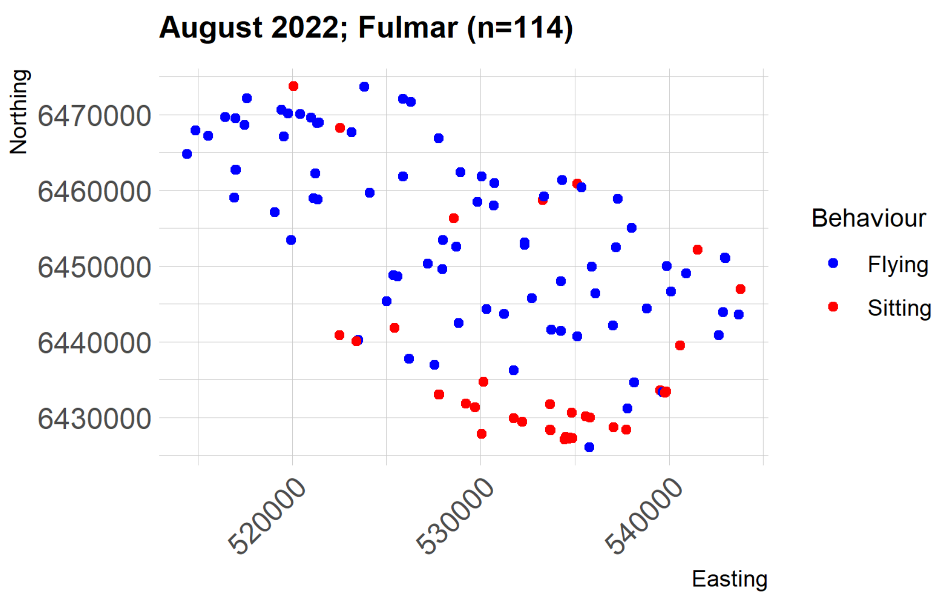

March 2022; Fulmar (n=28)
Northing
6470000
6460000
6450000
6440000
515000
520000
525000
530000
535000
540000
Easting
Behaviour
Flying
Sitting

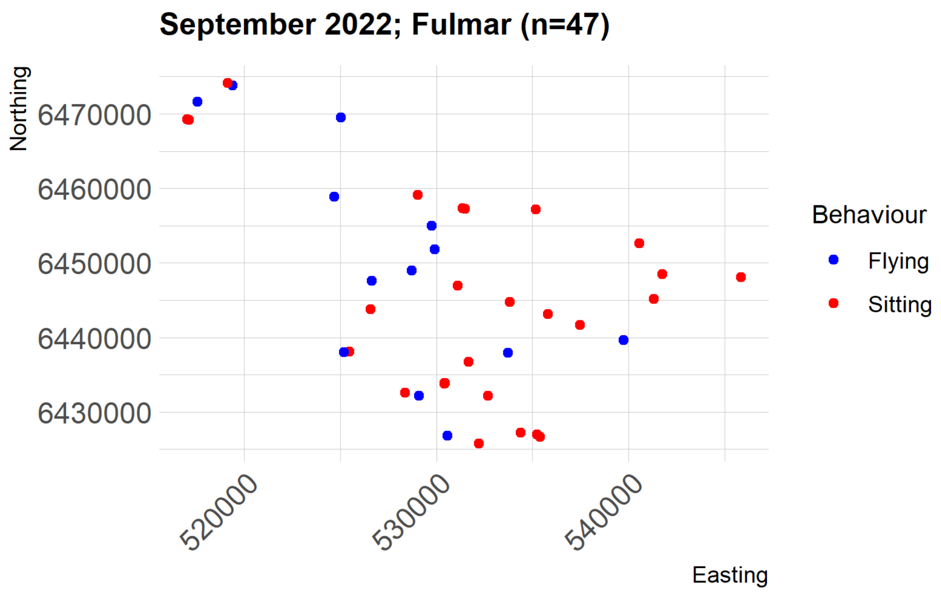

April 2022; Fulmar (n=70)
Northing
6470000
6460000
6450000
6440000
6430000
520000
525000
530000
535000
540000
Easting
Behaviour
Flying
Sitting

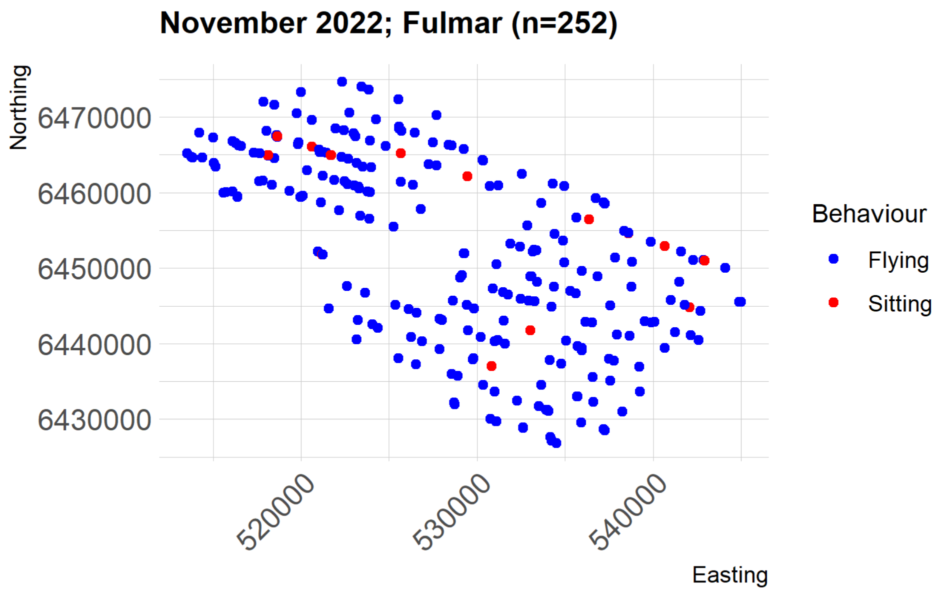

May 2022; Fulmar (n=51)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

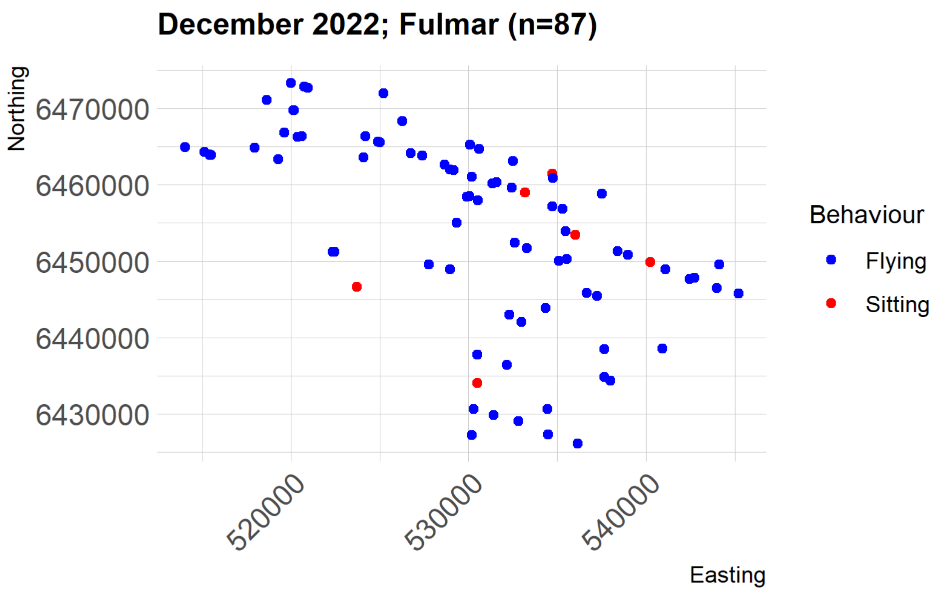

June 2022; Fulmar (n=204)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

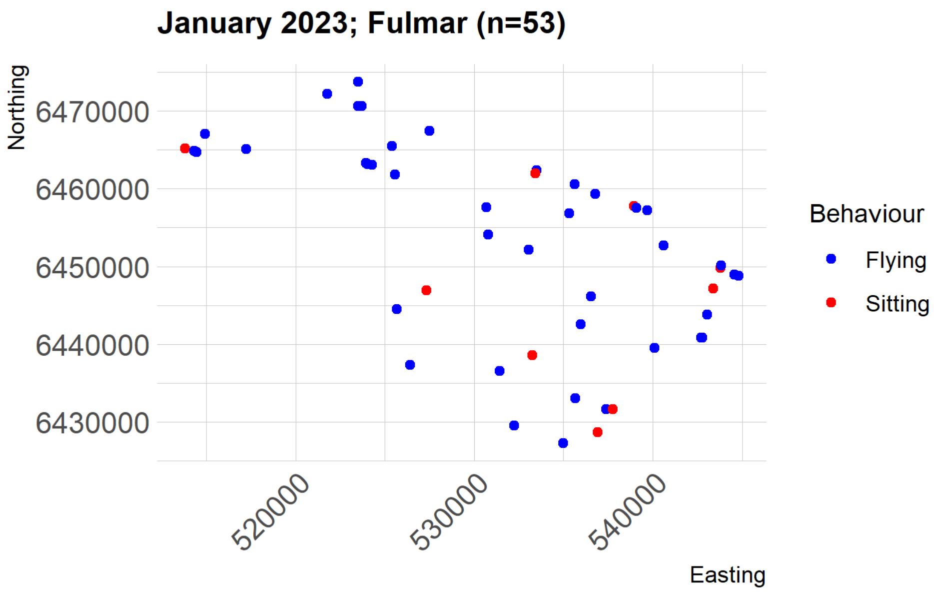

July 2022; Fulmar (n=1631)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

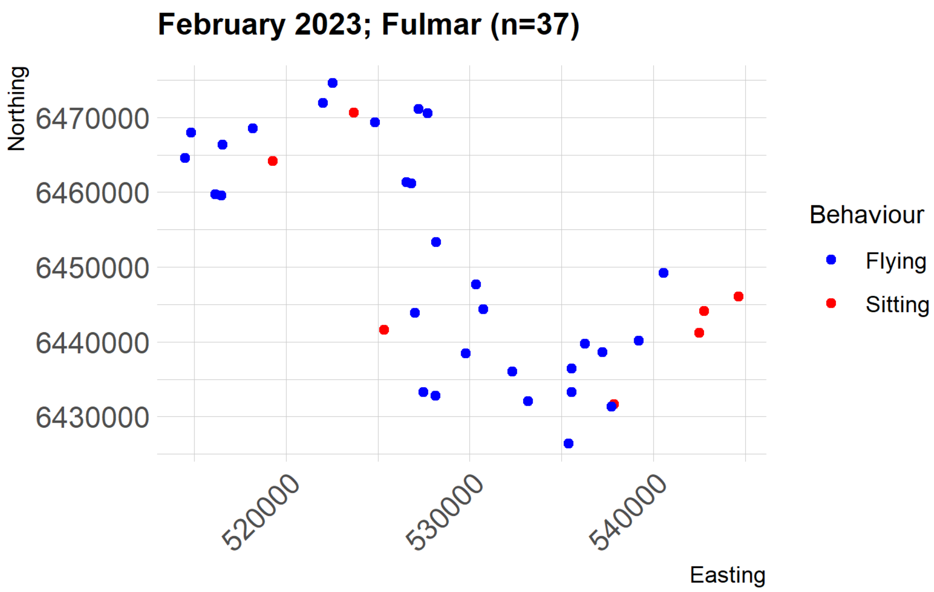

August 2022; Fulmar (n=114)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

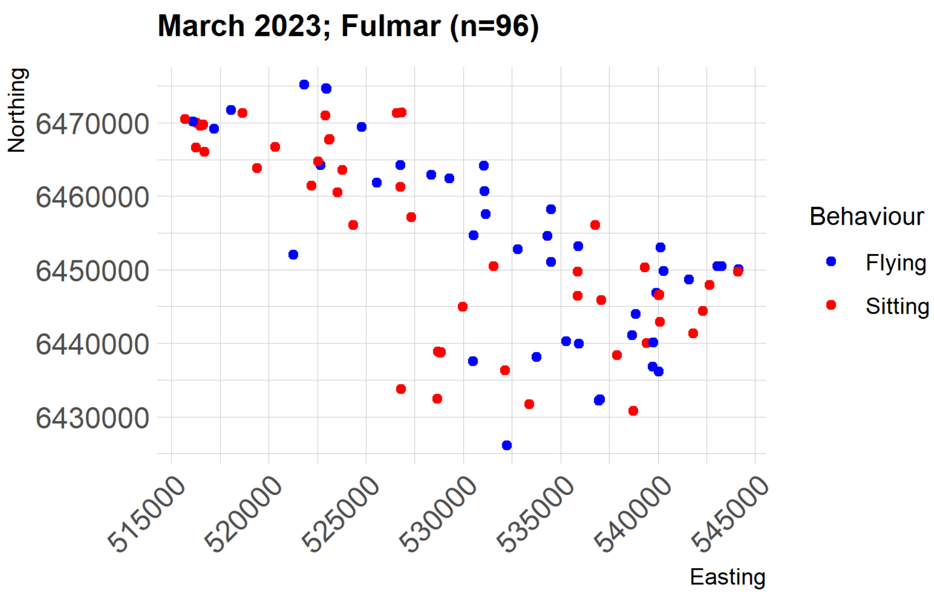

September 2022; Fulmar (n=47)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

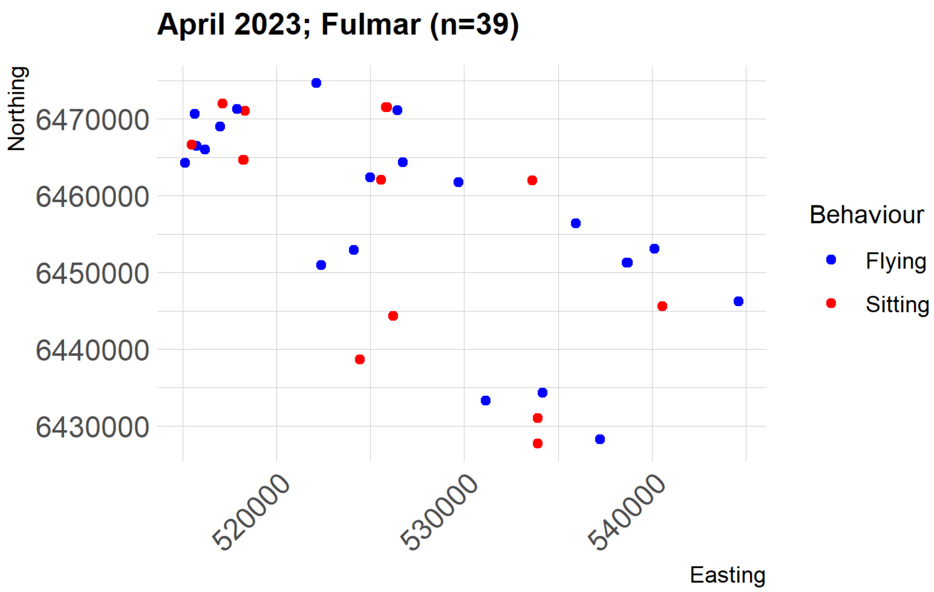

November 2022; Fulmar (n=252)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

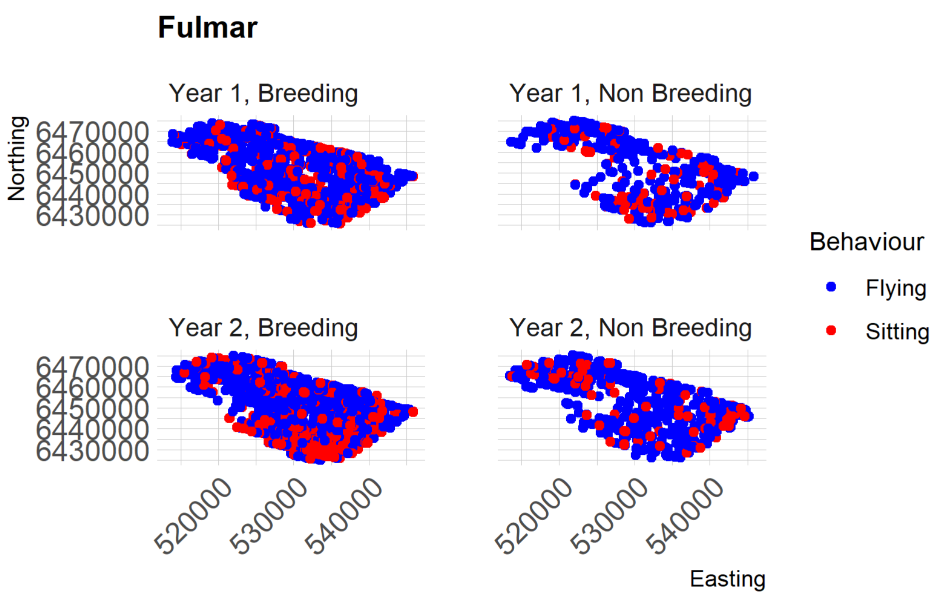

December 2022; Fulmar (n=87)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting


January 2023; Fulmar (n=53)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting


February 2023; Fulmar (n=37)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

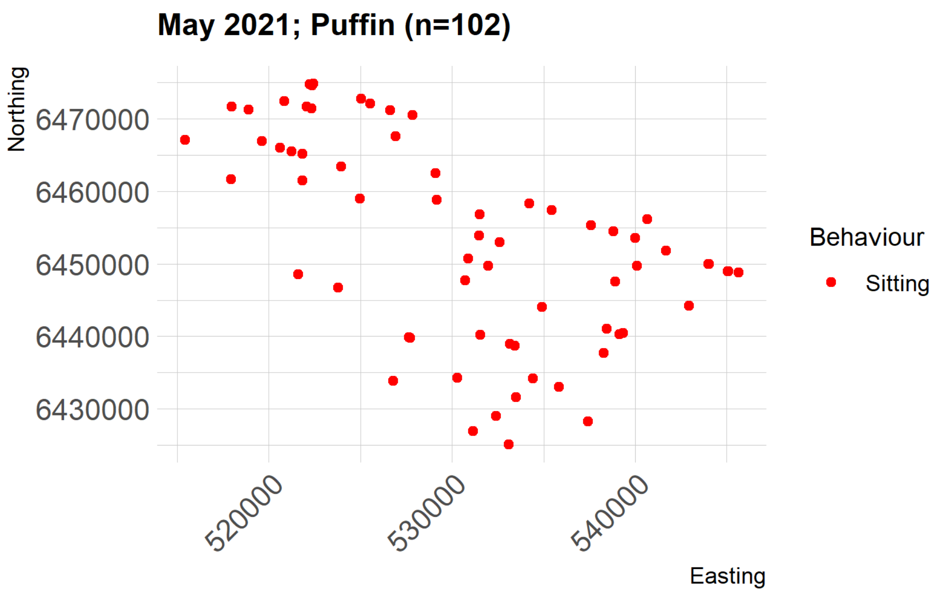

March 2023; Fulmar (n=96)
Northing
6470000
6460000
6450000
6440000
6430000
515000
520000
525000
530000
535000
540000
545000
Easting
Behaviour
Flying
Sitting

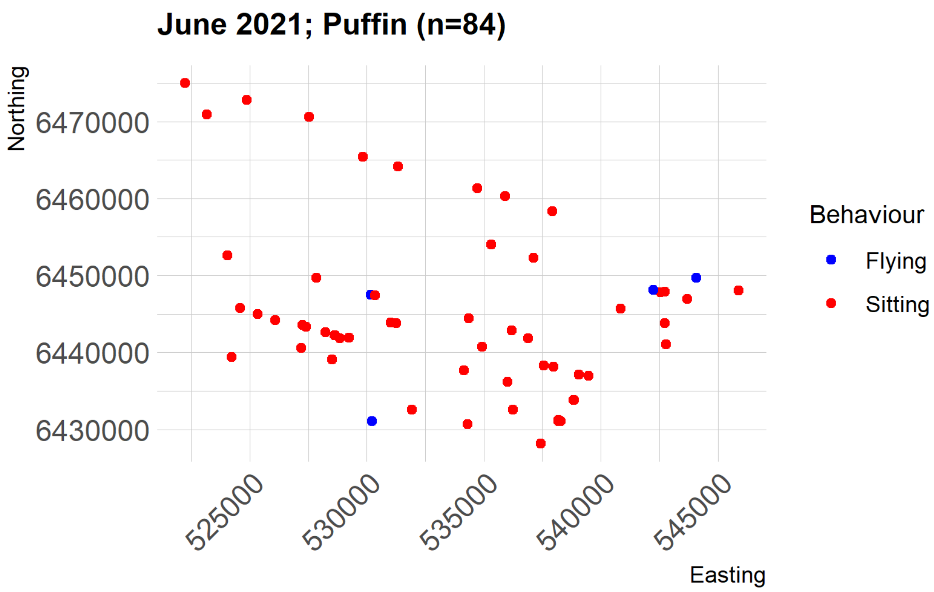

April 2023; Fulmar (n=39)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

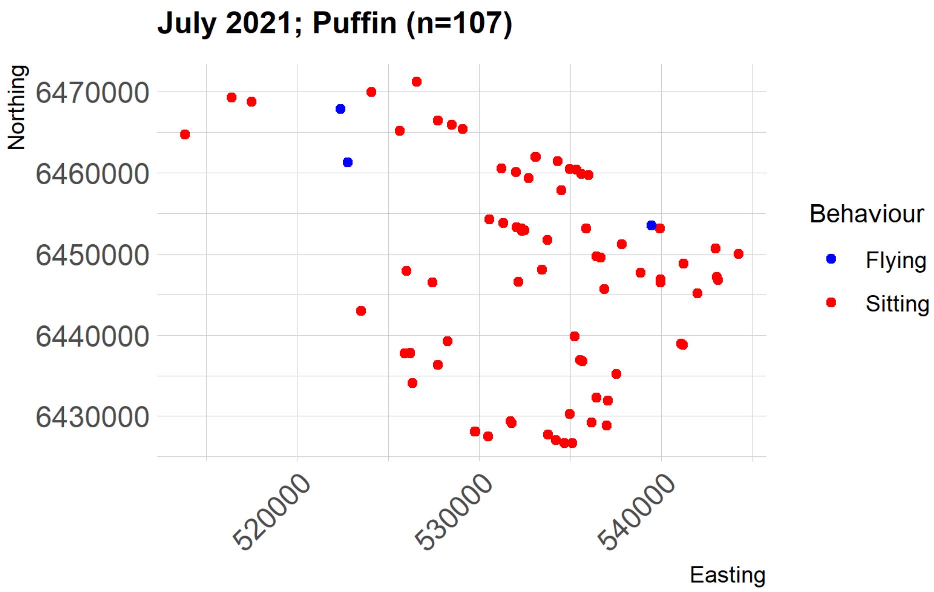

Fulmar
Year 1, Breeding
Year 1, Non Breeding
Year 2, Breeding
Year 2, Non Breeding
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

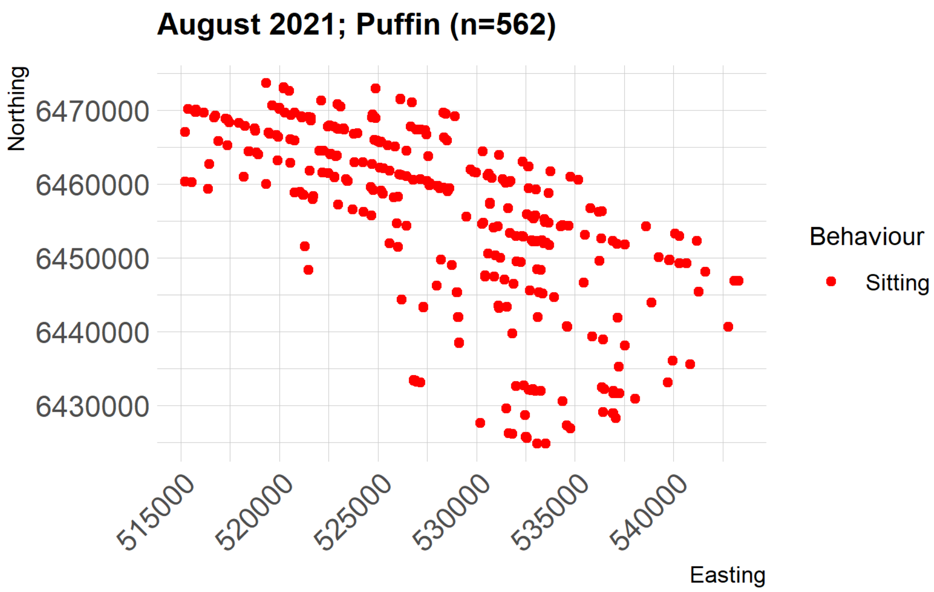
I.3.5 Puffin

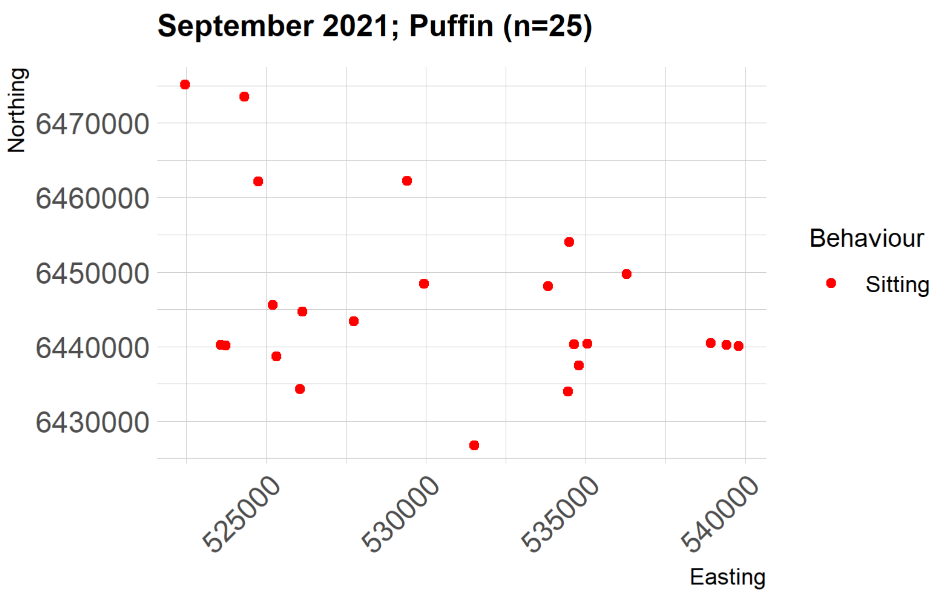
We suggest modelling puffin (N=2054) by survey for the surveys that had sufficient observations (13/24). All but three surveys with enough observations occur during the breeding season, so we could also model puffin by breeding season (April - 15 August). The August surveys from both years fall within the breeding season (01 Aug 2021, 10 Aug 2022). We suggest modelling all behaviours (flying & sitting) together as puffin do not fly high enough to be at risk for collision with turbines.

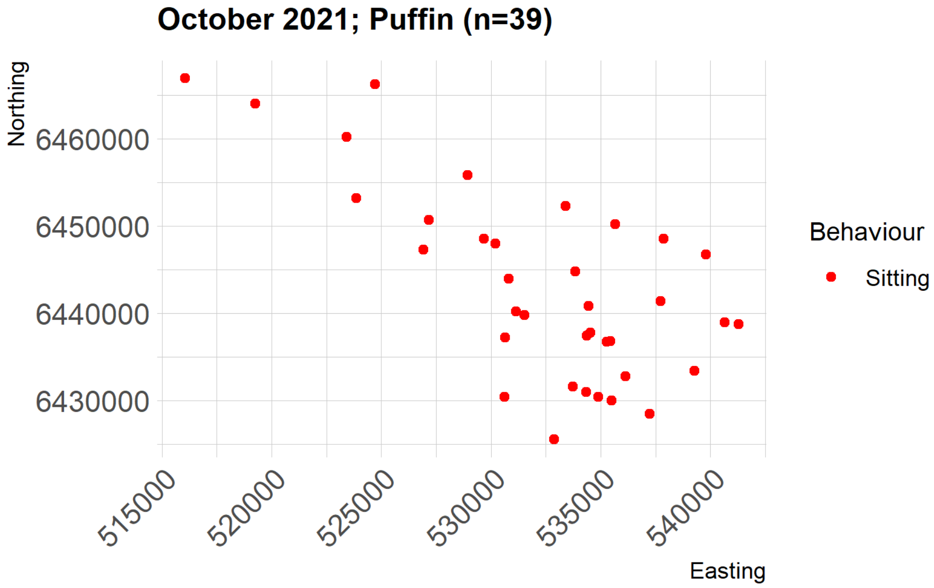


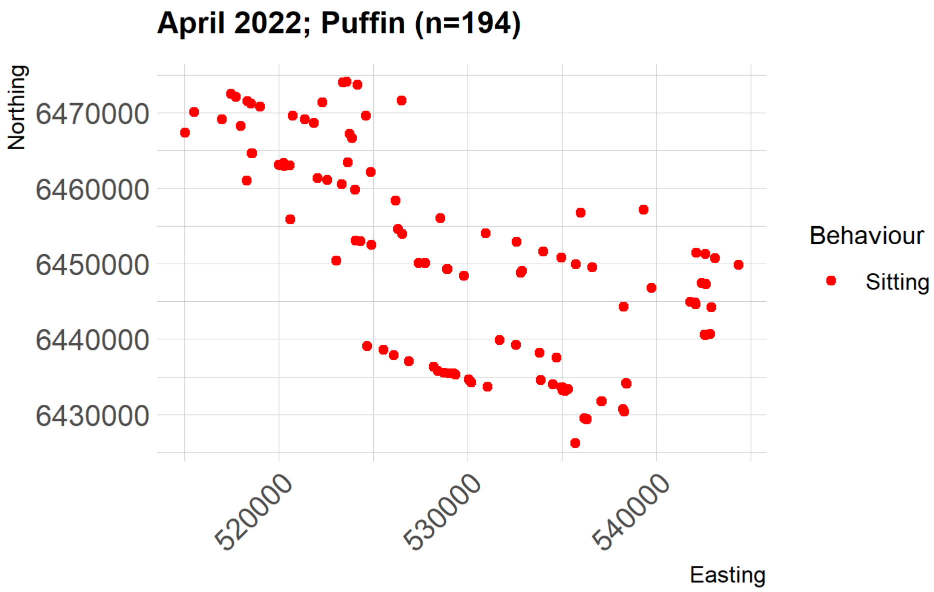

June 2021; Puffin (n=84)
Northing
6470000
6460000
6450000
6440000
6430000
525000
530000
535000
540000
545000
Easting
Behaviour
Flying
Sitting

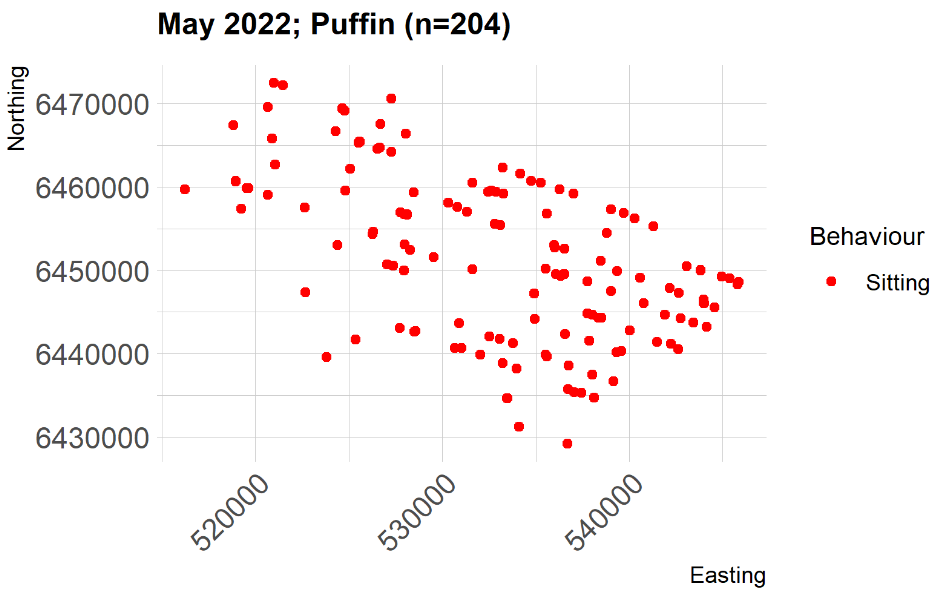

July 2021; Puffin (n=107)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

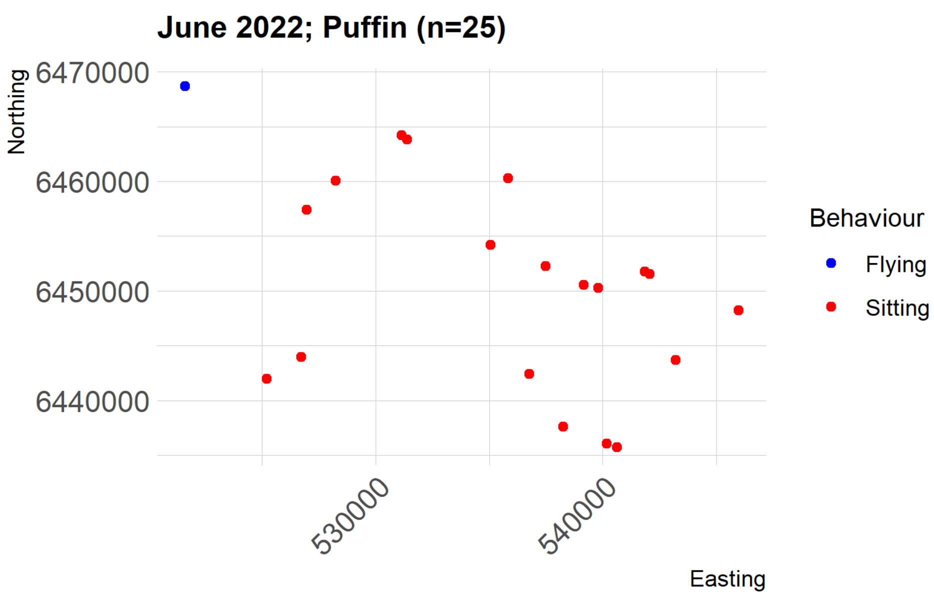

August 2021; Puffin (n=562)
Northing
6470000
6460000
6450000
6440000
6430000
515000
520000
525000
530000
535000
540000
Easting
Behaviour
Sitting

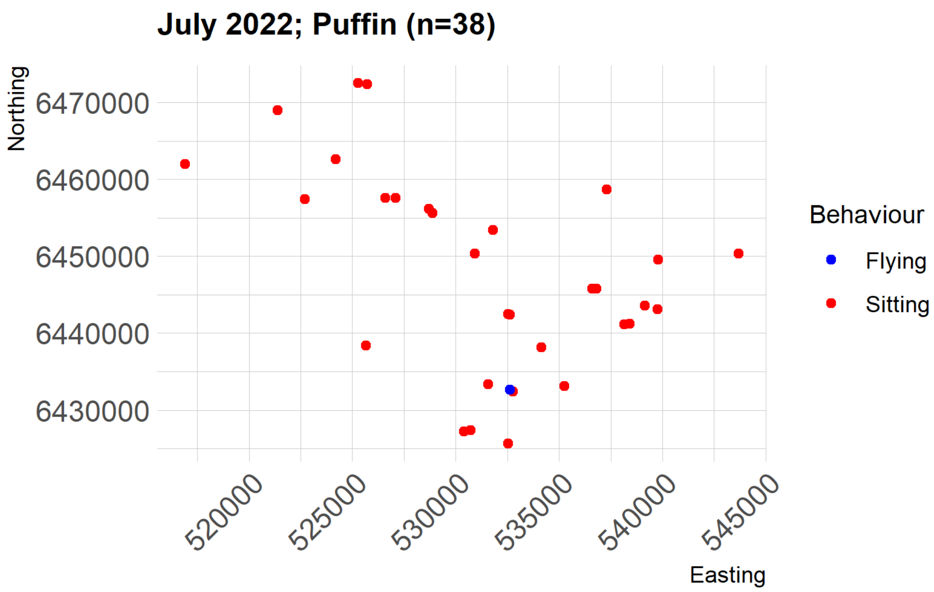

September 2021; Puffin (n=25)
Northing
6470000
6460000
6450000
6440000
6430000
525000
530000
535000
540000
Easting
Behaviour
Sitting

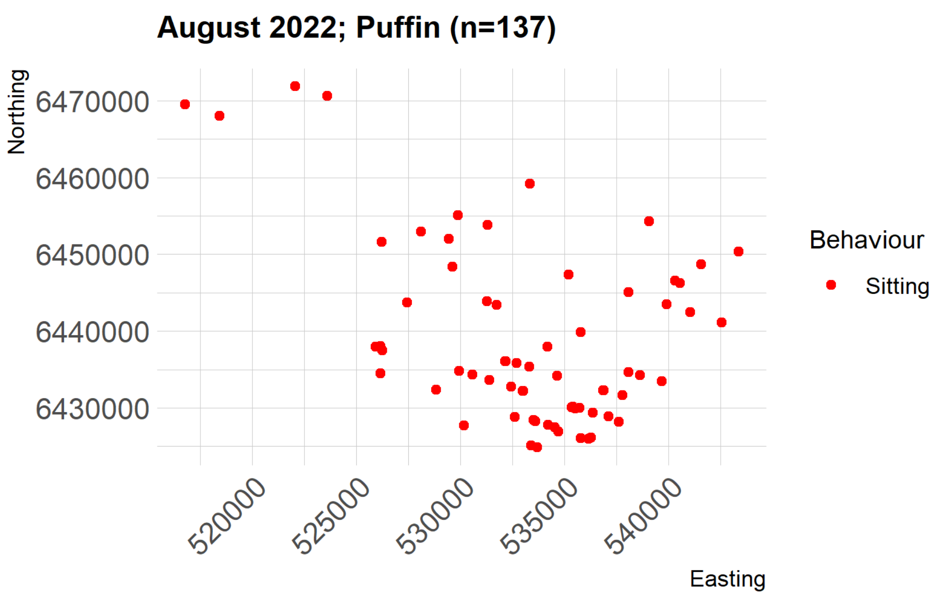

October 2021; Puffin (n=39)
Northing
6460000
6450000
6440000
6430000
515000
520000
525000
530000
535000
540000
Easting
Behaviour
Sitting

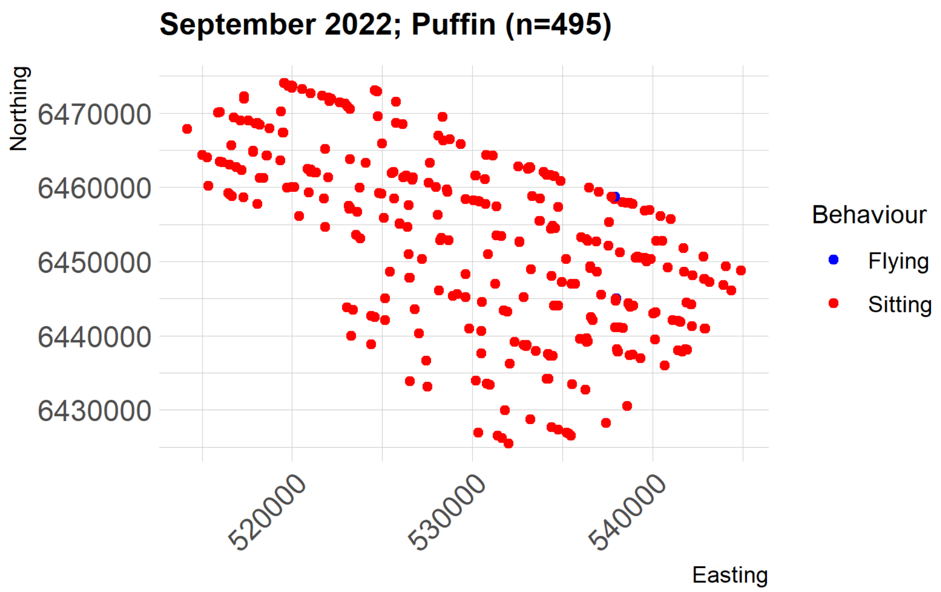

April 2022; Puffin (n=194)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Sitting

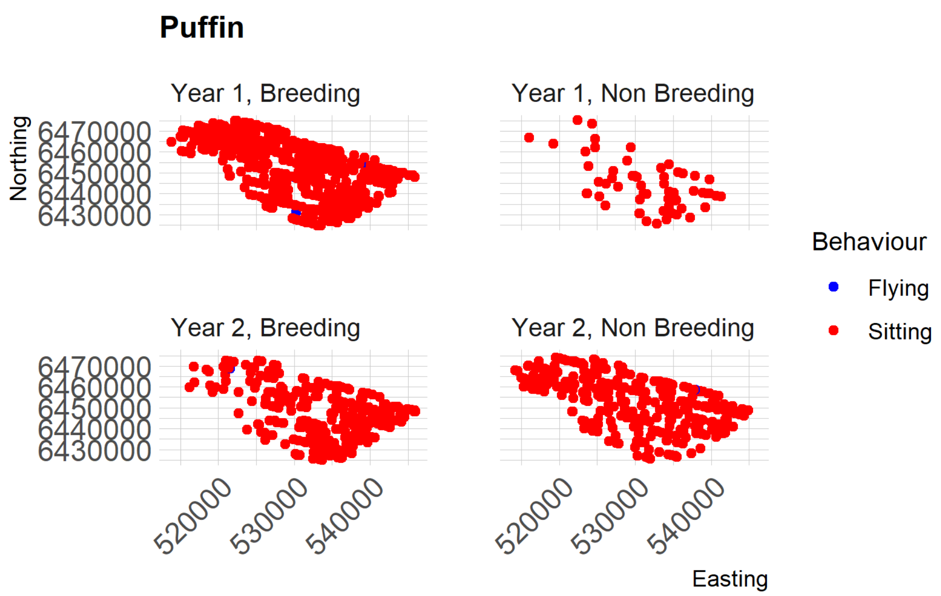

May 2022; Puffin (n=204)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Sitting


June 2022; Puffin (n=25)
Northing
6470000
6460000
6450000
6440000
530000
540000
Easting
Behaviour
Flying
Sitting


July 2022; Puffin (n=38)
Northing
6470000
6460000
6450000
6440000
6430000
520000
525000
530000
535000
540000
545000
Easting
Behaviour
Flying
Sitting

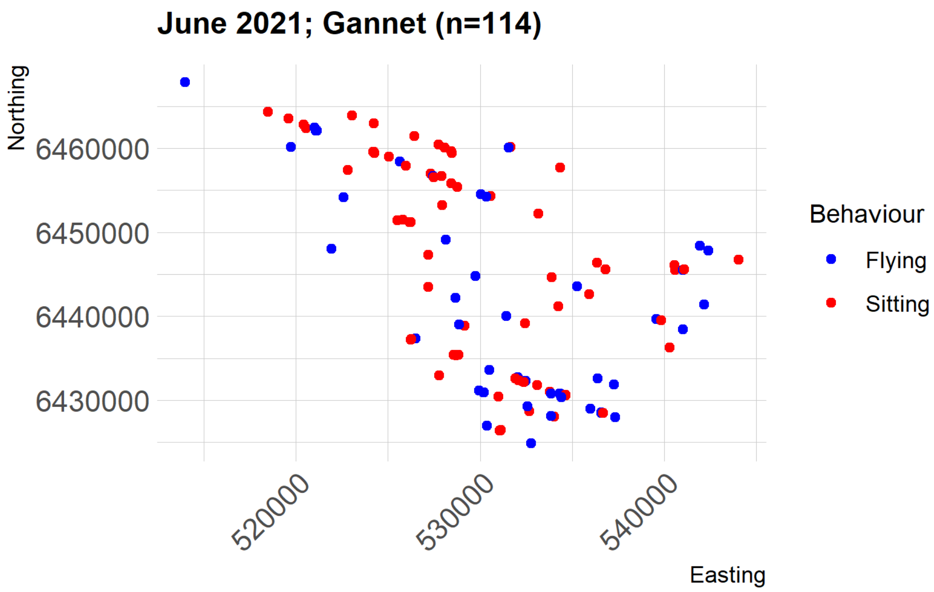

August 2022; Puffin (n=137)
Northing
6470000
6460000
6450000
6440000
6430000
520000
525000
530000
535000
540000
Easting
Behaviour
Sitting

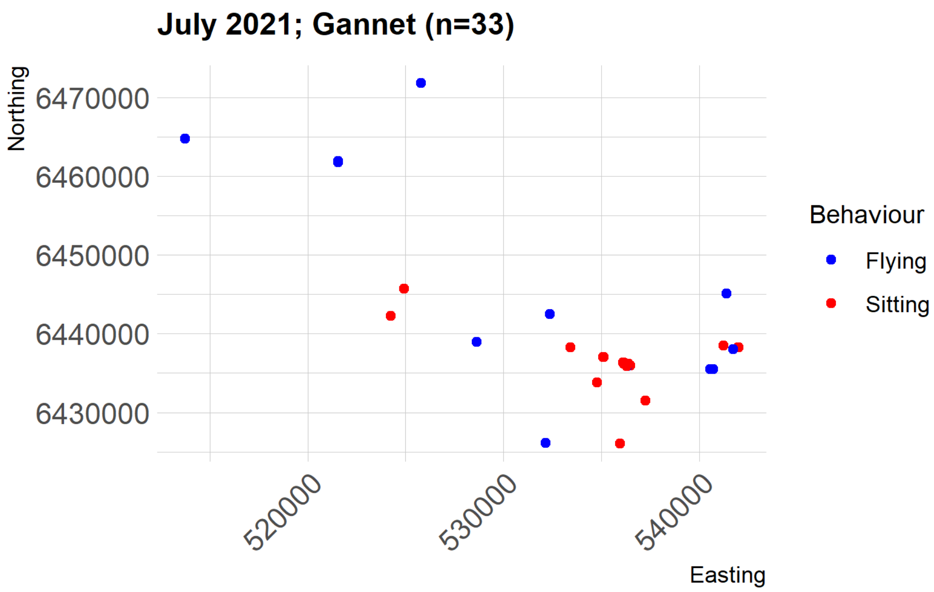

September 2022; Puffin (n=495)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

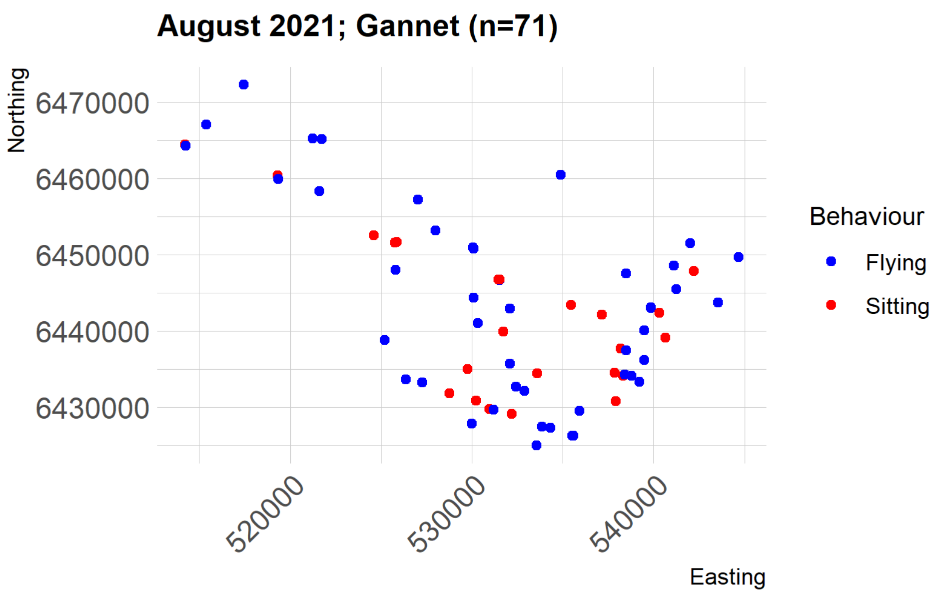

Puffin
Year 1, Breeding
Year 1, Non Breeding
Year 2, Breeding
Year 2, Non Breeding
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

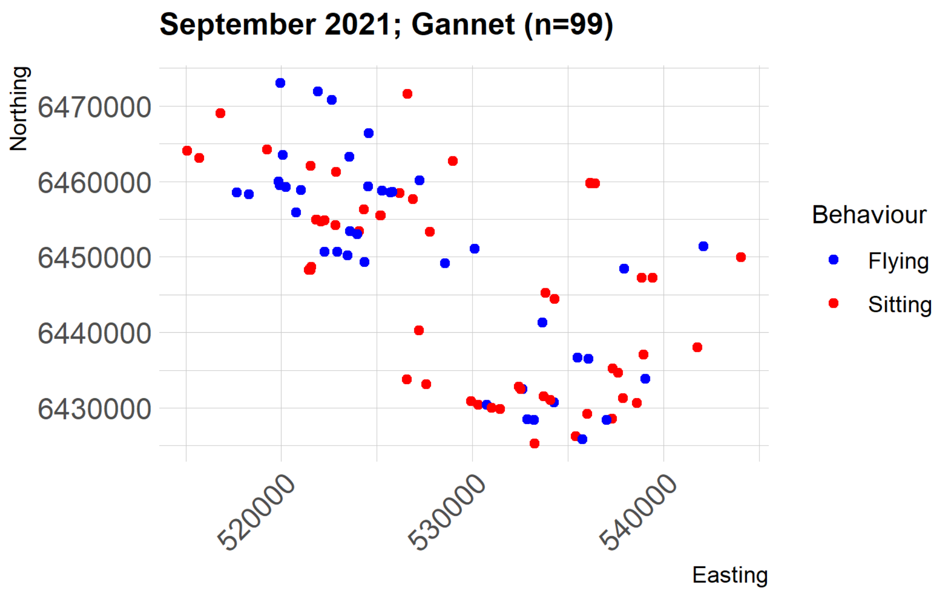
I.3.6 Gannet

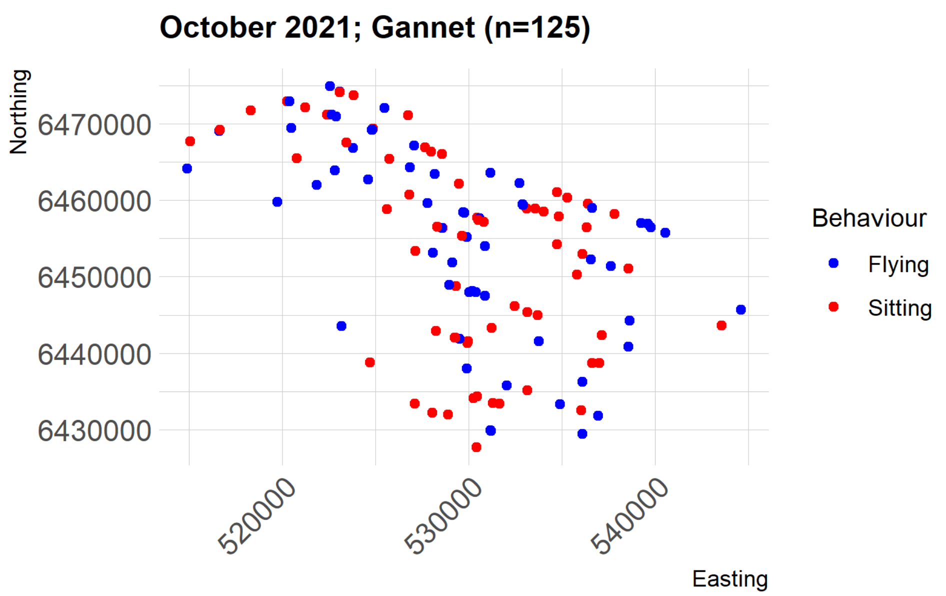
We suggest modelling gannet (N=1042) by survey for the surveys that had sufficient observations (10/24). All but two surveys with enough observations occur during the breeding season, so we could also model gannet by breeding season (15 March - September). Neither of the March surveys (02 Mar 2022, 10 Mar 2023) occured within the the breeding season, and are therefore considered "Non-Breeding." For months with enough observations of flying birds we will also suggest including behaviour in the model.

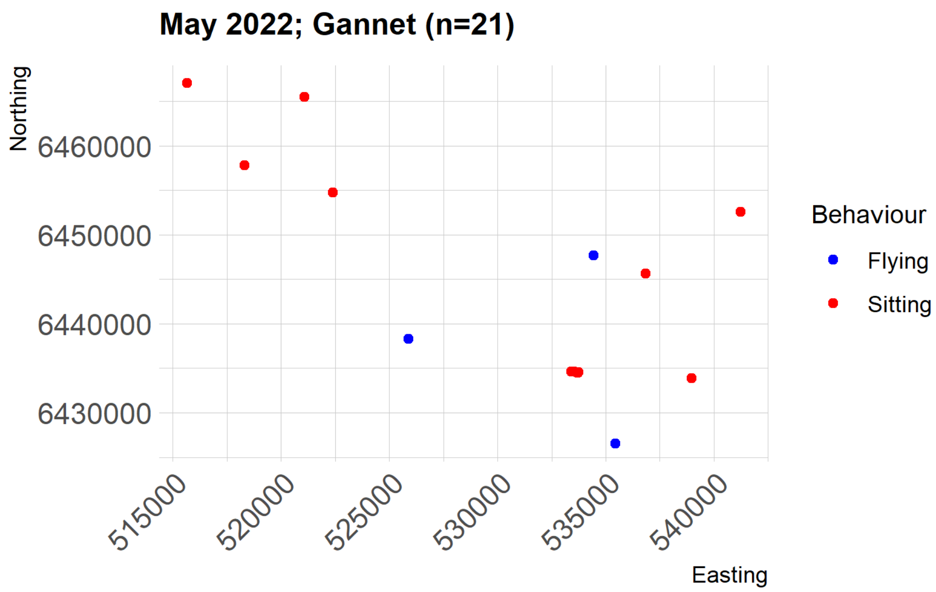


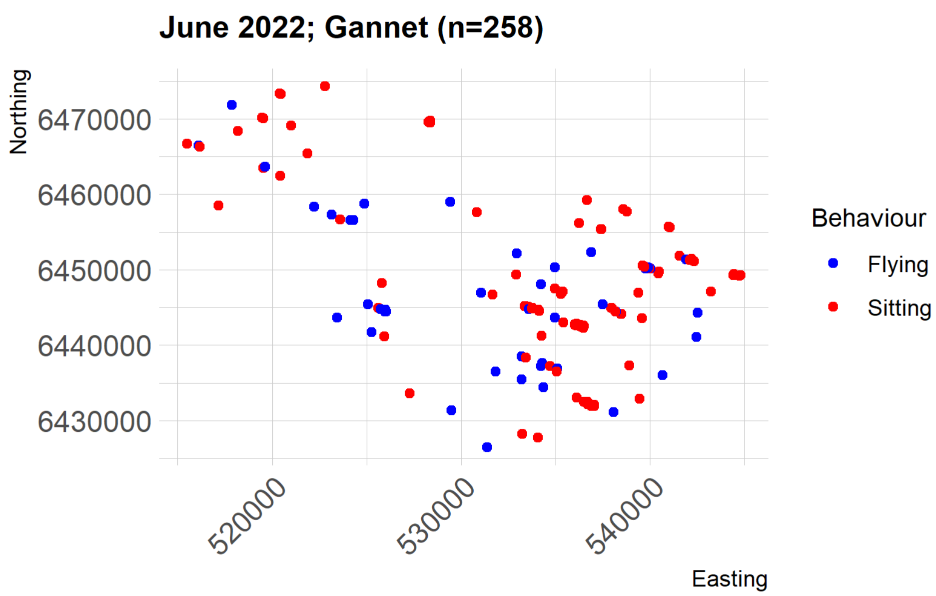

July 2021; Gannet (n=33)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

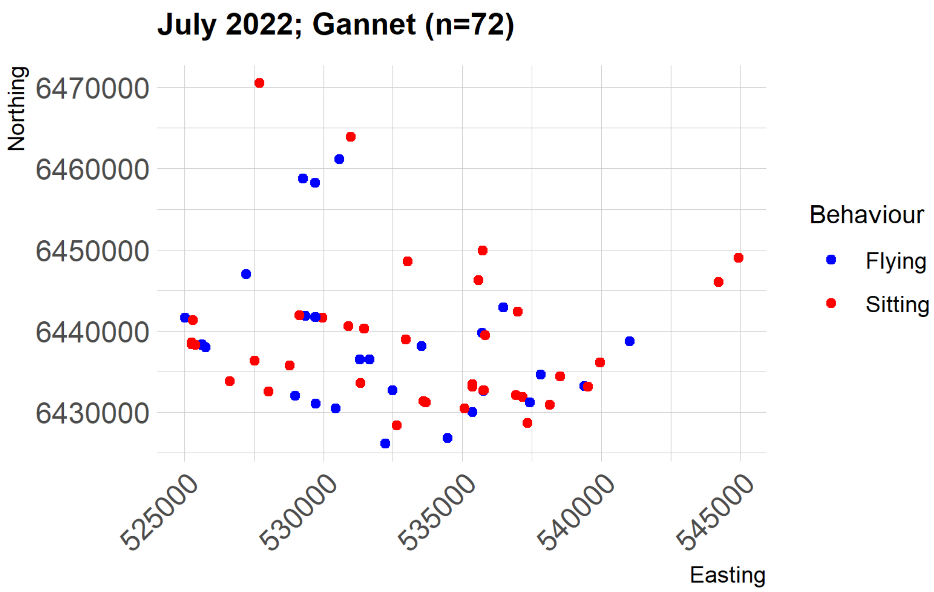

August 2021; Gannet (n=71)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

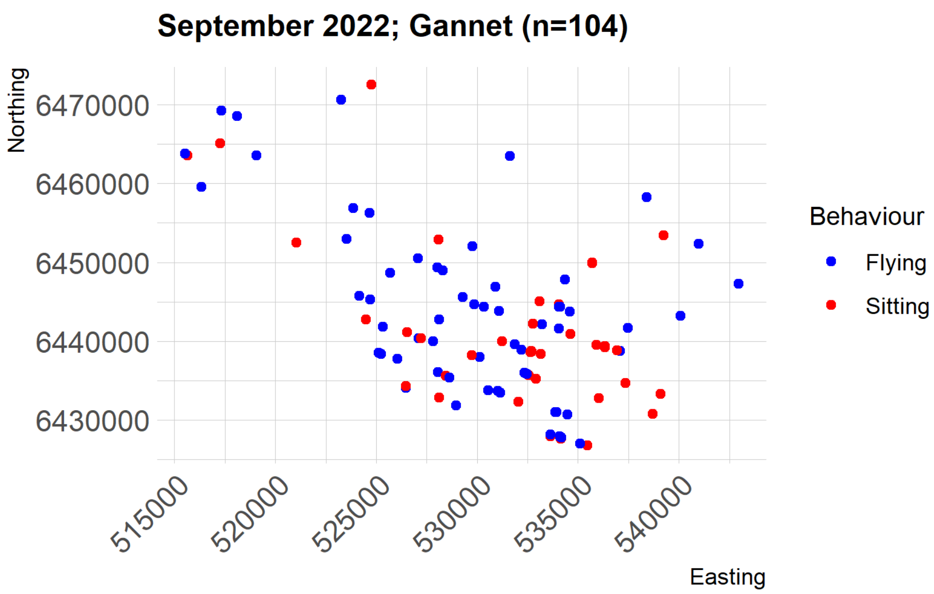

September 2021; Gannet (n=99)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

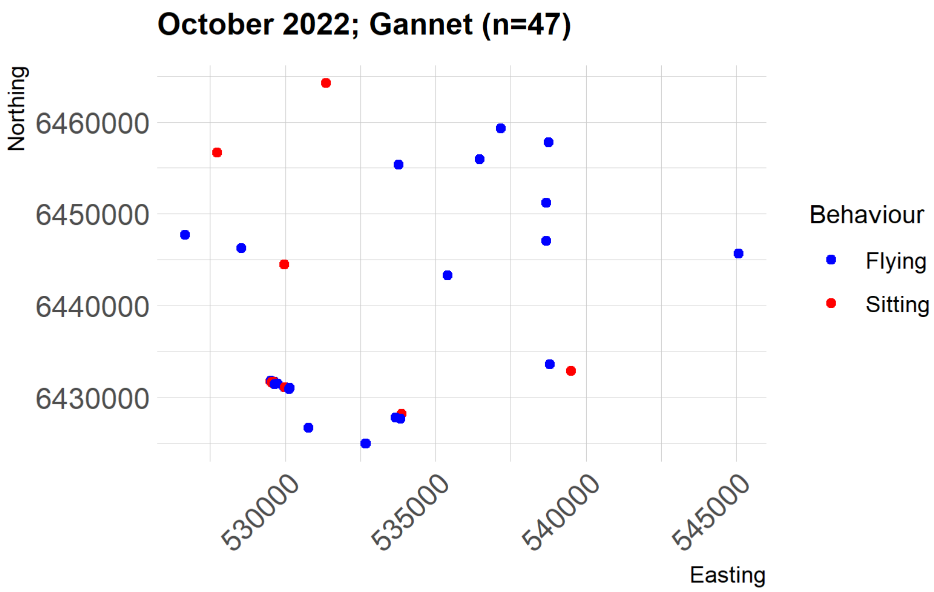

October 2021; Gannet (n=125)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

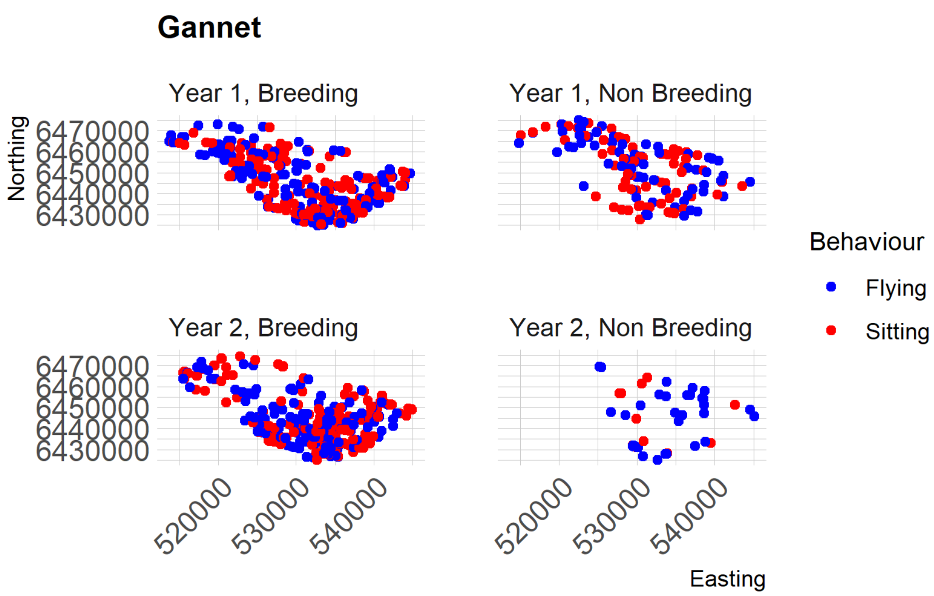

May 2022; Gannet (n=21)
Northing
6460000
6450000
6440000
6430000
515000
520000
525000
530000
535000
540000
Easting
Behaviour
Flying
Sitting


June 2022; Gannet (n=258)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting


July 2022; Gannet (n=72)
Northing
6470000
6460000
6450000
6440000
6430000
525000
530000
535000
540000
545000
Easting
Behaviour
Flying
Sitting

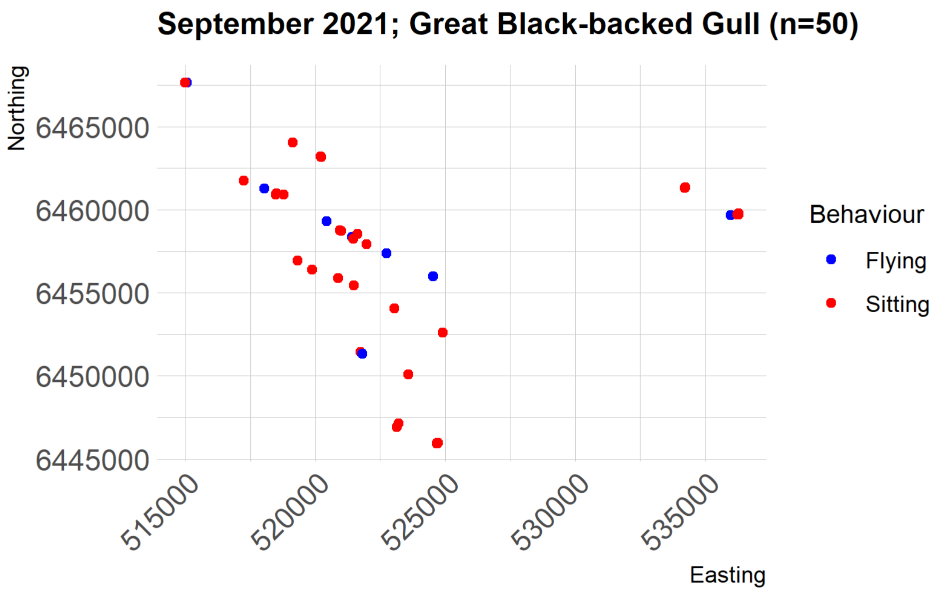

September 2022; Gannet (n=104)
Northing
6470000
6460000
6450000
6440000
6430000
515000
520000
525000
530000
535000
540000
Easting
Behaviour
Flying
Sitting

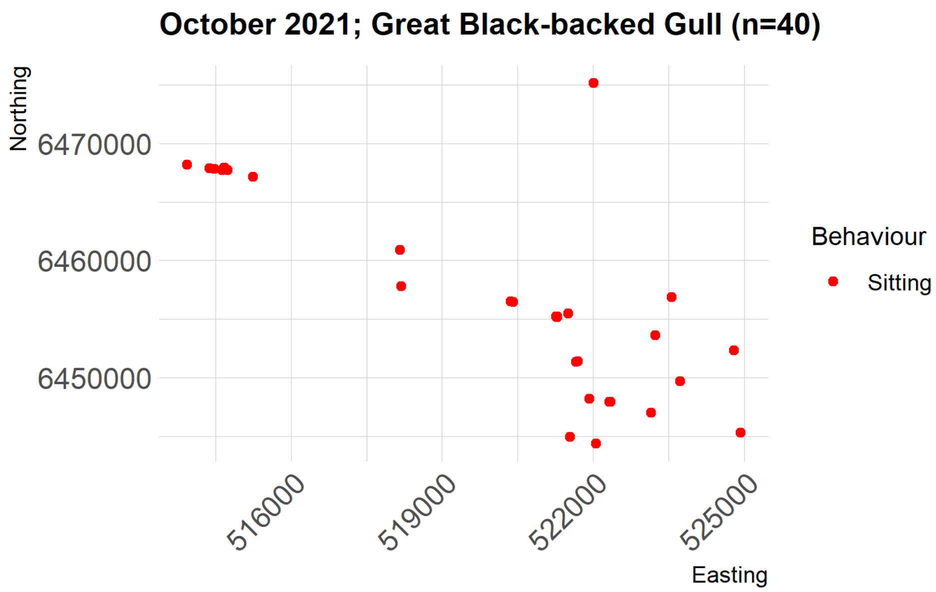

October 2022; Gannet (n=47)
Northing
6460000
6450000
6440000
6430000
530000
535000
540000
545000
Easting
Behaviour
Flying
Sitting

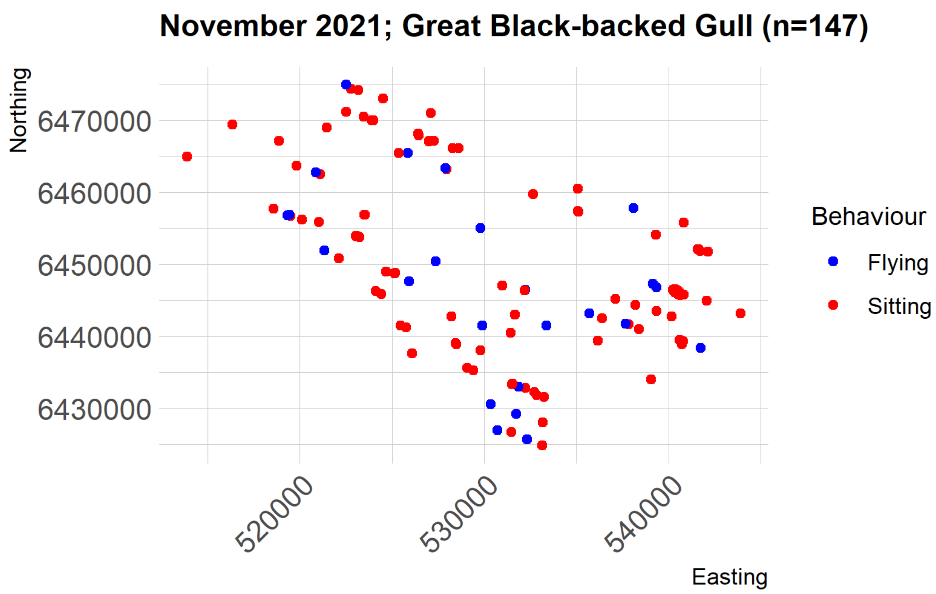

Gannet
Year 1, Breeding
Year 1, Non Breeding
Year 2, Breeding
Year 2, Non Breeding
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

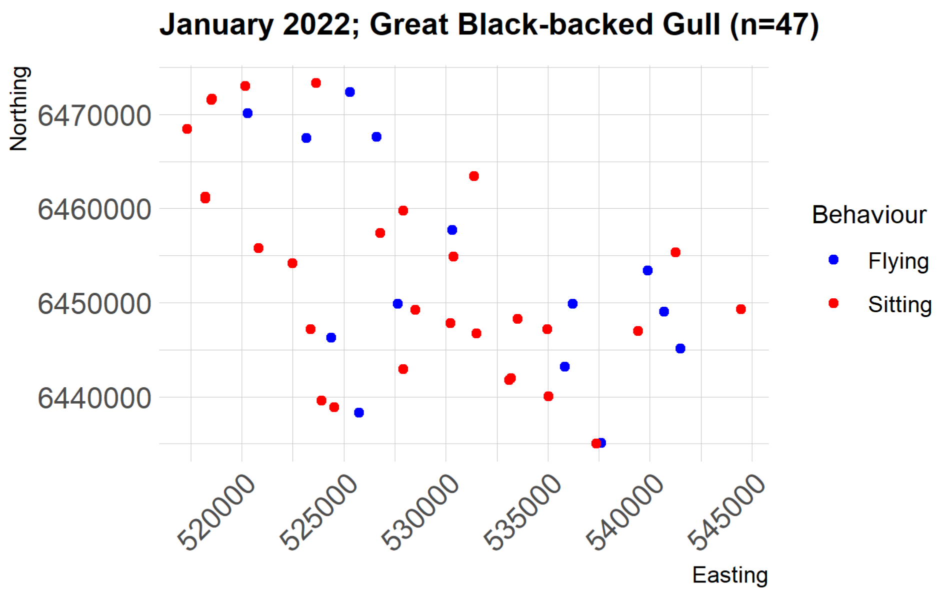
I.3.7 Great Black-backed Gull

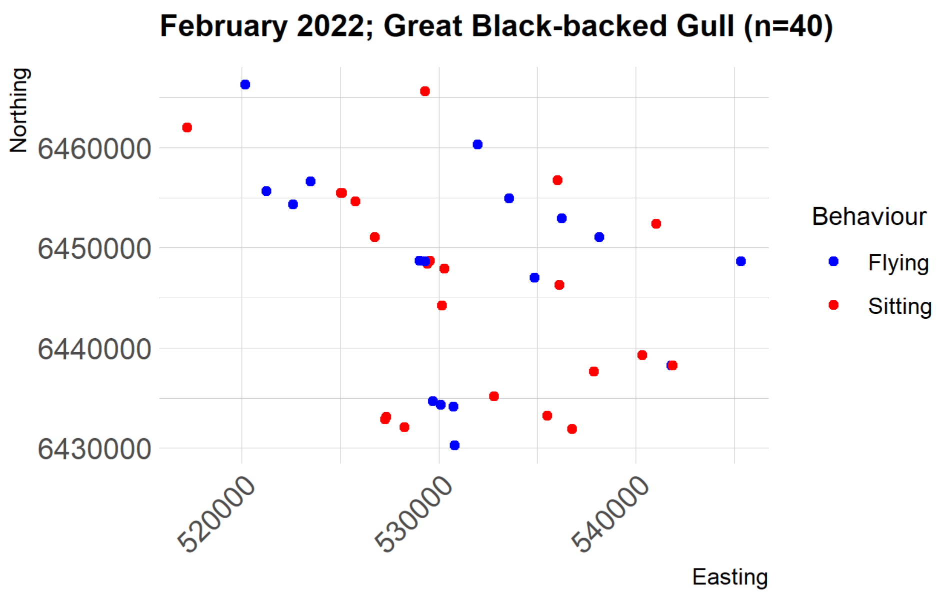
We suggest modelling Great Black-backed Gull (N=540) by survey-year because only 8/24 surveys had enough observations to model, and all occurred outside of the breeding season (April - August). Some surveys have minimal observations of flying birds, therefore behaviour would be better captured on the larger temporal unit of 'survey year'. We also noticed that there are about 64 observations of this species where the behaviour is listed as "perched". Depending on what the birds are perched upon we may be able to use these observations. For example, if the gulls are perched on turbines, we would not be able to model them with a distance to turbine covariate.

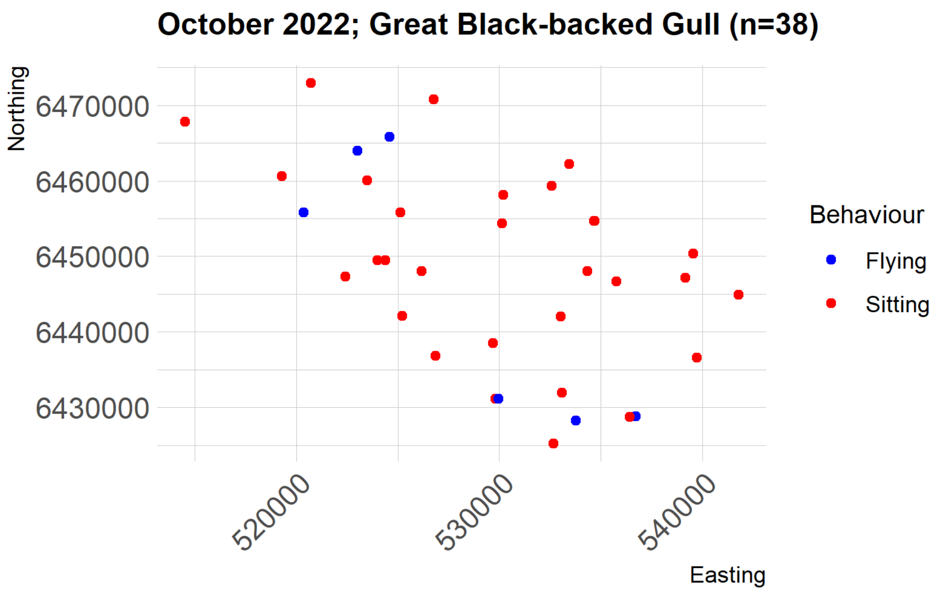


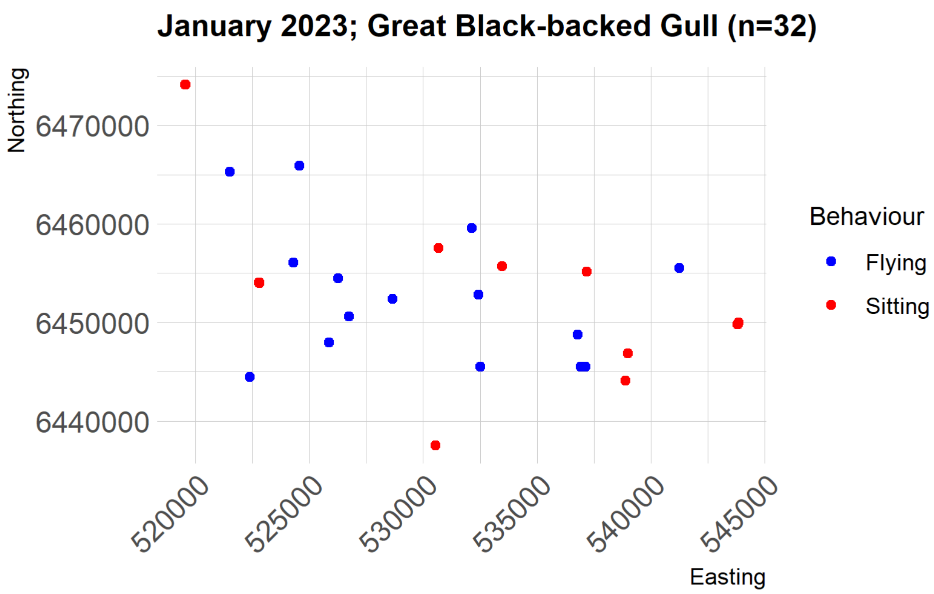

October 2021; Great Black-backed Gull (n=40)
Northing
6470000
6460000
6450000
516000
519000
522000
525000
Easting
Behaviour
Sitting

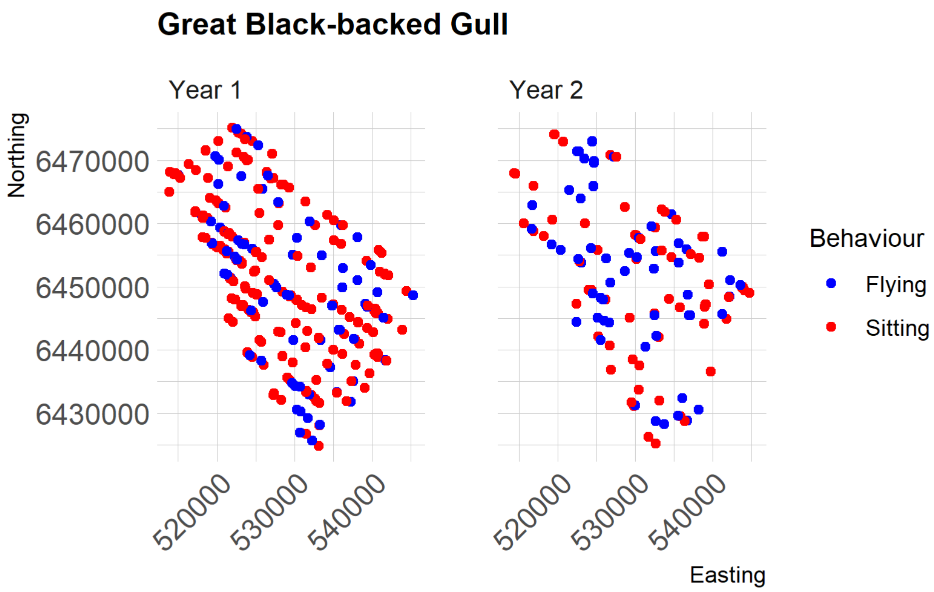

November 2021; Great Black-backed Gull (n=147)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting


January 2022; Great Black-backed Gull (n=47)
Northing
6470000
6460000
6450000
6440000
520000
525000
530000
535000
540000
545000
Easting
Behaviour
Flying
Sitting


February 2022; Great Black-backed Gull (n=40)
Northing
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

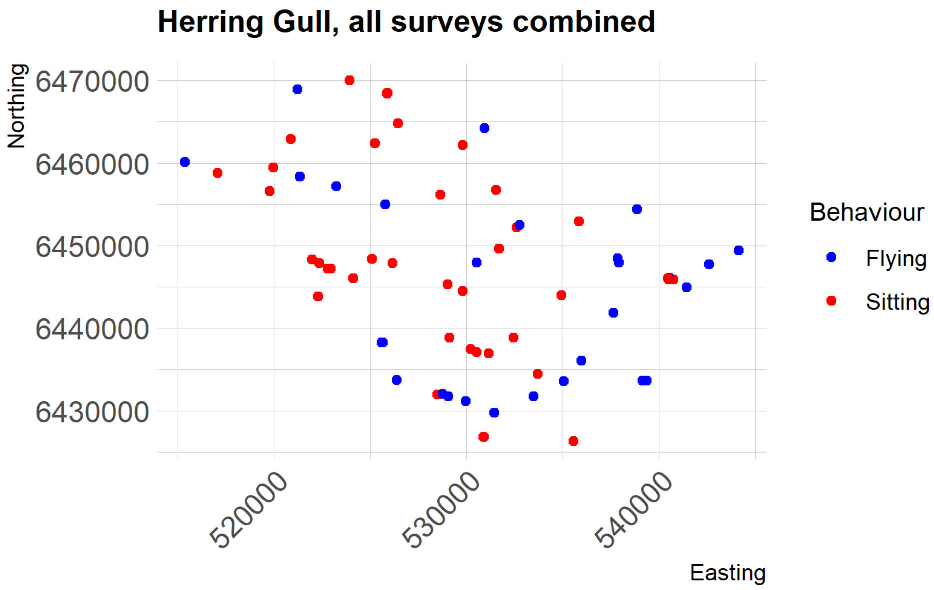

October 2022; Great Black-backed Gull (n=38)
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting


January 2023; Great Black-backed Gull (n=32)
Northing
6470000
6460000
6450000
6440000
520000
525000
530000
535000
540000
545000
Easting
Behaviour
Flying
Sitting


Great Black-backed Gull
Year 1
Year 2
Northing
6470000
6460000
6450000
6440000
6430000
520000
530000
540000
Easting
Behaviour
Flying
Sitting

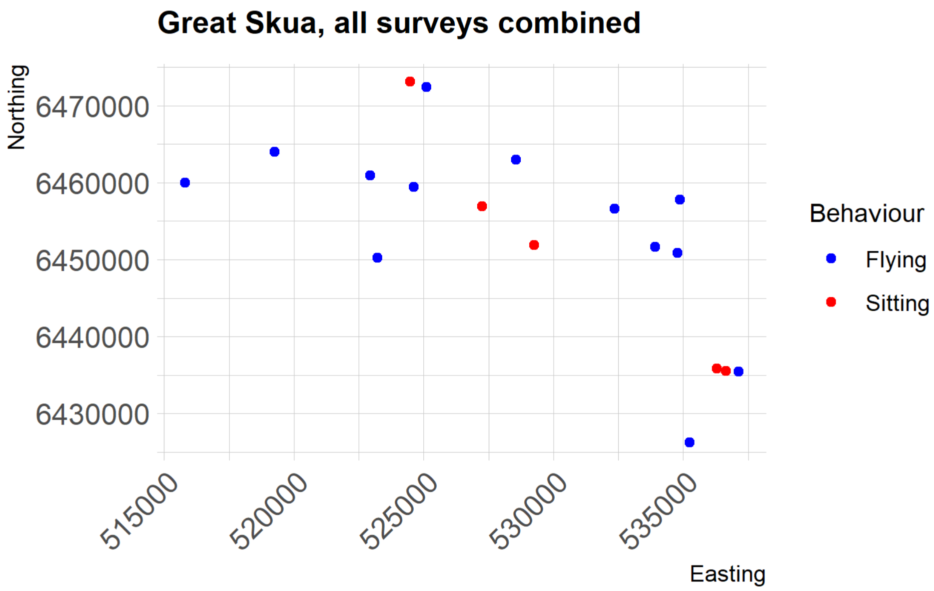
I.3.8 Herring Gull

With minimal observations, we will attempt to model herring gull (N=99) for all behaviours and surveys combined. This would constitute a "hot spot" analysis, which would identify potential areas of importance for this species within the study area, however this may be at a temporal scale too large to be very useful.



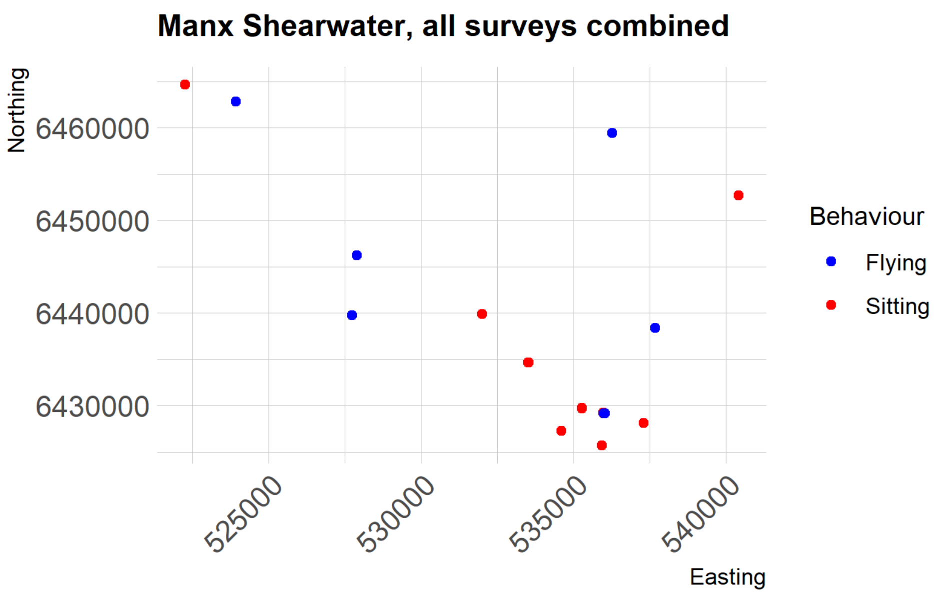
I.3.9 Great Skua

Great skua observations were too few to be modeled (N=22).



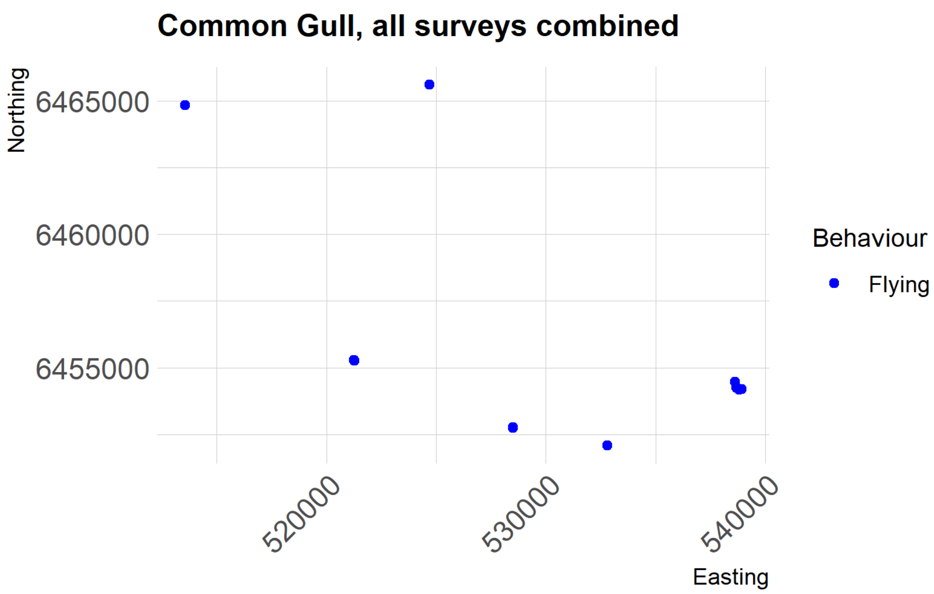
I.3.10 Manx Shearwater

Common tern observations were too few to be modelled (N=28).



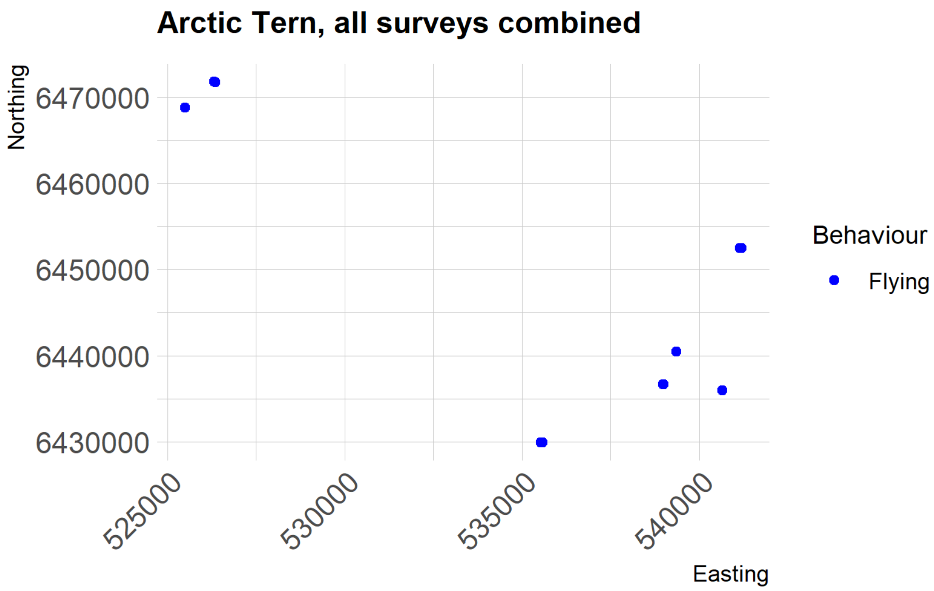
I.3.11 Common Gull

Common gull observations were too few to be modelled (N=13).



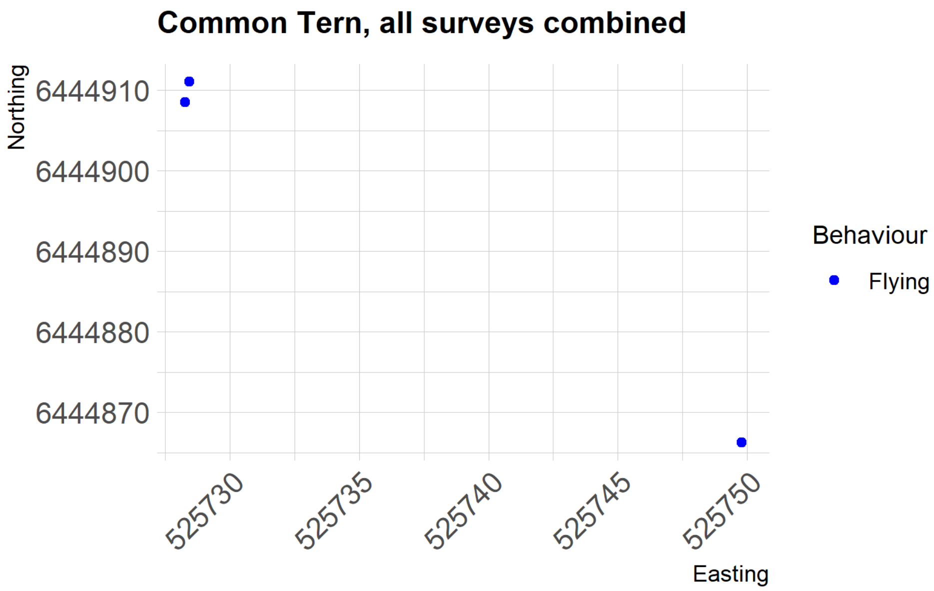
I.3.12 Arctic Tern

Arctic tern observations were too few to be modelled (N=15).



I.3.13 Common Tern

Common tern observations were too few to be modelled (N=3).



II Random Forests

II.1 Kittiwake

II.1.1 Model assessment

The top five models had RMSE values between 15.662 and 15.656 with R squared values between 0.0485 and 0.0499. These values were higher than the MRSea model, which had an R-squared value of 0.0426. The best model had an mtry value of 7, with a minimum node size of 80 (Figure II-1, Table II-1).

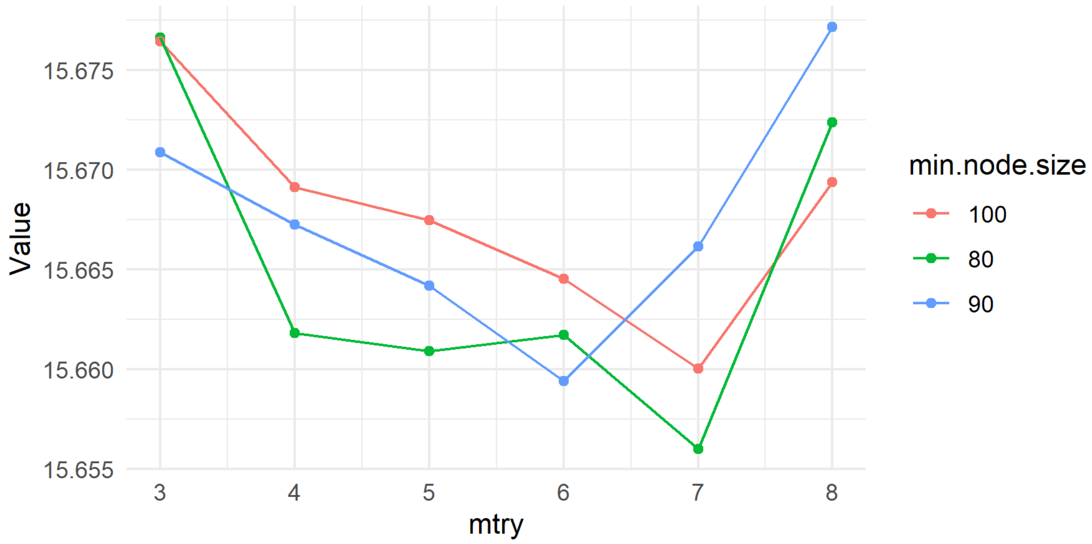


Figure II-1: Root mean squared error when varying mtry and min.node.size parameters in random forests

Table II-1: The top 5 models selected by the random forests analysis showing root mean squared error (RMSE) and R-squared values as calculated by 5-fold cross validation

mtry	min.node.size	RMSE	Rsquared
7	80	15.656	0.0499
6	90	15.659	0.0480
7	100	15.660	0.0485
5	80	15.661	0.0472
6	80	15.662	0.0485

II.1.2 Variable importance

The top predictor variables were daily sea surface temperature, northing (y.pos), and monthly mean sea surface temperature, followed by survey ID. This shows that most of the signal in the data come from variables that represent temporal variability (Table II-2).

Table II-2: The top 5 predictor variables from the random forests model and overall importance

Variable	Overall importance
SST_daily	100.00
y.pos	97.63
SST_mean	79.54
SurveyID^14	58.23
SurveyID^15	54.89

II.1.3 Distributional response

The partial relationship of distance to turbines on densities of kittiwake demonstrates an inverse relationship in that predicted densities decreased from 0 to 20km distance from turbines. The relationship reaches an inversion point at approximately 23km, where between 20 and 40km, there is a positive relationship between distance to turbine and density (Figure II-2). This suggests that turbines could be attracting kittiwake as the highest densities seems to be in areas nearer to turbines. However, this could simply be that kittiwake were more abundant around turbines for other reasons that we are unable to measure.

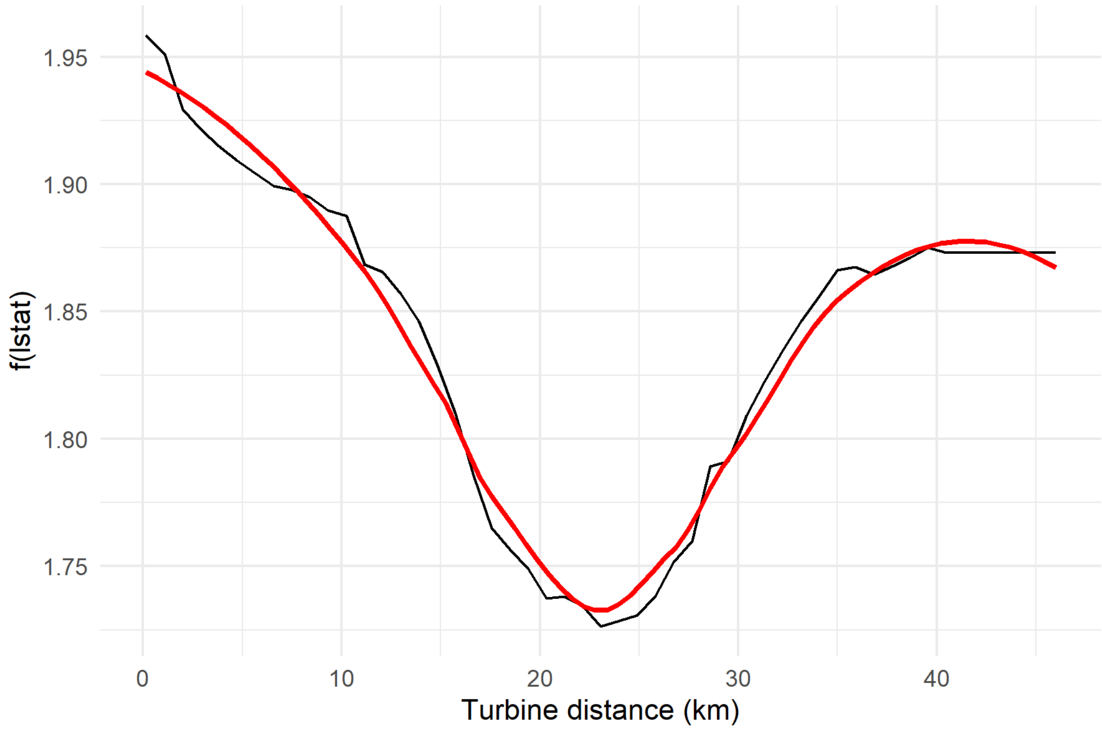


Figure II-2: Partial dependence plot of distance to turbine from the random forests model. The red line depicts the loess curve

II.1.4 Population estimates

Population estimates derived from the random forests baseline scenario (i.e., Existing turbines) fell well within the range of estimates derived from MRSea, sometimes only differing by 2 - 3%. We are therefore confident that the random forests kittiwake model is reliable, particularly in the context of the MRSea analysis.

Populations of kittiwake in the survey area were predicted to increase in all three scenarios due to the relationship between distance to turbine and density. The highest population increase was predicted for scenario 3, with percent change from baseline varying from -7 to 47% across all surveys (Table II-3).

Table II-3: Population estimates from the random forests models for the 4km buffer zone plus windfarm footprint survey area for the baseline scenario (i.e., based on currently installed turbines) and three potential turbine scenarios. Percent change from baseline is calculated for each scenario

SurveyID	Baseline	Scenario 1 (% change)	Scenario 2 (% change)	Scenario 3 (% change)
May 2021	1244	1398 (12.38)	1371 (10.21)	1418 (13.99)
June 2021	2142	2287 (6.77)	2213 (3.31)	2281 (6.49)
July 2021	6276	6850 (9.15)	6856 (9.24)	6954 (10.8)
August 2021	1982	2061 (3.99)	2039 (2.88)	2068 (4.34)
September 2021	919	853 (-7.18)	848 (-7.73)	854 (-7.07)
October 2021	614	606 (-1.3)	598 (-2.61)	597 (-2.77)
November 2021	953	1317 (38.2)	1344 (41.03)	1397 (46.59)
December 2021	223	230 (3.14)	231 (3.59)	235 (5.38)
February 2022	215	231 (7.44)	231 (7.44)	234 (8.84)
March 2022	274	269 (-1.82)	269 (-1.82)	269 (-1.82)
April 2022	643	636 (-1.09)	631 (-1.87)	631 (-1.87)
May 2022	3670	3817 (4.01)	3819 (4.06)	3818 (4.03)
June 2022	5027	5137 (2.19)	5118 (1.81)	5117 (1.79)
July 2022	5777	6330 (9.57)	6288 (8.85)	6360 (10.09)
August 2022	1048	1146 (9.35)	1137 (8.49)	1162 (10.88)
September 2022	346	343 (-0.87)	339 (-2.02)	343 (-0.87)
October 2022	251	252 (0.4)	252 (0.4)	253 (0.8)
February 2023	259	289 (11.58)	281 (8.49)	300 (15.83)
April 2023	650	633 (-2.62)	631 (-2.92)	631 (-2.92)

II.1.5 Distributions

The broad distribution of kittiwake in the baseline random forests models were nearly identical to those in the MRSea models. For Kittiwake, the models predict the highest densities in the southwest part of the survey area. In almost all cases, the densities of kittiwake increase through the site in relation to the locations of proposed turbines. This distributional response suggests a potential increase in the numbers of birds in the windfarm site (Figures II-3 – II-18).

II.1.5.1 Baseline scenarios

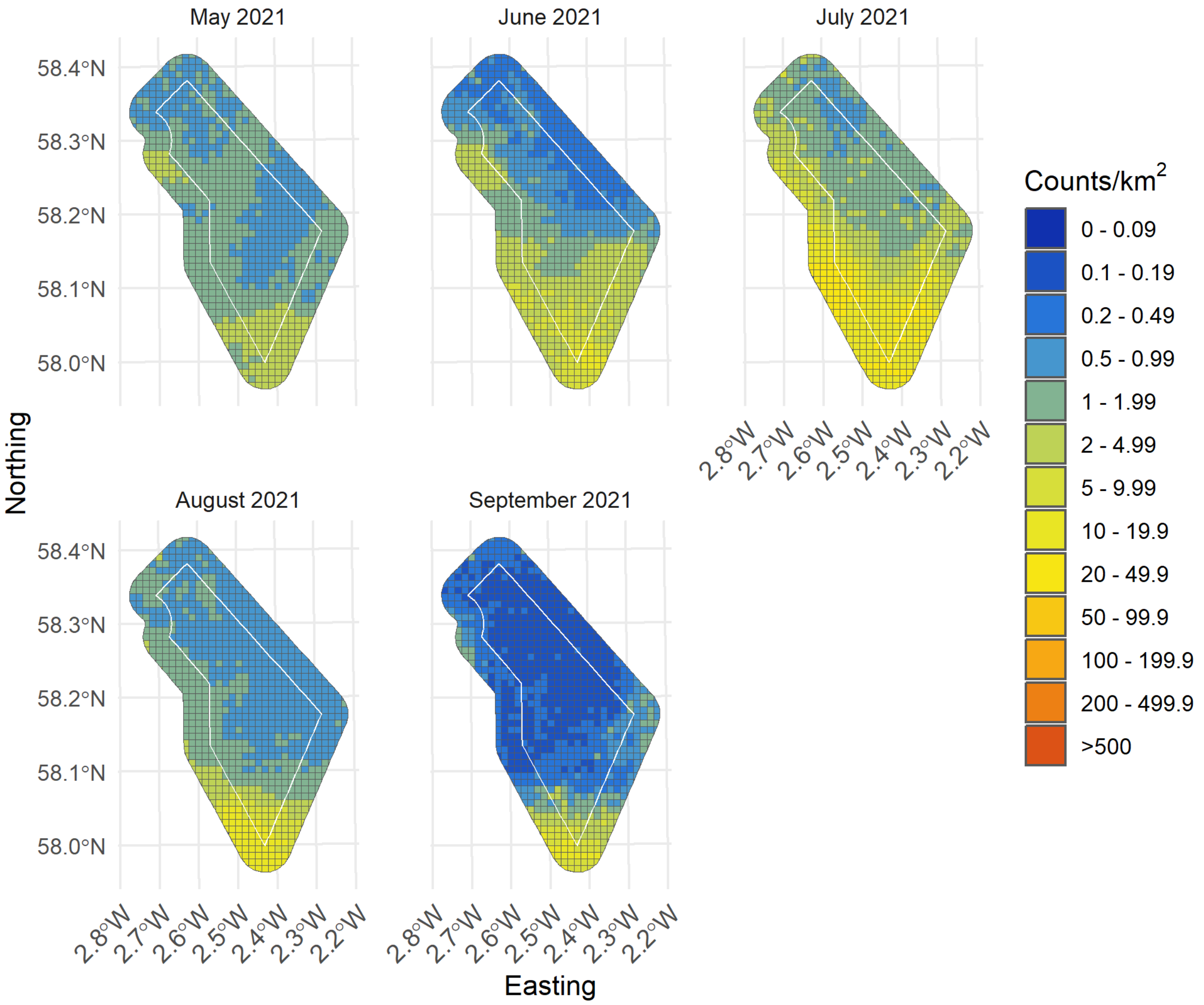


Figure II-3: Random forests baseline predictions of Kittiwake from May 2021 to September 2021

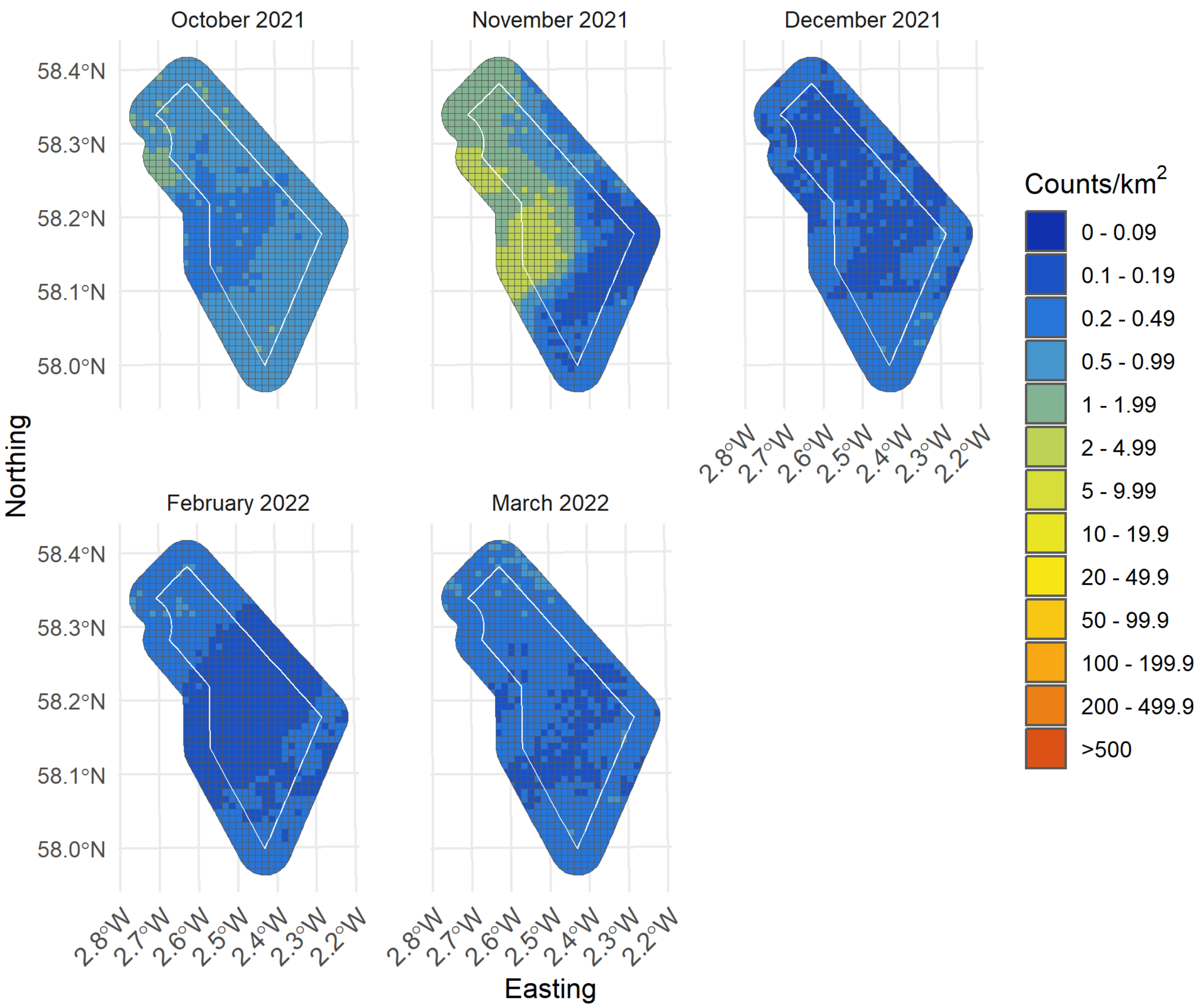


Figure II-4: Random forests baseline predictions of Kittiwake from October 2021 to March 2022

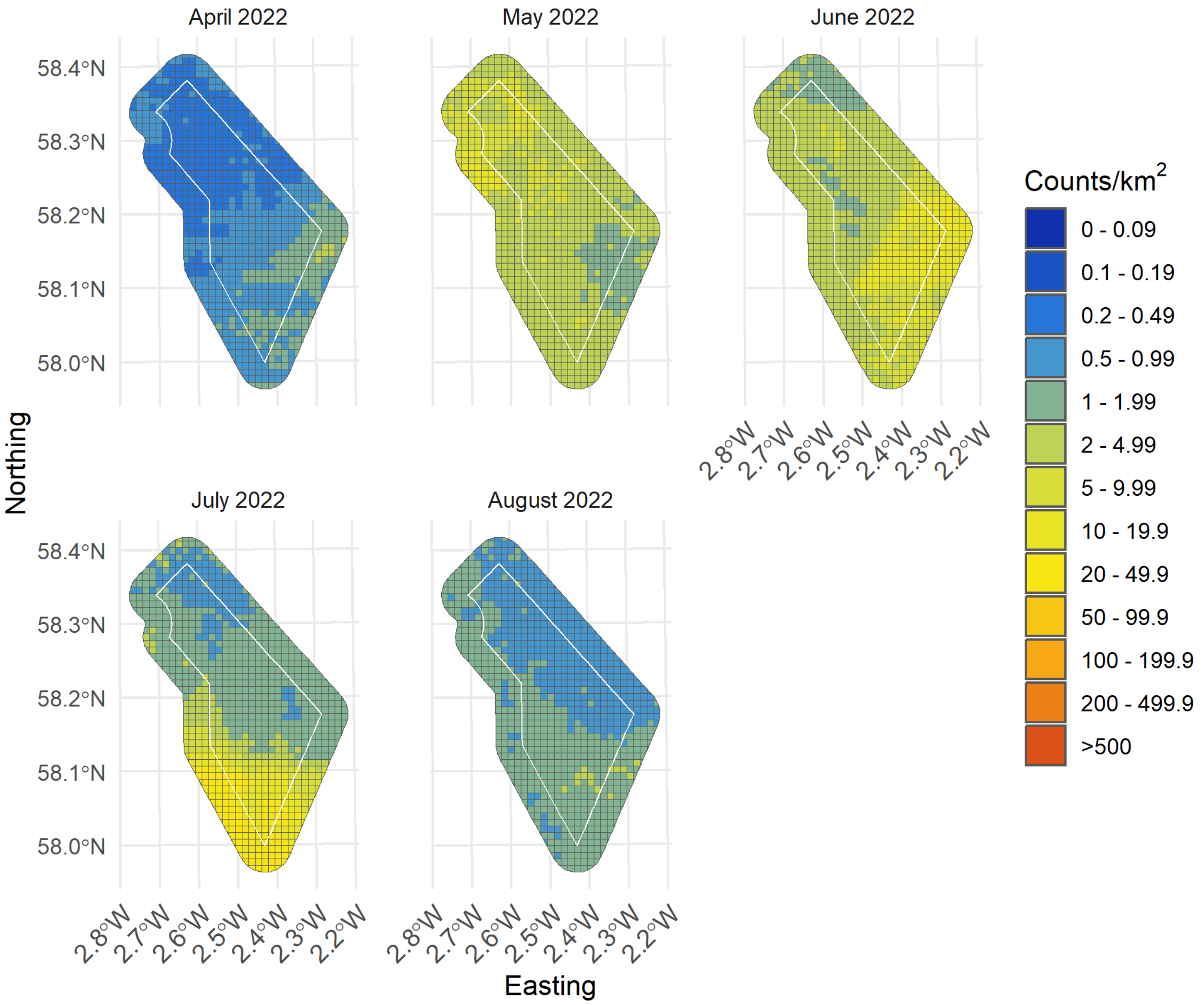


Figure II-5: Random forests baseline predictions of Kittiwake from April 2022 to August 2022

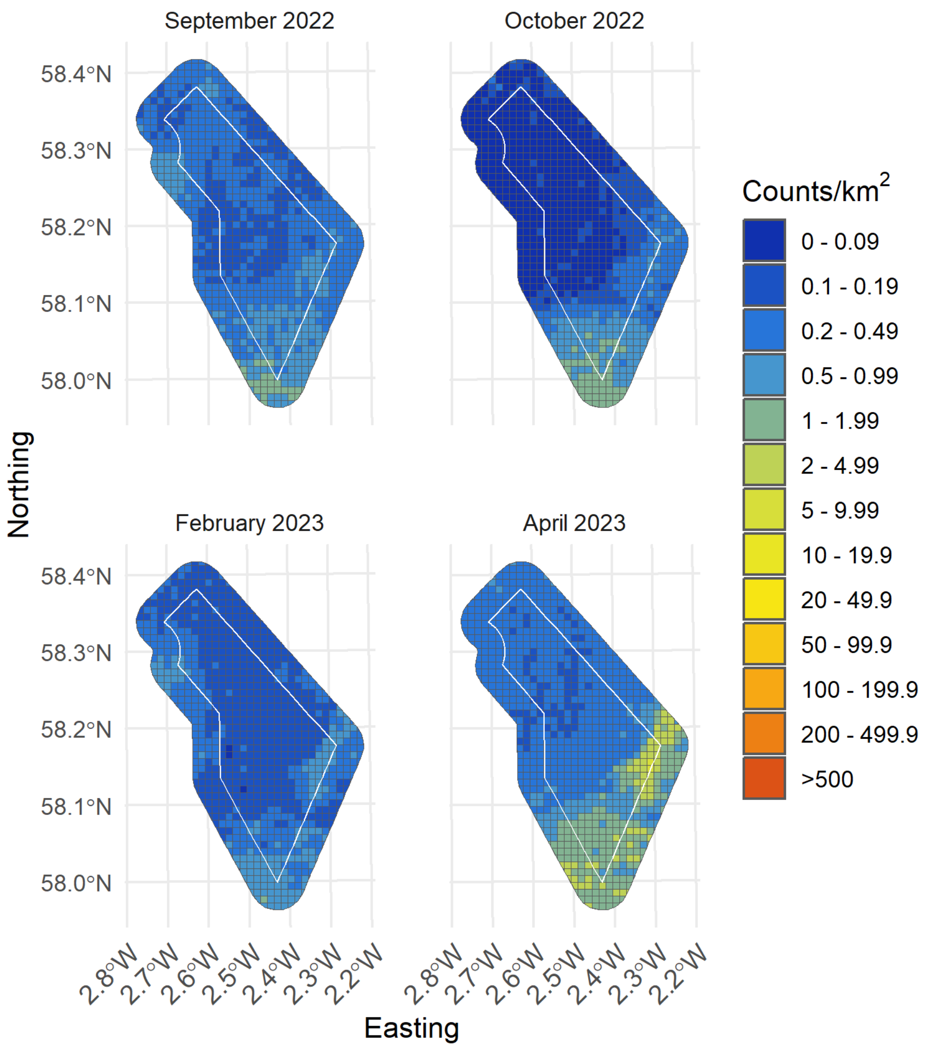


Figure II-6: Random forests baseline predictions of Kittiwake from September 2022 to April 2023

II.1.5.2 Turbine scenario 1

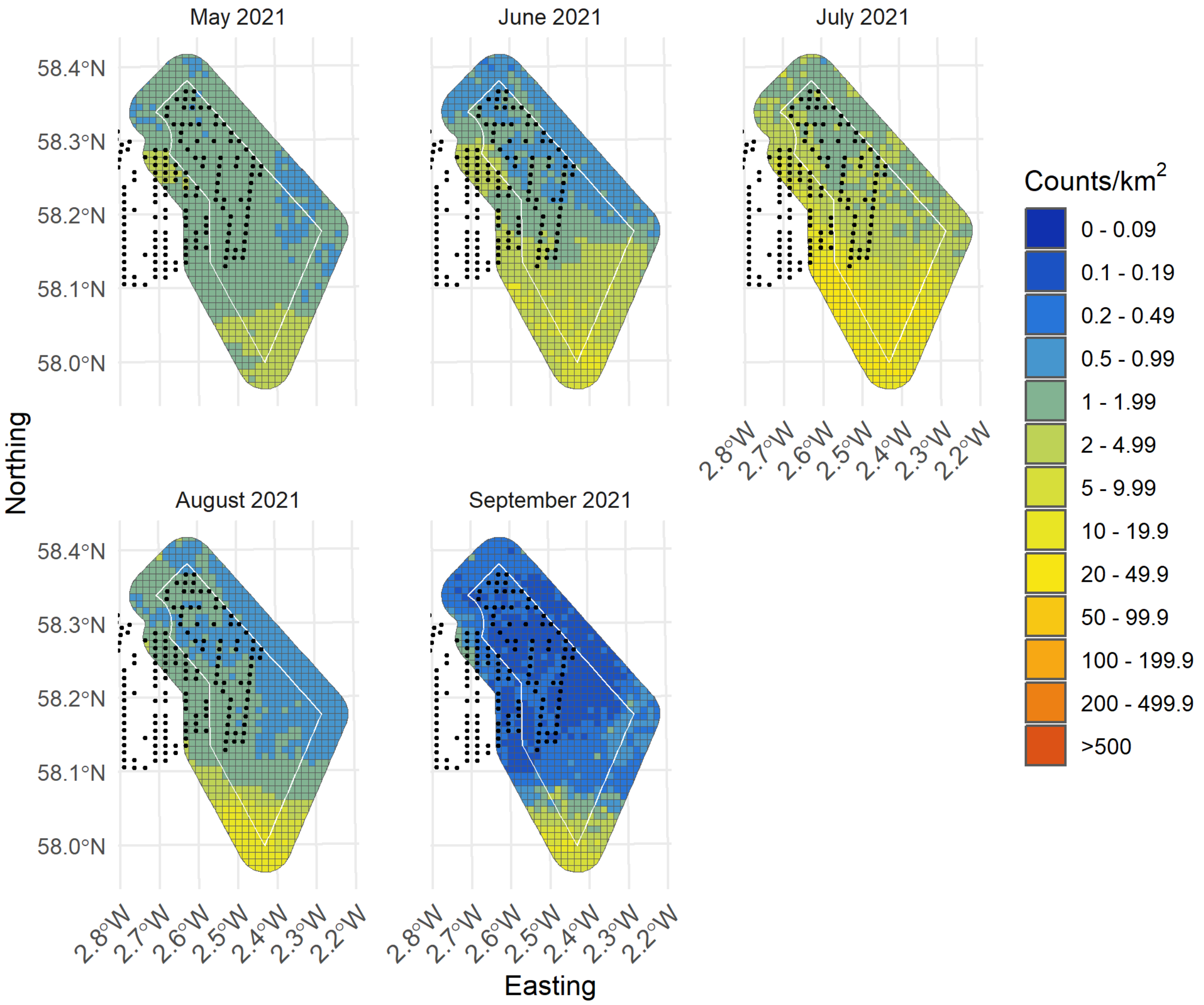


Figure II-7: Random forests predictions from turbine scenario 1 of Kittiwake from May 2021 to September 2021 with turbines presented as black dots

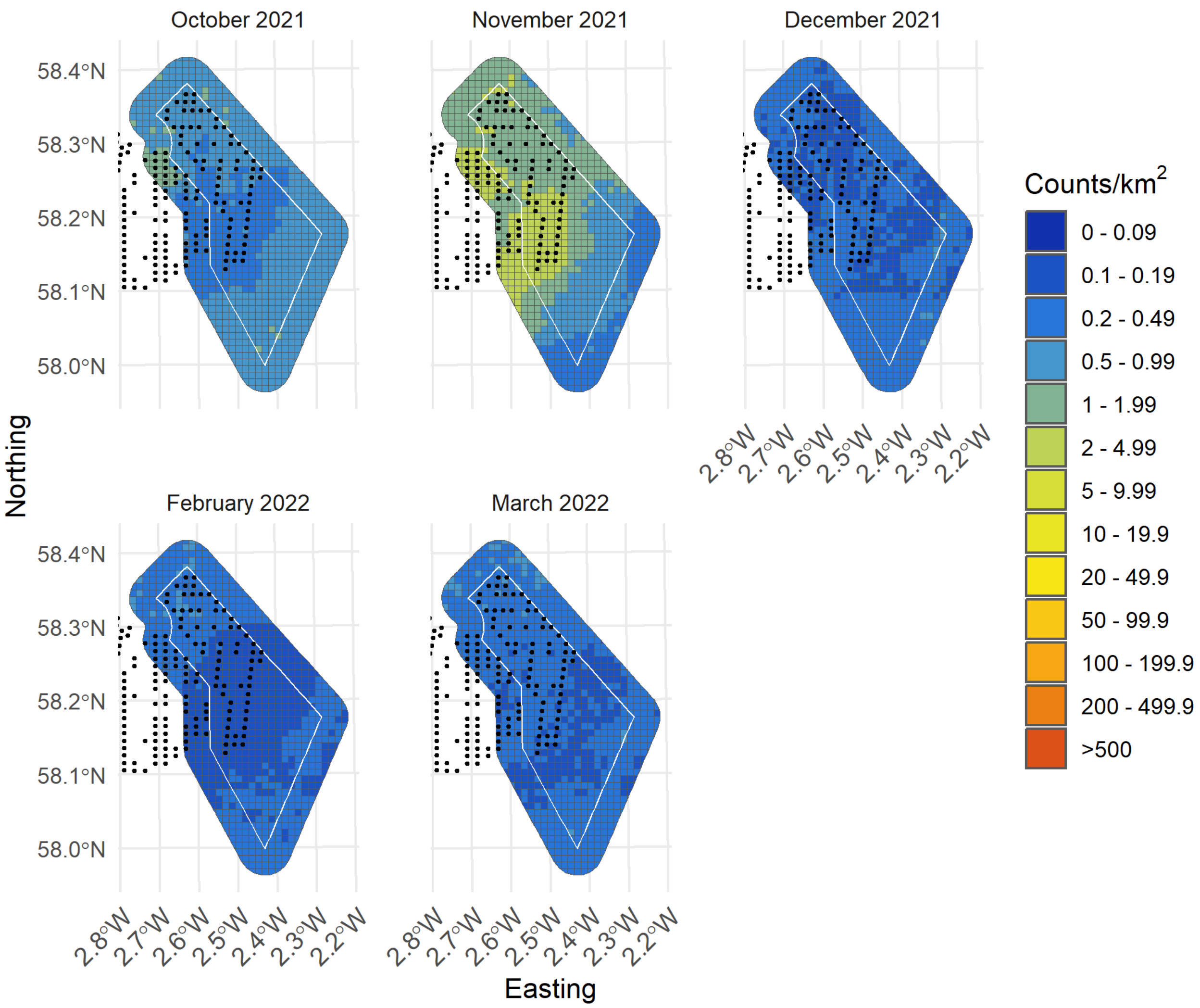


Figure II-8: Random forests predictions from turbine scenario 1 of Kittiwake from October 2021 to March 2022 with turbines presented as black dots

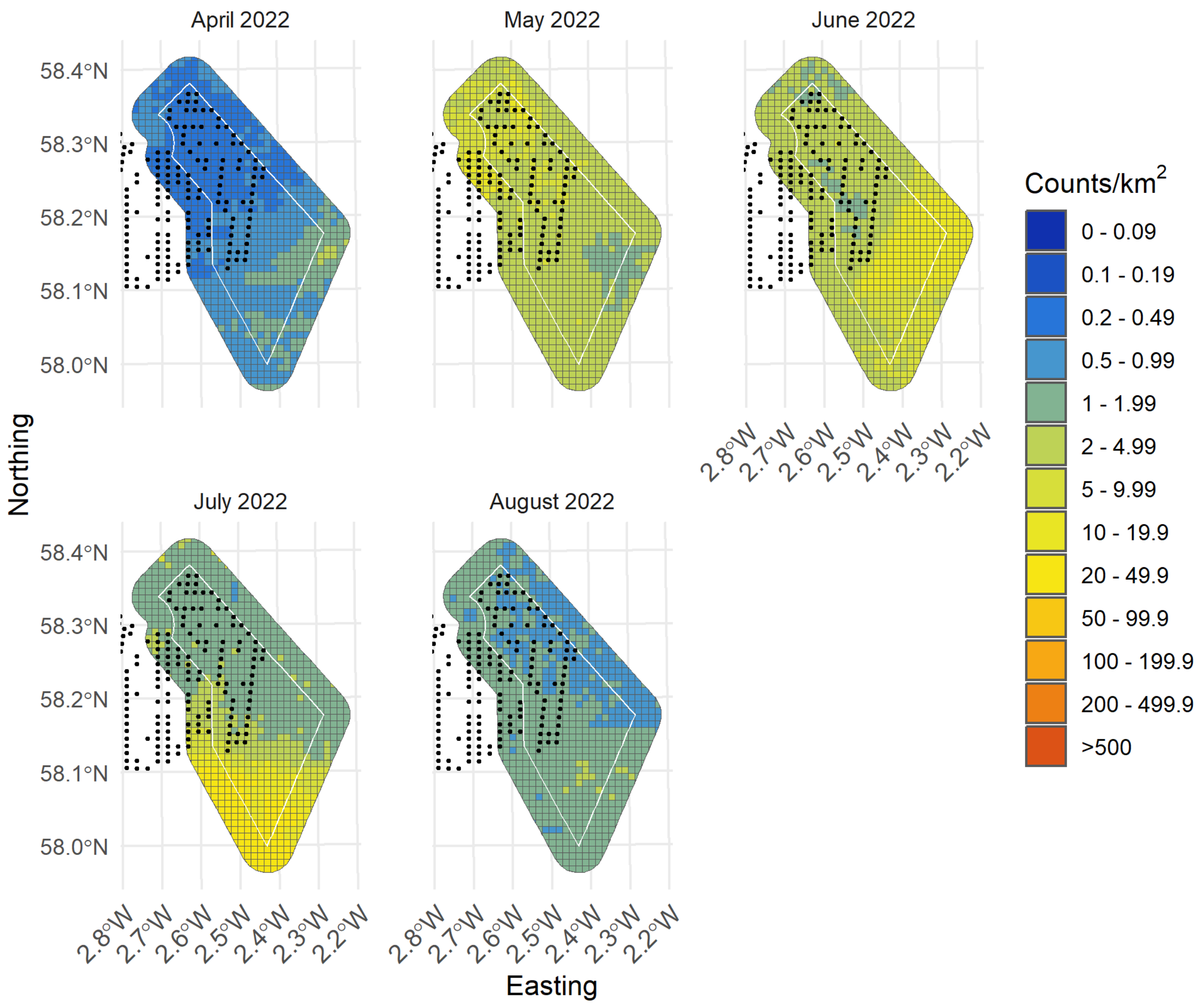


Figure II-9: Random forests predictions from turbine scenario 1 of Kittiwake from April 2022 to August 2022 with turbines presented as black dots

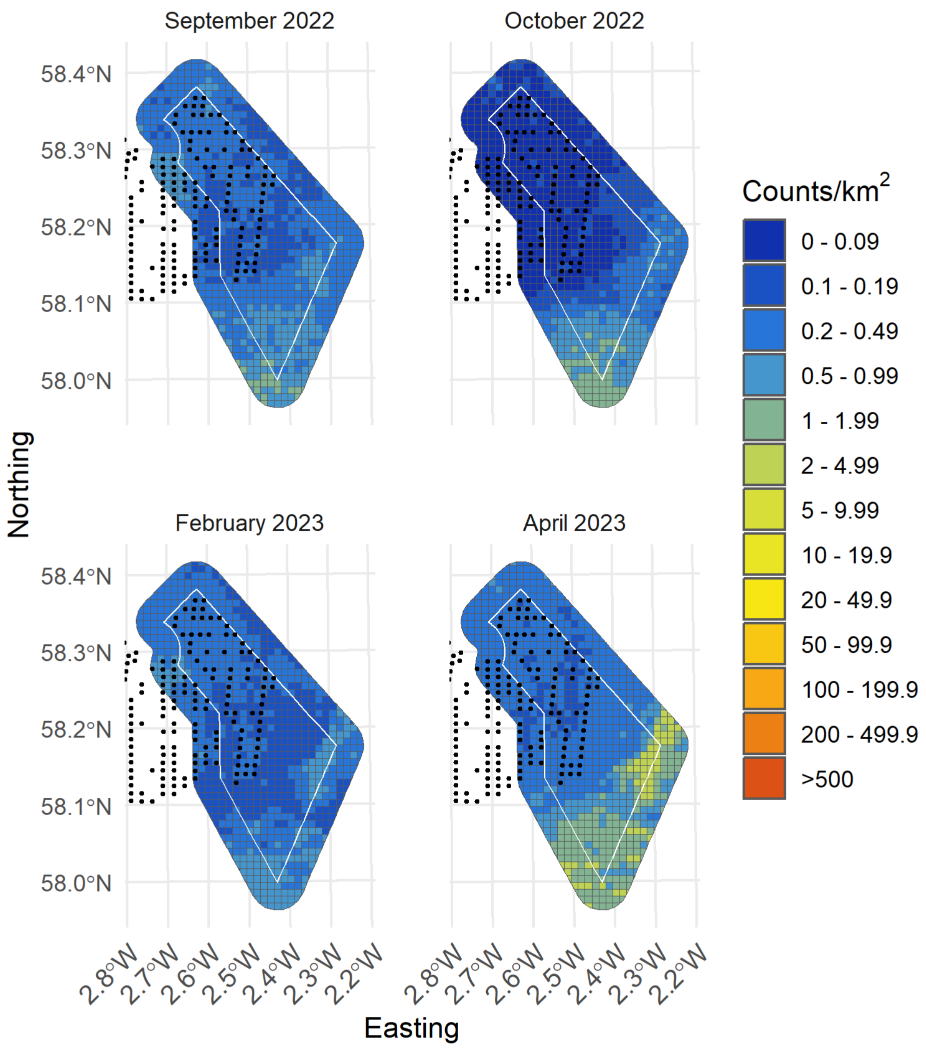


Figure II-10: Random forests predictions from turbine scenario 1 of Kittiwake from September 2022 to April 2023 with turbines presented as black dots

II.1.5.3 Turbine scenario 2

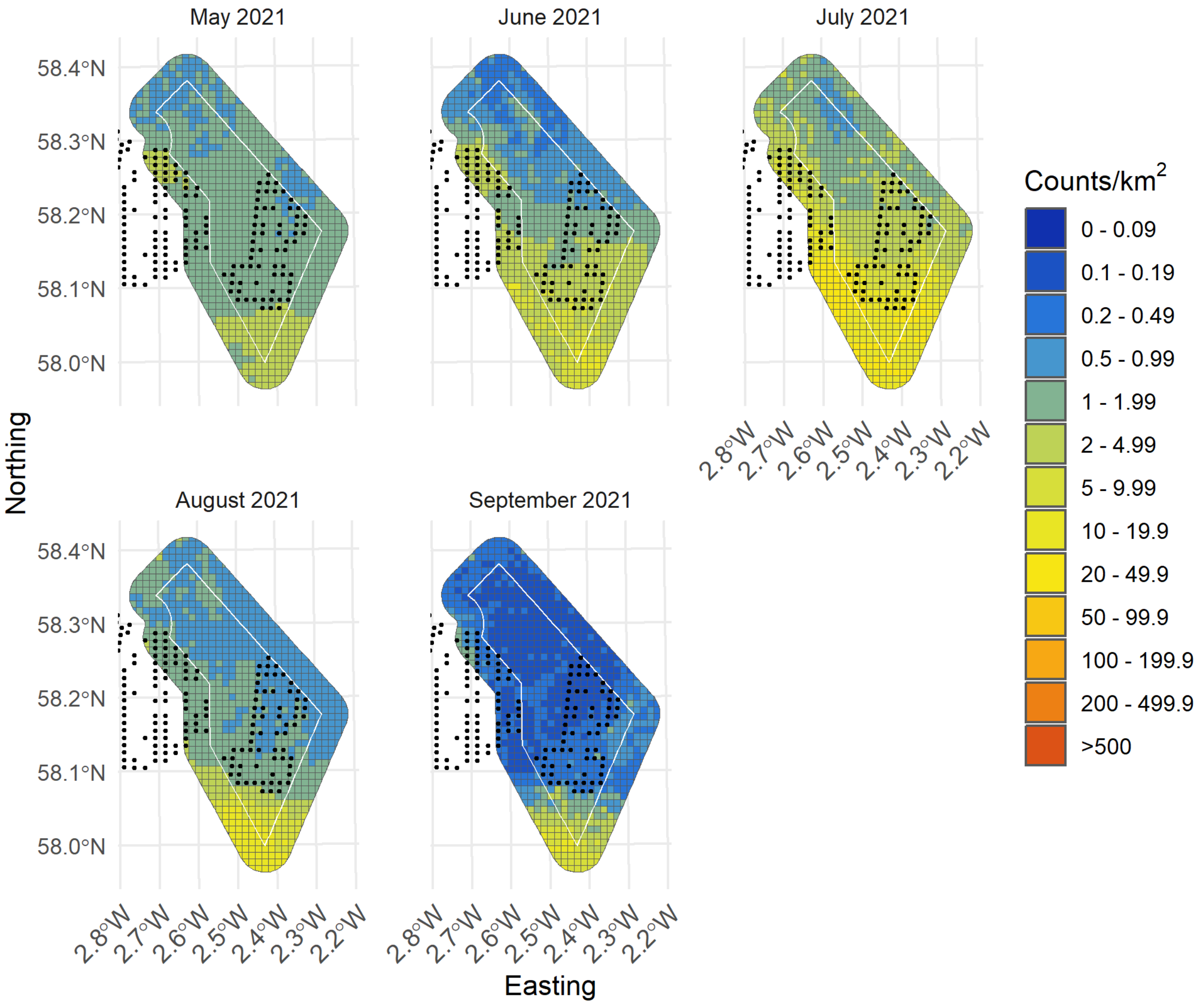


Figure II-11: Random forests predictions from turbine scenario 2 of Kittiwake from May 2021 to September 2021 with turbines presented as black dots

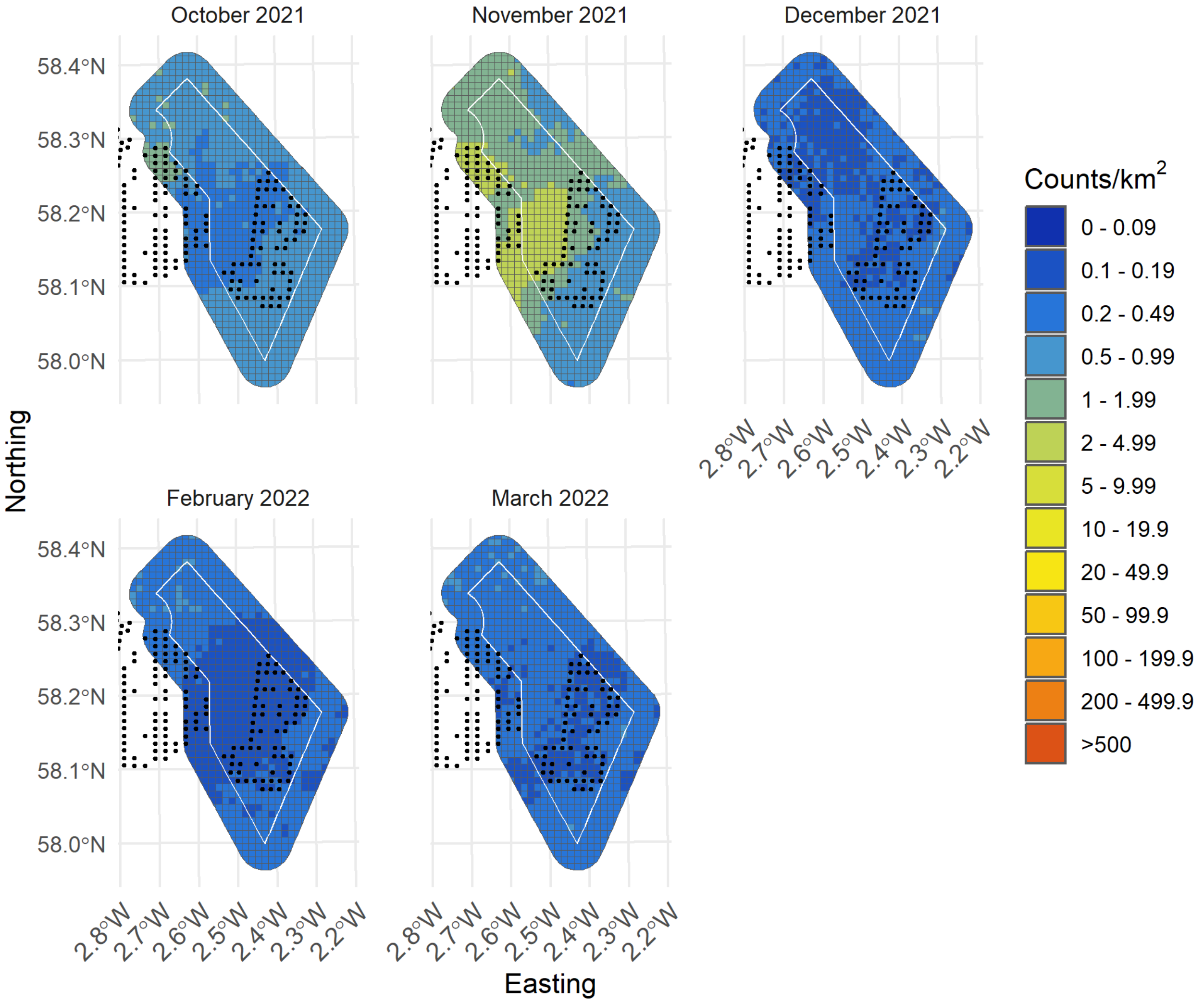


Figure II-12: Random forests predictions from turbine scenario 2 of Kittiwake from October 2021 to March 2022 with turbines presented as black dots

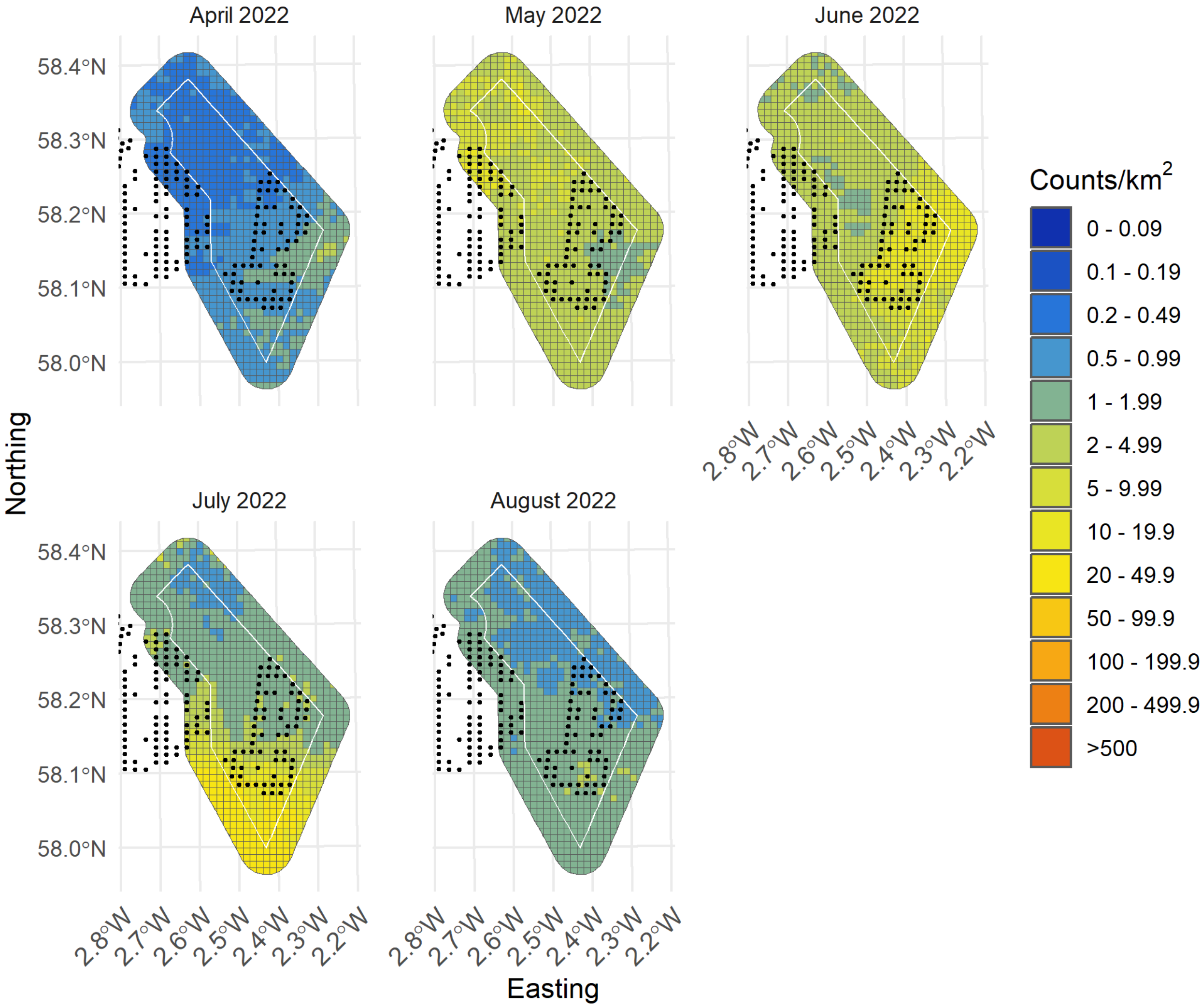


Figure II-13: Random forests predictions from turbine scenario 2 of Kittiwake from April 2022 to August 2022 with turbines presented as black dots

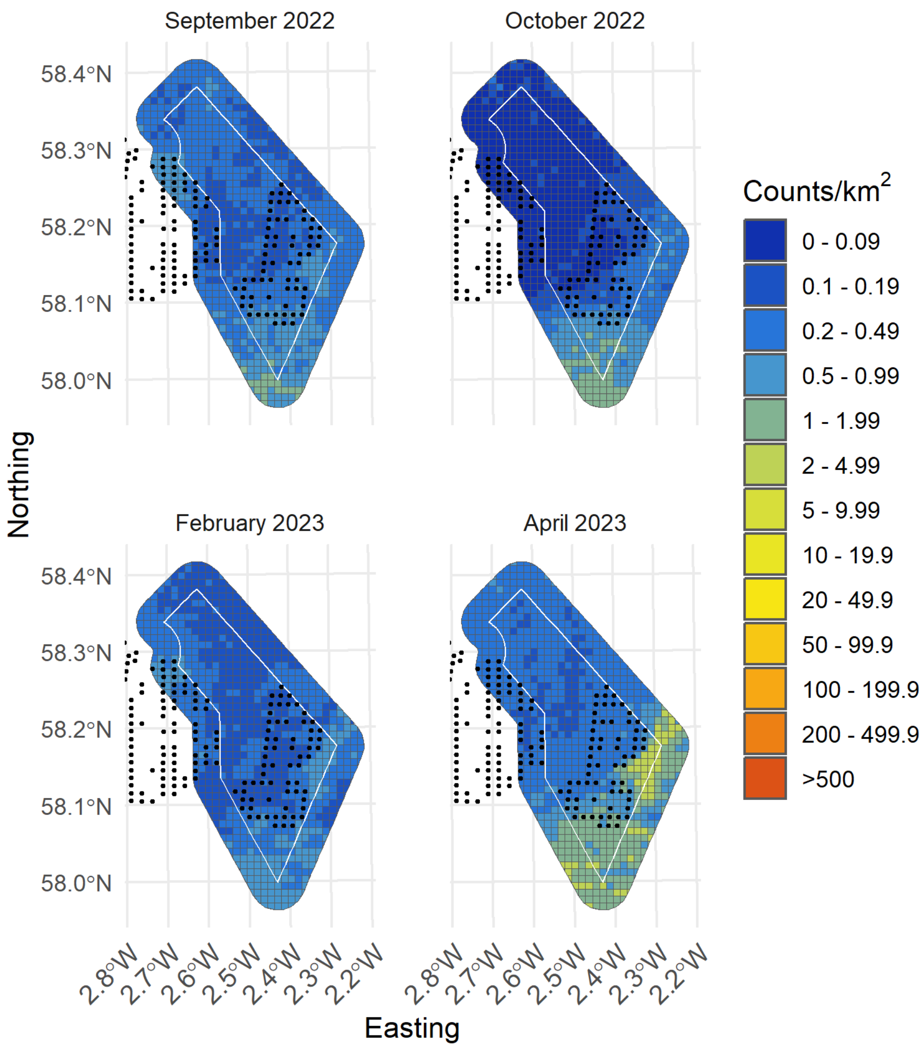


Figure II-14: Random forests predictions from turbine scenario 2 of Kittiwake from September 2022 to April 2023 with turbines presented as black dots

II.1.5.4 Turbine scenario 3

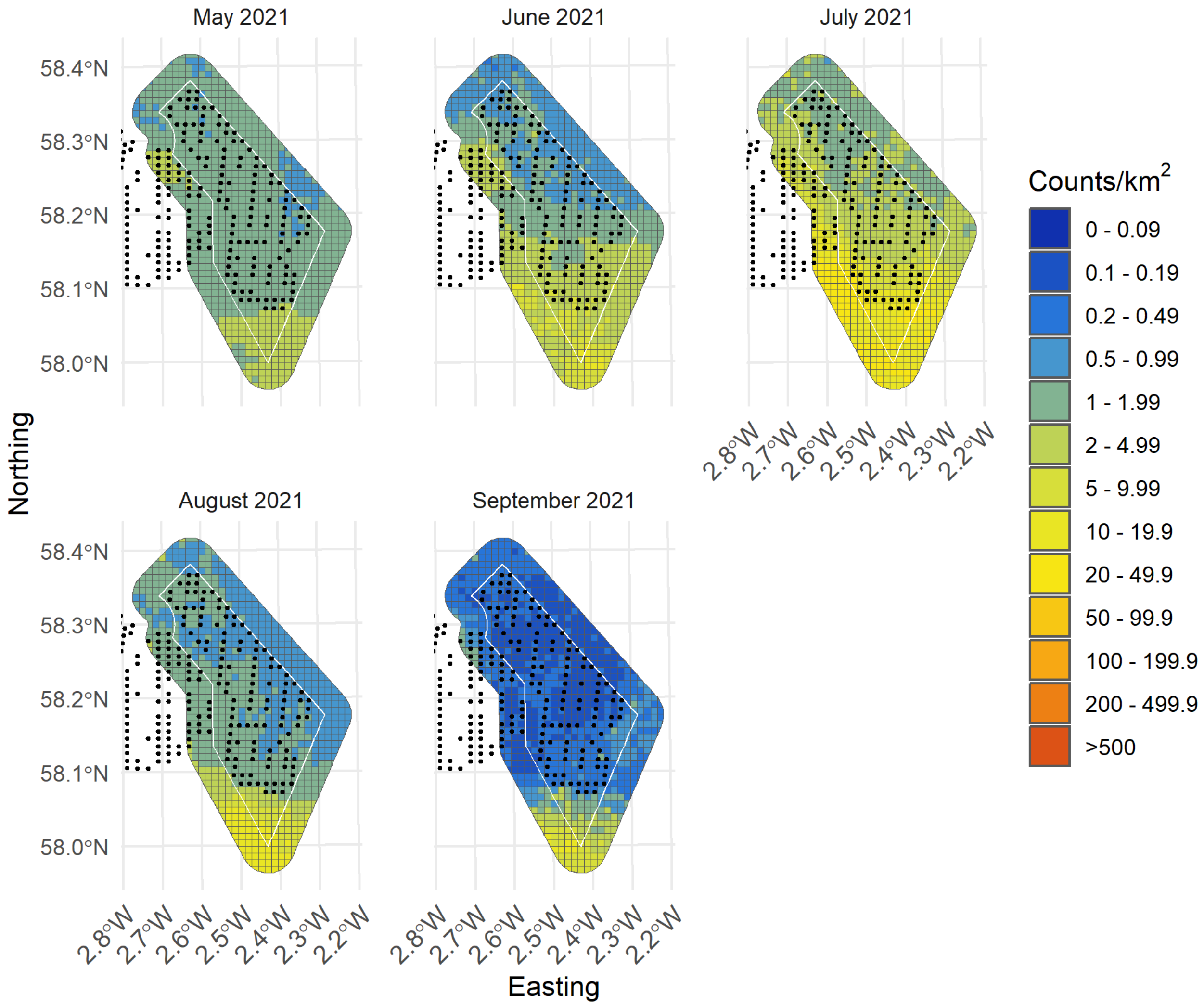


Figure II-15: Random forests predictions from turbine scenario 3 of Kittiwake from May 2021 to September 2021 with turbines presented as black dots

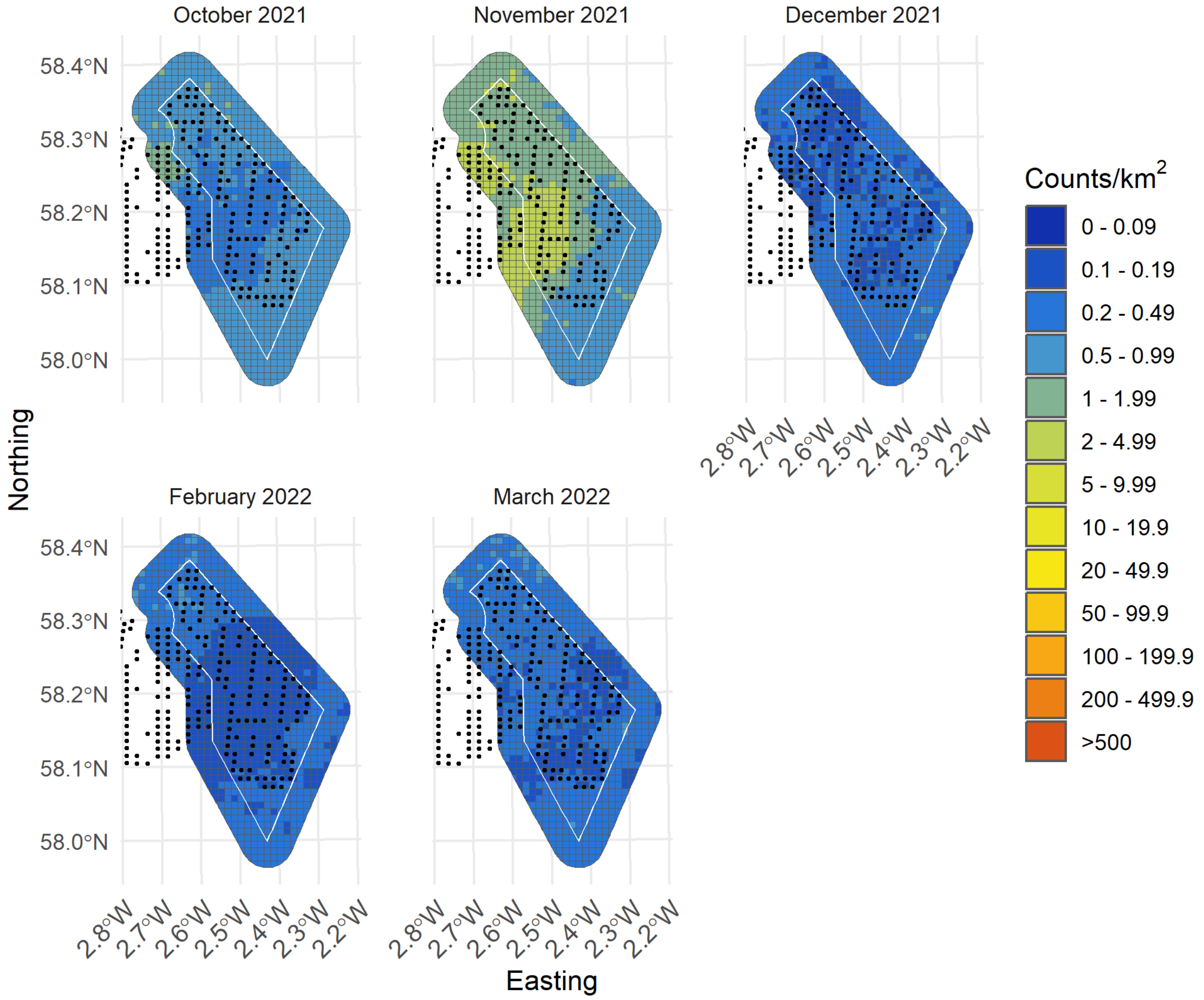


Figure II-16: Random forests predictions from turbine scenario 3 of Kittiwake from October 2021 to March 2022 with turbines presented as black dots

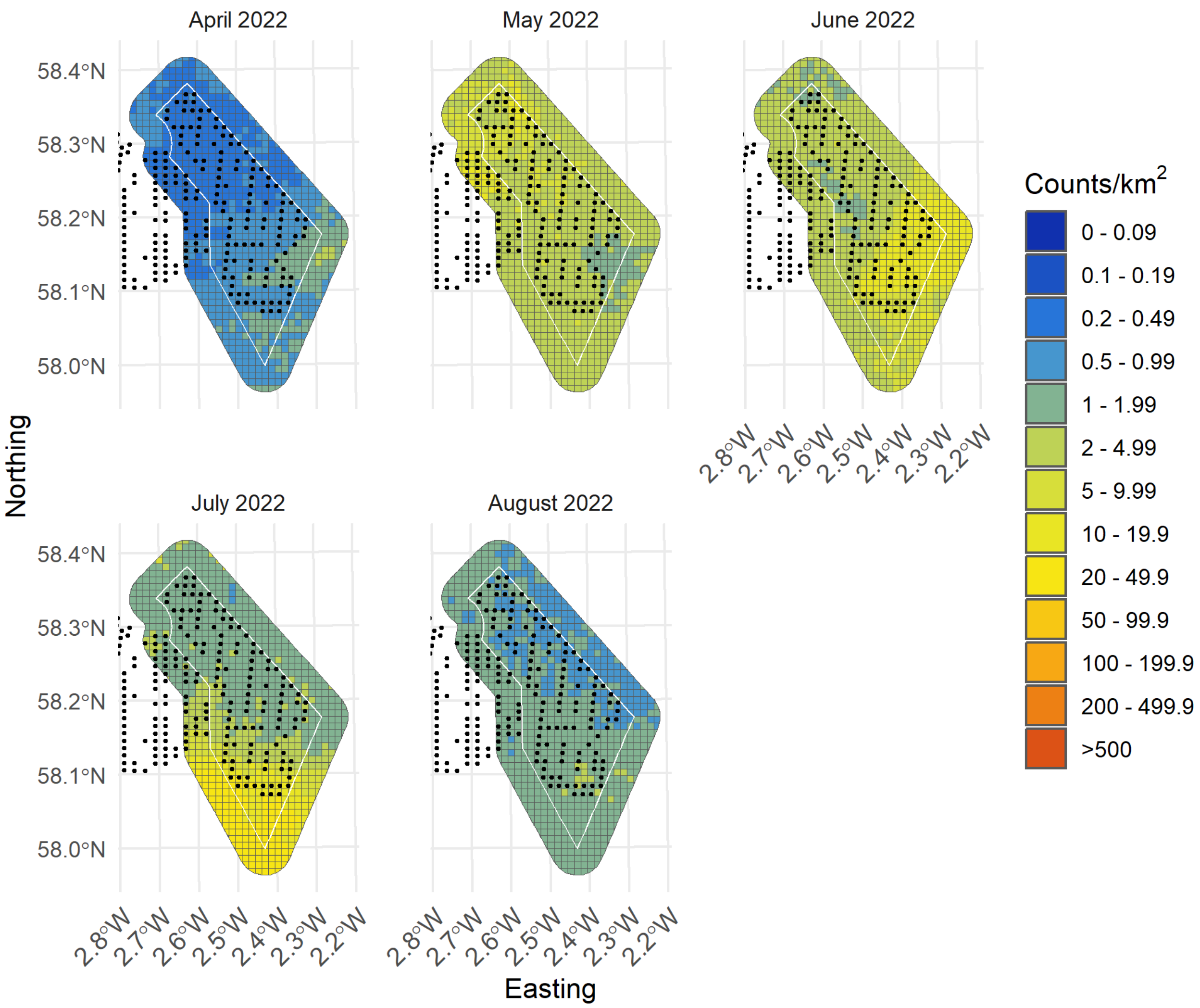


Figure II-17: Random forests predictions from turbine scenario 3 of Kittiwake from April 2022 to August 2022 with turbines presented as black dots

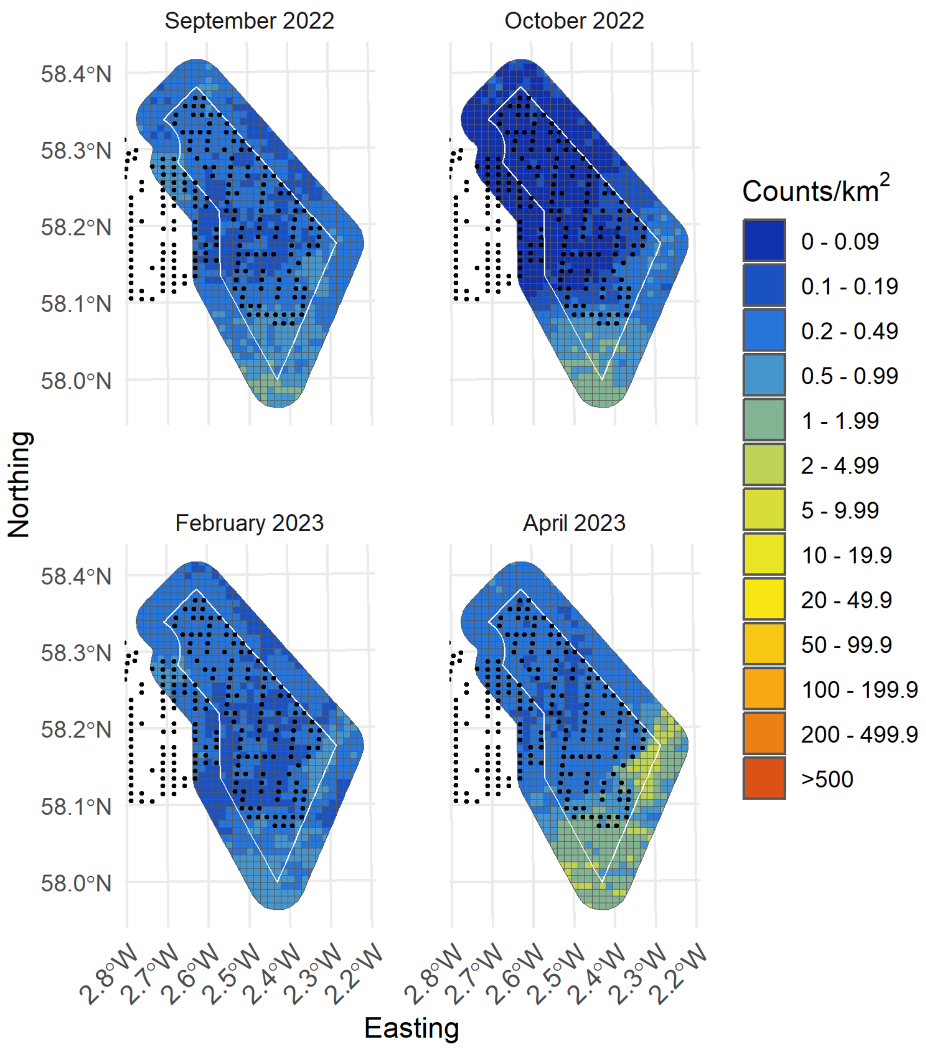


Figure II-18: Random forests predictions from turbine scenario 3 of Kittiwake from September 2022 to April 2023 with turbines presented as black dots

II.2 Fulmar

II.2.1 Model assessment

The top five models had RMSE values between 9.888 and 9.867 with R squared values between 0.144 and 0.141. These values were similar to those from the MRSea model, which had an R-squared value of 0.1437. The best model had an mtry value of 8, with a minimum node size of 80 (Figure II-19, Table II-4).

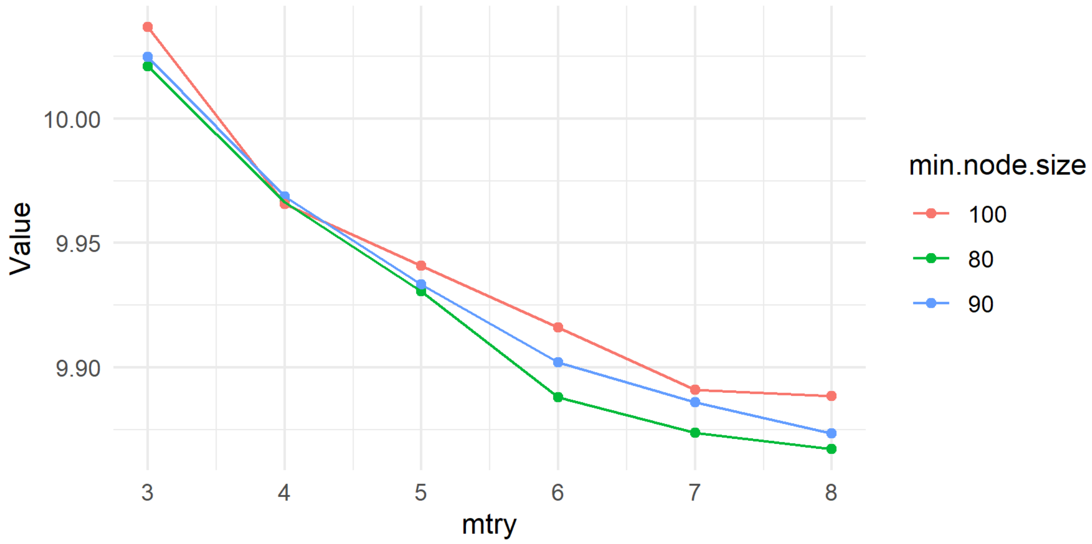


Figure II-19: Root mean squared error when varying mtry and min.node.size parameters in random forests

Table II-4: The top 5 models selected by the random forests analysis showing root mean squared error (RMSE) and R-squared values as calculated by 5-fold cross validation

mtry	min.node.size	RMSE	Rsquared
8	80	9.867	0.1445
8	90	9.873	0.1425
7	80	9.874	0.1435
7	90	9.886	0.1408
6	80	9.888	0.1405

II.2.2 Variable importance

The top predictor variables were monthly mean sea surface temperature, northing (y.pos), and survey ID. This shows that most of the signal in the data come from variables that represent temporal variability (Table II-5).

Table II-5: The top 5 predictor variables from the random forests model and overall importance

Variable	Overall importance
y.pos	100.00
SST_mean	83.20
SurveyID^18	74.27
SurveyID^8	74.09
SurveyID^12	61.67

II.2.3 Distributional response

The partial relationship of distance to turbines on densities of fulmar demonstrates a displacement effect with fewer birds expected in areas close to turbines. The relationship reaches a slight inflection point at approximately 15km, where the curve begins to flatten (Figure II-20). This does not mean that there is a displacement effect out to 15km or more, simply that birds have been predicted further away from turbines, which could be driven by other factors.

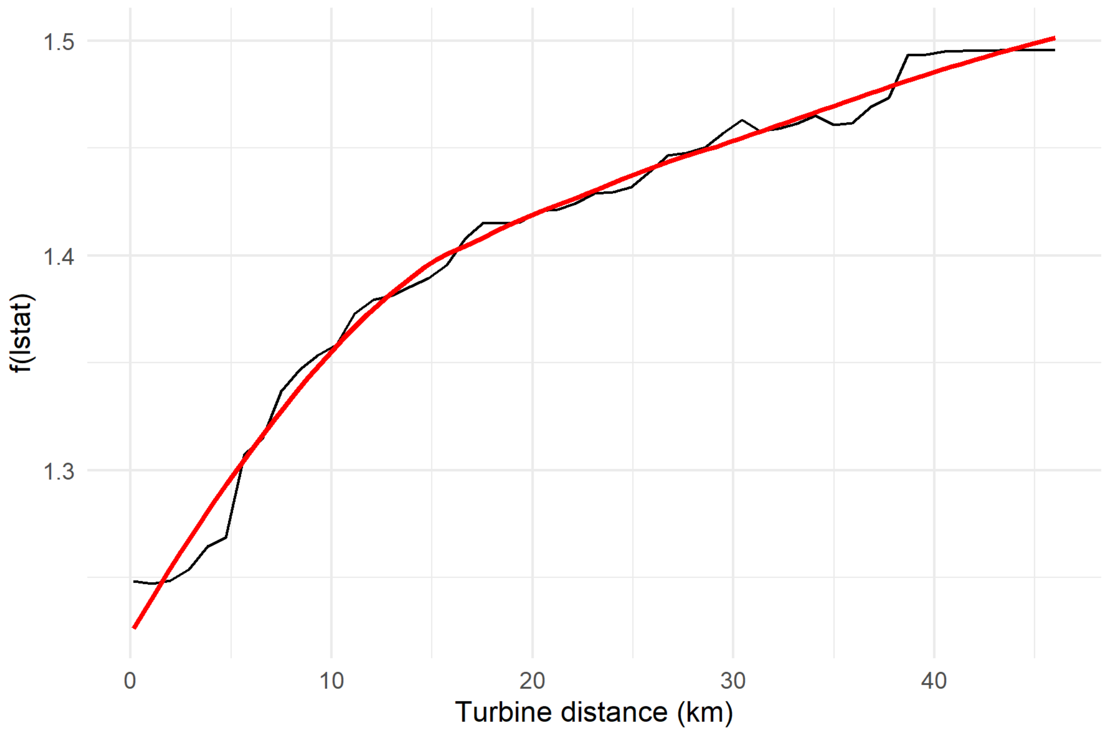


Figure II-20: Partial dependence plot of distance to turbine from the random forests model. The red line depicts the loess curve

II.2.4 Population estimates

Population estimates derived from the random forests baseline scenario (i.e., Existing turbines) fell well within the range of estimates derived from MRSea, sometimes only differing by 2 - 3%. We are therefore confident that the random forests fulmar model is reliable, particularly in the context of the MRSea analysis.

Populations of fulmar in the survey area were predicted to decrease in all three scenarios due to the relationship between distance to turbine and density. The biggest population decrease was predicted for scenario 3, with percent change from baseline varying from -1 to -22% across all surveys (Table II-6).

Table II-6: Population estimates from the random forests models for the 4km buffer zone plus windfarm footprint survey area for the baseline scenario (i.e., based on currently installed turbines) and three potential turbine scenarios. Percent change from baseline is calculated for each scenario

SurveyID	Baseline	Scenario 1 (% change)	Scenario 2 (% change)	Scenario 3 (% change)
May 2021	562	542 (-3.56)	539 (-4.09)	535 (-4.8)
June 2021	675	632 (-6.37)	628 (-6.96)	623 (-7.7)
July 2021	5121	4974 (-2.87)	4943 (-3.48)	4907 (-4.18)
August 2021	1723	1740 (0.99)	1732 (0.52)	1738 (0.87)
September 2021	671	603 (-10.13)	603 (-10.13)	598 (-10.88)
November 2021	964	793 (-17.74)	774 (-19.71)	766 (-20.54)
December 2021	1137	919 (-19.17)	906 (-20.32)	885 (-22.16)
January 2022	409	375 (-8.31)	366 (-10.51)	349 (-14.67)
February 2022	513	451 (-12.09)	447 (-12.87)	434 (-15.4)
March 2022	186	193 (3.76)	195 (4.84)	195 (4.84)
April 2022	396	376 (-5.05)	372 (-6.06)	369 (-6.82)
May 2022	340	323 (-5)	325 (-4.41)	323 (-5)
June 2022	1240	1225 (-1.21)	1231 (-0.73)	1215 (-2.02)
July 2022	9756	9052 (-7.22)	8772 (-10.09)	8596 (-11.89)
August 2022	716	694 (-3.07)	693 (-3.21)	690 (-3.63)
September 2022	301	287 (-4.65)	286 (-4.98)	283 (-5.98)
November 2022	1591	1648 (3.58)	1645 (3.39)	1647 (3.52)
December 2022	573	506 (-11.69)	503 (-12.22)	496 (-13.44)
January 2023	330	305 (-7.58)	304 (-7.88)	302 (-8.48)
February 2023	231	230 (-0.43)	230 (-0.43)	230 (-0.43)
March 2023	593	538 (-9.27)	535 (-9.78)	515 (-13.15)
April 2023	257	250 (-2.72)	251 (-2.33)	250 (-2.72)

II.2.5 Distributions

The broad distribution of fulmar in the baseline random forests models were nearly identical to those in the MRSea models. For fulmar, the models predict the highest densities in the southwest part of the survey area for most months. In all cases, the densities of fulmar decrease through the site in relation to the locations of proposed turbines. This distributional response suggests a potential displacement effect which could push beyond the boundary of the wind farm site (Figures II-21 – II-40).

II.2.5.1 Baseline scenarios

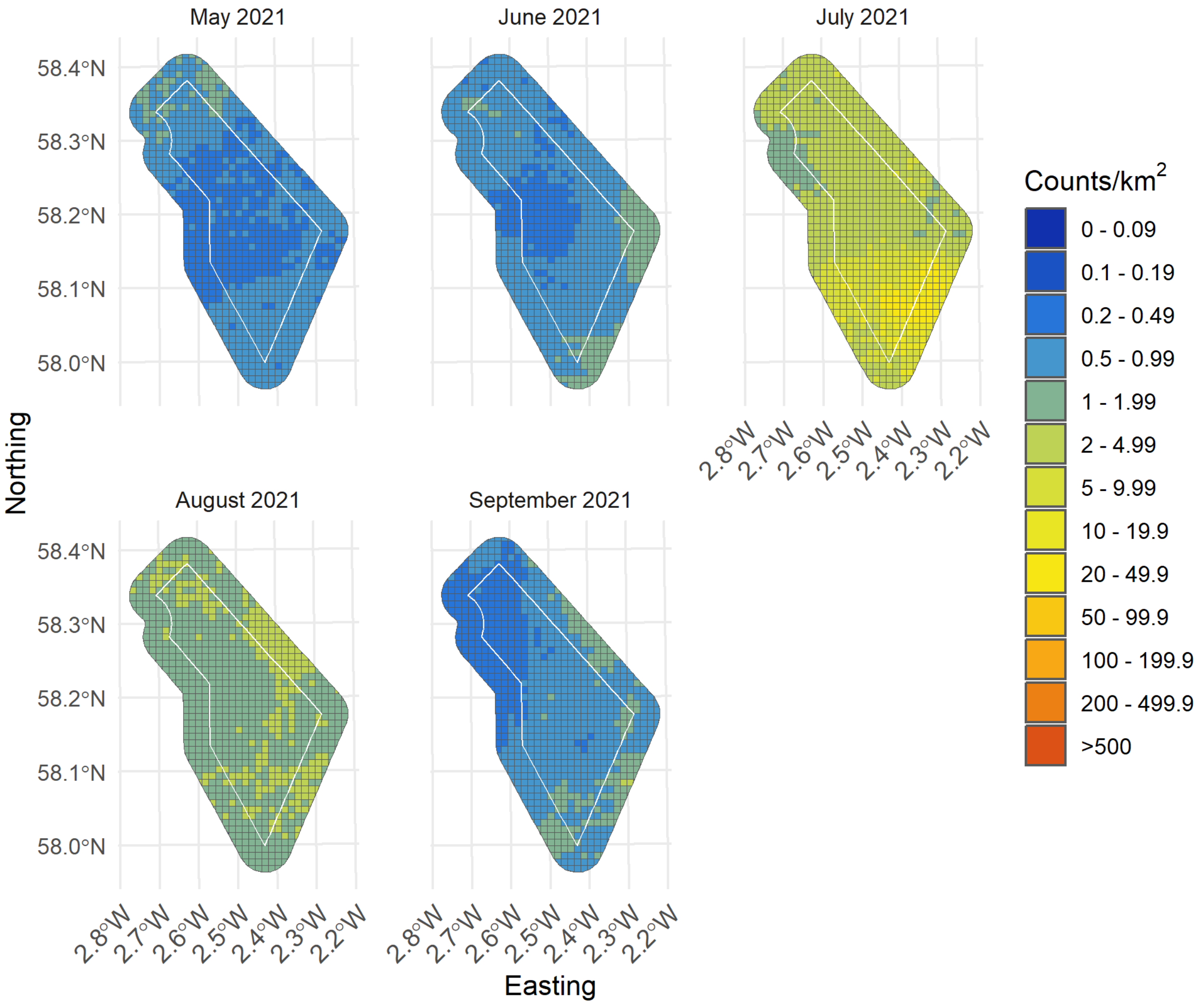


Figure II-21: Random forests baseline predictions of Fulmar from May 2021 to September 2021

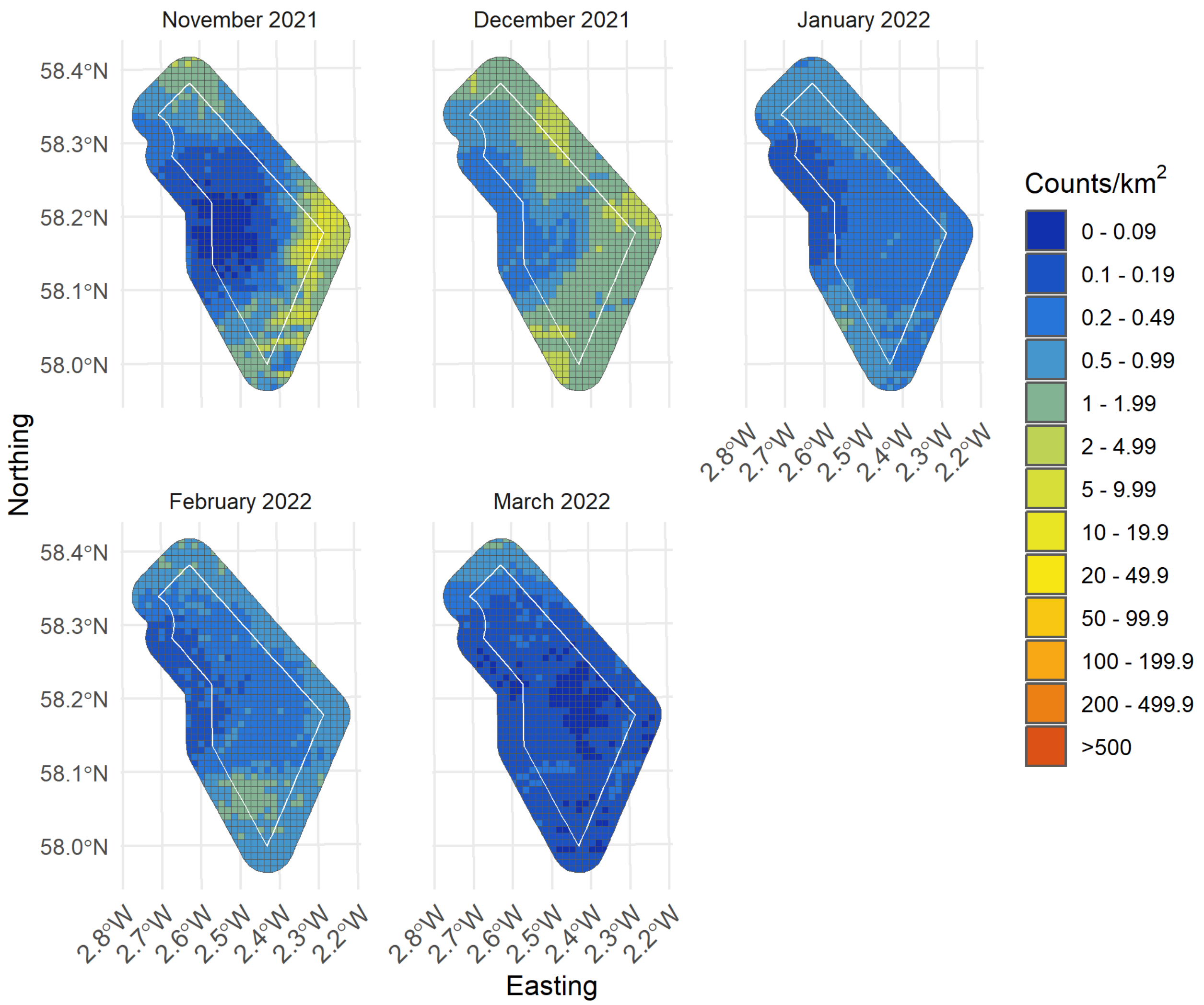


Figure II-22: Random forests baseline predictions of Fulmar from November 2021 to March 2022

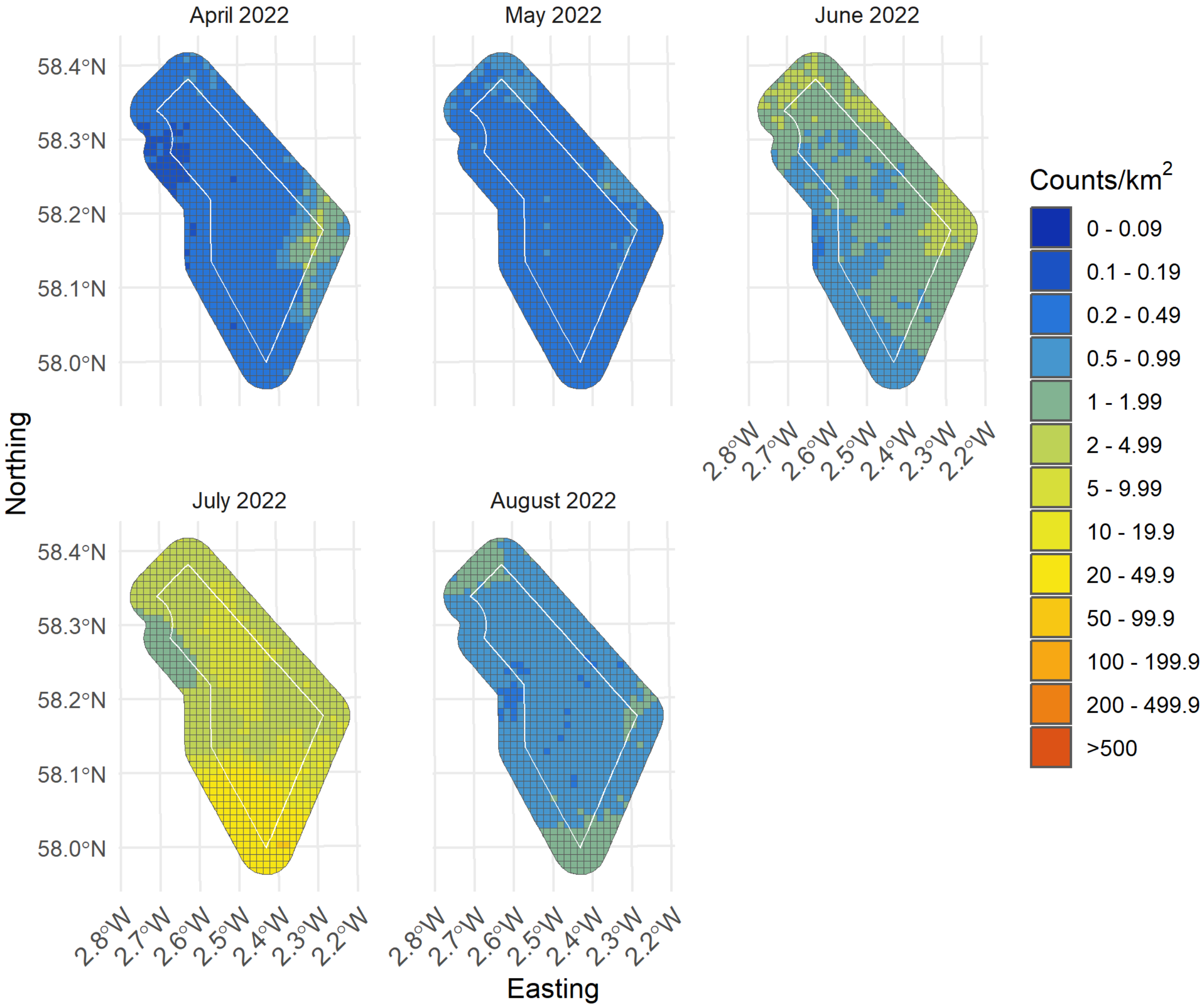


Figure II-23: Random forests baseline predictions of Fulmar from April 2022 to August 2022

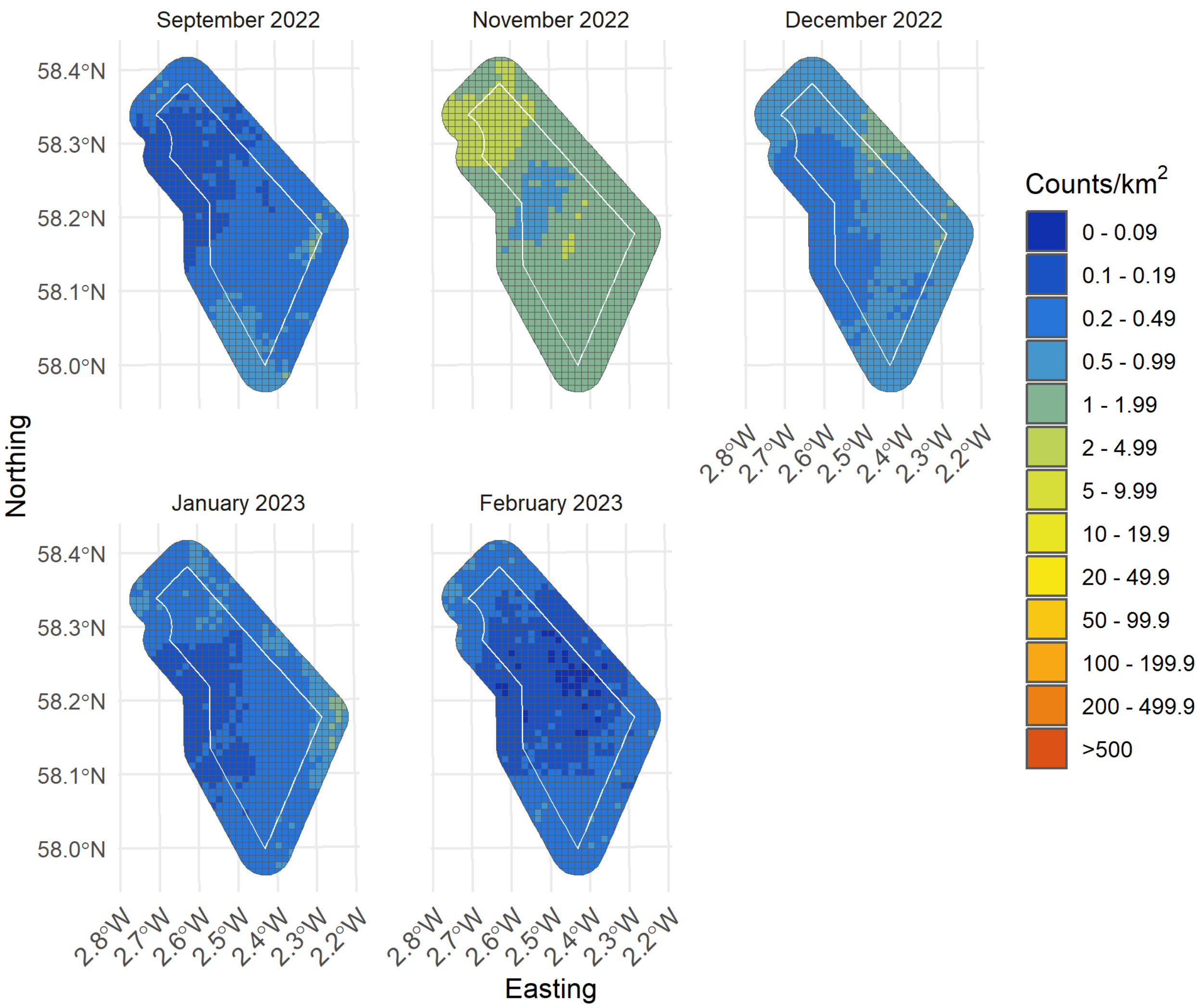


Figure II-24: Random forests baseline predictions of Fulmar from September 2022 to February 2023

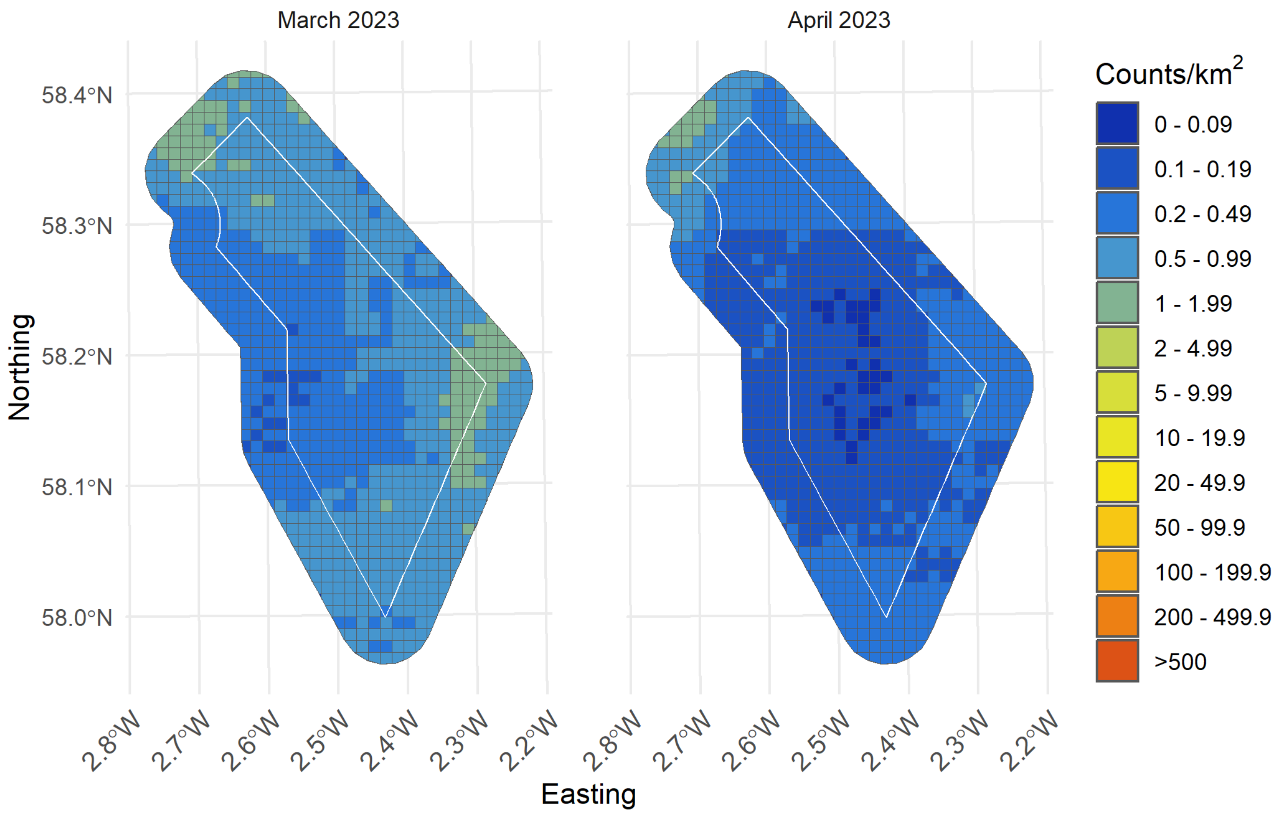


Figure II-25: Random forests baseline predictions of Fulmar from March 2023 to April 2023

II.2.5.2 Turbine scenario 1

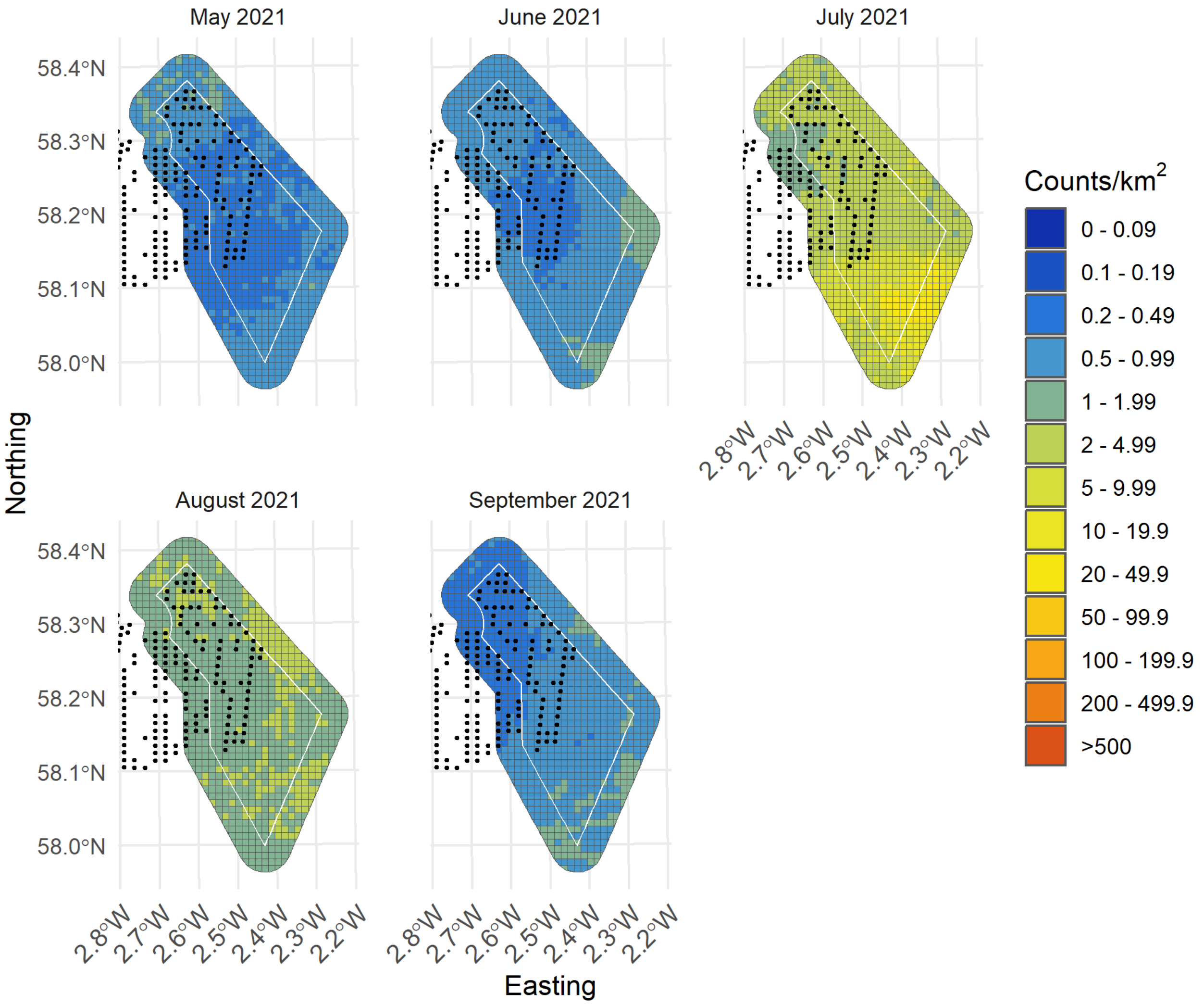


Figure II-26: Random forests predictions from turbine scenario 1 of Fulmar from May 2021 to September 2021 with turbines presented as black dots

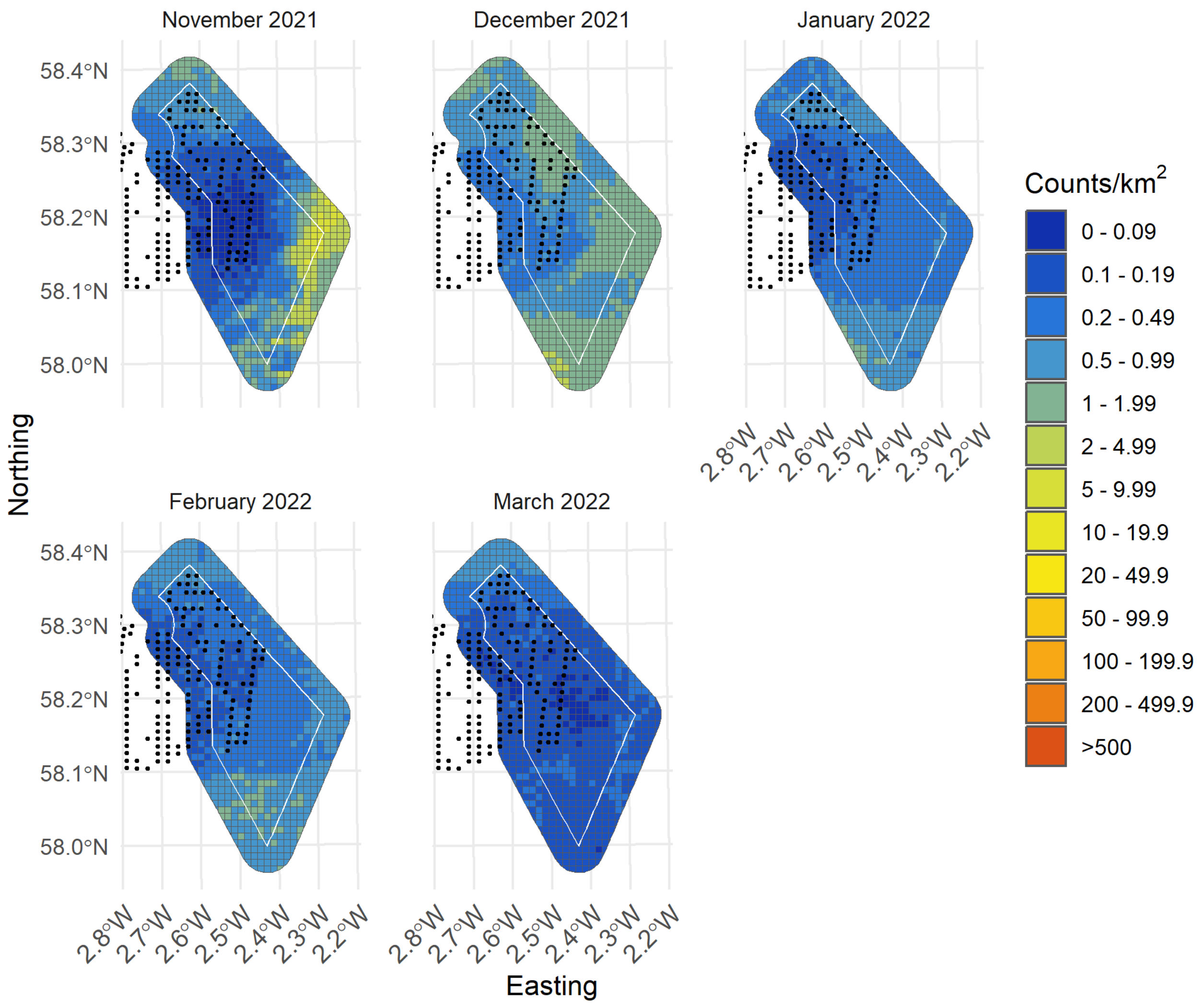


Figure II-27: Random forests predictions from turbine scenario 1 of Fulmar from November 2021 to March 2022 with turbines presented as black dots

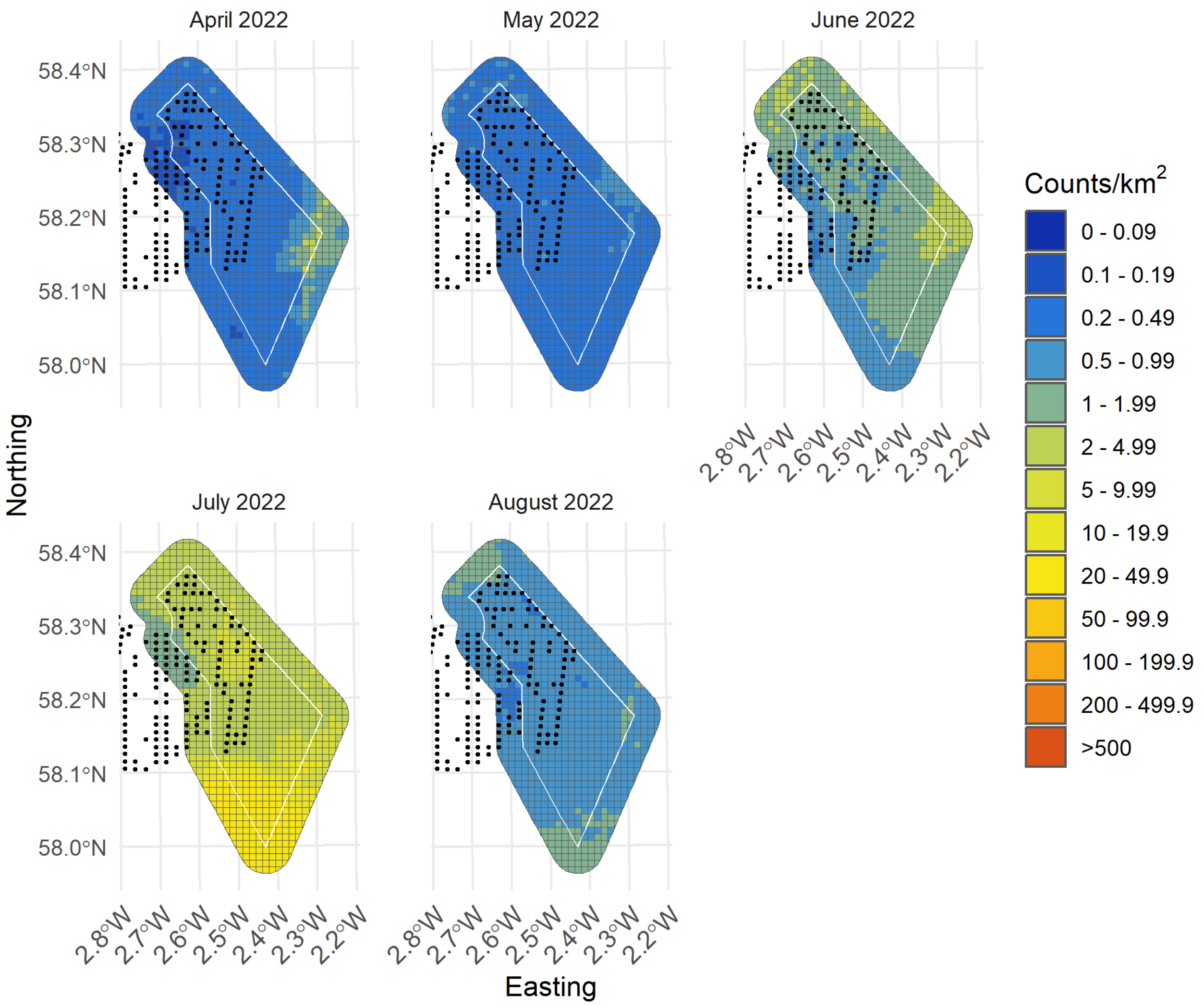


Figure II-28: Random forests predictions from turbine scenario 1 of Fulmar from April 2022 to August 2022 with turbines presented as black dots

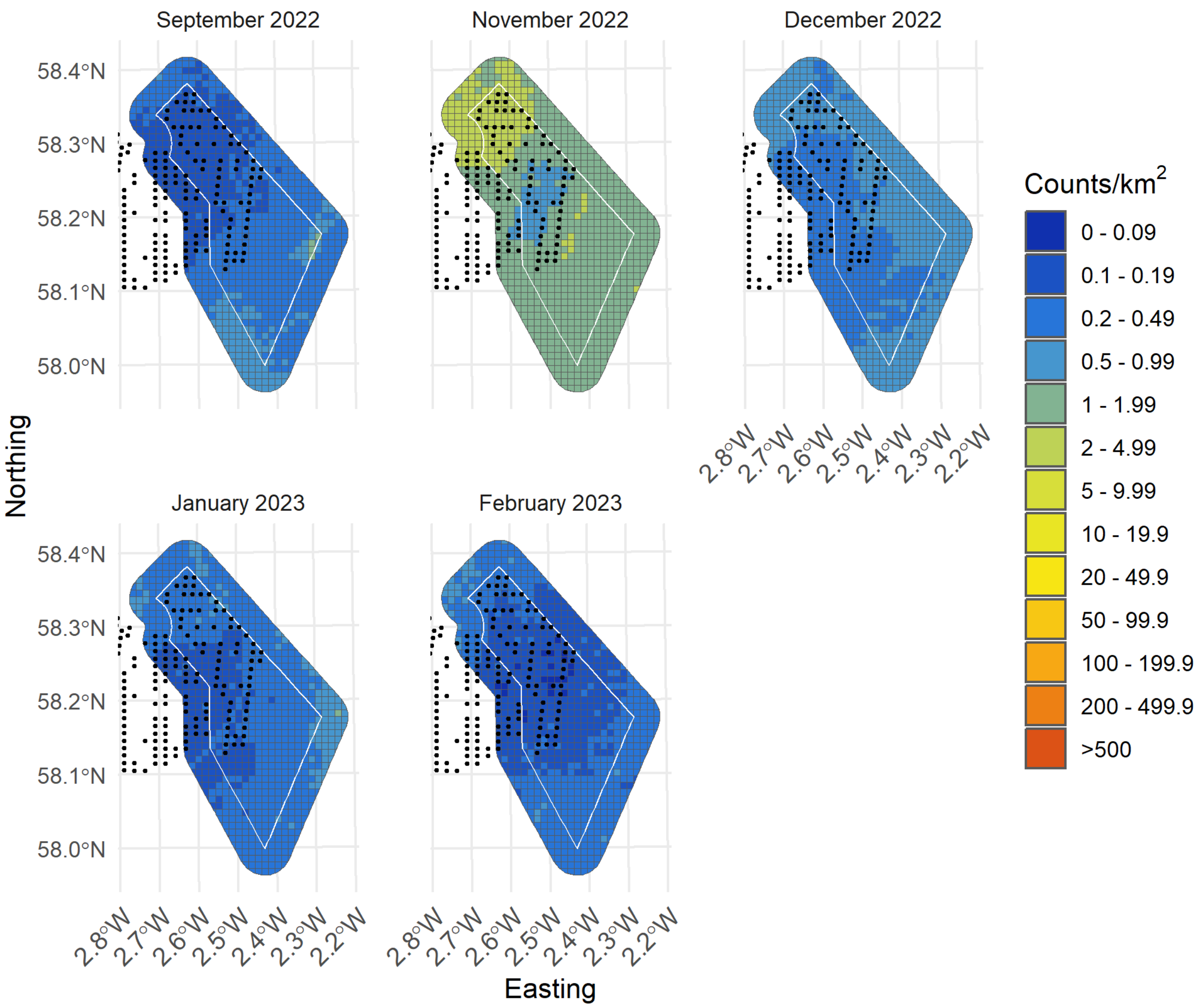


Figure II-29: Random forests predictions from turbine scenario 1 of Fulmar from September 2022 to February 2023 with turbines presented as black dots

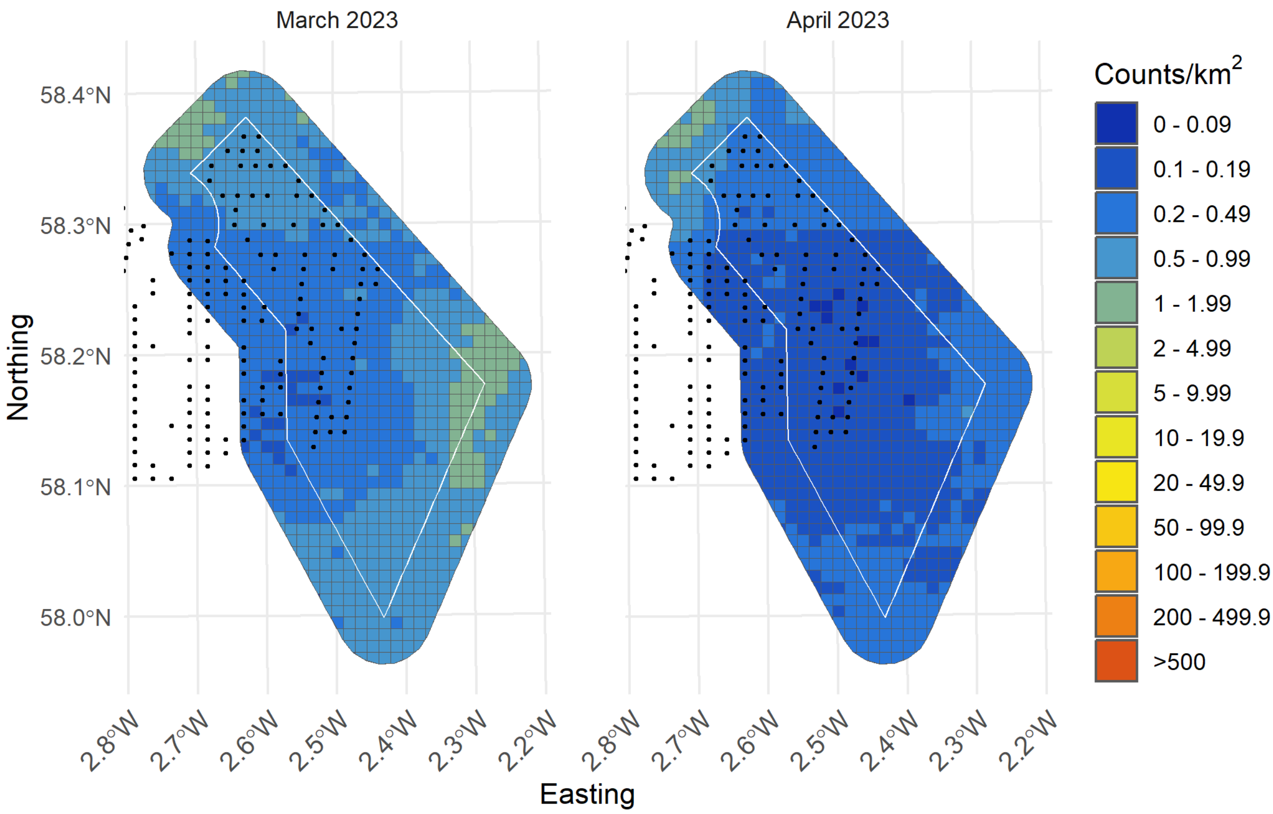


Figure II-30: Random forests predictions from turbine scenario 1 of Fulmar from March 2023 to April 2023 with turbines presented as black dots

II.2.5.3 Turbine scenario 2

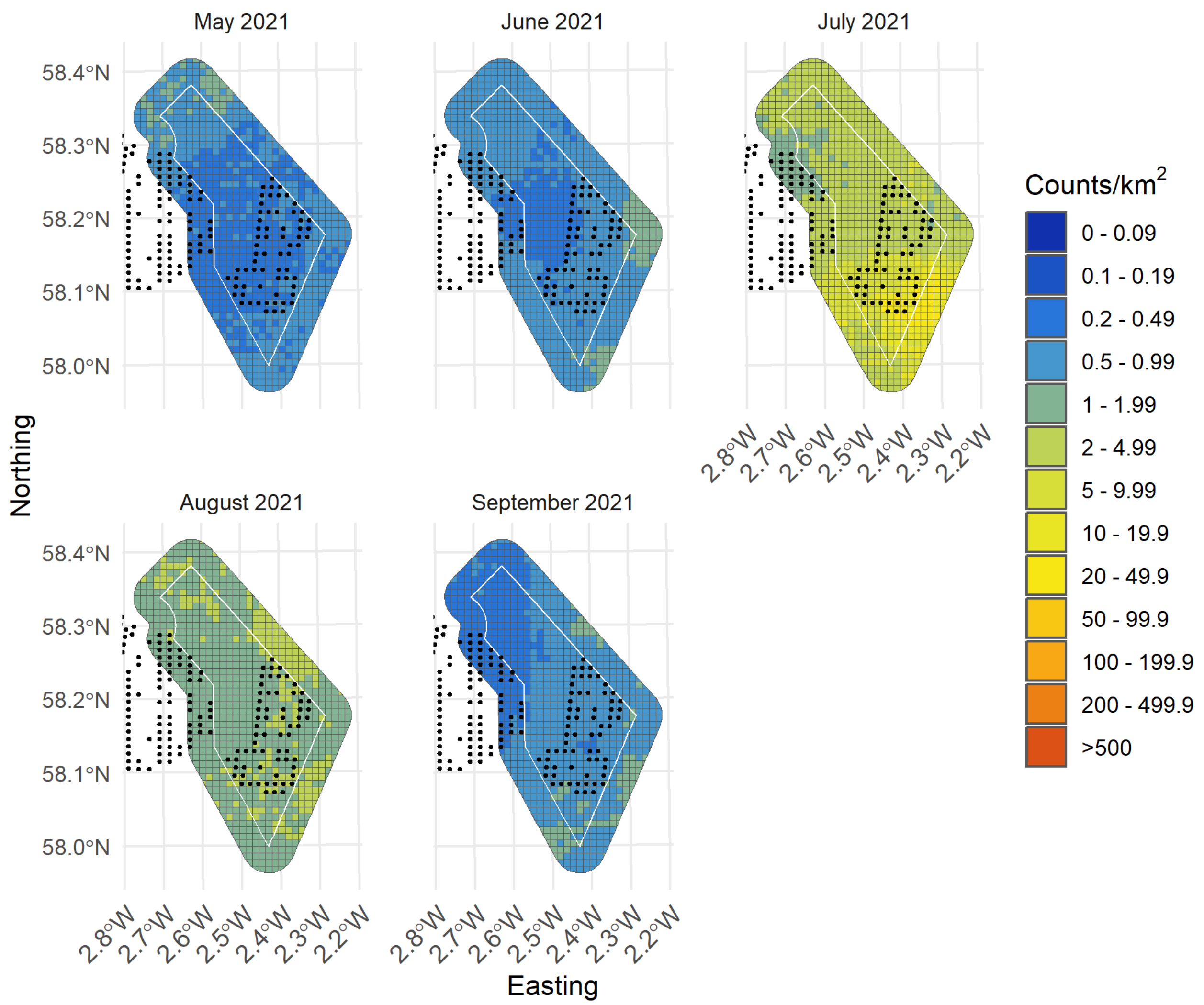


Figure II-31: Random forests predictions from turbine scenario 2 of Fulmar from May 2021 to September 2021 with turbines presented as black dots

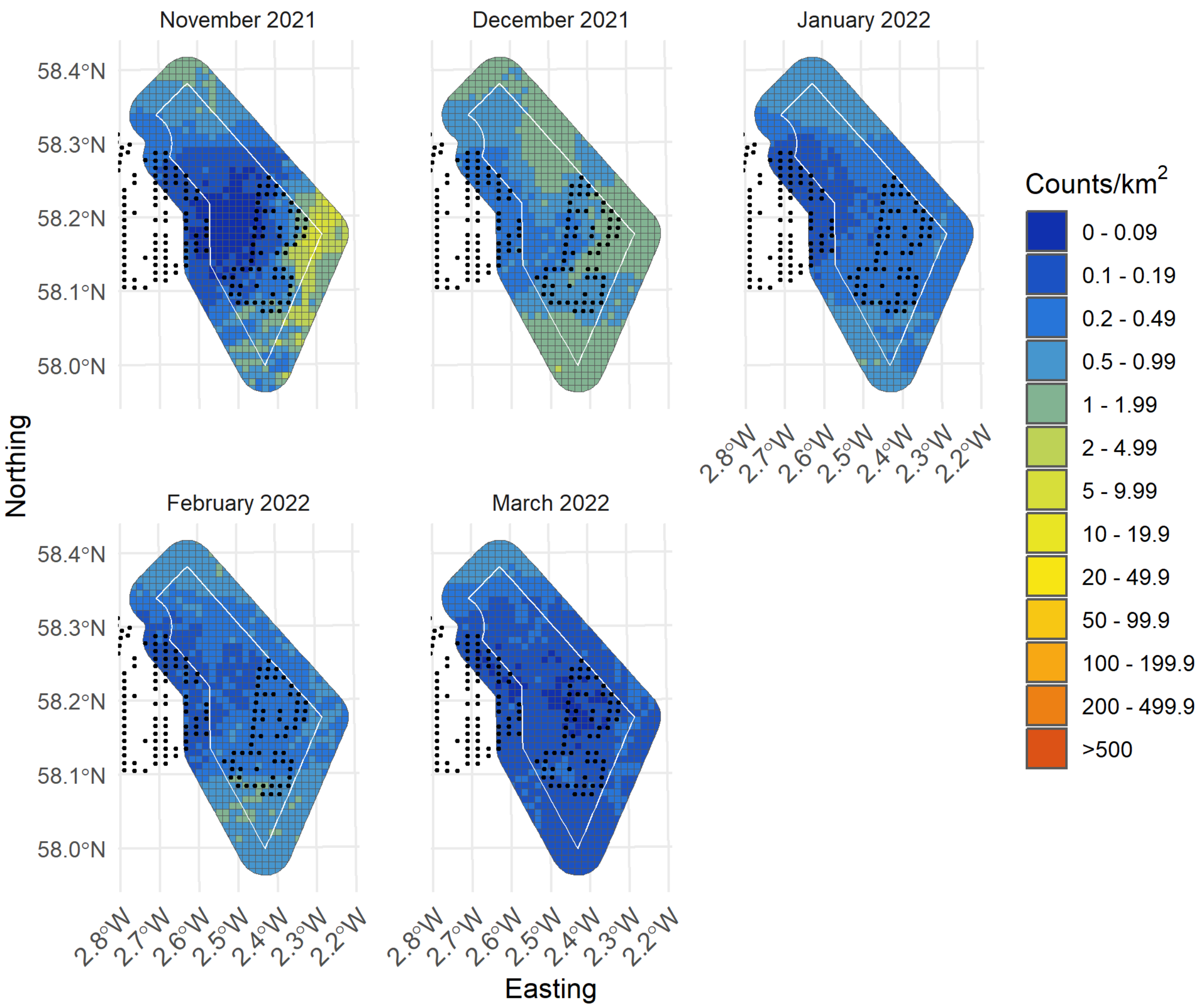


Figure II-32: Random forests predictions from turbine scenario 2 of Fulmar from November 2021 to March 2022 with turbines presented as black dots

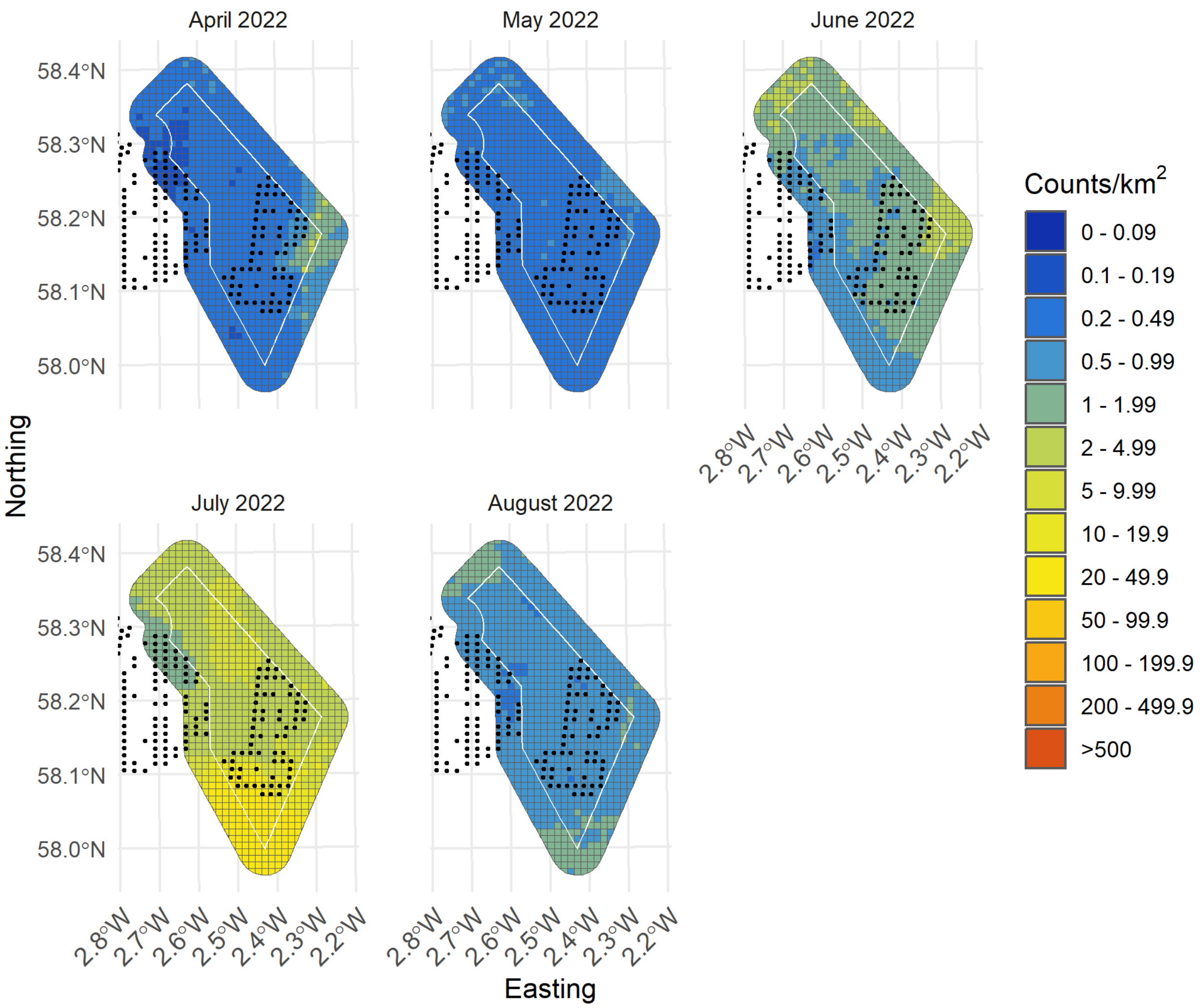


Figure II-33: Random forests predictions from turbine scenario 2 of Fulmar from April 2022 to August 2022 with turbines presented as black dots

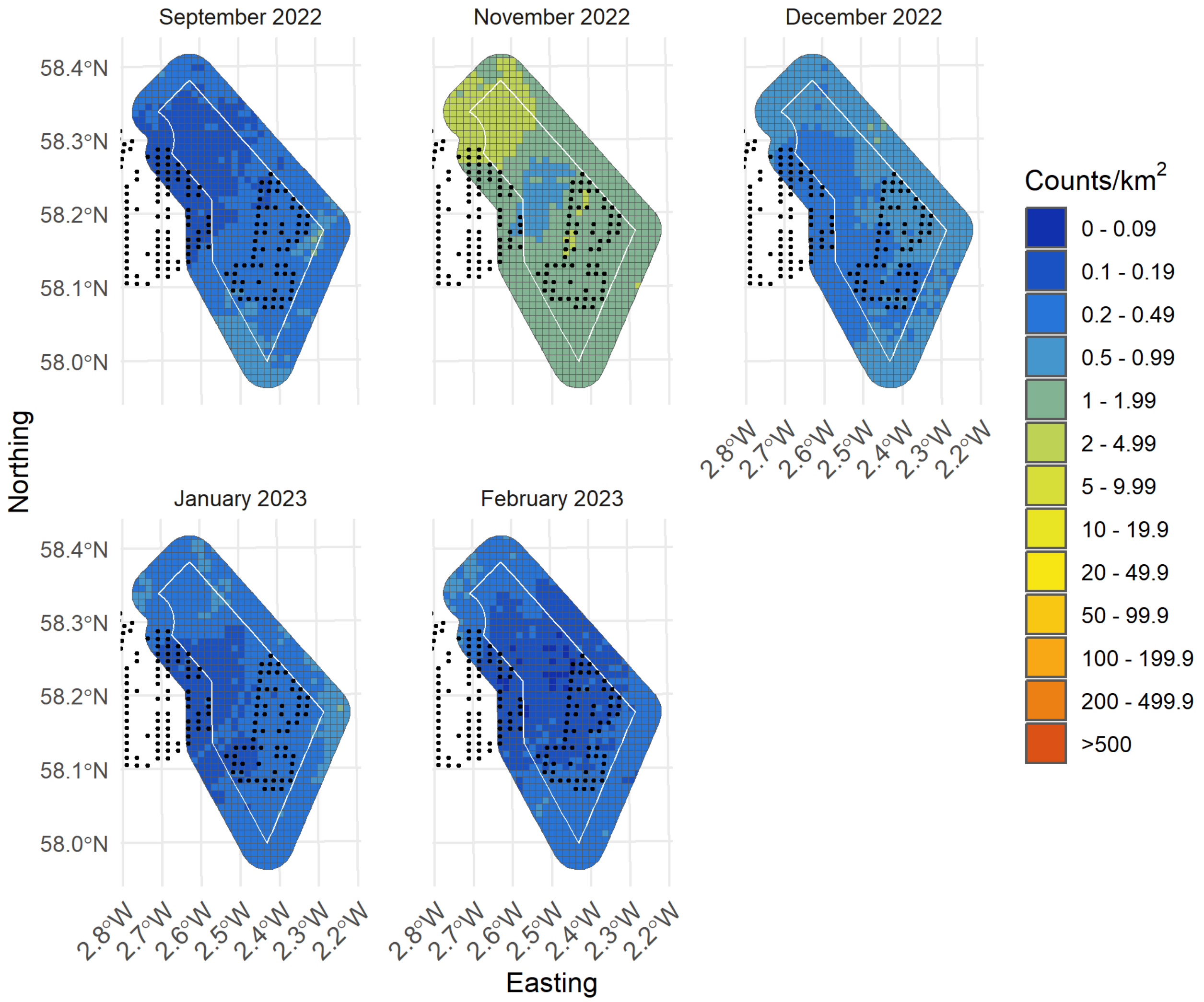


Figure II-34: Random forests predictions from turbine scenario 2 of Fulmar from September 2022 to February 2023 with turbines presented as black dots

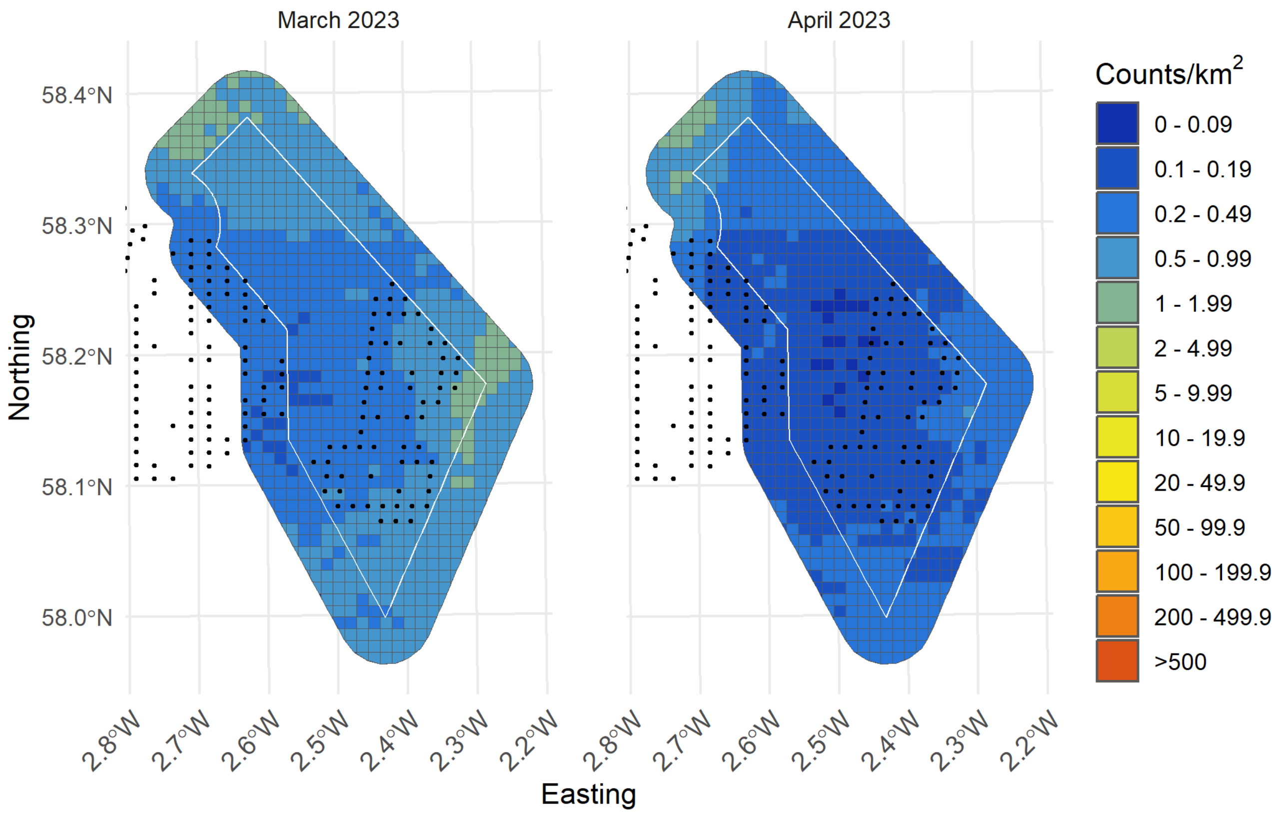


Figure II-35: Random forests predictions from turbine scenario 2 of Fulmar from March 2023 to April 2023 with turbines presented as black dots

II.2.5.4 Turbine scenario 3

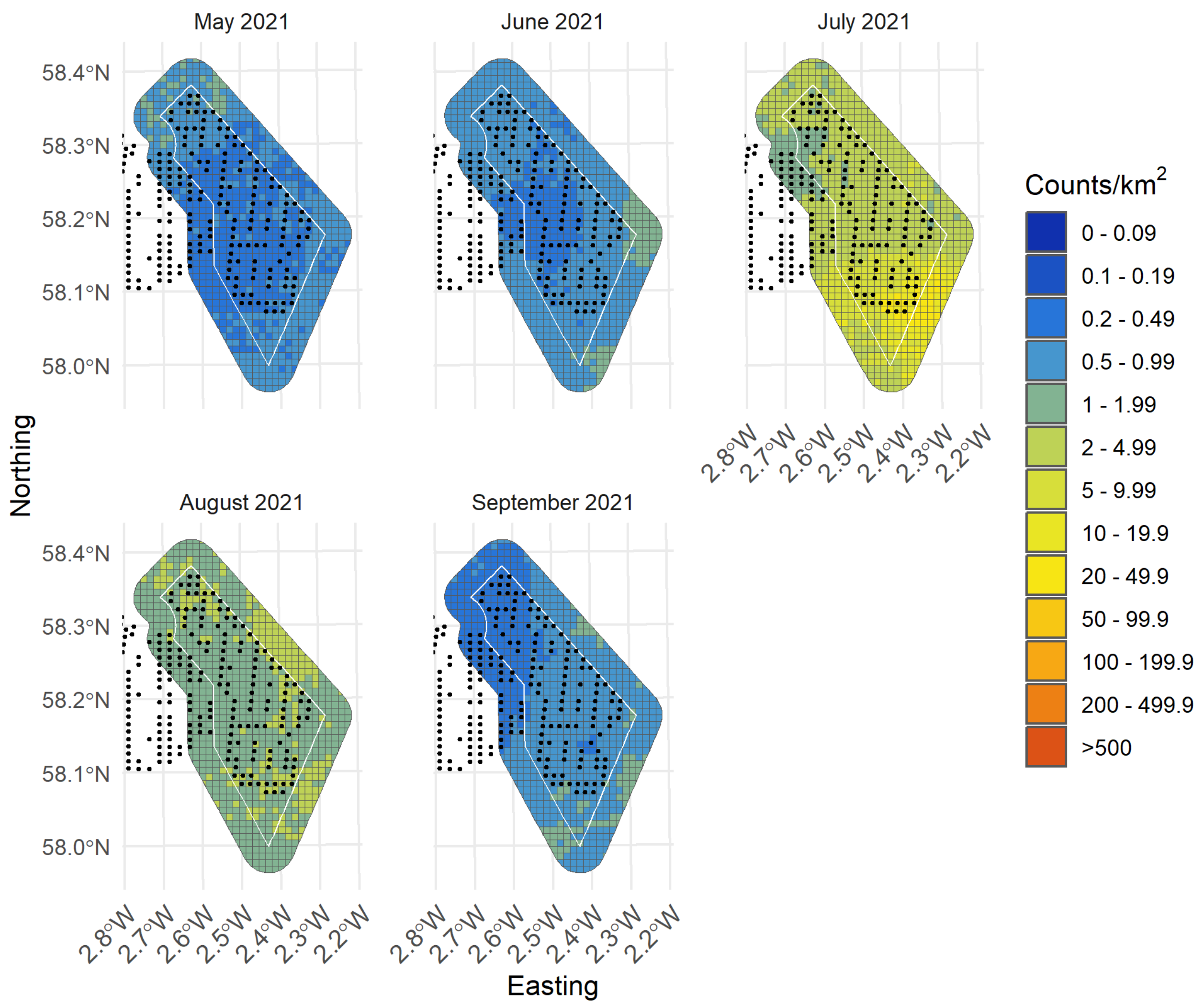


Figure II-36: Random forests predictions from turbine scenario 3 of Fulmar from May 2021 to September 2021 with turbines presented as black dots

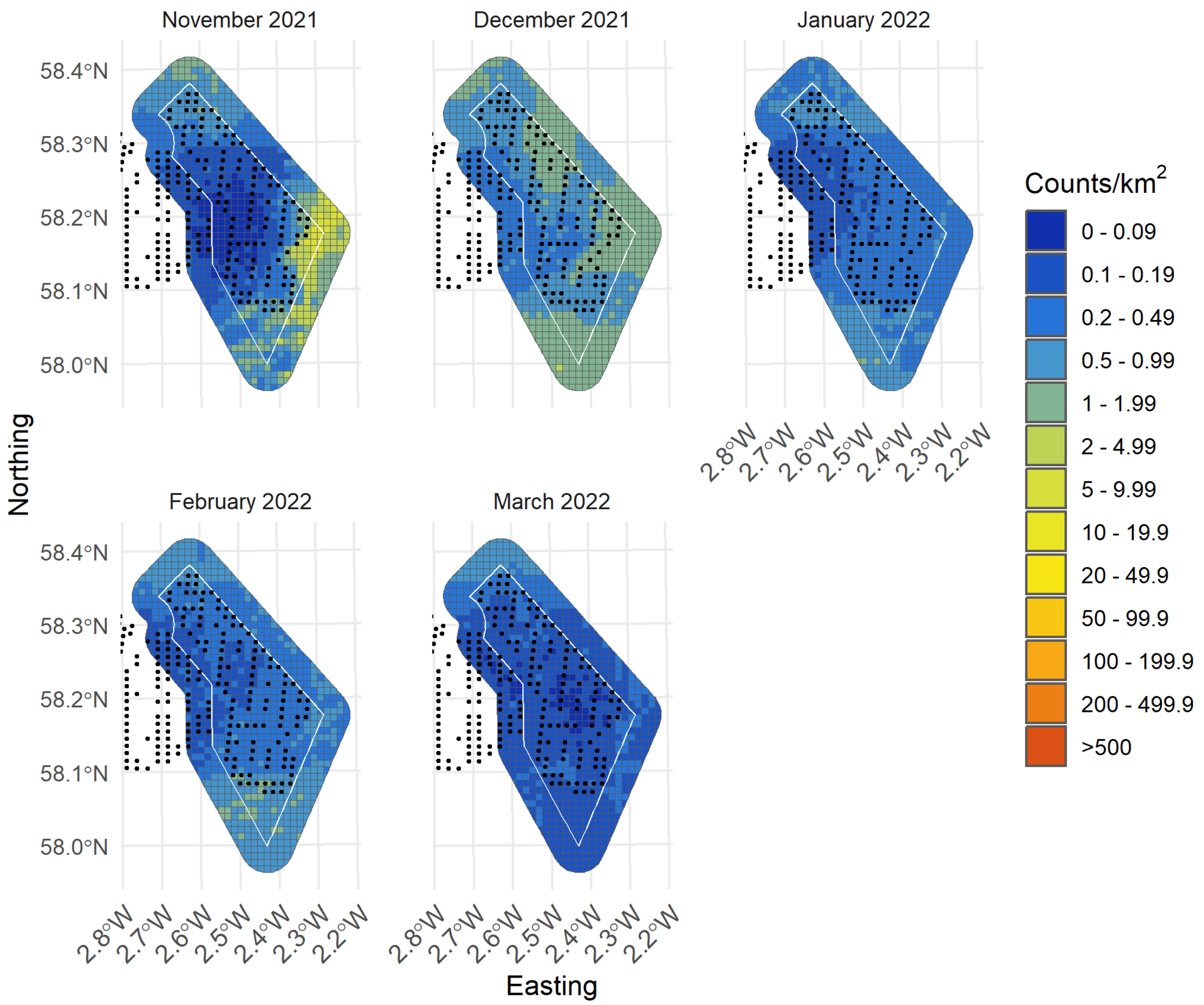


Figure II-37: Random forests predictions from turbine scenario 3 of Fulmar from November 2021 to March 2022 with turbines presented as black dots

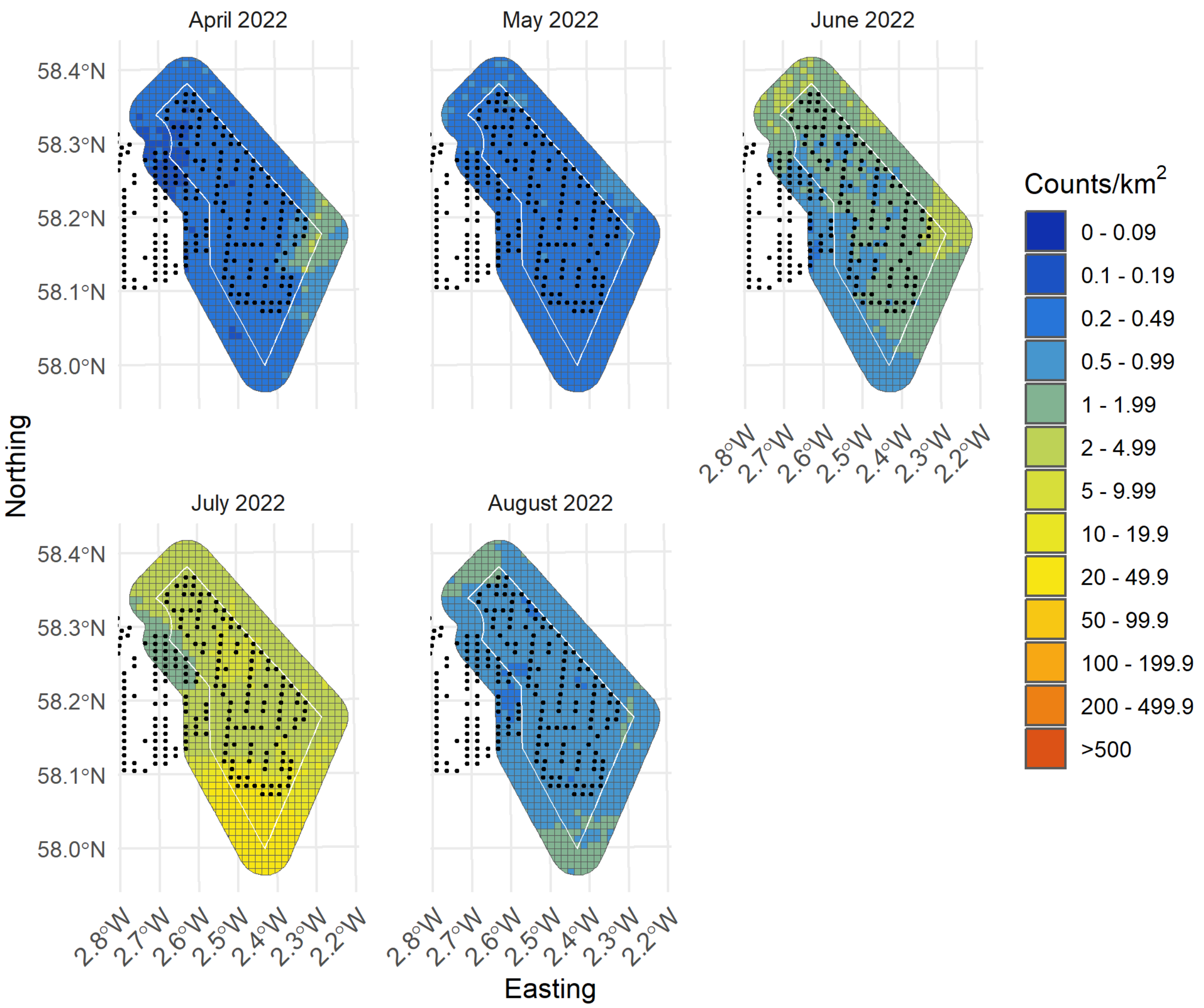


Figure II-38: Random forests predictions from turbine scenario 3 of Fulmar from April 2022 to August 2022 with turbines presented as black dots

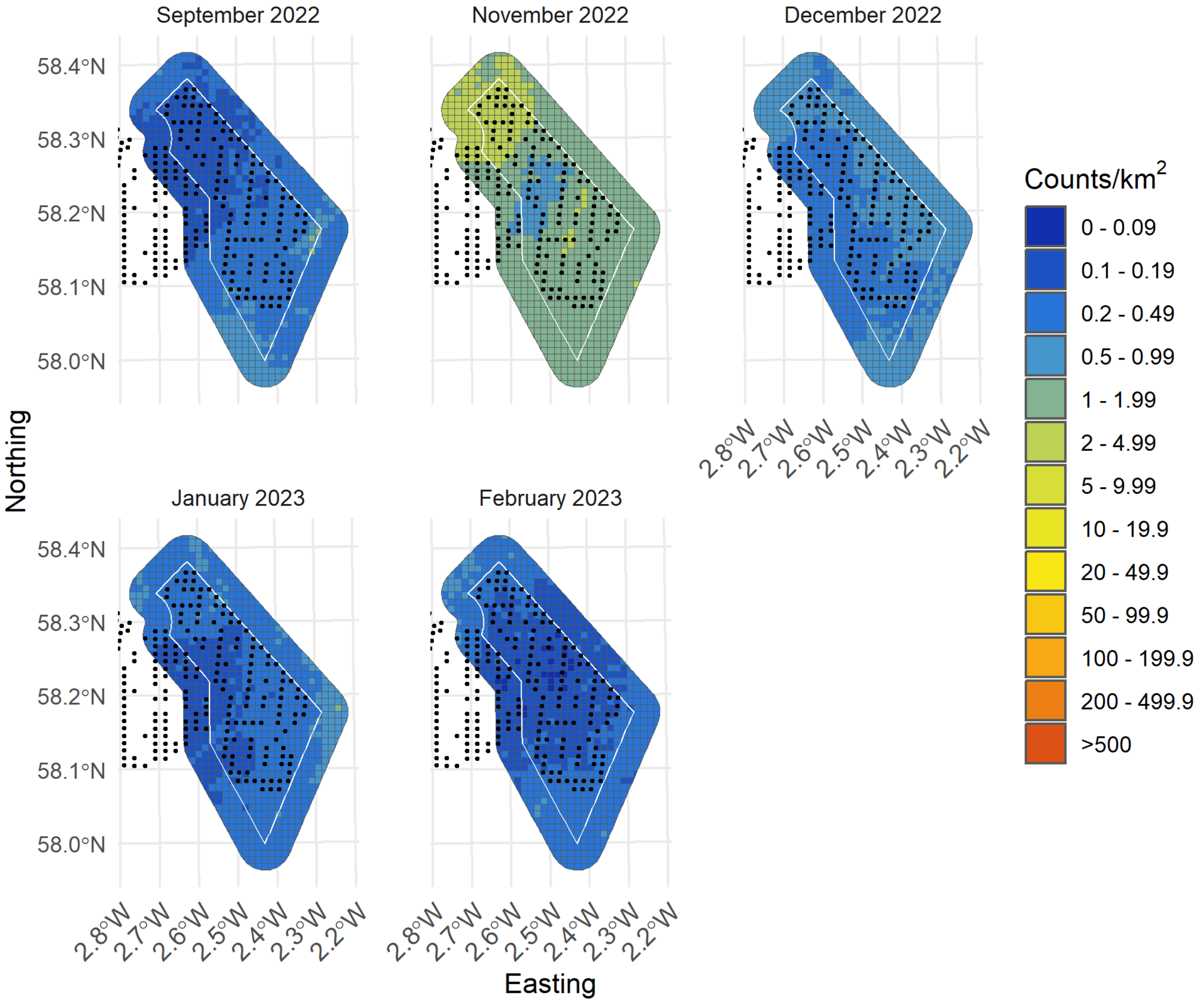


Figure II-39: Random forests predictions from turbine scenario 3 of Fulmar from September 2022 to February 2023 with turbines presented as black dots

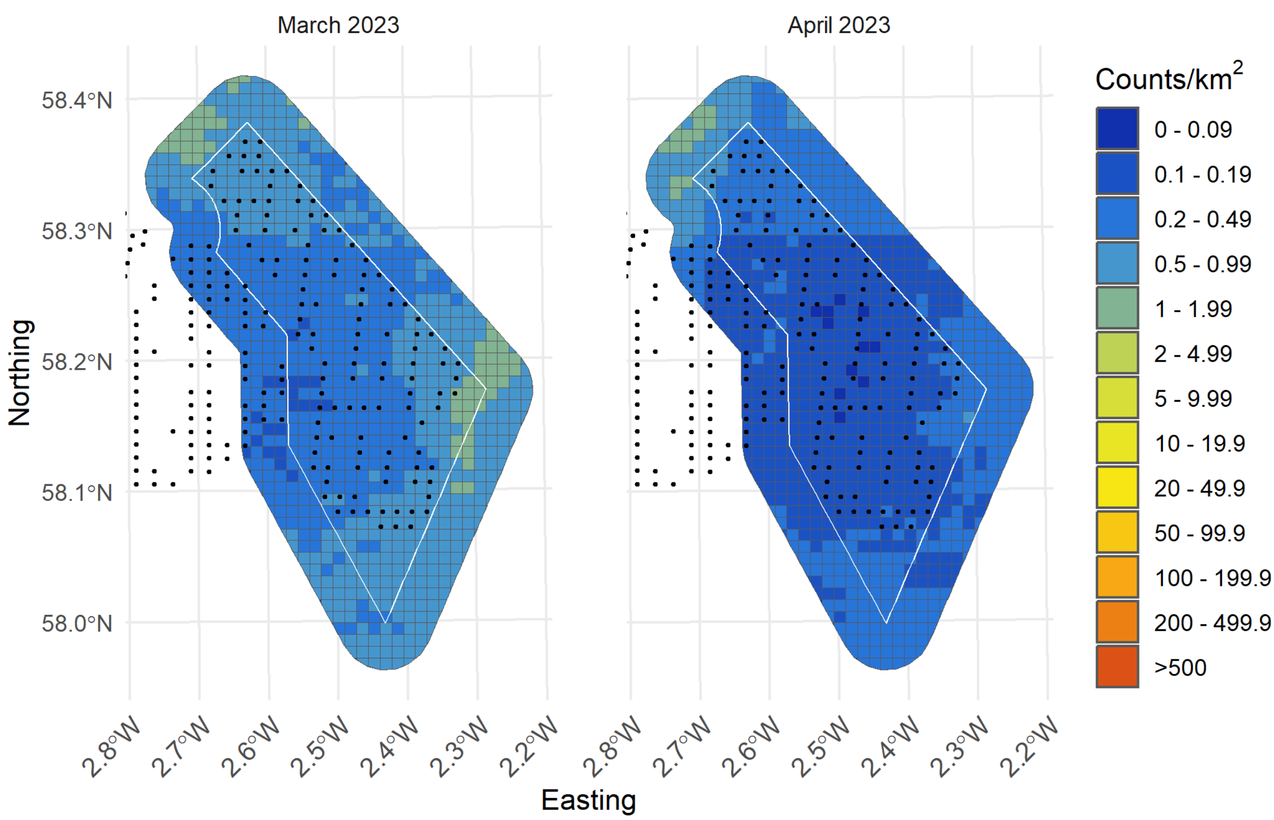


Figure II-40: Random forests predictions from turbine scenario 3 of Fulmar from March 2023 to April 2023 with turbines presented as black dots

II.3 Gannet

II.3.1 Model assessment

The top five models had RMSE values that were all around 3.835 with R squared values between 0.0442 and 0.0448. These R squared values were better than those from the MRSea model, which had an R-squared value of 0.0395. The best model had an mtry value of 6, with a minimum node size of 80 (Figure II-41, Table II-7).

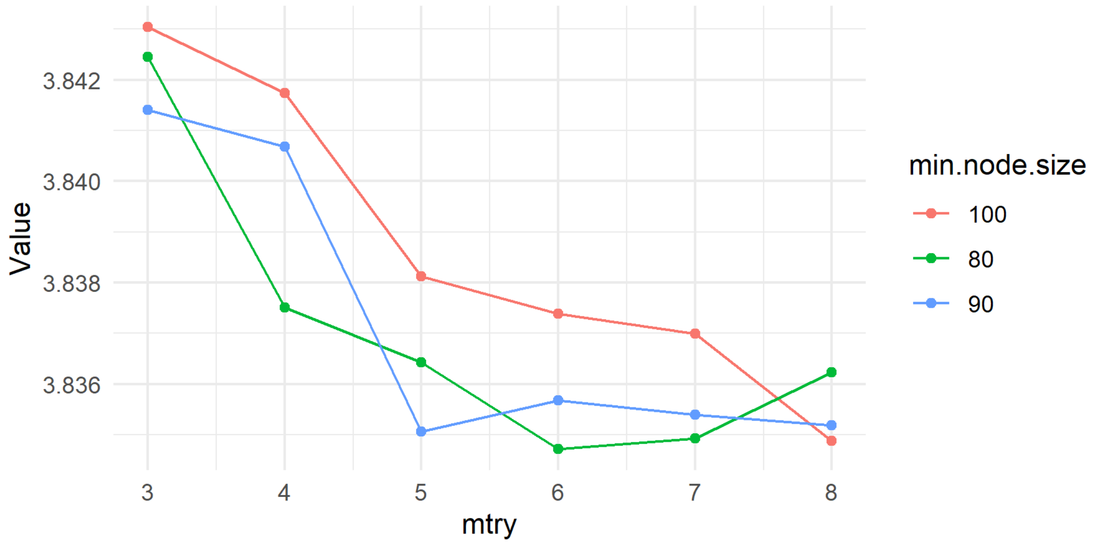


Figure II-41: Root mean squared error when varying mtry and min.node.size parameters in random forests

Table II-7: The top 5 models selected by the random forests analysis showing root mean squared error (RMSE) and R-squared values as calculated by 5-fold cross validation

mtry	min.node.size	RMSE	Rsquared
6	80	3.835	0.0443
8	100	3.835	0.0448
7	80	3.835	0.0445
5	90	3.835	0.0442
8	90	3.835	0.0442

II.3.2 Variable importance

The top predictor variables were distance to colony, density of sandeel, northing (y.pos), bathymetry and presence of sandeel. This shows that most of the signal in the data come from variables that represent ecological parameters related to prey (Table II-8).

Table II-8: The top 5 predictor variables from the random forests model and overall importance

Variable	Overall importance
dist2col	100.00
sandeel_pr_density	93.43
y.pos	85.49
Bathymetry	71.91
sandeel_pr_presence	69.76

II.3.3 Distributional response

The partial relationship of distance to turbines on densities of gannet demonstrates that there was an inverse relationship between density and distance to turbine from 0 to approximately 18km away from the turbines. After this, the curve has an inflection point where the relationship reverses, and more birds are predicted further away. (Figure II-42). The relationship between 0 - 18km only contributes to a partial dependence of approximately 0.03 and when comparing to the overall scale of the figure of 0.64 to 0.72, it suggests little evidence of a displacement effect. However, this relationship could be driven by other environmental covariates that we were unable to measure.

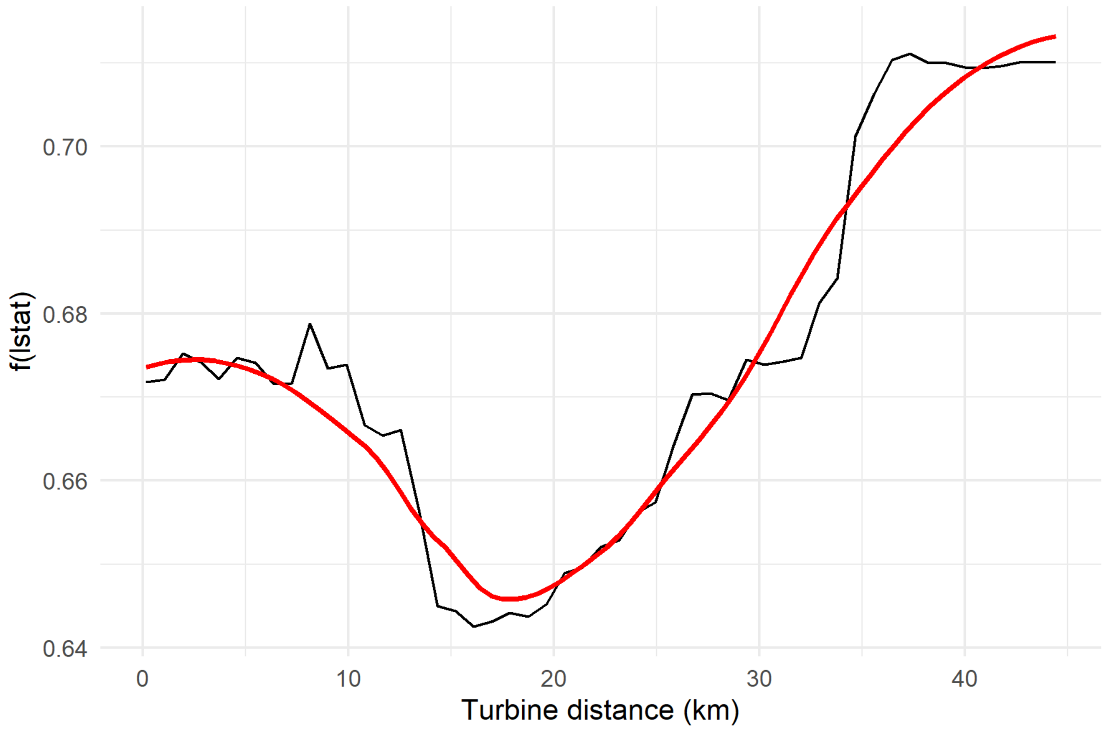


Figure II-42: Partial dependence plot of distance to turbine from the random forests model. The red line depicts the loess curve

II.3.4 Population estimates

Population estimates derived from the random forests baseline scenario (i.e., Existing turbines) fell well within the range of estimates derived from MRSea, sometimes only differing by 0.6 to 2%. We are therefore confident that the random forests gannet model is reliable, particularly in the context of the MRSea analysis.

There was no consistency in population increases or decreases between turbine scenarios vesus baseline. However, the largest variations were in scenario 3 (ranging from -9.75% to 19.94%) (Table II-9).

Table II-9: Population estimates from the random forests models for the 4km buffer zone plus windfarm footprint survey area for the baseline scenario (i.e., based on currently installed turbines) and three potential turbine scenarios. Percent change from baseline is calculated for each scenario

SurveyID	Baseline	Scenario 1 (% change)	Scenario 2 (% change)	Scenario 3 (% change)
June 2021	688	761 (10.61)	756 (9.88)	757 (10.03)
July 2021	293	275 (-6.14)	273 (-6.83)	276 (-5.8)
August 2021	455	477 (4.84)	470 (3.3)	477 (4.84)
September 2021	632	739 (16.93)	729 (15.35)	758 (19.94)
October 2021	790	721 (-8.73)	738 (-6.58)	713 (-9.75)
May 2022	147	154 (4.76)	152 (3.4)	155 (5.44)
June 2022	1465	1456 (-0.61)	1440 (-1.71)	1445 (-1.37)
July 2022	730	767 (5.07)	758 (3.84)	760 (4.11)
September 2022	655	728 (11.15)	720 (9.92)	727 (10.99)
October 2022	276	273 (-1.09)	274 (-0.72)	274 (-0.72)

II.3.5 Distributions

The broad distribution of gannet in the baseline random forests models were nearly identical to those in the MRSea models. For gannet, the models predict the highest densities in the southern part of the survey area for most months (with the exception of October 2021). The densities of gannet change variably between surveys in relation to the locations of proposed turbines. This distributional response suggests a variable displacement effect (Figures II-43 – II-50).

II.3.5.1 Baseline scenarios

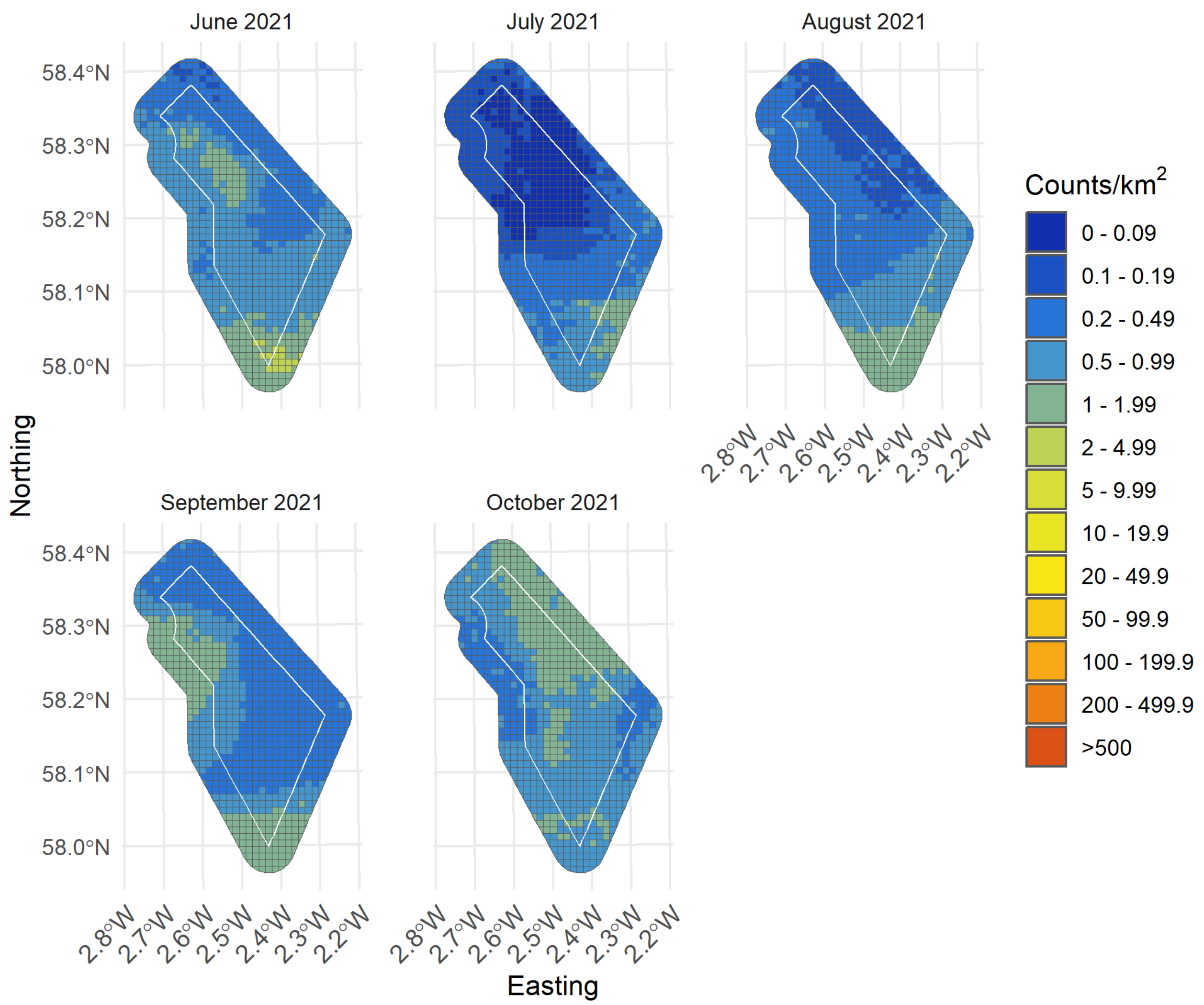


Figure II-43: Random forests baseline predictions of Gannet from June 2021 to October 2021

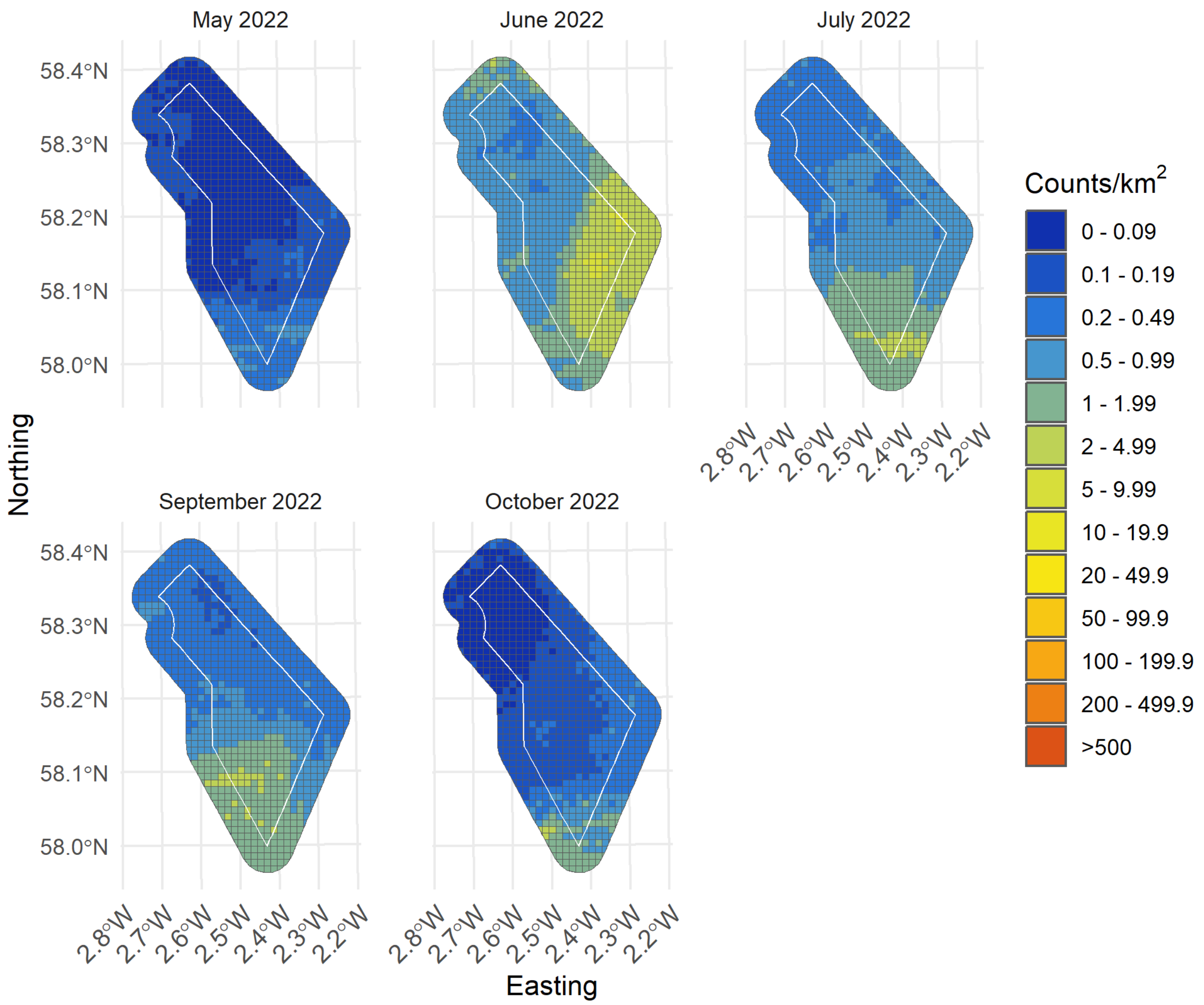


Figure II-44: Random forests baseline predictions of Gannet from May 2022 to October 2022

II.3.5.2 Turbine scenario 1

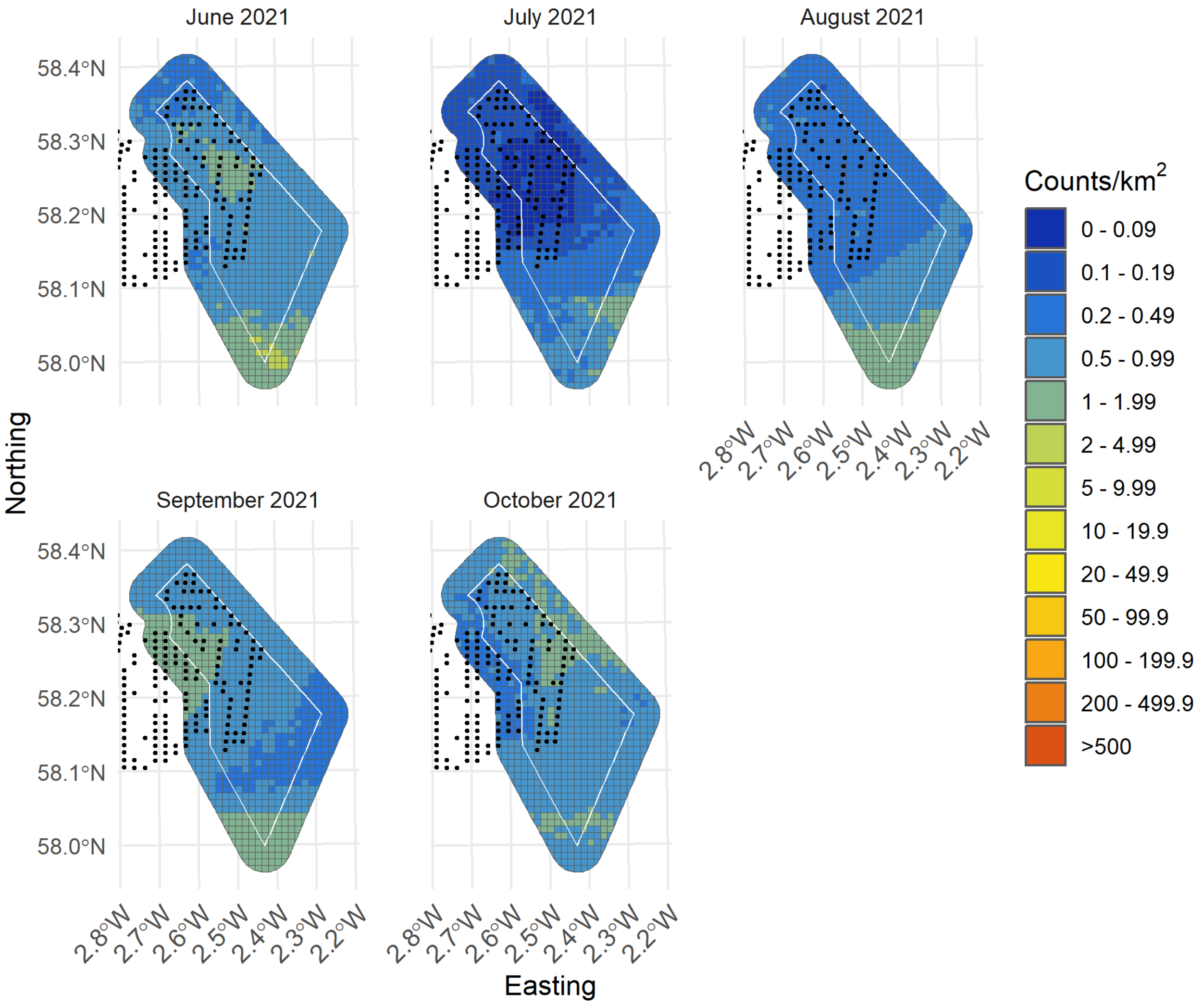


Figure II-45: Random forests predictions from turbine scenario 1 of Gannet from June 2021 to October 2021 with turbines presented as black dots

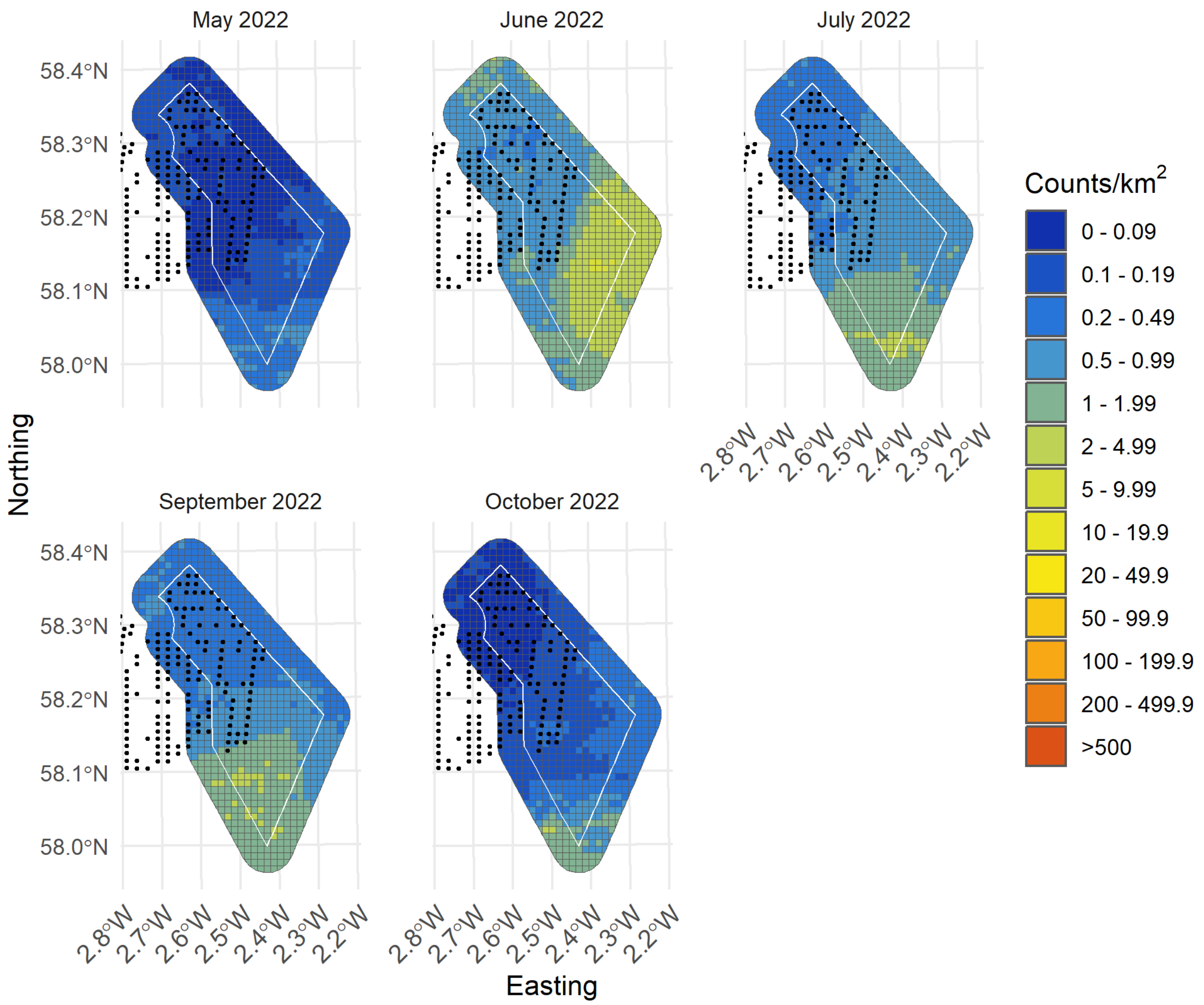


Figure II-46: Random forests predictions from turbine scenario 1 of Gannet from May 2022 to October 2022 with turbines presented as black dots

II.3.5.3 Turbine scenario 2

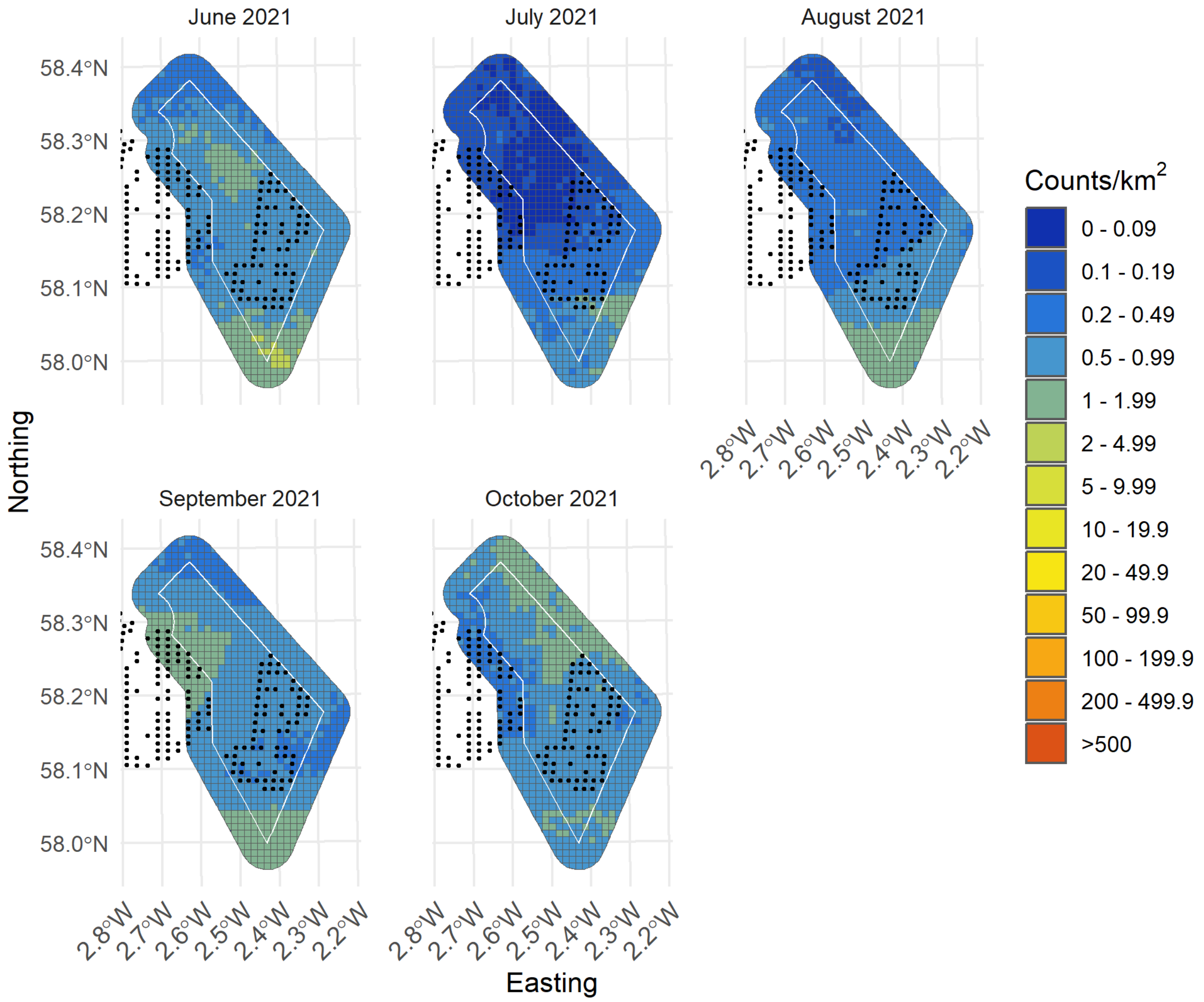


Figure II-47: Random forests predictions from turbine scenario 2 of Gannet from June 2021 to October 2021 with turbines presented as black dots

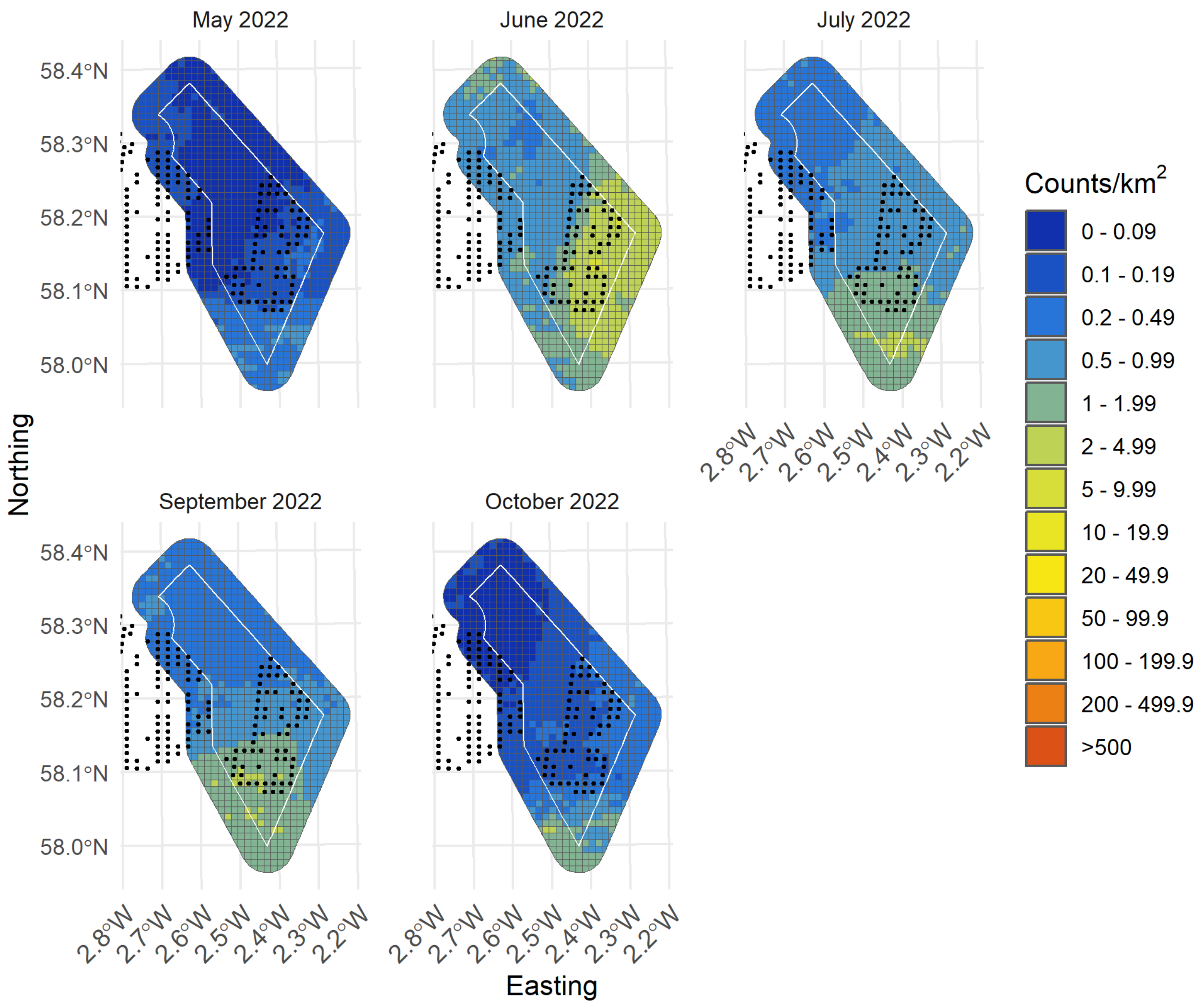


Figure II-48: Random forests predictions from turbine scenario 2 of Gannet from May 2022 to October 2022 with turbines presented as black dots

II.3.5.4 Turbine scenario 3

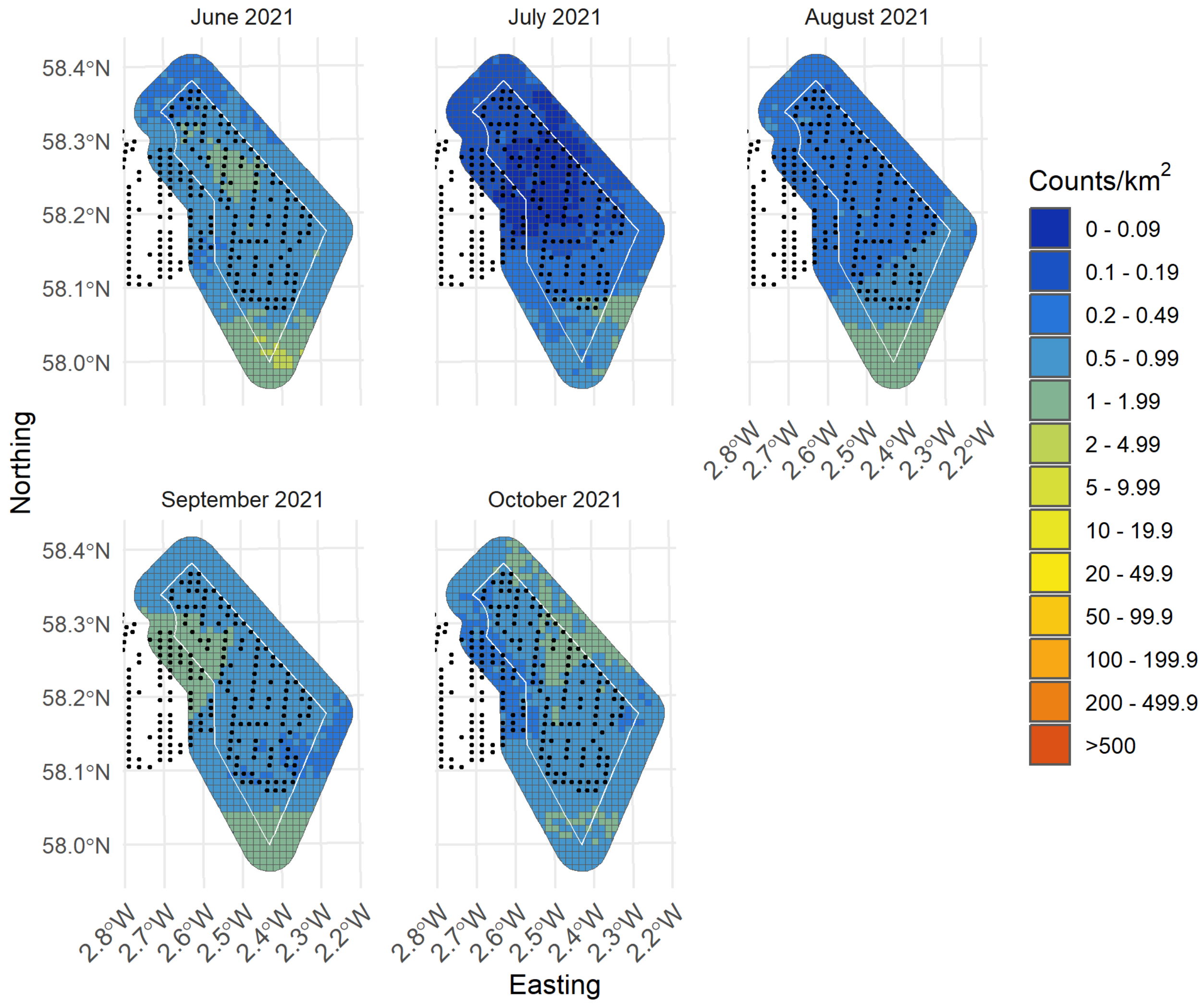


Figure II-49: Random forests predictions from turbine scenario 3 of Gannet from June 2021 to October 2021 with turbines presented as black dots

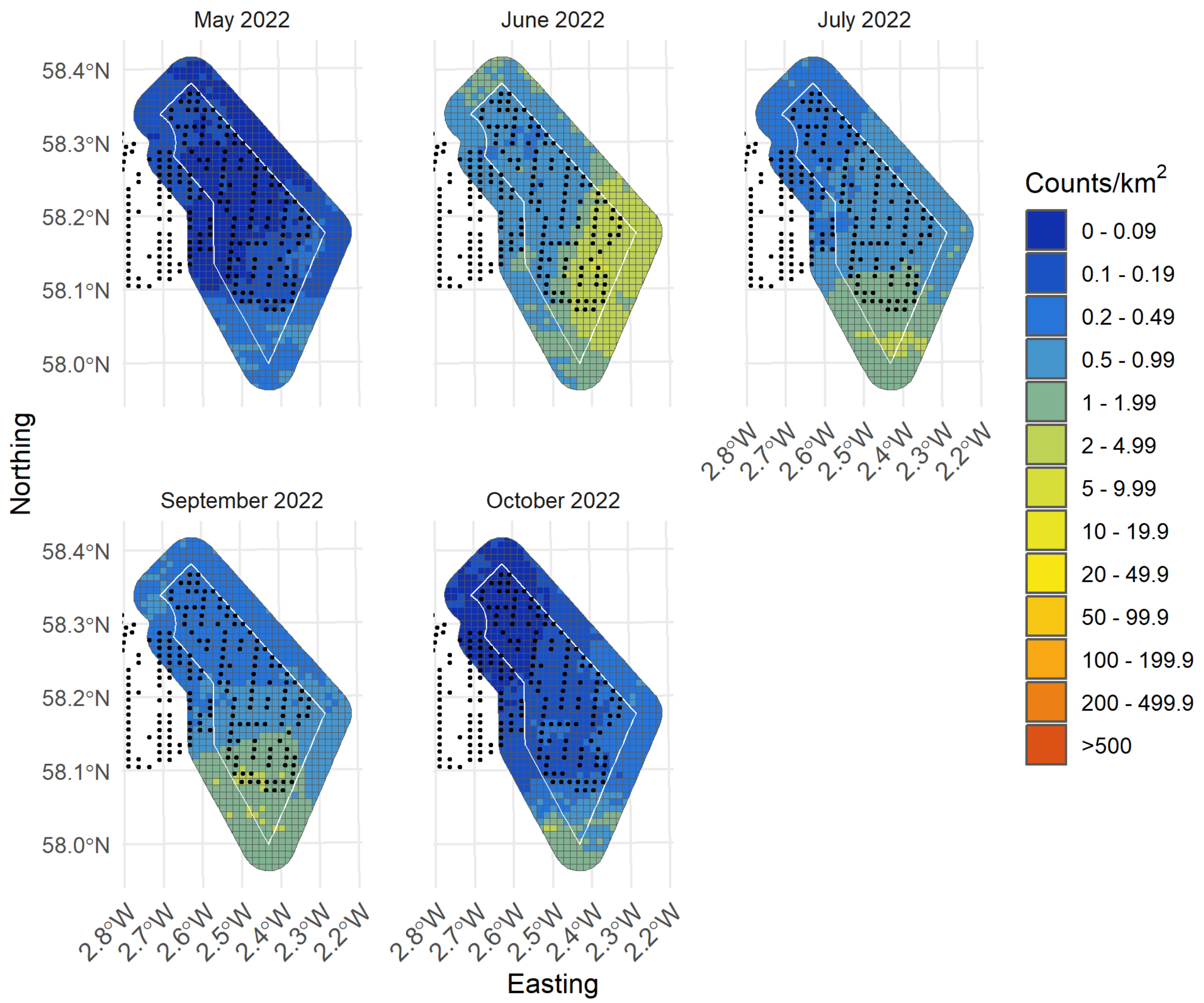


Figure II-50: Random forests predictions from turbine scenario 3 of Gannet from May 2022 to October 2022 with turbines presented as black dots

II.4 Puffin

II.4.1 Model assessment

The top five models had RMSE values between 3.999 and 4.011 with R squared values between 0.163 and 0.159. The R squared values were slightly lower than those from the MRSea model, which had an R-squared value of 0.1669. The best model had an mtry value of 8, with a minimum node size of 80 (Figure II-51, Table II-10).

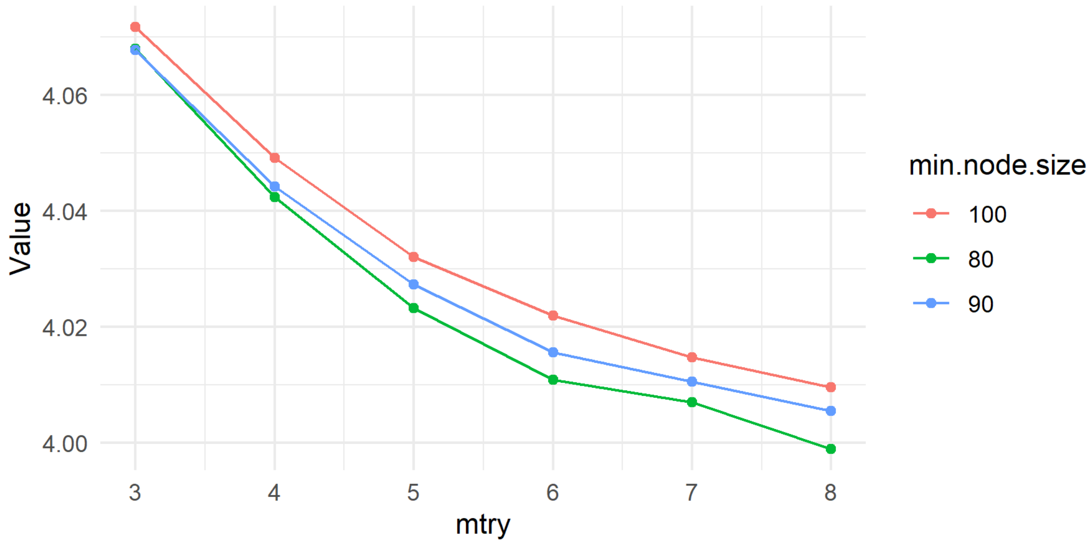


Figure II-51: Root mean squared error when varying mtry and min.node.size parameters in random forests

Table II-10: The top 5 models selected by the random forests analysis showing root mean squared error (RMSE) and R-squared values as calculated by 5-fold cross validation

mtry	min.node.size	RMSE	Rsquared
8	80	3.999	0.1633
8	90	4.005	0.1606
7	80	4.007	0.1603
8	100	4.010	0.1589
7	90	4.011	0.1587

II.4.2 Variable importance

The top predictor variables were survey ID, standard deviation of monthly sea surface temperature, and monthly mean sea surface temperature. This shows that most of the signal in the data come from variables that represent temporal variability (Table II-11).

Table II-11: The top 5 predictor variables from the random forests model and overall importance

Variable	Overall importance
SurveyID.Q	100.00
SST_sd	98.34
SurveyID^13	79.72
SST_mean	74.95
SurveyID^23	70.23

II.4.3 Distributional response

The partial relationship of distance to turbines on densities of puffin demonstrates a displacement effect with fewer birds expected in areas close to turbines. The relationship reaches a slight inflection point at approximately 20km, where the curve begins to flatten (Figure II-52). This does not indicate that there is a displacement effect out to 20km, simply that the model predicts more birds further away from the turbines, which could be due to relationships with other covariates.

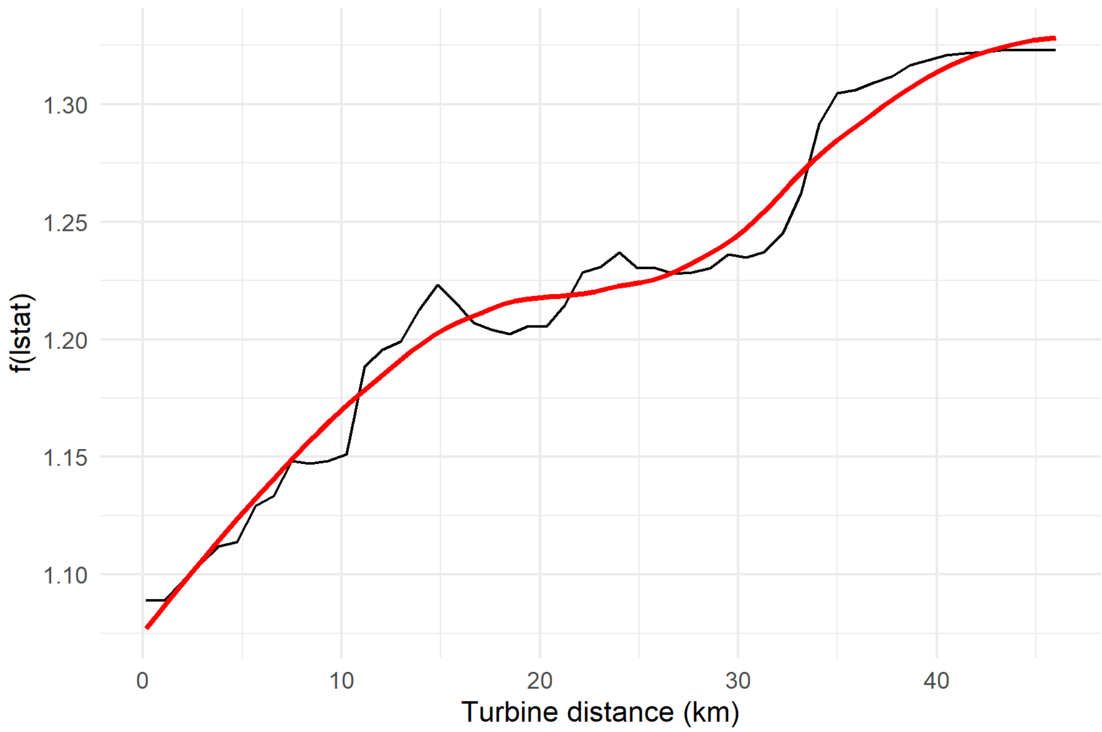


Figure II-52: Partial dependence plot of distance to turbine from the random forests model. The red line depicts the loess curve

II.4.4 Population estimates

Population estimates derived from the random forests baseline scenario (i.e., Existing turbines) fell well within the range of estimates derived from MRSea, sometimes only differing by 0.6 - 2%. We are therefore confident that the random forests puffin model is reliable, particularly in the context of the MRSea analysis. There were notably higher populations in August 2021 and September 2022, likely associated with post-breeding dispersal.

Populations of puffin in the survey area were mostly predicted to decrease in all three scenarios due to the relationship between distance to turbine and density, with the exception of the September 2021, where the population was predicted to increase by 11.69% in scenarios 2 and 3. The biggest population change was predicted for scenario 3, with percent change from baseline varying from -14% to 11.69% across all surveys (Table II-12).

Table II-12: Population estimates from the random forests models for the 4km buffer zone plus windfarm footprint survey area for the baseline scenario (i.e., based on currently installed turbines) and three potential turbine scenarios. Percent change from baseline is calculated for each scenario

SurveyID	Baseline	Scenario 1 (% change)	Scenario 2 (% change)	Scenario 3 (% change)
May 2021	633	577 (-8.85)	577 (-8.85)	573 (-9.48)
June 2021	520	492 (-5.38)	492 (-5.38)	491 (-5.58)
July 2021	687	601 (-12.52)	593 (-13.68)	591 (-13.97)
August 2021	4268	4050 (-5.11)	4062 (-4.83)	3932 (-7.87)
September 2021	154	168 (9.09)	172 (11.69)	172 (11.69)
October 2021	266	270 (1.5)	270 (1.5)	270 (1.5)
April 2022	1151	1142 (-0.78)	1148 (-0.26)	1131 (-1.74)
May 2022	1256	1197 (-4.7)	1196 (-4.78)	1193 (-5.02)
June 2022	165	159 (-3.64)	160 (-3.03)	160 (-3.03)
July 2022	238	239 (0.42)	239 (0.42)	238 (0)
August 2022	858	792 (-7.69)	788 (-8.16)	785 (-8.51)
September 2022	3096	2927 (-5.46)	2931 (-5.33)	2901 (-6.3)

II.4.5 Distributions

The broad distribution of puffin in the baseline random forests models were nearly identical to those in the MRSea models. For most surveys, the models predicted the highest densities in the eastern and southern parts of the survey area, with the exception of August 2021 and September 2022 where high densities of birds were identified through the whole region. In all cases, the densities of puffin decrease through the site in relation to the locations of proposed turbines. This distributional response suggests a potential displacement effect which could push individuals beyond the boundary of the wind farm site (Figures II-53 – II-64).

II.4.5.1 Baseline scenarios

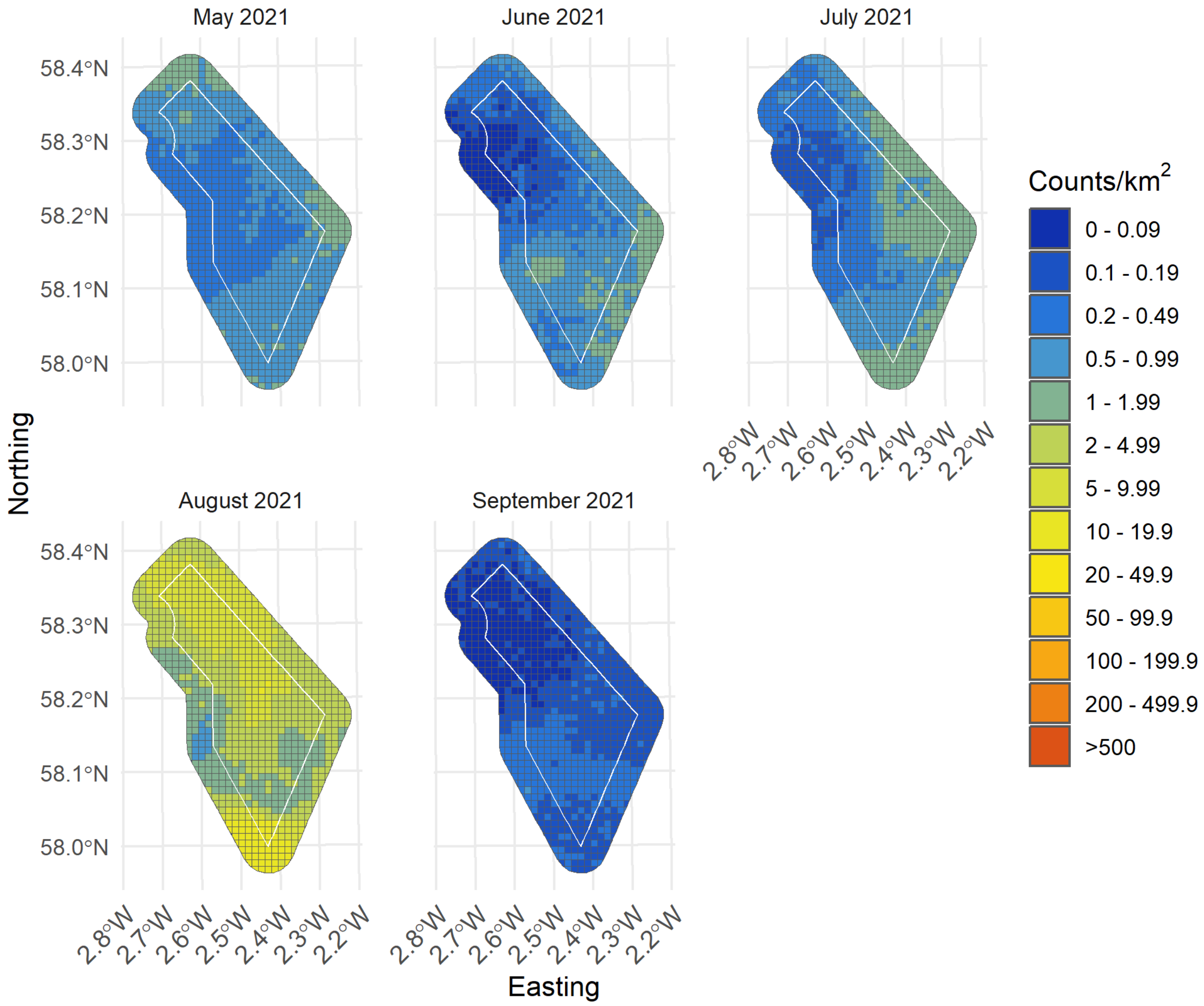


Figure II-53: Random forests baseline predictions of Puffin from May 2021 to September 2021

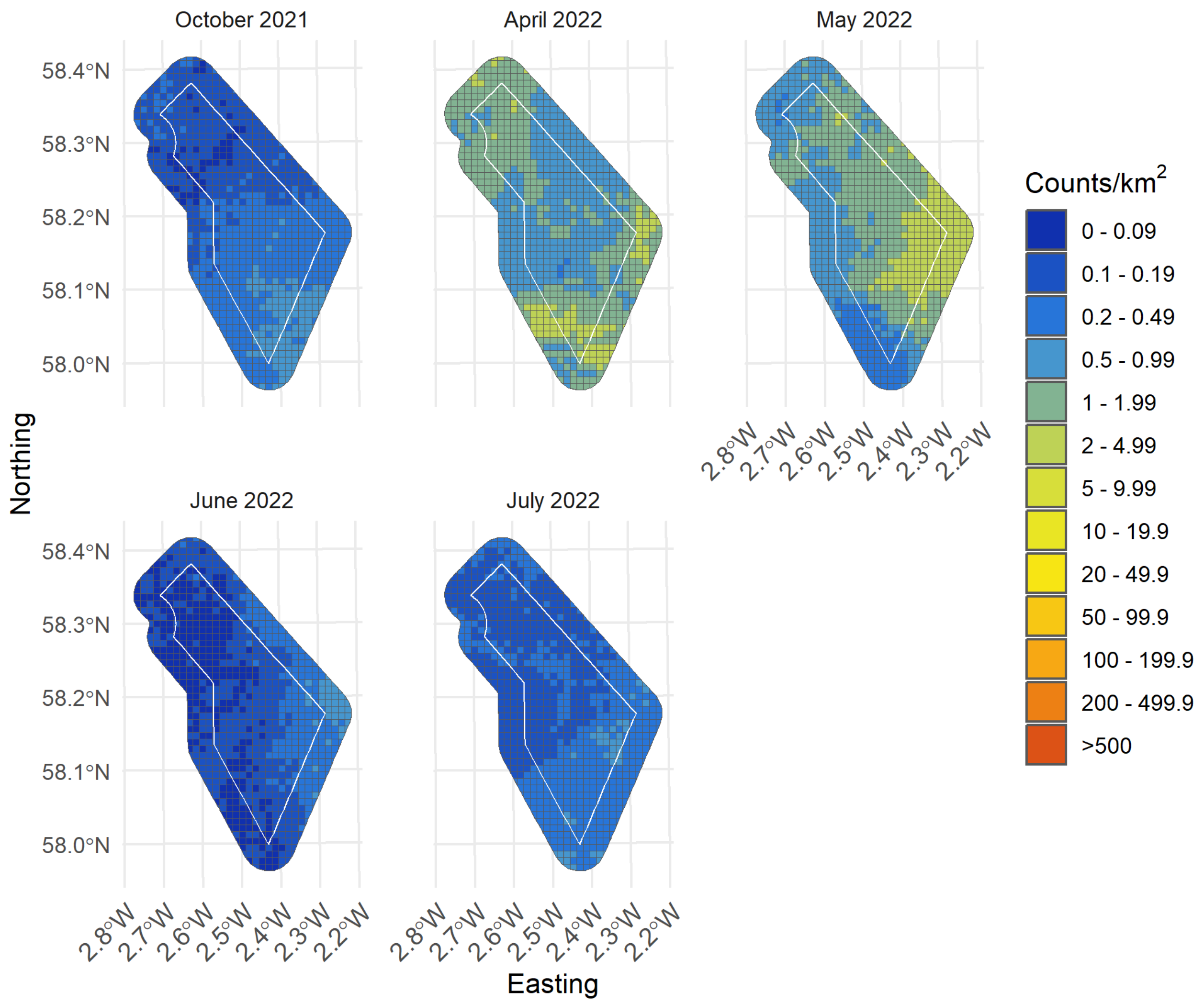


Figure II-54: Random forests baseline predictions of Puffin from October 2021 to July 2022

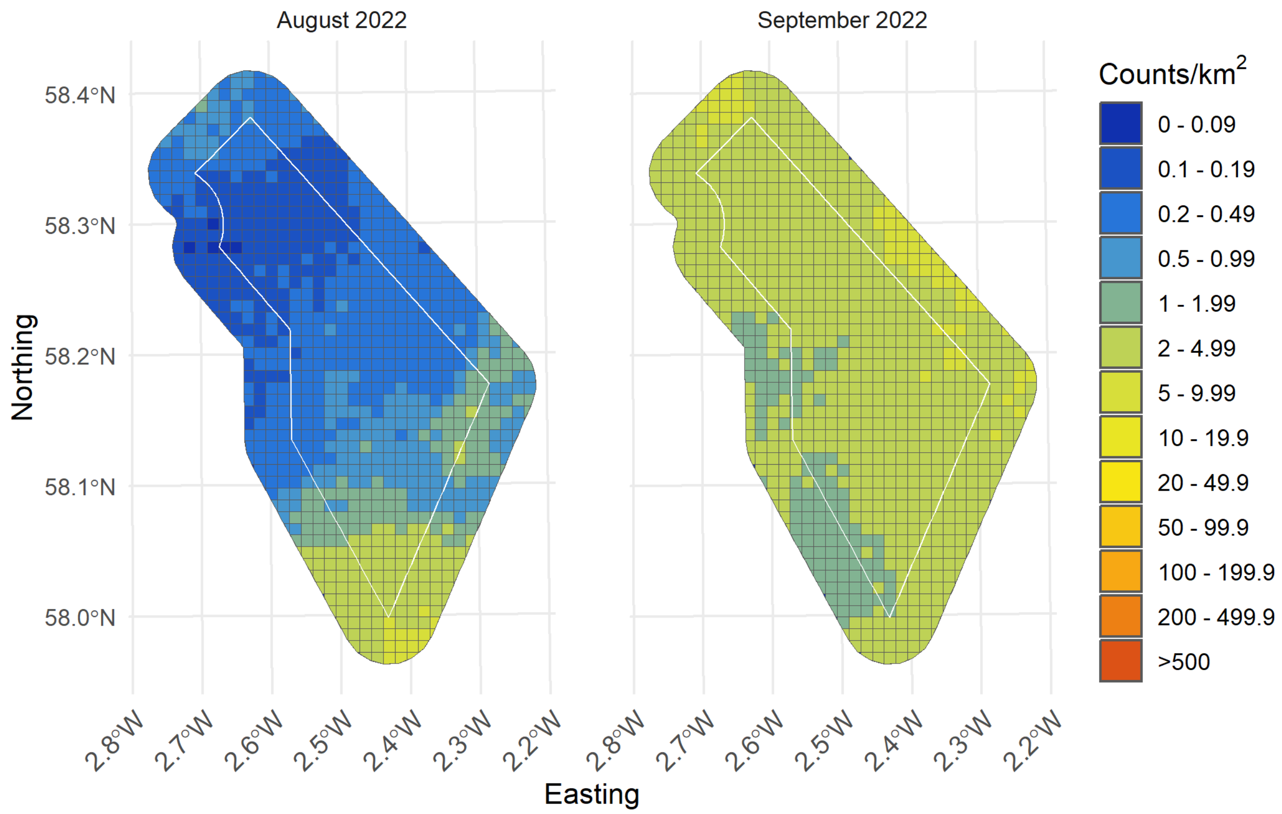


Figure II-55: Random forests baseline predictions of Puffin from August 2022 to September 2022

II.4.5.2 Turbine scenario 1

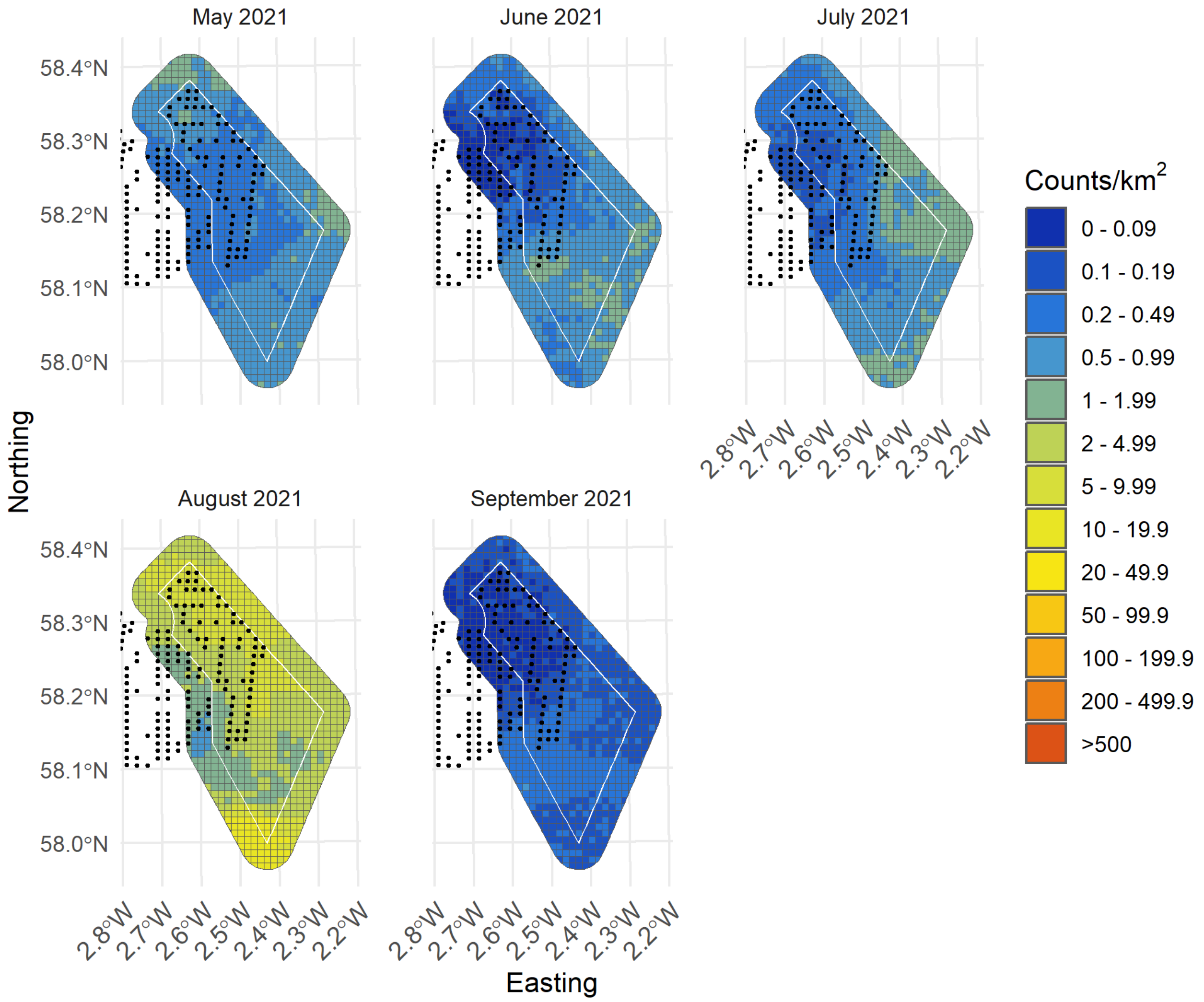


Figure II-56: Random forests predictions from turbine scenario 1 of Puffin from May 2021 to September 2021 with turbines presented as black dots

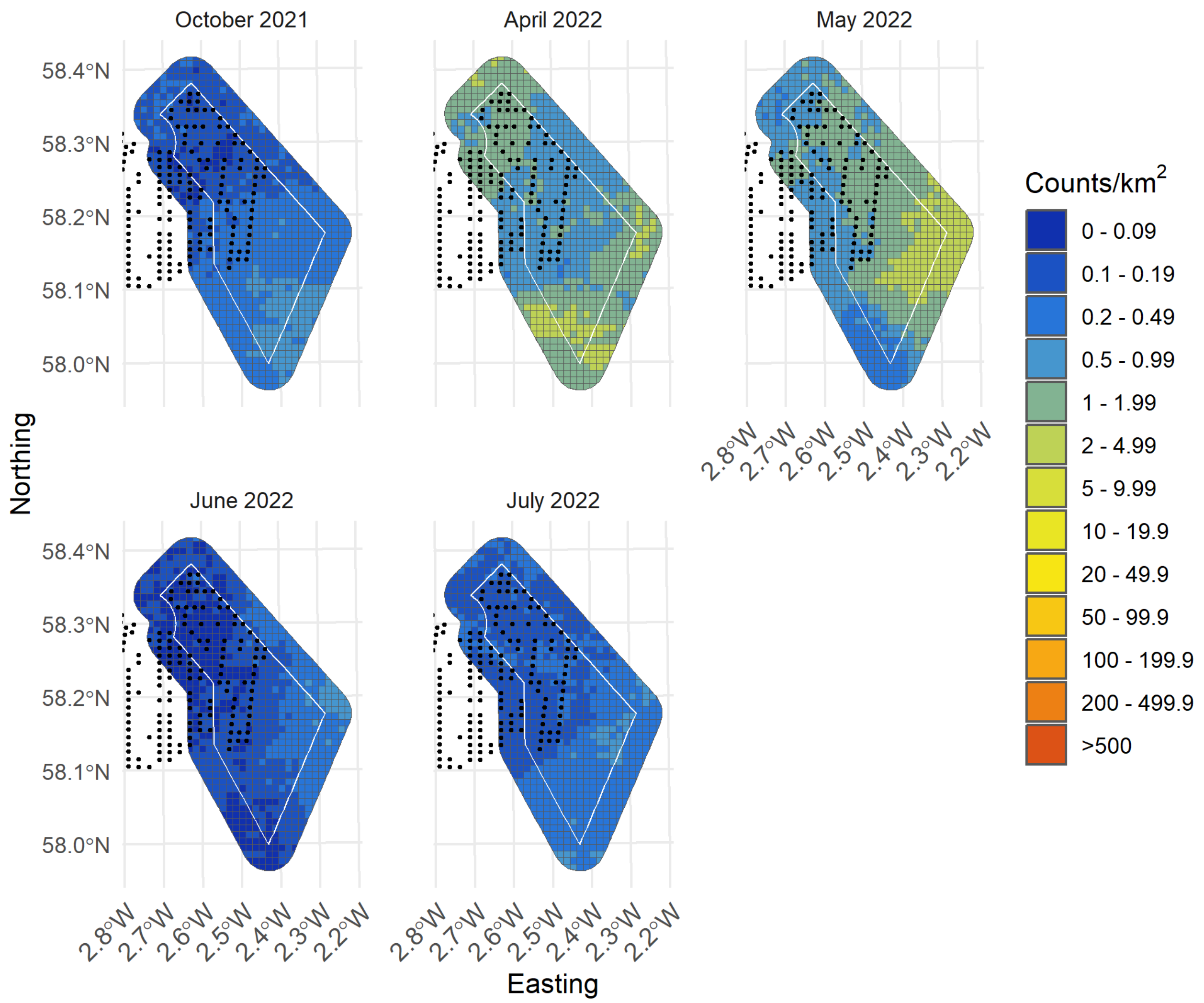


Figure II-57: Random forests predictions from turbine scenario 1 of Puffin from October 2021 to July 2022 with turbines presented as black dots

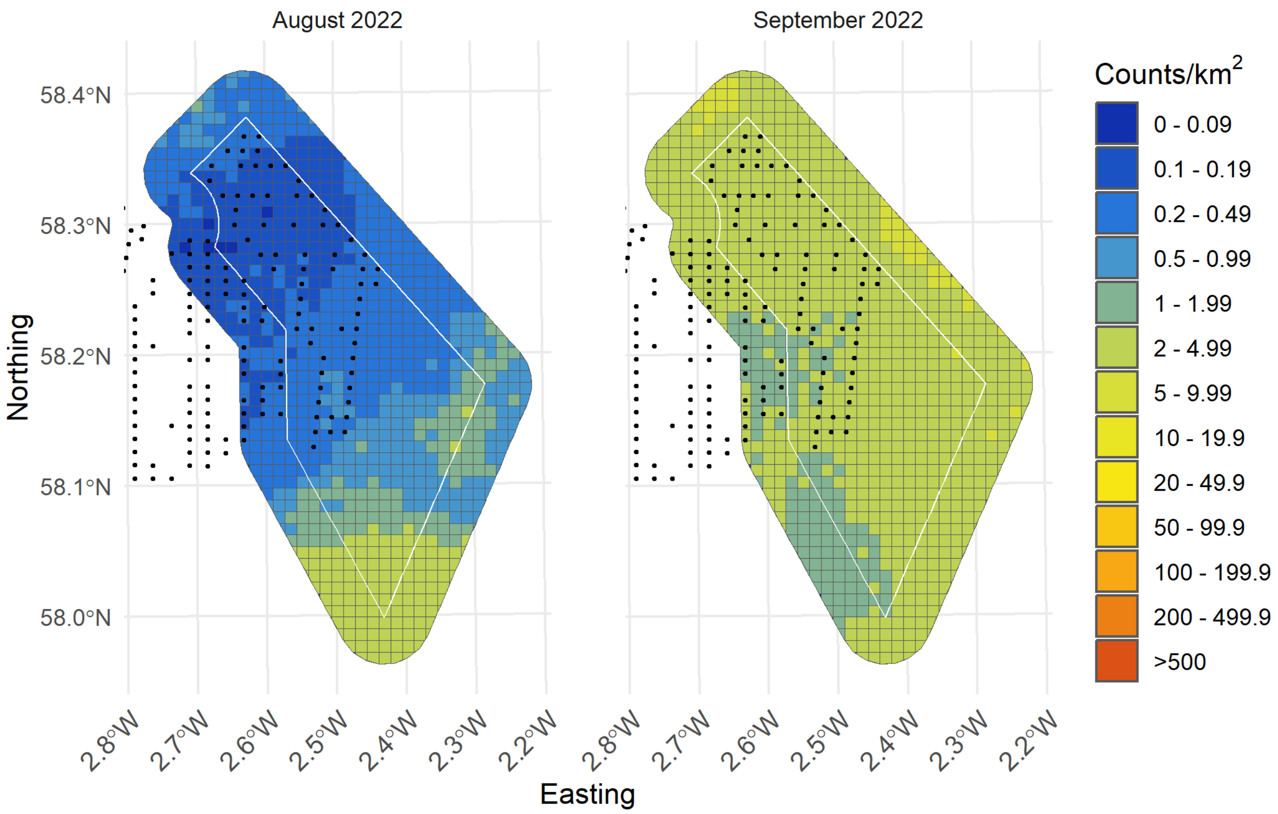


Figure II-58: Random forests predictions from turbine scenario 1 of Puffin from August 2022 to September 2022 with turbines presented as black dots

II.4.5.3 Turbine scenario 2

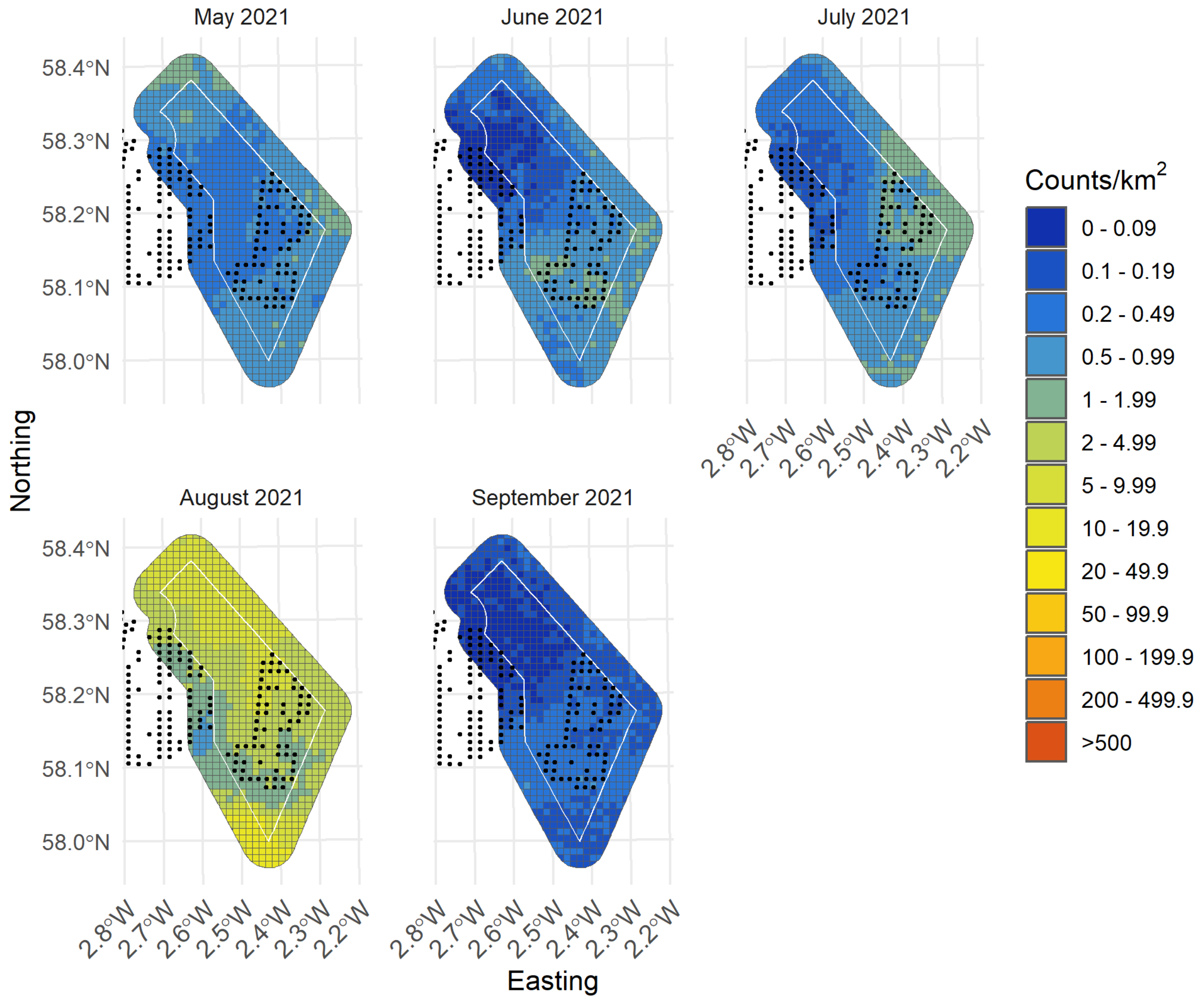


Figure II-59: Random forests predictions from turbine scenario 2 of Puffin from May 2021 to September 2021 with turbines presented as black dots

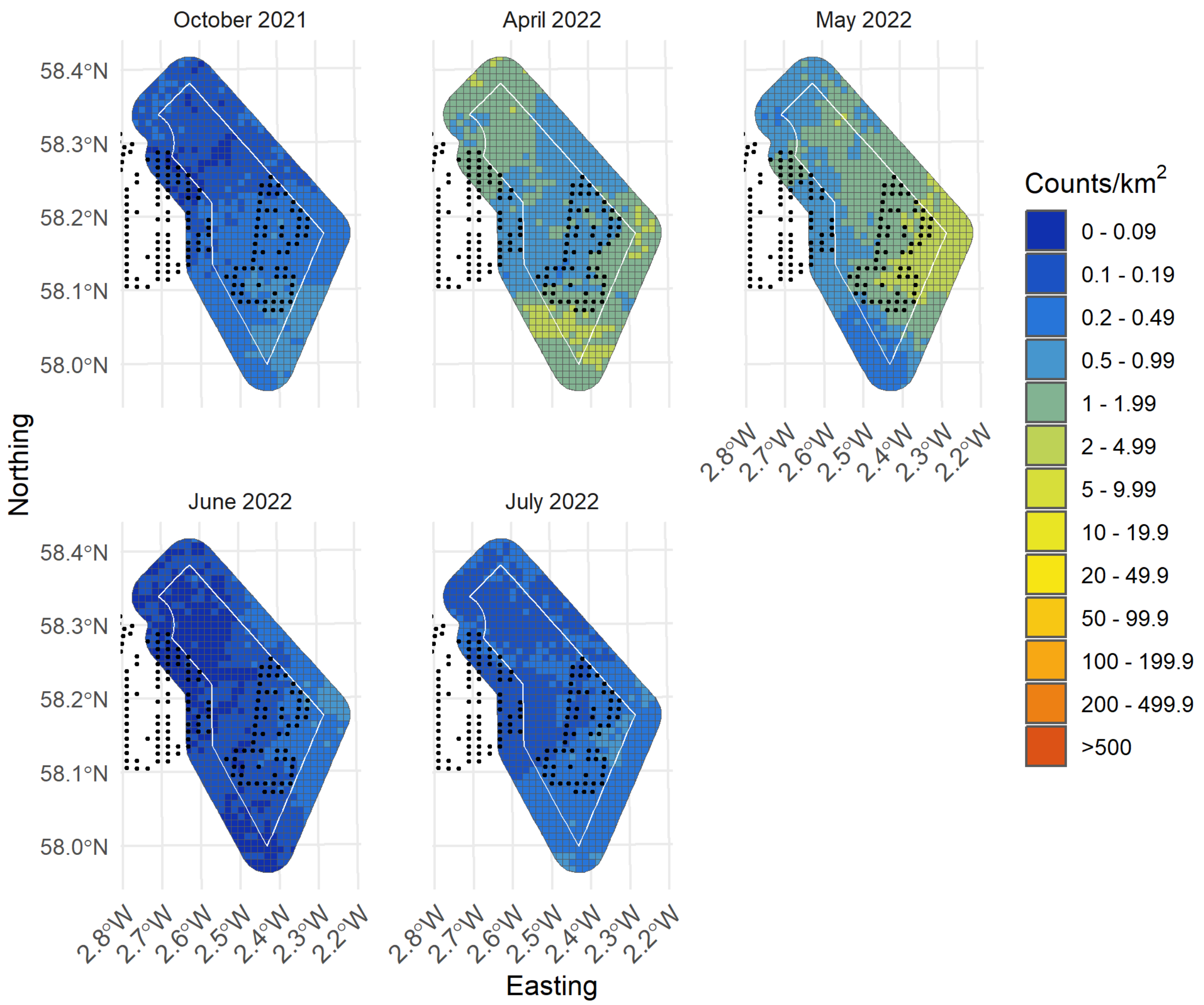


Figure II-60: Random forests predictions from turbine scenario 2 of Puffin from October 2021 to July 2022 with turbines presented as black dots

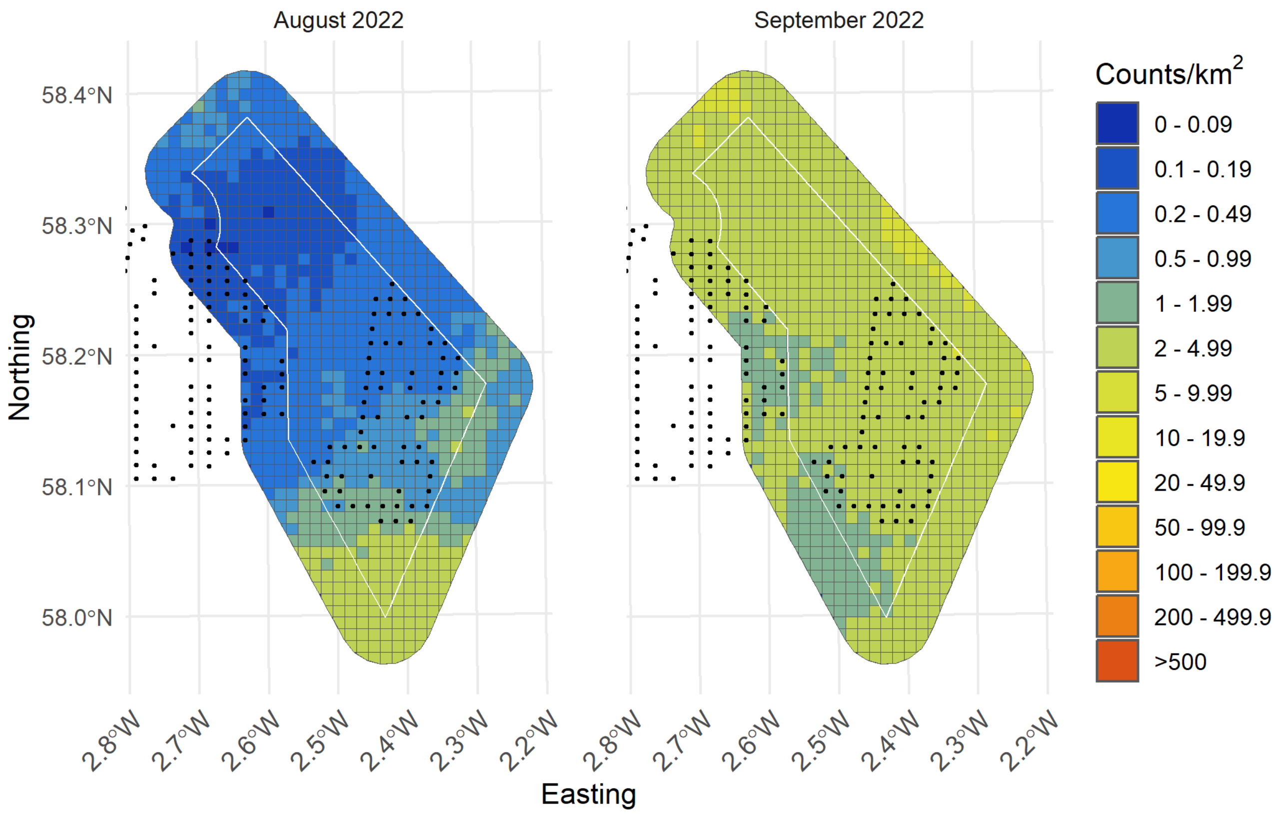


Figure II-61: Random forests predictions from turbine scenario 2 of Puffin from August 2022 to September 2022 with turbines presented as black dots

II.4.5.4 Turbine scenario 3

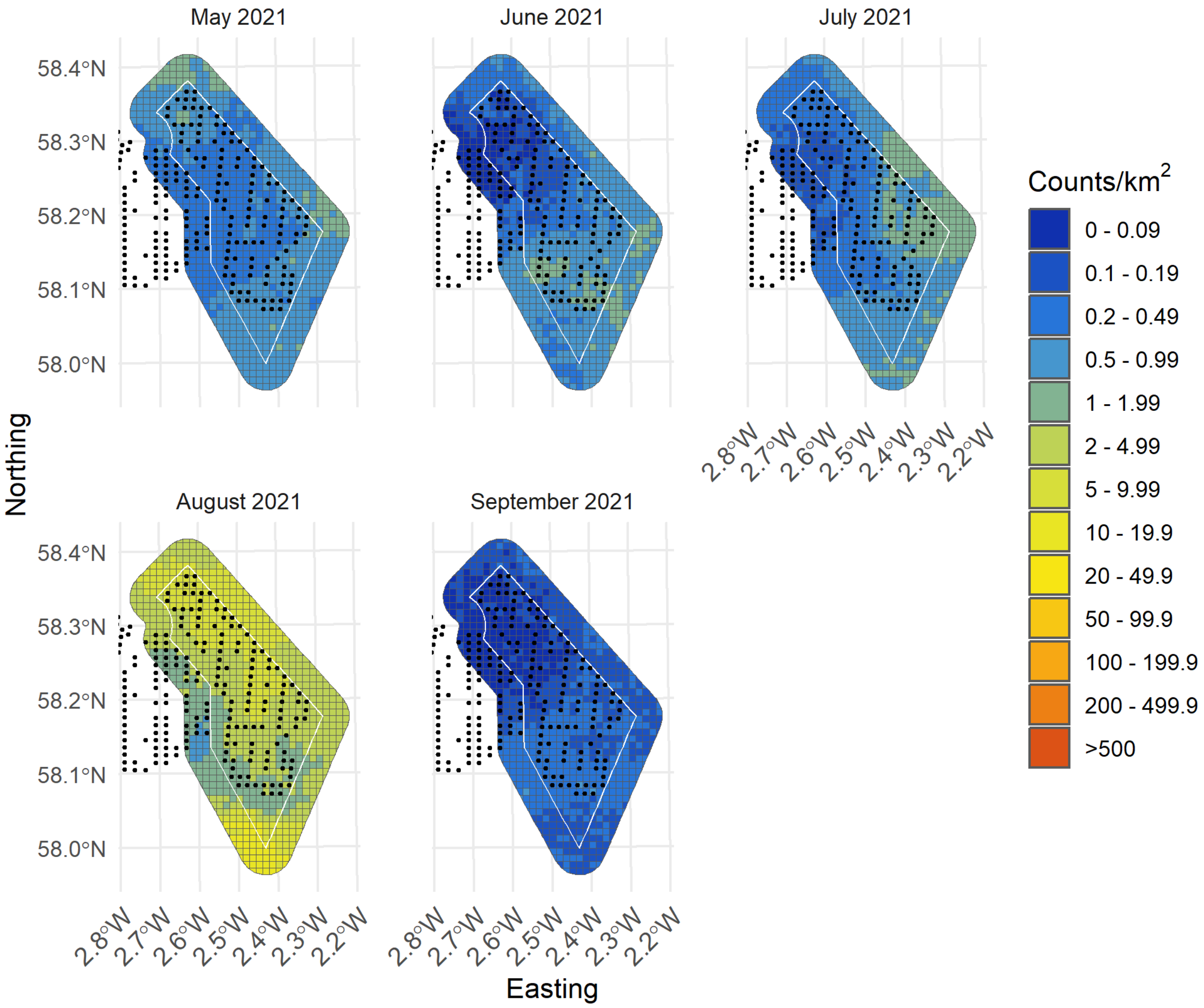


Figure II-62: Random forests predictions from turbine scenario 3 of Puffin from May 2021 to September 2021 with turbines presented as black dots

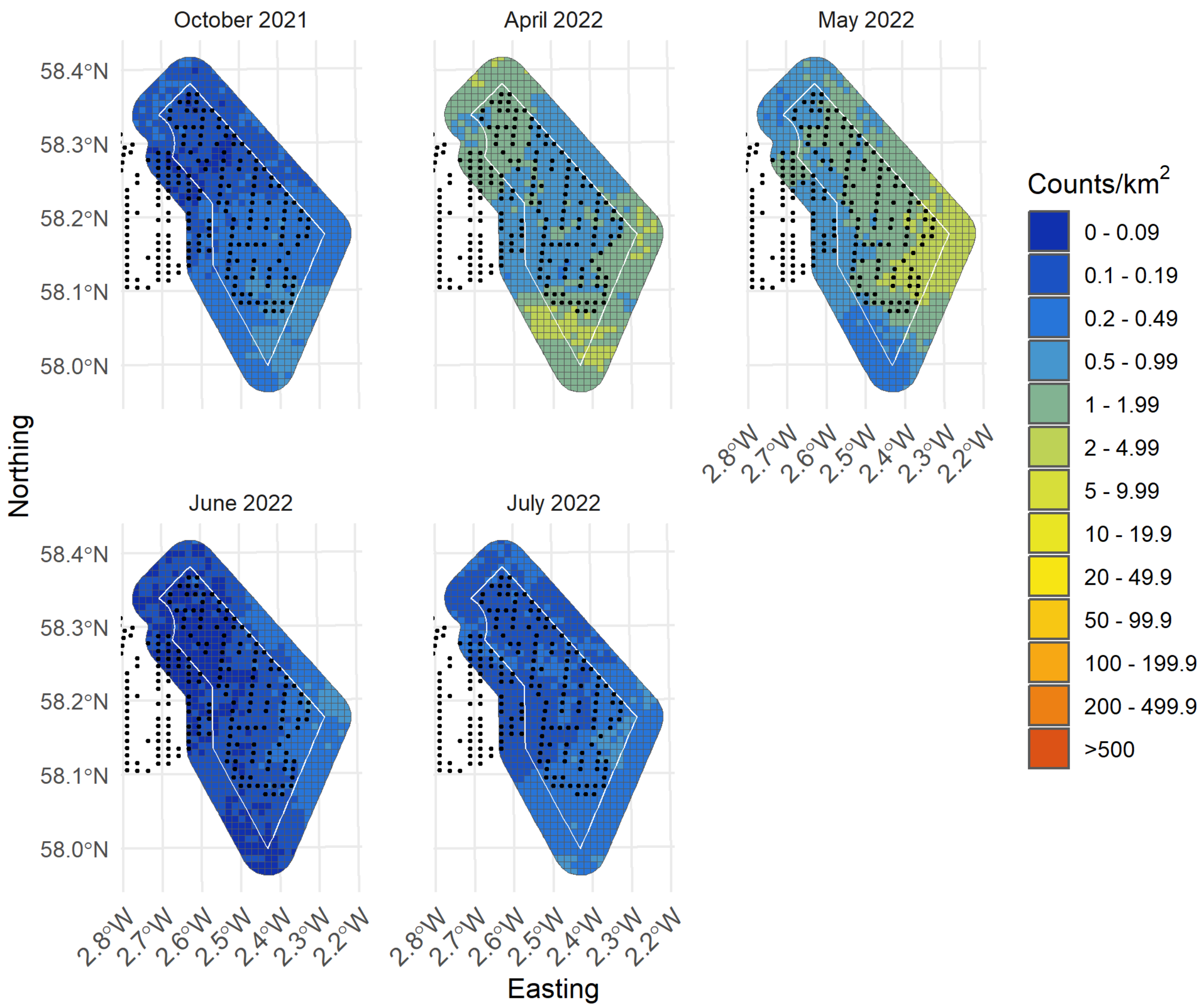


Figure II-63: Random forests predictions from turbine scenario 3 of Puffin from October 2021 to July 2022 with turbines presented as black dots

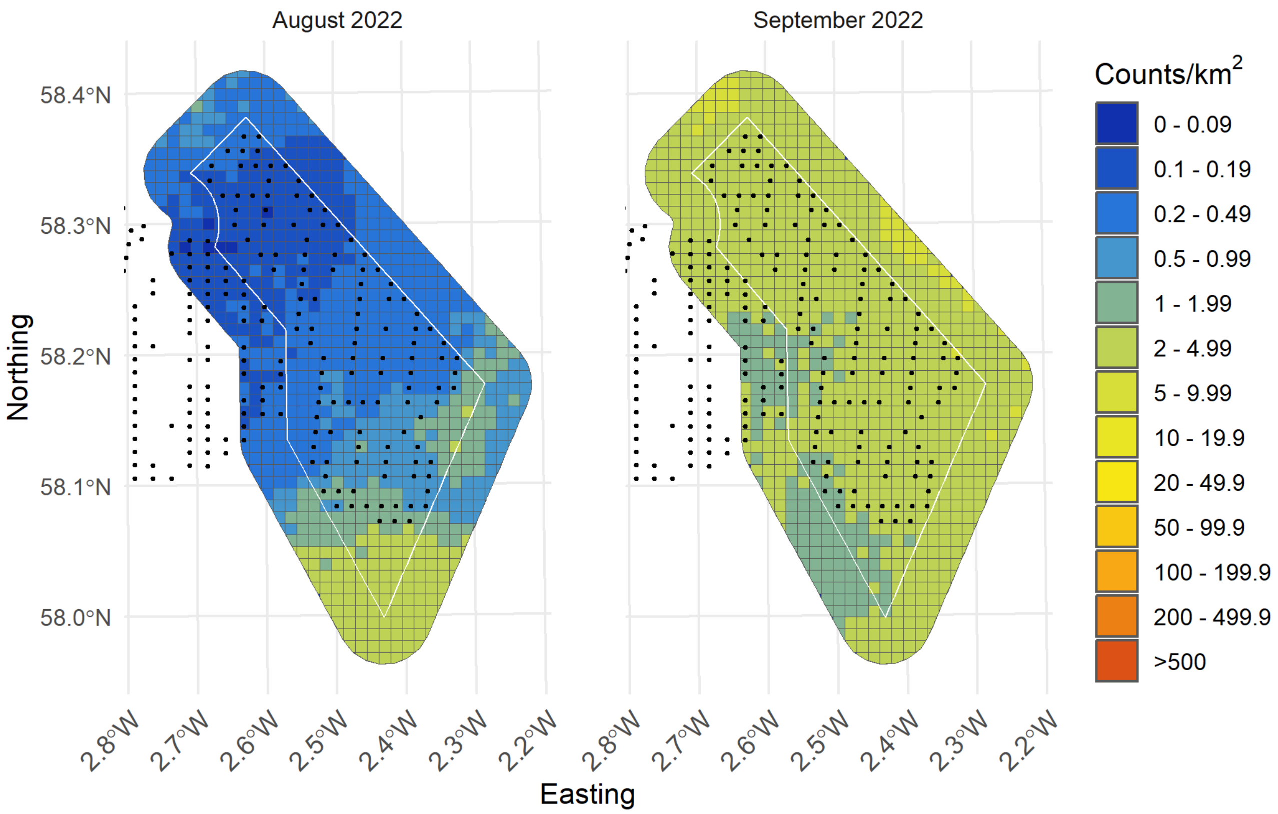


Figure II-64: Random forests predictions from turbine scenario 3 of Puffin from August 2022 to September 2022 with turbines presented as black dots

II.5 Razorbill

II.5.1 Model assessment

The top five models had RMSE values between 4.125 and 4.127 with R squared values between 0.0892 and 0.0903. The R squared values were higher than those from the MRSea model, which had an R-squared value of 0.0776. The best model had an mtry value of 6, with a minimum node size of 80 (Figure II-65, Table II-13).

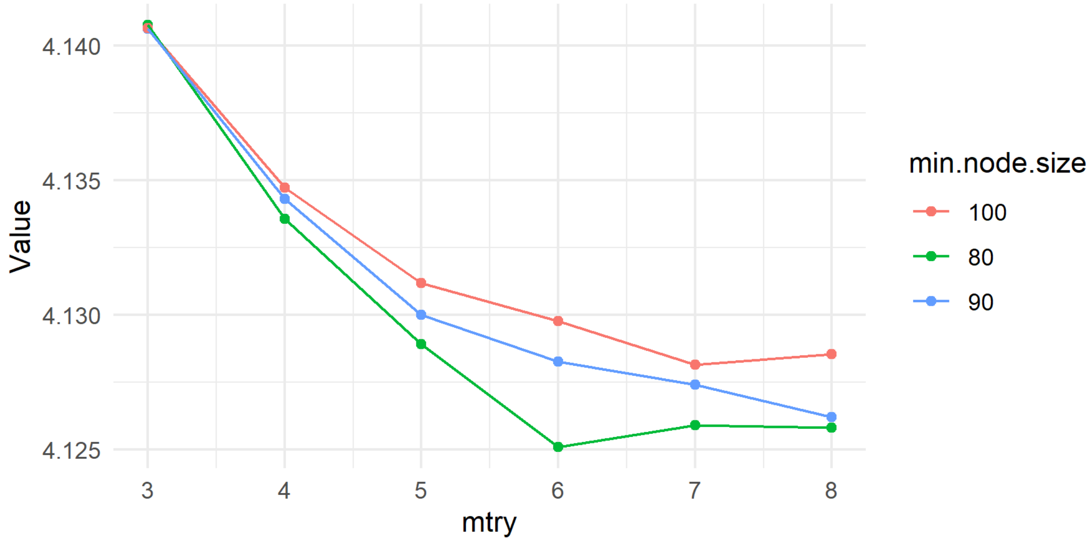


Figure II-65: Root mean squared error when varying mtry and min.node.size parameters in random forests

Table II-13: The top 5 models selected by the random forests analysis showing root mean squared error (RMSE) and R-squared values as calculated by 5-fold cross validation

mtry	min.node.size	RMSE	Rsquared
6	80	4.125	0.0903
8	80	4.126	0.0900
7	80	4.126	0.0899
8	90	4.126	0.0897
7	90	4.127	0.0892

II.5.2 Variable importance

The top predictor variables were monthly mean sea surface temperature, northing| (y.pos) and survey ID. This shows that most of the signal in the data come from variables that represent temporal variability (Table II-14).

Table II-14: The top 5 predictor variables from the random forests model and overall importance

Variable	Overall importance
SST_mean	100.00
y.pos	96.52
SurveyID^4	90.29
SurveyID^7	89.97
SurveyID^10	89.24

II.5.3 Distributional response

The partial relationship of distance to turbines on densities of razorbill demonstrates that higher densities of birds are predicted in areas close to turbines. The relationship reaches an inflection point at approximately 18km (Figure II-66). However, this effect seems quite small overall with the relative contribution towards predictions ranging between 1.05 and 1.13 (0.08). This does not necessarily mean that razorbill are attracted to turbines, but that they are predicted to be occurring at higher densities closer to turbines, which could be due to relationships with other environmental features.

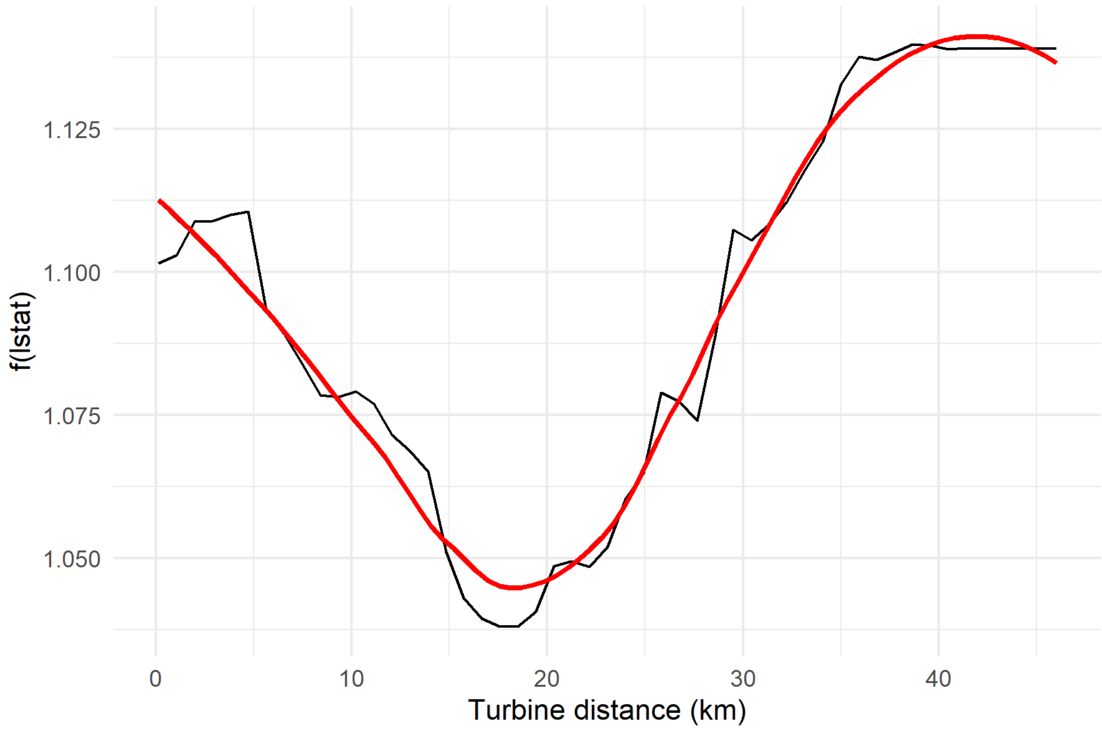


Figure II-66: Partial dependence plot of distance to turbine from the random forests model. The red line depicts the loess curve

II.5.4 Population estimates

Population estimates derived from the random forests baseline scenario (i.e., Existing turbines) fell well within the range of estimates derived from MRSea, sometimes only differing by 1 - 3%. We are therefore confident that the random forests razorbill model is reliable, particularly in the context of the MRSea analysis.

Populations of razorbill in the survey area increased and decreased variably in all three scenarios due to the relationship between distance to turbine and density. The biggest population change was predicted for scenario 3, with percent change from baseline varying from -5% to 46% across all surveys (Table II-15).

Table II-15: Population estimates from the random forests models for the 4km buffer zone plus windfarm footprint survey area for the baseline scenario (i.e., based on currently installed turbines) and three potential turbine scenarios. Percent change from baseline is calculated for each scenario

SurveyID	Baseline	Scenario 1 (% change)	Scenario 2 (% change)	Scenario 3 (% change)
May 2021	1120	1253 (11.88)	1254 (11.96)	1273 (13.66)
June 2021	1188	1157 (-2.61)	1154 (-2.86)	1155 (-2.78)
July 2021	1869	1758 (-5.94)	1762 (-5.72)	1764 (-5.62)
August 2021	709	706 (-0.42)	705 (-0.56)	707 (-0.28)
September 2021	530	522 (-1.51)	519 (-2.08)	526 (-0.75)
February 2022	717	719 (0.28)	704 (-1.81)	727 (1.39)
April 2022	981	1044 (6.42)	1042 (6.22)	1047 (6.73)
May 2022	2560	2983 (16.52)	2888 (12.81)	3058 (19.45)
June 2022	226	239 (5.75)	242 (7.08)	243 (7.52)
July 2022	1257	1215 (-3.34)	1206 (-4.06)	1200 (-4.53)
August 2022	381	428 (12.34)	431 (13.12)	432 (13.39)
September 2022	3737	3572 (-4.42)	3660 (-2.06)	3544 (-5.16)
October 2022	138	193 (39.86)	193 (39.86)	198 (43.48)
November 2022	217	218 (0.46)	218 (0.46)	218 (0.46)
February 2023	434	460 (5.99)	458 (5.53)	461 (6.22)
March 2023	547	569 (4.02)	567 (3.66)	565 (3.29)
April 2023	552	567 (2.72)	565 (2.36)	565 (2.36)

II.5.5 Distributions

The broad distribution of razorbill in the baseline random forests models were nearly identical to those in the MRSea models. For most surveys, the models predicted the highest densities in the southern parts of the survey area, with the exception of May and July 2022 where high densities of birds were identified through the whole region. The distributional response was varied between the surveys, but is most extreme in scenario 3 (Figures II-67 – II-82).

II.5.5.1 Baseline scenarios

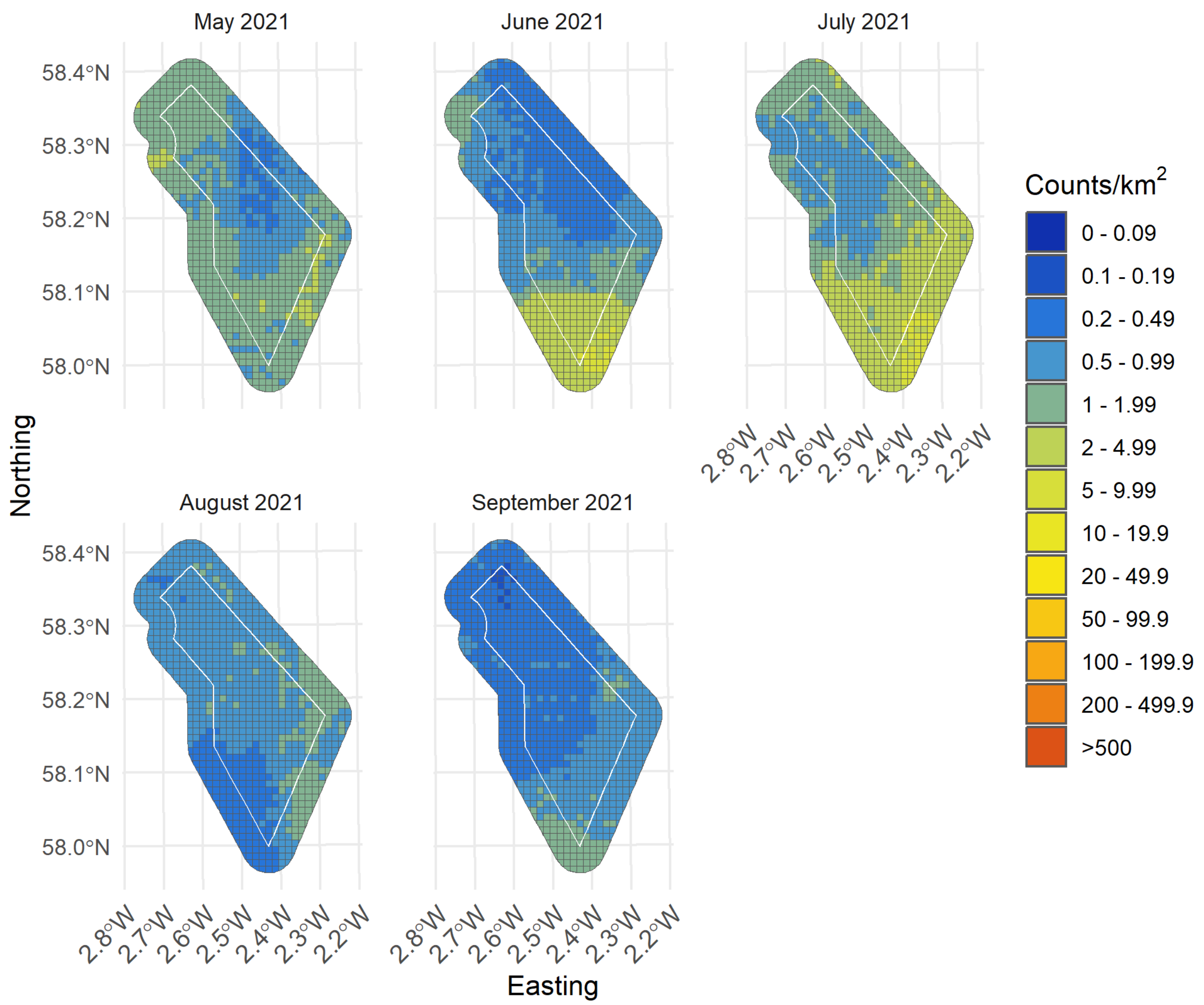


Figure II-67: Random forests baseline predictions of Razorbill from May 2021 to September 2021

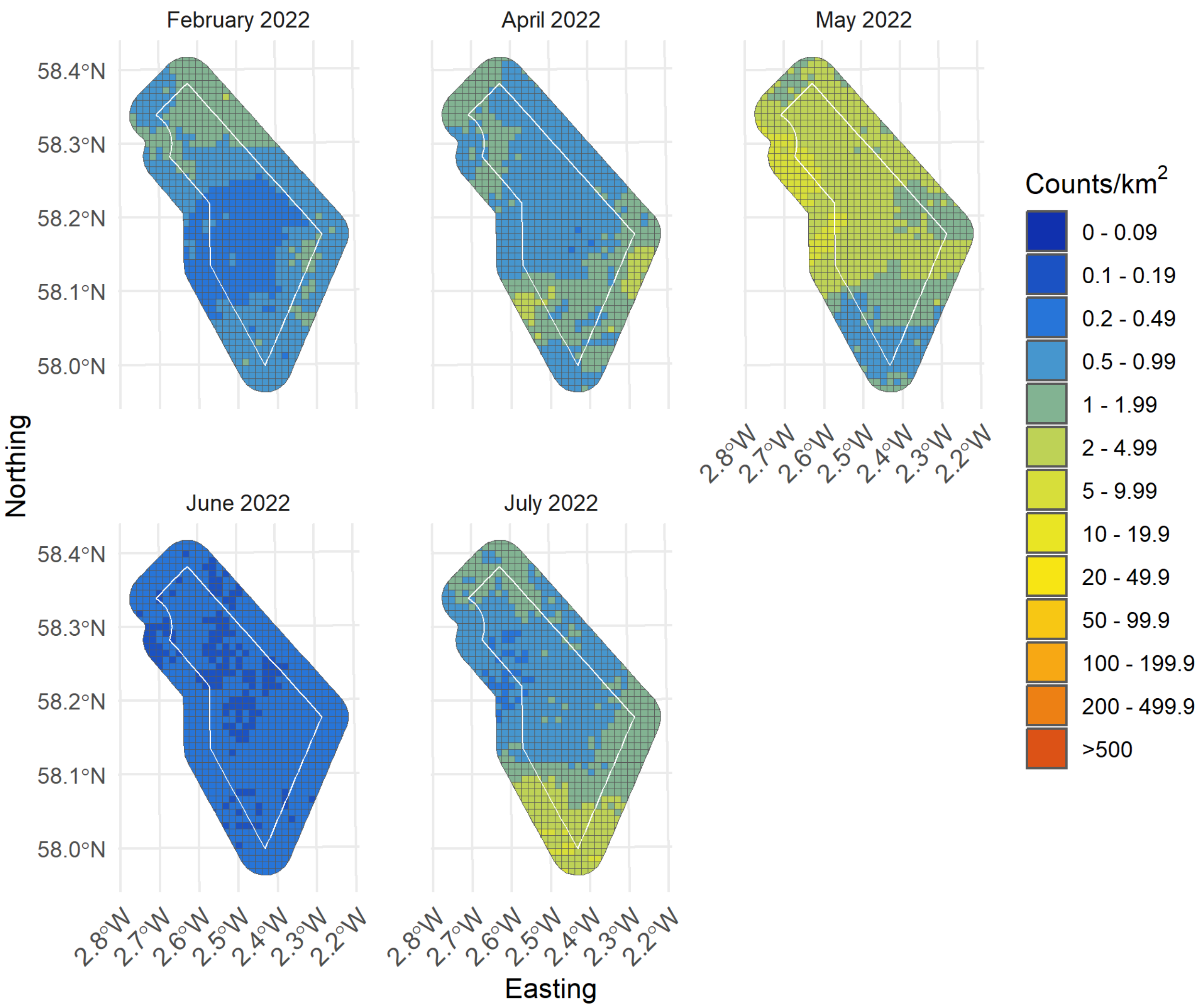


Figure II-68: Random forests baseline predictions of Razorbill from February 2022 to July 2022

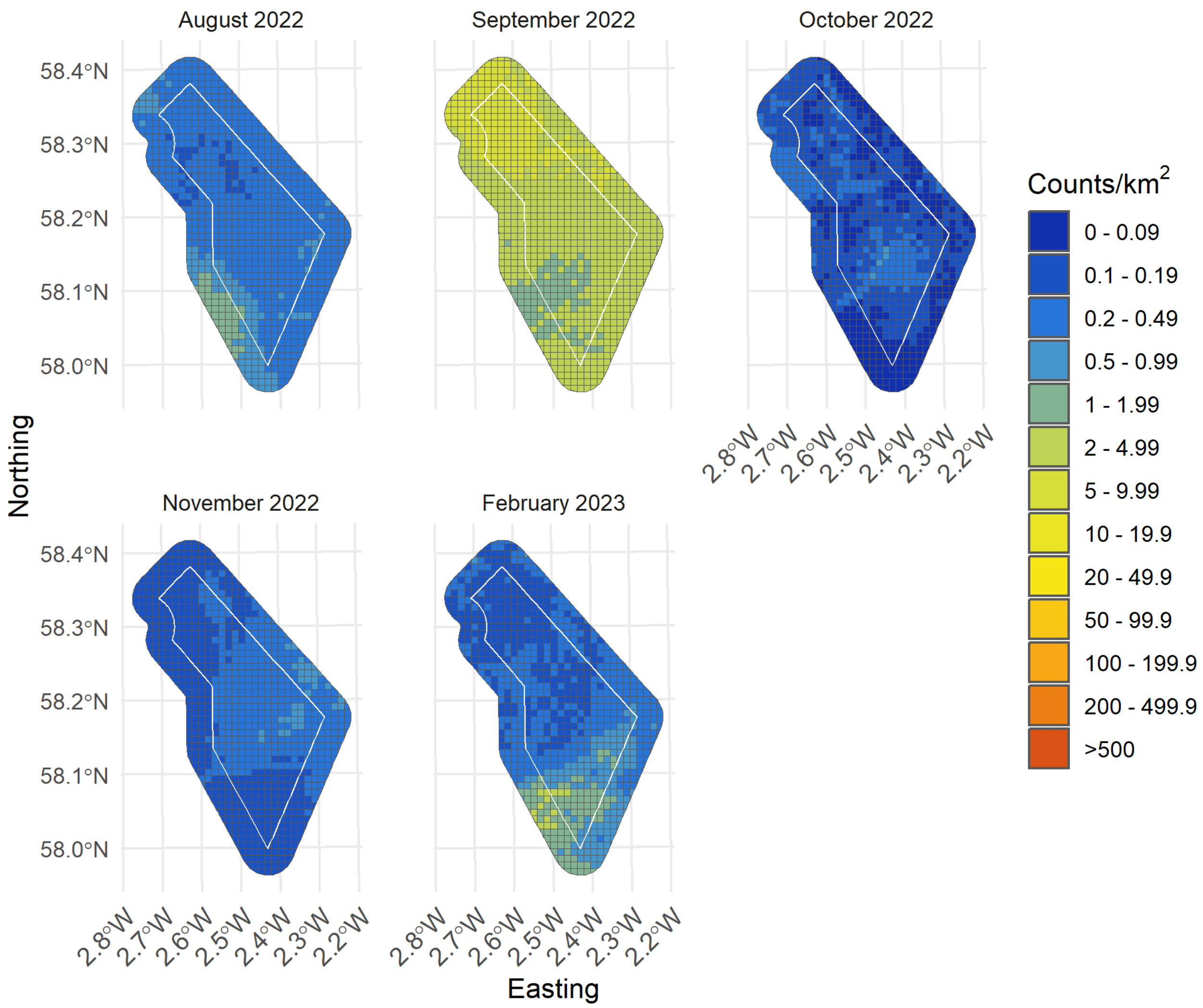


Figure II-69: Random forests baseline predictions of Razorbill from August 2022 to February 2023

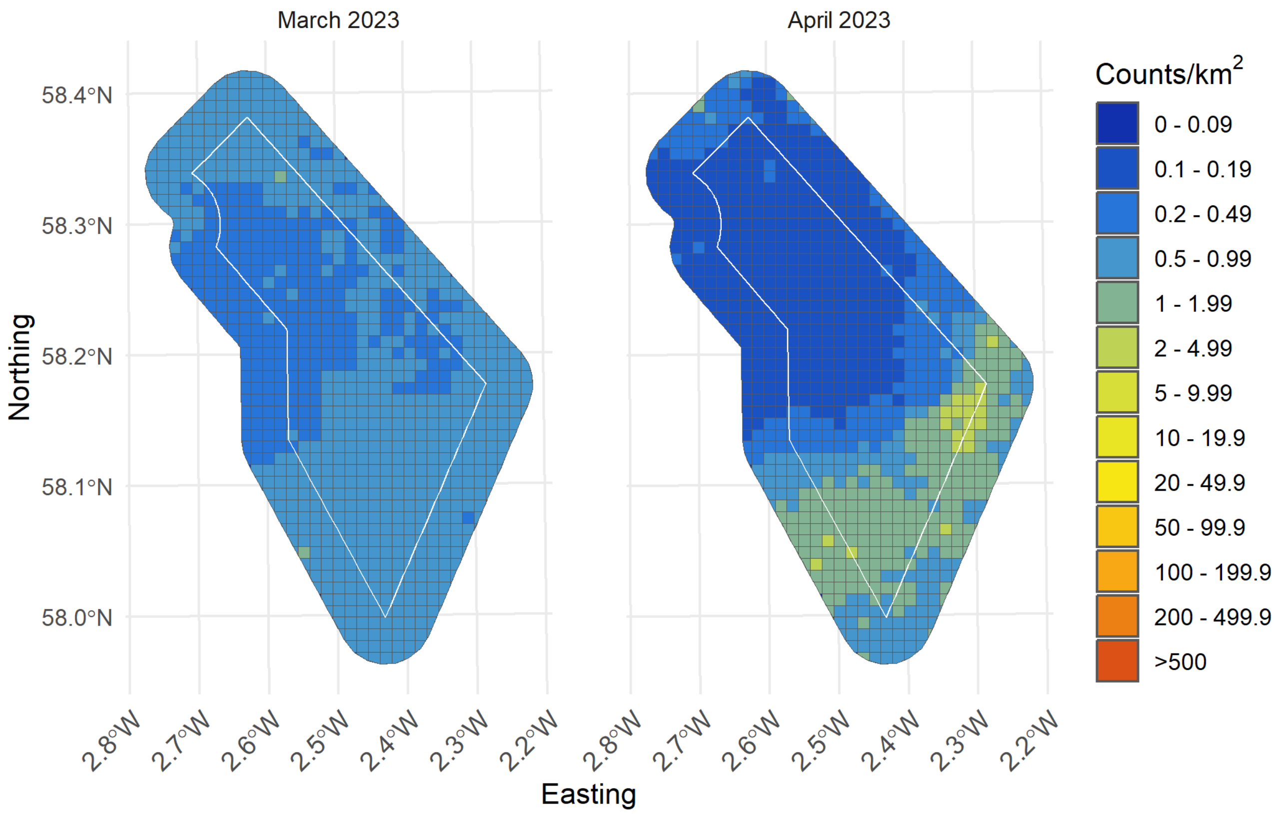


Figure II-70: Random forests baseline predictions of Razorbill from March 2023 to April 2023

II.5.5.2 Turbine scenario 1

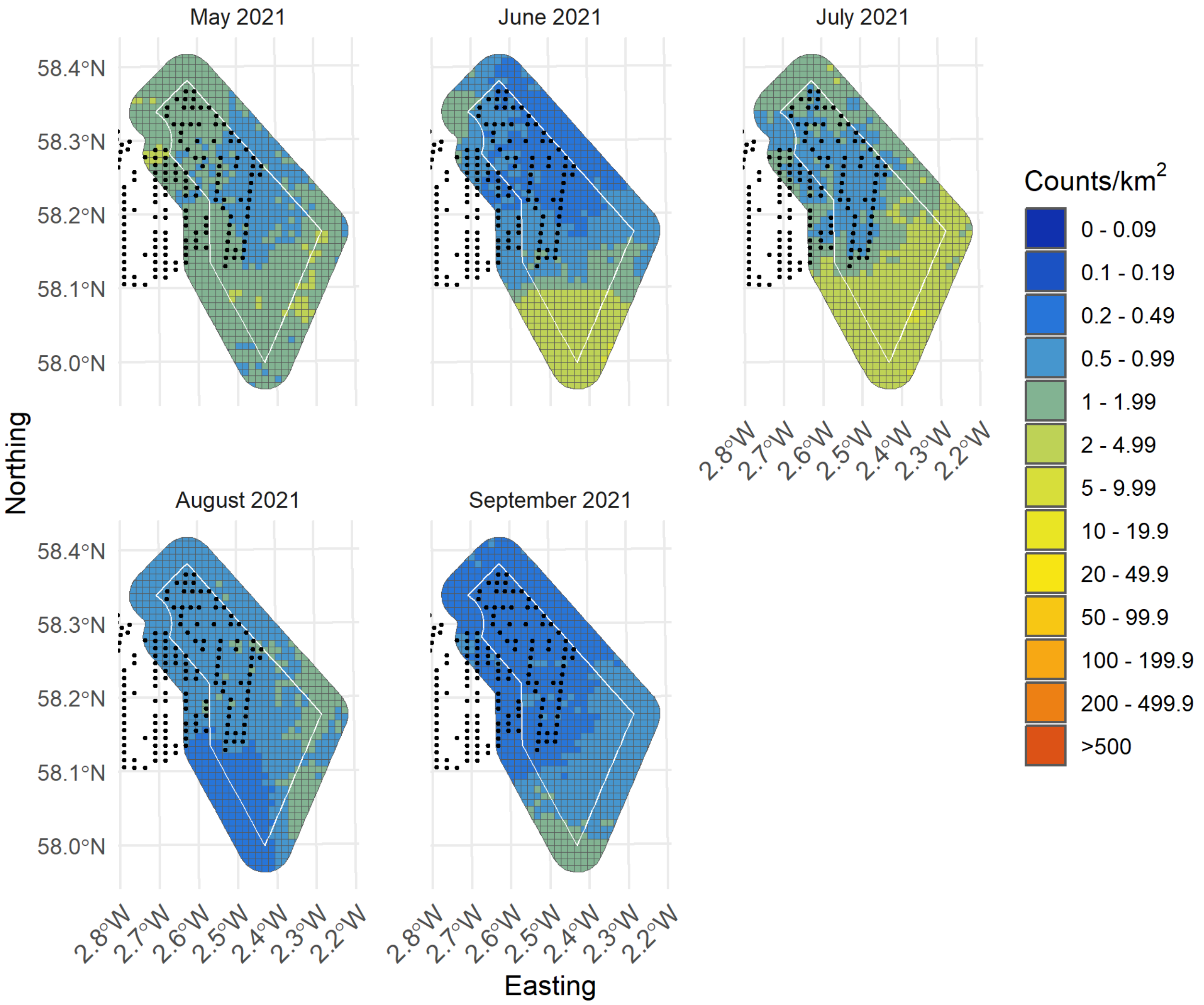


Figure II-71: Random forests predictions from turbine scenario 1 of Razorbill from May 2021 to September 2021 with turbines presented as black dots

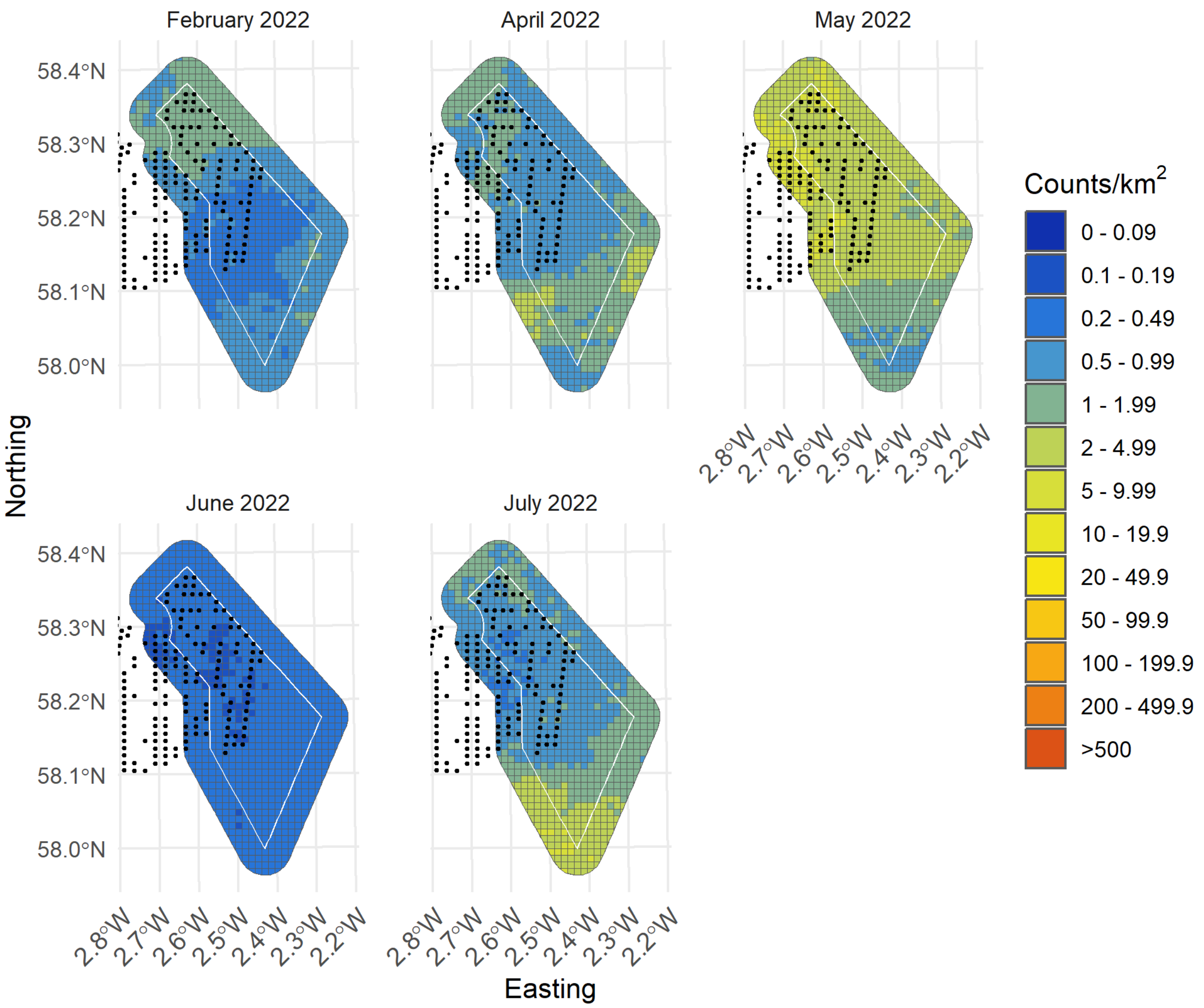


Figure II-72: Random forests predictions from turbine scenario 1 of Razorbill from February 2022 to July 2022 with turbines presented as black dots

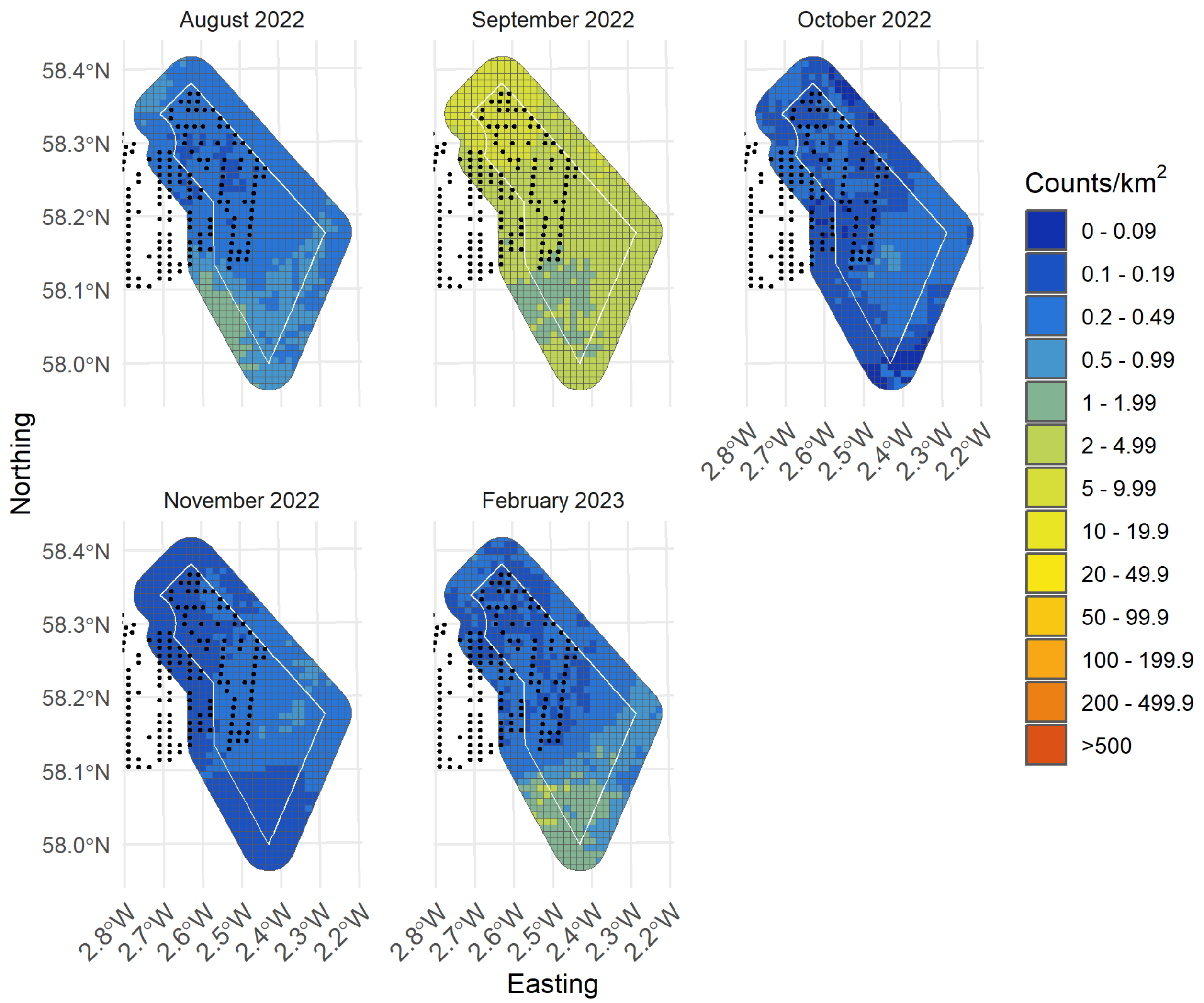


Figure II-73: Random forests predictions from turbine scenario 1 of Razorbill from August 2022 to February 2023 with turbines presented as black dots

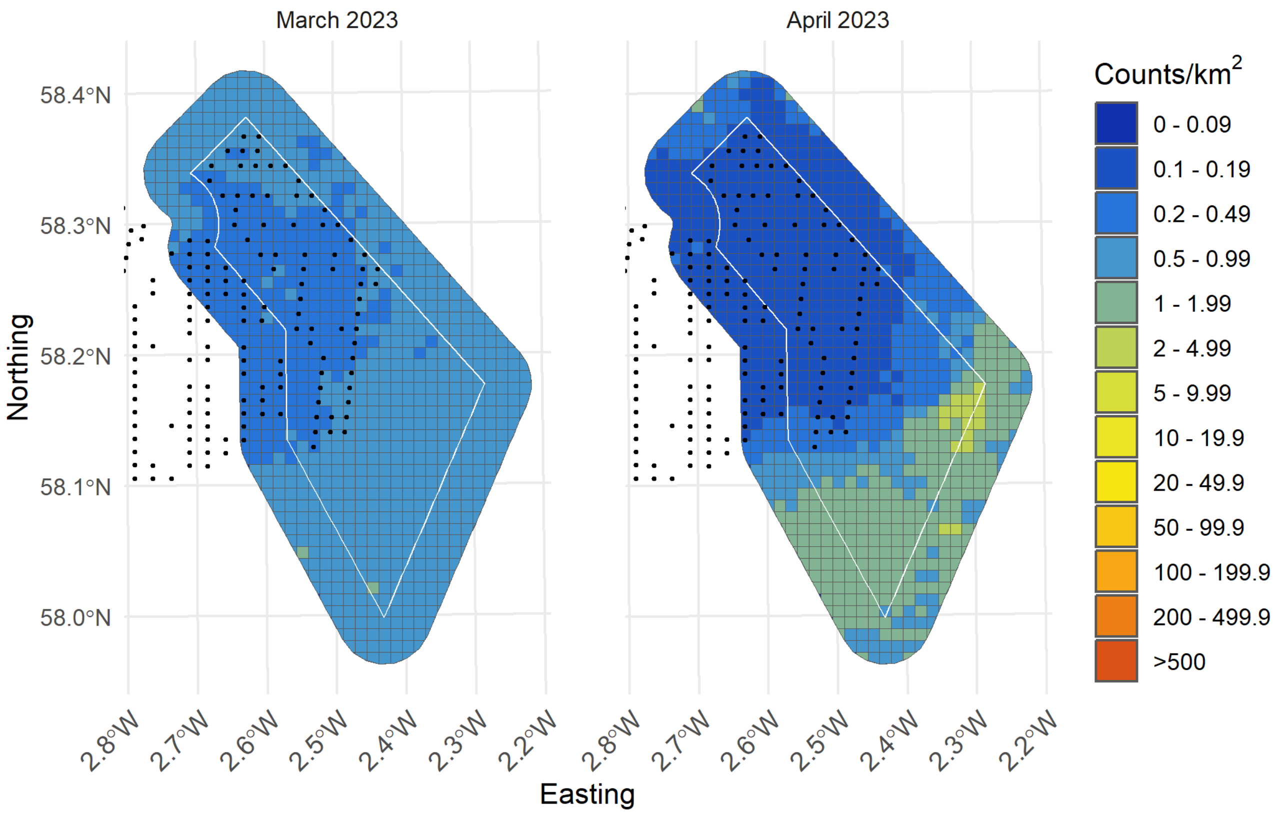


Figure II-74: Random forests predictions from turbine scenario 1 of Razorbill from March 2023 to April 2023 with turbines presented as black dots

II.5.5.3 Turbine scenario 2

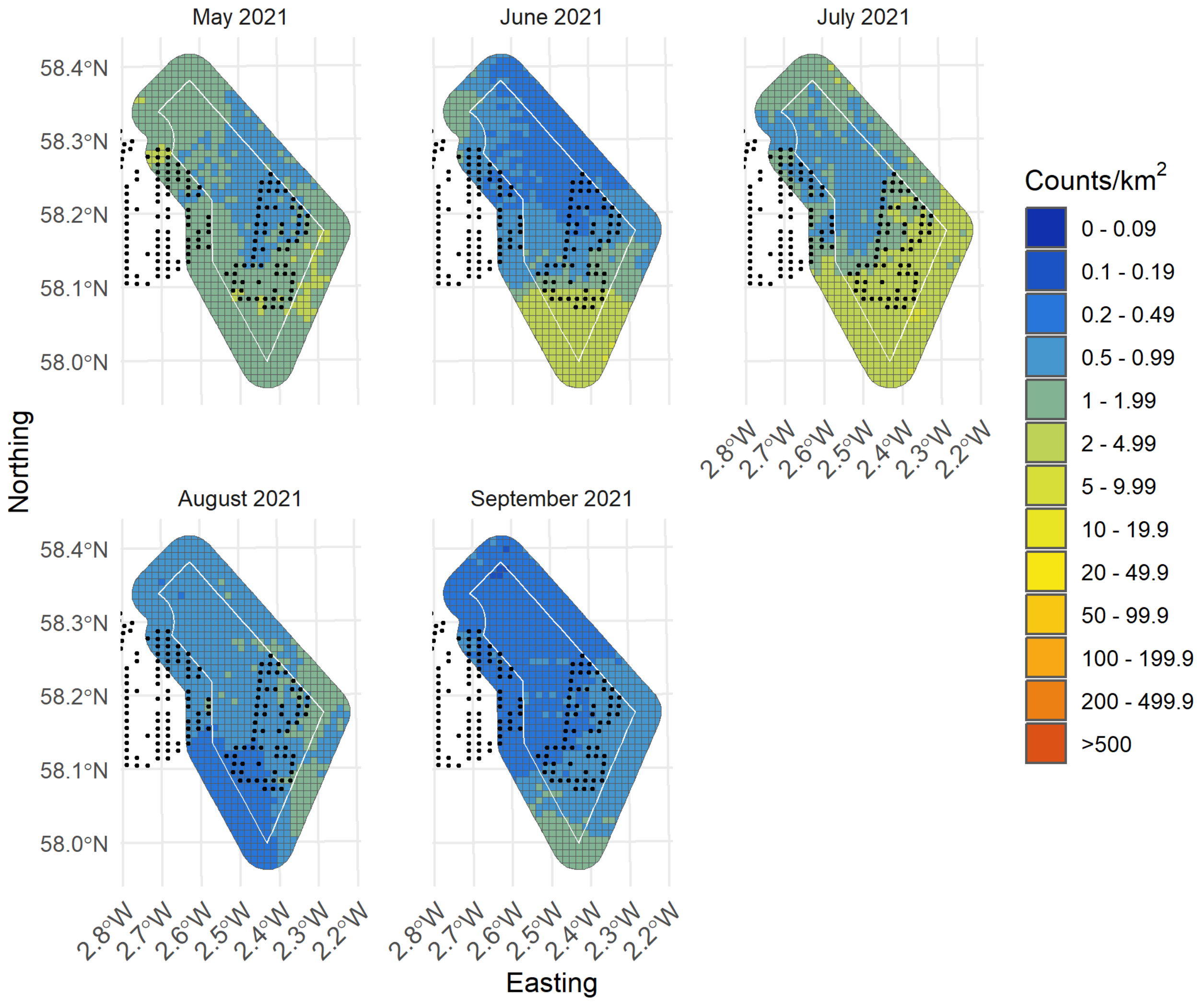


Figure II-75: Random forests predictions from turbine scenario 2 of Razorbill from May 2021 to September 2021 with turbines presented as black dots

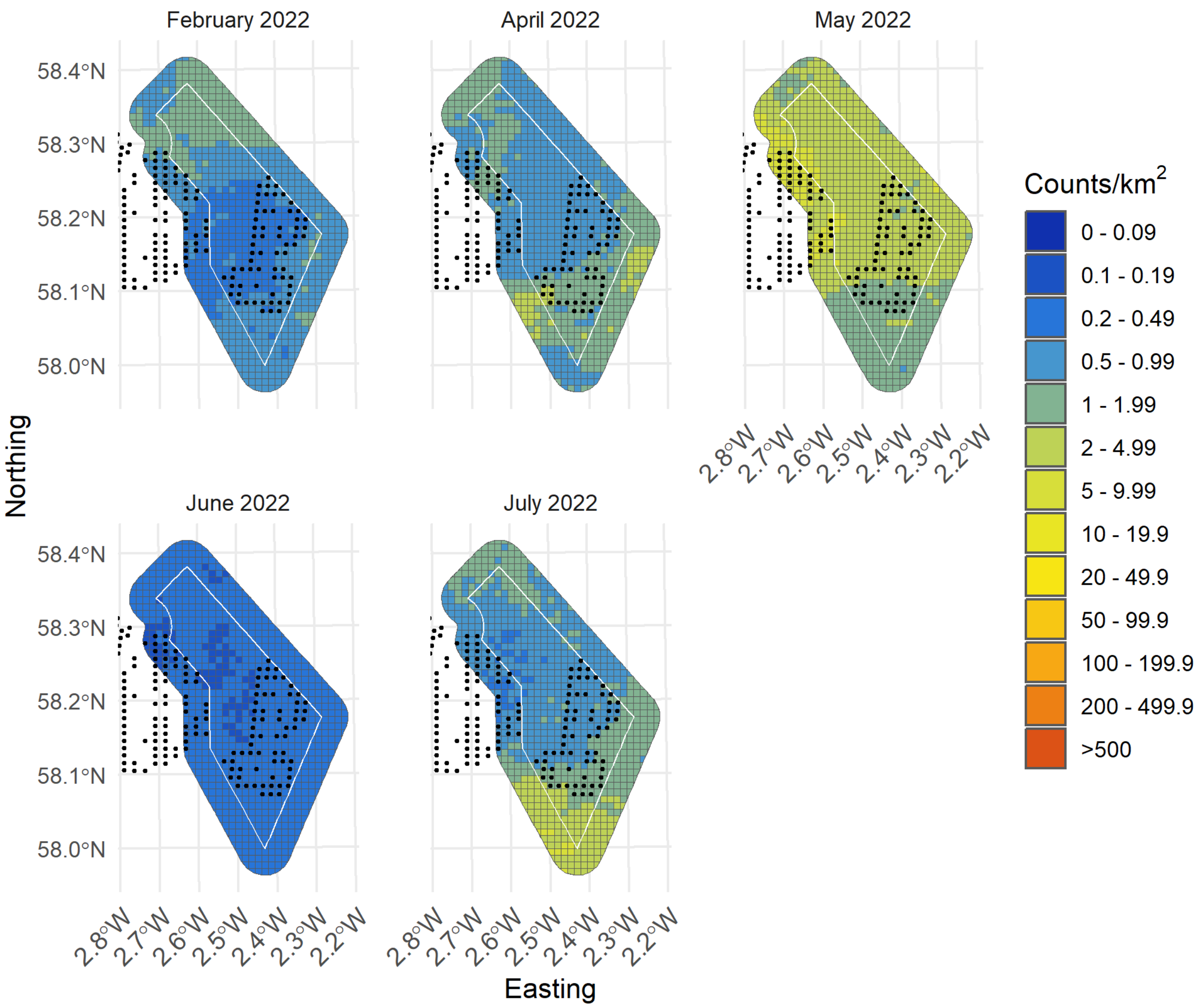


Figure II-76: Random forests predictions from turbine scenario 2 of Razorbill from February 2022 to July 2022 with turbines presented as black dots

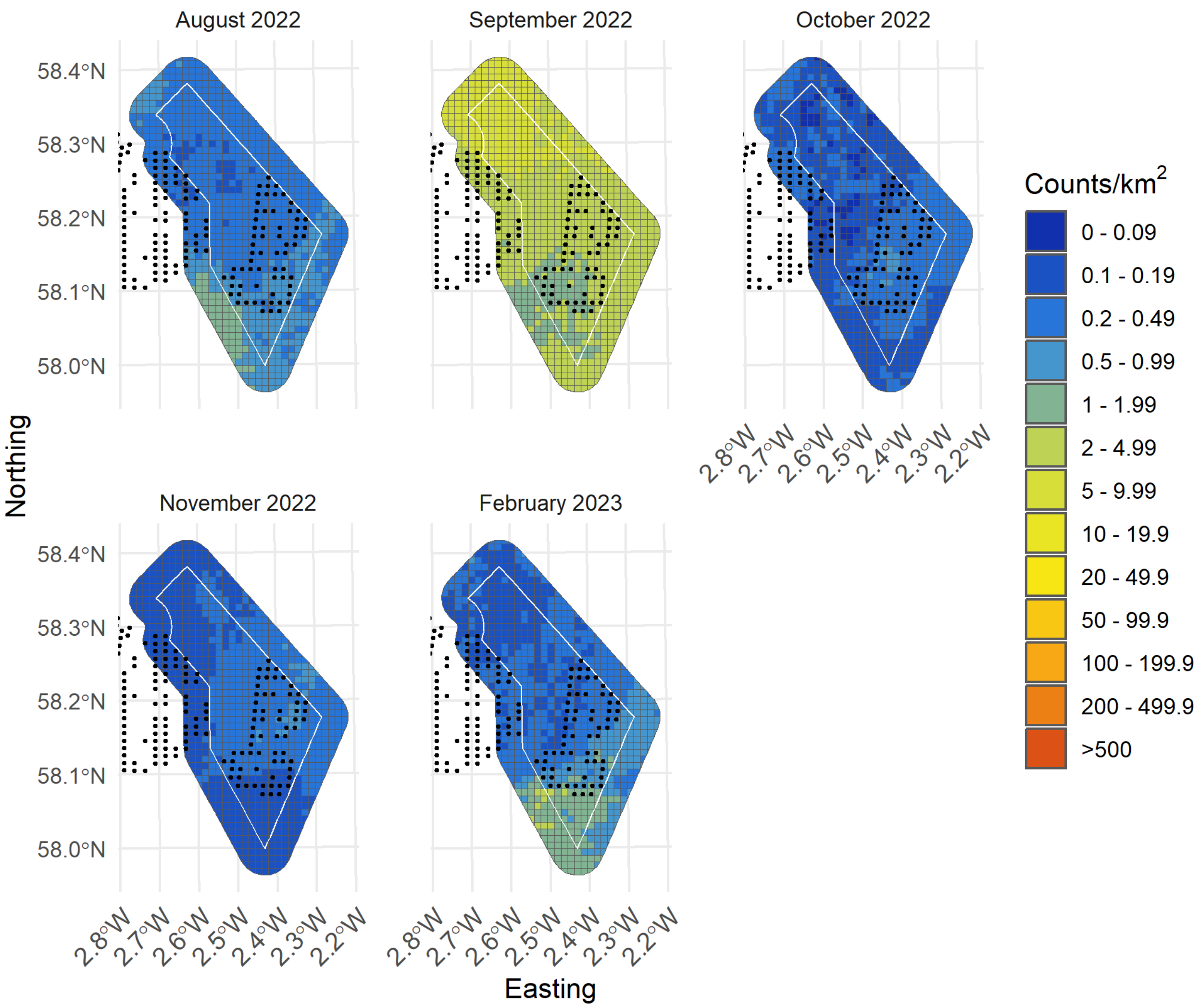


Figure II-77: Random forests predictions from turbine scenario 2 of Razorbill from August 2022 to February 2023 with turbines presented as black dots

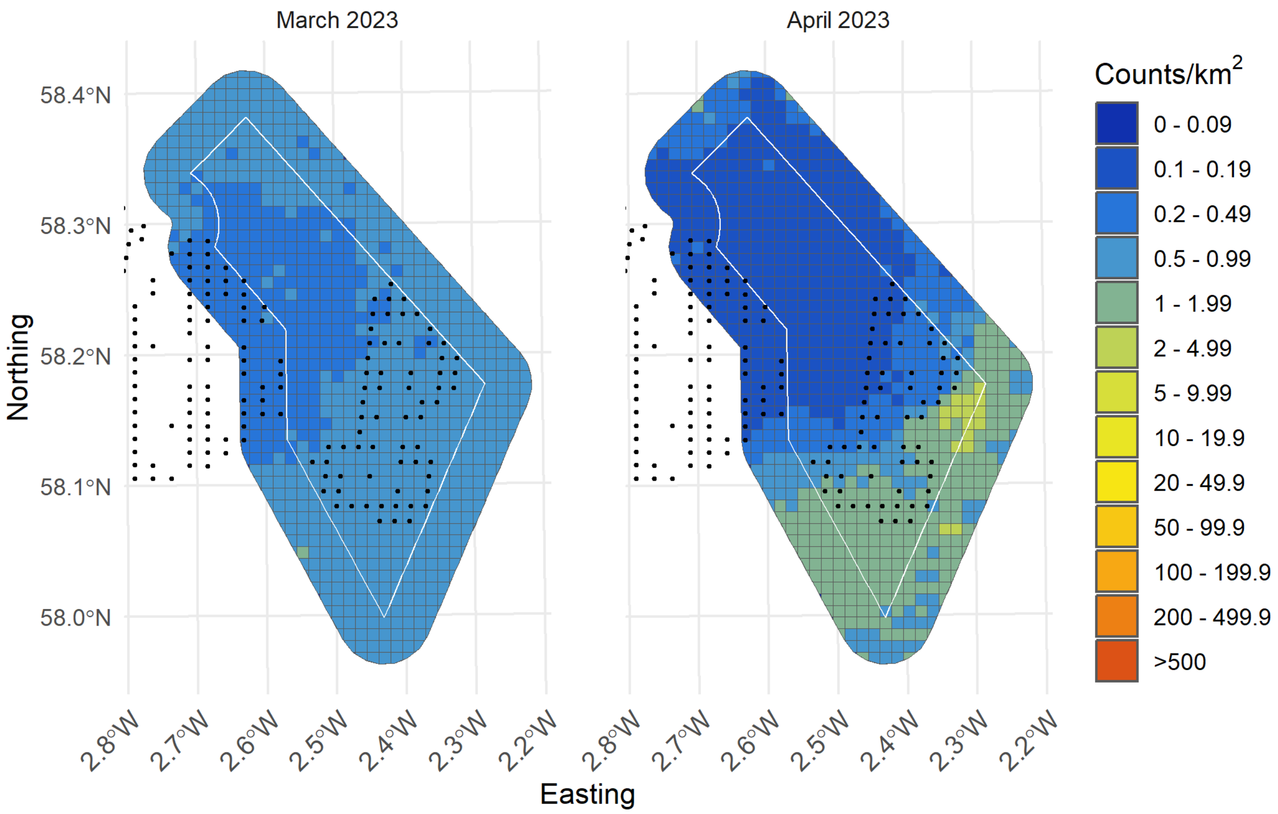


Figure II-78: Random forests predictions from turbine scenario 2 of Razorbill from March 2023 to April 2023 with turbines presented as black dots

II.5.5.4 Turbine scenario 3

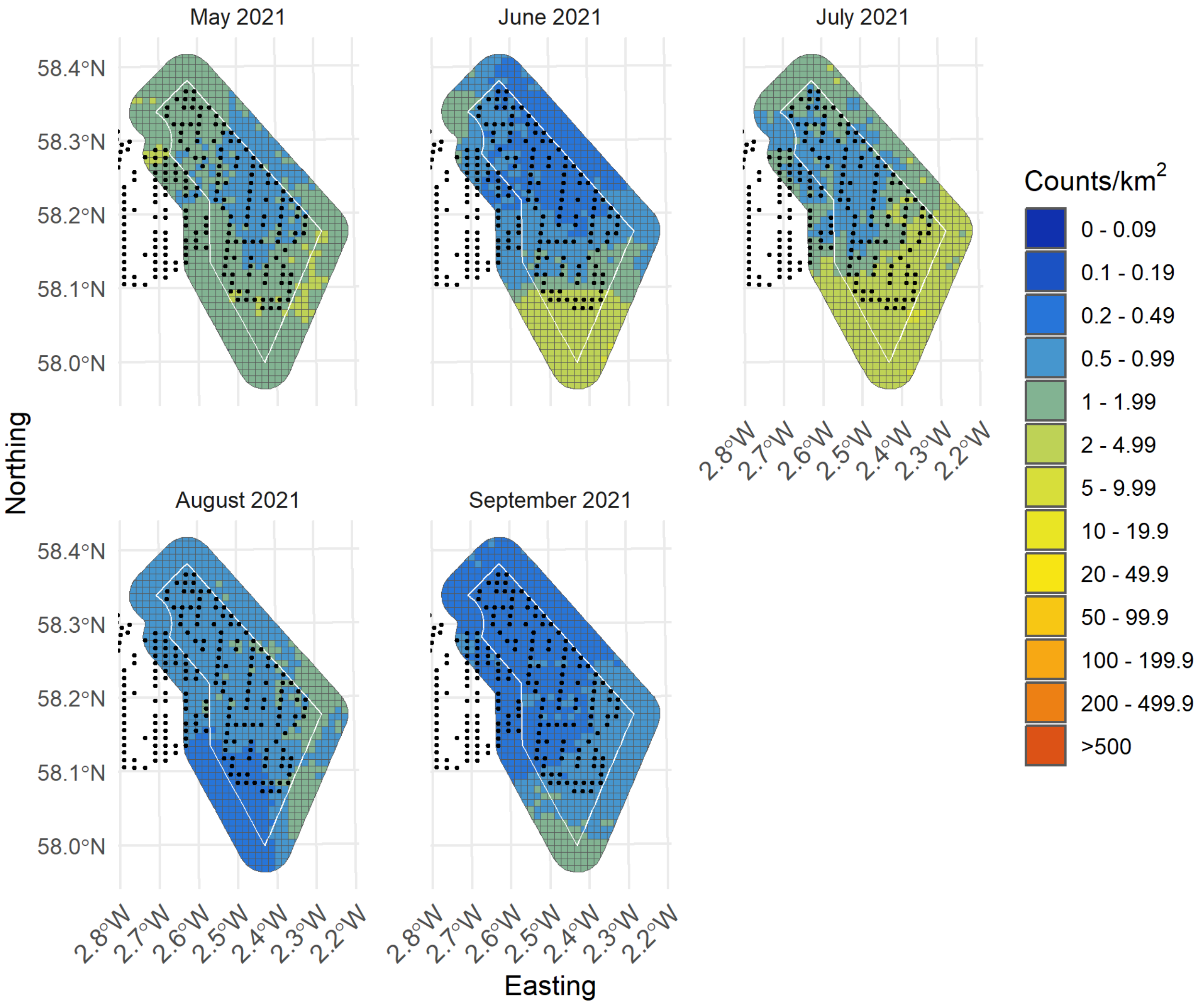


Figure II-79: Random forests predictions from turbine scenario 3 of Razorbill from May 2021 to September 2021 with turbines presented as black dots

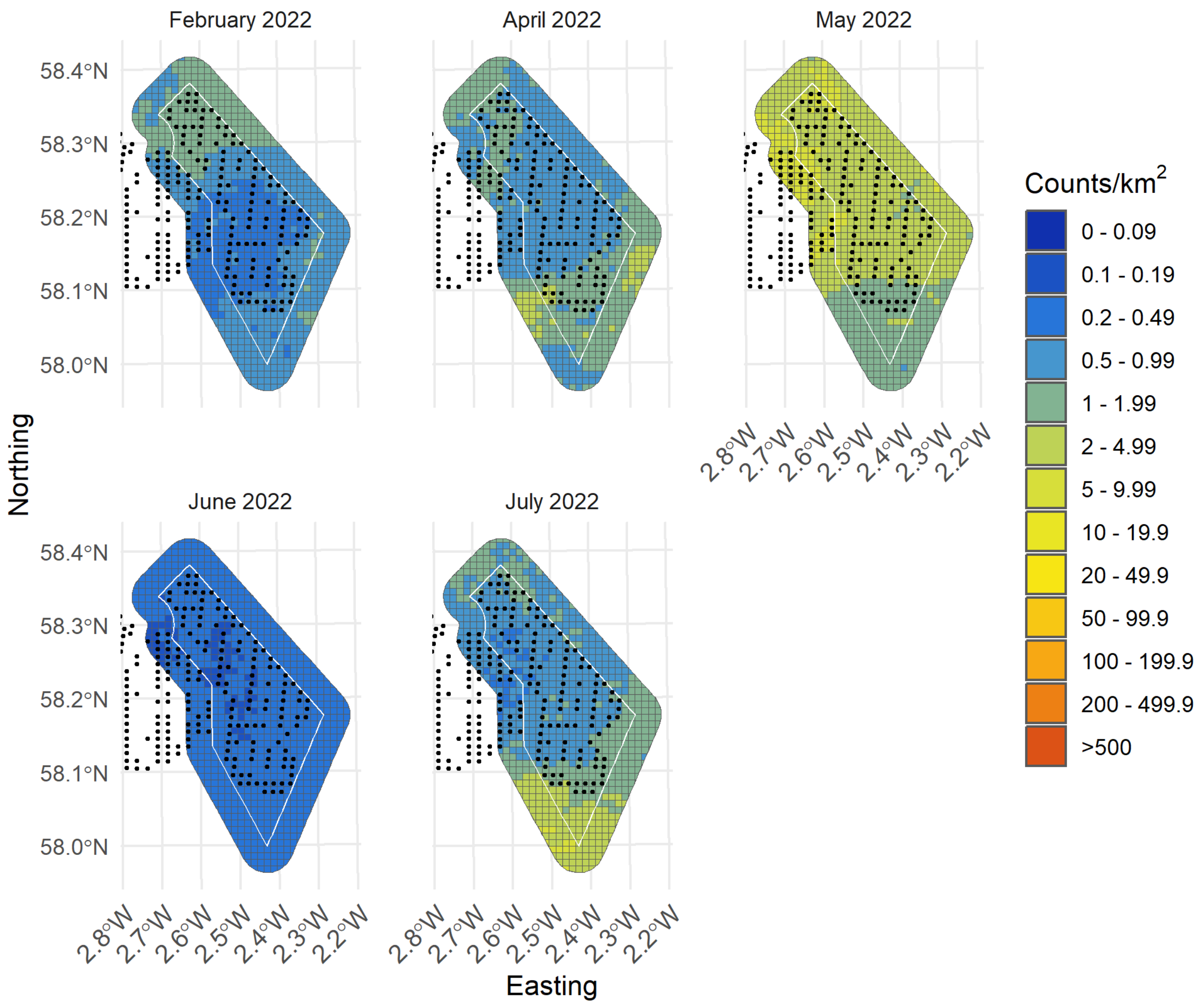


Figure II-80: Random forests predictions from turbine scenario 3 of Razorbill from February 2022 to July 2022 with turbines presented as black dots

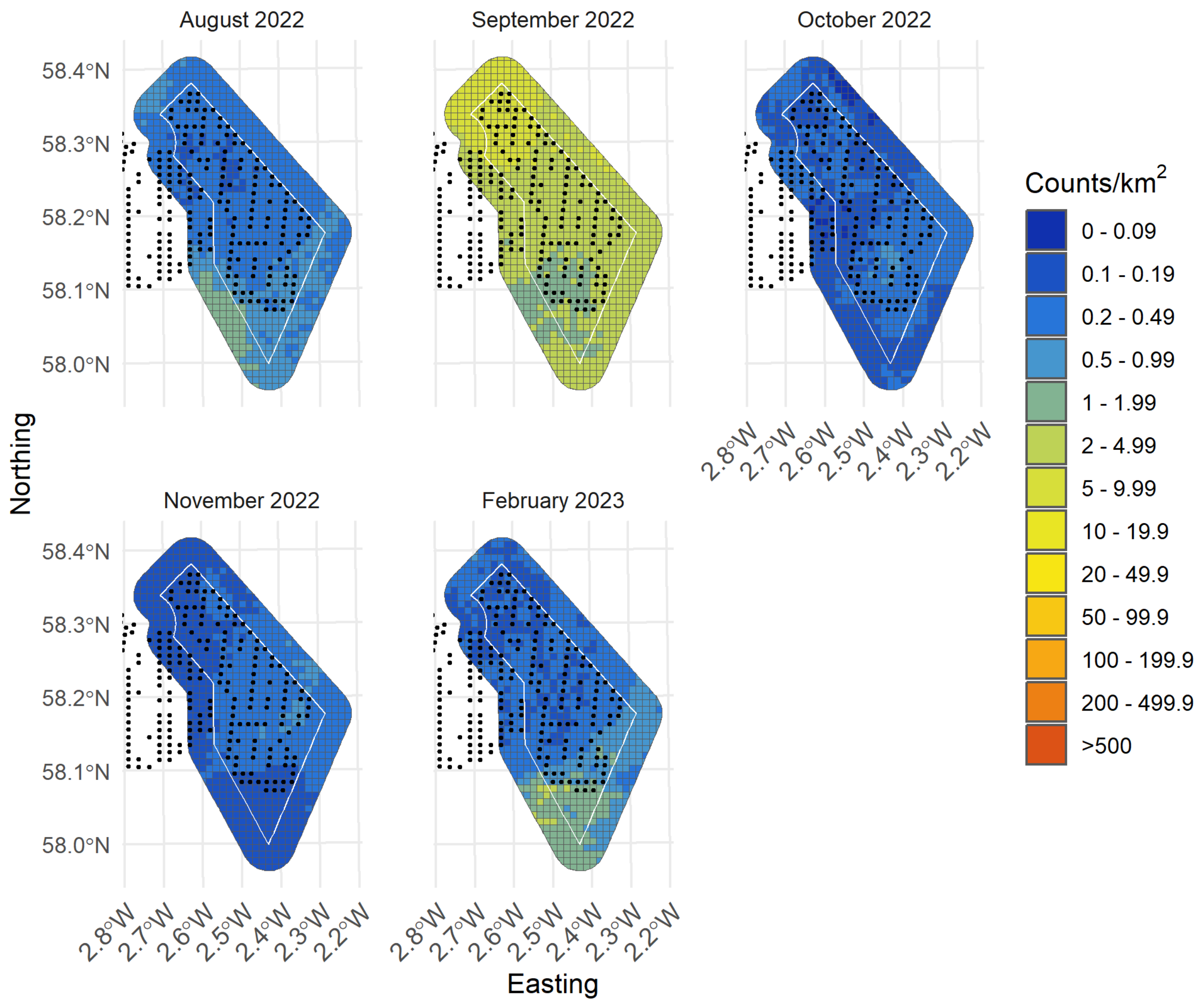


Figure II-81: Random forests predictions from turbine scenario 3 of Razorbill from August 2022 to February 2023 with turbines presented as black dots

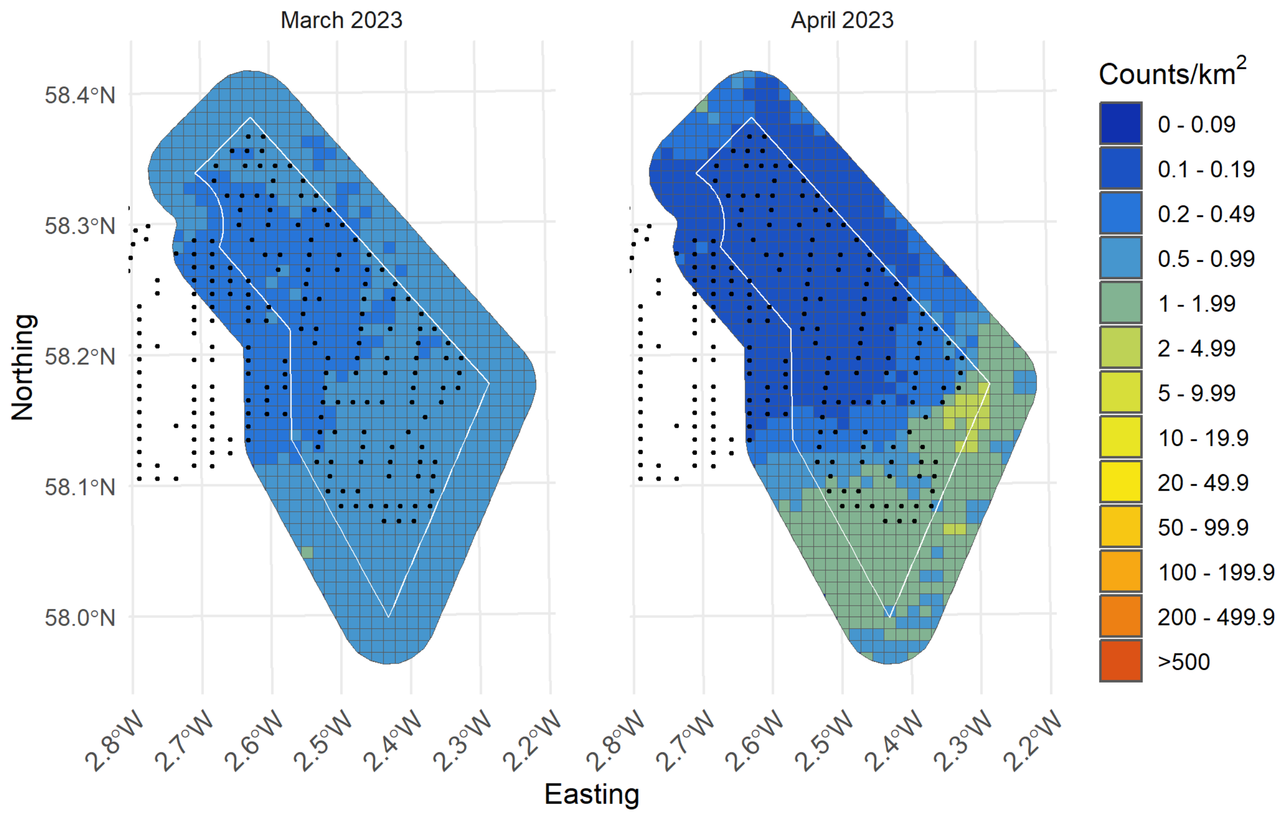


Figure II-82: Random forests predictions from turbine scenario 3 of Razorbill from March 2023 to April 2023 with turbines presented as black dots

II.6 Guillemot

II.6.1 Model assessment

The top five models had RMSE values between 33.386 and 33.095 with R squared values between 0.215 and 0.199. The R squared values were higher than those from the MRSea model, which had an R-squared value of 0.1843. The best model had an mtry value of 8, with a minimum node size of 80 (Figure II-83, Table II-16).

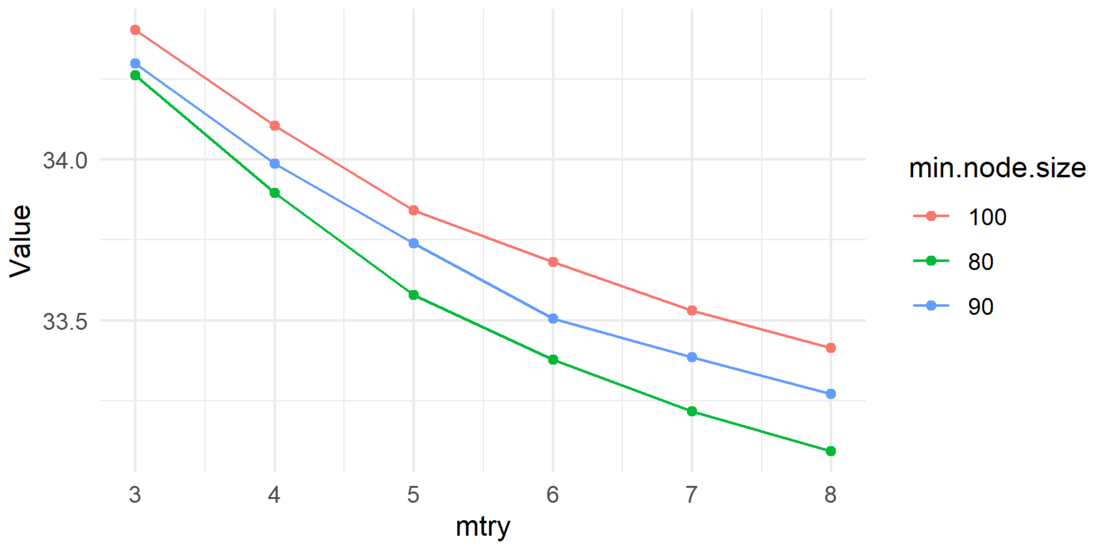


Figure II-83: Root mean squared error when varying mtry and min.node.size parameters in random forests

Table II-16: The top 5 models selected by the random forests analysis showing root mean squared error (RMSE) and R-squared values as calculated by 5-fold cross validation

mtry	min.node.size	RMSE	Rsquared
8	80	33.095	0.2146
7	80	33.218	0.2088
8	90	33.271	0.2048
6	80	33.378	0.2011
7	90	33.386	0.1986

II.6.2 Variable importance

The top predictor variables were northing (y.pos), survey ID, monthly mean sea surface temperature, and standard deviation of monthly sea surface temperature. This shows that most of the signal in the data come from variables that represent temporal variability (Table II-17).

Table II-17: The top 5 predictor variables from the random forests model and overall importance

Variable	Overall importance
y.pos	100.00
SurveyID^10	71.47
SST_mean	62.77
SurveyID^12	62.52
SST_sd	61.24

II.6.3 Distributional response

The partial relationship of distance to turbines on densities of guillemot demonstrates a slight inverse relationship between 0 - 10 km (Figure II-84). The relationship with the partial dependence from 0 to 20 km varies between 6.3 and 6.5, whereas the overall scale is from 6.3 to 7.6 which indicates that it is not as strong an effect within that x-axis range. If interpreting this as the figure depicts, there is relatively no distributional response associated with turbines for guillemot.

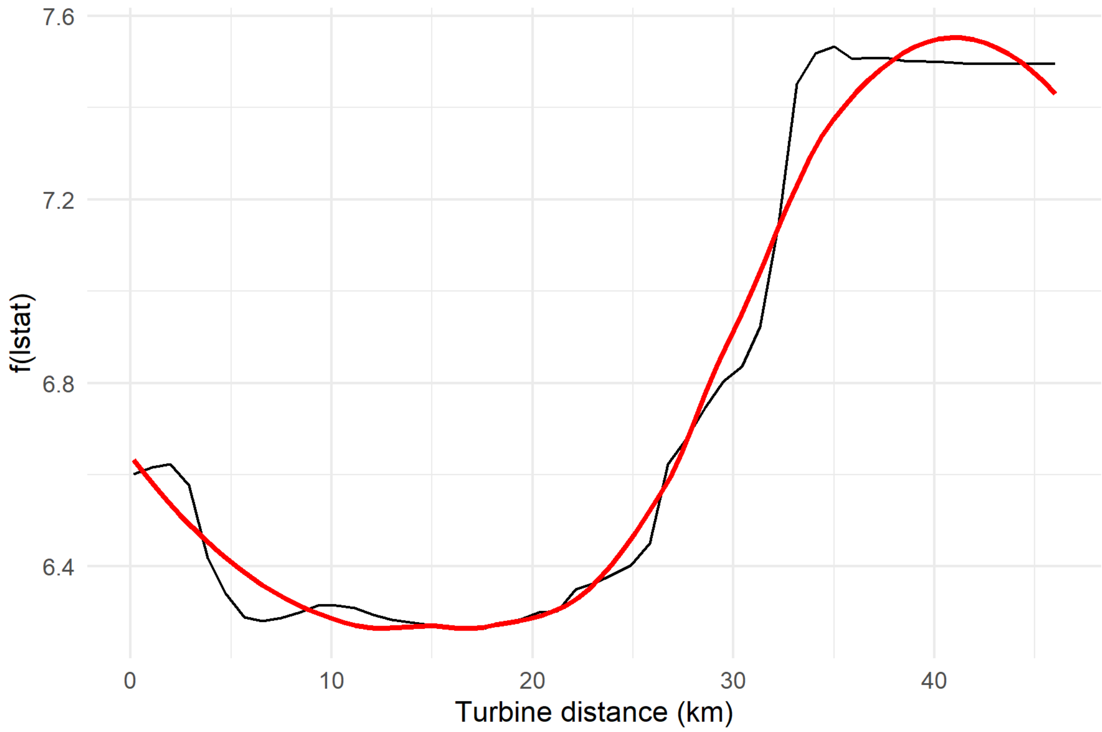


Figure II-84: Partial dependence plot of distance to turbine from the random forests model. The red line depicts the loess curve

II.6.4 Population estimates

Population estimates derived from the random forests baseline scenario (i.e., Existing turbines) mostly fell well within the range of estimates derived from MRSea, sometimes only differing by 2 - 3%. The largested discrepancy was the July 2021 survey which predicted a median population estimate 16704 birds, while the random forests analysis predicted 22587 birds. However, this estimate of 22587 fell within the confidence limits of the MRSea analysis. We are therefore confident that the random forests guillemot model is reliable, particularly in the context of the MRSea analysis.

Populations of guillemot in the survey area were mostly predicted to decrease in all three scenarios due to the relationship between distance to turbine and density, with the exception of May 2022, where the population was predicted to increase by 26.99%, and December 2022 where the population was predicted to increase by 11.94% in scenario 3. The biggest population change was predicted for scenario 3, with percent change from baseline varying from -19% to 27% across all surveys (Table II-18).

Table II-18: Population estimates from the random forests models for the 4km buffer zone plus windfarm footprint survey area for the baseline scenario (i.e., based on currently installed turbines) and three potential turbine scenarios. Percent change from baseline is calculated for each scenario

SurveyID	Baseline	Scenario 1 (% change)	Scenario 2 (% change)	Scenario 3 (% change)
May 2021	13202	13169 (-0.25)	13162 (-0.3)	13158 (-0.33)
June 2021	8588	8492 (-1.12)	8471 (-1.36)	8481 (-1.25)
July 2021	22587	21133 (-6.44)	21119 (-6.5)	21132 (-6.44)
August 2021	3408	3065 (-10.06)	3060 (-10.21)	3054 (-10.39)
September 2021	6125	5815 (-5.06)	5807 (-5.19)	5805 (-5.22)
October 2021	2250	2084 (-7.38)	2090 (-7.11)	2077 (-7.69)
November 2021	2799	2722 (-2.75)	2720 (-2.82)	2697 (-3.64)
December 2021	1452	1426 (-1.79)	1427 (-1.72)	1425 (-1.86)
January 2022	751	744 (-0.93)	739 (-1.6)	731 (-2.66)
February 2022	1980	2095 (5.81)	2095 (5.81)	2076 (4.85)
March 2022	1114	1120 (0.54)	1122 (0.72)	1116 (0.18)
April 2022	2952	2752 (-6.78)	2758 (-6.57)	2747 (-6.94)
May 2022	20986	25207 (20.11)	24726 (17.82)	26650 (26.99)
June 2022	14410	13994 (-2.89)	13930 (-3.33)	13862 (-3.8)
July 2022	17230	16888 (-1.98)	16726 (-2.93)	16716 (-2.98)
August 2022	1142	1098 (-3.85)	1094 (-4.2)	1091 (-4.47)
September 2022	13217	10686 (-19.15)	10683 (-19.17)	10676 (-19.23)
October 2022	1100	1086 (-1.27)	1074 (-2.36)	1092 (-0.73)
November 2022	189	186 (-1.59)	191 (1.06)	192 (1.59)
December 2022	653	728 (11.49)	726 (11.18)	731 (11.94)
January 2023	998	986 (-1.2)	988 (-1)	979 (-1.9)
February 2023	2843	2969 (4.43)	2966 (4.33)	2936 (3.27)
March 2023	3446	3825 (11)	3843 (11.52)	3847 (11.64)
April 2023	3182	3120 (-1.95)	3099 (-2.61)	3096 (-2.7)

II.6.5 Distributions

The broad distribution of guillemot in the baseline random forests models were nearly identical to those in the MRSea models. For most surveys, the models predicted the highest densities in the south western parts of the survey area, although the species was mostly ubiquitous throughout the site in all surveys. In nearly all cases, the densities of guillemot decrease through the site in relation to the locations of proposed turbines. This distributional response suggests a potential displacement effect which could push individuals beyond the boundary of the wind farm site (Figures II-85 – II-104).

II.6.5.1 Baseline scenarios

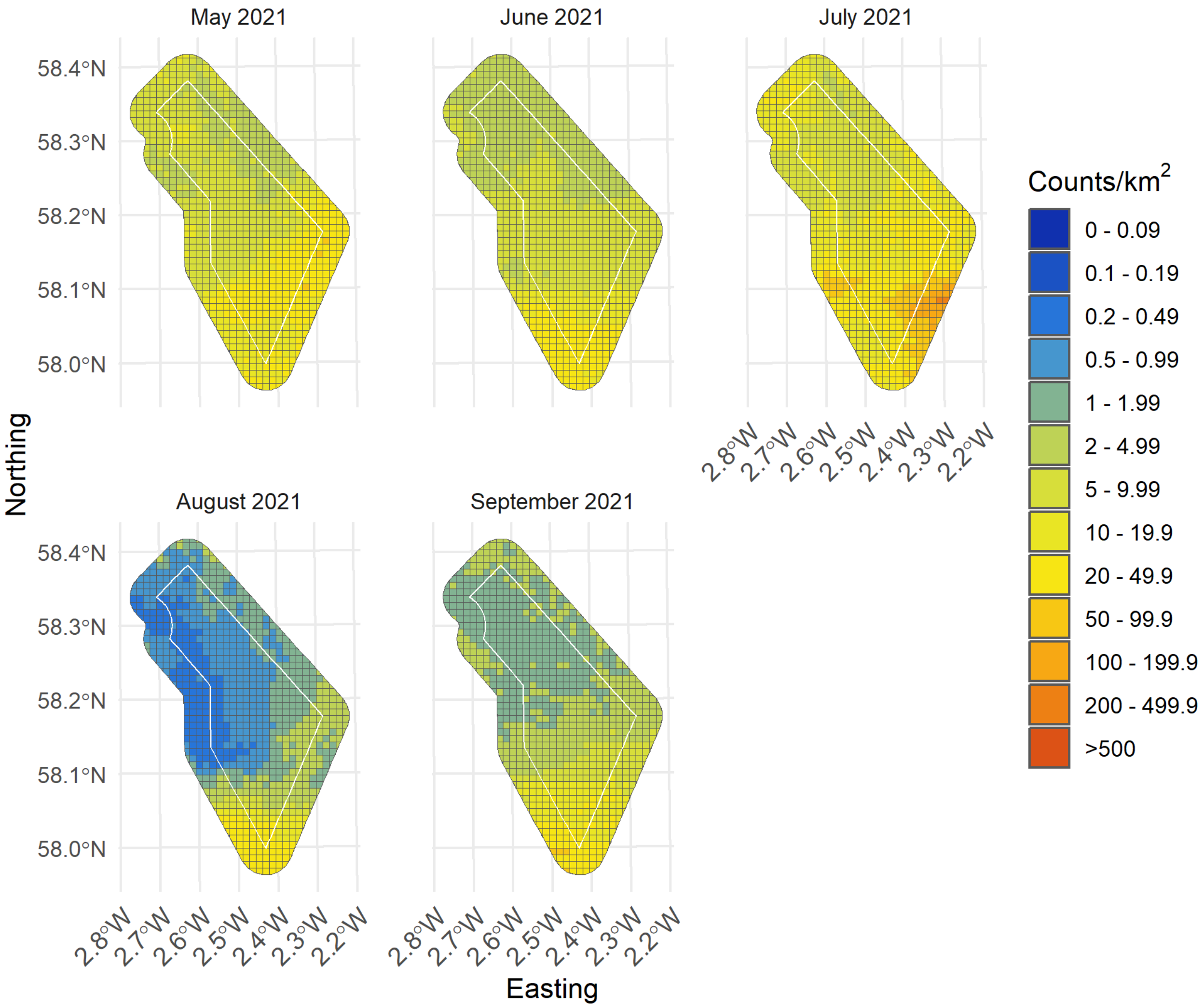


Figure II-85: Random forests baseline predictions of Guillemot from May 2021 to September 2021

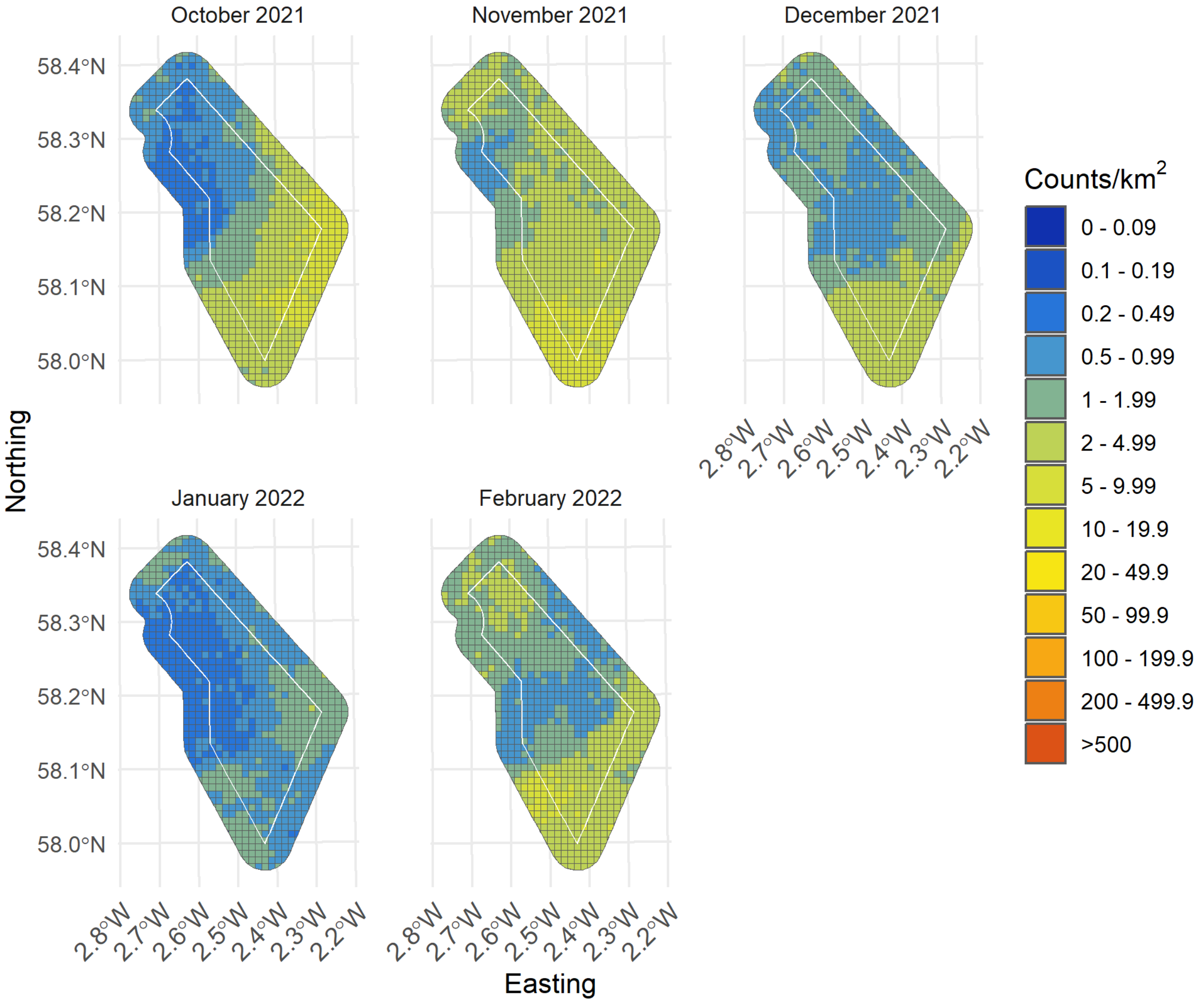


Figure II-86: Random forests baseline predictions of Guillemot from October 2021 to February 2022

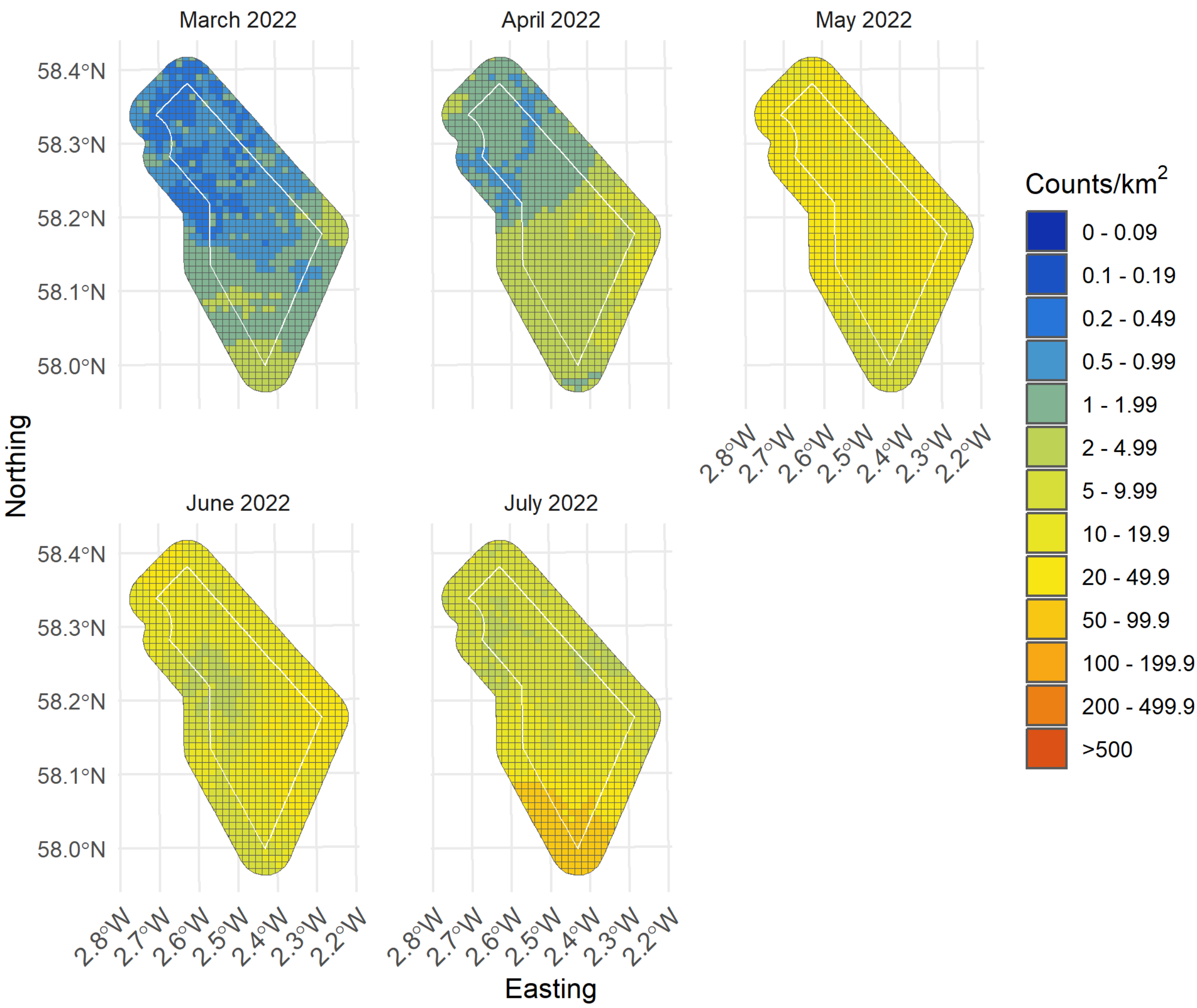


Figure II-87: Random forests baseline predictions of Guillemot from March 2022 to July 2022

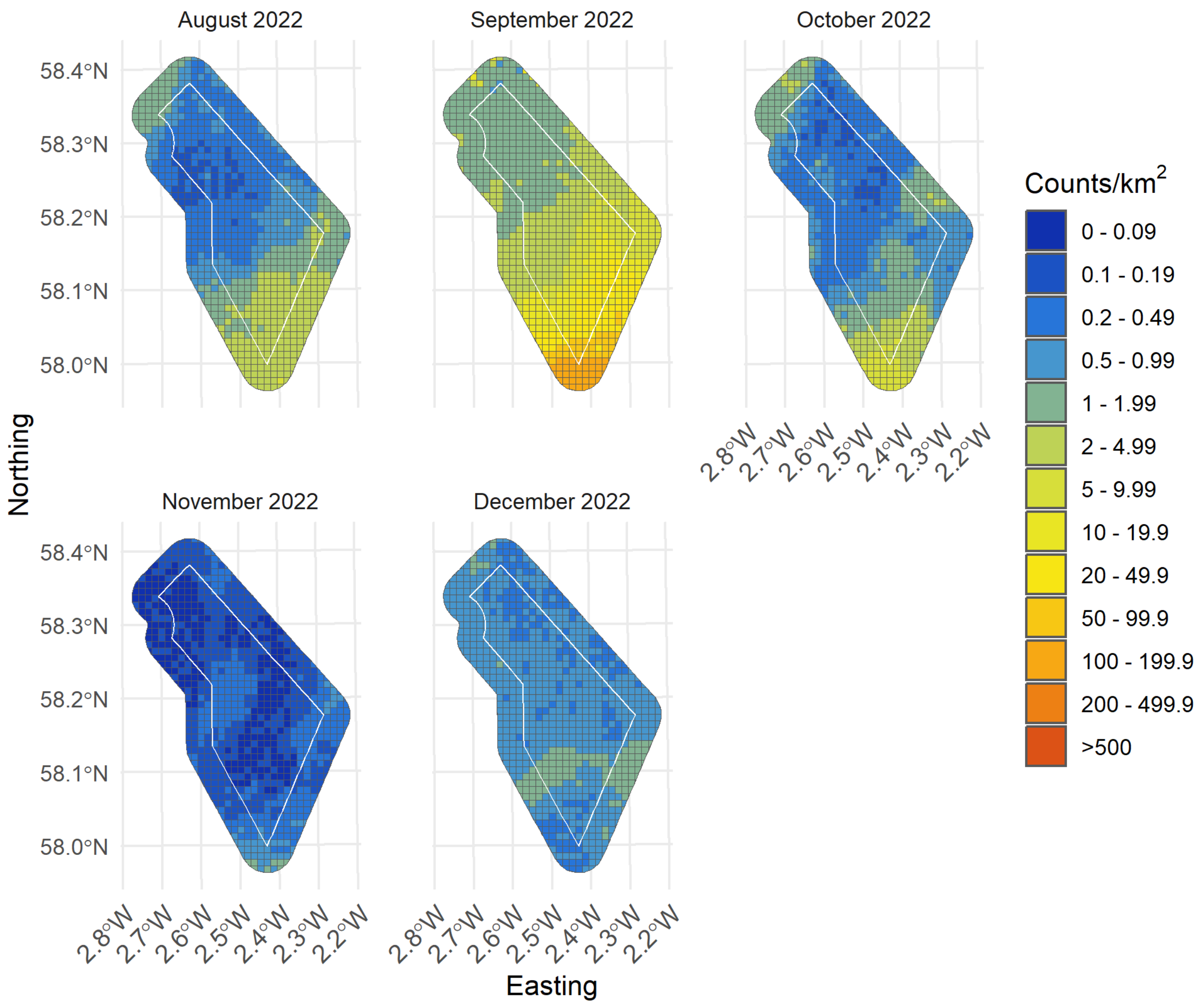


Figure II-88: Random forests baseline predictions of Guillemot from August 2022 to December 2022

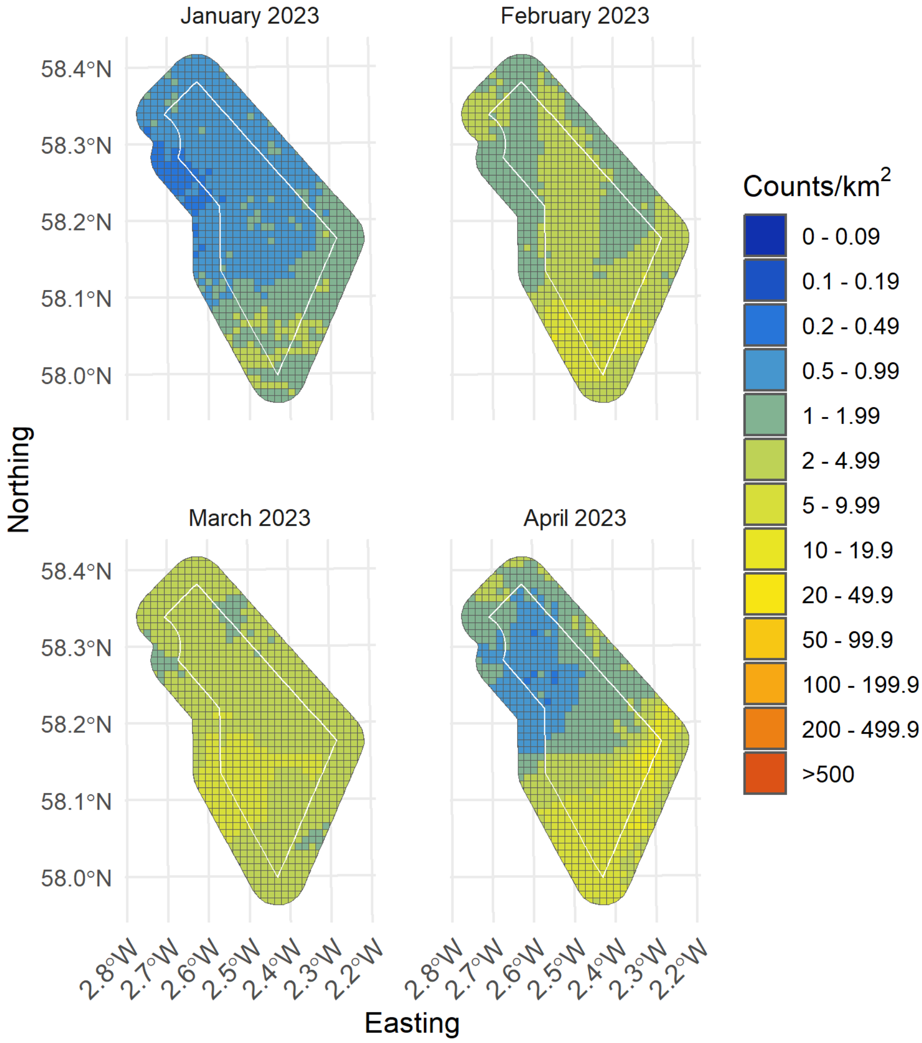


Figure II-89: Random forests baseline predictions of Guillemot from January 2023 to April 2023

II.6.5.2 Turbine scenario 1

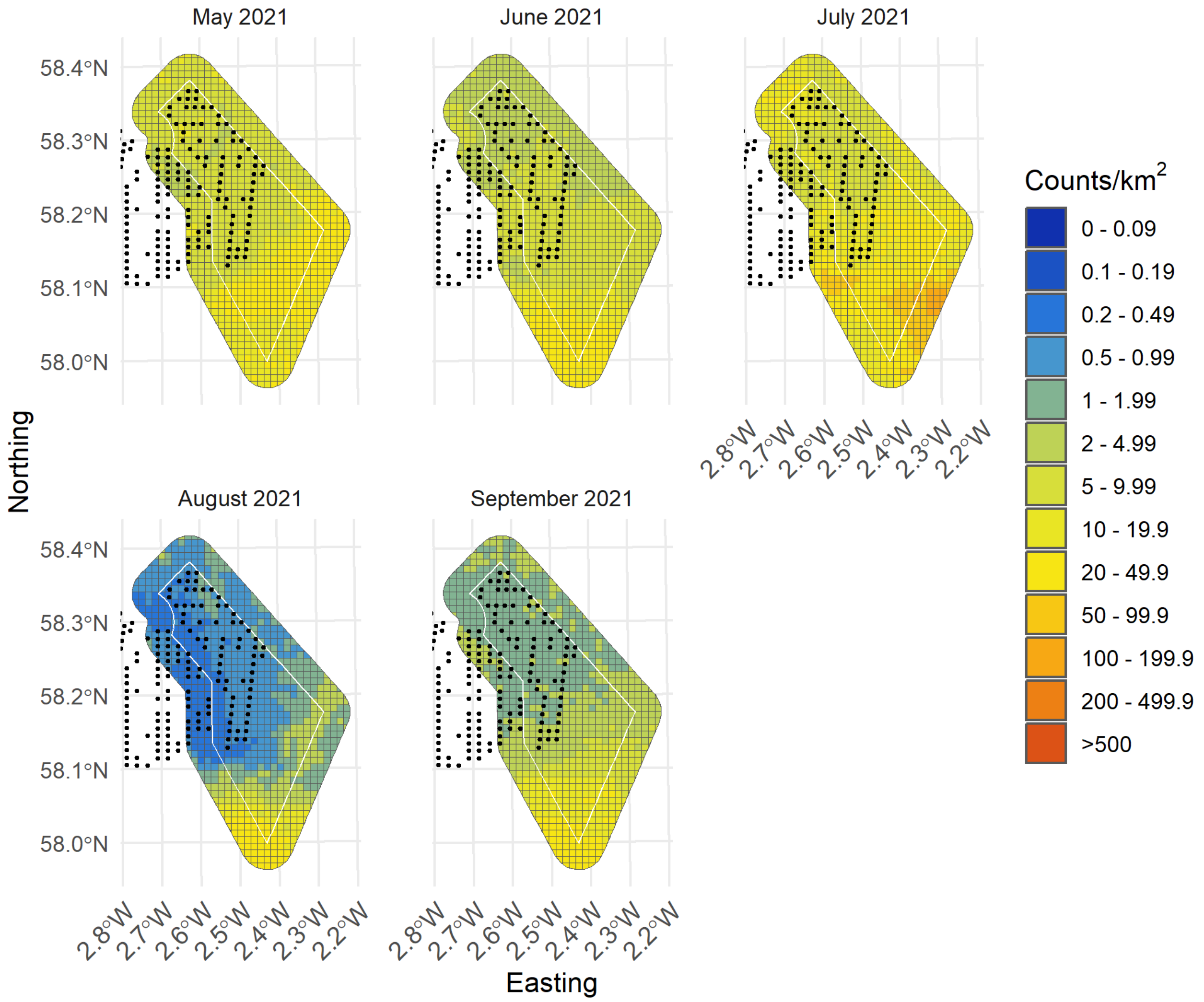


Figure II-90: Random forests predictions from turbine scenario 1 of Guillemot from May 2021 to September 2021 with turbines presented as black dots

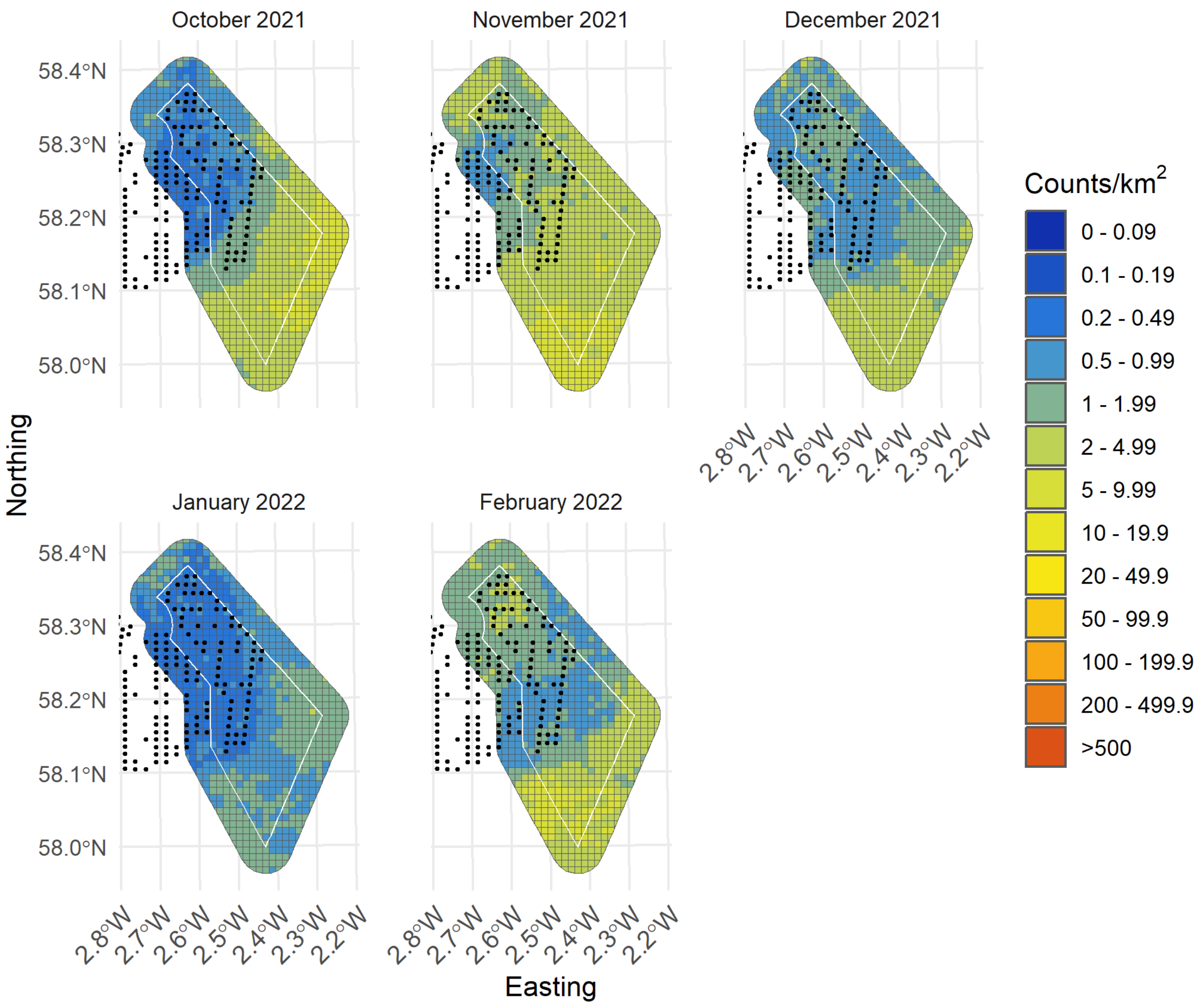


Figure II-91: Random forests predictions from turbine scenario 1 of Guillemot from October 2021 to February 2022 with turbines presented as black dots

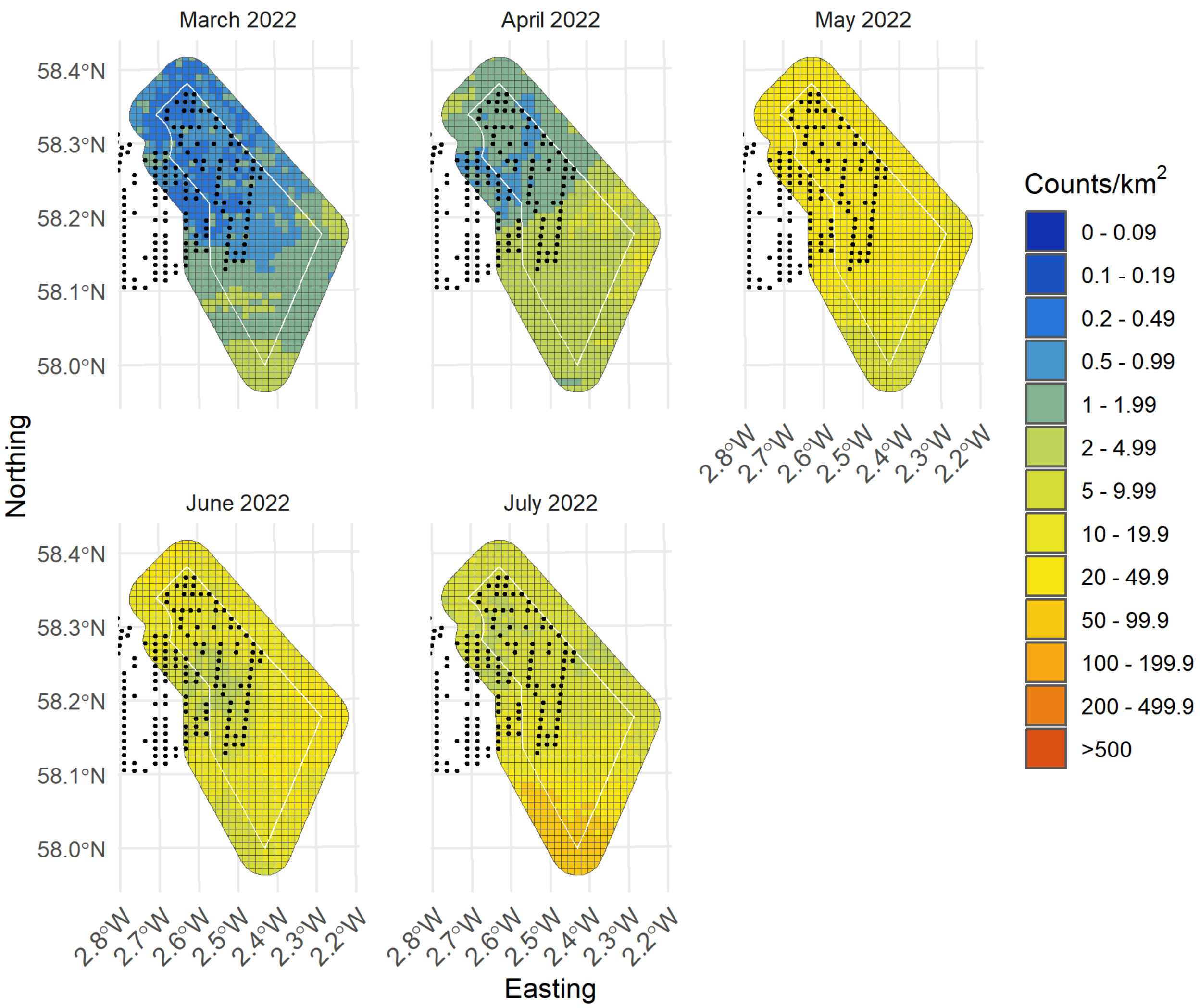


Figure II-92: Random forests predictions from turbine scenario 1 of Guillemot from March 2022 to July 2022 with turbines presented as black dots

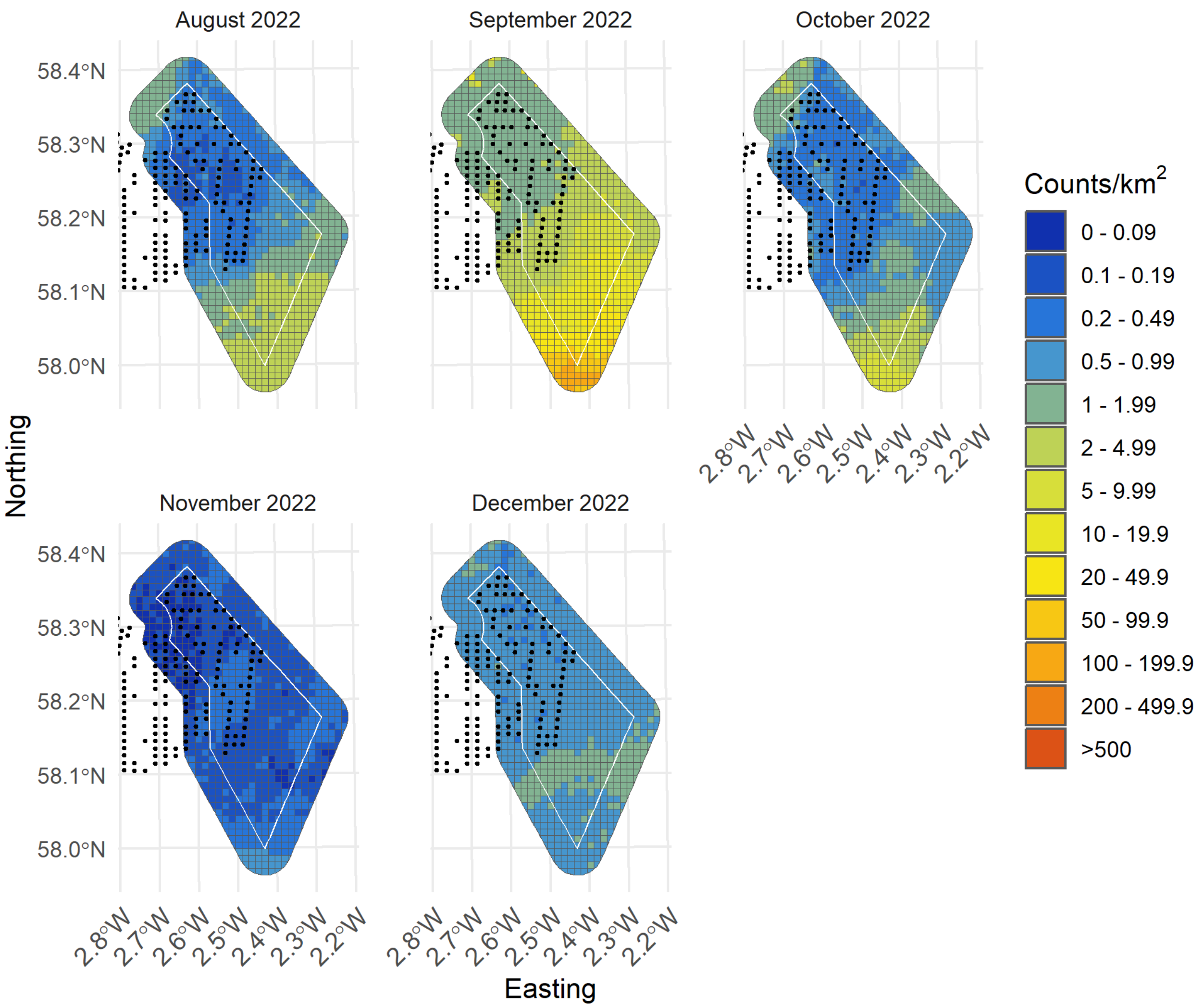


Figure II-93: Random forests predictions from turbine scenario 1 of Guillemot from August 2022 to December 2022 with turbines presented as black dots

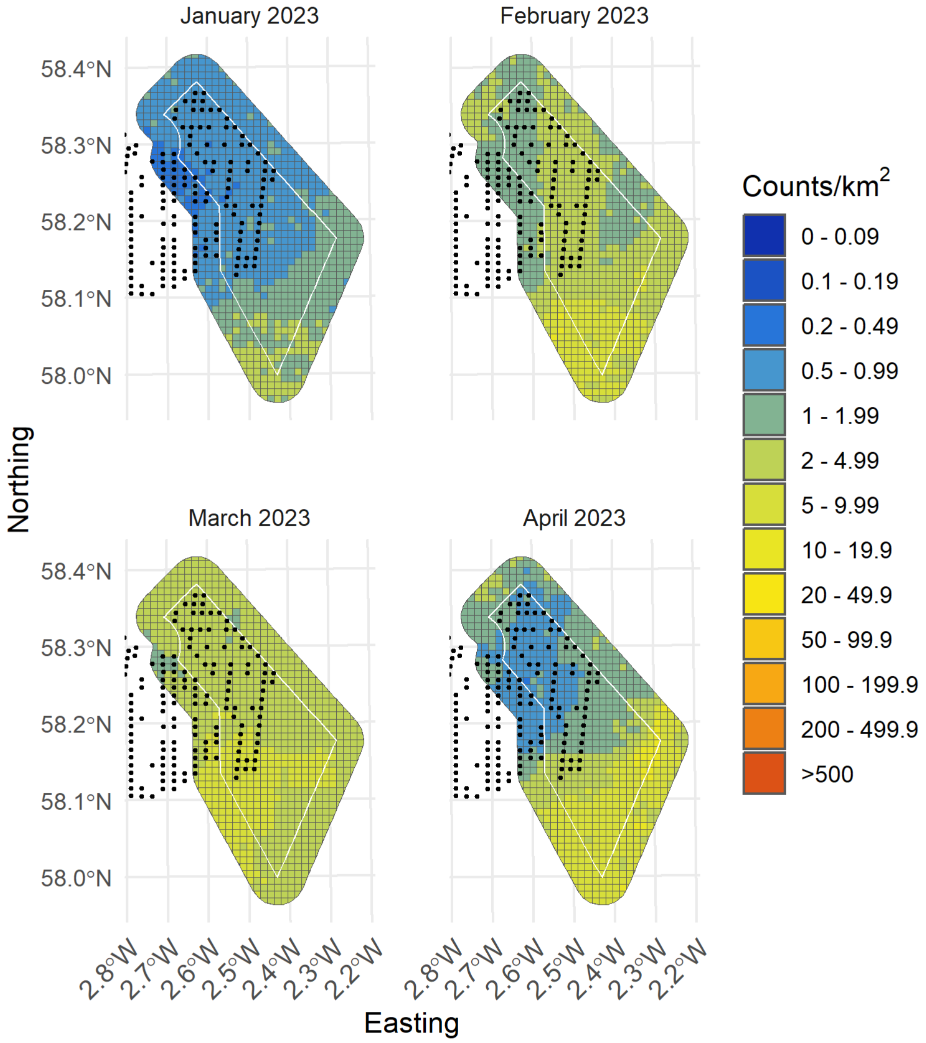


Figure II-94: Random forests predictions from turbine scenario 1 of Guillemot from January 2023 to April 2023 with turbines presented as black dots

II.6.5.3 Turbine scenario 2

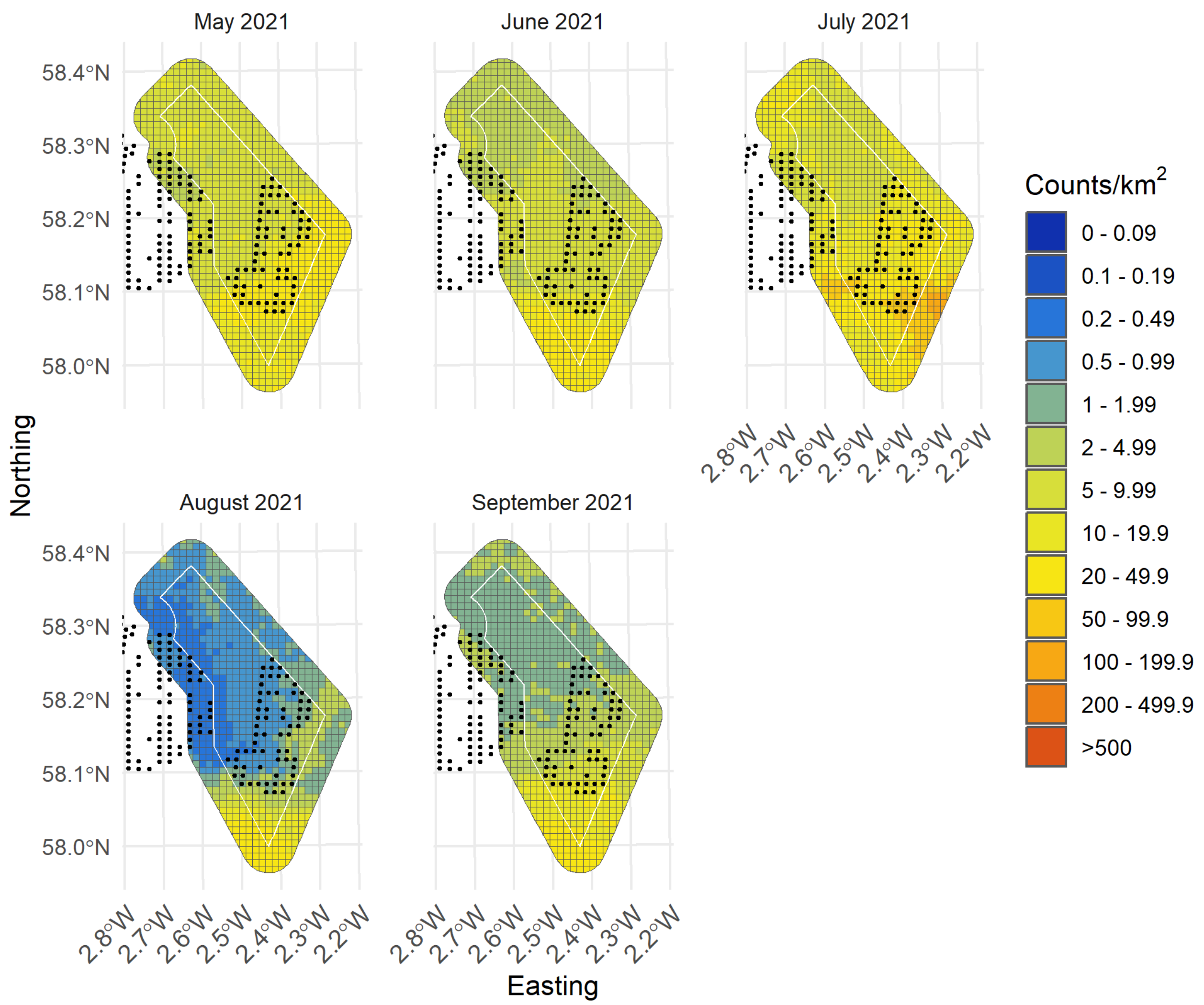


Figure II-95: Random forests predictions from turbine scenario 2 of Guillemot from May 2021 to September 2021 with turbines presented as black dots

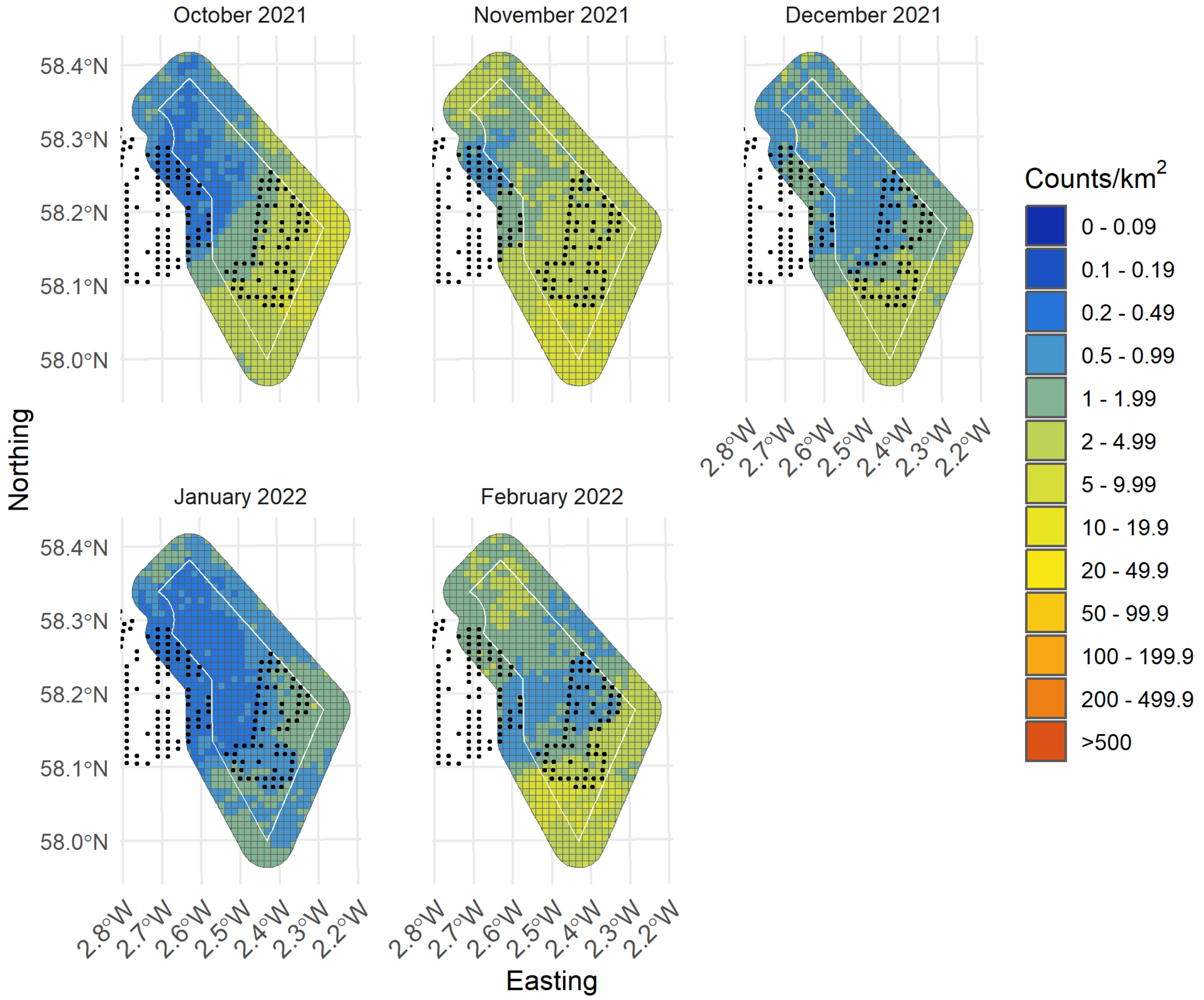


Figure II-96: Random forests predictions from turbine scenario 2 of Guillemot from October 2021 to February 2022 with turbines presented as black dots

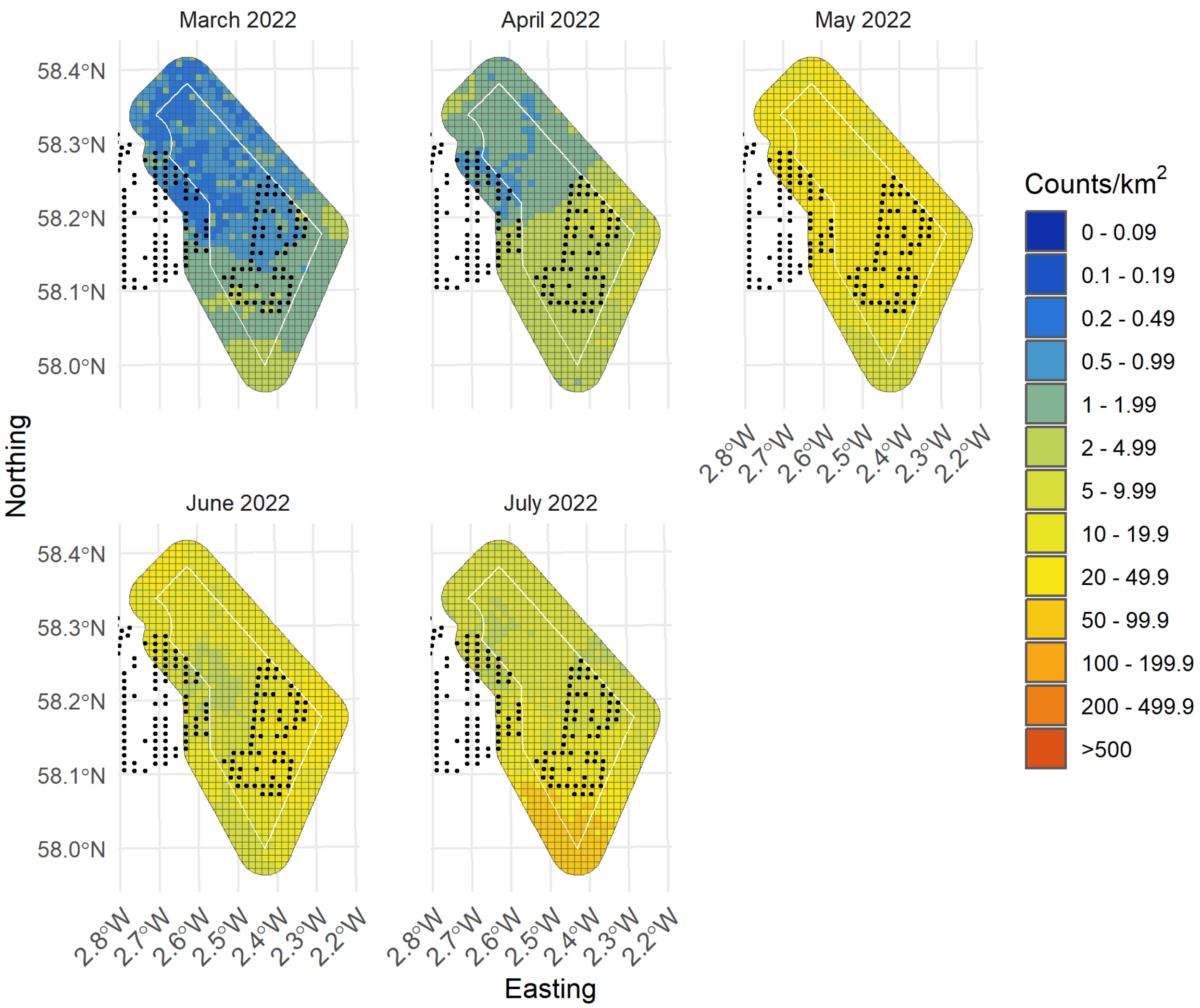


Figure II-97: Random forests predictions from turbine scenario 2 of Guillemot from March 2022 to July 2022 with turbines presented as black dots

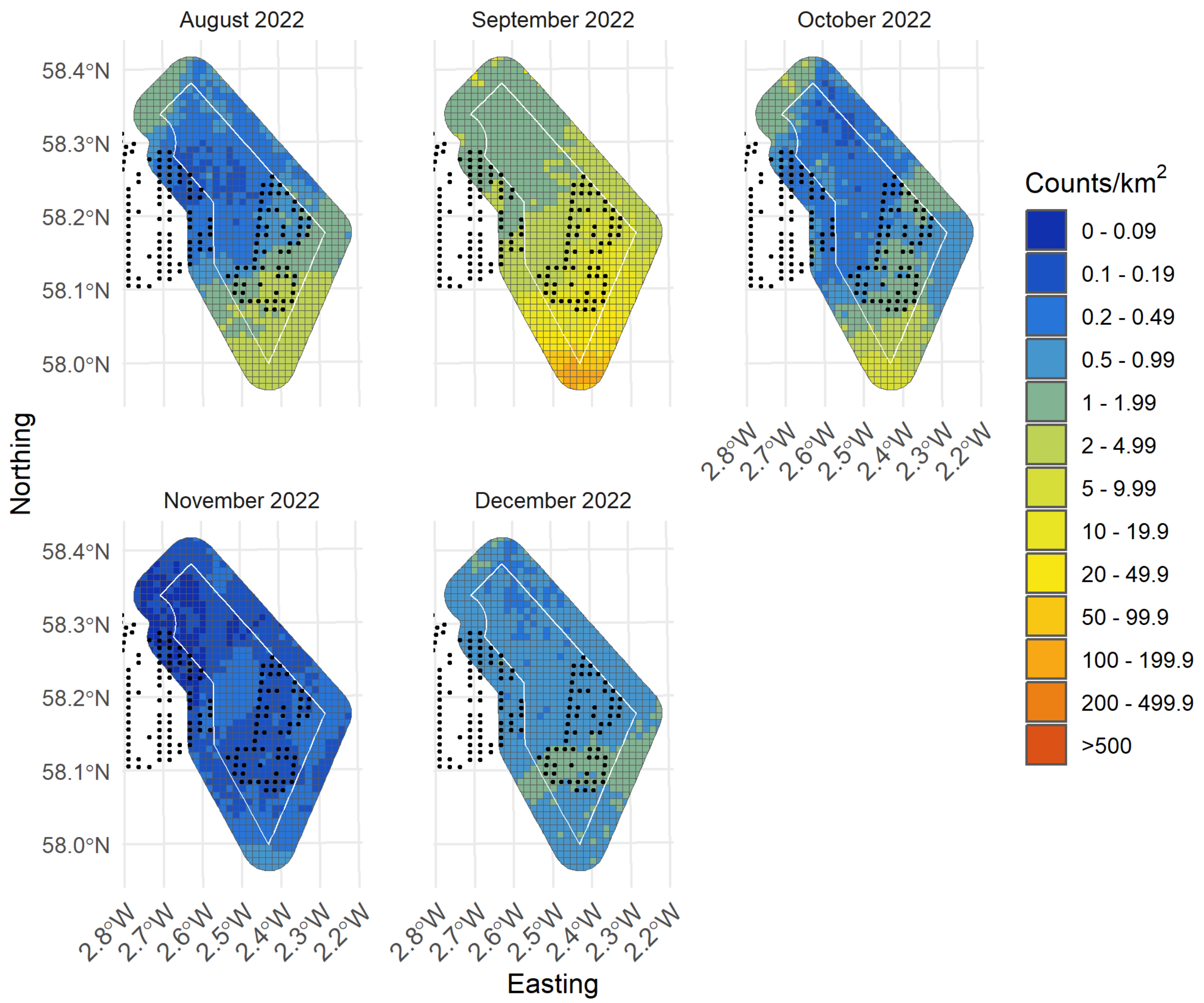


Figure II-98: Random forests predictions from turbine scenario 2 of Guillemot from August 2022 to December 2022 with turbines presented as black dots

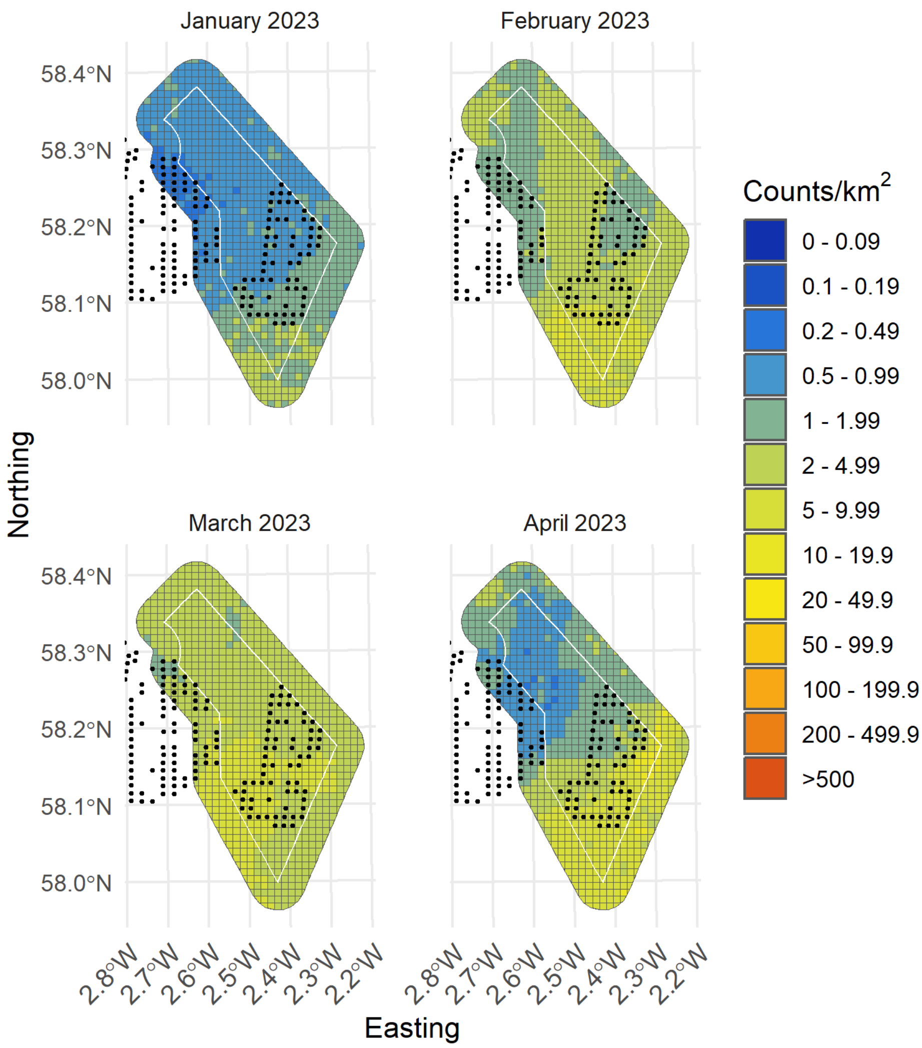


Figure II-99: Random forests predictions from turbine scenario 2 of Guillemot from January 2023 to April 2023 with turbines presented as black dots

II.6.5.4 Turbine scenario 3

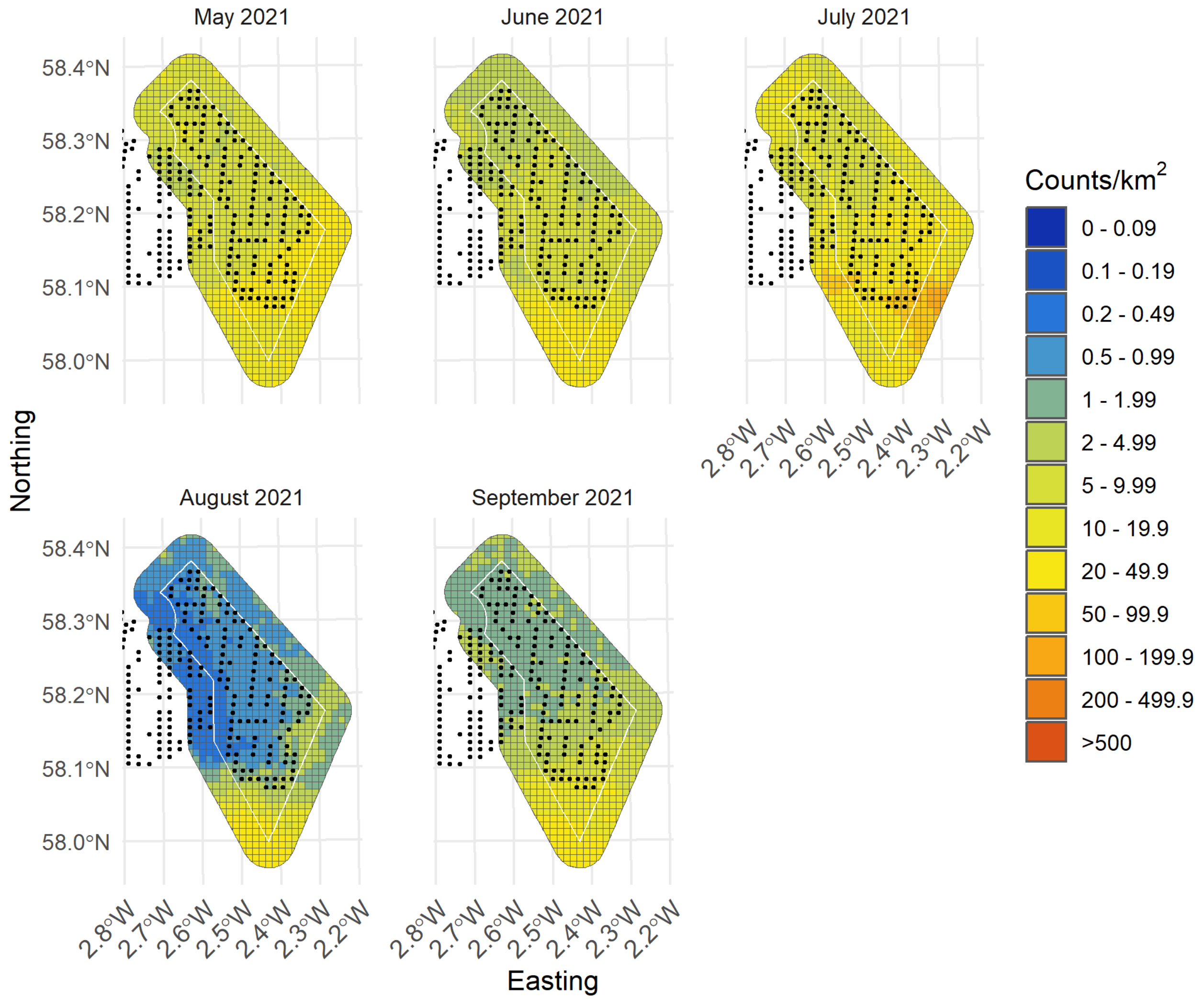


Figure II-100: Random forests predictions from turbine scenario 3 of Guillemot from May 2021 to September 2021 with turbines presented as black dots

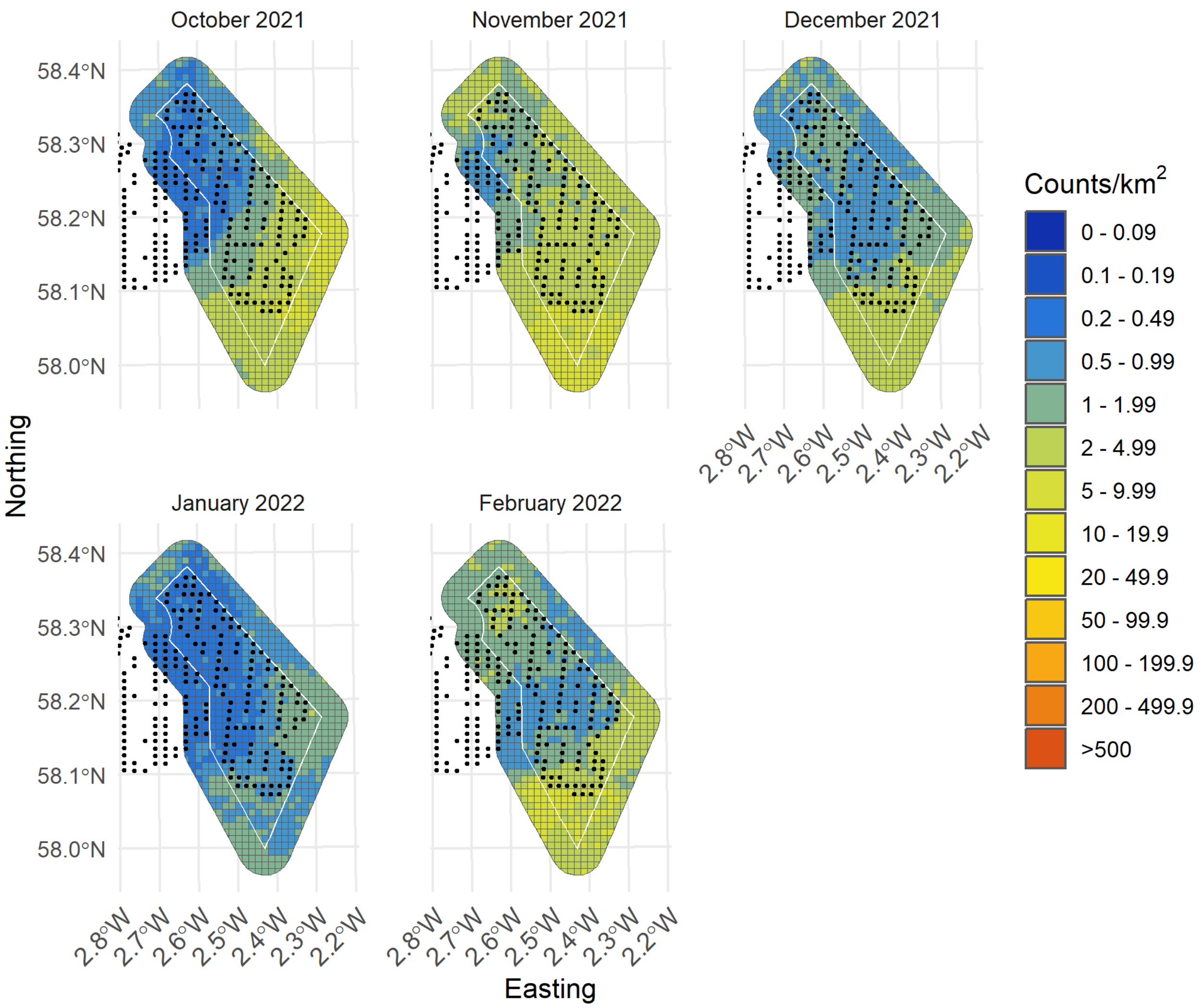


Figure II-101: Random forests predictions from turbine scenario 3 of Guillemot from October 2021 to February 2022 with turbines presented as black dots

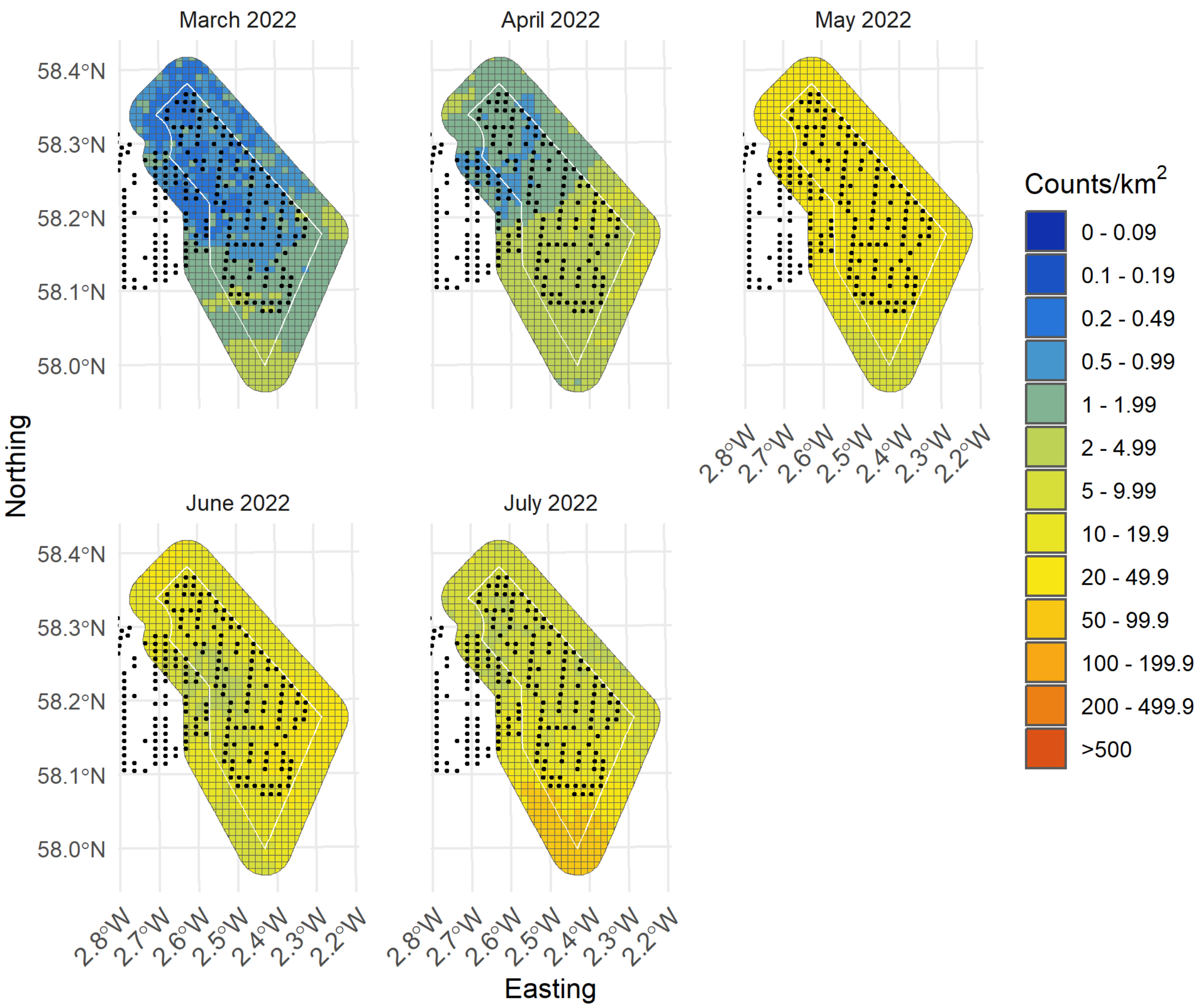


Figure II-102: Random forests predictions from turbine scenario 3 of Guillemot from March 2022 to July 2022 with turbines presented as black dots

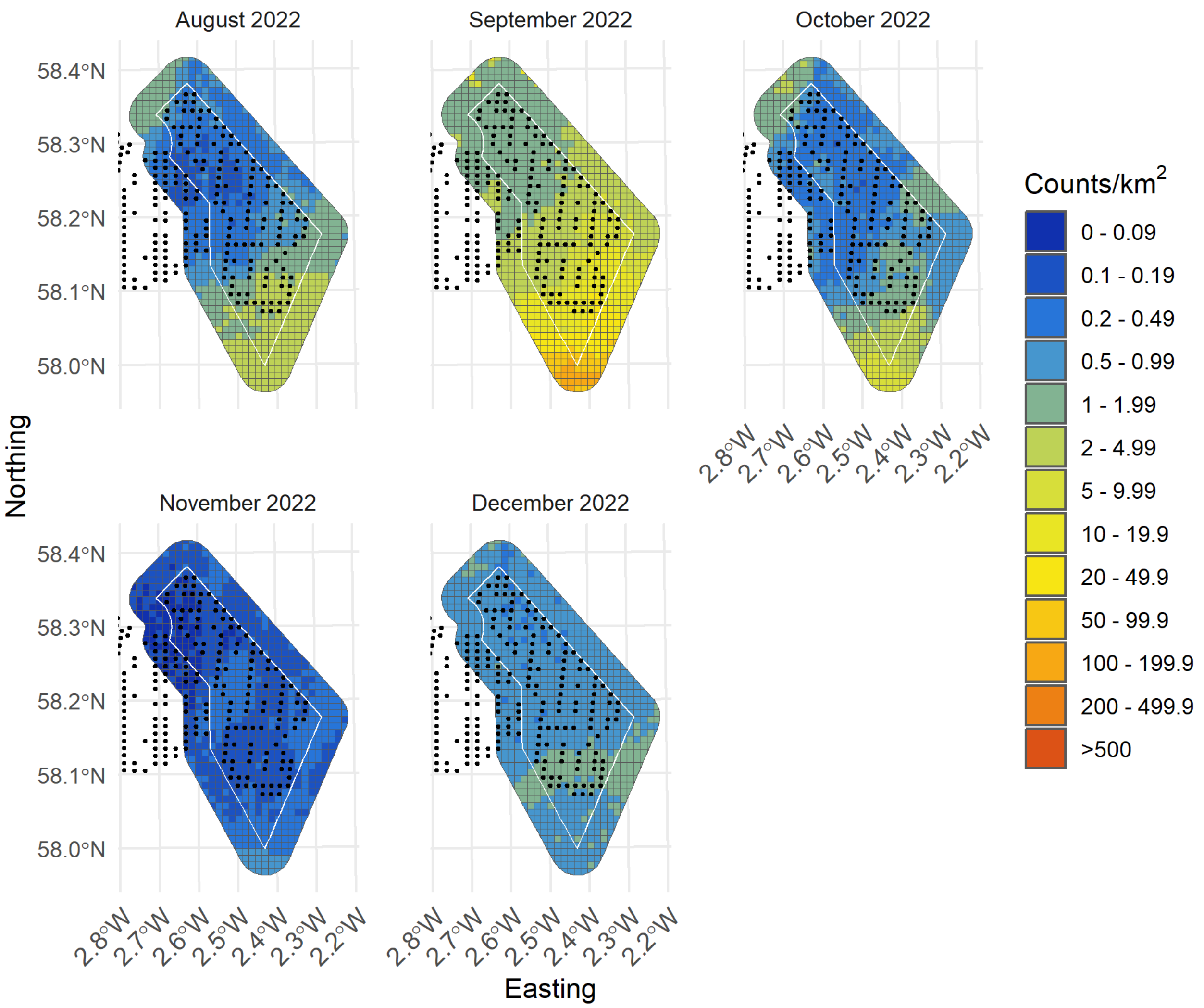


Figure II-103: Random forests predictions from turbine scenario 3 of Guillemot from August 2022 to December 2022 with turbines presented as black dots

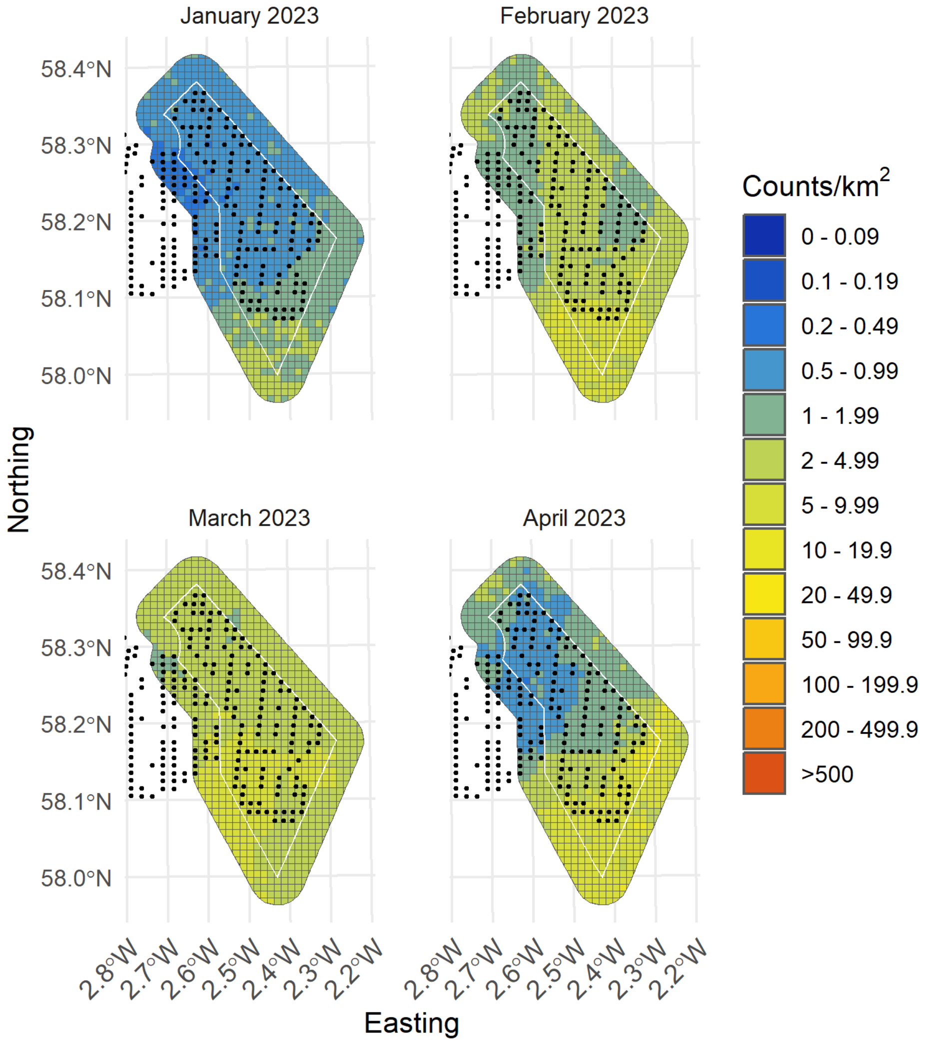


Figure II-104: Random forests predictions from turbine scenario 3 of Guillemot from January 2023 to April 2023 with turbines presented as black dots

III Hot spot analysis

III.1 Results

The hot-spot analysis revealed no persistent hot spots across the site for the entire set of density surface models. Most of the site was found to have moderate relative densities (see table 2 for a definition of the classes). The south-west portion of the site was the only region that had hot spots, though they were variable in nature. This is in line with the broad distributional patterns noted in the density surface models (Figure III-1).

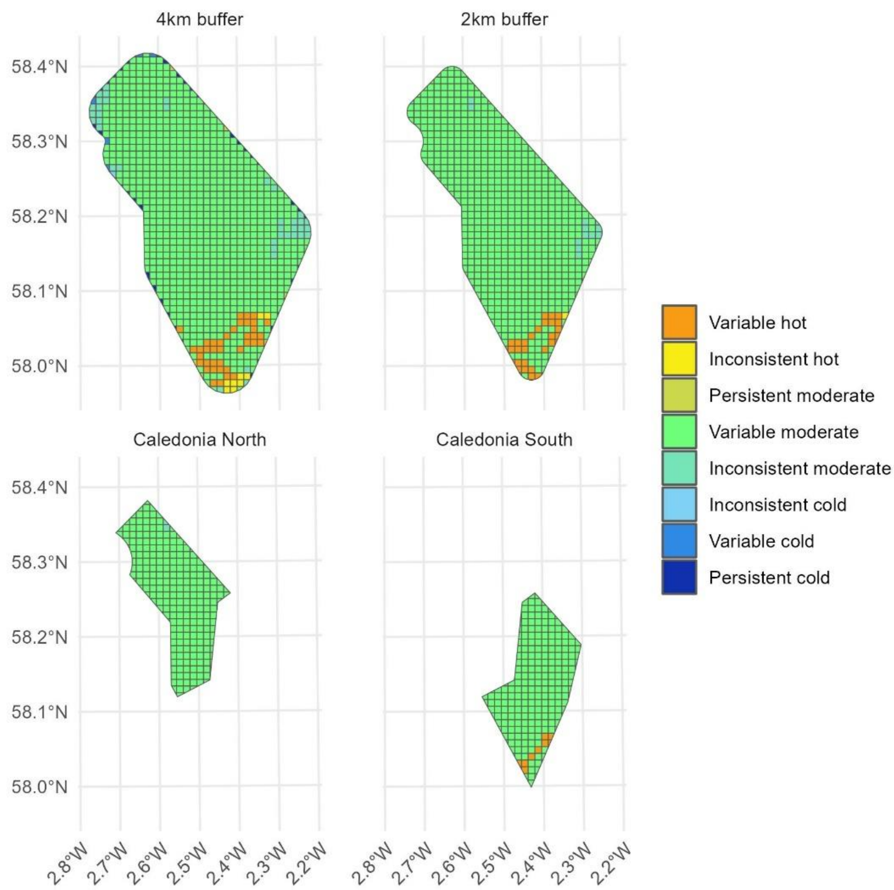


Figure III-1: Classified grid cells from the hot-spot analysis in the Caledonia Offshore wind farm area.

Data has the power to change how we interact with the natural world; we can explore its complexities and nuances then make changes for the better

Caledonia Offshore Wind Farm
5th Floor, Atria One
144 Morrison Street
Edinburgh
EH3 8EX

www.caledoniaoffshorewind.com

OCEAN WINDS