Effects of wind farms on the spatial distribution of Guillemots

Alain F. Zuur
Highland Statistics Ltd.
www.highstat.com

June 2018

Abstract

In this report we present the results of a statistical analysis applied on data on Common Guillemots (*Uria aalge*) from various European offshore wind farms.

Results for the Princess Amalia Wind Farm (PAWP) show that there are clear spatial patterns in the distribution of birds. However, these spatial patterns differ per survey and there is no consistent pattern in these spatial distributions. We were not able to detect a significant effect of the wind farm on these spatial distributions. Results for the Windpark Egmond aan Zee (OWEZ) are similar. Results for Robin Rigg showed that spatial patterns also differ per survey. However, there seems to be a (weak) wind farm avoidance effect in the spatial patterns for Robin Rigg.

The analysis of PAWP, OWEZ and Robin Rigg data indicate that spatial patterns in abundances differ per survey. The data from the Alpha Ventus, Blighbank en Thorntonbank, Horns Rev and Sheringham Shoal Offshore wind farms do not allow for the analysis of individual surveys due to the large number of zeros for Guillemots. Analysis of combined surveys may be misleading.

Given the fact that for three wind parks, spatial patterns differ per survey, and for all other wind farms considered in this report we cannot analyse data per survey, it may be an option to focus future research on a different species, or base the analyses on multiple species simultaneously.

1 Introduction

Many countries around the North Sea are building offshore wind parks to harvest wind energy to create energy. For example, Denmark is aiming to produce 50% of its required energy consumption via wind energy by 2020. Building large wind parks in the sea may affect sea life in, and around the parks.

In this report we analyse whether wind parks have any effect on the spatial distribution of Common Guillemots (*Uria aalge*). Data on sea birds are typically observed from a boat that navigates in transects around a wind farm. The resulting data sets typically contain many zeros, spatial correlation, temporal correlation and non-linear patterns. As a result the statistical analysis of such a data set is a major challenge and requires advanced statistical methods.

The original plan was to analyse data from at least 5 wind parks, but due to some wind farms having mainly zero observations for Guillemots, we ended up with the analysis of data from only 3 wind parks, namely Princess Amalia Wind Farm (PAWP) and Windpark Egmond aan Zee (OWEZ) in Dutch waters and the Robin Rigg wind park in the UK.

2 Data exploration for two wind farms; PAWP and OWEZ

In this section a short data exploration is applied on the PAWP and OWEZ data following the protocols described in Zuur et al. (2010; 2016a).

2.1 Spatial locations

The offshore PAWP an OWEZ wind farms are located in Dutch coastal waters and Figure 1 shows the spatial position of the sampling locations around these wind farms. Although the first impression is that we have a lot of observations quite close to each other, these data are actually from 9 years; see Figure 2. The later graph shows that during the summer months hardly any guillemots were sampled.

For the statistical analyses of these data we can employ different strategies. First of all, we have to decide whether we want to analyse the combined data from both wind farms, or whether we should split up the data and analyse the data of each wind farm separately. Furthermore, in both approaches we have the option to analyse the data from all surveys, or analyse the data from each survey separately.

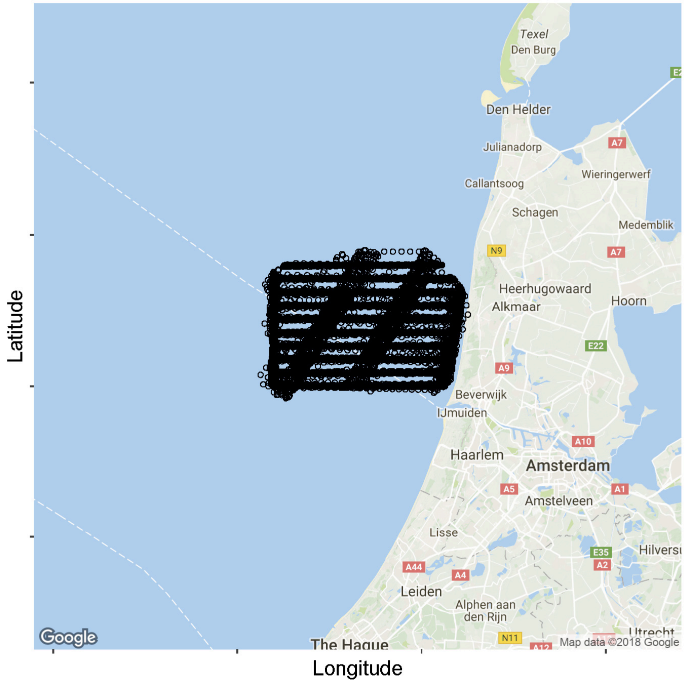

Figure 1. Spatial position of sampling locations for the offshore PAWP and OWEZ wind farms.

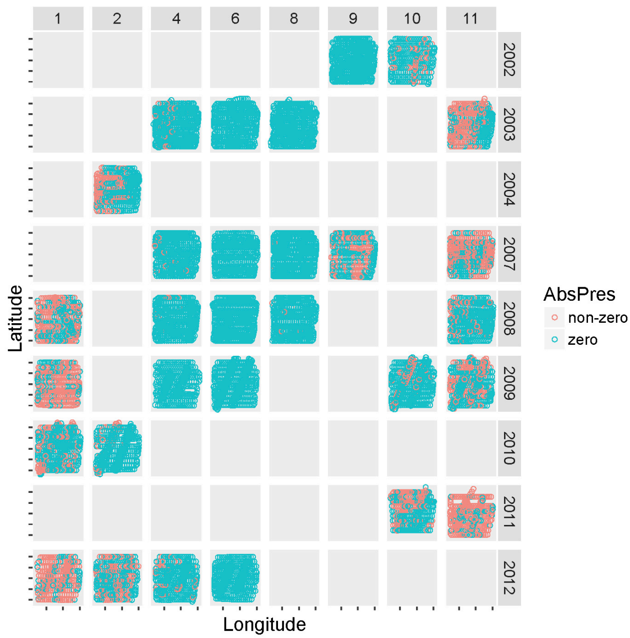

Figure 2. Spatial position of sampling locations by year and month for PAWP and OWEZ. A green dot means that zero birds were observed whereas a red dot indicates that the count was larger than 0.

As to the first issue, the wind farms were built at different times, hence the potential disturbance effect of the PAWP and OWEZ wind farms may be different over time. On the other hand, once both wind farms were operational it is plausible that they had a combined (disturbance) effect on bird distributions. The problem with splitting up the data per wind farm is to decide which data actually belongs to a particular wind farm.

We will initially analyse the data from all surveys. However, results of analyses presented in Section 4 show that spatial distributions differ considerably per survey, and we will therefore end up analysing the data per survey in later sections. Another motivation for analysing the data per survey is that the sampling design of at least one other wind farm has changed over time (i.e. for the Blighbank and Thorntonbank study areas sampling changed from 10 minute resolution to 2 minute resolution). Analysing data by survey, and comparing summary statistics per survey minimizes (though not eliminates) the effect of changes in sampling design. A major advantage of analysing the data for each survey separately is that it will be easier to observe small-scale differences and computing time will also be shorter.

2.2 Zero inflation for PAWP and OWEZ

The third issue we need to address is which surveys we can include in the analyses as some surveys consist for 100% of zeros. The numerical information below shows the number of observations and percentage of zeros for each survey for the combined PAWP and OWEZ data. The first column shows the year (first 4 digits) and month (last 2 digits).

	N	%zeros
200209	464	100.0
200210	247	83.0
200304	503	93.6
200306	461	100.0
200308	472	100.0
200311	338	53.3
200402	393	70.7
200704	455	97.4
200706	445	100.0
200708	554	99.3
200709	301	63.8
200711	393	38.9
200801	295	49.8
200804	449	98.0
200806	510	100.0
200808	475	98.7
200811	411	86.9
200901	284	29.2
200904	316	100.0
200906	430	100.0
200910	415	90.4
200911	404	62.6
201001	403	72.0
201002	382	89.8
201110	302	61.3
201111	411	24.6
201201	300	28.0
201202	335	62.1
201204	332	89.5
201206	333	100.0

We decided to omit the surveys with more than 75% of zeros from the analyses.

3 Statistical models

3.1 Models for a single survey

To introduce the statistical models that will be employed in this report, we first focus on the data from one survey. We arbitrarily selected the survey from November 2003. During this survey 349 observations were made (around the OWEZ and PAWP wind farms) and 54% of the observations were equal to 0. Figure 3 shows the sampling locations for this survey. We also plotted the area covered by the two wind parks (the red polygon is PAWP and the green polygon is OWEZ).

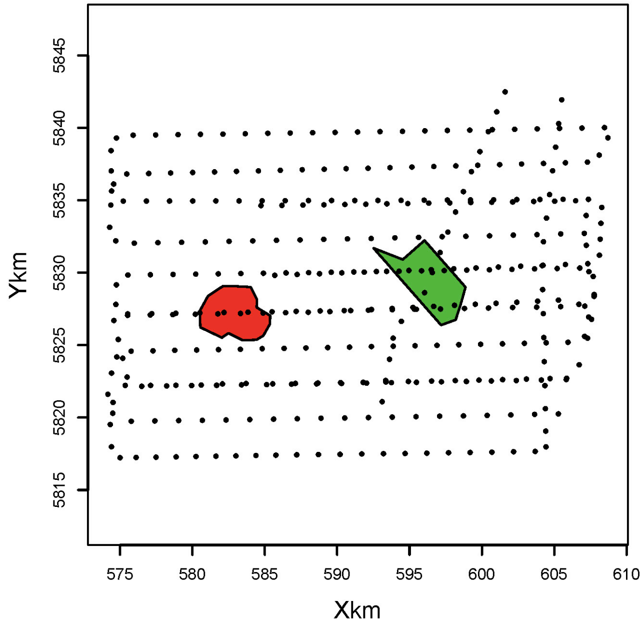

Figure 3. Sampling locations for the survey made in November 2003. The area covered by the PAWP wind park is plotted as a red polygon and the green area is the OWEZ wind park.

The observed data consists of counts of guillemots. We use the following Poisson generalized linear model (GLM; Zuur et al. 2013) as starting point of the analysis of data from one survey:

$$\begin{aligned} &Birds_i \sim Poisson(\mu_i) \\ &E(Birds_i) = \text{var}(Birds_i) = \mu_i \\ &\log(\mu_i) = \beta_1 + \beta_2 \times LogArea_i + \text{Other Covariates}_i \end{aligned} \tag{3.1}$$

This model states that we assume that the number of observed guillemots at site i is Poisson distributed with mean μ_i, and μ_i is modelled as an exponential function of covariates.

The sampling effort differs per site and this is quantified with the variable ‘Area’. One option is to use the natural log of Area as an offset (Zuur et al. 2013). This means that the model can be written as μ_i = Area$_i$ × exp(Intercept + Covariates$_i$) due to properties of exponential and logarithmic functions. However, such an approach assumes that if we double the sampling effort, then we also double the expected number of birds and this may not apply for these data. We therefore decided to use the natural log of sampling effort simply as a covariate; see Equation (3.1).

As to the 'Covariates' component, for the moment we will drop it. In the next subsection, we will formulate models for the analysis of data from multiple surveys. Then we will consider terms like 'construction period' and 'survey' (as a random effect) for the 'Covariates' component.

Results of analyses applied on the November 2003 survey indicate that the Poisson GLM is overdispersed (see Section 5) due to spatial correlation and zero-inflation. In a Poisson GLM the mean is equal to the variance; see Equation (3.1). Quite often this relationship does not hold for ecological data. If the variance is larger than the mean then we have overdispersion. Possible causes for overdispersion are violation of the independence assumption (e.g. spatial correlation) or zero-inflation.

We therefore consider three more models for the survey data from November 2003, namely the Poisson GLM with spatial correlation in Equation (3.2), the zero-inflated Poisson GLM in Equation (3.3) and the zero-inflated Poisson GLM *with* spatial correlation GLM in Equation (3.4).

$$\begin{aligned} & Birds_i \sim Poisson(\mu_i) \\ & E(Birds_i) = \mathrm{var}(Birds_i) = \mu_i \\ & \log(\mu_i) = \beta_1 + \beta_2 \times LogArea_i + v_i \end{aligned} \tag{3.2}$$

The crucial difference between models (3.1) and (3.2) is that in the later one we are using a random effect v_i. Instead of assuming that this random effect is independently and normally distributed, as is usual in mixed effects modelling (Pinheiro and Bates, 2000), we will assume that these are spatially correlated. This is achieved via the covariance matrix of the random effects v_is (Rue et al. 2009).

In Equation (3.3) the zero-inflated Poisson (ZIP) model is presented. In this model, an extra term π is used to model the excessive number of zeros; see Zuur et al. (2009, 2012, 2016b) for details.

$$\begin{aligned} & Birds_i \sim ZIP(\mu_i, \pi) \\ & E(Birds_i) = (1-\pi) \times \mu_i \\ & \log(\mu_i) = \beta_1 + \beta_2 \times LogArea_i \end{aligned} \tag{3.3}$$

Finally, the ZIP model *with* spatial correlation is presented in Equation (3.4). This model allows for excessive number of zeros *and* spatial correlation. This is a rather advanced model. The potential problem with this model is that sites with zero abundance may be geographically close to each other. If that is the case then the zero-inflation parameter π and the spatial term may potentially fight for the same information, and that may cause numerical estimation problems when fitting this model.

$$\begin{aligned} &Birds_i \sim ZIP(\mu_i, \pi) \\ &E(Birds_i) = (1-\pi) \times \mu_i \\ &\log(\mu_i) = \beta_1 + \beta_2 \times LogArea_i + v_i \end{aligned} \tag{3.4}$$

Estimation of all models is done with the package R-INLA (Rue et al. 2009), which is executed from within the statistical software package R (R Core Team, 2018). The technical background of INLA is quite complicated and is not discussed here. See Zuur et al. (2017) for a non-technical explanation.

3.2 Technical information

Before we continue we provide a little bit more technical information on the procedures used by R-INLA. The spatial models in Equations (3.2) and (3.4) contain a spatial correlated random effect. However, R-INLA does not estimate the v_is and its covariance matrix directly. Instead, it puts a fine grid on top of the sampling locations; see Figure 4. This grid is called a mesh. The mesh contains a large number of nodes. At each of these nodes R-INLA will estimate a w_j value. The mesh in Figure 4 has 3715 nodes, and therefore we have 3715 of these w_js values. Once R-INLA has estimated all the w_j values (and the corresponding covariance matrix), it can easily calculate the spatial correlated random effects u_i. The w_js are also spatially correlated and a special correlation function is used to model its covariance matrix.

Technically, the Matérn correlation function is used in combination with the SPDE approach, which stands for continuous domain stochastic partial differential equations. Full details can be found in Blangiardo and Cameletti (2015).

Main point summary: Instead of estimating the correlation between the spatial random effects u_i directly, INLA works at a deeper level, namely with a large number of w_j values defined on a fine grid (called: mesh).

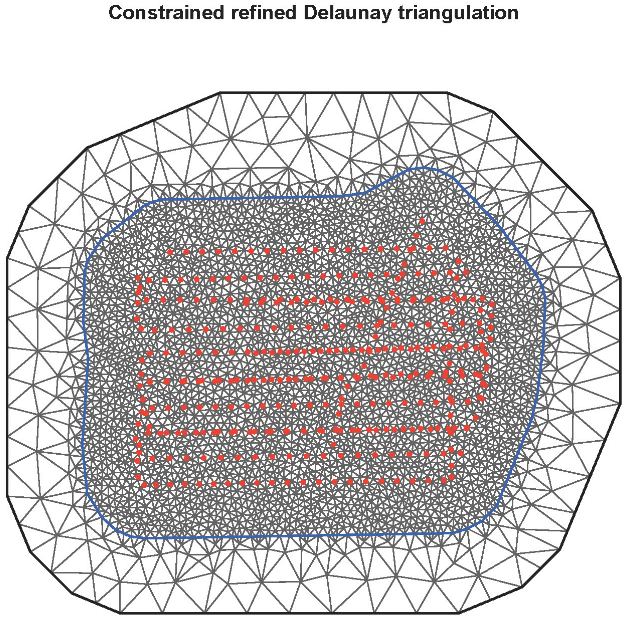

Figure 4. The mesh used by the spatial ZIP model. The mesh has 3715 vertices. The mesh has an inner part (fine mesh) and also an outer part to avoid numerical problems due to the boundary.

3.3 Models for multiple surveys

Before applying models (3.1) to (3.4) on the November 2003 survey data we will analyse the data of all surveys from both wind farms. Three models were applied; a model with no spatial correlation, a model in which there is only spatial correlation, and a model with spatial-temporal correlation. And we used the Poisson, zero-inflation and negative binomial versions of these models. The motivation for using the negative binomial (NB) distribution is this distribution is also capable of dealing with excessive number of zeros. Hence, in fact we applied 9 models. Results are presented in Section 4.

The model with the spatial-temporal correlation is using the so-called 'replicate' correlation. In this correlation structure all surveys share the same correlation parameters, but the spatial random fields can change from survey to survey.

Main point summary: Three types of models are applied on the bird data from all surveys of the PAWP and OWEZ wind farms; a model without spatial correlation, a model with spatial correlation, and a model with spatial-temporal correlation. And we combined these models with the Poisson, zero-inflation and negative binomial distributions.

4 Results for all surveys and both wind parks

When comparing models in R-INLA we can use the DIC and/or WAIC. The use of these statistics is similar to the use of the AIC in frequentist analysis; the lower the DIC (or WAIC) the better the model.

Below, the DIC and WAIC values of the models without spatial correlation (using a Poisson, ZIP and NB distribution), with spatial correlation, and with spatial-temporal correlation are presented.

	dic	waic
Poisson GLM	20704.36	20764.92
ZIP GLM	17370.20	17420.35
NB GLM	13945.30	13948.36
Poisson GLM + SRF	16395.43	18212.21
ZIP GLM + SRF	14856.01	15840.24
NB GLM + SRF	13766.82	13776.66
Poisson GLM + replicate SRF	**11900.04**	**12028.75**
ZIP GLM + replicate SRF	12282.33	12603.47
NB GLM + replicate SRF	12807.43	12796.74

The DIC and WAIC values of all models with spatial-temporal correlation are lower. In the replicate correlation model, each survey is allowed to have a different spatial pattern, although all spatial patterns share the same statistical parameters that define the spatial correlation. Figure 5 shows the spatial random fields for each selected survey.

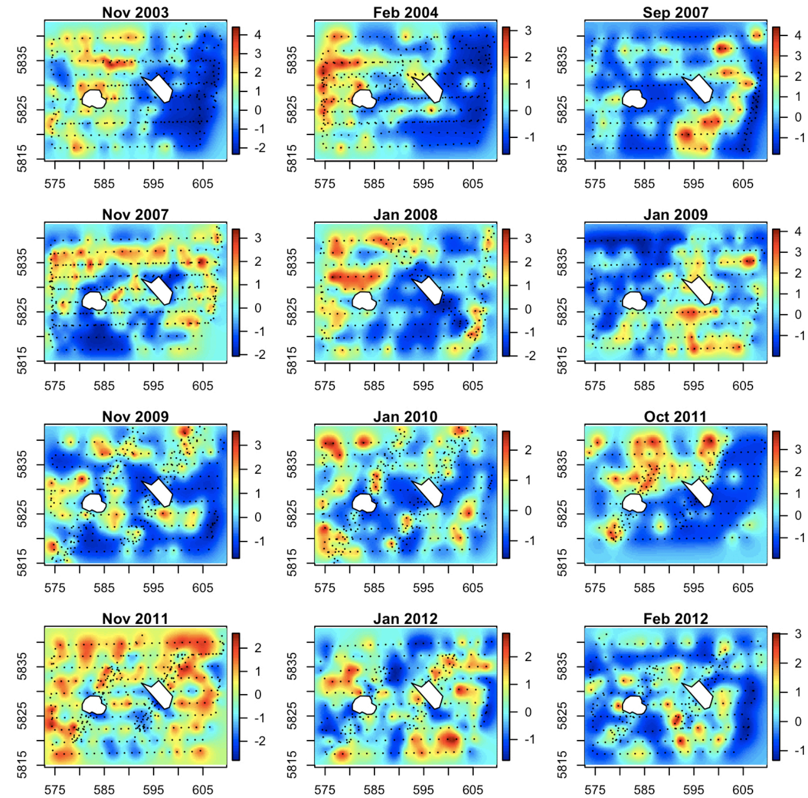

Figure 5. Spatial random fields obtained by the Poisson GLM with spatial-temporal correlation applied on the combined data of the OWEZ and PAWP wind parks. Red values correspond to areas with higher abundances and blue values to areas with lower abundances.

Note that the spatial patterns differ considerably from survey to survey. Because of this we decided to analyse the data from each survey separately. The advantage of doing this is that it will give us more detailed information, and we can more easily obtain summary statistics. Additionally, although the DIC and WAIC values indicate that the Poisson distribution is the best model, it may well be that for individual surveys a zero-inflated distribution is better.

Main point summary: DIC and WAIC values indicate that the spatial correlation differs per survey. The spatial random fields in Figure 5 indicate that the spatial patterns of birds differ per survey.

5 Results for one survey (November 2003)

In the previous section we showed that the spatial distribution of each survey is rather different. This means that the spatial distribution of the guillemots differs per survey, and it also justifies analysing the data from each individual survey. Before we present the results of all surveys, we will present the full results of the analysis of one survey, namely that of the November 2003 survey. The reason for doing this is to give an idea which steps we carried out for each survey, but are not shown here.

5.1 Poisson GLM

We started the statistical analysis by applying the Poisson GLM in Equation (3.1) on the November 2003 data using R-INLA. After fitting any statistical model, one has to apply model validation. In such a step we plot residuals versus fitted values, plot residuals versus each covariate in the model and each covariate not in the model, and we also need to assess the residuals for spatial dependency.

Another model validation step is to simulate 1,000 data sets from the model. The simulated 1,000 data sets from the model should be similar to the observed data. But if the simulated data sets contain fewer numbers of zeros than the observed data set, then we know that the model cannot cope with the excessive number of zeros.

Figure 6 shows a variogram of the Pearson residuals obtained from the Poisson GLM. A horizontal band of points would indicate spatial independency but in this case there is clear violation of independence. Figure 7 summarises the simulation study. We simulated 1,000 data sets from the model, and for each simulated data set we calculate the number of zeros. These are plotted as a frequency plot. The red dot represents the number of zeros in the November 2003 survey. If the model would be able to cope with the excessive number of zeros then the red dot would lie within the black bars of the frequency plot. Clearly, that is not the case here. Hence, the Poisson GLM cannot cope with the excessive number of zeros. On itself this is not surprising as the model only contains $LogArea_i$ as a covariate.

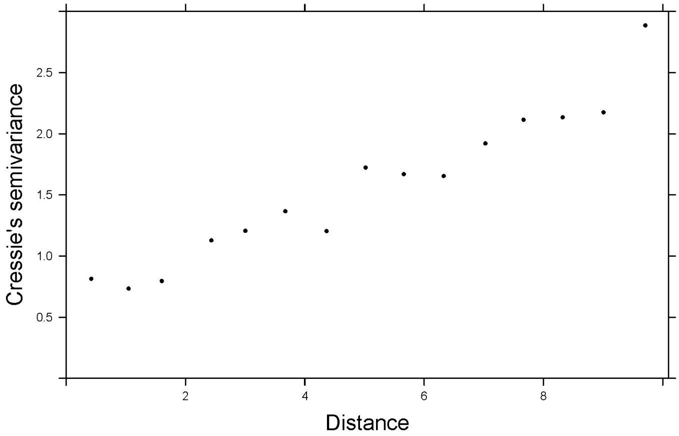

Figure 6. Variogram of Pearson residuals obtained by the Poisson GLM in Equation (3.1).

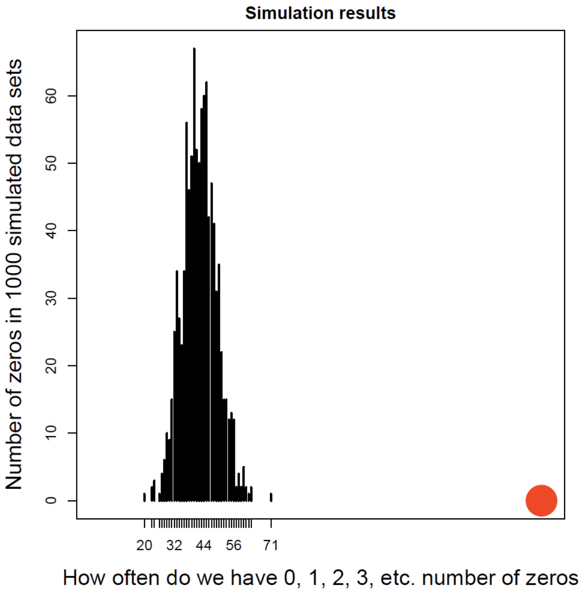

Figure 7. Results of simulation study. The red dot represents the number of zeros in the observed data in survey from November 2003. The frequency plot (black lines) are based on the number of zeros in 1,000 simulated data sets.

We then applied the other three models that were discussed in the previous section. The tools used in R-INLA are based on Bayesian statistics and comparing models in a Bayesian context can be done by comparing DICs or WAICs. The DICs and WAICs values for all four models are as follows.

	dic	waic
Poisson GLM	2261.25	2279.00
Spatial Poisson GLM	1049.27	1195.36
ZIP GLM	1649.90	1664.50
Spatial ZIP GLM	**986.67**	**993.71**

The lower these statistics, the better is the model. The results indicate that the spatial ZIP model is best model, followed by the spatial Poisson GLM. We will apply the spatial ZIP model in the next subsection.

Main point summary: Model validation of the Poisson GLM applied on the November 2003 data indicated that Pearson residuals are spatially correlated and the model cannot cope with the excessive number of zeros. Results of all models indicate that the spatial ZIP model is the best, as judged by DIC and WAIC values.

5.2 Poisson GLM with spatial correlation and zero-inflation

5.2.1 Interpretation of the spatial random field (w_js)

In Section 3.2 we gave a rather short explanation about the spatial random field. Recall that we use a mesh on which a large number of w_j values are estimated. Once R-INLA has estimated all the 3715 w_j values, we can use special functions in R to plot them; see Figure 8. We will now explain how to read this graph. The fitted values of the ZIP model with spatial correlation are calculated as

$$(1 - \pi) \times \exp(\text{Intercept} + \text{LogArea effect} + \text{Spatial correlation})$$

This is based on the expression for the mean in Equation (3.4). Due to the properties of the exponential function, this can also be written as

$$(1 - \pi) \times \exp(\text{Intercept} + \text{LogArea effect}) \times \exp(\text{Spatial correlation})$$

The posterior means of the intercept and slope for LogArea are as follows.

	mean	sd	0.025quant	0.975quant
Intercept	1.8717	0.2376	1.4051	2.3379
LogArea	0.3011	0.1843	-0.0607	0.6626

The posterior mean of $\pi = 0.18$. This means that the fitted values can be written as

$$0.82 \times \exp(1.87 + 0.30 \times \text{LogArea}_i) \times \exp(\text{Spatial correlation})$$

The 'Spatial correlation' term represents the u_is in Equation (3.4). R-INLA replaces these by $\mathbf{A} \times \mathbf{w}$, where the vector $\mathbf{w}$ contains all the w_j values and $\mathbf{A}$ is a matrix consisting of zeros and ones. Figure 8 is a visual representation of all the w_j values. Wherever the w_js are zero (i.e. green areas), we have exp(0) = 1, and there is no spatial effect. And if $w_j = 3$ (red areas) we need to multiply the other components in the model with exp(3) ≈ 20. Finally, if $w_j = -3$ (dark blue areas) we need to multiply the other components in the model with exp(-3) ≈ 0.05. In other words, the spatial random field (i.e. all the w_js) in Figure 8 shows where we have high abundances (the red areas) and where we have low abundances (the blue areas).

The standard deviation of the spatial random field (not shown here) is about 0.5, which means that any w_j value in Figure 8 that is larger than 1-ish and smaller than -1-ish are important.

A standard INLA analysis would stop at this point and the focus of the discussion would be what the spatial random field represents (e.g. real dependency or missing covariates).

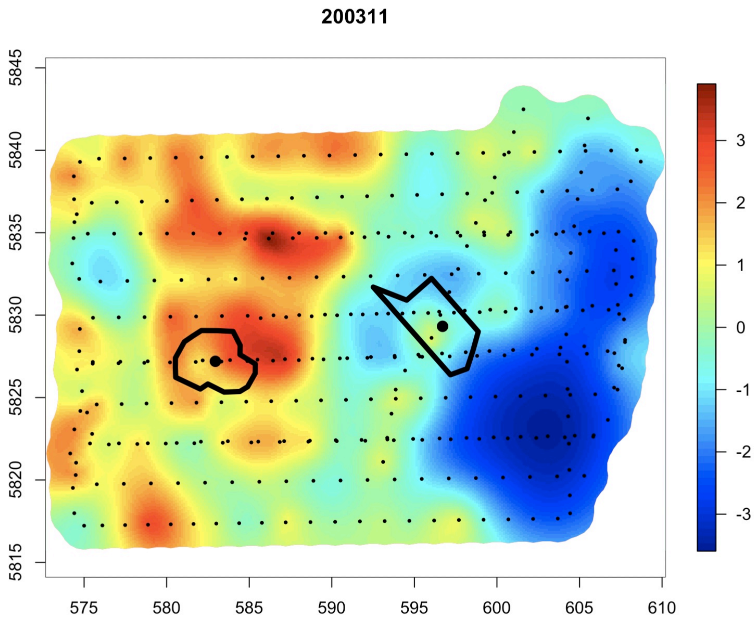

Figure 8. Spatial random field obtained by the spatial ZIP model applied on the survey data from November 2003. The coordinates are UTM coordinates.

5.2.2 Converting the spatial random field into avoidance and preference

The main motivation for conducting the statistical analysis in this report is to assess whether there are any spatial disturbances due to the wind farms. In order to quantify 'spatial disturbance' we somehow need to convert the colours in Figure 8 into numbers versus distance. Actually, it is not the colours (as these are just a visual presentation) but the w_j values underlying the colours.

One possible strategy is to draw circles with different discrete radiuses around the center of a wind farm (see Figure 9 for PAWP) and determine the percentage of w_j values that are negative as this means 'avoidance'. And it is also interesting to calculate the percentage of red values and compare this to distance to the wind farm. Unfortunately, determining these percentages is not a trivial exercise and we will discuss how to do this in the next section.

The only problem with this specific survey is that it is a so-called T_0 survey. That means that the survey was taken before construction of the wind farm; hence there is not supposed to be any disturbances yet. However, we can consider the centre of the 'wind farm to be' as a reference point and quantify how much/far the red areas and blue areas are away from this reference point. Ultimately, we want to repeat this exercise for every survey and look at changes over time in these 'distance' patterns for T_0, T_1 (during construction) and T_2 (post-construction) surveys.

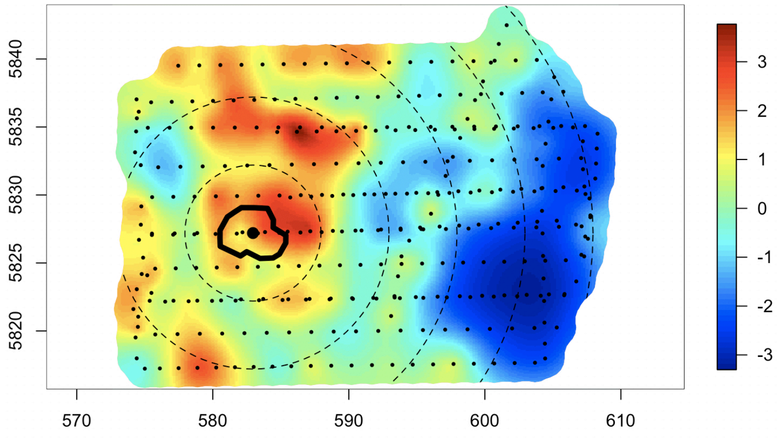

Figure 9. Spatial random field obtained by the spatial ZIP model. Circles with a radius of 5 km, 10 km, 15 km, etc. around the centre of PAWP have been added.

5.3 Excursion sets for PAWP

In this section we discuss how to determine the percentage of w_j values in each distance band that are negative, and also the percentage that is positive. Figure 10 shows the same picture as in Figure 9, except that we have added the spatial positions of the w_j values as small yellow dots.

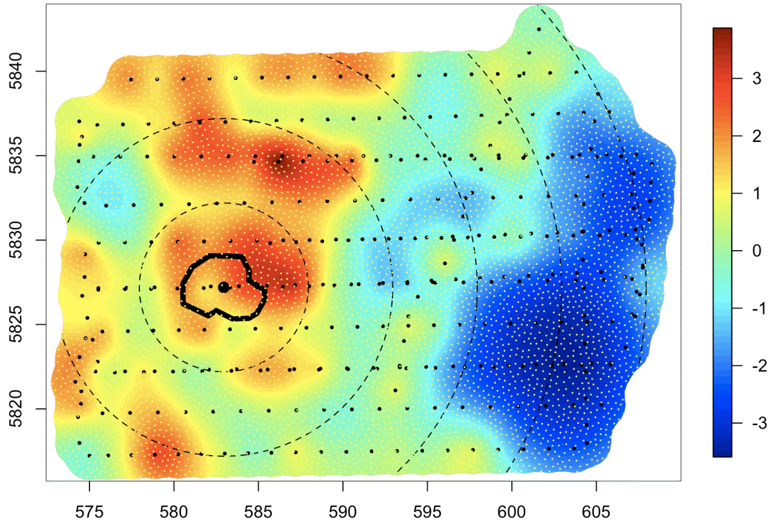

Figure 10. As Figure 9. The position of the w_js have been added as small yellow dots.

For each w_j we also get its posterior standard deviation and using this we can determine whether a w_j is important. Wang et al. (2017) and Bolin and Lindgren (2015) explain how to do this. A brief outline of this process is given next.

The spatial random field can a be written as $f(x)$, where x are the sampling locations. The challenge is now to find all the x values (or: all the locations) in the study area for which $f(x) > u$ with a given probability $1 - \alpha$. For example, we may be interested to know all the locations for which $f(x) > 0$ with probability 0.9. These are all the points were we have above average counts due to spatial patterns.

In a frequentist analysis this would be formulated as which of the w_js in Figure 9 are significantly different from 0 (and positive). But with 3715 w_js, we would need to do a correction for multiple testing, for example a Bonferroni correction.

In Bayesian analysis this works slightly different, see Bolin and Lindgren (2015) for details. First we calculate the marginal probability that a w_j is larger than 0. The phrase ‘marginal’ means that we are ignoring the multiple testing problem. We just close our eyes and repeat this process 3715 times. The locations were a w_j is positive and important are plotted as slightly larger yellow dots (within the red areas) in Figure 11.

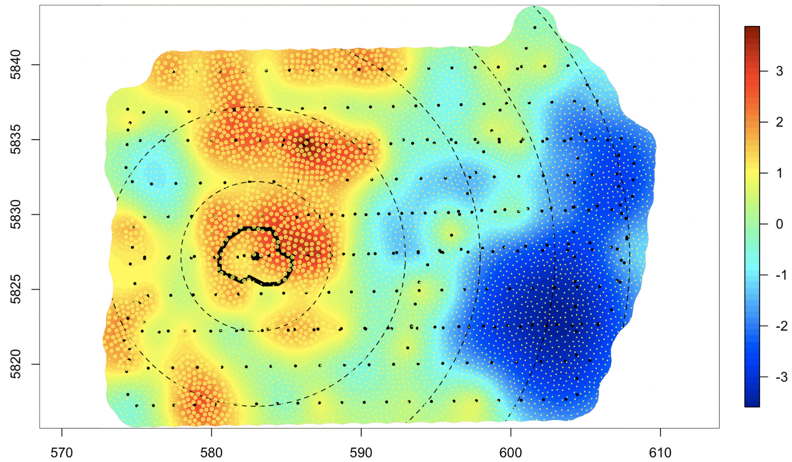

Figure 11. As Figure 9. Yellow dots within the red areas refer to locations where the w_j is positive and important.

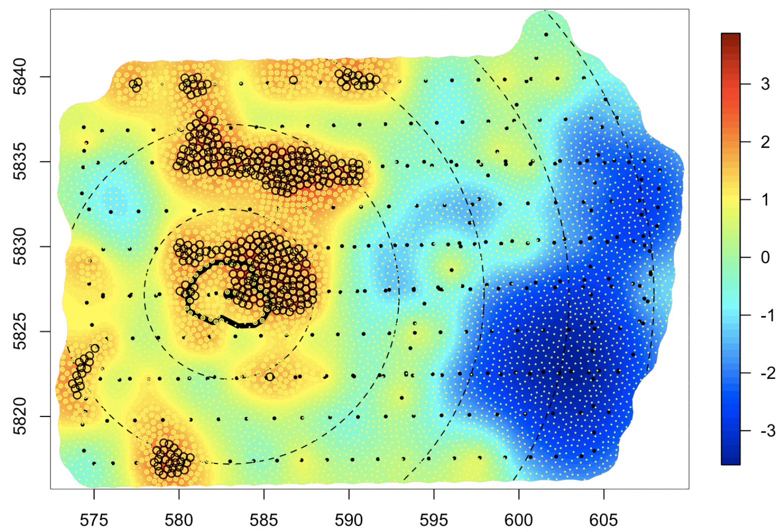

Figure 12. As Figure 9. Black dots refer to the exclusion set for the positive values. At these locations the joint probability that the w_j values are positive and important is at least 0.90.

To deal with the Bayesian equivalent of multiple testing, Bolin and Lindgren (2015) introduce an exclusion set. These are the locations where the *joint* probability that the w_j values *at all these* locations is larger than 0 is at least 0.9. The locations for which the joint probability is positive and larger than 0 are plotted as black dots (within the read areas) in Figure 12. We can consider these areas as ‘preference areas’. Instead of looking at positive values of the w_js, we can also look at negative values. This would identify locations with lower abundances of guillemots; call it ‘avoidance areas’.

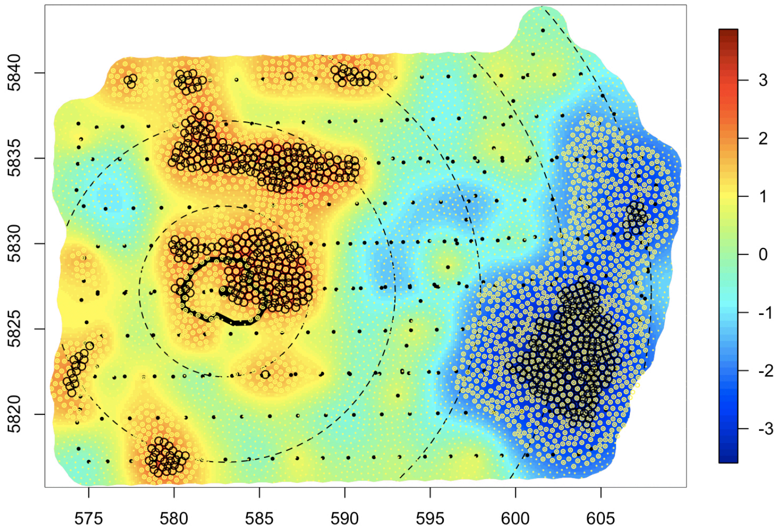

Figure 13. As Figure 9 to Figure 12. The yellow dots within the blue areas refer to points were we have an important reduction of the spatial random field. The black dots within the blue areas form the exclusion set for the negative values.

Now that we know at which sampling areas of the spatial random field we have important and positive values, and important negative values, we proceed to the next stage and calculate the percentages of such points for each distance band around the reference point (the center of the ‘to be build wind farm’); see Figure 14. We simply calculated the number of nodes (or: w_js) that are important within a distance band and divided this by the total number of nodes within a distance band (while taking into account the boundary of the study area). The left panel shows the percentage of important nodes (or: w_js) for the negative values (blue areas in Figure 9) in each distance band, and the right panel shows the percentage of important nodes that are positive within each distance band.

These graphs give the same information as Figure 8, namely we have above average values close (i.e. 0 – 5 km and 5 – 10 km) to the reference point. And the further away from the reference point (15 – 20 km and 20 – 25 km) the lower the abundances.

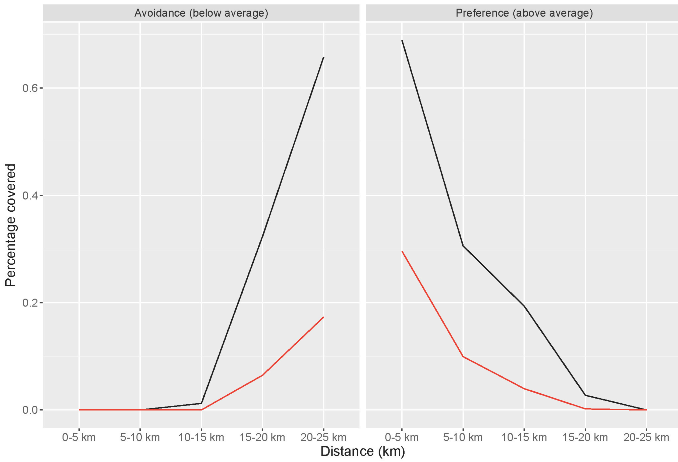

Figure 14. Percentages of important nodes in each distance band for PAWP. The black lines are for the marginal probabilities and the red lines are for the joint probabilities. The left panel is for the negative values (blue areas in Figure 9) and the right panel is for the positive values (red areas in Figure 9).

We will repeat the whole excursion set analysis for OWEZ, and potentially we can also combine the results and calculate the percentage of important nodes for the distance bands of both wind farms (i.e. calculate the percentage of important nodes in both the 0 – 5 km intervals). We will not do the last part as the two wind parks were built at different times.

6 Results for each individual survey

We repeated the analysis described in Section 5 for all surveys and results are presented in Figure 15 to Figure 26. Red values correspond to above average values and blue values to below average values. Small circles correspond to important marginal probabilities and the black open circles to important joint probabilities. And in Figure 27 all these 12 graphs are presented in one figure. It seems that the spatial random field changes considerably from survey to survey. We also observed this in the analysis of the combined data.

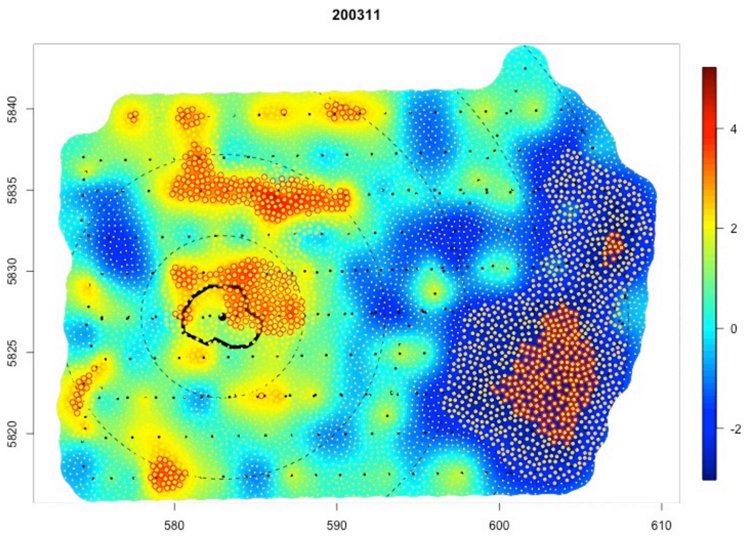

Figure 15. Results for the November 2003 survey.

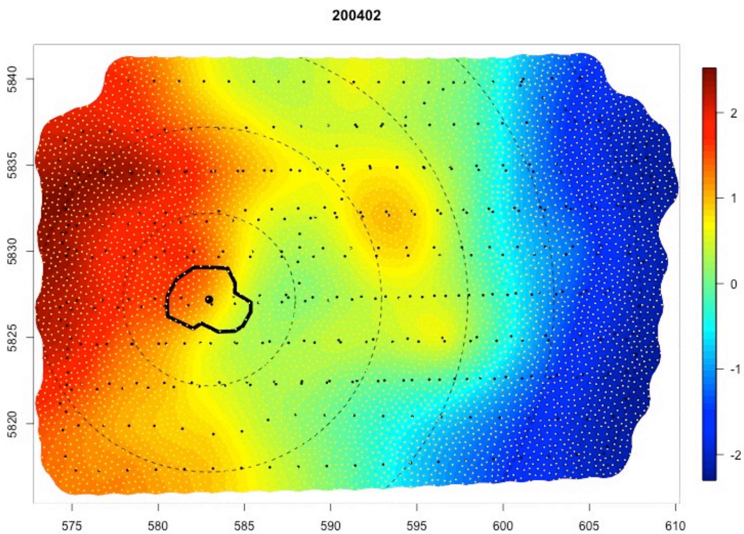

Figure 16. Results for the February 2004 survey.

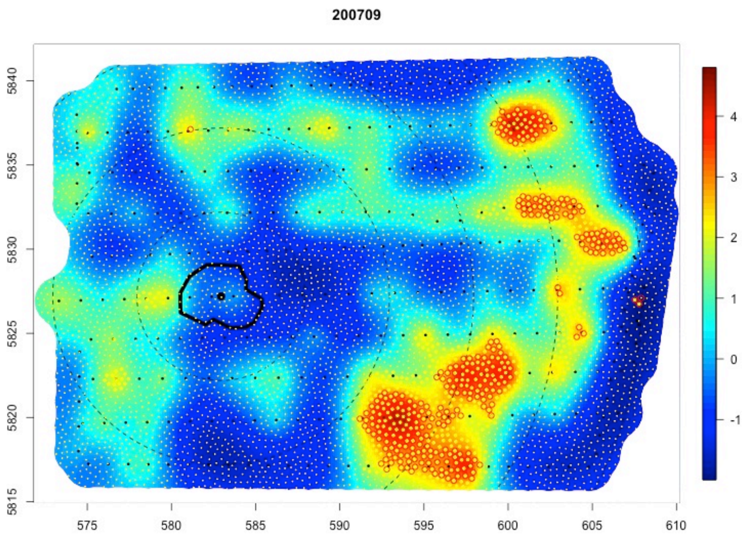

Figure 17. Results for the September 2007 survey.

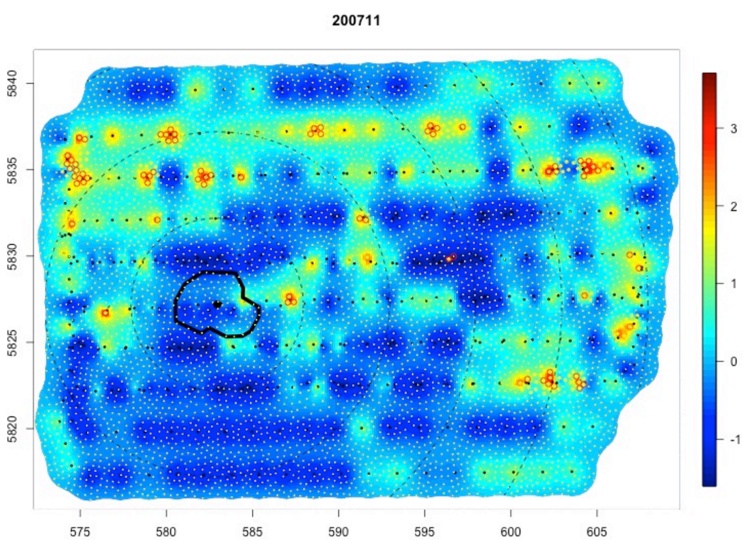

Figure 18. Results for the November 2011 survey.

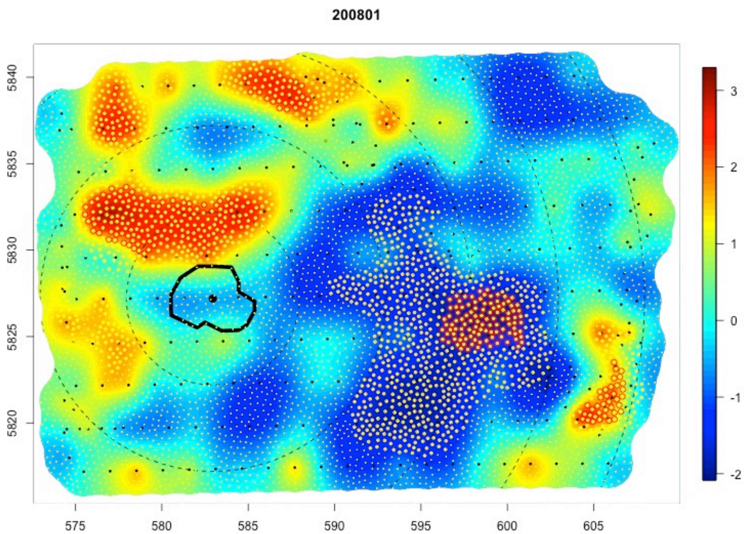

Figure 19. Results for the January 2008 survey.

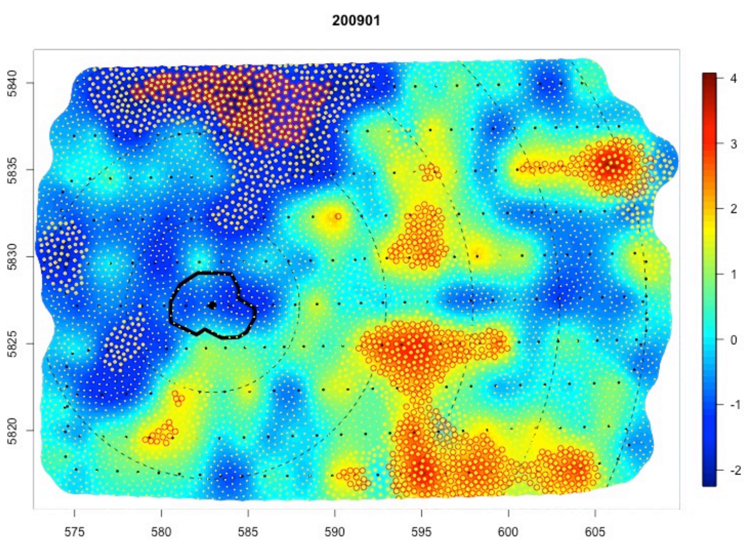

Figure 20. Results for the January 2009 survey.

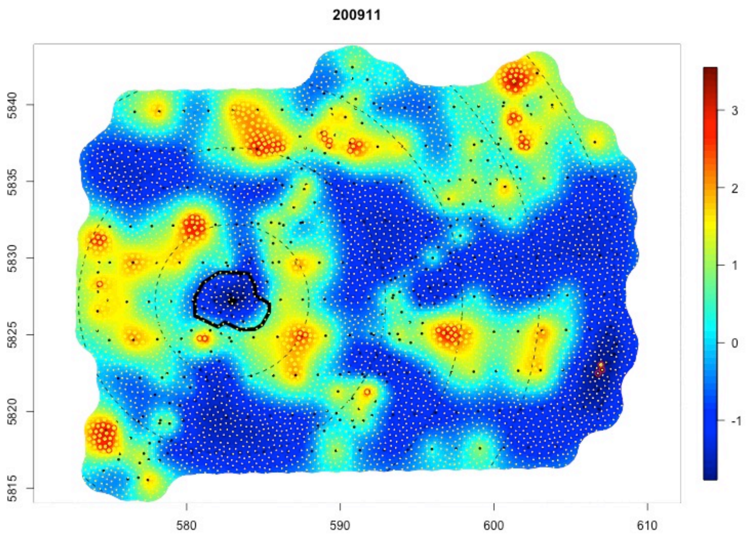

Figure 21. Results for the November 2009 survey.

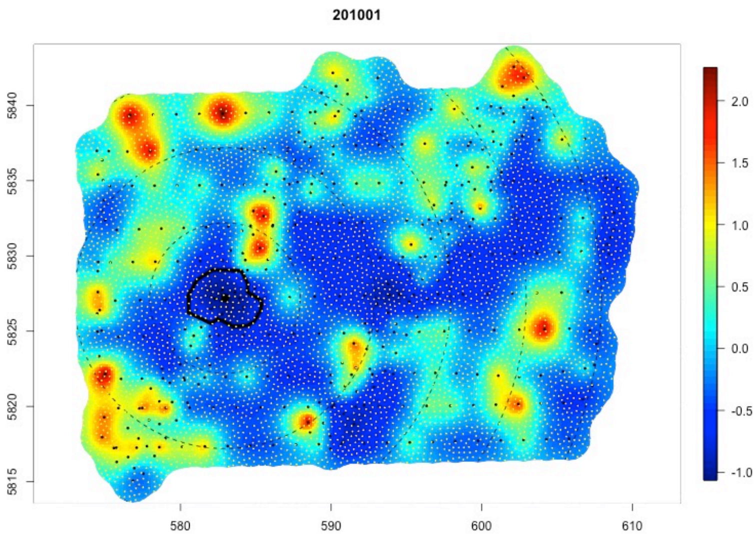

Figure 22. Results for the January 2010 survey.

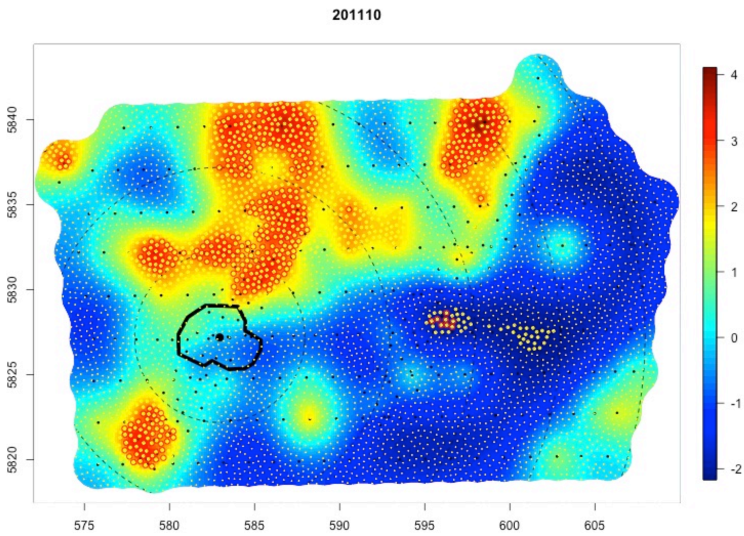

Figure 23. Results for the October 2011 survey.

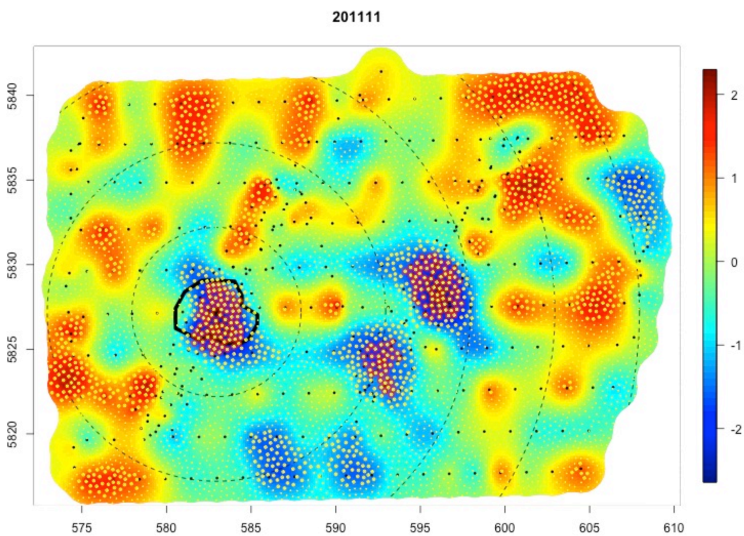

Figure 24. Results for the November 2011 survey.

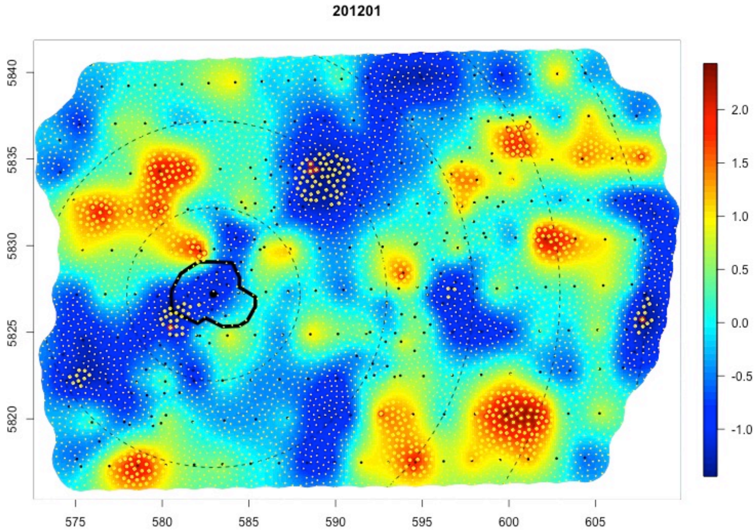

Figure 25. Results for the January 2012 survey.

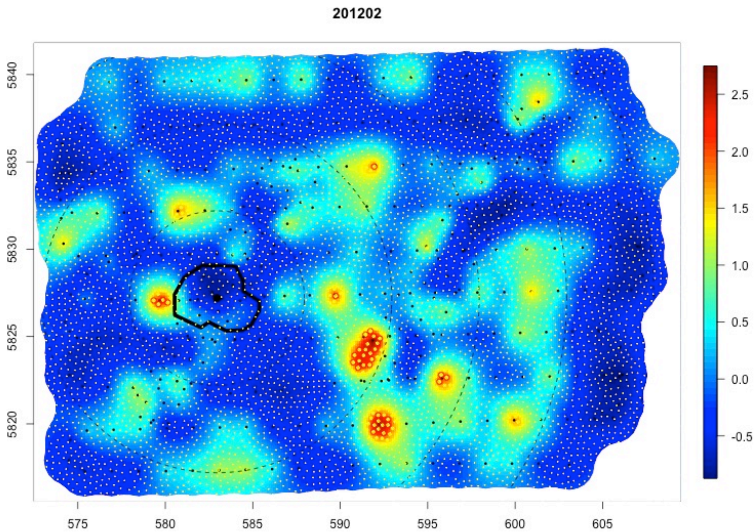

Figure 26. Results for the February 2012 survey.

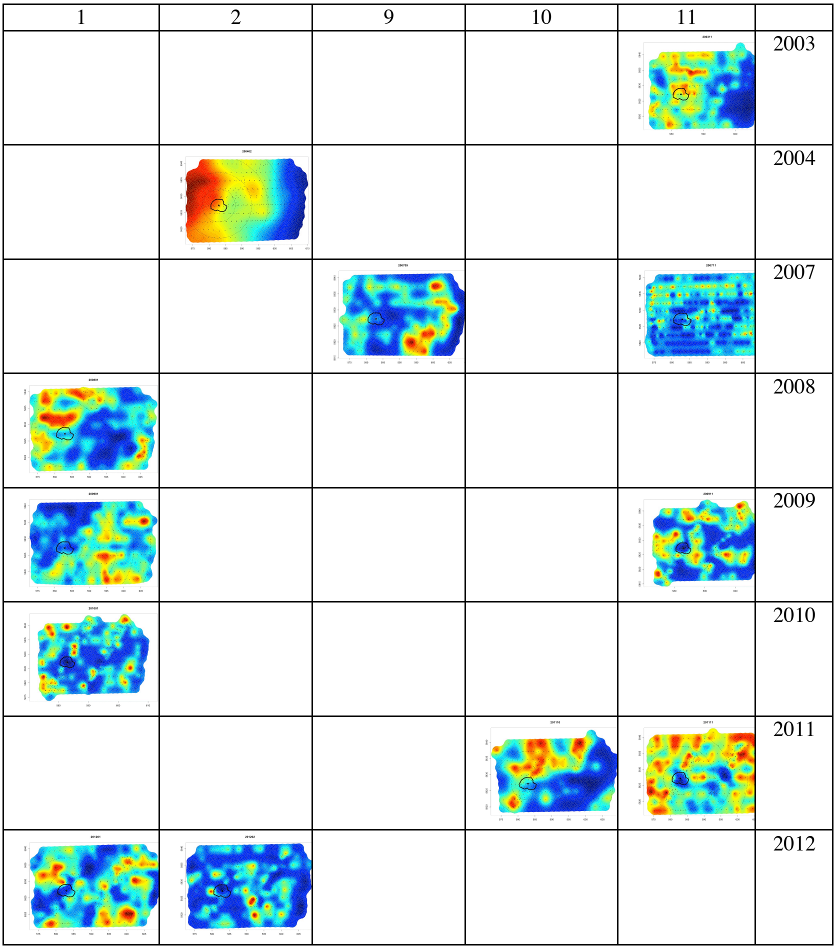

Figure 27. Results for all surveys. The colour red corresponds to above average values and blue to below average values.

7 Linear mixed effects applied on summary statistics

In the previous section we presented the spatial random fields obtained from the analyses of individual surveys. These spatial fields represent the abundances of guillemots. It is quite clear that the spatial patterns differ per survey. Instead of visually comparing these spatial random fields, we defined distance bands around each wind farm, and we calculated the percentage of important w_j in each band for each wind farm. We did this for marginal probabilities and for joint probabilities. We have plotted these in Figure 28 for PAWP and in Figure 29 for OWEZ.

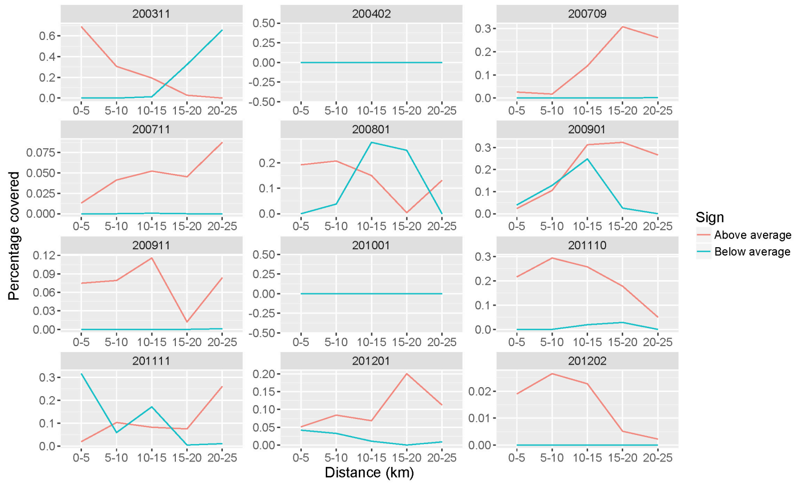

Figure 28. Percentage of coverage per distance band for the marginal probabilities for PAWP.

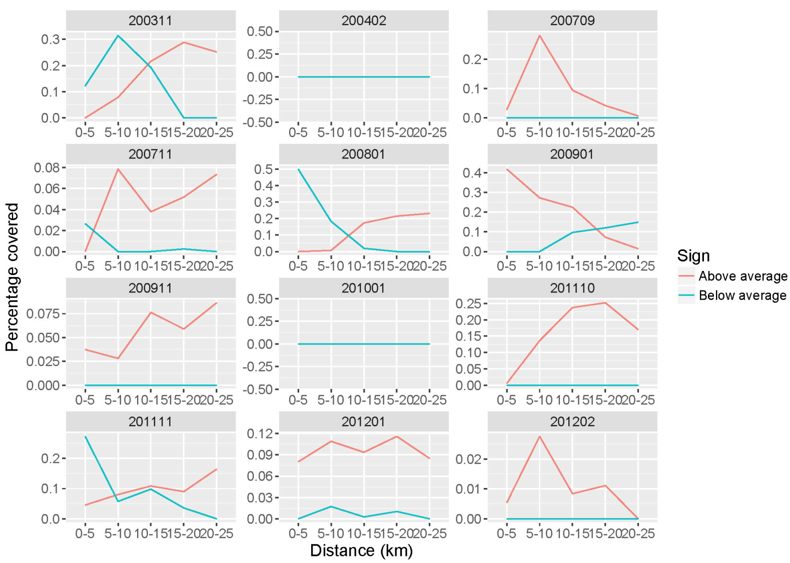

Figure 29. Percentage of coverage per distance band for the marginal probabilities for OWEZ.

Figure 28 shows the percentage of coverage for the marginal distributions plotted versus distance from the centre of the PAWP wind farm. If the PAWP wind farm would have been a disturbance factor for guillemots, then all the red lines would be increasing for larger distances (the further away from the wind farm, the higher the abundance), and the blue lines would all decrease (the closer to a wind farm to lower the abundance). However, this is not the case. To formalise this visual observation we applied a linear mixed effect model on these percentages. To be more precise we modelled the (marginal) percentages as a function of distance. And we used survey as a random effect and we also used distance as a random slope. In normal words, such a model investigates whether there is a relationship between the (marginal) percentages and distance, while allowing for different intercepts and slopes per survey.

We also applied this model on the joint probabilities, and we carried out the analysis for PAWP and for OWEZ.

$$
\begin{aligned}
&Percentage_{ij} \sim N\left(\mu_i, \sigma^2\right) \\
&E\left(Percentage_{ij}\right) = \mu_i = Intercept + \text{Distance}_{ij} + \\
&\qquad\qquad Survey_i + \text{Distance}_{ij} \times Survey_i
\end{aligned} \tag{3.4}
$$

Results for the ‘above average’ data for the marginal probabilities for PAWP are as follows.

```
                Value      SE DF   t-val   p-value
(Intercept)  0.09980 0.05710 47   1.7477   0.0870
Distance    -0.00087 0.01558 47  -0.0561   0.9555
```

This means that in Figure 28, when looking at only the red lines, there is no significant distance effect. The same was done for the blue lines, and for the joint probabilities. In none of the models we had a significant distance effect.

Exactly the same was done for the OWEZ summary statistics, and again there was no significant distance effect in any of the models.

Main point summary: In this section we analysed the percentage of important nodes per distance band. A visual inspection and the application of a linear mixed effects model indicate that there is no distance effect. This means that we did not find a consistent disturbance or preference effect of the two wind farms.

8 Analysis of Robin Rigg data

In this chapter we analyse the guillemot data from the Robin Rigg wind farm, which is based in the UK.

8.1 Data exploration for the Robin Rigg data

The statistical analyses carried out in this supplement are similar to those applied on the PAWP and OWEZ data.

8.1.1 Spatial locations

Figure 30 shows the sampling locations. One of the main differences between this wind farm and the OWEZ/PAWP wind farms is that the Robin Rigg wind farm is in an estuary. The study area of this wind farm is also smaller than the combined study area of OWEZ and PAWP.

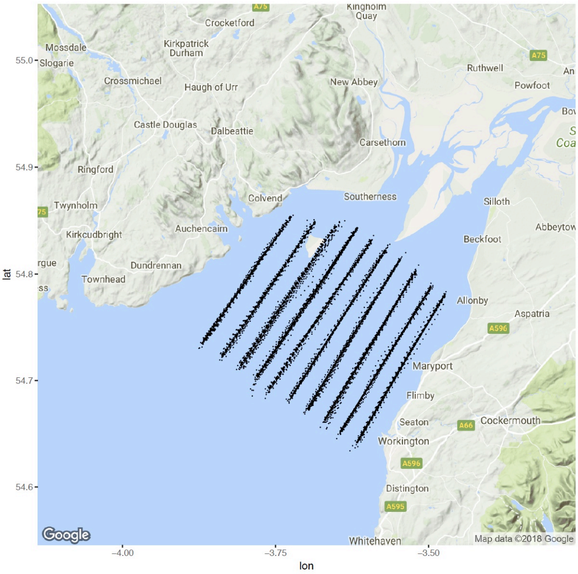

Figure 30. Spatial position of sampling locations for the Robin Rigg wind farm.

Figure 31 also shows the positions of the sampling locations, but this time we superimposed the center of the wind farm as a red dot, and the positions of the outer turbines are represented as a red polygon.

The data used in this supplement are all post-construction, and surveys were taken on a monthly basis from March 2010 to February 2013. Figure 32 shows the sampling locations by month and year. Two differences between the Robin Rigg data and the OWEZ/PAWP data sets are that for Robin Rigg we have more surveys, and these are also regular spaced in time (one every month). This allows for the implementation of an additional statistical model.

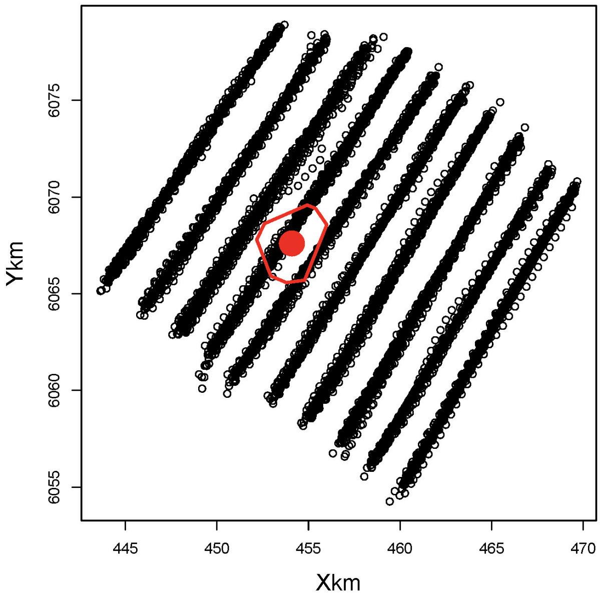

Figure 31. Spatial position of sampling locations for the Robin Rigg wind farm. The red lines represents the outer turbines.

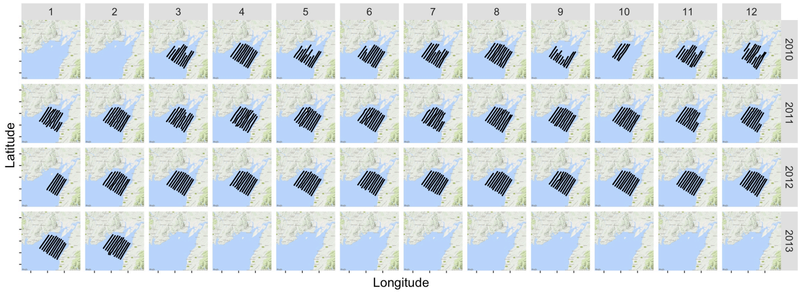

Figure 32. Spatial position of sampling locations by year and month for the Robin Rigg wind farm.

81.2 Zero inflation

One thing that the Robin Rigg data set has in common with the PAWP and OWEZ data sets is zero inflation. Figure 33 shows the zeros and ones per survey.

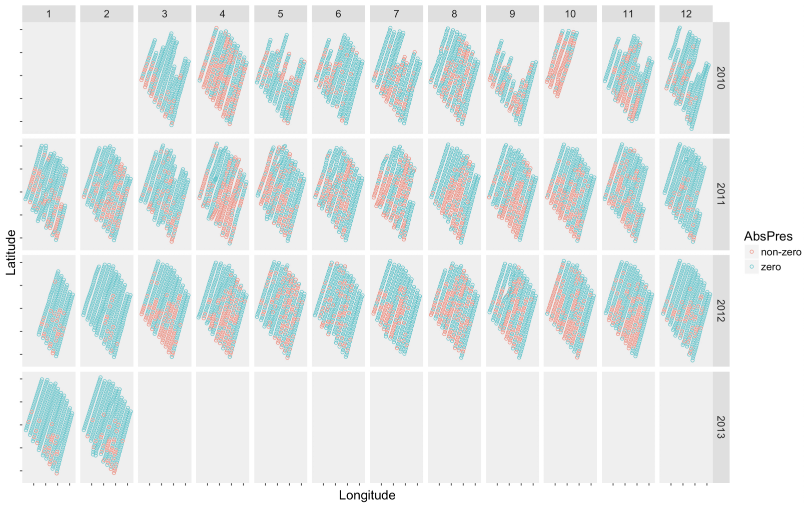

Figure 33. Spatial position of sampling locations by year and month for the Robin Rigg wind farm. A green dot means that zero birds were observed whereas a red dot indicates that the count was larger than 0.

The numerical information below gives the survey number, number of observations per survey and the % of zeros per survey. For the PAWP and OWEZ data we dropped all surveys for which we had more than 75% of zeros. Here we will keep all the data so that we don't get problems with the regular spaced time series nature of the data.

```
    Sample size %zeros
87          260   83.1
88          302   48.7
89          226   77.9
90          268   82.5
91          276   76.4
92          306   72.2
93          183   69.4
94          118   28.8
95          260   69.2
96          258   86.4
97          268   75.7
98          302   71.9
99          287   81.5
100         314   52.9
101         305   64.9
102         297   70.4
103         274   54.7
104         308   67.2
105         302   63.6
```

```
106          303   64.4
107          272   69.9
108          271   89.3
109          185   84.9
110          299   95.7
111          299   66.6
112          303   67.7
113          303   69.3
114          298   80.2
115          299   73.6
116          294   58.8
117          299   71.9
118          301   60.1
119          300   67.0
120          307   84.0
121          296   90.9
122          309   87.7
```

8.2 Statistical models

We applied the same models that were applied on the OWEZ and PAWP data, namely a Poisson GLMM, a zero-inflation GLMM and a negative binomial GLMM. The reason for using the second and third models is the presence of the large number of zeros. Each of these models was applied with the following dependency structures.

1. A GLMM assuming that there is no spatial correlation.
2. A GLMM assuming that there is spatial correlation, but this spatial correlation does not change over time.
3. A GLMM assuming that there is spatial correlation that changes randomly over time. This was the replicate structure.
4. A GLMM assuming that there is spatial correlation that changes over time following an auto-regressive pattern of order 1 (AR1).

The first three models were discussed in detail for the PAWP and OWEZ data. The fourth model, the GLMM with spatial-temporal AR1 correlation, was not applied on the OWEZ and PAWP data due to the irregular temporal nature of those data sets. For the Robin Rigg data, such a model can be applied because we have survey data for each month between 3 March 2010 and 2 February 2013. The model formulation is as follows.

$$
\begin{aligned}
&Birds_{it} \sim Poisson(\mu_{it}) \\
&E(Birds_{it}) = \mathrm{var}(Birds_{it}) = \mu_{it} \\
&\log(\mu_{it}) = \beta_1 + \beta_2 \times LogArea_{it} + x_i + Survey_t \\
&x_{it} = \phi \times x_{i,t-1} + v_t
\end{aligned}
$$

We assume that the numbers of birds sampled at location i in survey t are Poisson distributed with the mean μ_{it}. The mean is then modelled as a function of an intercept, sampling effort, a survey effect (modelled as a random effect), and a spatial-temporal correlated random field. For survey t, the spatial random field is x_{it} and is equal to ϕ times the spatial random field from survey $t - 1$ plus pure spatial noise. A high value of ϕ (i.e. close to 1) means that there is strong temporal correlation.

In non-technical jargon, the GLMM with an AR1 spatial-temporal correlation assumes that there is spatial correlation, and the spatial patterns for survey j depend partly on the spatial pattern for survey $j - 1$.

The model can easily be extended with a zero-inflated distribution or a negative binomial distribution to deal with the excessive number of zeros.

8.2.2 Technical information

Figure 34 shows the mesh that was used; it has 2023 vertices. The same priors for the hyperparameters were used as for the PAWP and OWEZ models. The spatial-temporal models require more computing time than for the OWEZ and PAWP data sets. The reason for this is that we have 36 surveys, and for each of them we need to estimate 2023 w's for the spatial random field.

Initial attempts to run models with the AR1 spatial-temporal dependency structure failed due to the large number of nodes (ws) in the mesh. We therefore followed Zuur et al. (2017) who defined knots (time points) and defined the AR1 correlation structure on these knots. We used time knots separated by 4 surveys. In normal words this means that we have a spatial random field for every fourth defined time point (knot), and the (time-) distances between survey and knot are used as weighting factors to calculate the fitted values for each survey.

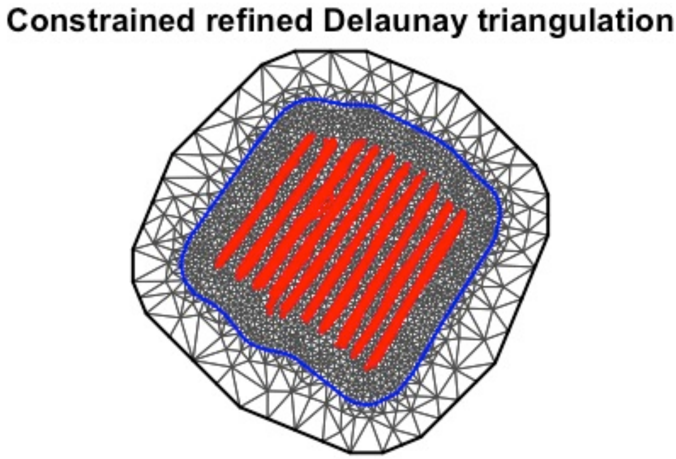

Figure 34. The mesh used by the spatial ZIP model. The mesh has 2023 vertices. The mesh has an inner part (with a finer resolution) and also an outer part to avoid numerical problems due to the boundary.

8.4 Results for all surveys

The DIC and WAIC values of the models without spatial correlation (using a Poisson, ZIP and NB distribution), with spatial correlation, with spatial-temporal correlation using the replicate option, and the spatial-temporal AR-1 correlation structures are presented below.

	DIC	WAIC
Poisson GLMM + no correl.	38848.00	39071.26
ZIP GLMM + no correlation	28392.35	28570.28
NB GLMM + no correlation	22008.70	22019.16
Poisson GLMM + SRF	30240.28	32168.53
ZIP GLMM + SRF	24598.46	25938.54
NB GLMM + SRF	20661.66	20686.35
Poisson GLMM + Replicate SRF	20459.53	22780.97
ZIP GLMM + Replicate SRF	**19646.35**	**20073.20**
NB GLMM + Replicate SRF	**19413.98**	**19337.86**
Poisson GLMM + AR1 SRF	23472.25	30716.84
ZIP GLMM + AR1 SRF	20807.90	21620.99
NB GLMM + AR1 SRF	20277.94	20308.53

The best model, as judged by DIC and WAIC values, is the NB GLMM with the replicate spatial-temporal correlation. And the second best model is the ZIP GLMM with the replicate spatial-temporal correlation.

Recall that in the replicate correlation model, each survey is allowed to have a different spatial pattern, although all spatial patterns share the same statistical parameters that define the spatial correlation. We did try to increase the time resolution for the AR1 models (by not using knots), but the models did not converge.

In the remaining part of this section we will present the results of the NB GLMM with replicate spatial-temporal correlation, though we could also have presented the ZIP GLMM with replicate spatial-correlation.

Figure 35 shows the spatial random fields for each survey obtained by the NB GLMM with replicate spatial-temporal correlation.

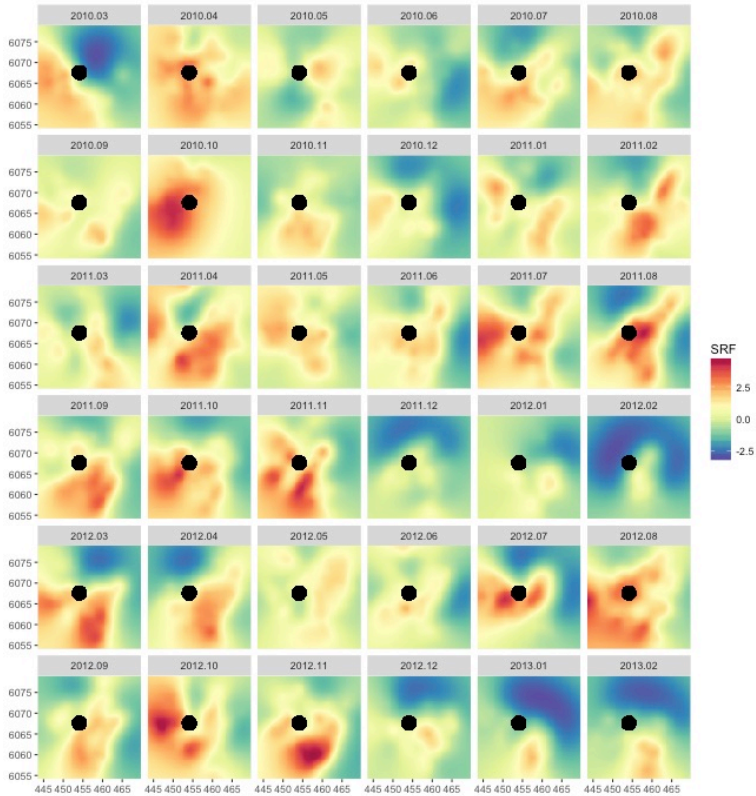

Figure 35. Spatial random fields obtained by the NB GLMM with spatial-temporal replicate correlation applied on Robin Rigg data. Red values correspond to areas with higher abundances and blue values to areas with lower abundances. To obtain the fitted values of the model we have to (i) add the intercept, sampling area effect, random effect survey and the spatial random fields, and then (ii) the exponential function has to be applied. This model does not use time knots.

Note that in some time periods the spatial patterns look similar, whereas in other time periods the spatial patterns look considerably different from survey to survey. This is exactly what the replicate correlation allows for. With the AR1 correlation (and a large ϕ), each survey would have been similar to the previous one. But that is clearly not the case here.

Because we have spatial-temporal correlation that changes by survey we decided to analyse the data from each survey separately in the next section; just as we did for the OWEZ and PAWP data sets.

Model validation was applied on the modelling results of the NB GLMM with the spatial-temporal replicate correlation. Simulation results showed that this model can cope with the large number of zeros. The Pearson residuals still contain a small amount of spatial correlation.

Main point summary: DIC and WAIC values indicate that the spatial correlation differs randomly per survey for the Robin Rigg data. This does not mean that there is no wind farm effect. The results merely state that there are spatial patterns in Guillemots abundances, and these change randomly from survey to survey.

8.5 Results for each individual survey

Just as for the OWEZ and PAWP data, we continued the analysis by applying a model with spatial correlation on data of each individual survey. We have 36 surveys for this data set, and that means that we applied 36 times a model with spatial correlation. We will not present all the 36 spatial random fields here (they look rather similar to those presented in Figure 35).

Recall from the OWEZ and PAWP analyses that R-INLA will estimate a w value on each node of the mesh in Figure 34. We will get 2,023 of those. Some will be positive and others will be negative. Positive values correspond to above average values of the spatial random field and therefore above average fitted guillemot values. To determine whether there is a wind farm effect we drew multiple circles with different radiuses around the center of the wind farm and counted the number of important (as in: 'significant') ws in each distance band. And we converted this into a proportion. We did this for marginal probabilities and for joint probabilities (which does a sort of post-hoc correction).

Although the negative binomial distribution gave the best model when we analysed all surveys in the previous section, it may well be that (for some) individual surveys the ZIP distribution is better. We therefore ran the individual survey analyses with the negative binomial distribution, and also with the ZIP distribution, and we will present results of both models.

The study area for Robin Rigg is smaller than the combined OWEZ and PAWP study area. We therefore had to use smaller distance bands around the center of the wind farm, namely 0 – 3 km, 3 – 6 km, 6 – 9 km and 9 –

12 km. The first distance band (0 – 3 km) covers the entire wind park plus a small area beyond the boundary.

The patterns observed for the marginal and joint probabilities look similar, and here we focus on the marginal probabilities.

We will focus on the positive *w*s values as these are associated with non-zero and non-low counts. Results for the positive *w* values using the ZIP distribution are presented in Figure 36. Recall that positive *w* values correspond to above-average values for birds. Each panel in Figure 36 shows the percentage of important positive *w*'s versus the distance bands for a specific survey. We have conditioned the graphs on month and year. For example, the fourth panel from the left in the top row is for the April 2010 survey. The line in the graph shows that in the 0 – 3 km band, we have 40% of nodes with important (i.e. significant) positive *w* values. In normal words this means that a lot of birds where observed within 3 km of the center of the wind farm. Closer inspection of the corresponding modelling results show that these birds were actually observed just outside the wind farm perimeter, but still within 3 km of the center. A similar pattern was observed in August 2011.

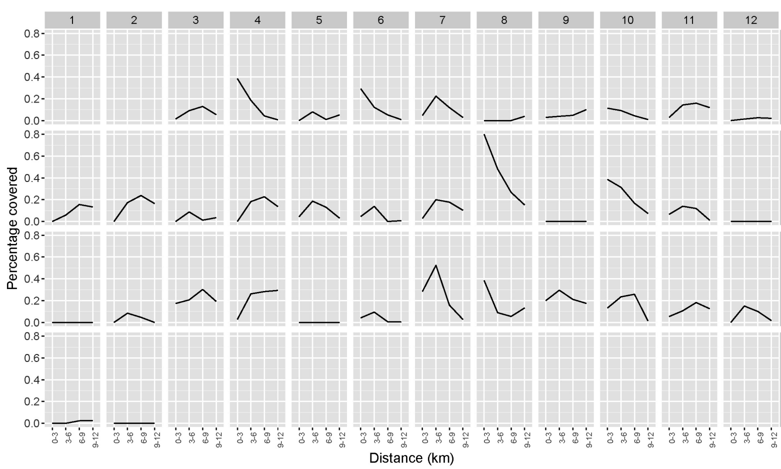

Figure 36. Percentage of coverage of positive (and important) values of the spatial random field per distance band for the marginal probabilities for Robin Rigg for each month and year. Results are obtained with a ZIP GLM with spatial correlation.

The interesting thing in Figure 36 is that for a lot of surveys (e.g. the first 7 surveys in 2011) the first distance band has a low percentage of positive (and important) *w* values, and then this percentage increases. In normal words this means that within the first 3 km we have low (spatially correlated) abundances and the further away we go from the wind farm the higher the abundances (and it then seem to reach a plateau or decreases again). It is tempting to associate this pattern with a wind farm disturbance effect.

We then took all 36 curves in Figure 36 and analysed them with a linear mixed effects model in which we model the percentage of important and positive *w* values as a function of distance (modelled as a categorical covariate) and a random effect survey. Results of this model show that the distance effect is significant ($X^2 = 12.028$, df = 3, $p = 0.007$).

The estimated parameters of this model are as follows.

	Value	SE	DF	t-val	p-val
(Intercept)	0.099	0.020	105	4.892	0.000
fBandsNew3-6 km	0.039	0.021	105	1.844	0.067
fBandsNew6-9 km	0.004	0.021	105	0.227	0.820
fBandsNew9-12 km	-0.035	0.021	105	-1.664	0.098

These results indicate that the percentage of important *w*s in the 0 – 3 km distance band is 9.9%. In the 3 – 6 km distance band this is 3.9% higher. The corresponding *p*-value is 0.067, indicating that this is only a weak effect. Note that these p-values do not indicate whether the 6 – 9 km band is different from the 3 – 6 km band. They only indicate whether each band is different from the 0 – 3 km band. Another criticism is that we use frequentist tools to analyse a summary statistic obtained with Bayesian tools.

We repeated the analyses with the NB distribution and results broadly similar. The percentage of important and positive w values per distance band are plotted for each survey in Figure 37.

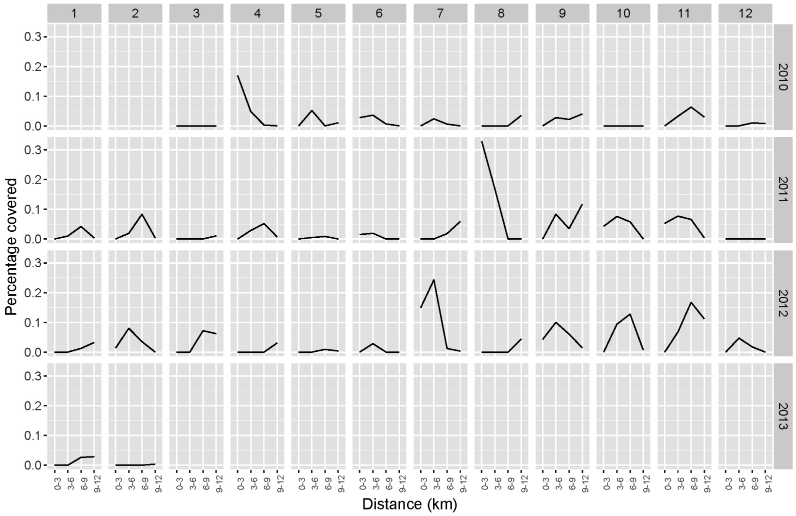

Figure 37. Percentage of coverage per distance band for the marginal probabilities for Robin Rigg. Results are obtained with a NB GLM with spatial correlation.

In this case the linear mixed effects model gives $X^2 = 6.85$ ($df = 3$, $p = 0.076$), indicating again a weak effect of distance.

Main point summary: In this section we analysed the percentage of important and positive *w*s values of the spatial random field per distance band. A visual inspection and the application of a linear mixed effects model on summary statistics (from 36 surveys) indicate that there is a weak distance effect. Higher values of the spatial random field are mostly observed further away from the wind farm (i.e. not in the 0 – 3 km distance band). However, the observed effects have low statistical significance and are obtained via a pragmatic statistical analysis. On the other hand, finding a pattern in this type of data may be considered as a statistical miracle.

References

Blangiardo M, Cameletti M (2015). *Spatial and Spatio-temporal Bayesian Models with R – INLA*. Chichester: John Wiley & Sons.

Bolin D, Lindgren F (2015). Excursion and contour uncertainty regions for latent Gaussian models. *J. R. Statist. Soc. B* 77, Part 1, pp. 85–106

Pinheiro J, Bates D (2000). *Mixed Effects Models in S and S-Plus*. New York: Springer.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rue H, Martino S, Chopin N (2009). Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations. *Journal of the Royal Statistical Society B* 71: 319–392.

Wang X, Ryan YY, Faraway JJ (2017) Bayesian Regression Modeling with INLA. Chapman & Hall/CRC.

Zuur AF, Ieno EN, Walker N, Saveliev AA, Smith GM (2009a). *Mixed Effects Models and Extensions in Ecology*. New York: Springer.

Zuur AF, Ieno EN, Elphick CS (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* 1: 3–14.

Zuur AF, Saveliev AA, Ieno EN (2012). *Zero Inflated Models and Generalized Linear Mixed Models with R*. Newburgh, UK: Highland Statistics.

Zuur AF, JM Hilbe, EN Ieno (2013). *Beginner's Guide to GLM and GLMM with R*. Newburgh, UK: Highland Statistics.

Zuur AF, Ieno EN (2016a). A protocol for conducting and presenting results of regression-type analyses. *Methods in Ecology and Evolution* 7: 636–664.

Zuur AF, Ieno EN (2016b). *Beginner's Guide to Zero-Inflated Models with R*. Newburgh, UK: Highland Statistics.

Zuur AF, Ieno EN, Saveliev AA (2017). Beginner's Guide to Spatial, Temporal and Spatial-Temporal Ecological Data Analysis with R-INLA. Newburgh, UK: Highland Statistics.

Supplement A: Alpha Venus wind farm

In this supplement we show the problems with the analysis of Guillemots abundances for the Alpha Venus wind farm.

A.1 Number of zeros per survey

The table below shows the number of observations for Guillemot abundances per survey and the percentage of zeros per survey.

	Sample size	%zeros
11	808	82.1
12	811	99.8
13	814	99.3
14	812	100.0
15	808	98.6
16	812	93.8
17	808	99.5
18	741	100.0
19	805	98.9
20	808	99.3

Note that the percentage of zeros for each survey is large. Perhaps the first survey can still be analysed but there is no point in analysing Guillemot abundances for the other surveys.

Supplement B: Blighbank wind farm

In this supplement we show the problems with the data of Guillemots abundances for Blighbank wind farm.

B.1 Number of zeros per survey

The table below shows the survey number (year and month), sample size and percentage of zeros of guillemot abundances per survey. Also note that the sampling scheme changed from 10 minutes surveys to 2 minutes surveys in November 2013. The number of spatial sampling locations required for models with spatial correlation is certainly larger than the 30-ish that are available for the 2010 – 2013 data.

	Sample size	%zeros
201010	29	93.1
201011	16	87.5
201012	33	60.6
20111	39	51.3
201110	31	74.2
201111	33	48.5
201112	31	41.9
20112	30	40.0
20113	33	75.8
20114	31	100.0

20115	30	96.7
20116	30	100.0
20117	29	100.0
20118	28	100.0
20119	42	100.0
20121	65	21.5
201210	36	80.6
201211	17	52.9
201212	36	47.2
20122	39	20.5
20123	43	60.5
20124	37	89.2
20125	37	97.3
20126	20	100.0
20127	3	100.0
20129	35	100.0
20131	4	100.0
201311	131	91.6
201312	222	72.5
20132	154	92.9
20133	215	83.7
20134	127	96.1
20135	210	99.5
20137	191	100.0
20138	143	100.0
20139	3	66.7
20141	147	84.4
201410	139	99.3
201411	184	96.7
201412	28	67.9
20142	5	80.0
20143	4	100.0
20144	176	98.9
20146	3	100.0
20147	12	100.0
20148	4	100.0
20149	150	98.0
20151	2	100.0
201512	4	100.0
20152	181	94.5
20154	152	100.0
20159	3	100.0
20161	4	75.0
201612	3	100.0
20162	3	100.0
20163	2	50.0

Surveys with more than 100-ish observations can potentially be used for an R-INLA analysis. But only the December survey from 2013 has enough non-zeros. This means that potentially only data from 1 survey can be analysed.

Supplement C: Thorntonbank wind farm

In this supplement we show the problems with the data of Guillemots abundances for Thorntonbank wind farm.

C.1 Number of zeros per survey

The table below shows the survey number, number of observations and percentages of zero abundances.

	Sample size	%zeros
201210	43	65.1
201211	42	52.4
201212	8	75.0
20131	117	86.3
201311	26	96.2
201312	181	67.4
20132	200	88.5
20133	139	90.6
20134	156	99.4
20135	165	100.0
20136	8	100.0
20137	179	100.0
20138	156	100.0
20139	176	97.7
20141	26	73.1
201410	164	99.4
201411	186	98.9
201412	133	88.0
20142	178	53.9
20143	178	87.6
20144	195	98.5
20146	120	100.0
20147	13	100.0
20148	135	100.0
20149	169	100.0
20151	152	90.1
201512	135	95.6
20152	1	100.0
20154	199	100.0
20155	42	100.0
20157	51	100.0
20158	165	100.0
20159	171	100.0
20161	185	56.8

```
201610           1  100.0
201612         181   93.4
20162          181   68.0
20163          190   88.9
20164           41  100.0
20166           42  100.0
20167           38  100.0
20168           37  100.0
20169          148  100.0
```

There are potentially 2 or 3 surveys that could be used for a spatial analysis.