Seabirds & Offshore Wind Farms

Monitoring Results 2011

Nicolas Vanermen, Eric W.M. Stienen, Thierry Onkelinx, Wouter Courtens,

Marc Van de walle, Pieter Verschelde & Hilbran Verstraete

Study commissioned by the Royal Belgian Institute for Natural Sciences,

Management Unit of the North Sea Mathematical Models

Research Institute for Nature and Forest

Kliniekstraat 25

B-1070 Brussels

Report INBO.R.2012.25

June 2012

1 Introduction

In order to meet the targets set by the European Directive 2009/29/EG on renewable energy, the European Union is aiming at a total offshore capacity of 43 GW by the year 2020. Meanwhile, the offshore wind industry is growing fast and by the end of 2011, 1371 offshore wind turbines were already fully grid-connected in European waters, totalling 3.8 GW (European Wind Energy Association, 2011). The Belgian government has reserved a concession zone comprising almost 7% of the waters under its jurisdiction for wind farming (an area measuring 238 km^2). In 2008, C-Power installed six wind turbines (30 MW) at the Thorntonbank, located 27 km offshore, and in 2009, Belwind constructed 55 turbines (165 MW) at the Blighbank, 40 km offshore. In the first coming years at least 175 more turbines will be installed in this part of the North Sea (MUMM, 2011).

Possible effects of offshore wind farming on seabirds range from direct mortality through collision, to more indirect effects like habitat change (including positive effects of increased food availability and resting opportunities), habitat loss and barrier-effects (Exo *et al*. 2003, Langston & Pullan 2003, Fox *et al*. 2006, Drewitt & Langston 2006, Stienen *et al*. 2007). Whereas several studies investigated the effects of offshore turbines on migrating or local seabird communities (Desholm 2005, Petterson 2005, Petersen *et al.* 2006, Larsen & Guillemette 2007), only a few papers focussed on the monitoring protocol to assess these effects (Maclean *et al.* 2006 & 2007, Pérez-Lapeña *et al.* 2010 & 2011).

The Research Institute for Nature and Forest (INBO) is in charge of monitoring the effects of these wind farms on the local seabird distribution. Therefore, it designed a BACI monitoring program and delineated impact and control areas for both wind farm projects. INBO performs monthly seabird surveys in these areas, and developed an impact assessment methodology accounting for the statistical problems inherent to ‘seabirds at sea’ (SAS) data.

2 Methodology

Based on a peer review we revised our methodology (as compared to the one presented in Vanermen *et al*. 2011), the most crucial difference being the application of zero-inflated negative binomial modelling, instead of quasi likelihood estimation. We performed power analyses to investigate how the power of our impact study is affected by survey length, monitoring intensity and data characteristics. Lastly, we applied the proposed methodology for assessing seabird displacement effects caused by the early presence of the C-Power and Belwind wind farms.

2.1 BACI monitoring set-up

Stewart-Oaten & Bence (2001) reviewed several approaches for environmental impact assessment, differing in goals and time series available. When 'before' data are available and the inclusion of a suitable control is possible, BACI is the suggested approach. While the importance of temporal replication in BACI assessments is widely recognized, there is disagreement on the role of spatial replication, i.e. inclusion of several control locations (Bernstein & Zalinski 1983, Stewart-Oaten *et al*., 1986, Underwood 1994, Underwood & Chapman 2003, Stewart-Oaten & Bence 2001). In a 'seabirds at sea' (SAS) context, including more than one control area is unfeasible, considering the obvious logistic and financial limitations. However, Stewart-Oaten & Bence (2001) argue that when the goal of the assessment is to detect a particular change at a non-random place (e.g. the Thorntonbank wind farm), variation among control sites is irrelevant to the assessment problem. The authors conclude that multiple controls are not needed, but can be useful for insurance, model checking and causal assessment.

Migrating birds show deflections in flight orientation from up to a distance of 1 to 5 km (Petterson 2005, Petersen *et al.* 2006), but little is known on the avoidance of swimming birds. Yet, a significant post-construction decrease in densities of divers, scoters and Long-tailed Ducks was shown by Petersen *et al.* (2006) out to a distance of 3 km away from the Nysted wind farm in Denmark. Considering this, we applied a buffer zone of 3 km around the future wind farms to define the 'impact area' (Figure 1), being the zone where effects of turbine presence can be expected. Next, an equally large control area was delineated, harbouring comparable numbers of seabirds, showing similar environmental conditions, and enclosing a high number of historical count data (Vanermen *et al*. 2010). Considering the large day-to-day variation in observation conditions and seabird densities, the distance from the control to the impact area was chosen to be small enough to be able to survey both areas on the same day by means of a research vessel. As a result, control and impact area are only 1.5 km apart, equalling half the mean distance sailed during a ten-minute transect count (the applied unit in our seabird database).

Considering the fact that the construction of the wind farms is far from completed (55 out of 110 turbines at the Blighbank and 6 out of 54 turbines at the Thorntonbank at the time of data collection), the impact area regarded at this stage is limited to the zone where turbines are already present, surrounded by a buffer zone of 3 km (see Figure 1). Also, data collected during the construction periods are not included for impact assessment. During construction activities, access to the wind farm areas was often restricted, hampering adequate monitoring. Moreover, construction activities may cause other effects to occur than the ones during the operational phase. Recently, access to the wind farms has greatly improved, e.g. during construction of phase 2 & 3 of the C-Power wind farm.

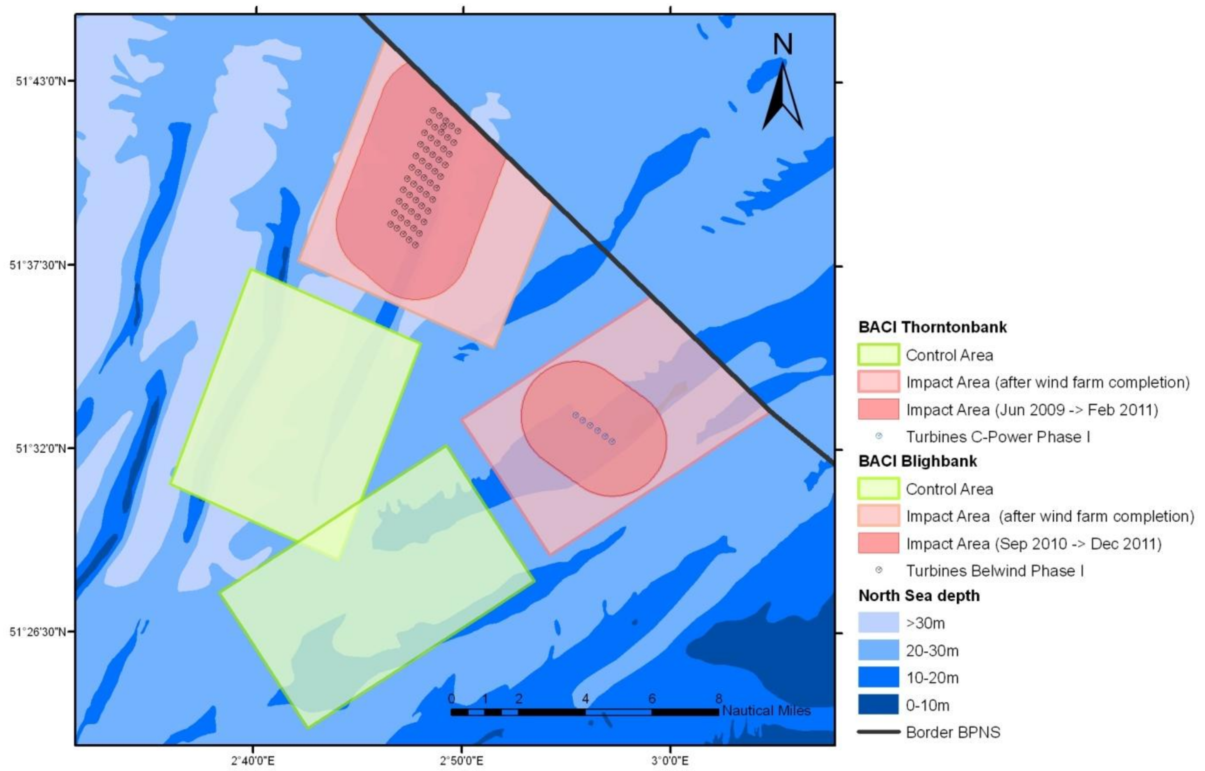

Figure 1. BACI set-up for the monitoring at the Thorntonbank & Blighbank wind farms.

The first turbines at the Thorntonbank were erected in 2008, and the reference period includes all data collected up until March 2008. INBO started monthly monitoring of the study area in 2005, but has data available dating back to 1993. In total, 64 surveys were included in the reference dataset - with two counts per area per survey this results in a sample size (N) of 128. Construction activities continued until May 2009, and meanwhile access to the area was restricted. Impact data hence include all observations collected from June 2009 to February 2011 (after which construction activities for phase 2 took place), totalling 33 impact surveys (N=66).

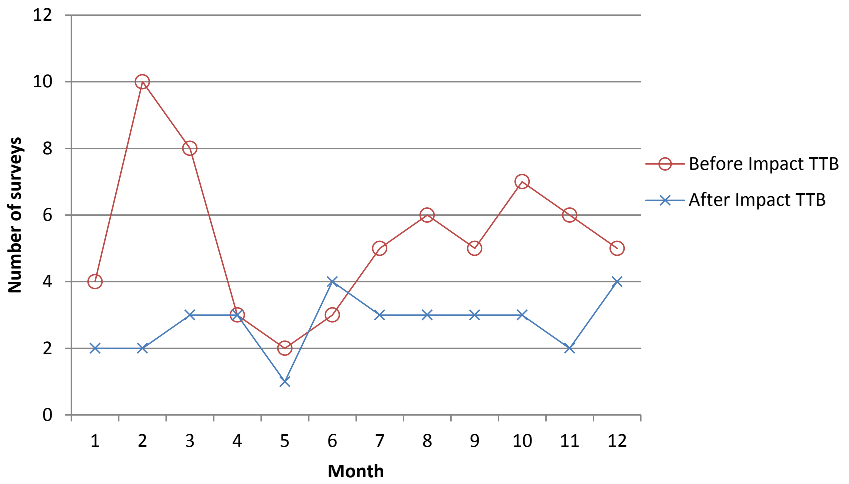

Figure 2. Count effort at the Thorntonbank study area, with indication of the number of surveys performed before and after the construction of the first turbines.

At the Blighbank construction activities started in September 2009, prior to which INBO performed 73 reference surveys (N=146). The last of 55 turbines was built in September 2010, and from that month on, impact monitoring was performed inside the wind farm. The impact period includes all data collected from September 2010 to December 2011 (totalling 16 surveys – N=32).

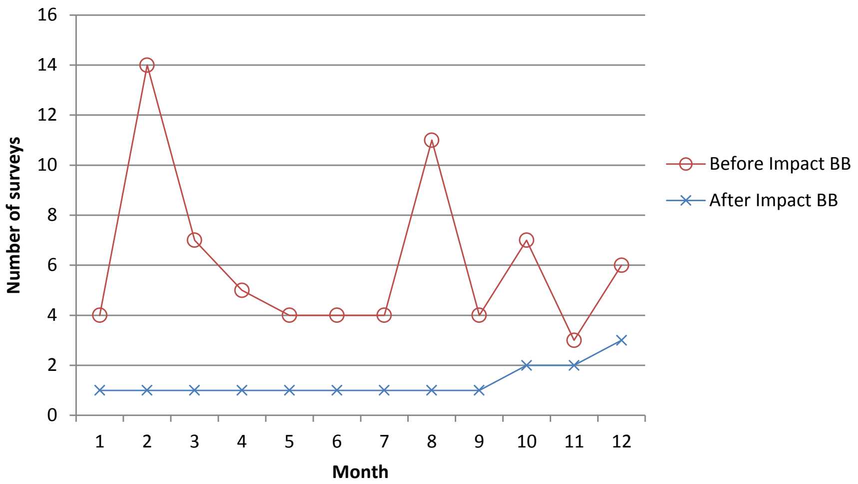

Figure 3. Count effort at the Blighbank study area, with indication of the number of surveys performed before and after the construction of the first turbines.

2.2 Ship-based seabird counts

Both in the impact and control areas, monitoring was performed through ship-based seabird counts. These are conducted according to a standardized and internationally applied method (Tasker *et al*. 1984, Komdeur *et al*. 1992). While steaming, all birds in touch with the water (swimming, dipping, diving) located within a 300 m wide transect along one side of the ship's track are counted ('transect count'). For flying birds, this transect is divided in discrete blocks of time. During one minute the ship covers a distance of approximately 300 m, and right at the start of each minute we count all birds flying within a quadrant of 300 by 300 m inside the transect ('snapshot count'). Taking into account the distance travelled, these count results can be transformed to seabird densities. The applied count unit in our seabird database is the result of so-called 'ten-minute tracks'.

Stewart-Oaten *et al*. (1986) state that in BACI-assessments, any information gained from replicates taken at the same time is not useful, and that it is better to consider one summarised value (observation X_{ijk}) for each time (t_{ij}), in period i (Before/After) and at place k (Control/Impact). Accordingly, we summed our transect count data per area (Control/Impact) and per monitoring day, resulting in day-totals. This way, we avoided pseudo-replication, and minimized overall variance. It is also advised to take samples in the impact and control area simultaneously (Stewart-Oaten *et al*. 1986), and so we included only those days at which both areas were visited, minimizing variation due to short-term temporal changes in seabird abundance and in weather and observation conditions. Today, the monitoring routes always include both of these areas, but this was not always the case in our historical data.

We used data on thirteen seabird species occurring regularly in the Thorntonbank and Blighbank wind farm areas (see Table 1).

Table 1. Species included in the assessment of displacement effects caused by wind turbines.

Species	Thorntonbank	Blighbank
Northern Fulmar (*Fulmarus glacialis*)	X	X
Northern Gannet (*Morus bassanus*)	X	X
Great Skua (*Stercorarius skua*)		X
Little Gull (*Hydrocoloeus minutus*)	X	X
Common Gull (*Larus canus*)	X	X
Lesser Black-backed Gull (*Larus argentatus*)	X	X
Herring Gull (*Larus fuscus*)	X	X
Great Black-backed Gull (*Larus marinus)*	X	X
Black-legged Kittiwake (*Rissa tridactyla*)	X	X
Sandwich Tern (*Sterna sandvicensis*)	X	
Common Tern (*Sterna hirundo*)	X	
Common Guillemot (*Uria aalge*)	X	X
Razorbill (*Alca torda*)	X	X

2.3 Data-analysis: Reference modelling

The data collected prior to the construction of the turbines were modelled during the so-called ‘reference modelling’. There are several ways in which SAS-data can be modelled, using generalized linear models (Leopold *et al*., 2004, Maclean *et al*. 2006 & 2007), quasi-likelihood estimation (McDonald *et al*. 2000), generalized additive models (Clarke *et al*. 2003, Karnovsky *et al*. 2006, Huettmann & Diamond 2006, Certain *et al*. 2007), or combining one of these with geostatistics (Pebesma *et al*. 2000, Pérez-Lapeña *et al*. 2010 & 2011). When a counted subject is randomly dispersed, count results correspond to a Poisson-distribution (McCullagh & Nelder 1989). However, as seabirds often occur strongly aggregated, we applied a negative binomial (NB) distribution, being the standard parametric model used to account for over-dispersion (Potts & Elith 2006). Another common problem in ecological data is an excess in zero counts (Fletcher *et al*. 2005). We tested if our data were in fact zero-inflated, and performed preliminary tests to compare the performance of a NB model with a zero-inflated NB model (ZINB), both in terms of predictive value as of resulting power (Zeileis *et al*. 2008, Wenger & Freeman 2008). Zero-inflated models consist of two components, a count component modelling the positive count data (in this case according to a negative binomial distribution), and a zero-component modelling the excess of zeros.

Despite the data aggregation to day totals, it seemed that for several species the count data were still zero-inflated. Preliminary tests learned that in this case, the ZINB models performed better compared to NB models, both in terms of the predicted model probability as in terms of power. On the other hand, when comparing the ZINB with NB model results for non-zero-inflated data, coefficient estimates and corresponding P-values are highly similar, and power results are unaffected by the choice of model (further illustrated in the §3.1.2, and Figure 6). During this explorative part of the study (reference modelling, data simulation and power analyses) we therefore chose to apply one type of model, being the zero-inflated type, as a base for all data simulations and consequent power calculations, making it easier to compare and interpret the obtained results.

Whether counts were performed in the control or impact area is defined in the count component of the models by the factor variable ‘CI’ (Control-Impact). We also added seasonality as an explanatory variable since seabird occurrence is subject to large seasonal fluctuations. Seasonal patterns can be described through a sine curve, which can be modelled as the linear sum of a sine and a cosine term (Stewart-Oaten & Bence 2001, Onkelinx *et al*. 2008), including ‘month’ as a continuous variable. We did not allow for interaction between area (CI) and seasonality since differences in seasonal patterns are not likely to occur at such a small scale.

As described above, the response variable equals the total number of birds observed (inside the transect) during one monitoring day in either the control or impact area. To correct for varying monitoring effort, the number of km² counted is included in the model as an offset-variable. The count component of the ZINB model is thus of the following form:

$$\log(response) = offset(\log(km^2)) + a_1 + a_2.\sin\left(2\Pi\frac{month}{12}\right) + a_3.\cos\left(2\Pi\frac{month}{12}\right) + a_6.CI \quad \text{(Eq. 1)}$$

In Eq.1, seasonality is modelled as a sine curve with a period of 12 months. Several migratory species however show two peaks in density per year. For these species another sine curve with a period of 6 months is added, and the reference model can thus be written as:

$$\log(response) = offset(\log(km^2)) + a_1 + a_2.\sin\left(2\Pi\frac{month}{12}\right) + a_3.\cos\left(2\Pi\frac{month}{12}\right) + a_4.\sin\left(2\Pi\frac{month}{6}\right) + a_5.\cos\left(2\Pi\frac{month}{6}\right) + a_6.CI$$

(Eq. 2)

Lastly, the zero-component of the ZINB model is built up solely from an intercept (b_1), linked to response by a logit-function. Back-transformation of this intercept results in the additive chance of encountering a zero-value (e.g. an intercept of 1 corresponds to a chance of 73.1%).

The resulting reference model is selected through backward model selection, first testing for the area-effect CI, and then testing for the seasonality-effect, considering an ANOVA test-statistic, and comparing the AIC-values of the different models.

2.4 Power analysis

The power analysis as presented in this report is based on the reference data collected in the Thorntonbank study area (see also §2.1). The power is estimated by simulating random datasets with pre-defined characteristics, e.g. the model parameters as found during the reference modelling (§2.3), and imposing a hypothetical change on the post-construction numbers. This change in numbers is supposed to occur throughout the impact area, immediately after the impact, and to persist as long as turbines are present ('press disturbance' – Underwood 1992, Underwood & Chapman 2003).

The model to determine a turbine impact is a simple extension of the count component of the selected reference model:

$$response \sim Seasonality + CI + BA + BA:CI \quad \text{(Eq. 3)}$$

Or – when the factor variable CI was already rejected from the reference model – the impact model looks somewhat different:

$$response \sim Seasonality + BA + T \quad \text{(Eq. 4)}$$

In both equations, 'Seasonality' is the sine wave described earlier and the two-level factor variable BA stands for Before/After the impact. In Eq.3, a turbine effect is indicated by the amount of interaction between BA with CI, while in Eq.4, this effect is indicated by factor T (which stands for turbine presence versus absence).

2.4.1 Power analysis: effect of model parameters

To be able to isolate the effect of the several model parameters, we first modelled the reference data applying the same reference ('base') model for all species (Eq. 1). This revealed empirical ranges of the intercept (a_1), the amplitude of seasonality (= $\sqrt{a_2^2 + a_3^2}$), the CI-effect (a_6), the amount of zero-inflation (b_1) and theta (θ). The latter is part of the variance function of a negative binomial distribution:

$$V(\mu) = \mu + \frac{\mu^2}{\theta} \quad \text{(Eq. 5)}$$

Next, we varied all of these coefficient values within the given ranges, and calculated the power for each scenario. At this stage, the monitoring set-up is held constant, with a reference and impact period of both 5 years, one survey per month (with an effort of 10 km² per area), a decrease in numbers of 50% and a significance level of 10%. This significance level represents the chance of wrongly concluding that the turbines are causing an impact, while in fact they are not ('type I error'). Each scenario is simulated 1000 times, and the power thus equals the percentage of times the z-test reveals a P-value less than 10% for the BA:CI or T-term, indicating a turbine effect.

2.4.2 Power analysis: effect of survey duration and degree of seabird displacement

In a second step we calculated powers based on species-specific reference models (as explained in §2.3), varying monitoring set-up characteristics, i.e. the decrease in numbers in the impact area to be detected (25, 50 & 75%) and the monitoring period (5 years before versus 1, 3, 5, 7, 9, 11, 13 & 15 years after impact).

2.5 Data-analysis: Impact modelling

During the impact modelling we analysed all collected count data to investigate whether the presence of wind turbines is causing seabird displacement. As outlined in §2.4, the applied impact model is a simple extension in the count component of the reference model (Eq. 3 & 4). While we applied a ZINB model for all species during the explorative phase, we now considered each species separately to decide whether to use the ZINB or NB model. Two criteria can be used to do so:

- The P-value of the zero-component intercept: the null hypothesis of the z-test testing for the effect of the intercept is that b_1 equals zero. Back-transformation of an intercept value of zero however corresponds to a chance of 50%, which can be classified as a high degree of zero-inflation.
- A Vuong test (Vuong, 1989): a test that compares non-nested models, as is the case here with a NB model and its zero-inflated analogue. The sign (+/-) of the test-statistic indicates which model is superior over the other in terms of probability. However, in most cases, the corresponding P-value appeared to be indecisive.

Hence, none of these two options gave satisfactory results. Therefore, we defined our own criterion and calculated the lower boundary of the confidence interval of the zero-component intercept: when this lower boundary exceeds -2.2 (corresponding to an additive chance of 10% to encounter zero birds), we decided to hold on to the ZINB model. The choice made as such largely corresponds to what one would expect based on the sign (+/-) of the Vuong test-statistic.

2.6 Statistics

All modelling was performed in R.2.14.0 (R Development Core Team 2011), making use of the following packages:

- MASS (Venables & Ripley 2002)
- pscl (Zeileis *et al*. 2008, Jackman 2011)

3 Results

3.1 Reference modelling & Power analyses

3.1.1 Base modelling: coefficient estimates

First, we applied the same ‘base model’ (Eq. 1) to all species, providing us with empirical coefficient ranges. Based upon these, we defined unique coefficient combinations, which are applied in the ‘test models’. As such, the intercept a_1 of the count component was varied stepwise from -4 to 0. The amplitude was varied by setting a_3 to zero and varying a_2 from 1 to 4, again in discrete steps of one unit. Figure 4 displays the empirical model coefficients, as well as the ones used for the ‘test models’. In order to be able to fully exclude the effect of seasonality, we also combined an amplitude of 0 with an intercept varying from -4 to 2.

Next, we defined an empirical range for theta, as well as for b_1, indicating zero-inflation. The base modelling revealed an interaction between the theta-value and the amount of zero-inflation. For data showing no zero-inflation ($b_1 < -5$), theta was small, varying between 0.18 and 0.66, while in data subject to zero-inflation ($b_1 > 0.5$), theta-values were clearly higher, ranging from 0.48 to 1.40. This is interesting, because it suggests that in the latter case, over-dispersion is (at least partly) captured by the zero-component. Thus we combined a b_1-value of -10 (zero-inflation=0%) with a theta varying by 0.2, 0.4 & 0.6, and a b_1-value of 1 (zero-inflation=±75%) with a theta varying by 0.6 & 1.2.

Combining all of these parameters, we end up with 135 theoretical scenarios. This enables us to isolate and explore the effect of the different model parameters on the power of our impact analysis, given a certain monitoring set-up (i.e. to detect a decrease in numbers of 50% after 10 years of monitoring, i.e. 5 year before and 5 years after the impact).

Until now, the area-coefficient a_6 was fixed at zero, but the base models showed this coefficient to vary between -1.02 and 1.25. As a last step, we calculated the effect of the CI-factor on the resulting power by varying a_6 with -1, 0 and 1.

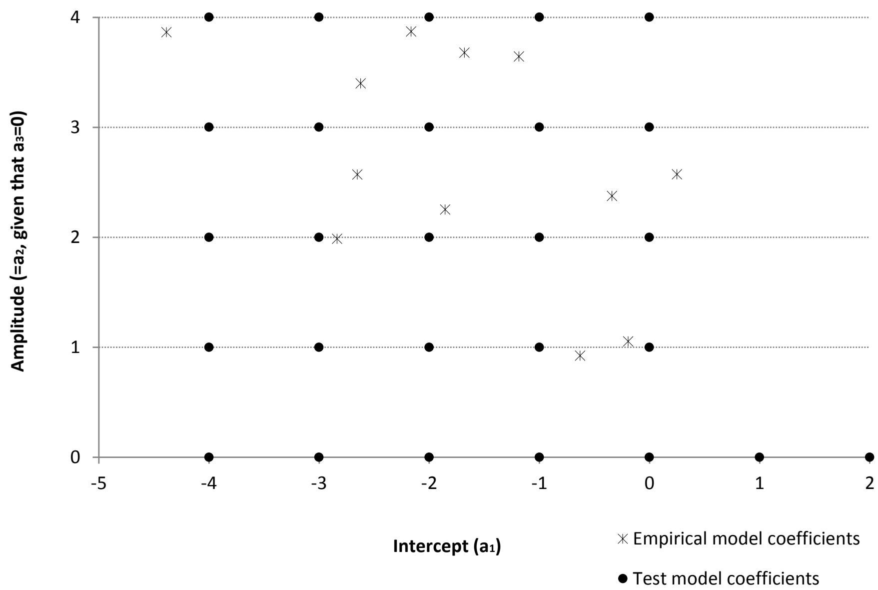

Figure 4. Values for the intercept (a_1) and amplitude (equalling a_2 as a_3 is set to zero) as used in the test models, and indication of the empirical values as found in the reference data collected in the Thorntonbank study area.

Since all of these model coefficient values are linked to the response variable by a logarithmic link function, they are difficult to interpret. Therefore we visualize the corresponding predicted densities for 8 unique combinations of intercept and amplitude (Figure 5).

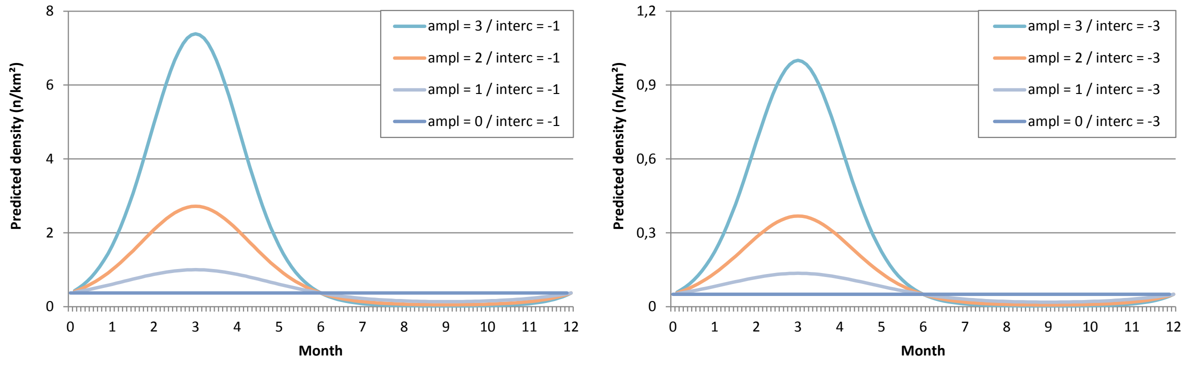

Figure 5. Predicted densities (n/km²) when applying to 8 unique combinations of intercept and amplitude values as used in the test models (see also Figure 4).

3.1.2 Power analysis: effect of model parameters

We calculated the power for 135 scenarios with varying intercept, amplitude, theta and amount of zero-inflation, as determined in §3.1.1.

Zero-inflation has a clear negative effect on the power of the impact study (Figure 6). It is also shown that when non-zero-inflated data are simulated (intercept of the zero-component = -10), equal powers are obtained when comparing NB and ZINB models. When we do include zero-inflation in the data simulation (b_1=0 or b_1=1, corresponding to a zero-inflation of 50 & 73%), the ZINB model clearly performs better. We hypothesise that this is due to fact that over-dispersion can now be captured by the zero-component, instead of being fully absorbed by the theta value.

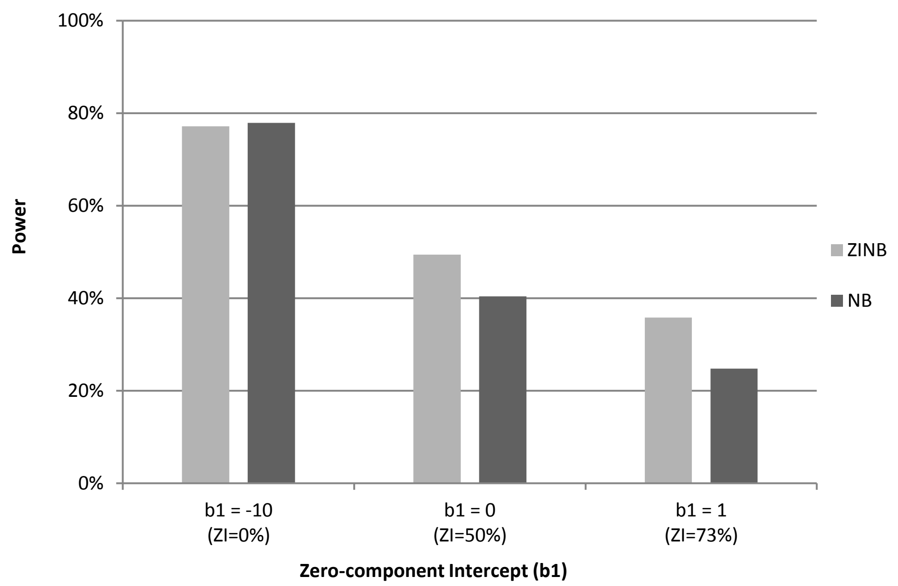

Figure 6. Comparison of the power to detect a 50% decrease in numbers based on a negative binomial (NB) and a zero-inflated model (ZINB), for several levels of zero-inflation (a_1=-1, a_2=1, a_3=0, a_6=0, θ=0.5).

The results show that θ is another important parameter influencing the power of our impact analysis (Figure 7). A theta of 0.2 or less inevitably results in low power after five years of post-impact monitoring, and assuming no zero-inflation is present, a value of 0.4 is needed to obtain a power of 80%.

Base modelling showed that for some species, the reference data combine a seemingly favourable theta with a certain amount of zero-inflation. The power-curve "θ=0.6 / ZI=73%" in Figure 7 shows that all benefits gained from a favourable theta are lost due to zero-inflation. As θ continues to rise, power results start to catch up ("θ=1.2 / ZI=73%"), but still do not exceed the powers found for the scenarios "θ=0.2 / ZI=0%" and "θ=0.4 / ZI=0%".

Based on Figure 7, we also see that the intercept is positively correlated with resulting power, which is particularly true for intercepts ranging from -4 to 0. Increase in power levels off when the intercept exceeds zero, corresponding to a seabird density of 1 bird/km². Due to strong seasonality, the intercepts estimated for our reference data were in fact all below or around zero (Figure 4).

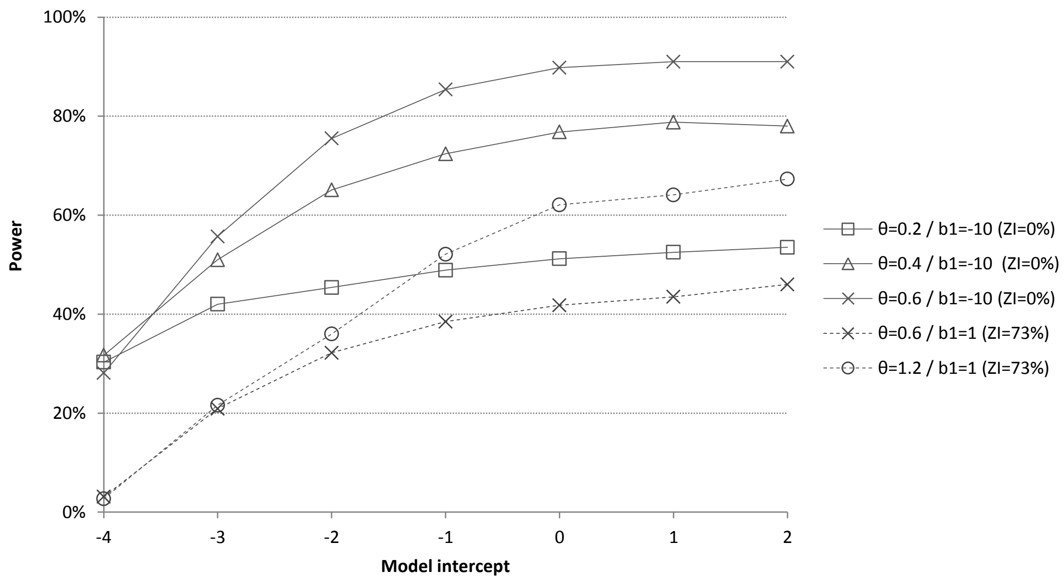

Figure 7. Effect of the model intercept, theta (θ) and the amount of zero-inflation (ZI) on the power of the impact analysis (for test models with a seasonal amplitude equalling zero).

The amplitude of the modelled seasonality pattern appears to have a rather limited effect on the power to detect a change in numbers. We found a positive correlation between the amplitude and power in case of very low intercepts (<-3), and a slightly negative correlation in case of higher intercepts (Figure 8).

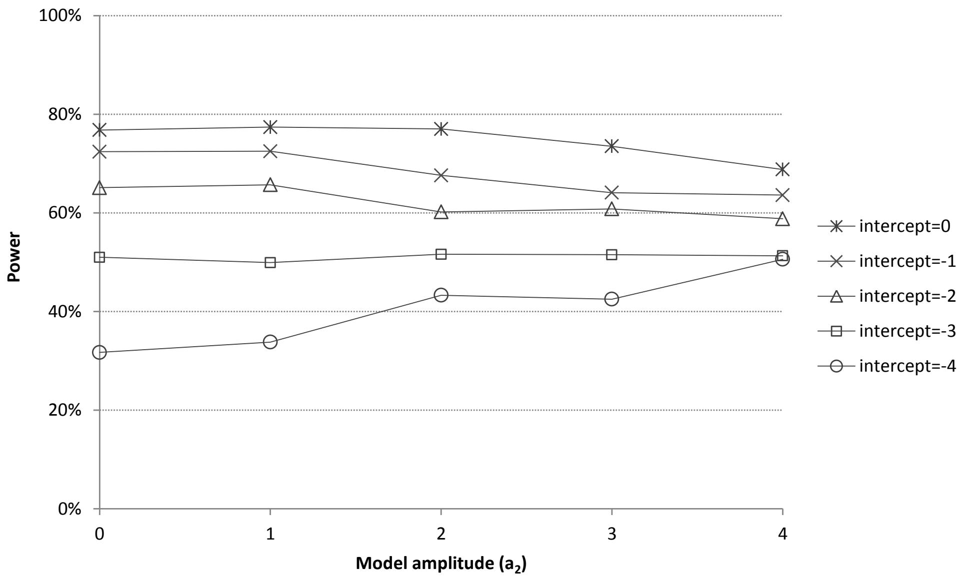

Figure 8. Effect of the seasonal amplitude (equalling a_2 as a_3 is set to zero) and the model intercept (a_1) on the power of the impact analysis (when θ=0.4).

Finally, we investigated the effect of the area factor (CI). For the same relative decrease in numbers (50%), we simulated datasets with varying CI-coefficients a_6 (-1, 0 & 1), and calculated the power based on two different types of impact models. One model takes in account the imposed CI effect (see Eq. 3), while the other one ignores it (Eq. 4). Figure 9 shows the importance of including the CI-factor into the model. When doing so, the power results are much more stable (and hence reliable) compared to the results when the CI-effect is ignored. Of course, when the CI-factor does not attribute significantly to the reference model (P>0.10), it can and should be excluded, as the resulting gain in 2 degrees of freedom will always reflected by better power.

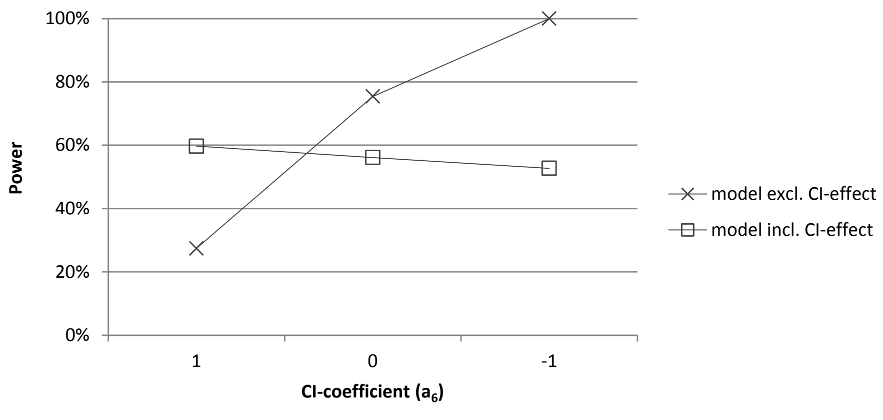

Figure 9. Comparison of power results for two types of models (including or excluding an area effect – see Eq. 3 & 4) for several levels of CI-coefficient a_6 (a_1=-1, a_2=1, a_3=0, θ= 0.5).

3.1.3 Species-specific reference models (Thorntonbank)

We built species-specific reference models (as set out in §2.4.2) and Table 2 shows all estimated coefficients. Considering their specific seasonal occurrence in the study area, we used a double sine curve to explain seasonal variation in numbers for four species, i.e. Northern Gannet, Little Gull, Sandwich Tern and Common Tern. The occurrence of all other species was described by using a single sine curve. In only two out of twelve species, we retained a significant area-effect i.e. for Common Gull (a_6=1.26) and Black-legged Kittiwake (a_6=-0.87).

Back-transformation of the intercept values b_1 of the model's zero component (IntZero) shown in Table 2 learns that zero-inflation occurs in the data of Northern Fulmar (54.0%), Sandwich Tern (52.2%) and Common Tern (74.8%). For the two latter species, theta values are high (3.68 & 11.05), suggesting that most of the over-dispersion is captured by the zero-component. In all other species zero-inflation is very close to 0%. Figure 10 displays the seasonally varying model predictions for all 12 seabird species.

Table 2. Model coefficients of the selected reference models at the Thorntonbank.

	IntCount	Sin (1yr)	Cos (1yr)	Sin (1/2yr)	Cos (1/2yr)	CI	IntZero	θ
Northern Fulmar	-0.83	-1.08	0.17				0.16	0.27
Northern Gannet	-0.82	-0.65	0.26	-0.60	-0.54		-10.55	0.37
Little Gull	-3.35	1.67	3.75	-1.28	-0.84		-3.46	0.22
Common Gull	-4.39	2.00	3.30			1.26	-10.85	0.21
Lesser Black-backed Gull	0.07	1.09	-2.33				-11.09	0.22
Herring Gull	-2.75	1.77	0.78				-7.70	0.20
Great Black-backed Gull	-1.52	-0.30	2.30				-10.19	0.18
Black-legged Kittiwake	-0.36	-1.10	2.13			-0.87	-12.94	0.26
Sandwich Tern	-8.90	0.48	-11.00	1.18	-6.39		0.09	3.64
Common Tern	-10.54	-1.25	-13.61	-0.93	-7.24		1.09	11.03
Common Guillemot	-1.29	0.56	3.63				-11.59	0.65
Razorbill	-2.50	-0.16	3.39				-11.12	0.32

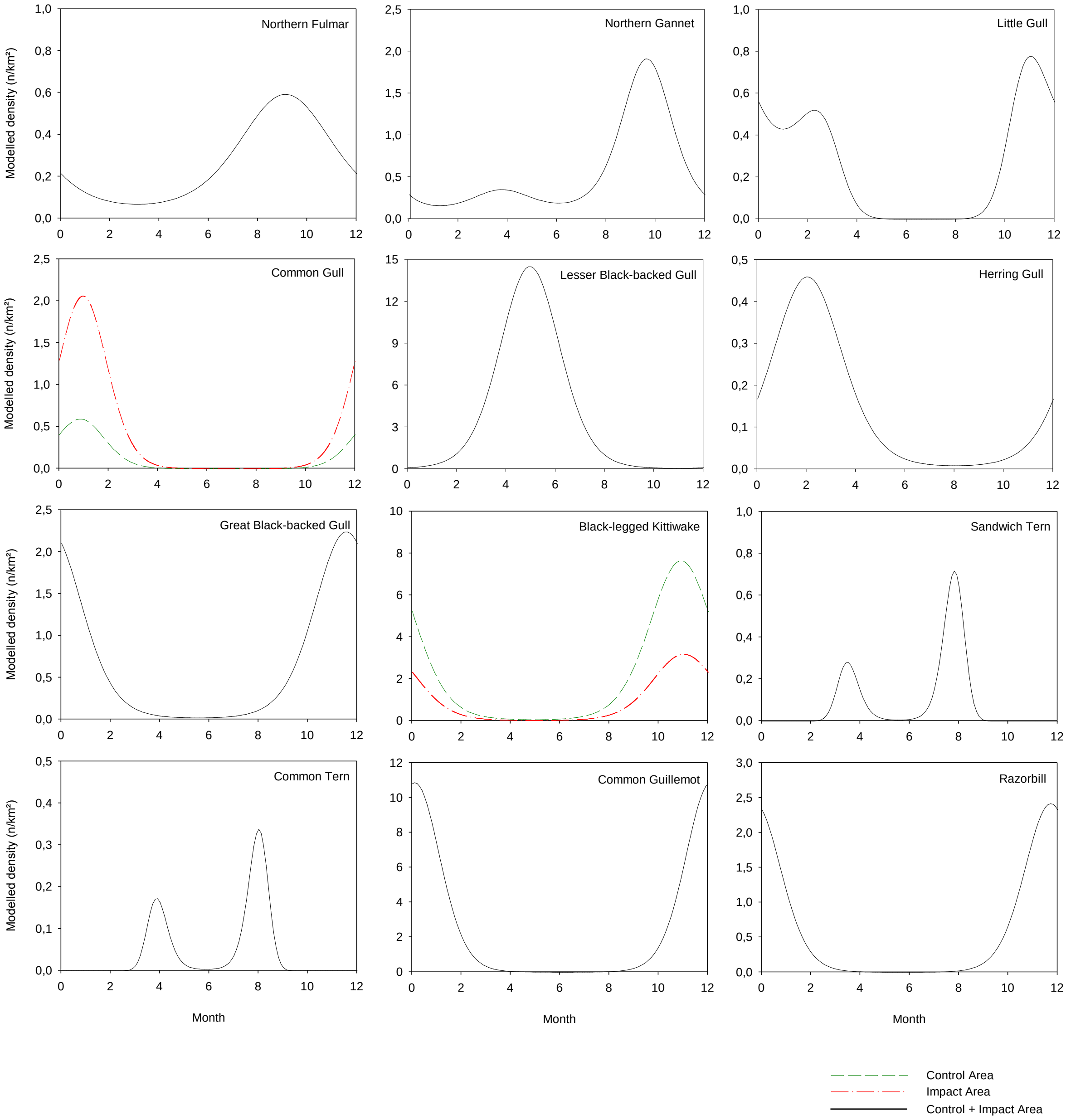

Figure 10. Modelled densities of 12 seabird species, based on data collected at the Thorntonbank study area prior to the construction of the wind farm.

3.1.4 Power analysis: effect of survey duration and degree of seabird displacement

Based on the selected reference models, we studied how power is related to survey duration (Figure 11). We found that for none of the 12 seabird species under study, we will be able to detect a change in numbers of 25% with a power of more than 55%, not even after 15 years of impact monitoring. In contrast, a change in numbers of 50% should be detectable within less than 10 years with a chance of >90% in two seabird species i.e. Northern gannet and Common guillemot. Within the same time frame we will be able to detect a decrease of 75% with a power >90% in all species except for Common Gull.

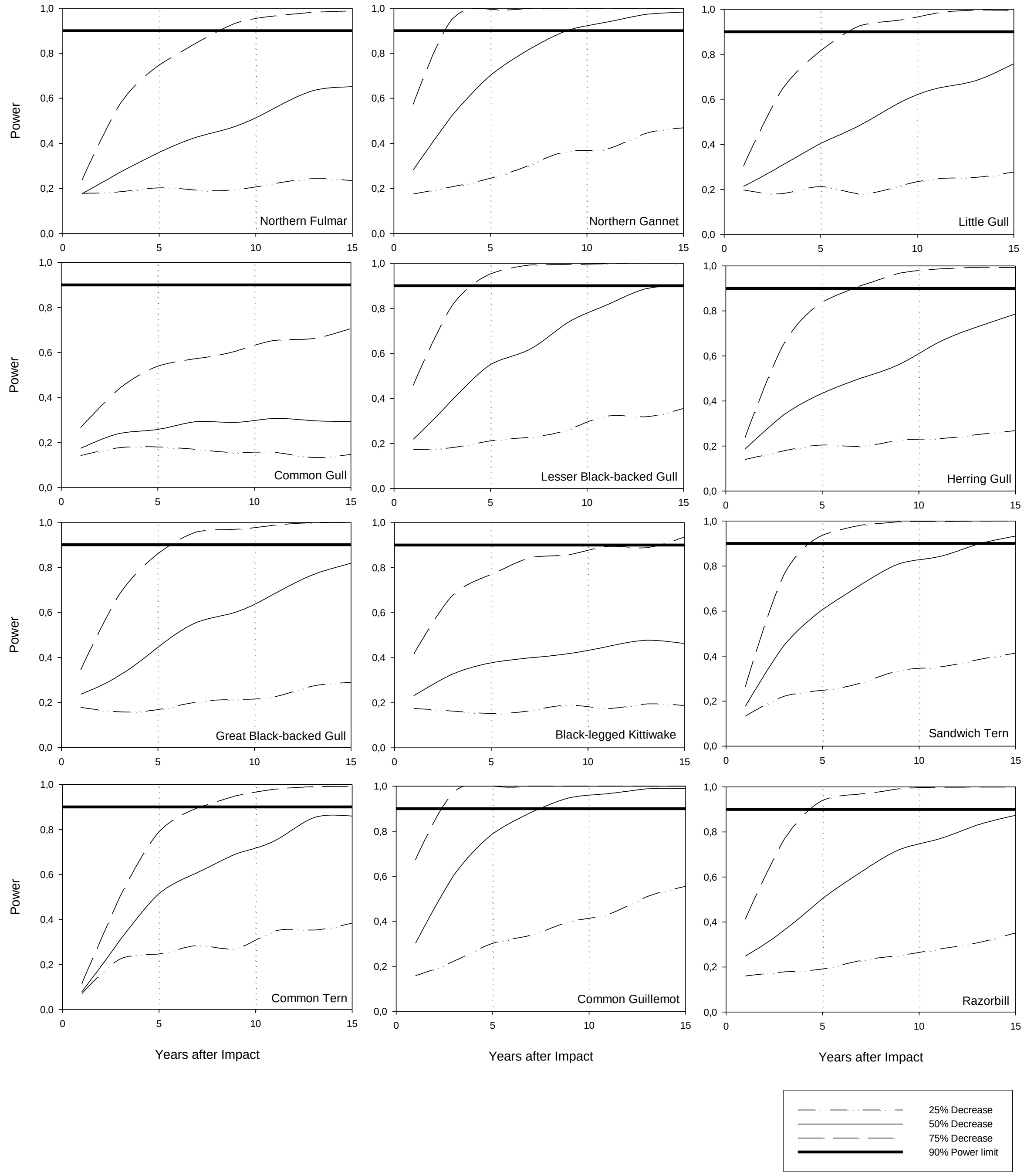

Figure 11. Power results for 12 seabird species for an impact study with a monitoring intensity of one survey of 10km² per month per area, and 5 years of reference monitoring (significance level = 0.10).

3.2 Impact modelling

3.2.1 Thorntonbank

The impact modelling at the Thorntonbank study area only reveals attraction effects, i.e. for Little Gull, Great Black-backed Gull, Black-legged Kittiwake and both tern species.

Figure 12 shows typical BACI-graphs displaying 4 geometric mean density values. These graphs give a first indication of attraction or avoidance effects, but these might as well be hidden. For example, based on the BACI-graphs, it is relatively obvious that there must have been an effect on the occurrence of Little Gull, Sandwich Tern & Common Tern. However, this is much less obvious based on the graphs of Great Black-backed Gull and Black-legged Kittiwake, showing that the impact modelling process reveals effects that otherwise could be hard to detect.

Table 3. Impact modelling results for the Thorntonbank wind farm.

		T – effect		BA:CI – effect	
		Coeff	P-Value		
Northern Fulmar	ZINB	-13,63	0,986		
Northern Gannet	NB	-0,71	0,127		
Little Gull	NB	**1,22**	**0,084.**		
Common Gull	NB			-1,43	0,101
Lesser Black-backed Gull	NB	-0,13	0,809		
Herring Gull	NB	0,37	0,566		
Great Black-backed Gull	NB	**1,49**	**0,023***		
Black-legged Kittiwake	NB			**2,01**	**0,005***
Sandwich Tern	ZINB	**2,43**	**0,001****		
Common Tern	ZINB	**2,42**	**0,028***		
Common Guillemot	NB	-0,17	0,710		
Razorbill	NB	0,43	0,480		

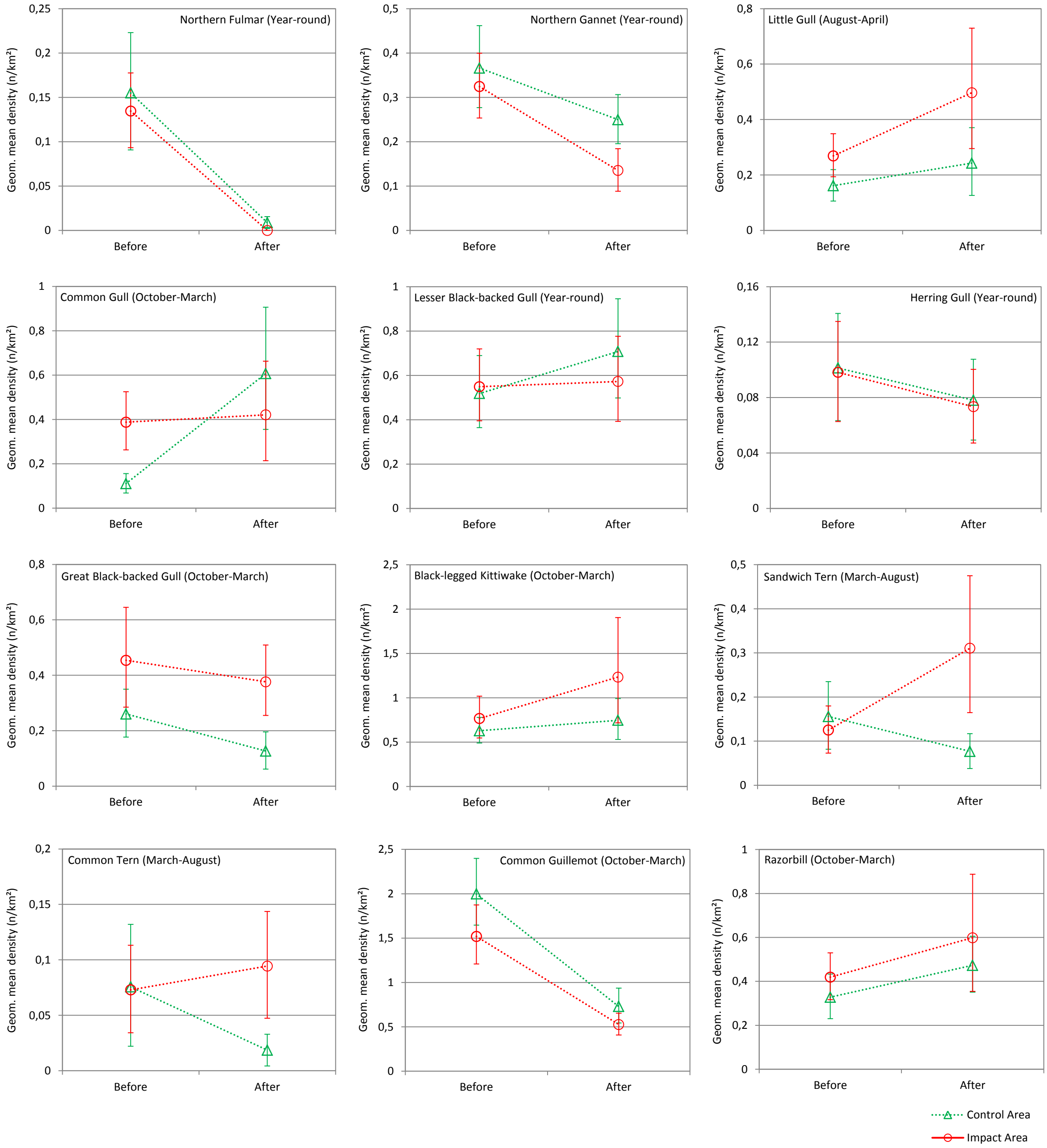

Figure 12. Geometric mean seabird densities (+/- std. errors) in the reference and impact area before and after the turbines were built at the Thorntonbank.

3.2.2 Blighbank

Reference modelling revealed a significant area effect for three species, i.e. Little, Common and Great Black-backed Gull. All three showed higher densities in the impact area compared to the reference area. The data of Great Skua, Little Gull and Common Gull appear to be zero-inflated (75-80%). As in the reference data at the Thorntonbank, a positive intercept in the zero-component is accompanied with a high theta value in the count component, suggesting that overdispersion is captured by the zero-component of the model. For the non-zero-inflated data, theta varies between 0.10 and 0.58. Analogous to the reference data at the Thorntonbank, the two most favourable theta values are found in the count data of Common guillemot (0.58) and Northern Gannet (0.40), while the least favourable theta (0.10) is put away for Great Back-backed Gull. The only species where we modelled a double-peaked seasonality is Northern Gannet (Figure 13).

Table 4. Model coefficients of the selected reference models at the Blighbank.

	IntCount	Sin (1yr)	Cos (1yr)	Sin (1/2yr)	Cos (1/2yr)	CI	IntZero	θ
Northern Fulmar	-1.71	0.94	0.84				-8.23	0.14
Northern Gannet	-1.50	-0.16	1.50	0.01	-0.96		-10.13	0.40
Great Skua	-1.88						1.09	4.76
Little Gull	-12.30	11.26	-1.09			1.83	1.29	1.63
Common Gull	-3.24	1.24	2.82			0.71	1.44	97828.37
Lesser Black-backed Gull	-1.08	0.52	-0.67				-9.48	0.17
Herring Gull	-4.58	2.51	1.42				-7.34	0.33
Great Black-backed Gull	-2.80	1.64	1.73			2.24	-9.90	0.10
Black-legged Kittiwake	-1.13	0.18	2.56				-11.21	0.27
Common Guillemot	-1.69	1.15	3.00				-11.32	0.58
Razorbill	-4.07	1.79	3.45				-7.99	0.29

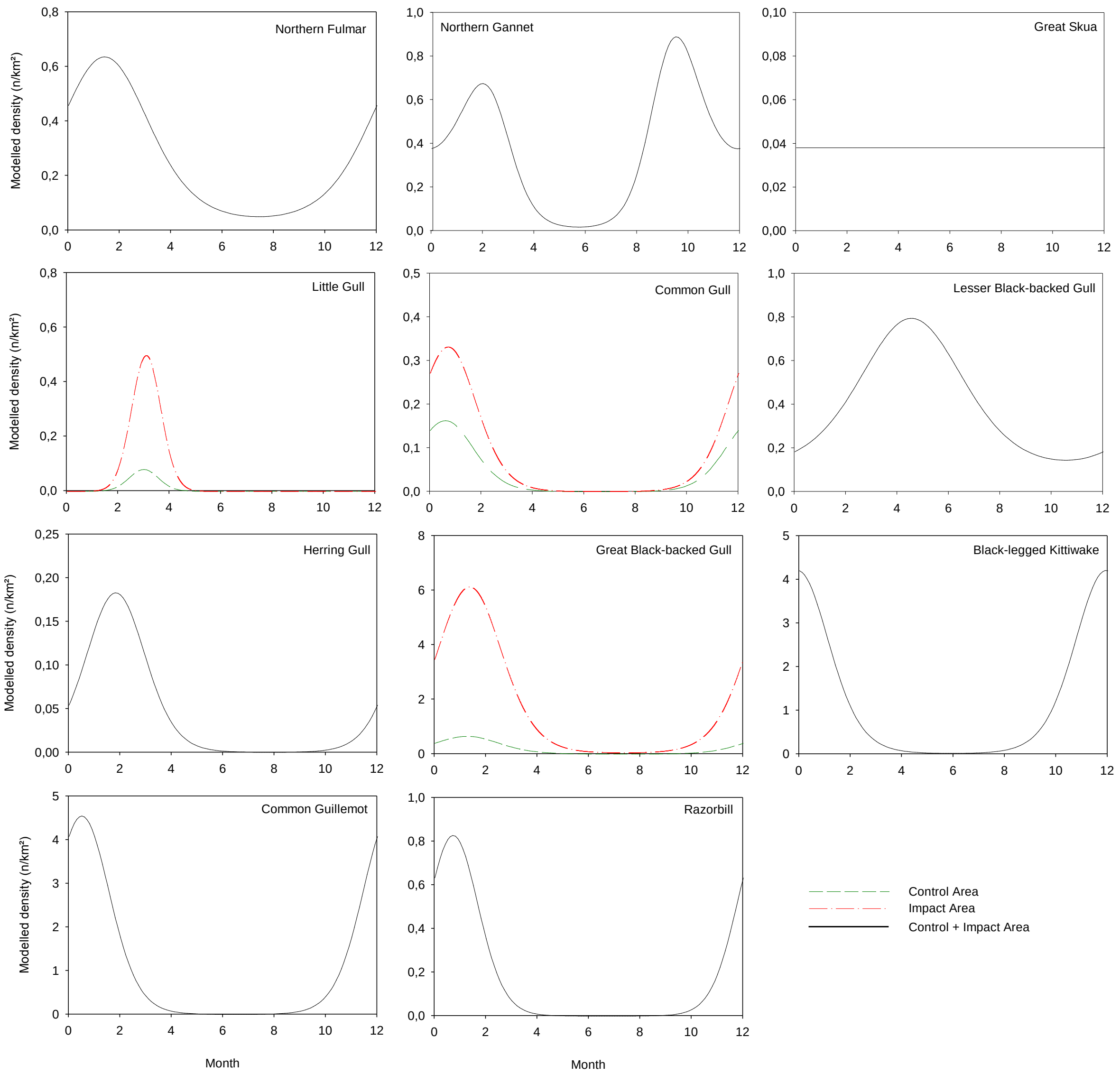

Figure 13. Modelled densities of 11 seabird species, based on data collected at the Blighbank study area prior to the construction of the wind farm.

In the impact data, zero-inflation persisted in the count results of Great Skua and Common Gull, while this was no longer the case for Little Gull. On the other hand, we did use a ZINB model for Herring Gull, since a NB model was unable to fit.

After the turbines were built, numbers of Common Guillemot and Northern Gannet significantly decreased in the wind farm area, while numbers of Common Gull increased. These trends are also obvious when looking at the BACI-graphs in Figure 14. Based on the BACI-graph of Herring Gull, we could have expected a positive turbine effect, but this was not detected by our statistical modelling (P=0.209).

Table 5. Impact modelling results for the Blighbank wind farm.

		T – effect		BA:CI – effect	
		Coeff	P-Value	Coeff	P-Value
Northern Fulmar	NB	-28.60	1.000		
Northern Gannet	NB	**-1.50**	**0.016***		
Great Skua	ZINB	-14.86	0.995		
Little Gull	NB			-0.79	0.643
Common Gull	ZINB			**3.04**	**0.026***
Lesser Black-backed Gull	NB	0.14	0.871		
Herring Gull	ZINB	1.34	0.209		
Great Black-backed Gull	NB			-0.55	0.653
Black-legged Kittiwake	NB	0.56	0.444		
Common Guillemot	NB	**-1.15**	**0.046***		
Razorbill	NB	-1.29	0.127		

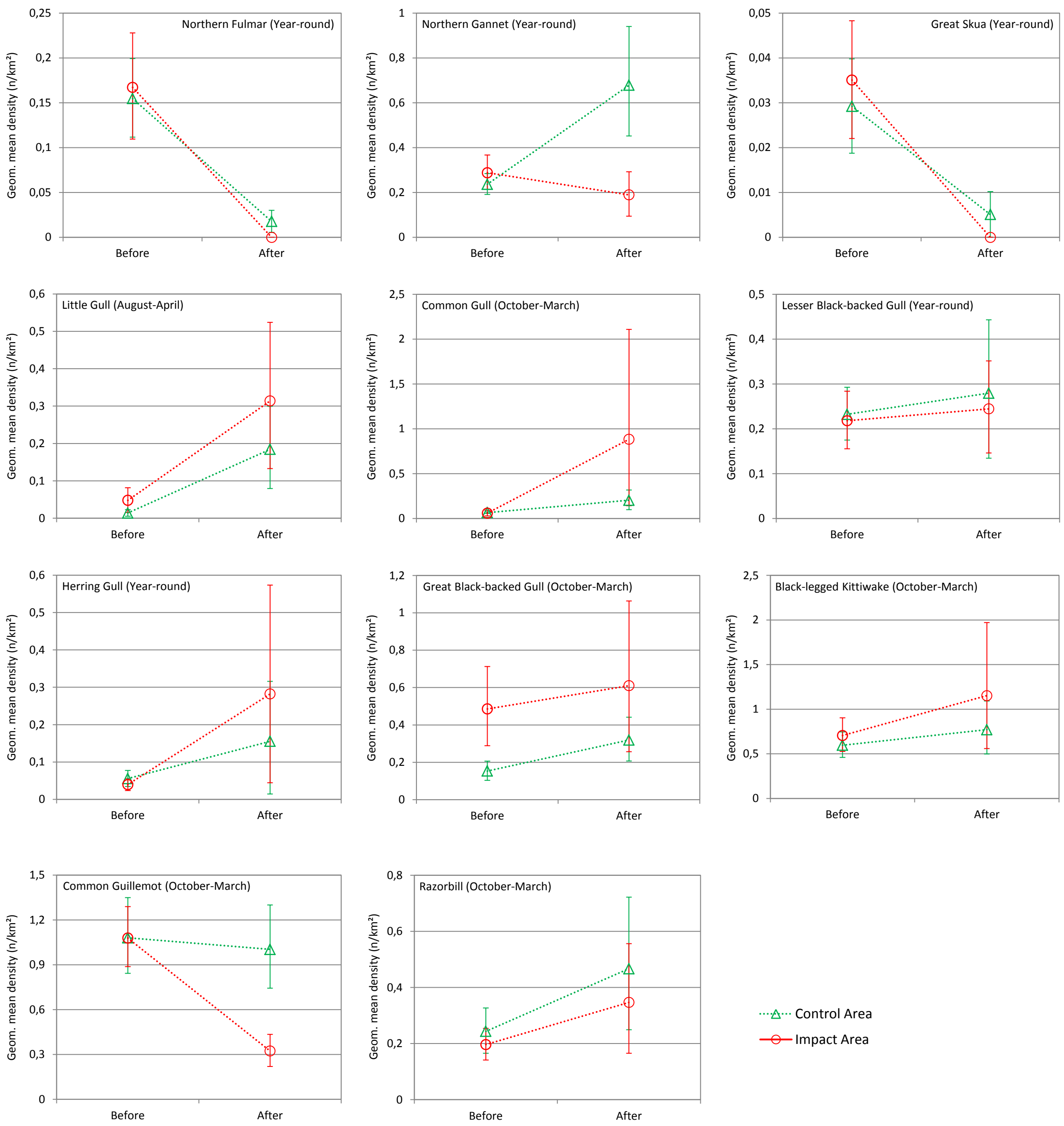

Figure 14. Geometric mean seabird densities (+/- std. errors) in the reference and impact area before and after the turbines were built at the Blighbank.

4 Discussion

Impact assessment

The impact modelling at the Thorntonbank study area only reveals attraction effects, i.e. for Little Gull, Great Black-backed Gull, Black-legged Kittiwake and both tern species. These findings are highly provisory since it is mathematically impossible to count *inside* a one dimensional wind farm (i.e. one line of wind mills). At best, any conclusions drawn from the study presented here are valid *for a wind farm buffer zone* (in this study set to 3 km).

At the OWEZ wind farm in the Netherlands, Little Gulls are rarely seen inside the wind farm and seemed to avoid the area between the turbines, and the same was concluded for Sandwich Tern (Leopold *et al*., 2010). At the Horns Rev wind farm in Denmark, Petersen *et al*. (2006) found slightly increased (non-significant) post-construction numbers of Little Gull inside the wind farm, and a significant increase in numbers just outside its boundaries (up to 2 km). The same authors found a total absence of Common Tern inside the wind farm, avoidance up to 1 km outside its boundaries, but a clear post-construction increase in numbers in the immediate vicinity of the farm (1 to 8km). This is in correspondence to what was found in this study, and meanwhile, the findings at Horns Rev stress the need to perform separate analyses for the wind farm and the buffer zone around it!

Nevertheless, if the attraction effects as found now should persist during the following wind farm phases, this is of serious conservational importance. Both tern species as well as Little Gull are included on the Annex I list of the Birds Directive (EC/2009/147), and high proportions of the biogeographical populations of all three species migrate through the Southern North Sea (Stienen *et al*. 2007).

After the turbines were built at the Blighbank, numbers of Common Guillemot and Northern Gannet significantly decreased in the wind farm area. In correspondence, avoidance by gannets and auks is reported by Petersen *et al*. (2006) at the Horns Rev wind farm in Denmark, and by Leopold *et al*. (2010) in the OWEZ wind farm in the Netherlands.

In contrast, numbers of Common Gull significantly increased, and the BACI-graphs suggest attraction of Herring Gull as well. While gulls are known at least not to avoid the wind farms, attraction effects could not be proven during the Danish and Dutch monitoring program (Petersen *et al*. 2006, Leopold *et al*. 2010). Spatial distribution of gulls is strongly influenced by fishery activities, which makes it very difficult to discern and correctly interpret any changes in distribution patterns. In this respect, the main effect of wind farms on gull distribution patterns is likely to result from the prohibition for trawlers to fish inside their boundaries (Leopold *et al.* 2010).

Nevertheless, despite the absence of beam trawlers, all gull species were regularly observed between the turbines. Gulls are probably attracted by the wind farm from a sheer physical point of view, with the farm functioning as a stepping stone, a resting place or a reference feature in the wide open sea. During recent surveys in 2012, good numbers of auks and even Harbour porpoises were encountered inside the wind farm. From an ecological point of view, the presence of auks is very interesting, and we wonder if these self-fishing species are already habituating to the presence of the turbines, and if they will profit from a (hypothetical) increase in food availability (Degrear *et al*. 2011).

Data handling

Traditionally, the applied count unit in SAS-research is the result of a 5- or 10-minute track, geo-referenced in the middle point (following Tasker *et al*. 1984, Komdeur *et al*. 1992). However, when collected during the same day, these rather short transect counts are likely to be pseudo-replicates which are not independent (Stewart-Oaten *et al*. 1986, Pebesma *et al*. 2000, Karnovsky *et al*. 2006). Therefore we condensated our transect count data to day totals per area.

Based on these binned data, we applied a negative binomial (NB) distribution to predict seabird densities in the study area. In case of highly over-dispersed data, the use of a NB distribution is to be preferred over a quasi-poisson distribution, as used in Vanermen *et al*. (2010) (Zuur *et al*. 2009). Moreover, simulating a (continuous) quasi-poisson distribution, implies the simulation results to be rounded to the nearest integer, which in the end may result in false power results. Seasonal variation was modelled by fitting a sine curve to our data, enabling us to include 'month' as a continuous variable in the models. This method performed much better compared to the inclusion of 'month' as a factor variable, which splits the data in twelve subsets, resulting in highly unreliable coefficient estimates. In order to explain spatial variation in seabird distribution and abundance, environmental variables are often included in the assessment modelling (e.g. Garthe 1997, Pebesma *et al*. 2000, Karnovsky *et al*. 2006, Huettmann & Diamond 2006, Maclean *et al.* 2006 & 2007, Oppel *et al*. in press). However, in this study, any variation in seabird numbers induced by environmental gradients is excluded through the aggregation of our data per day and per area, while the difference between both areas is described by a two-level factor variable ('CI'). The last challenge in the modelling process was dealing with zero-inflation, as SAS-data – and ecological data in general – are often characterised by an excess in zero-counts (Fletcher *et al*. 2005). We investigated if this was also the case in our data by fitting a zero-inflated model (ZINB), built out of a negative binomial count component (predicting abundance given that birds are present) and a logistical zero component (predicting presence/absence). Due to the data condensation overall variance was lowered, but still few species showed zero-inflated count data. In this case, we strongly recommend using the ZINB model. It was shown that for data subject to an excess in zero-counts, the ZINB model results in better power compared to the NB model.

Statistical power

Modelling the reference data collected in the Thorntonbank study area resulted in empirical ranges of model coefficients. Based upon these we defined numerous scenarios, varying model parameters as well as monitoring set-up characteristics. For each scenario we performed 1000 simulations, allowing us to investigate how the different model parameters affect the power of detecting a change in numbers. Each of these parameters appears to interact with one another, so unambiguous conclusions are difficult to draw. Nevertheless, it could be shown that for the given monitoring set-up (5 years before / 5 year after the impact with a survey effort of 10 km² per month per area), count data subject to zero-inflation and/or characterised by a low theta (<0.4) will hardly be of any value in impact monitoring. Ideally, the data show no zero-inflation ($b_1<-5$), a positive intercept ($a_1>0$), a favourable theta (>0.4) and no significant area effect.

Clearly, after binning the data to day totals, the nature and characteristics of the count results can no longer be changed, but still there are some ways to enhance the power. By far the easiest way to do so is to apply a higher significance threshold (alpha). In this context, a higher alpha increases the chance of wrongly concluding that the turbines are causing an impact, while in fact they are not ('type I error'). However, a stringent significance level goes at the expense of the power, resulting that certain impact effects may go unnoticed (Underwood & Chapman 2003). Most impact studies are meant to function as an early warning system, in order to detect potential negative effects as soon as possible. For decision-making, ecological studies commonly set the probability of a type I error (α) to 0.05, and the probability of a type II error (β) to 0.20. However, this choice tends to be arbitrary and such values imply that the acceptable risk of committing a type II error is four times higher than the risk of a type I error (Pérez-Lapeña *et al*. 2011). In this paper, we use 90% as a boundary for 'sufficient' power (β) and the acceptable risk of making a type I error α was set to 10%, thus equalling acceptable levels for both risks (α=β). Nevertheless, it would still be better for these values to be determined by predefined management objectives (Pérez-Lapeña *et al*. 2011). An approach to set acceptable values for α and β based on costs (in economic, political, environmental and social terms) is proposed by Mapstone (1995).

In a negative binomial distribution the variance function equals $V(\mu) = \mu + \frac{\mu^2}{\theta}$, and so variance is negatively correlated with theta (θ). According to Underwood & Chapman (2003), power is strongly affected by the variability in the measurements. Indeed, we found that power strongly increases with increasing theta. A low theta value depicts over-dispersion, which in this case might arise from year-to-year variation in observed seabird numbers or from strong spatial aggregation of seabirds (e.g. the presence of a fishing vessel inside the study area). It is also closely related to the amount of unexplained data variance, which proves that building a good reference model, i.e. a model explaining as much biologically relevant variation as possible, is of key importance to the final impact assessment results.

Another finding of this study is the importance of selecting a well-considered control area. Ideally, this area hosts highly comparable seabird numbers to the wind farm site, allowing us to perform the impact assessment with more degrees of freedom, reflected by better power.

As was shown, power is strongly enhanced by counting for a longer period of time, due to the increase in sample size (Underwood & Chapman 2003, Pérez-Lapeña *et al*. 2011). One could argue that the timeframe needed to reach a certain power can be halved by performing two monitoring surveys each month. This is in fact true, but surveys still need to be sufficiently spread over time to avoid temporal autocorrelation. Contrastingly, doubling the effort by counting 20 km² per survey per area - instead of 10 km² - does not result in enhanced power, at least not in a direct way. However, it can yield more reliable count results, which in turn can influence the parameter estimates. If let's say, doubling the count effort per survey has a positive effect on the theta value, or lowers the amount of zero-inflation, this will inevitably be reflected in a higher power. It would be very interesting to know how this count effort per survey is linked to the variation/robustness in parameter estimates.

As a last step we calculated powers based on species-specific reference models of twelve seabird species, as observed at the Thorntonbank study area prior to the construction of turbines in 2008. To detect a of 50% decrease in numbers, a power of 90% is reached within 10 years for two seabird species only, i.e. Northern Gannet and Common guillemot. Within the same time frame, power to detect a 75% decrease in numbers exceeds 90% for all species, except for Common Gull. Poorest results are seen in Common Gull and Black-legged Kittiwake, both exhibiting a significant difference in abundance between control and impact area during reference years. All of these results are based on a monitoring set-up in which there is one monthly survey, with an effort of 10 km² in both the control and impact area.

Maclean *et al.* (2006 & 2007) conducted a comparable study on long-time series of aerial survey count data of five seabird species (Red-throated Diver, Common Scoter, Sandwich Tern, Lesser & Great Black-backed Gull), collected in the UK North Sea waters. The (hypothetical) monitoring set-up in that study is quite different from the one presented here. The authors calculated the power of detecting changes within a study area of varying size (2x2 km^2, 5x5 km^2, etc.), with the hypothetical wind farm located in the centre. The study investigates the effect of the gradient of decline (uniform / gradually), spatial scale, survey intensity, survey duration, inclusion of spatial variables and inclusion of reference areas. Maclean *et al*. (2007) concluded that "the statistical power to detect a 50% change in bird numbers remains low (<85%) for all species irrespective of the length of time over which monitoring is carried out", for a significance level of 0.20. The power results presented here are thus clearly higher. We hypothesize that this is largely due to the binning of data, in which day totals instead of single transect counts were used as a base for modelling.

Acknowledgements

First of all we want to thank C-Power NV & Belwind NV for financing this research, and the Management Unit of the North Sea Mathematical Models (MUMM) for assigning it to us. MUMM staff members Robin Brabant & Steven Degraer are thanked for their logistic support and cooperation throughout the monitoring program. During the early stages of the statistical processing, my colleagues Dirk Bauwens and Paul Quataert provided great and crucial advice. We thank all volunteers (especially Walter Wackenier) and colleagues who assisted during the seabird counts, as well as the crew of the research vessel 'Zeeleeuw', and not in the least the Flanders Marine Institute (VLIZ), for their extremely valuable logistic support.

References

Bernstein, B.B. & Zalinski, J. 1983. An optimum sampling design and power tests for environmental biologists. *Journal of Environmental Management* **16**: 35-43.

Certain, G., Bellier, E., Planque, B. & Bretagnolle, V. 2007. Characterising the temporal variability of the spatial distribution of animals: an application to seabirds at sea. *Ecography* **30**: 695-708.

Clarke, E.D., Spear, L.B., McCracken, M.L., Marques, F.F.C., Borchers, D.L., Buckland, S.T., Ainley, D.G. 2003. Validating the use of generalized additive models and at sea surveys to estimate size and temporal trends of seabird populations. *Journal of Applied Ecology* **40**: 278-292.

Degraer, S., Brabant, R. & Rumes, B. (Eds.) 2011. Offshore wind farms in the Belgian part of the North Sea: Selected findings from the baseline and targeted monitoring. Brussels: Management Unit of the North Sea Mathematical Models – Marine Ecosystem Management Unit.

Desholm, M. 2005. Wind farm related mortality among avian migrants – a remote sensing study and model analysis. Ph. D. University of Copenhagen.

Drewitt, A.L. & Langston, R.H.W. 2006. Assessing the impact of wind farms on birds. *Ibis* **148**: 29-42.

Exo, K.-M., Hüppop, O. & Garthe, S. 2003. Birds and offshore wind farms: a hot topic in marine ecology. *Wader Study Group Bull*. **100**: 50-53.

European Wind Energy Association. 2011. The European offshore wind industry - key 2011 trends and statistics. [online] Available at: <http://www.ewea.org/index.php?id=1861> [Accessed 20 June 2012]

Fletcher, D., MacKenzie, D. & Villouta, E. 2005. Modelling skewed data with many zeros: A simple approach combining ordinary and logistic regression. *Environmental and Ecological Statistics* **12**: 45-54.

Fox, A.D., Desholm, M., Kahlert, J., Christensen, T.K. & Petersen, I.K. 2006. Information needs to support environmental impact assessment of the effects of European marine offshore wind farms on birds. *Ibis* **148**: 129-144.

Garthe, S. 1997. Influence of hydrography, fishing activity, and colony location on summer seabird distribution in the south-eastern North Sea. *ICES Journal of Marine Science* **54**: 566-77.

Huettmann, F. & Diamond, A.W. 2006. Large-scale effects on the spatial distribution of seabirds in the Northwest Atlantic. *Landscape Ecology* **21**: 1089-1108.

Jackman, S. 2011. pscl: A package of classes and methods for R developed in the political science computational laboratory. Stanford University.

Karnovsky, N.J., Spear, L.B., Carter, H.R., Ainley, D.G., Amey, K.D., Ballance, L.T., Briggs, K.T., Ford, R.G., Hunt Jr., G.L., Keiper, C., Mason, J.W., Morgan, K.H., Pitman, R.L. & Tynan, C.T. 2006. At-sea distribution, abundance and habitat affinities of Xantus's Murrelets. *Marine Ornithology* **38**: 89-104.

Komdeur, J., Bertelsen, J., Cracknell, G. (eds) 1992. Manual for aeroplane and ship surveys of waterfowl and seabirds. IWRB Spec. Publ. 19. Slimbridge: IWRB.

Langston, R.H.W. & Pullan, J.D. 2003. Windfarms and Birds: An analysis of the effects of windfarms on birds, and guidance on environmental assessment criteria and site selection issues. T-PVS/Inf(2003)12. Report written by Birdlife International on behalf of the Bern Convention.

Larsen, J.K. & Guillemette, M. 2007. Effects of wind turbines on flight behaviour of wintering common eiders : implications for habitat use and collision risk. *Journal of Applied Ecology* **44**: 516-522.

Leopold, M.F., Camphuysen, C.J., ter Braak, C.J.F., Dijkman, E.M., Kersting, K. & van Lieshout, S.M.J. 2004. Baseline studies North Sea Wind Farms: Lot 8 Marine Birds in and around the future sites Nearshore Windfarm (NSW) and Q7. Alterra-report 1048. Wageningen: Alterra.

Leopold, M.F., Dijkman, E.M., Teal, L. & the OWEZ-team (2010). Local Birds in and around the Offshore Wind Farm Egmond aan Zee (T0 & T1, 2002-2010). Report Number C187/11. Wageningen: IMARES.

McCullagh, P. & Nelder, J.A. 1989. Generalized linear models (2nd edition). London: Chapman and Hall.

McDonald, T.L., Erickson, W.P. & McDonald, L.L. 2000. Analysis of Count Data from Before-After Control-Impact studies. *Journal of Agricultural, Biological and Environmental Statistics* **5**: 262-279.

Maclean, I.M.D., Skov, H., Rehfisch, M.M. & Piper, W. 2006. Use of aerial surveys to detect bird displacement by offshore windfarms. BTO Research Report No. 446 to COWRIE. Thetford: British Trust for Ornithology.

Maclean, I.M.D., Skov, H. & Rehfisch, M.M. 2007. Further use of aerial surveys to detect bird displacement by offshore wind farms. BTO Research Report No. 482 to COWRIE. Thetford: British Trust for Ornithology.

Mapstone, B.D. 1995. Scalable decision rules for environmental impact studies: effect size, type I, and type II errors. *Ecological Applications* **5**: 401-410.

MUMM. 2011. Windmolenparken in zee. [online] Available at: <http://www.mumm.ac.be/NL/Management/Sea-based/windmills.php> [Accessed 10 December 2011]

Onkelinx, T., Van Ryckegem, G., Bauwens, D., Quataert, P. & Van den Bergh, E. 2008. Potentie van ruimtelijke modellen als beleidsondersteunend instrument met betrekking tot het voorkomen van watervogels in de Zeeschelde. Report INBO.R.2008.34. Brussels: Research Institute for Nature and Forest.

Oppel, S., Meirinho, A., Ramírez, I., Gardner, B., O'Connel, A.F., Miller, P.I. & Louzao, M. in press. Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. Biological Conservation (2011).

Pebesma, E.J., Bio, A.F. & Duin, R.N.M. 2000. Mapping Sea Bird Densities on the North Sea: combining geostatistics and generalized linear models. In: Kleingeld, W.J. & Krige, D.G. (eds). Geostatistics 2000 Cape Town, Proceedings of the Sixth International Geostatistics Congress held in Cape Town, South Africa, April 2000.

Pérez-Lapeña, B., Wijnberg, K.M., Hulscher, S.J.M.H. & Stein, A. 2010. Environmental impact assessment of offshore wind farms: a simulation-based approach. *Journal of Applied Ecology* **47**: 1110-1118.

Pérez-Lapeña, B., Wijnberg, K.M., Stein, A. & Hulscher, S.J.M.H. 2011. Spatial factors affecting statistical power in testing marine fauna displacement. *Ecological Applications* **21**: 2756–2769.

Pettersson, J. 2005. The impact of offshore wind farms on bird life in Kalmar Sound, Sweden. A final report based on studies 1999-2003. Lund: Lunds Universitet.

Petersen, I.K., Christensen, T.K., Kahlert, J., Desholm, M. & Fox, A.D. 2006. Final results of bird studies at the offshore wind farms at Nysted and Horns Rev, Denmark. Denmark: National Environmental Research Institute.

Potts, J.M. & Elith, J. 2006. Comparing species abundance models. *Ecological Modelling* **199**: 153-163.

R Development Core Team 2011. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Schneider, D.C. 1982. Fronts and Seabird Aggregations in the Southeastern Bering Sea. *Marine Ecology - Progress Series* **10**: 101-103.

Schneider, D.C. & Duffy, D.C. 1985. Scale-dependent variability in seabird abundance. *Marine Ecology - Progress Series* **25**: 211-218.

Schneider, D.C. 1990. Spatial autocorrelation in marine birds. *Polar Research* **8**: 89-97.

Stewart-Oaten, A., Murdoch, W.W. & Parker, K.R. 1986. Environmental impact assessment: “pseudoreplication” in time? *Ecology* **67**: 929-940

Stewart-Oaten, A., Bence, J.R. 2001. Temporal and spatial variation in environmental impact assessment. *Ecological Monographs* **71**: 305-339.

Stienen, E.W.M., Van Waeyenberge, J., Kuijken, E. & Seys, J. 2007. Trapped within the corridor of the Southern North Sea: the potential impact of offshore wind farms on seabirds. In: de Lucas, M., Janss, G.F.E. & Ferrer, M. (eds). Birds and Wind Farms - Risk assessment and Mitigation: 71-80. Madrid: Quercus.

Tasker, M.L., Jones, P.H., Dixon, T.J. & Blake, B.F. 1984. Counting seabirds at sea from ships: a review of methods employed and a suggestion for a standardised approach. *Auk* **101**: 567-577.

Underwood, A.J. 1994. On beyond BACI: sampling designs that might reliably detect environmental disturbances. *Ecological Applications* **4**: 3-15.

Underwood, A.J. & Chapman, M.G. 2003. Power, precaution, Type II error and sampling design in assessment of environmental impacts. *Journal of Experimental Biology and Ecology* **296**: 49-70.

Vanermen, N., Stienen, E.W.M., Courtens, W. & Van de walle, M. 2006. Referentiestudie van de avifauna van de Thorntonbank. Report IN.A.2006.22. Brussels: Research Institute for Nature and Forest.

Vanermen, N. & Stienen, E.W.M. 2009. Seabirds & Offshore Wind Farms: Monitoring results 2008. In: Degraer, S. & Brabant, R. (eds). Offshore wind farms in the Belgian part of the North Sea – State of the art after two years of environmental monitoring: 151-221. Brussels: Management Unit of the North Sea Mathematical Models.

Vanermen, N., Stienen, E.W.M., Onkelinx, T., Courtens, W., Van de walle, M. & Verstraete, H. 2010. Monitoring seabird displacement: a modelling approach. In: Degraer, S., Brabant, R. & Rumes, B. (eds). Offshore wind farms in the Belgian part of the North Sea – Early environmental impact assessment and spatio-temporal variability: 133-152. Brussels: Management Unit of the North Sea Mathematical Models.

Venables, W. N. & Ripley, B. D. 2002. Modern Applied Statistics with S. Fourth Edition. New York: Springer.

Vuong, Q.H. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* **57**: 307-333.

Wenger, S.J. & Freeman, M.C. 2008. Estimating species occurrence, abundance and detection probability using zero-inflated distributions. *Ecology* **89**(10): 2983-2959.

Zeileis, A., Keibler, C. & Jackman, S. 2008. Regression Models for Count Data in R. *Journal of Statistical Software* **27** (8).

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A. & Smith G.M. 2009. Mixed Effects Models and Extensions in Ecology with R. New York: Springer.

APPENDIX

```
> #Selected Impact Models THORNTONBANK

> summary(zeroinfl.NSV3.imp)

Call:
zeroinfl(formula = NSV ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12))
+ cos(2 * pi * (MONTH/12)) + BA + Molens | 1, data = TTB.DD, dist = "negbin")

Pearson residuals:
       Min         1Q     Median         3Q        Max
-0.3302414 -0.3152794 -0.2531322 -0.0003061  6.8209347

Count model coefficients (negbin with log link):
                          Estimate Std. Error z value Pr(>|z|)
(Intercept)               -0.78244    0.71943  -1.088 0.276777
sin(2 * pi * (MONTH/12))  -1.08341    0.39752  -2.725 0.006422 **
cos(2 * pi * (MONTH/12))   0.06835    0.46427   0.147 0.882965
BATRUE                    -3.37950    0.99990  -3.380 0.000725 ***
MolensTRUE               -13.62809  756.28675  -0.018 0.985623
Log(theta)                -1.24553    1.16563  -1.069 0.285273

Zero-inflation model coefficients (binomial with logit link):
            Estimate Std. Error z value Pr(>|z|)
(Intercept)    0.218      1.223   0.178    0.859
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Theta = 0.2878
Number of iterations in BFGS optimization: 17
Log-likelihood: -133.6 on 7 Df
```

```
> summary(NB.JVG3.imp)

Call:
glm.nb(formula = JVG ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + sin(2 * pi * (MONTH/6)) + cos(2 *
    pi * (MONTH/6)) + BA + Molens, data = TTB.DD, link = log,
    init.theta = 0.4066831065)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1.58674  -1.03777  -0.68537  -0.03652   2.62636

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               -0.8003     0.1629  -4.912    9e-07 ***
sin(2 * pi * (MONTH/12))  -0.5391     0.1764  -3.056  0.00225 **
cos(2 * pi * (MONTH/12))   0.3966     0.2035   1.949  0.05131 .
sin(2 * pi * (MONTH/6))   -0.4087     0.1916  -2.133  0.03296 *
cos(2 * pi * (MONTH/6))   -0.3619     0.1843  -1.964  0.04954 *
BATRUE                    -0.4280     0.3455  -1.239  0.21540
MolensTRUE                -0.7134     0.4673  -1.527  0.12686
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(0.4067) family taken to be 1)

    Null deviance: 212.28  on 193  degrees of freedom
Residual deviance: 173.08  on 187  degrees of freedom
AIC: 696.26

Number of Fisher Scoring iterations: 1


              Theta:  0.4067
          Std. Err.:  0.0625

 2 x log-likelihood:  -680.2550
```

```
> summary(NB.DWM3.imp)

Call:
glm.nb(formula = DWM ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + sin(2 * pi * (MONTH/6)) + cos(2 *
    pi * (MONTH/6)) + BA + Molens, data = TTB.DD, link = log,
    init.theta = 0.2157172552)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1.26173  -0.92469  -0.31892  -0.03319   1.89072

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               -3.3321     0.5729  -5.816 6.02e-09 ***
sin(2 * pi * (MONTH/12))   2.5318     0.5021   5.042 4.60e-07 ***
cos(2 * pi * (MONTH/12))   3.3727     0.7932   4.252 2.12e-05 ***
sin(2 * pi * (MONTH/6))   -1.9536     0.4950  -3.947 7.93e-05 ***
cos(2 * pi * (MONTH/6))   -0.9429     0.4165  -2.264   0.0236 *
BATRUE                    -0.2862     0.5704  -0.502   0.6159
MolensTRUE                 1.2222     0.7066   1.730   0.0837 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(0.2157) family taken to be 1)

    Null deviance: 185.19  on 193  degrees of freedom
Residual deviance: 103.55  on 187  degrees of freedom
AIC: 469.3

Number of Fisher Scoring iterations: 1


              Theta:  0.2157
          Std. Err.:  0.0402
Warning while fitting theta: alternation limit reached

 2 x log-likelihood:  -453.2970
```

```
> summary(NB.STM4.imp)

Call:
glm.nb(formula = STM ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + BA + CI + BA:CI, data = TTB.DD,
    link = log, init.theta = 0.2495300467)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.4291  -0.7511  -0.2862  -0.1097   2.3953

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               -4.3860     0.5373  -8.163 3.27e-16 ***
sin(2 * pi * (MONTH/12))   2.0756     0.3540   5.864 4.53e-09 ***
cos(2 * pi * (MONTH/12))   3.1376     0.4648   6.750 1.47e-11 ***
BATRUE                     1.6879     0.6345   2.660  0.00781 **
CITRUE                     1.3493     0.5279   2.556  0.01060 *
BATRUE:CITRUE             -1.4287     0.8712  -1.640  0.10105
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(0.2495) family taken to be 1)

    Null deviance: 204.113  on 193  degrees of freedom
Residual deviance:  91.194  on 188  degrees of freedom
AIC: 405.58

Number of Fisher Scoring iterations: 1


              Theta:  0.2495
          Std. Err.:  0.0509

 2 x log-likelihood:  -391.5780
```

```
> summary(NB.KLM3.imp)

Call:
glm.nb(formula = KLM ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + BA + Molens, data = TTB.DD, link = "log",
    init.theta = 0.2583893342)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.6534  -1.0147  -0.7314  -0.2596   2.9587

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               0.08589    0.18860   0.455    0.649
sin(2 * pi * (MONTH/12))  0.93281    0.20551   4.539 5.65e-06 ***
cos(2 * pi * (MONTH/12)) -2.15226    0.23464  -9.173  < 2e-16 ***
BATRUE                   -0.36595    0.41307  -0.886    0.376
MolensTRUE               -0.12718    0.52476  -0.242    0.809
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(0.2584) family taken to be 1)

    Null deviance: 253.94  on 193  degrees of freedom
Residual deviance: 169.41  on 189  degrees of freedom
AIC: 877

Number of Fisher Scoring iterations: 1


              Theta:  0.2584
          Std. Err.:  0.0341

 2 x log-likelihood:  -865.0030
```

```
> summary(NB.ZM3.imp)

Call:
glm.nb(formula = ZM ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + BA + Molens, data = TTB.DD, link = log,
    init.theta = 0.2709723123)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.3196  -0.7289  -0.4303  -0.2661   2.5304

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               -2.6537     0.2633 -10.080  < 2e-16 ***
sin(2 * pi * (MONTH/12))   1.6188     0.2832   5.715 1.09e-08 ***
cos(2 * pi * (MONTH/12))   0.6439     0.2966   2.171   0.0299 *
BATRUE                    -0.2330     0.5187  -0.449   0.6533
MolensTRUE                 0.3742     0.6514   0.574   0.5657
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(0.271) family taken to be 1)

    Null deviance: 142.54  on 193  degrees of freedom
Residual deviance: 103.69  on 189  degrees of freedom
AIC: 348.89

Number of Fisher Scoring iterations: 1


              Theta:  0.2710
          Std. Err.:  0.0623

 2 x log-likelihood:  -336.8870
```

```
> summary(NB.GM3.imp)

Call:
glm.nb(formula = GM ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + BA + Molens, data = TTB.DD, link = log,
    init.theta = 0.2215195383)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.3990  -0.8869  -0.5718  -0.3066   2.5332

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               -1.4042     0.2258  -6.219 5.00e-10 ***
sin(2 * pi * (MONTH/12))  -0.4443     0.2400  -1.851  0.06412 .
cos(2 * pi * (MONTH/12))   1.9657     0.2901   6.777 1.23e-11 ***
BATRUE                    -1.5799     0.5343  -2.957  0.00311 **
MolensTRUE                 1.4853     0.6533   2.274  0.02298 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(0.2215) family taken to be 1)

    Null deviance: 197.29  on 193  degrees of freedom
Residual deviance: 130.05  on 189  degrees of freedom
AIC: 547.24

Number of Fisher Scoring iterations: 1


              Theta:  0.2215
          Std. Err.:  0.0377

 2 x log-likelihood:  -535.2360
```

```
> summary(NB.DTM4.imp)

Call:
glm.nb(formula = DTM ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + BA + CI + BA:CI, data = TTB.DD,
    link = log, init.theta = 0.2521873069)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5748  -0.8633  -0.5863  -0.3841   4.4846

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)              -0.09948    0.27714  -0.359 0.719630
sin(2 * pi * (MONTH/12)) -0.71786    0.22055  -3.255 0.001135 **
cos(2 * pi * (MONTH/12))  2.75634    0.27510  10.019  < 2e-16 ***
BATRUE                   -1.76164    0.50011  -3.523 0.000427 ***
CITRUE                   -1.44933    0.39590  -3.661 0.000251 ***
BATRUE:CITRUE             2.00641    0.71184   2.819 0.004823 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(0.2522) family taken to be 1)

    Null deviance: 235.41  on 193  degrees of freedom
Residual deviance: 149.88  on 188  degrees of freedom
AIC: 748.6

Number of Fisher Scoring iterations: 1


              Theta:  0.2522
          Std. Err.:  0.0354
Warning while fitting theta: alternation limit reached

 2 x log-likelihood:  -734.5970
```

```
> summary(zeroinfl.GS3.imp)

Call:
zeroinfl(formula = GS ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) + cos(2 * pi * (MONTH/12))
+ sin(2 * pi * (MONTH/6)) +
    cos(2 * pi * (MONTH/6)) + BA + Molens | 1, data = TTB.DD, dist = "negbin")

Pearson residuals:
       Min         1Q     Median         3Q        Max
-0.5985091 -0.3565730 -0.0250004 -0.0001934  3.9707844

Count model coefficients (negbin with log link):
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)              -7.21613    1.74128  -4.144 3.41e-05 ***
sin(2 * pi * (MONTH/12))  0.48406    0.53033   0.913 0.361377
cos(2 * pi * (MONTH/12)) -9.18166    2.32097  -3.956 7.62e-05 ***
sin(2 * pi * (MONTH/6))   1.15188    0.55164   2.088 0.036790 *
cos(2 * pi * (MONTH/6))  -4.05951    1.23647  -3.283 0.001027 **
BATRUE                   -1.20911    0.64357  -1.879 0.060278 .
MolensTRUE                2.43086    0.70371   3.454 0.000552 ***
Log(theta)               -0.01584    0.72927  -0.022 0.982673

Zero-inflation model coefficients (binomial with logit link):
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -0.1741     0.6098  -0.285    0.775
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Theta = 0.9843
Number of iterations in BFGS optimization: 20
Log-likelihood: -106.9 on 9 Df
```

```
> summary(zeroinfl.VD3.imp)

Call:
zeroinfl(formula = VD ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) + cos(2 * pi * (MONTH/12))
+ sin(2 * pi * (MONTH/6)) +
    cos(2 * pi * (MONTH/6)) + BA + Molens | 1, data = TTB.DD, dist = "negbin")

Pearson residuals:
       Min         1Q     Median         3Q        Max
-5.280e-01 -2.808e-01 -5.227e-03 -2.756e-05  3.863e+00

Count model coefficients (negbin with log link):
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               -8.9539     2.9633  -3.022  0.00251 **
sin(2 * pi * (MONTH/12))  -1.0098     1.1384  -0.887  0.37505
cos(2 * pi * (MONTH/12)) -11.7248     4.0681  -2.882  0.00395 **
sin(2 * pi * (MONTH/6))   -0.4623     1.1455  -0.404  0.68655
cos(2 * pi * (MONTH/6))   -5.2275     1.8071  -2.893  0.00382 **
BATRUE                    -2.2586     0.9473  -2.384  0.01712 *
MolensTRUE                 2.4153     1.1017   2.192  0.02835 *
Log(theta)                 0.8496     1.4034   0.605  0.54493

Zero-inflation model coefficients (binomial with logit link):
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   0.7171     0.5879    1.22    0.223
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Theta = 2.3386
Number of iterations in BFGS optimization: 20
Log-likelihood: -60.11 on 9 Df
```

```
> summary(NB.ZK3.imp)

Call:
glm.nb(formula = ZK ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + BA + Molens, data = TTB.DD, link = "log",
    init.theta = 0.7706574986)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.2800  -0.8894  -0.3907  -0.1102   2.6088

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               -1.2452     0.1760  -7.077 1.48e-12 ***
sin(2 * pi * (MONTH/12))   0.5645     0.1532   3.684  0.00023 ***
cos(2 * pi * (MONTH/12))   3.5207     0.2496  14.104  < 2e-16 ***
BATRUE                    -1.4954     0.3274  -4.567 4.94e-06 ***
MolensTRUE                -0.1666     0.4472  -0.373  0.70951
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(0.7707) family taken to be 1)

    Null deviance: 464.45  on 193  degrees of freedom
Residual deviance: 151.50  on 189  degrees of freedom
AIC: 761.06

Number of Fisher Scoring iterations: 1


              Theta:  0.771
          Std. Err.:  0.121

 2 x log-likelihood:  -749.060
```

```
> summary(NB.ALK3.imp)

Call:
glm.nb(formula = ALK ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + BA + Molens, data = TTB.DD, link = "log",
    init.theta = 0.3388846558)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.6284  -0.7568  -0.3622  -0.1317   3.4090

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               -2.6015     0.2872  -9.059   <2e-16 ***
sin(2 * pi * (MONTH/12))   0.1614     0.2319   0.696    0.486
cos(2 * pi * (MONTH/12))   3.6379     0.3870   9.400   <2e-16 ***
BATRUE                    -0.4514     0.4727  -0.955    0.340
MolensTRUE                 0.4285     0.6073   0.706    0.480
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(0.3389) family taken to be 1)

    Null deviance: 217.36  on 193  degrees of freedom
Residual deviance: 115.62  on 189  degrees of freedom
AIC: 511.96

Number of Fisher Scoring iterations: 1


              Theta:  0.3389
          Std. Err.:  0.0596

 2 x log-likelihood:  -499.9550
```

```
> #SUMMARIES Impact models BLIGHBANK

> summary(NB.NSV3.imp)

Call:
glm.nb(formula = NSV ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + BA + Molens, data = BB_DD, link = "log",
    init.theta = 0.148194239)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1.09999  -0.77895  -0.55848  -0.01942   1.81198

Coefficients:
                          Estimate Std. Error z value Pr(>|z|)
(Intercept)              -1.716e+00  2.524e-01  -6.800 1.05e-11 ***
sin(2 * pi * (MONTH/12))  9.090e-01  3.215e-01   2.827  0.00470 **
cos(2 * pi * (MONTH/12))  9.082e-01  3.694e-01   2.459  0.01395 *
BATRUE                   -2.867e+00  9.897e-01  -2.896  0.00377 **
MolensTRUE               -2.860e+01  7.864e+05   0.000  0.99997
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(0.1482) family taken to be 1)

    Null deviance: 119.220  on 177  degrees of freedom
Residual deviance:  85.433  on 173  degrees of freedom
AIC: 342.8

Number of Fisher Scoring iterations: 1


              Theta:  0.1482
          Std. Err.:  0.0330

 2 x log-likelihood:  -330.7980
```

```
> summary(NB.JVG3.imp)

Call:
glm.nb(formula = JVG ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + sin(2 * pi * (MONTH/6)) + cos(2 *
    pi * (MONTH/6)) + BA + Molens, data = BB_DD, link = log,
    init.theta = 0.4125317842)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.6598  -1.0153  -0.5723  -0.0017   3.3087

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)              -1.49789    0.18444  -8.121 4.62e-16 ***
sin(2 * pi * (MONTH/12)) -0.14119    0.19082  -0.740 0.459347
cos(2 * pi * (MONTH/12))  1.60591    0.25527   6.291 3.16e-10 ***
sin(2 * pi * (MONTH/6))   0.06762    0.20486   0.330 0.741319
cos(2 * pi * (MONTH/6))  -0.82039    0.22841  -3.592 0.000328 ***
BATRUE                    0.95934    0.44698   2.146 0.031852 *
MolensTRUE               -1.50284    0.62651  -2.399 0.016452 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(0.4125) family taken to be 1)

    Null deviance: 203.64  on 177  degrees of freedom
Residual deviance: 149.61  on 171  degrees of freedom
AIC: 609.85

Number of Fisher Scoring iterations: 1


              Theta:  0.4125
          Std. Err.:  0.0696

 2 x log-likelihood:  -593.8540
```

```
> summary(zeroinfl.GJ0.imp)

Call:
zeroinfl(formula = GJ ~ offset(log(KM2)) + BA + Molens | 1, data = BB_DD, dist = "negbin")

Pearson residuals:
    Min      1Q  Median      3Q     Max 
-0.4310 -0.3428 -0.2756 -0.1857  4.1149 

Count model coefficients (negbin with log link):
             Estimate Std. Error z value Pr(>|z|)    
(Intercept)   -1.8842     0.4977  -3.786 0.000153 ***
BATRUE        -2.0034     1.1410  -1.756 0.079121 .  
MolensTRUE   -14.8584  2323.9275  -0.006 0.994899    
Log(theta)     1.5998     3.1537   0.507 0.611959    

Zero-inflation model coefficients (binomial with logit link):
            Estimate Std. Error z value Pr(>|z|)  
(Intercept)   1.0819     0.6284   1.722   0.0852 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Theta = 4.9522 
Number of iterations in BFGS optimization: 18 
Log-likelihood: -71.97 on 5 Df
```

```
> summary(NB.DWM4.imp)

Call:
glm.nb(formula = DWM ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + BA + CI + BA:CI, data = BB_DD,
    link = log, init.theta = 0.1009550181)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1.17461  -0.46538  -0.13820  -0.01035   1.82505

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               -8.1001     1.3445  -6.024 1.70e-09 ***
sin(2 * pi * (MONTH/12))   5.1082     1.2426   4.111 3.94e-05 ***
cos(2 * pi * (MONTH/12))   1.1776     0.7526   1.565 0.117651
BATRUE                     5.0209     1.4153   3.548 0.000389 ***
CITRUE                     1.5092     0.9531   1.584 0.113293
BATRUE:CITRUE             -0.7942     1.7146  -0.463 0.643206
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(0.101) family taken to be 1)

    Null deviance: 82.051  on 177  degrees of freedom
Residual deviance: 33.883  on 172  degrees of freedom
AIC: 164.15

Number of Fisher Scoring iterations: 1


              Theta:  0.1010
          Std. Err.:  0.0356
Warning while fitting theta: alternation limit reached

 2 x log-likelihood:  -150.1490
```

```
> summary(zeroinfl.STM4.imp)

Call:
zeroinfl(formula = STM ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) + cos(2 * pi *
(MONTH/12)) + BA + CI + BA:CI |
    1, data = BB_DD, dist = "negbin")

Pearson residuals:
     Min       1Q   Median       3Q      Max
-0.36015 -0.31621 -0.16539 -0.03687  5.41832

Count model coefficients (negbin with log link):
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               -4.4229     0.8609  -5.137 2.79e-07 ***
sin(2 * pi * (MONTH/12))   1.8159     0.5689   3.192  0.00141 **
cos(2 * pi * (MONTH/12))   3.3250     0.8290   4.011 6.05e-05 ***
BATRUE                     1.1057     0.9542   1.159  0.24652
CITRUE                     0.2306     0.7885   0.292  0.76998
BATRUE:CITRUE              3.0404     1.3631   2.230  0.02572 *
Log(theta)                -1.1637     0.8718  -1.335  0.18193

Zero-inflation model coefficients (binomial with logit link):
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.06849    0.94534   0.072    0.942
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Theta = 0.3123
Number of iterations in BFGS optimization: 70
Log-likelihood: -102.7 on 8 Df
```

```
> summary(NB.KLM3.imp)

Call:
glm.nb(formula = KLM ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + BA + Molens, data = BB_DD, link = log,
    init.theta = 0.1889525948)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.2458  -0.9206  -0.7296  -0.2216   3.2427

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               -1.0918     0.2101  -5.197 2.03e-07 ***
sin(2 * pi * (MONTH/12))   0.6623     0.2557   2.590   0.0096 **
cos(2 * pi * (MONTH/12))  -0.7035     0.2858  -2.461   0.0138 *
BATRUE                    -0.1421     0.6441  -0.221   0.8254
MolensTRUE                 0.1406     0.8634   0.163   0.8706
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(0.189) family taken to be 1)

    Null deviance: 136.59  on 177  degrees of freedom
Residual deviance: 124.92  on 173  degrees of freedom
AIC: 534.82

Number of Fisher Scoring iterations: 1


              Theta:  0.1890
          Std. Err.:  0.0341

 2 x log-likelihood:  -522.8190
```

```
> summary(zeroinfl.ZM3.imp)

Call:
zeroinfl(formula = ZM ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) + cos(2 * pi * (MONTH/12))
+ BA + Molens | 1,
    data = BB_DD, dist = "negbin")

Pearson residuals:
     Min       1Q   Median       3Q      Max
-0.37357 -0.31579 -0.17313 -0.04402 30.04299

Count model coefficients (negbin with log link):
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               -4.8567     0.7612  -6.380 1.77e-10 ***
sin(2 * pi * (MONTH/12))   2.3782     0.8081   2.943  0.00325 **
cos(2 * pi * (MONTH/12))   2.3220     0.7117   3.262  0.00110 **
BATRUE                     1.8931     1.0856   1.744  0.08120 .
MolensTRUE                 1.3441     1.0691   1.257  0.20868
Log(theta)                -1.9645     0.3154  -6.228 4.71e-10 ***

Zero-inflation model coefficients (binomial with logit link):
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   -7.633    134.661  -0.057    0.955
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Theta = 0.1402
Number of iterations in BFGS optimization: 67
Log-likelihood: -107.7 on 7 Df
```

```
> summary(NB.GM4.imp)

Call:
glm.nb(formula = GM ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + BA + CI + BA:CI, data = BB_DD,
    link = log, init.theta = 0.1281982346)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.2549  -0.8260  -0.5394  -0.2798   2.7581

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)              -2.78390    0.43709  -6.369 1.90e-10 ***
sin(2 * pi * (MONTH/12))  1.56109    0.33991   4.593 4.38e-06 ***
cos(2 * pi * (MONTH/12))  2.01988    0.38578   5.236 1.64e-07 ***
BATRUE                   -0.08526    0.90114  -0.095    0.925
CITRUE                    2.17483    0.54569   3.985 6.73e-05 ***
BATRUE:CITRUE            -0.54725    1.21866  -0.449    0.653
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(0.1282) family taken to be 1)

    Null deviance: 151.878  on 177  degrees of freedom
Residual deviance:  99.226  on 172  degrees of freedom
AIC: 450.91

Number of Fisher Scoring iterations: 1


              Theta:  0.1282
          Std. Err.:  0.0234

 2 x log-likelihood:  -436.9140
```

```
> summary(NB.DTM3.imp)

Call:
glm.nb(formula = DTM ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + BA + Molens, data = BB_DD, link = log,
    init.theta = 0.3001119269)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5878  -0.8940  -0.5717  -0.3286   4.3551

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               -1.1839     0.1980  -5.978 2.26e-09 ***
sin(2 * pi * (MONTH/12))   0.3094     0.2176   1.422    0.155
cos(2 * pi * (MONTH/12))   2.7880     0.2807   9.933  < 2e-16 ***
BATRUE                    -0.7920     0.5502  -1.439    0.150
MolensTRUE                 0.5601     0.7311   0.766    0.444
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(0.3001) family taken to be 1)

    Null deviance: 221.85  on 177  degrees of freedom
Residual deviance: 140.13  on 173  degrees of freedom
AIC: 655.32

Number of Fisher Scoring iterations: 1


              Theta:  0.3001
          Std. Err.:  0.0450

 2 x log-likelihood:  -643.3200
```

```
> summary(NB.ZK3.imp)

Call:
glm.nb(formula = ZK ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + BA + Molens, data = BB_DD, link = "log",
    init.theta = 0.6771545637)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.0940  -0.7830  -0.4222  -0.1829   2.6045

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               -1.6287     0.1853  -8.788  < 2e-16 ***
sin(2 * pi * (MONTH/12))   1.1346     0.1811   6.264 3.75e-10 ***
cos(2 * pi * (MONTH/12))   2.8436     0.2490  11.422  < 2e-16 ***
BATRUE                    -0.2857     0.3950  -0.723   0.4694
MolensTRUE                -1.1458     0.5732  -1.999   0.0456 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(0.6772) family taken to be 1)

    Null deviance: 357.23  on 177  degrees of freedom
Residual deviance: 142.73  on 173  degrees of freedom
AIC: 645.87

Number of Fisher Scoring iterations: 1


              Theta:  0.677
          Std. Err.:  0.116

 2 x log-likelihood:  -633.866
```

```
> summary(NB.ALK3.imp)

Call:
glm.nb(formula = ALK ~ offset(log(KM2)) + sin(2 * pi * (MONTH/12)) +
    cos(2 * pi * (MONTH/12)) + BA + Molens, data = BB_DD, link = "log",
    init.theta = 0.2957021084)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1.67192  -0.66779  -0.23553  -0.07611   2.20746

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               -3.9717     0.4499  -8.828  < 2e-16 ***
sin(2 * pi * (MONTH/12))   1.6747     0.3742   4.475 7.64e-06 ***
cos(2 * pi * (MONTH/12))   3.3734     0.5258   6.415 1.41e-10 ***
BATRUE                     1.5165     0.6221   2.438   0.0148 *
MolensTRUE                -1.2944     0.8481  -1.526   0.1269
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Negative Binomial(0.2957) family taken to be 1)

    Null deviance: 170.882  on 177  degrees of freedom
Residual deviance:  81.774  on 173  degrees of freedom
AIC: 334.79

Number of Fisher Scoring iterations: 1


              Theta:  0.2957
          Std. Err.:  0.0692

 2 x log-likelihood:  -322.7930
```