EIA STUDY

Horns Rev offshore windfarm

SUMMARY

Effect of Horns Rev offshore windfarm of marine bird communities.

Author

150022854

1 ABSTRACT

This report summarizes the statistical procedures used and provides the findings, in line with the Environmental Impact Assessment (EIA) for the Horns Rev offshore windfarm off the coast of Denmark. In particular, this study is concerned with two questions:

1. Has the construction of the windfarm impacted on the numbers of birds present in the area?
2. Did the spatial distribution of birds changed after construction? *I.e.* have birds moved from their pre-construction habitual areas to different spots?

In order to achieve such objective, the animal abundance/presence has been modelled using statistical techniques, whose details can be found in subsequent sections. After such an analysis, we found that:

- The overall bird presence in the area did not change after the impact. In other words we did not found any relationship between the number of animals per unit area and the construction of the turbines.
- They appear to have dispersed in the map after the construction. As can be seen in Figure 1, before the impact they were concentrated in the south-east corner with some presence at the centre. After impact, their presence at the centre increased, with a big agglomeration in the south and some grouping at the northern boundary. There is evidence of turbines avoidance.

2 INTRODUCTION

The aim of this study, commissioned by DONG Energy, is to quantify the environmental impact of the Hons Rev offshore windfarm. The area of interest seems to be of significant importance for many birds species, which inhabit these waters. Nevertheless, the construction of the turbines could have potentially impacted on the birds, changing their feeding and breeding habits; in short, impacting on their presence in the area. To discover the potential changes, two questions needed answer: whether the overall bird abundance changed and whether their spatial distribution was modified.

Among the dataset provided, six variables[1] showed some interesting relationship with the response. Their scatterplots are provided in Figure 2 and Figure 3. Birds seem dispersed almost homogenously in the space, even though some concentrations at small to medium levels of the X coordinate and at Y coordinate extrema are present. In addition, they seem to prefer medium depths, while no clear pattern is visible for neither day nor month. One could argue for higher bird presence after impact, but from those scatterplots it's difficult to say, due to the high number of zero counts. Nevertheless, we see that for some covariates, the relationship with counts is highly non-linear. Moreover, the data has been collected by aerial survey through a transect scheme: for this reason we suspected high correlation between observations, grouped by transect. The runs test on a preliminary model provided a p-value less than 0.05 with a statistic of -130 signifying important positive correlation. We can see this also from Figure 4 where some blocks have high correlation (0.4) at lag 20.

In order to address the first feature (non-linear relationships) we decided to apply a Poisson-based GAM with regression splines. In particular, to better model the spatial component, we relied on a CReSS surface (Scott-Hayward, Mackenzie, Donovan , Walker, & Ashe, 2014). To address the second feature we instead relied on a GEE in order to model the correlation within transect. The blocking structure used was transect

[1] See Table 2 for a list and description of covariates. Note that *impact* will be subsequently removed from the model.

N°, a unique identifier for each transect. This choice is justified on Figure 4: the correlation declines hence the structure is suitable.

All the analysis has been conducted using the R software.

3 MODELLING METHODS[2]

Our initial model was structured as a Poisson-based GAM with log link and overdispersed variance. The Poisson distribution was dictated by the nature of the response, which were counts. At the same time, the overdispertion reasonably represent the data, as seen in Figure 5: the variance assumed under the model works for fitted values smaller than 200. Among the covariates, *impact* and *month* were included as factors, while *depth* and *day* as smooth terms. Those lasts were entered using B-spline basis functions, while for the spatial component (*x.pos* and *y.pos*) we relied upon a local radial Gaussian fitted using the CReSS method. In order to select the best knot number and positions we used automated model selection via SALSA (Walker, Mackenzie, Donovan, & O'Sullivan, 2010). In addition we allowed for an interaction between this spatial component and the construction phase (*i.e.* before/after the impact). Finally, we also included an offset term based on the area of each element in the observation grid. However, huge collinearity was present due to the interaction effect: for this reason we scaled the numeric component (the local radial basis). Table 1 shows the improvements. Nevertheless, as said, given that high correlation between residuals was present (see Figure 9 as example. The other orderings produced similar results), we fitted all this using a GEE to estimate the empirical standard errors. The blocking structure used was based on *TNO* (transect N°).

The covariate relationships are visible from Figure 6 to Figure 8 which present the partial plots. Partial residuals are omitted (they would squeeze the lines). The *day* covariate presents quite large CI: hence we cannot exclude a simple linear relationship. The large boundaries are probably due to the few values available for such covariate. Moreover, *impact* was found to be not significant based on its marginal ANOVA p-value of 0.202 but it could not be dropped due to its interaction effect. We also controlled for influential points with Cook's Distance (Figure 13): some large values are present between observations 15000 to 20000.

Finally, the fit of this model can be visualized in Figure 10. It presents both severe under predictions and over predictions, hence it has a very poor fit.

4 RESULTS AND LIMITATIONS

With this model, we were able to answer the two research questions. In particular, there is no evidence that the turbines impacted on birds presence. In fact, the marginal ANOVA p-value for the *impact* covariate was found to be lower than 0.05. At the same time, as we can see from the quilt plots in Figure 1, their spatial distribution changed and they seem to have moved away from the windfarm site. At the same time, they seem to prefer winter months and medium depths.

However, this model has some significant problems in prediction (Figure 10). We can better see this issue in Figure 11: under prediction in the first part and sever over prediction in the last observations. This bad fit could be solved by including more relevant covariates in the model. In addition, even though automatic knot allocation was used (with CV and different starting positions to avoid local minima), the relationship with *depth* seems to be bad modelled. We see this from Figure 12, where the disparity between the

[2] The relevant code for the study can be requested at af212@st-andrews.ac.uk.

relationship assumed (black line) and what it should be (grey line) is significant. Possibly, more rounds of CV could solve this issue.

5 APPENDIX

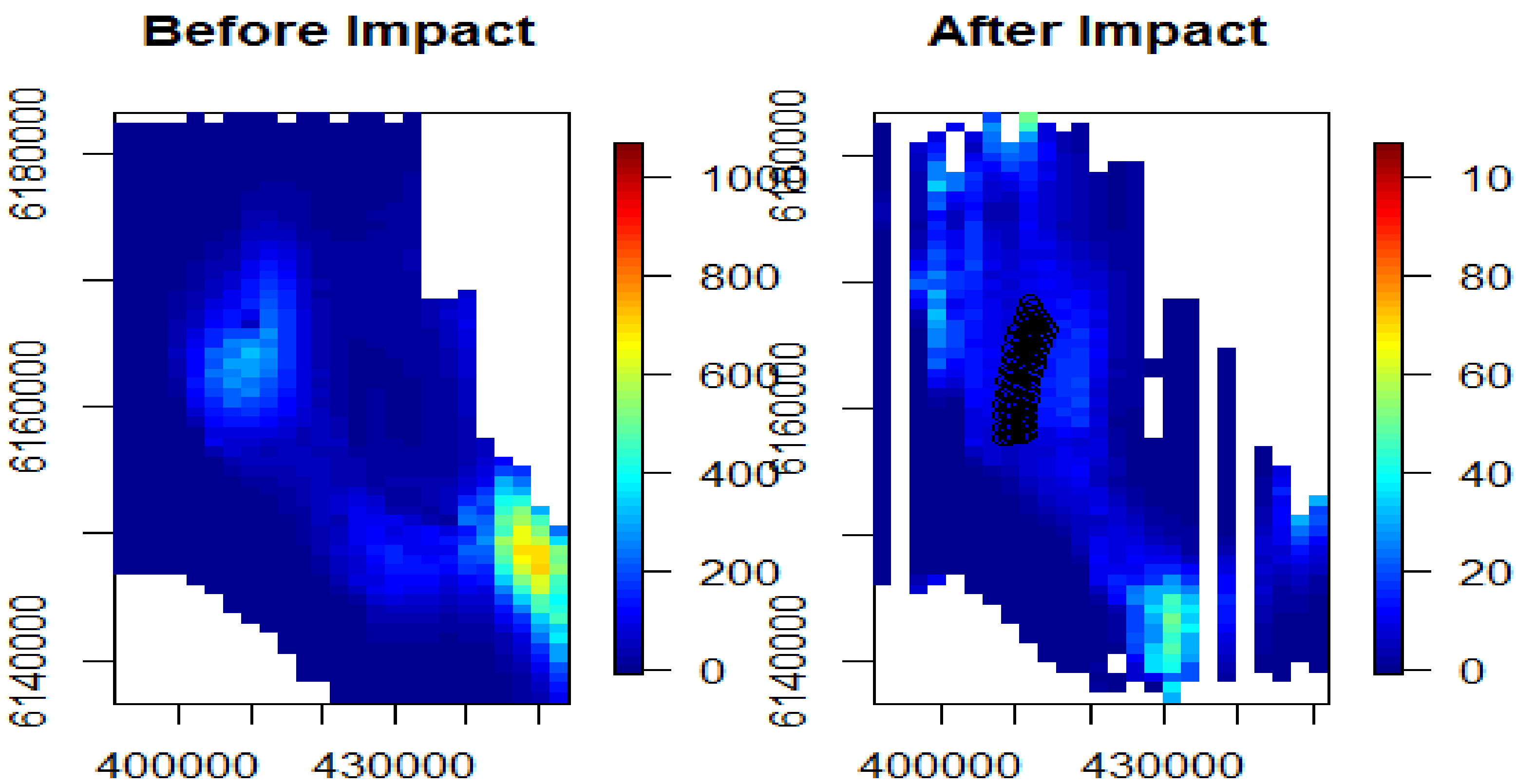

Figure 1: The area of the study with birds presence: before (left) and after construction (right). The black point in the right quilt plot represent the position of the turbines.

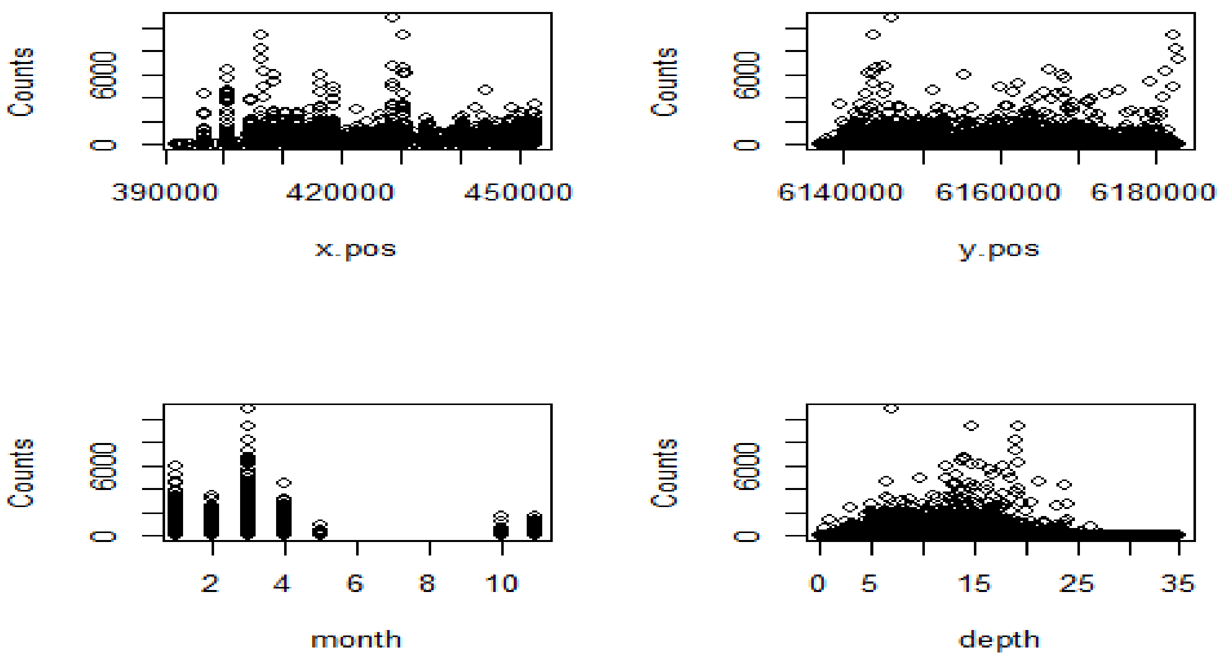

Figure 2: Scatterplots of bird Counts against various covariates

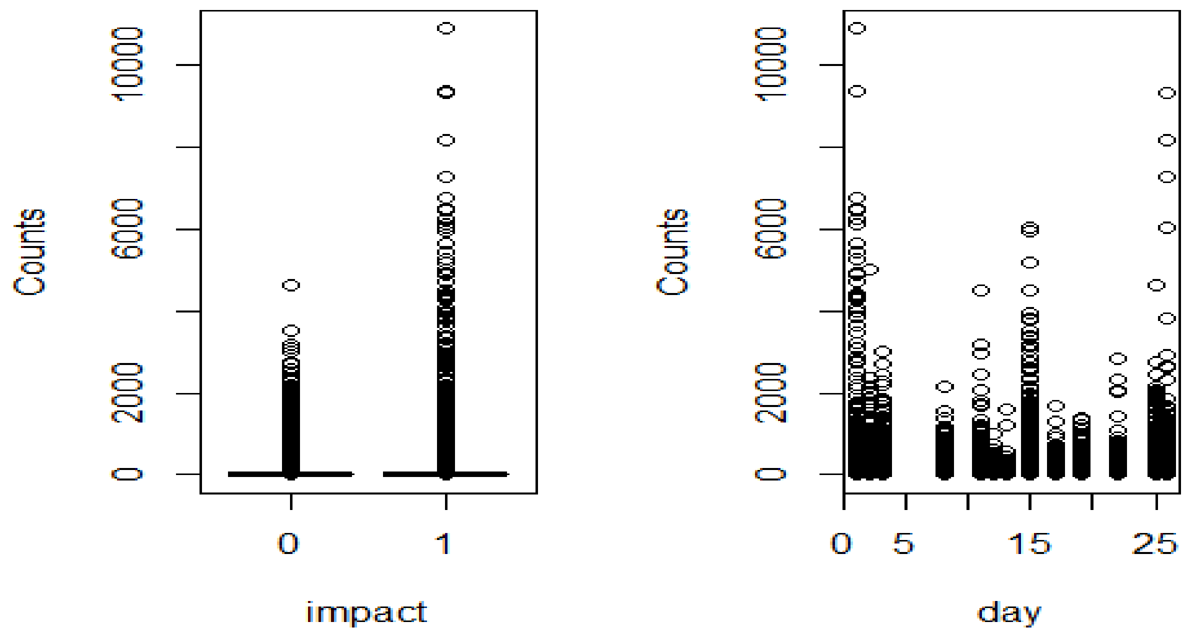

Figure 3: Scatterplots of bird Counts against various covariates

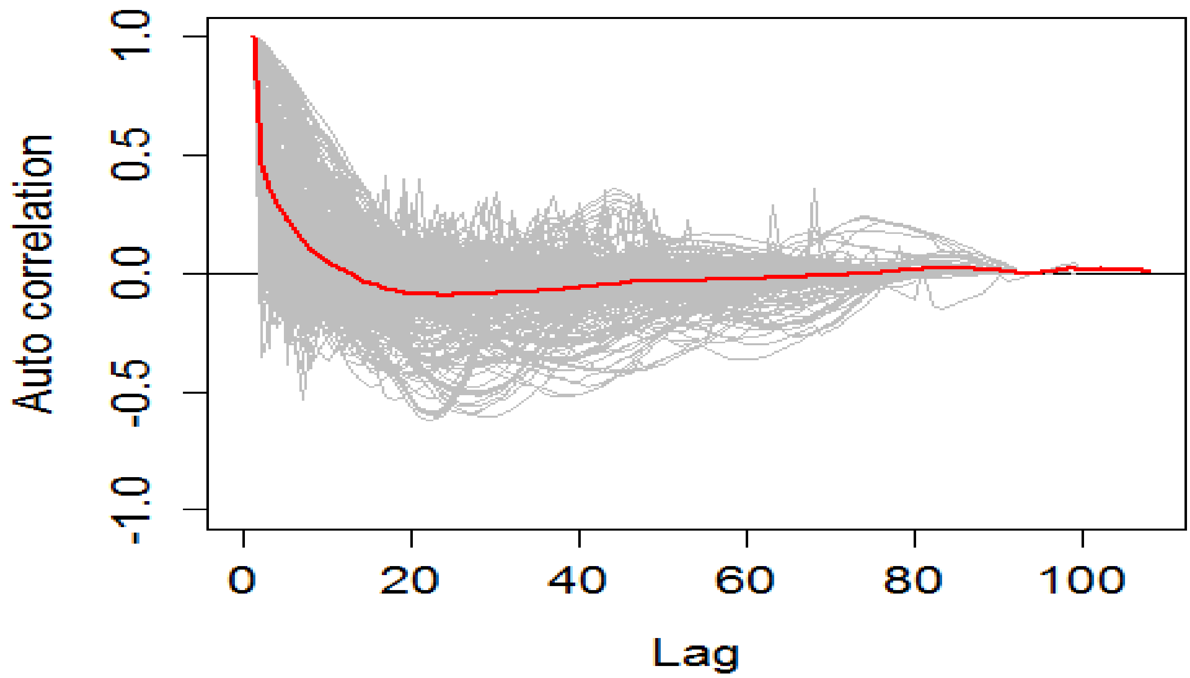

Figure 4: Plot of the correlations of residuals for each block (grey lines; blocking structure based on transect number). The red lines indicates the mean correlation.

Covariate	Adjusted GVIF (before scaling)	Adjusted GVIF (after scaling)
impact	24.723	3.192
month	1.090	1.090
s(Depth)	1.472	1.472
s(day)	1.143	1.143
s(x.pos, y.pos)	1.748	1.748
s(x.pos, y.pos):impact	1.941	1.751

Table 1: Comparison of adjusted GVIF before and after scaling the spatial covariate s(x.pos, y.pos). Even though 3.192 is probably still high, the improvement is remarkable. The s() function refers to smooth terms.

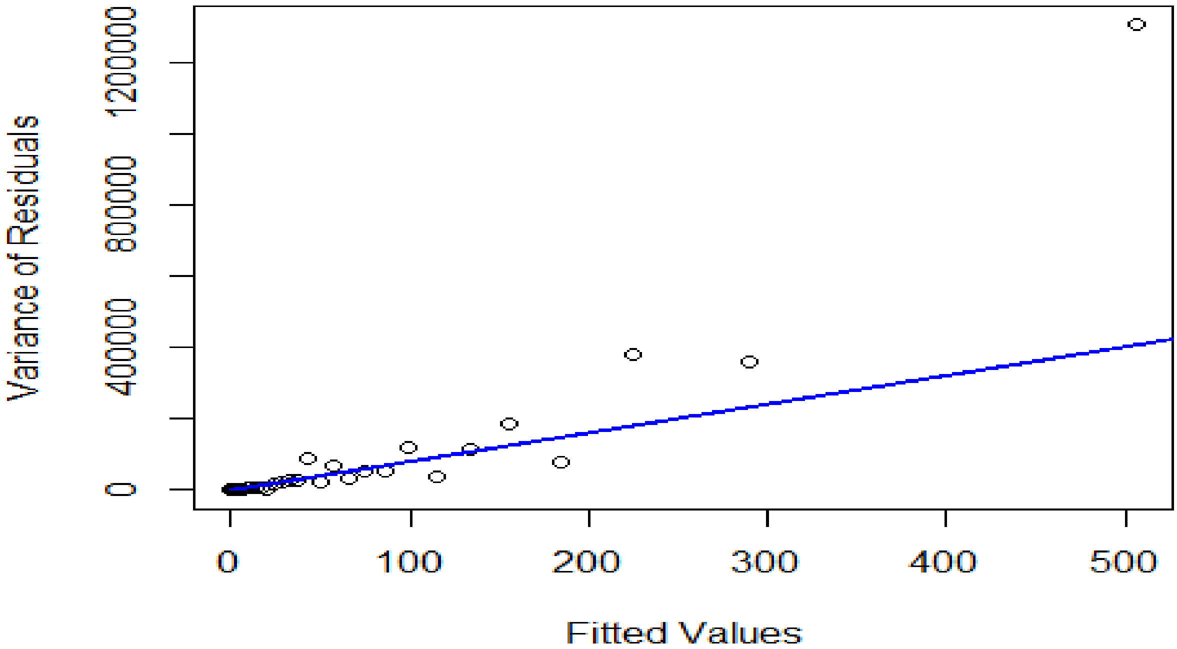

Figure 5: Mean-Variance relationship. The blue line represent the overdispersed variance assumed under the model, while the points are the variance of residual bins.

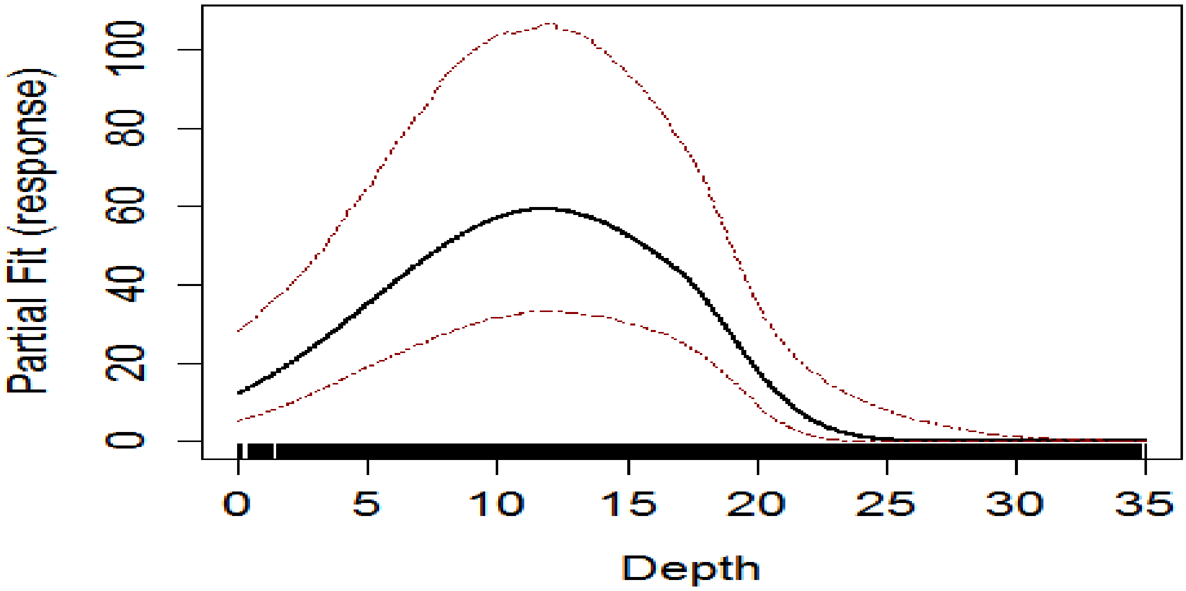

Figure 6: Partial fit for the Depth covariate with 95% CI. Scale of the response.

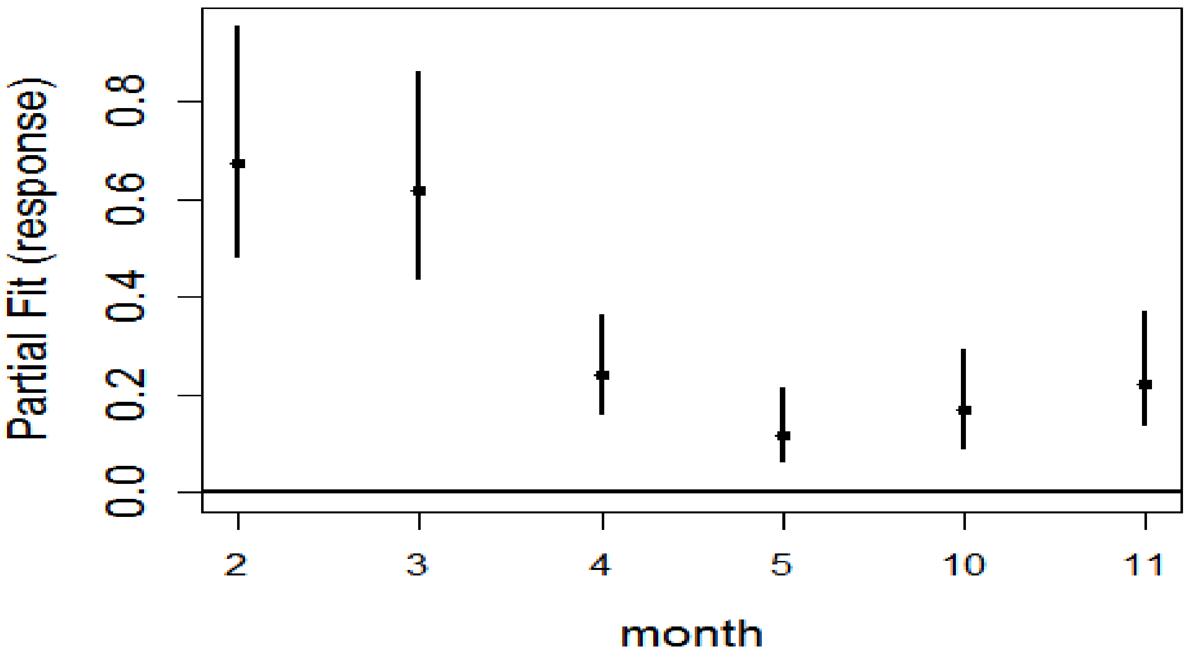

Figure 7: Partial fit for the month covariate with 95% CI. Scale of the response.

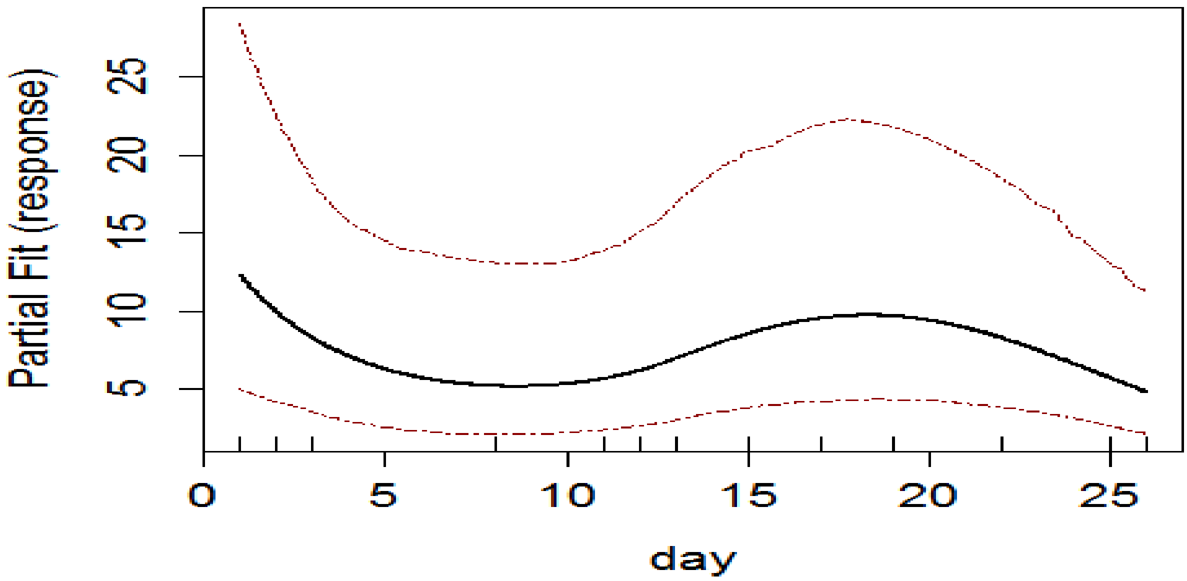

Figure 8: Partial fit for the day covariate with 95% CI. Scale of the response.

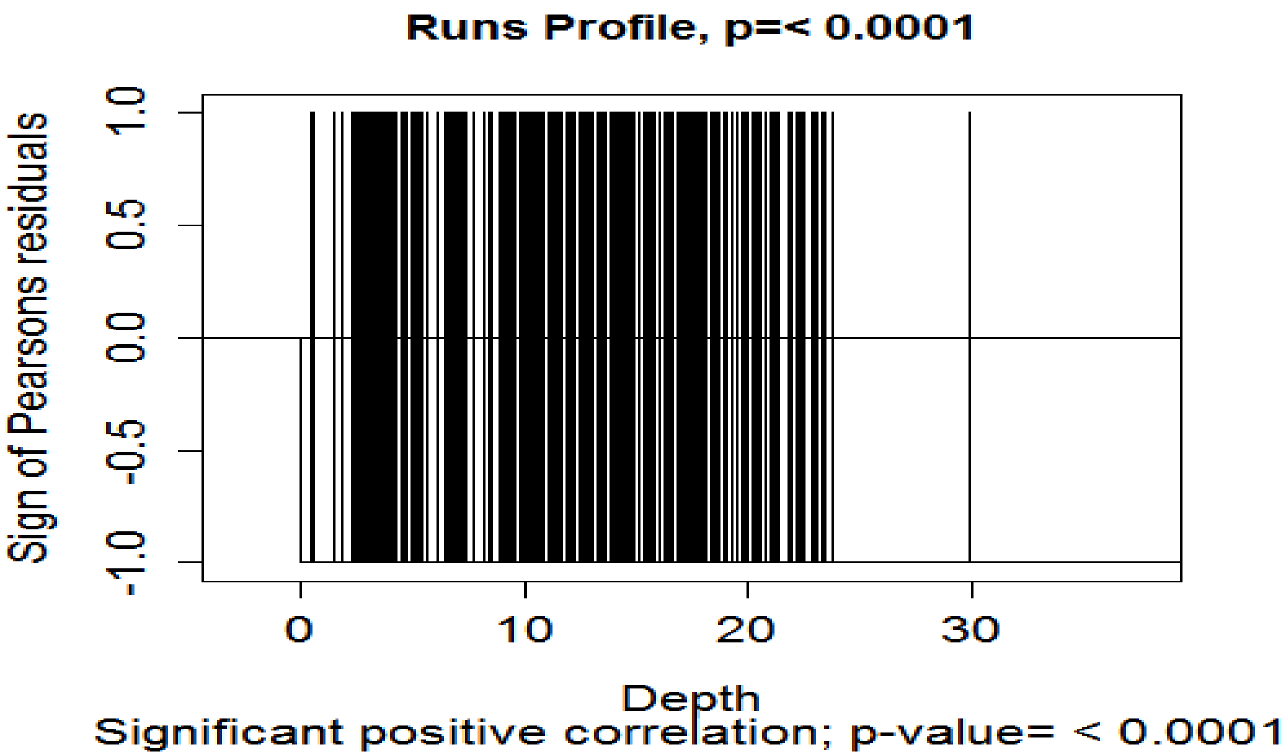

Figure 9: Runs plot and test for Pearsons residuals in Depth order.

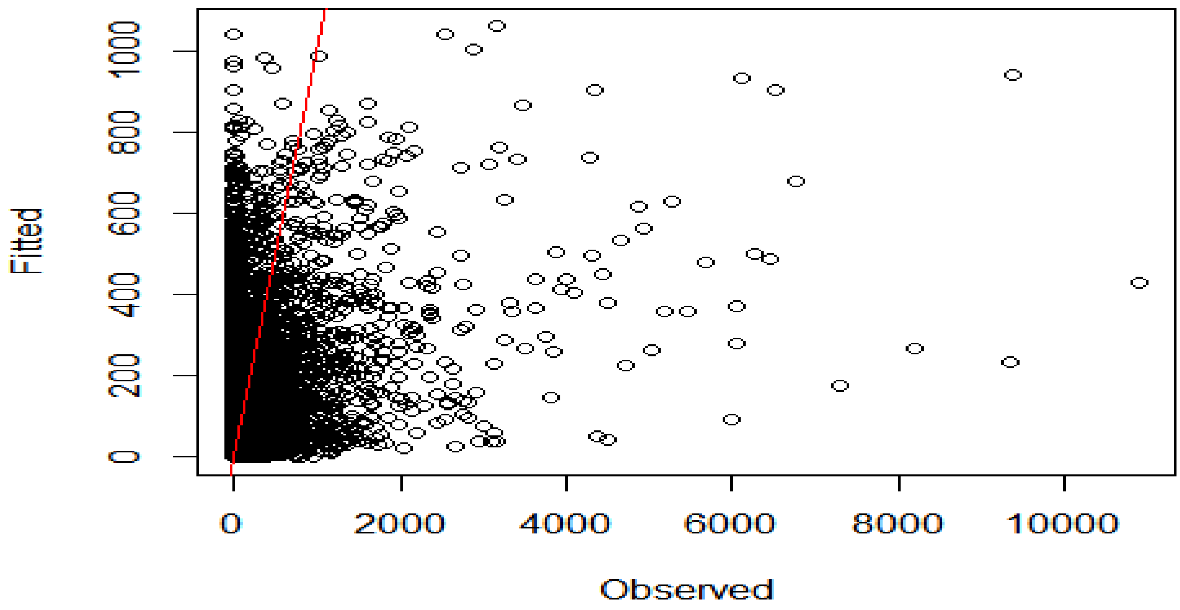

Figure 10: Observed vs Fitted plot. The red line is at 45°.

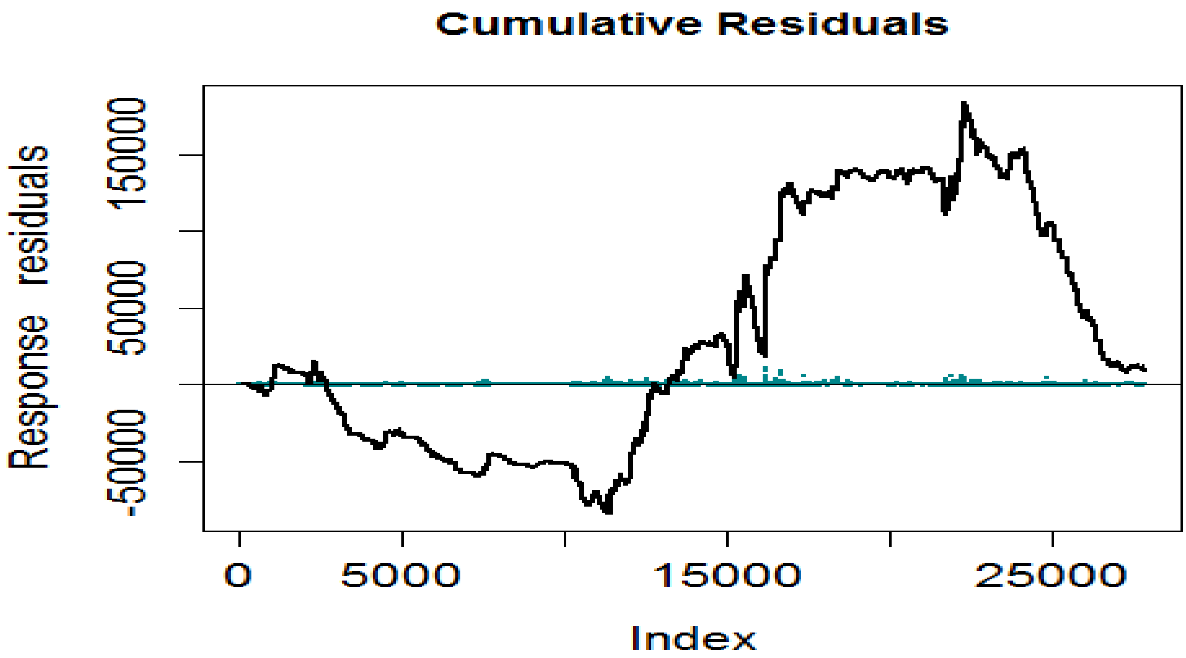

Figure 11: Cumulative residuals in observation order. The blue points represent residuals, while the black line is the cumulative sum of residuals.

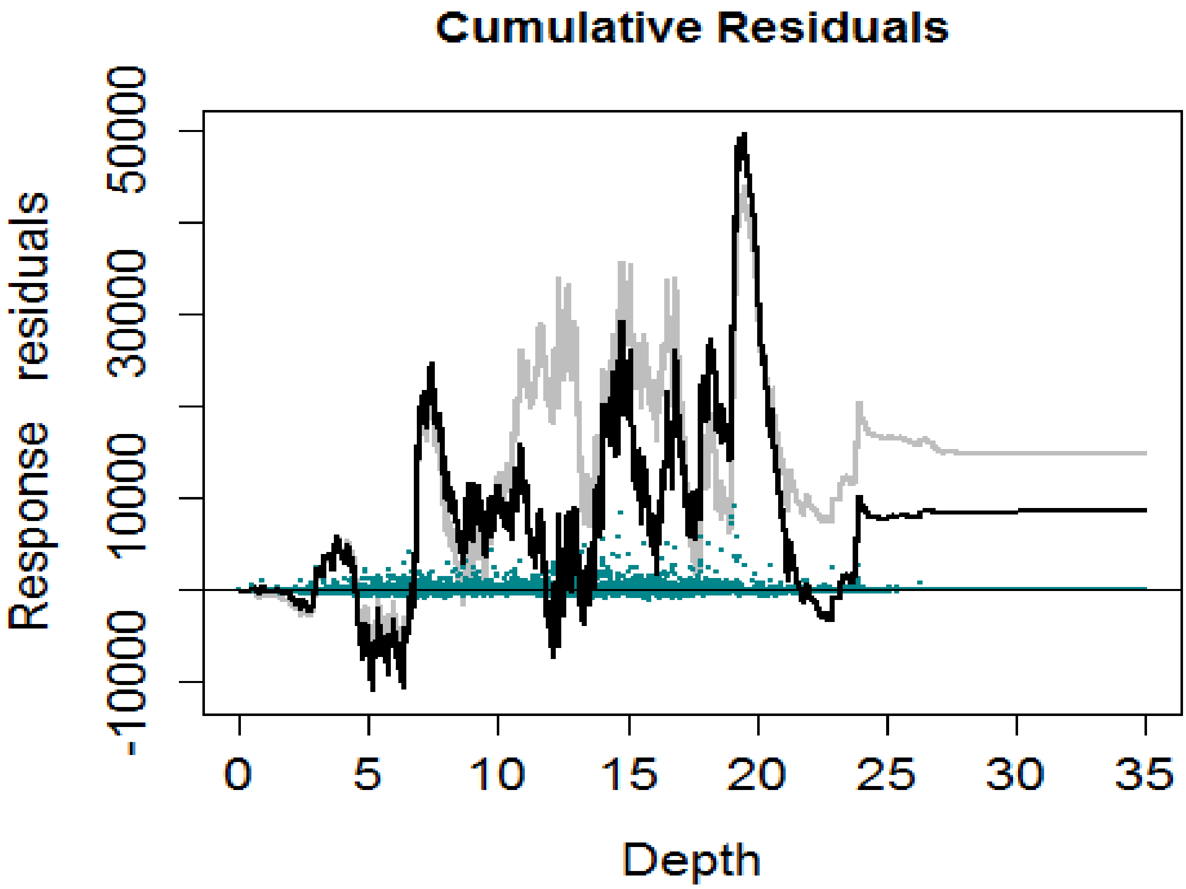

Figure 12: Cumulative residuals ordered by Depth value. Blue points are the residuals, the grey line represents what the cumulative residuals should be if the relationship was modelled correctly. The black line is the actual relationship.

Covariate	Description
Impact	Before= 0 / after= 1 the construction of the turbines
Day	Day of month of the observation
x.pos	X coordinate of the observation
y.pos	Y coordinate of the observation
Depth	Sea depth for that particular observation
TNO	Transect number: unique identifier for each transect

Table 2: Name and short description of covariates used in the model.

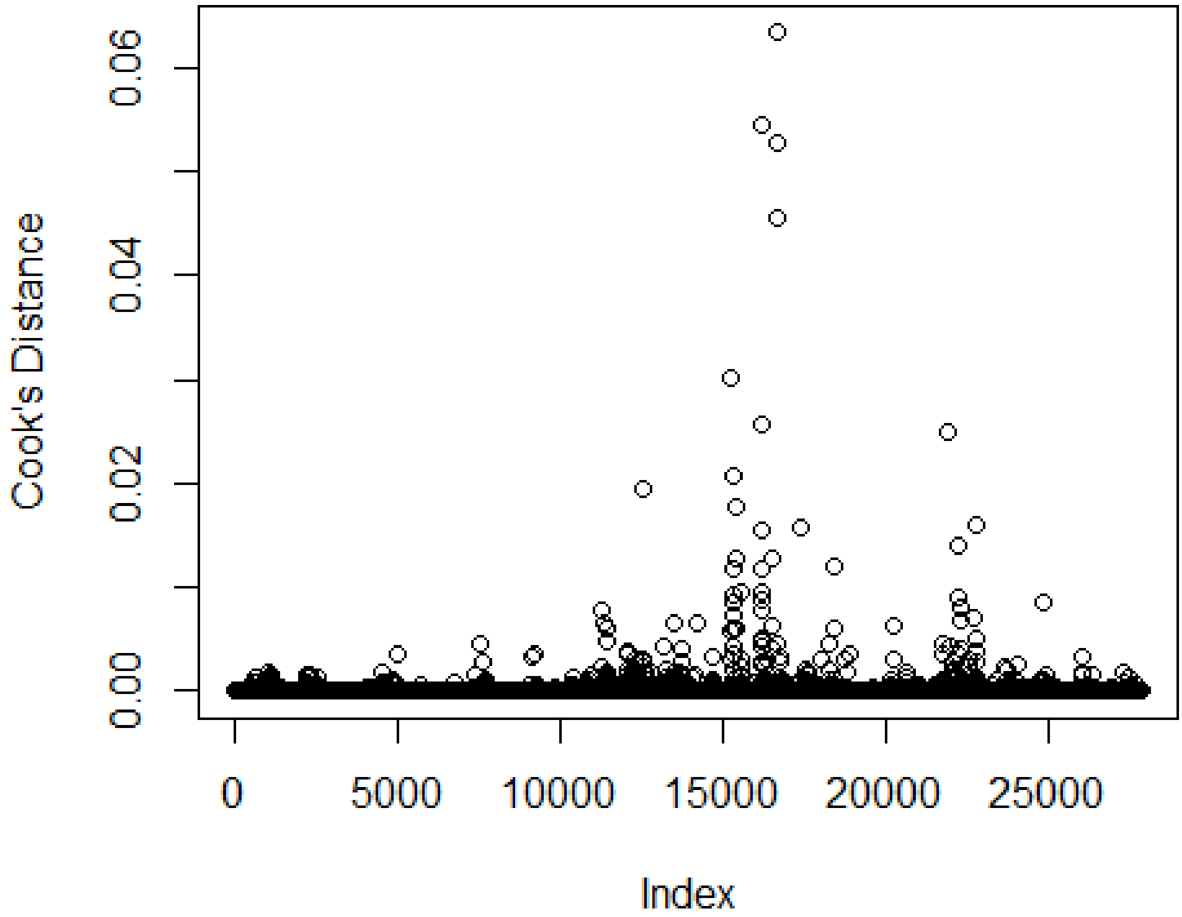

Figure 13: Cook's distance in observation order. Note the large values between Index 15000 and 20000.

6 REFERENCES

Scott-Hayward, L., Mackenzie, M., Donovan , C., Walker, C., & Ashe, E. (2014). Complex Region Spatial Smoother (CReSS). *Journal of Computational and Graphical Statistics, 23*(2).

Walker, C., Mackenzie, M., Donovan, C., & O'Sullivan, M. (2010). SALSA - a Spatially Adaptive Local Smoothing Algorithm. *Journal of Statistical Computation and Simulation, 81*(2), 179-191.