



**Marubeni**



# Appendix 11.4, Annex A: Offshore Ornithology MRSea Validation Methodology

**Array EIA Report**  
2024

Revision	Comments	Author	Checker	Approver
FINAL	Final	RPS	RPS	RPS

Approval for Issue		
For and on behalf of Ossian OWFL	Paul Darnbrough	28 June 2024

Prepared by:	RPS Energy
Prepared for:	Ossian Offshore Wind Farm Limited (OWFL)
Checked by:	Andrew Logie
Accepted by:	Fraser Malcolm
Approved by:	Paul Darnbrough

© Copyright RPS Group Plc. All rights reserved.

The report has been prepared for the exclusive use of our client.

The report has been compiled using the resources agreed with the client and in accordance with the scope of work agreed with the client. No liability is accepted by RPS for any use of this report, other than the purpose for which it was prepared. The report does not account for any changes relating to the subject matter of the report, or any legislative or regulatory changes that have occurred since the report was produced and that may affect the report. RPS does not accept any responsibility or liability for loss whatsoever to any third party caused by, related to or arising out of any use or reliance on the report.

RPS accepts no responsibility for any documents or information supplied to RPS by others and no legal liability arising from the use by others of opinions or data contained in this report. It is expressly stated that no independent verification of any documents or information supplied by others has been made.

RPS has used reasonable skill, care and diligence in compiling this report and no warranty is provided as to the report's accuracy.

# CONTENTS

- 1. MRSea Validation Methodology .....1
  - 1.1. MRSea Modelling .....1
  - 1.2. Guillemot Breeding Season.....1

## TABLES

- Table 1.1: Collinearity Test Between Predictor Variables Led to Exclusion of Distance Variables ..... 1
- Table 1.2: Highlights Correlation Between Variables After Distance Variables Were Removed..... 1

## FIGURES

- Figure 1.1: Model Validation of Initial Model, 1D and 2D Model Using Cross-Validation ..... 1
- Figure 1.2: Summary Outputs of the Best Fitting 2D Model ..... 2
- Figure 1.3: Auto-correlation Plot of “Transect” Blocking Structure..... 2
- Figure 1.4: ANOVA Table Providing Insight into the significance of Each Predictor Within the Best Fitting 2D Model ..... 2
- Figure 1.5: Cumulative Residuals of Different Environmental Predictors in the Initial Model With Linear Covariates, Showing Areas of Systematic Over- And Under-Prediction..... 3
- Figure 1.6: Runs Profiles of Pearsons Residuals for the Linear Model ..... 3
- Figure 1.7: Partial Model Fits at the Link Scale for Each Parameter Used in the Final Model ..... 4
- Figure 1.8: Partial Model Fits at the Link Scale for Each Parameter Used in the Final Model ..... 4
- Figure 1.9: Plot of Observed Versus Fitted Values of the Final Model ..... 4
- Figure 1.10: Fitted Values Versus Scaled Pearsons Residuals. Measures the Change in Covariance as a Result of Removing Each Data Point in Turn, which is Indicative of How Influential that Data Point is in the Precision of Parameter Estimates. The PRESS Statistic Plot Shows the Change in the PRESS (Predicted Residual Sum of Squares) as a Result of Omitting Each Data Point, which is Indicative of the Influence of that Data Point on the Sensitivity of Model Predictions..... 4
- Figure 1.11: The “COVRATIO” (Covariance Ratio)..... 5
- Figure 1.12: Plot of Mean Residual Variance Against Mean Fitted for Each 5% Quantile of Fitted Values ..... 5

# 1. MRSEA VALIDATION METHODOLOGY

## 1.1. MRSEA MODELLING

1. Covariates were first checked for collinearity, which revealed distance variables (distance to coast) had strong collinearity issues with other predictors, and were subsequently left out of further analysis Table 1.1 and Table 1.2). Variance inflation factors for these variables were above 10, where any that are above 2.5 are considered unacceptable (Johnston *et al.*, 2018).

Table 1.1: Collinearity Test Between Predictor Variables Led to Exclusion of Distance Variables

```
> car::vif(initialModel)
              GVIF Df GVIF^(1/(2*Df))
as.factor(Month)  1.028533 23      1.000612
depth            1.393959  1      1.180660
d2coast          14100.766982  1    118.746650
x.pos            9577.558472  1     97.865001
y.pos            898.744427  1     29.979066
```

Table 1.2: Highlights Correlation Between Variables After Distance Variables Were Removed

```
> car::vif(initialModel)
              GVIF Df GVIF^(1/(2*Df))
as.factor(Month)  1.002877 23      1.000062
depth            1.391795  1      1.179743
x.pos            1.923439  1      1.386881
y.pos            1.610617  1      1.269101
```

2. MRSea models were then constructed using these predictor variables. For each bird species, the basic model from which MRSea builds the more complex model has the following form:

$$\text{Species Count} \sim \text{Month} + \text{offset}(\log(\text{area})), \text{family} = \text{quasipoisson}$$

3. At each step of the modelling process, RPS performed diagnostic checks to confirm whether models are appropriate. Example diagnostic plots are shown in section 1.2 for the guillemot breeding season model.

## 1.2. GUILLEMOT BREEDING SEASON

```
> cv_1d_initial
[1] 17.35773
> cv_1d
[1] 16.3977
> cv_2d
[1] 13.6451
```

Figure 1.1: Model Validation of Initial Model, 1D and 2D Model Using Cross-Validation

```
Call:
gamMRSea(formula = response ~ as.factor(Month) + bs(depth, knots = splineParams[[2]]$knots,
degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) +
  LRF.g(radiusIndices, dists, radii, aR) + offset(log(area)),
family = quasipoisson(link = log), data = datsub, splineParams = splineParams)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-7.0199 -1.1549 -0.7175  0.2474 15.3781

Coefficients:
            Estimate Std. Error Robust S.E. t value Pr(>|t|)
(Intercept)    0.27586    0.77857    0.57672    0.478 0.632429
as.factor(Month)2021-04 -1.02895    0.34064    0.58520   -1.758 0.078734 .
as.factor(Month)2021-05  3.83423    0.15481    0.24061   15.935 < 2e-16 ***
as.factor(Month)2021-06  3.54757    0.23720    0.44603    7.954 2.03e-15 ***
as.factor(Month)2021-07  2.19852    0.14384    0.16052   13.696 < 2e-16 ***
as.factor(Month)2022-03  1.49012    0.16823    0.17002    8.764 < 2e-16 ***
as.factor(Month)2022-04  3.59704    0.18386    0.21140   17.015 < 2e-16 ***
as.factor(Month)2022-05 -0.53435    0.21308    0.18451   -2.896 0.003788 **
as.factor(Month)2022-06  3.15684    0.25610    0.22800   13.846 < 2e-16 ***
as.factor(Month)2022-07  4.53555    0.14631    0.15994   28.358 < 2e-16 ***
s(depth)1      -0.98576    1.04472    1.05698   -0.933 0.351039
s(depth)2      -0.03149    0.75539    0.53283   -0.059 0.952875
s(depth)3      -0.14129    0.77067    0.58023   -0.244 0.807623
s(depth)4      -0.40268    0.77165    0.58207   -0.692 0.489077
s(depth)5      0.51427    0.77726    0.64193    0.801 0.423076
s(x.pos, y.pos)b1 -3.04375    0.98438    1.60335   -1.898 0.057678 .
s(x.pos, y.pos)b2  5.27989    0.88813    1.50266    3.514 0.000444 ***
s(x.pos, y.pos)b3  4.80605    0.32274    0.78901    6.091 1.17e-09 ***
s(x.pos, y.pos)b4  2.03754    0.41240    0.70406    2.894 0.003813 **
s(x.pos, y.pos)b5 -3.23760    0.30243    0.46364   -6.983 3.09e-12 ***
s(x.pos, y.pos)b6 -2.20026    0.19375    0.42188   -5.215 1.87e-07 ***
s(x.pos, y.pos)b7 -3.37021    0.20309    0.38770   -8.693 < 2e-16 ***
s(x.pos, y.pos)b8 -3.11032    0.46013    0.57192   -5.438 5.52e-08 ***
s(x.pos, y.pos)b9 -3.75286    0.37778    0.63645   -5.897 3.84e-09 ***
s(x.pos, y.pos)b10 -5.68795    0.64886    0.79913   -7.118 1.18e-12 ***
s(x.pos, y.pos)b11 -2.04703    0.21451    0.27508   -7.442 1.08e-13 ***
s(x.pos, y.pos)b12 -1.43793    0.26955    0.29041   -4.951 7.50e-07 ***
s(x.pos, y.pos)b13 -2.01387    0.23255    0.27520   -7.318 2.73e-13 ***
s(x.pos, y.pos)b14 -2.31715    0.22900    0.38386   -6.036 1.64e-09 ***
s(x.pos, y.pos)b15 -1.34147    0.19077    0.34167   -3.926 8.69e-05 ***
s(x.pos, y.pos)b16  1.54217    0.25264    0.43792    3.522 0.000431 ***
s(x.pos, y.pos)b17  2.96493    0.21331    0.51049    5.808 6.53e-09 ***
s(x.pos, y.pos)b18 -1.88403    0.53189    0.38904   -4.843 1.30e-06 ***
s(x.pos, y.pos)b19 -2.21817    0.45751    0.32579   -6.809 1.05e-11 ***
s(x.pos, y.pos)b20 -1.56958    0.10191    0.12966  -12.106 < 2e-16 ***
s(x.pos, y.pos)b21 -1.26587    0.12190    0.30302   -4.177 2.97e-05 ***
s(x.pos, y.pos)b22 -0.87683    0.09434    0.16871   -5.197 2.07e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for quasipoisson family taken to be 3.51005)

Null deviance: 48646 on 9240 degrees of freedom
Residual deviance: 23682 on 9204 degrees of freedom
AIC: NA

Max Panel Size = 73; Number of panels = 306
Number of Fisher Scoring iterations: 6
```

Figure 1.2: Summary Outputs of the Best Fitting 2D Model

4. An auto-correlation function was then plotted to check whether this blocking structure was appropriate (volume 3, appendix 11.4, annex B, Table 1.1). The grey lines indicate residual correlation within transects, and the red line indicates the mean residual correlation. Residual correlation needs to be independent between blocks (red line), but can be present within blocks (grey lines). The fact that the red line quickly drops to zero shows independence of residual correlation among blocks, meaning the chosen blocking structure was appropriate.

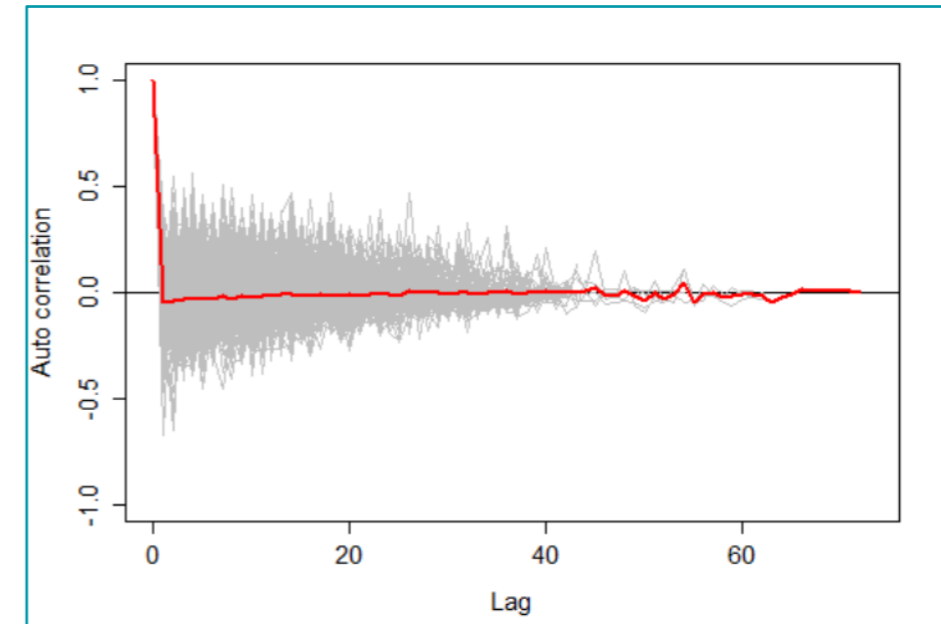


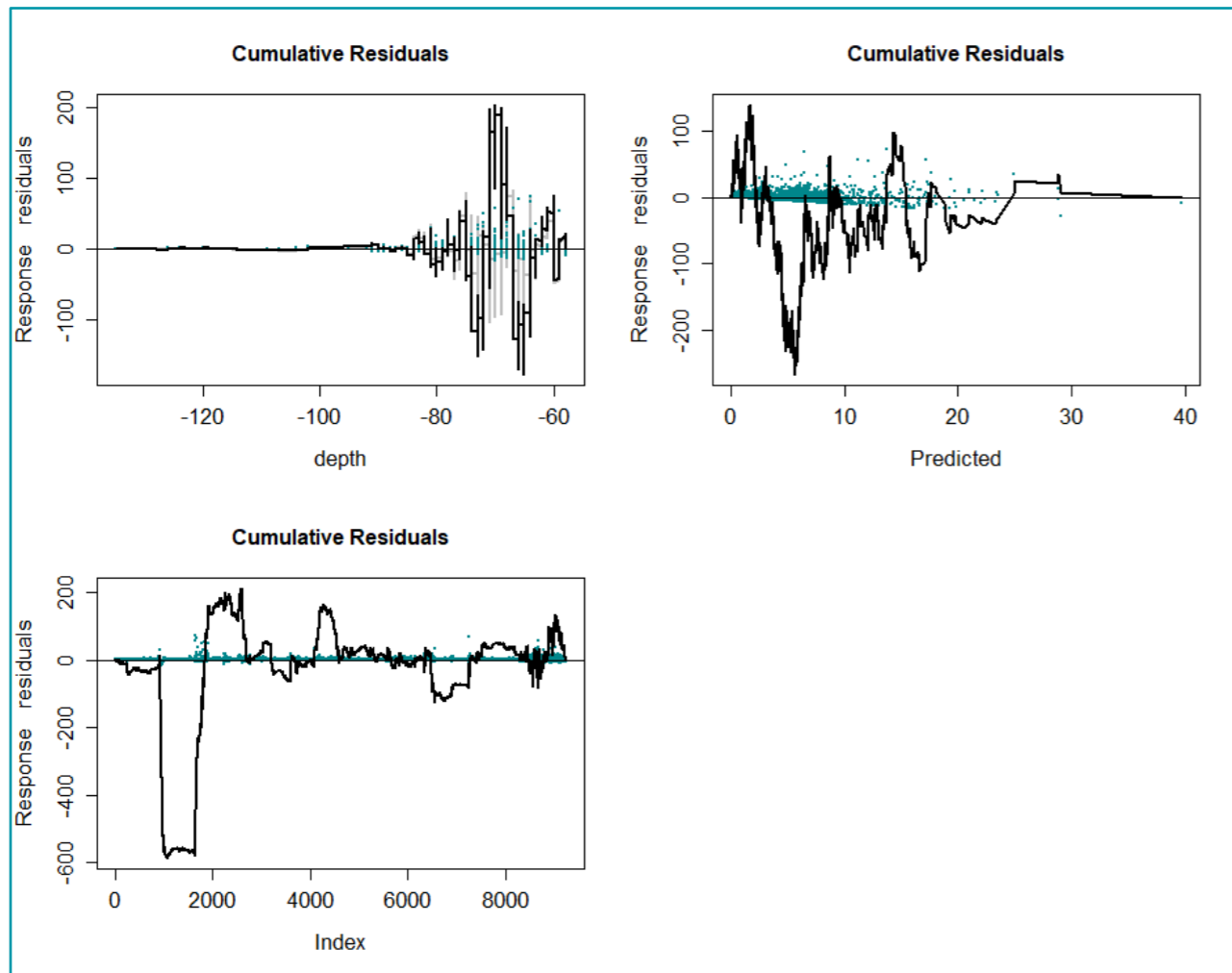
Figure 1.3: Auto-correlation Plot of "Transect" Blocking Structure

```
> anova(salsa2dbest)
Analysis of 'Wald statistic' Table
Model: quasipoisson, link: log
Response: response
Marginal Testing
Max Panel Size = 73; Number of panels = 306

            Df      X2 P(>|Chi|)
as.factor(Month)  9 1084.31 < 2e-16 ***
s(depth)         5   10.15  0.07106 .
s(x.pos, y.pos) 22  908.81 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

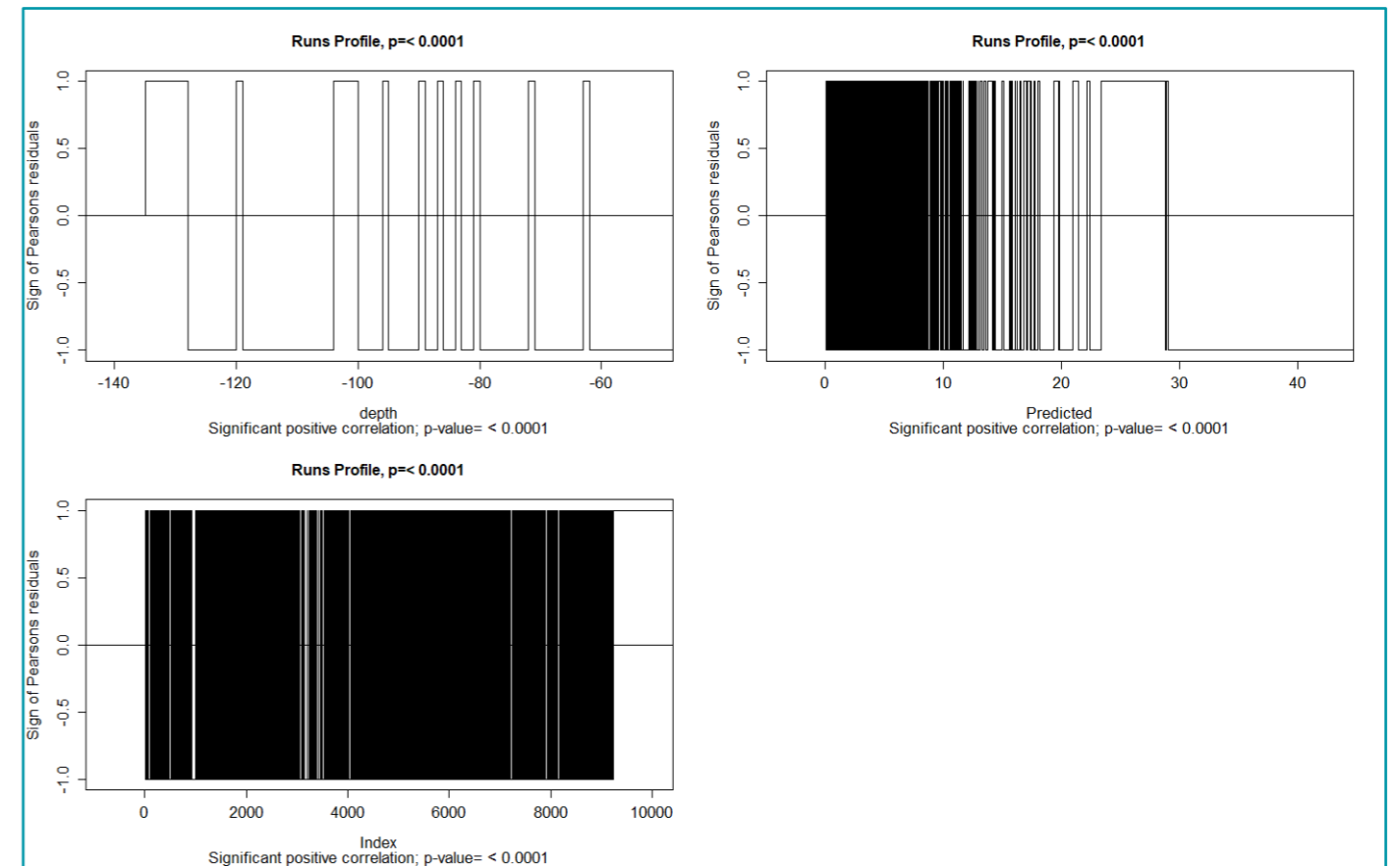
Figure 1.4: ANOVA Table Providing Insight into the significance of Each Predictor Within the Best Fitting 2D Model

5. The cumulative residuals of the model were checked, with the blue dots are the residuals, and the black line is the line of cumulative residual. The plot demonstrated that there was over- and under-prediction (Figure 1.5) and therefore a nonlinear model was warranted to break up some of this over and under prediction.



**Figure 1.5: Cumulative Residuals of Different Environmental Predictors in the Initial Model With Linear Covariates, Showing Areas of Systematic Over- And Under-Prediction**

6. Next, it was checked whether the runs profiles were random, which they were not (Figure 1.6). The low p-values indicate that there is correlation in the residuals, no matter how they are ordered. This means the transects need to be modelled as a blocking variable to control for residual variance, which is what the model does.



**Figure 1.6: Runs Profiles of Pearson's Residuals for the Linear Model**

7. In the first (1-dimensional) stage of the MRSea model, the basic model was expanded to include water depth as both linear and smoothed explanatory variables. In stage 1, RPS implemented tenfold cross validation (CV), which fits 1-dimensional splines for each environmental covariate through 90% of the data, and then uses this to predict the other 10% of the data. To get a robust estimate for each covariate, this process is repeated ten times.
8. In the second (2-dimensional) stage (the actual spatial model), the x-y coordinates were fitted to the best model from stage 1, using month as an interaction term, allowing for different density surfaces to be estimated for different months. In stage 2, RPS opted for a model selection procedure using 10-fold cross validation (CV) again.
9. Partial fits of the final model were then checked for any irregularities (Figure 1.7). Confidence limits of the environmental parameters were particularly wide for the depth variable, implying that their initial significance in the first stage was altered by fitting the spatial model in stage 2.

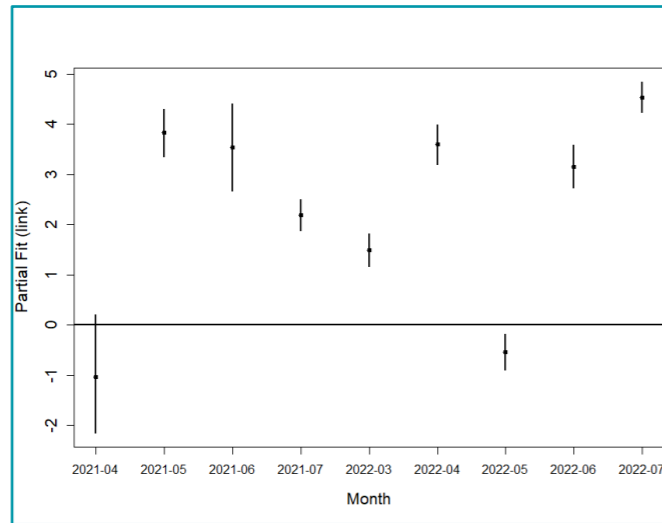


Figure 1.7: Partial Model Fits at the Link Scale for Each Parameter Used in the Final Model

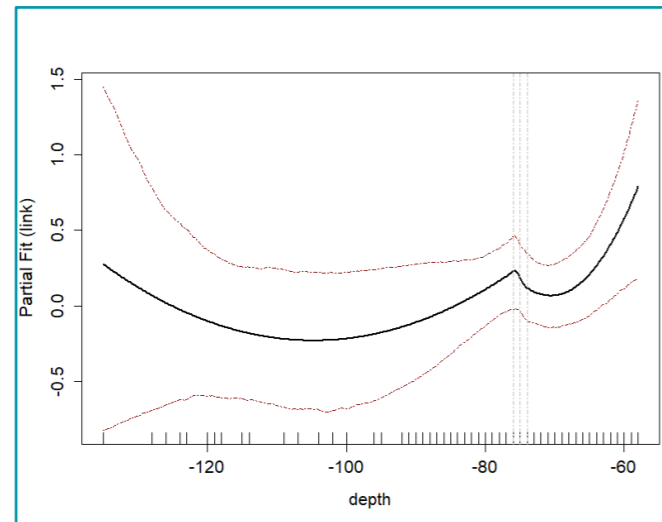


Figure 1.8: Partial Model Fits at the Link Scale for Each Parameter Used in the Final Model

10. When plotting observed versus fitted values however, the concordance correlation can be considered good (0.5497), the only deviation being that some of the higher data values were underpredicted by the model. However, this will not lead to systematic underestimation of abundances. Model fit was also good with a marginal R squared value of 0.3814. There were no unusual patterns revealed by locally weighted least squares regression.
11. The model made certain assumptions about the data. The most important assumption was that the effects of water depth was common to all months of data. Note that this does not imply that the relative distribution of birds is the same across all months, because the density and distribution landscape is altered for each month in stage 2 by the 2-dimensional model by using "survey date" as an interaction term.

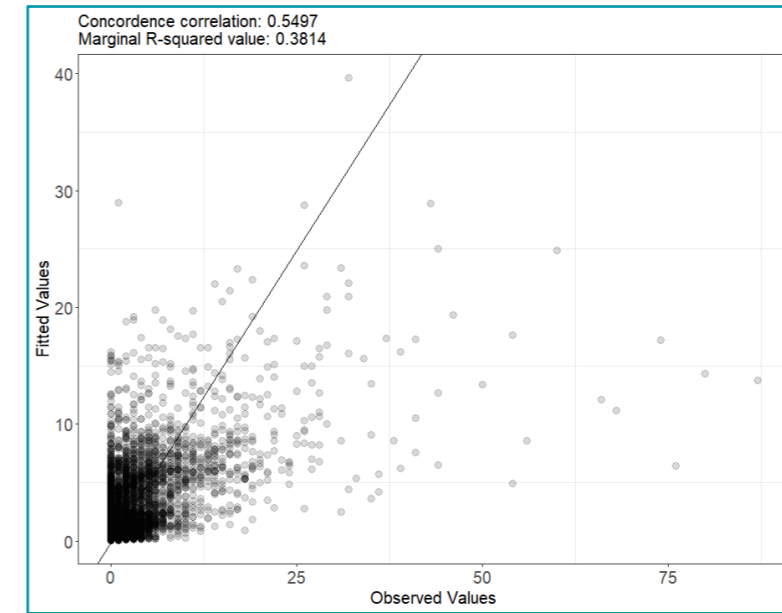


Figure 1.9: Plot of Observed Versus Fitted Values of the Final Model

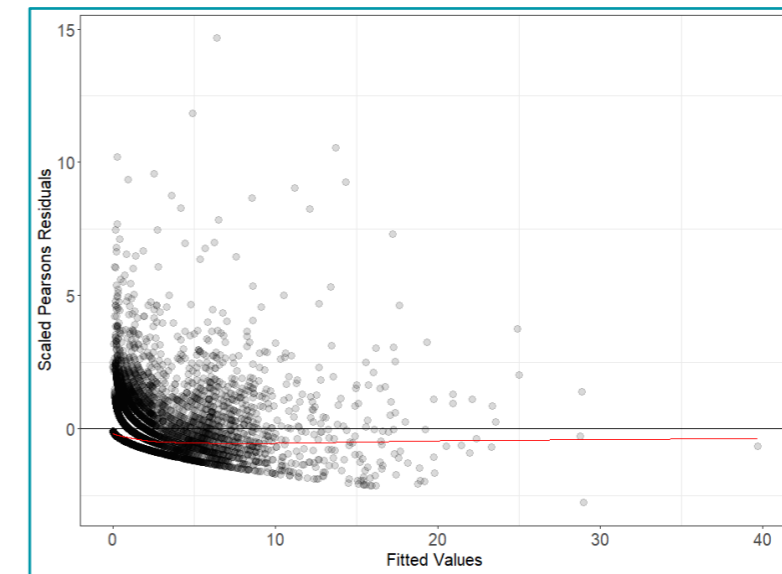


Figure 1.10: Fitted Values Versus Scaled Pearson's Residuals. Measures the Change in Covariance as a Result of Removing Each Data Point in Turn, which is Indicative of How Influential that Data Point is in the Precision of Parameter Estimates. The PRESS Statistic Plot Shows the Change in the PRESS (Predicted Residual Sum of Squares) as a Result of Omitting Each Data Point, which is Indicative of the Influence of that Data Point on the Sensitivity of Model Predictions

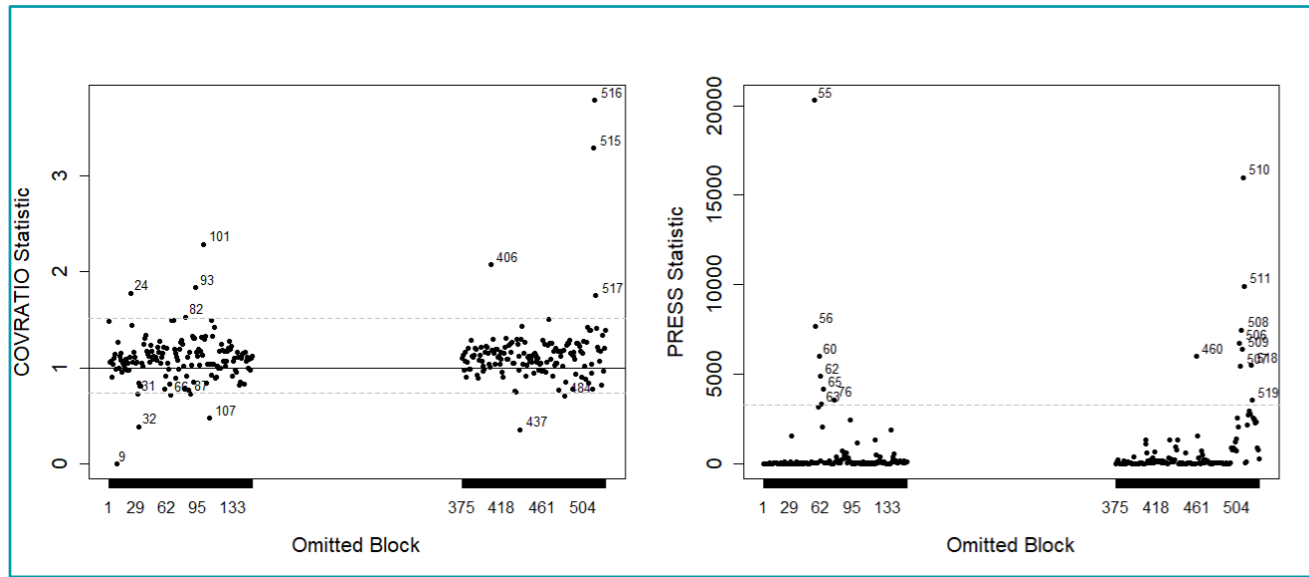


Figure 1.11: The “COVRATIO” (Covariance Ratio)

12. Lastly, mean residuals were plotted where the black line is the 1:1 line while the red line has a gradient equal to the estimated dispersion parameter. A quasi-Poisson model (as used) assumes variance is proportional to the mean and therefore ideally the points would be scattered about the red line. This figure therefore also indicates that the data are not truly quasi-Poisson, with greater residual variance than expected at higher fitted values.

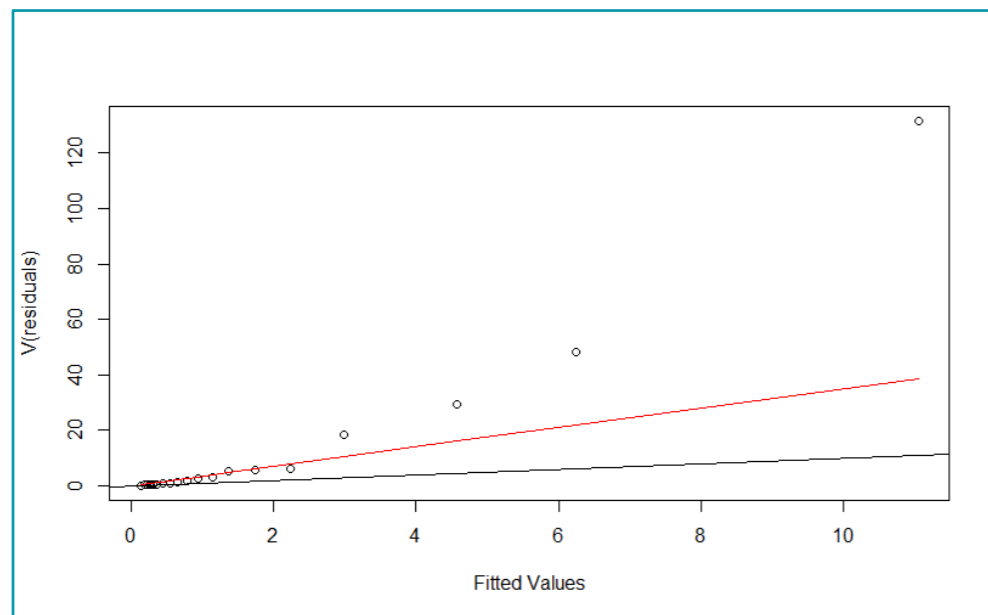


Figure 1.12: Plot of Mean Residual Variance Against Mean Fitted for Each 5% Quantile of Fitted Values



# Ossian



**Marubeni**



**Ossian Offshore Wind Farm Limited**

Inveralmond House  
200 Dunkeld Road  
Perth  
PH1 3AQ

**Project Office**

Fourth Floor  
10 Bothwell Street  
Glasgow  
G2 6NT

[ossianwindfarm.com](http://ossianwindfarm.com)