# Brunswick Point Peep Models

Sarah Popov

2022-11-29

Roberts Bank, Delta, BC, is situated in the great Pacific Flyway and serves as an important stopover for peeps migrating north in the spring. It therefore hosts a large seasonal population of peeps: namely, Western sandpiper (WESA) and Dunlin (DUNL), that rely on the seasonal nutritional bounty provided by the Fraser River delta.

This document describes the suite of models used to estimate yearly changes in spatial distribution and abundance of peeps in the Roberts Bank estuary.

There are two datasets used in this modelling pipeline:

1. `sr`, which contains species ratio data (WESA:DUNL)
2. `dat`, which contains bird counts + environmental covariates

For an interactive interface with the `dat` dataset used in this document, see https://popovs.shinyapps.io/peepr/.

**sr summary**

```
##   survey_date              wesa             dunl            total
##  Min.   :1997-04-21   Min.   :   0.0   Min.   :   0.0   Min.   :  48.0
##  1st Qu.:2005-04-20   1st Qu.: 126.8   1st Qu.:  42.0   1st Qu.: 287.2
##  Median :2009-04-23   Median : 355.0   Median :  93.0   Median : 524.5
##  Mean   :2009-09-05   Mean   : 470.1   Mean   : 176.7   Mean   : 646.8
##  3rd Qu.:2015-04-16   3rd Qu.: 719.0   3rd Qu.: 203.5   3rd Qu.: 943.8
##  Max.   :2022-05-04   Max.   :2605.0   Max.   :5000.0   Max.   :5047.0
##
##      p_wesa           p_dunl            year       julian_day
##  Min.   :  0.00   Min.   :  0.00   2006   : 23   Min.   :101.0
##  1st Qu.: 53.04   1st Qu.:  8.83   2019   : 23   1st Qu.:112.0
##  Median : 81.32   Median : 18.68   2005   : 18   Median :117.0
##  Mean   : 68.94   Mean   : 31.06   2007   : 18   Mean   :117.6
##  3rd Qu.: 91.17   3rd Qu.: 46.96   2008   : 18   3rd Qu.:123.0
##  Max.   :100.00   Max.   :100.00   2012   : 18   Max.   :137.0
##                                    (Other):200
##         dos.V1
##  Min.   :-2.2763470
##  1st Qu.:-0.7798519
##  Median :-0.0996268
##  Mean   :-0.0221924
##  3rd Qu.: 0.7166433
##  Max.   : 2.6212735
##
```

**dat summary**

```
##   survey_date          start_time          station
##  Min.   :1997-04-21   Length:1872        Length:1872
##  1st Qu.:2004-04-16   Class :character   Class :character
##  Median :2008-05-08   Mode  :character   Mode  :character
##  Mean   :2009-03-15
##  3rd Qu.:2014-04-19
##  Max.   :2022-05-04
##
##            station_n             station_s    mumblies_yn       mud_yn
##  Coal Port      :532   Bend           :  70   Mode :logical   Mode :logical
##  Pilings        :463   Coal Port      :  55   FALSE:1692      FALSE:1687
##  Brunswick Point:310   View corner    :  45   TRUE :180       TRUE :185
##  34th St pullout:202   34th St pullout:  17
##  View corner    :174   Pilings        :  15
##  Bend           :118   (Other)        :   1
##  (Other)        : 73   NA's           :1669
##   marsh_yn       tide_edge_yn    flying_yn        final_count
##  Mode :logical   Mode :logical   Mode :logical   Min.   :     0
##  FALSE:1806      FALSE:1797      FALSE:1858      1st Qu.:   500
##  TRUE :66        TRUE :75        TRUE :14        Median :  4125
##                                                  Mean   : 12775
##                                                  3rd Qu.: 14500
##                                                  Max.   :222500
##
##      p_wesa          elev_min         elev_max      elev_median
##  Min.   :  0.00   Min.   :-0.040   Min.   :3.960   Min.   :2.240
##  1st Qu.: 52.38   1st Qu.: 0.830   1st Qu.:4.300   1st Qu.:3.070
##  Median : 80.66   Median : 1.210   Median :4.430   Median :3.255
##  Mean   : 68.74   Mean   : 1.226   Mean   :4.418   Mean   :3.238
##  3rd Qu.: 90.97   3rd Qu.: 1.630   3rd Qu.:4.560   3rd Qu.:3.415
##  Max.   :100.00   Max.   : 2.420   Max.   :4.850   Max.   :3.735
##  NA's   :589
##    elev_mean       elev_range         flow       total_precip
##  Min.   :2.660   Min.   :1.870   Min.   : 996   Min.   : 0.000
##  1st Qu.:2.966   1st Qu.:2.710   1st Qu.:2420   1st Qu.: 0.000
##  Median :3.036   Median :3.210   Median :3090   Median : 0.000
##  Mean   :3.041   Mean   :3.192   Mean   :3224   Mean   : 2.099
##  3rd Qu.:3.120   3rd Qu.:3.600   3rd Qu.:3990   3rd Qu.: 2.000
##  Max.   :3.422   Max.   :4.640   Max.   :7830   Max.   :28.200
##                                  NA's   :5      NA's   :4
##    mean_temp           u                 v              windspd
##  Min.   : 3.80   Min.   :-19.8067   Min.   :-42.3467   Min.   : 5.292
##  1st Qu.: 9.20   1st Qu.: -3.8203   1st Qu.: -9.9586   1st Qu.:10.583
##  Median :10.50   Median : -0.9551   Median : -0.7814   Median :12.958
##  Mean   :10.65   Mean   : -0.7078   Mean   : -1.8559   Mean   :14.326
##  3rd Qu.:11.90   3rd Qu.:  2.5419   3rd Qu.:  7.9826   3rd Qu.:16.958
##  Max.   :17.50   Max.   : 15.0489   Max.   : 20.4541   Max.   :43.750
##
##     wind_deg        station_n_no    station_s_no    station_diff    julian_day
##  Min.   :-177.35   Min.   :1.000   Min.   :3.000   Min.   :0.00   Min.   :105
##  1st Qu.: -77.12   1st Qu.:4.000   1st Qu.:5.000   1st Qu.:1.00   1st Qu.:112
##  Median : -24.63   Median :5.000   Median :6.000   Median :1.00   Median :118
```

```
##  Mean   :  14.62   Mean   :5.551   Mean   :6.094   Mean   :1.01   Mean   :118
##  3rd Qu.: 119.16   3rd Qu.:8.000   3rd Qu.:8.000   3rd Qu.:1.00   3rd Qu.:124
##  Max.   : 179.93   Max.   :8.000   Max.   :8.000   Max.   :2.00   Max.   :135
##                                    NA's   :1669    NA's   :1669
##         dos.V1
##  Min.   :-1.8256543
##  1st Qu.:-0.8427712
##  Median :-0.0003000
##  Mean   : 0.0000000
##  3rd Qu.: 0.8421712
##  Max.   : 2.3867017
##
```

## Species composition model

The daily ratio of Western sandpiper (WESA) to Dunlin (DUNL) across the entire study period is first modelled using a dataset of known species ratios (species ratios are not measured during every survey).

The ratios are modelled using a binomial generalized linear mixed model (binomial GLMM). The resulting predicted ratios are then used to estimate the number of WESA vs. DUNL per day.

### Five models are built and compared

Response variable: * `y` - WESA:DUNL ratio

Predictor variables: * `dos` - day of season (recentered/scaled Julian date) * `year` - year of survey

```
## Data: sr
## Models:
## lme5: y ~ 1 + (1 | year)
## lme4: y ~ dos + (1 | year)
## lme2: y ~ dos + I(dos^2) + (1 | year)
## lme3: y ~ dos + (dos | year)
## lme1: y ~ dos + I(dos^2) + (dos + I(dos^2) | year)
##      npar   AIC   BIC logLik deviance   Chisq Df Pr(>Chisq)
## lme5    2 64756 64764 -32376    64752
## lme4    3 25804 25815 -12899    25798 38954.1  1     <2e-16 ***
## lme2    4 16295 16310  -8144    16287  9510.6  1     <2e-16 ***
## lme3    5 22521 22540 -11255    22511     0.0  1          1
## lme1    9 11083 11117  -5532    11065 11446.2  4     <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The best-fit model is `lme1`. The residuals from `lme1` are appended to the `sr` dataset and another model is re-fit in order to estimate overdispersion. Because the standard deviation of the residuals is $< 1$, the model is deemed an appropriate candidate for predicting daily species ratios.

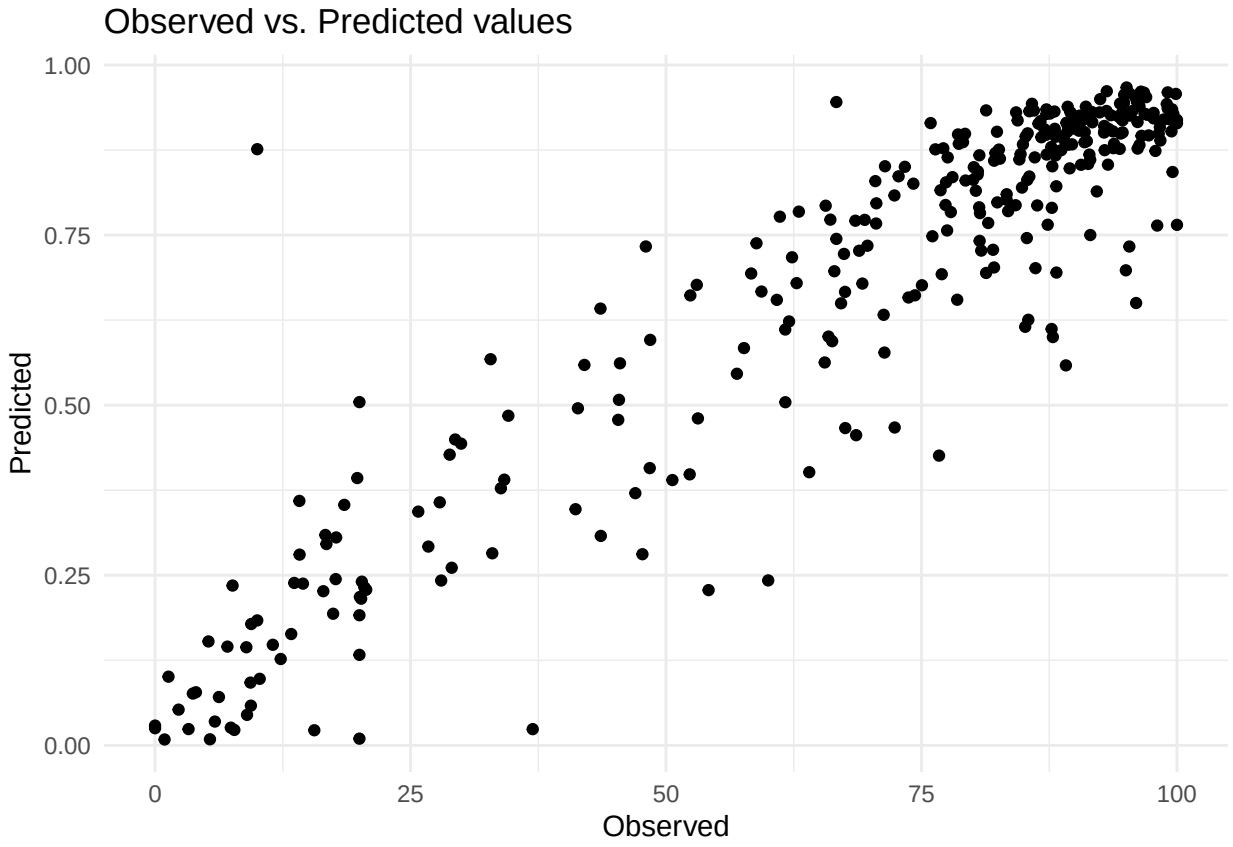
```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: y ~ dos + I(dos^2) + (dos + I(dos^2) | year) + (1 | resids)
##    Data: sr
```

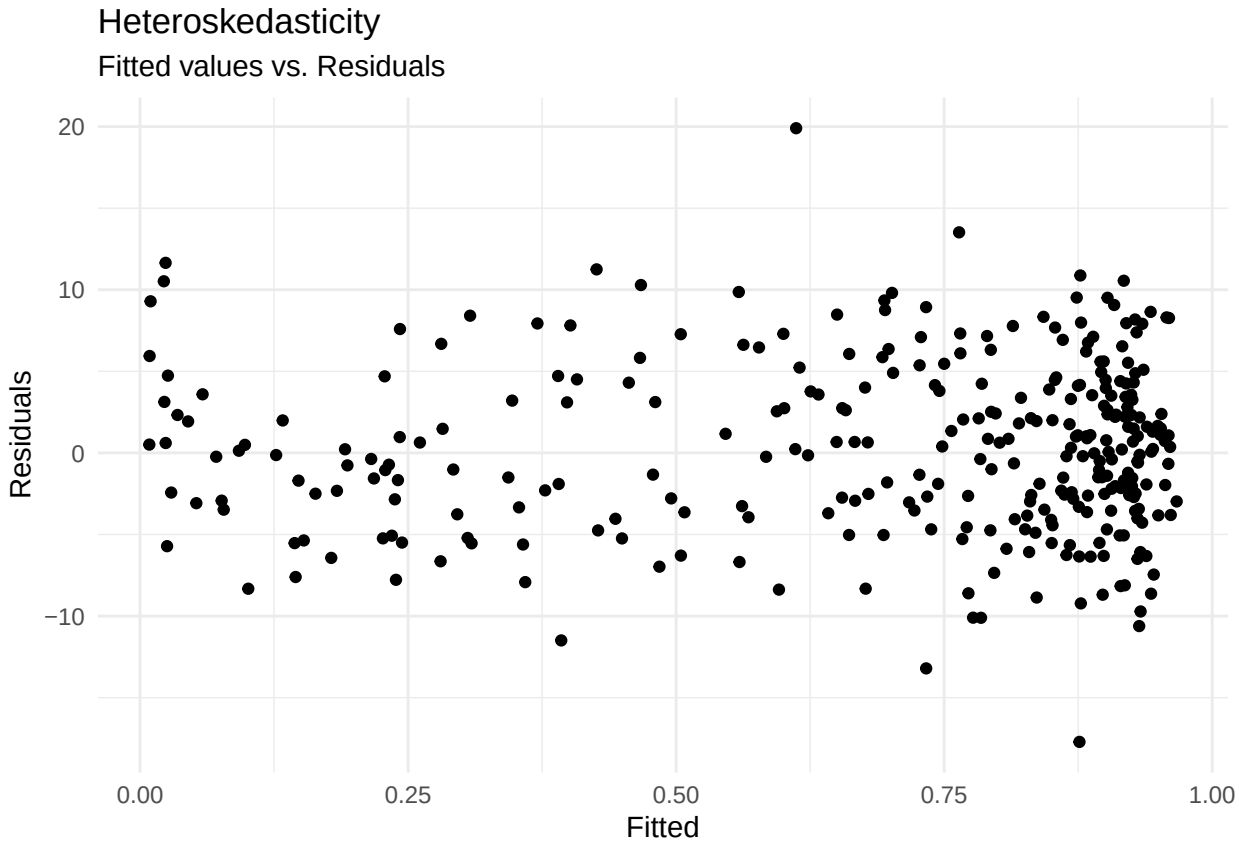
```
##       AIC       BIC    logLik  deviance  df.resid
##  3409.879  3447.499 -1694.939  3389.879       308
## Random effects:
##  Groups Name        Std.Dev. Corr
##  resids (Intercept) 0.94716
##  year   (Intercept) 0.65139
##         dos         0.06454  -1.00
##         I(dos^2)    0.41479  -0.65  0.68
## Number of obs: 318, groups:  resids, 318; year, 24
## Fixed Effects:
## (Intercept)          dos     I(dos^2)
##      1.8826       1.3770      -0.8085
```

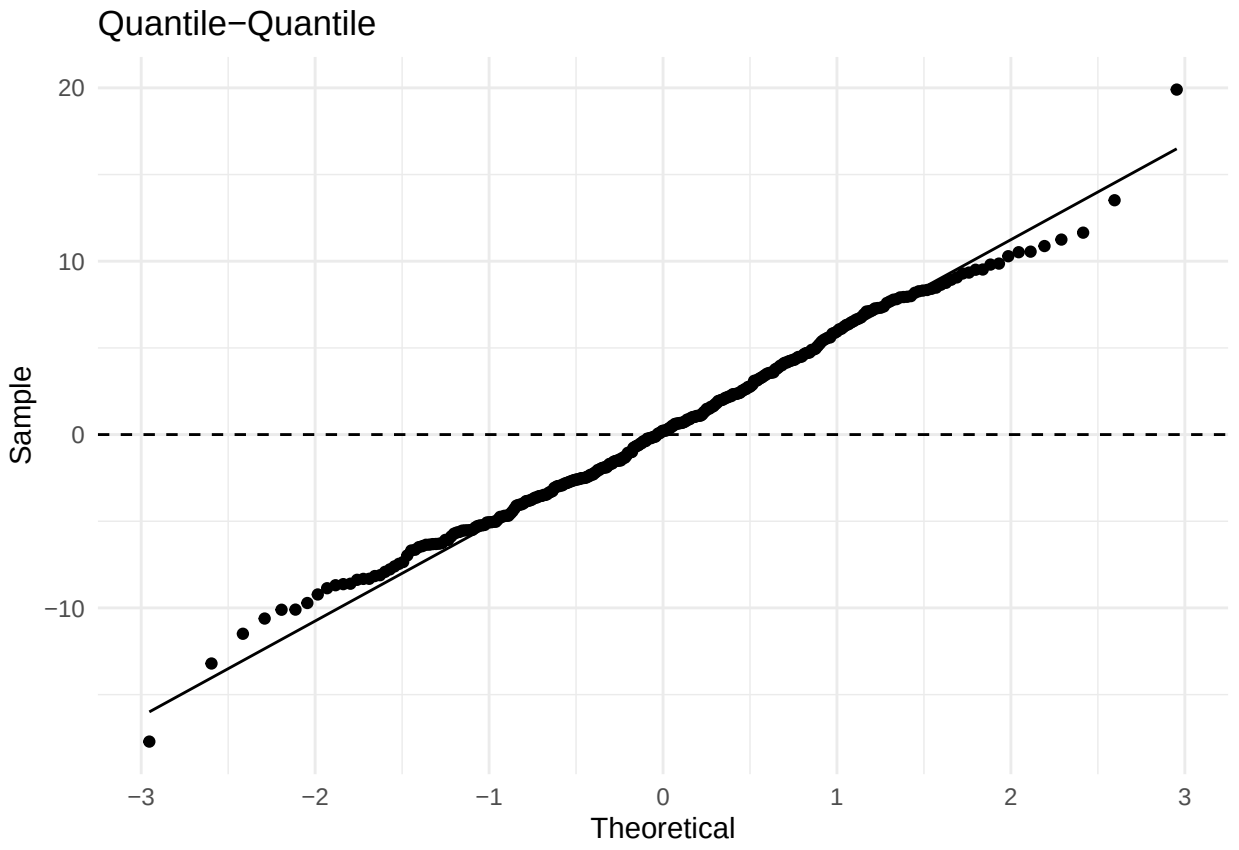
**Summary of best fit model (`lme1`)**

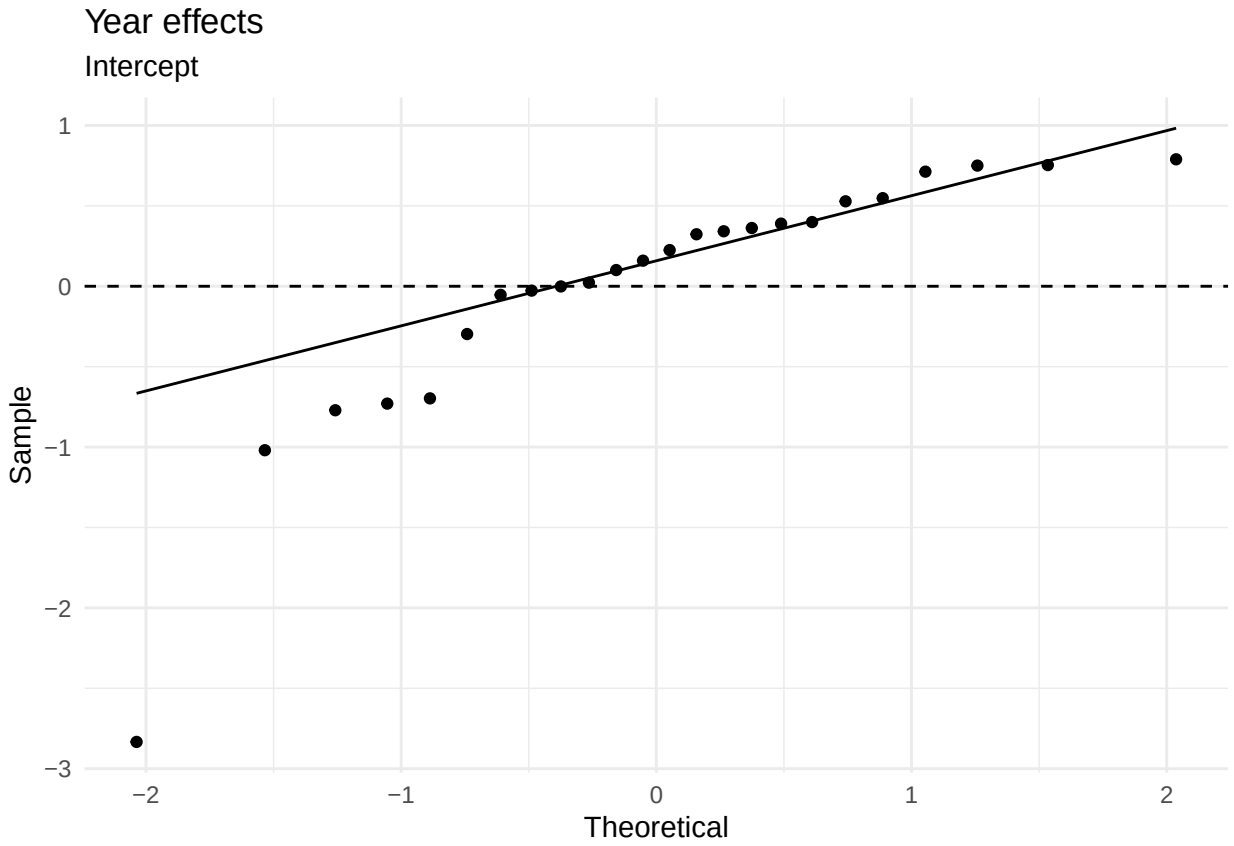
```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: y ~ dos + I(dos^2) + (dos + I(dos^2) | year)
##    Data: sr
##
##      AIC      BIC   logLik deviance df.resid
##  11082.7  11116.6  -5532.4  11064.7      309
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -23.5674  -3.5412   0.2172   3.6901  21.7396
##
## Random effects:
##  Groups Name        Variance Std.Dev. Corr
##  year   (Intercept) 0.6045   0.7775
##         dos         0.4629   0.6803    0.18
##         I(dos^2)    0.6650   0.8155   -0.71 -0.53
## Number of obs: 318, groups:  year, 24
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   1.7477     0.1593  10.974  < 2e-16 ***
## dos           1.3365     0.1398   9.558  < 2e-16 ***
## I(dos^2)     -0.8168     0.1676  -4.874 1.09e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) dos
## dos       0.177
## I(dos^2) -0.705 -0.522
```

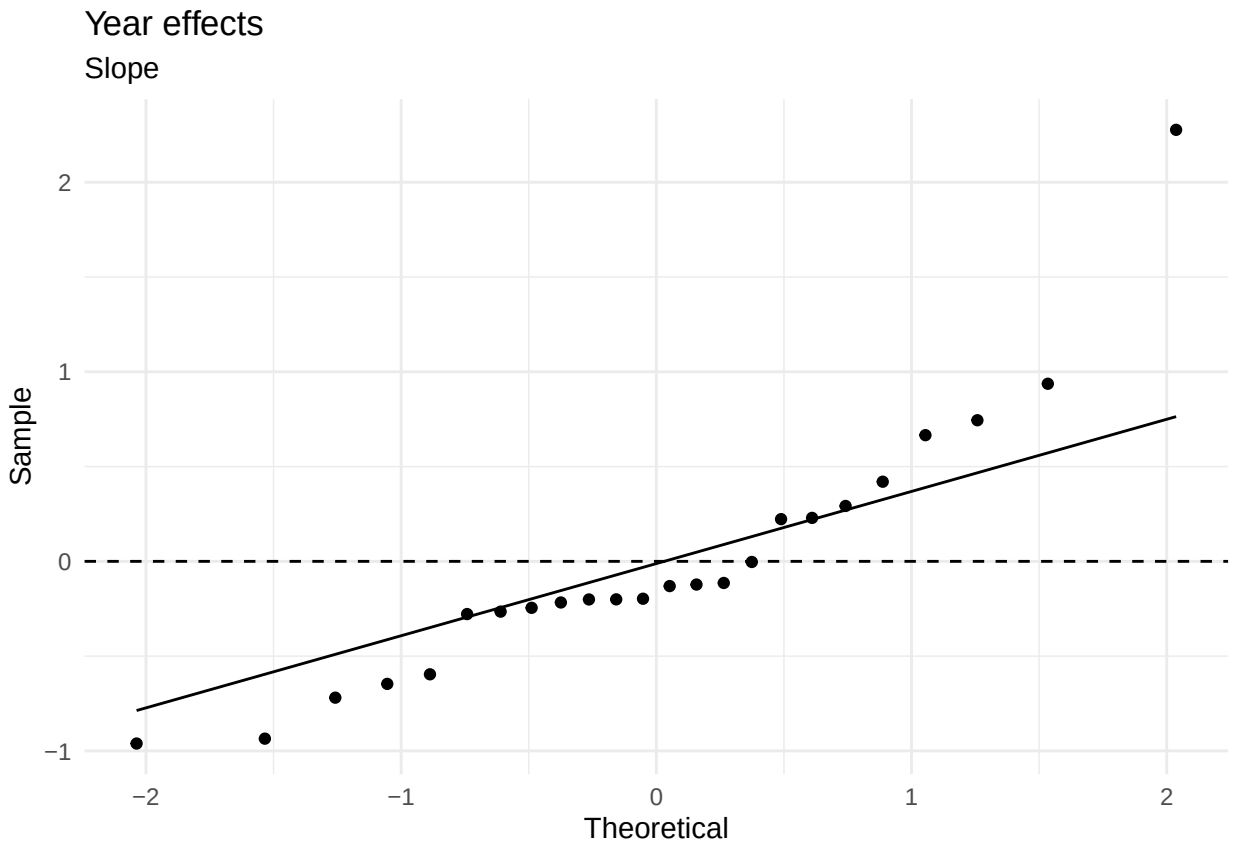
**Check assumptions of best fit model**

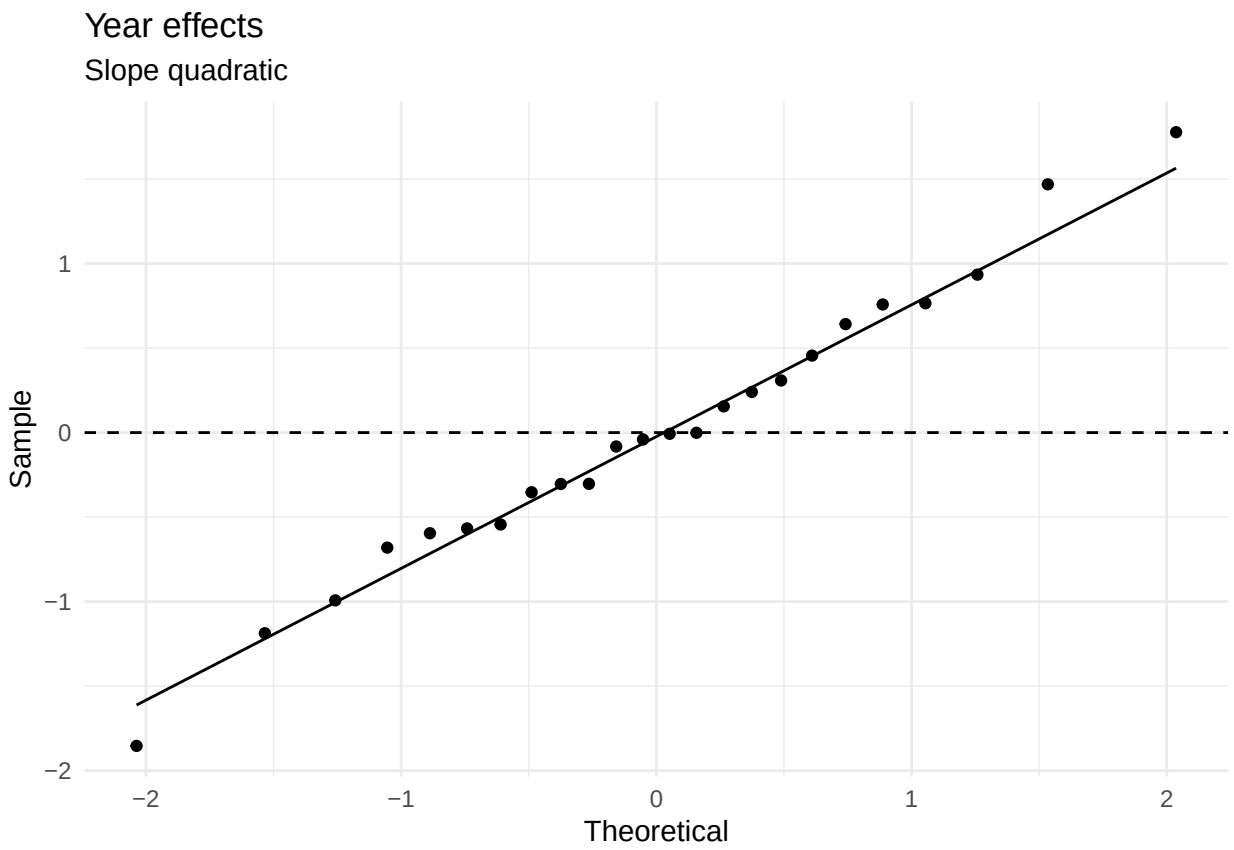



Heteroskedasticity
Fitted values vs. Residuals
Residuals
Fitted
20
10
0
−10
0.00
0.25
0.50
0.75
1.00


Quantile−Quantile
Sample
Theoretical
20
10
0
−10
−3
−2
−1
0
1
2
3


Year effects
Intercept
Sample
Theoretical
1
0
−1
−2
−3
−2
−1
0
1
2


Year effects
Slope
Sample
Theoretical
2
1
0
−1
−2
−1
0
1
2


Year effects
Slope quadratic
Sample
Theoretical
1
0
−1
−2
−2
−1
0
1
2

## Predict WESA/DUNL population

Using the derived binomial GLMM above, we will predict the amount of WESA and DUNL each day. For any years that are missing from the bGLMM we will assign the mean proportion of WESA:DUNL.