Wind Effects on Butterfly Abundance - GAMM Analysis

```r
library(tidyverse)
```

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.4     v readr     2.1.5
v forcats   1.0.0     v stringr   1.5.1
v ggplot2   3.5.2     v tibble    3.2.1
v lubridate 1.9.3     v tidyr     1.3.1
v purrr     1.0.2
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becon
```

```r
library(broom)
library(gratia)
```

```
Attaching package: 'gratia'

The following object is masked from 'package:stringr':

    boundary
```

```r
library(performance)
library(DHARMa)
```

```
This is DHARMa 0.4.7. For overview type '?DHARMa'. For recent changes, type news(package = 'I
```

```r
library(ggeffects)
library(here)
```

```
here() starts at /Users/kylenessen/Documents/Code/masters-analysis
```

```r
library(mgcv) # Load mgcv last to avoid conflicts
```

```
Loading required package: nlme

Attaching package: 'nlme'

The following object is masked from 'package:dplyr':

    collapse

This is mgcv 1.9-3. For overview type 'help("mgcv-package")'.
```

```r
theme_set(theme_minimal())
```

```r
# Load the prepared data
# Assuming df is already loaded with the structure described
# If not, load it here:
df <- read_csv(here("data", "analysis_dataset_final.csv"))
```

```
Rows: 2098 Columns: 16
-- Column specification --------------------------------------------------------
Delimiter: ","
chr   (4): deployment_id, image_filename, day_id, Observer
dbl  (10): total_butterflies, butterflies_direct_sun, temperature, view_id, ...
lgl   (1): AR_start
dttm  (1): timestamp

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
# Check data structure
glimpse(df)
```

```
Rows: 2,098
Columns: 16
$ deployment_id          <chr> "SC1", "SC1", "SC1", "SC1", "SC1", "SC1", "SC1"~
$ image_filename         <chr> "SC1_20231118063002.JPG", "SC1_20231118070001.J~
$ total_butterflies      <dbl> 56, 33, 44, 55, 51, 42, 48, 46, 46, 56, 40, 47,~
$ butterflies_direct_sun <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~
$ timestamp              <dttm> 2023-11-18 06:30:02, 2023-11-18 07:00:01, 2023~
$ day_id                 <chr> "SC1-20231118", "SC1-20231118", "SC1-20231118",~
$ AR_start               <lgl> TRUE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE,~
$ temperature            <dbl> 16, 17, 16, 17, 16, 17, 17, 17, 17, 18, 18, 17,~
$ Observer               <chr> "Skyler", "Skyler", "Skyler", "Skyler", "Skyler~
$ view_id                <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,~
$ wind_mean              <dbl> 2.34333333, 2.34333333, 2.61333333, 2.44666667,~
$ wind_max_gust          <dbl> 4.7, 4.7, 5.1, 4.7, 4.1, 4.1, 4.3, 4.3, 4.3, 4.~
$ wind_sd                <dbl> 0.3738738, 0.3738738, 0.3093189, 0.3919301, 0.3~
$ gust_differential_mean <dbl> 0.95000000, 0.95000000, 0.96000000, 0.90333333,~
$ cumulative_wind        <dbl> 70.3, 70.3, 78.4, 73.4, 66.3, 61.3, 63.8, 74.7,~
$ time_above_threshold   <dbl> 24, 24, 29, 26, 20, 14, 18, 26, 19, 28, 23, 25,~
```

```r
# Check for missing values
summary(df)
```

```
 deployment_id      image_filename     total_butterflies butterflies_direct_sun
 Length:2098        Length:2098        Min.   :  0.0     Min.   :  0.000
 Class :character   Class :character   1st Qu.:  5.0     1st Qu.:  0.000
 Mode  :character   Mode  :character   Median : 26.0     Median :  0.000
                                       Mean   : 72.5     Mean   :  4.662
                                       3rd Qu.:105.8     3rd Qu.:  1.000
                                       Max.   :770.0     Max.   :295.000
   timestamp                        day_id            AR_start
 Min.   :2023-11-18 06:30:01.00   Length:2098        Mode :logical
 1st Qu.:2023-12-21 16:57:31.75   Class :character   FALSE:1999
 Median :2024-01-03 15:05:01.00   Mode  :character   TRUE :99
 Mean   :2024-01-02 22:23:31.49
 3rd Qu.:2024-01-16 15:59:01.75
 Max.   :2024-02-03 17:30:01.00
  temperature      Observer            view_id        wind_mean
 Min.   : 3.00   Length:2098        Min.   :1.000   Min.   :0.00000
 1st Qu.:12.00   Class :character   1st Qu.:2.000   1st Qu.:0.05333
 Median :14.00   Mode  :character   Median :2.000   Median :0.64333
 Mean   :14.62                      Mean   :2.967   Mean   :0.74296
 3rd Qu.:17.00                      3rd Qu.:4.000   3rd Qu.:1.09583
```

```
 Max.   :30.00                      Max.   :5.000   Max.   :4.95000
 wind_max_gust       wind_sd        gust_differential_mean cumulative_wind
 Min.   : 0.000   Min.   :0.00000   Min.   :0.00000        Min.   :  0.00
 1st Qu.: 0.700   1st Qu.:0.05986   1st Qu.:0.04333        1st Qu.:  1.60
 Median : 1.300   Median :0.17162   Median :0.23667        Median : 19.25
 Mean   : 1.635   Mean   :0.19289   Mean   :0.29865        Mean   : 22.26
 3rd Qu.: 2.200   3rd Qu.:0.28679   3rd Qu.:0.40000        3rd Qu.: 32.88
 Max.   :12.800   Max.   :1.37730   Max.   :3.42667        Max.   :148.50
 time_above_threshold
 Min.   : 0.000
 1st Qu.: 0.000
 Median : 0.000
 Mean   : 2.131
 3rd Qu.: 0.000
 Max.   :30.000
```

```
# Check correlations among predictors
cor_matrix <- df %>%
  select(temperature, wind_mean, time_above_threshold, butterflies_direct_sun) %>%
  cor(use = "complete.obs")
print(cor_matrix)
```

```
                       temperature    wind_mean time_above_threshold
temperature             1.00000000 -0.182469624          -0.13910348
wind_mean              -0.18246962  1.000000000           0.77545421
time_above_threshold   -0.13910348  0.775454211           1.00000000
butterflies_direct_sun  0.04631443 -0.001783819           0.02046628
                       butterflies_direct_sun
temperature                       0.046314432
wind_mean                        -0.001783819
time_above_threshold              0.020466285
butterflies_direct_sun            1.000000000
```

```
# Load the data
df_full <- df

# Prepare data for modeling
# Select variables, ensure correct types, create AR.start, and handle missing values
df_model <- df_full %>%
  select(total_butterflies, temperature, wind_mean, butterflies_direct_sun, time_above_thresh
  mutate(
    day_id = as.factor(day_id),
```

```r
    Observer = as.factor(Observer)
  ) %>%
  group_by(day_id) %>%
  mutate(AR_start = row_number() == 1) %>%
  ungroup() %>%
  na.omit()

# Define the models
k_val <- 25

# Note: After changing AR_start to factor, all models need to be rerun
# to avoid errors with ggpredict()

# Model 1: Null Model
m_null <- bam(total_butterflies ~ s(day_id, bs = "re") + s(Observer, bs = "re"),
              data = df_model,
              family = tw(),
              method = "fREML",
              AR.start = df_model$AR_start)

# Model 2: Single Predictor Models
m_temp <- bam(total_butterflies ~ s(temperature, k = k_val) + s(day_id, bs = "re") + s(Obser
              data = df_model,
              family = tw(),
              method = "fREML",
              discrete = TRUE,
              AR.start = df_model$AR_start)

m_wind <- bam(total_butterflies ~ s(wind_mean, k = k_val) + s(day_id, bs = "re") + s(Observe
              data = df_model,
              family = tw(),
              method = "fREML",
              discrete = TRUE,
              AR.start = df_model$AR_start)

m_sun <- bam(total_butterflies ~ s(butterflies_direct_sun, k = k_val) + s(day_id, bs = "re")
             data = df_model,
             family = tw(),
             method = "fREML",
             discrete = TRUE,
             AR.start = df_model$AR_start)
```

```r
m_time <- bam(total_butterflies ~ s(time_above_threshold, k = k_val) + s(day_id, bs = "re") +
              data = df_model,
              family = tw(),
              method = "fREML",
              discrete = TRUE,
              AR.start = df_model$AR_start)

# Model 3: Additive Model
m_additive <- bam(total_butterflies ~ s(temperature, k = k_val) +
                                    s(wind_mean, k = k_val) +
                                    s(butterflies_direct_sun, k = k_val) +
                                    s(time_above_threshold, k = k_val) +
                                    s(day_id, bs = "re") +
                                    s(Observer, bs = "re"),
                  data = df_model,
                  family = tw(),
                  method = "fREML",
                  discrete = TRUE,
                  AR.start = df_model$AR_start)

# Model 4: Additive + Interaction Models
m_int_temp_wind <- bam(total_butterflies ~ s(temperature, k = k_val) +
                                          s(wind_mean, k = k_val) +
                                          s(butterflies_direct_sun, k = k_val) +
                                          s(time_above_threshold, k = k_val) +
                                          ti(temperature, wind_mean, k = 10) +
                                          s(day_id, bs = "re") +
                                          s(Observer, bs = "re"),
                       data = df_model,
                       family = tw(),
                       method = "fREML",
                       discrete = TRUE,
                       AR.start = df_model$AR_start)

# We can now inspect these models.
```

```r
# Model Diagnostics

# 1. Check for concurvity in the more complex models
# Concurvity is the GAM equivalent of multicollinearity.
# High values (close to 1) can be problematic.
concurvity(m_additive, full = FALSE)
```

```
$worst
                              para s(temperature) s(wind_mean)
para                      1.0000000      0.8451781    0.2916079
s(temperature)            0.8451781      1.0000000    0.2837381
s(wind_mean)              0.2916079      0.2837381    1.0000000
s(butterflies_direct_sun) 0.9195307      0.7709800    0.2765283
s(time_above_threshold)   0.9247267      0.7735880    0.9703350
s(day_id)                 1.0000000      0.8967660    0.7501991
s(Observer)               1.0000000      0.8513145    0.3950173
                          s(butterflies_direct_sun) s(time_above_threshold)
para                                      0.9195307               0.9247267
s(temperature)                            0.7709800               0.7735880
s(wind_mean)                              0.2765283               0.9703350
s(butterflies_direct_sun)                 1.0000000               0.8527999
s(time_above_threshold)                   0.8527999               1.0000000
s(day_id)                                 0.9404885               0.9631800
s(Observer)                               0.9235708               0.9267385
                          s(day_id) s(Observer)
para                      1.0000000   1.0000000
s(temperature)            0.8967660   0.8513145
s(wind_mean)              0.7501991   0.3950173
s(butterflies_direct_sun) 0.9404885   0.9235708
s(time_above_threshold)   0.9631800   0.9267385
s(day_id)                 1.0000000   1.0000000
s(Observer)               1.0000000   1.0000000

$observed
                              para s(temperature) s(wind_mean)
para                      1.0000000      0.6625146    0.1608988
s(temperature)            0.8451781      1.0000000    0.1825379
s(wind_mean)              0.2916079      0.1620679    1.0000000
s(butterflies_direct_sun) 0.9195307      0.6258758    0.1622306
s(time_above_threshold)   0.9247267      0.5999957    0.6076298
s(day_id)                 1.0000000      0.8170810    0.7057903
s(Observer)               1.0000000      0.7037214    0.2914298
                          s(butterflies_direct_sun) s(time_above_threshold)
para                                     0.08741979               0.8649132
s(temperature)                           0.08236921               0.7254332
s(wind_mean)                             0.02778797               0.5432099
s(butterflies_direct_sun)                1.00000000               0.7984062
s(time_above_threshold)                  0.08154680               1.0000000
s(day_id)                                0.18204362               0.9471478
s(Observer)                              0.09128754               0.8701670
```

```
                           s(day_id) s(Observer)
para                      2.798795e-05  0.01392749
s(temperature)            3.846065e-02  0.12301282
s(wind_mean)              4.132660e-02  0.12002814
s(butterflies_direct_sun) 1.214323e-02  0.02969339
s(time_above_threshold)   1.485889e-02  0.04370982
s(day_id)                 1.000000e+00  1.00000000
s(Observer)               2.162940e-02  1.00000000

$estimate
                               para s(temperature) s(wind_mean)
para                      1.0000000     0.22268132   0.09352529
s(temperature)            0.8451781     1.00000000   0.12099482
s(wind_mean)              0.2916079     0.07442152   1.00000000
s(butterflies_direct_sun) 0.9195307     0.24326589   0.09794363
s(time_above_threshold)   0.9247267     0.21651457   0.54964051
s(day_id)                 1.0000000     0.45860714   0.54351471
s(Observer)               1.0000000     0.27784180   0.18053080
                          s(butterflies_direct_sun) s(time_above_threshold)
para                                      0.8599215               0.7840271
s(temperature)                            0.7204348               0.6628564
s(wind_mean)                              0.2487714               0.6156513
s(butterflies_direct_sun)                 1.0000000               0.7292559
s(time_above_threshold)                   0.7963064               1.0000000
s(day_id)                                 0.8873320               0.9032166
s(Observer)                               0.8651186               0.7914298
                           s(day_id) s(Observer)
para                      0.01011404   0.2564329
s(temperature)            0.03204806   0.2659519
s(wind_mean)              0.03436996   0.1509167
s(butterflies_direct_sun) 0.02368181   0.2589513
s(time_above_threshold)   0.02668857   0.2548050
s(day_id)                 1.00000000   1.0000000
s(Observer)               0.04055501   1.0000000
```

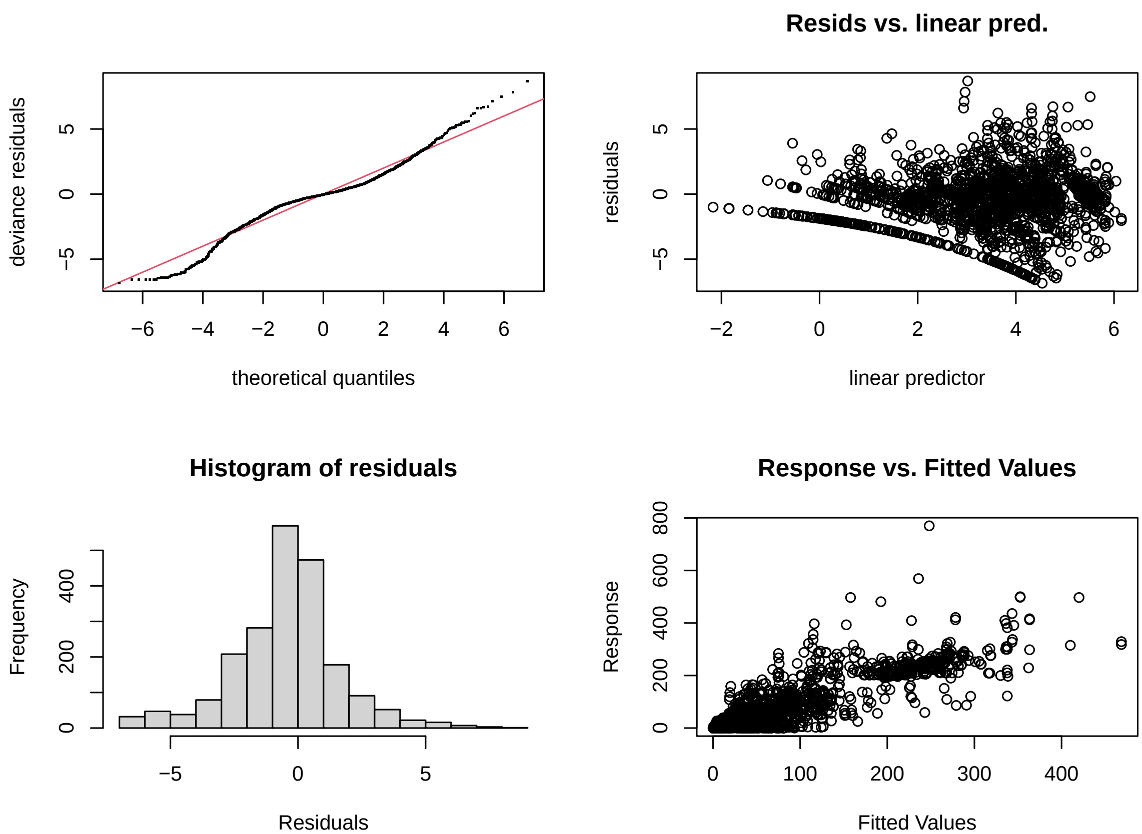
```r
concurvity(m_int_temp_wind, full = FALSE)
```

```
$worst
                               para s(temperature) s(wind_mean)
para                      1.0000000      0.8451781    0.2916079
s(temperature)            0.8451781      1.0000000    0.2837381
s(wind_mean)              0.2916079      0.2837381    1.0000000
```

```
s(butterflies_direct_sun) 0.9195307      0.7709800    0.2765283
s(time_above_threshold)   0.9247267      0.7735880    0.9703350
ti(temperature,wind_mean) 0.4868951      0.9999986    1.0000000
s(day_id)                 1.0000000      0.8967660    0.7501991
s(Observer)               1.0000000      0.8513145    0.3950173
                          s(butterflies_direct_sun) s(time_above_threshold)
para                                      0.9195307               0.9247267
s(temperature)                            0.7709800               0.7735880
s(wind_mean)                              0.2765283               0.9703350
s(butterflies_direct_sun)                 1.0000000               0.8527999
s(time_above_threshold)                   0.8527999               1.0000000
ti(temperature,wind_mean)                 0.4644367               0.9118858
s(day_id)                                 0.9404885               0.9631800
s(Observer)                               0.9235708               0.9267385
                          ti(temperature,wind_mean) s(day_id) s(Observer)
para                                      0.4868951 1.0000000   1.0000000
s(temperature)                            0.9999986 0.8967660   0.8513145
s(wind_mean)                              1.0000000 0.7501991   0.3950173
s(butterflies_direct_sun)                 0.4644367 0.9404885   0.9235708
s(time_above_threshold)                   0.9118858 0.9631800   0.9267385
ti(temperature,wind_mean)                 1.0000000 0.7537288   0.5476079
s(day_id)                                 0.7537288 1.0000000   1.0000000
s(Observer)                               0.5476079 1.0000000   1.0000000

$observed
                               para s(temperature) s(wind_mean)
para                      1.0000000      0.6831604    0.1609006
s(temperature)            0.8451781      1.0000000    0.1825403
s(wind_mean)              0.2916079      0.1704362    1.0000000
s(butterflies_direct_sun) 0.9195307      0.6420576    0.1622323
s(time_above_threshold)   0.9247267      0.6194861    0.6075949
ti(temperature,wind_mean) 0.4868951      0.4431387    0.9743921
s(day_id)                 1.0000000      0.8269818    0.7058021
s(Observer)               1.0000000      0.7201849    0.2914433
                          s(butterflies_direct_sun) s(time_above_threshold)
para                                     0.02419883               0.8698967
s(temperature)                           0.03000484               0.7291871
s(wind_mean)                             0.01166099               0.5363753
s(butterflies_direct_sun)                1.00000000               0.8025345
s(time_above_threshold)                  0.02492866               1.0000000
ti(temperature,wind_mean)                0.03900198               0.6336190
s(day_id)                                0.13043146               0.9489516
s(Observer)                              0.02942697               0.8749344
```

```
                          ti(temperature,wind_mean)    s(day_id) s(Observer)
para                                    0.006781943 2.808914e-05  0.01418196
s(temperature)                          0.203115104 4.000917e-02  0.12263097
s(wind_mean)                            0.402237745 4.148478e-02  0.11835026
s(butterflies_direct_sun)               0.032004435 1.227514e-02  0.03013886
s(time_above_threshold)                 0.275812747 1.513723e-02  0.04393100
ti(temperature,wind_mean)               1.000000000 8.832540e-02  0.22388916
s(day_id)                               0.359894238 1.000000e+00  1.00000000
s(Observer)                             0.014382665 2.228416e-02  1.00000000

$estimate
                               para s(temperature) s(wind_mean)
para                      1.0000000     0.22268132   0.09352529
s(temperature)            0.8451781     1.00000000   0.12099482
s(wind_mean)              0.2916079     0.07442152   1.00000000
s(butterflies_direct_sun) 0.9195307     0.24326589   0.09794363
s(time_above_threshold)   0.9247267     0.21651457   0.54964051
ti(temperature,wind_mean) 0.4868951     0.55367547   0.96868862
s(day_id)                 1.0000000     0.45860714   0.54351471
s(Observer)               1.0000000     0.27784180   0.18053080
                          s(butterflies_direct_sun) s(time_above_threshold)
para                                      0.8599215               0.7840271
s(temperature)                            0.7204348               0.6628564
s(wind_mean)                              0.2487714               0.6156513
s(butterflies_direct_sun)                 1.0000000               0.7292559
s(time_above_threshold)                   0.7963064               1.0000000
ti(temperature,wind_mean)                 0.4239859               0.6898828
s(day_id)                                 0.8873320               0.9032166
s(Observer)                               0.8651186               0.7914298
                          ti(temperature,wind_mean)  s(day_id) s(Observer)
para                                    0.009631861 0.01011404   0.2564329
s(temperature)                          0.098035495 0.03204806   0.2659519
s(wind_mean)                            0.191446443 0.03436996   0.1509167
s(butterflies_direct_sun)               0.022811273 0.02368181   0.2589513
s(time_above_threshold)                 0.059548869 0.02668857   0.2548050
ti(temperature,wind_mean)               1.000000000 0.09493085   0.2503535
s(day_id)                               0.165360791 1.00000000   1.0000000
s(Observer)                             0.020540210 0.04055501   1.0000000
```

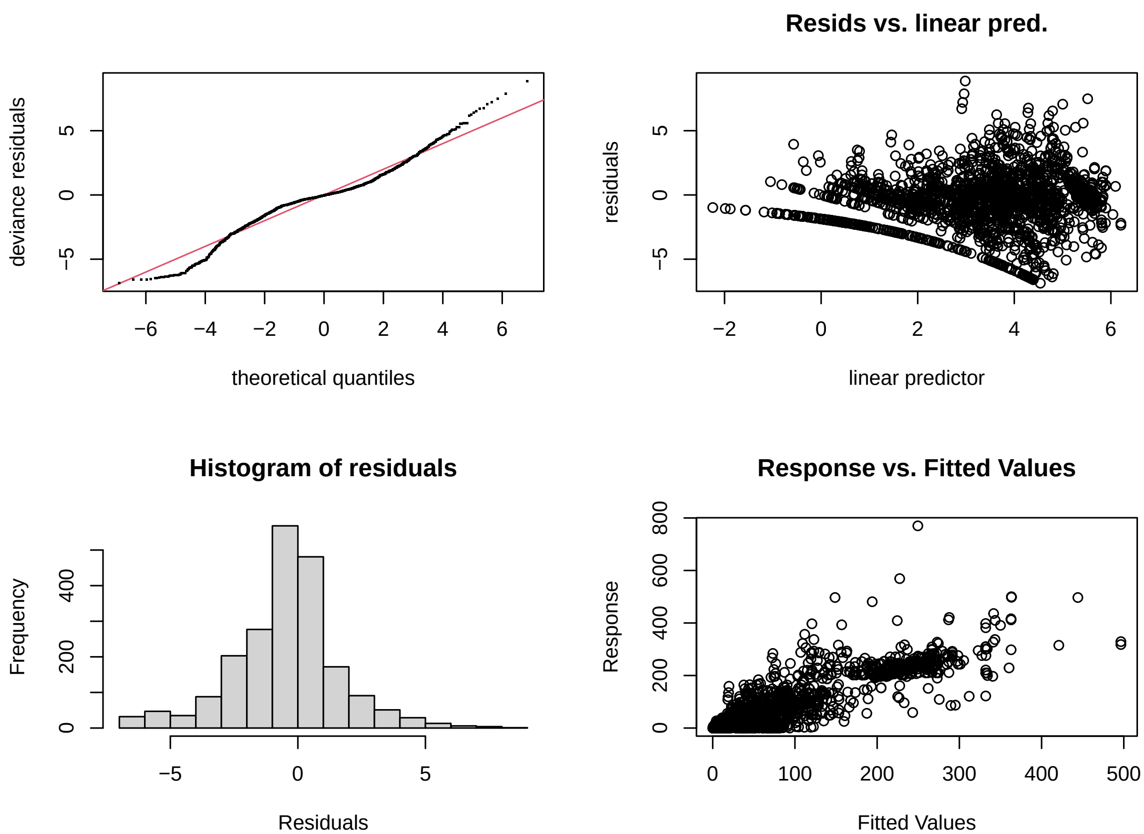
```
# 2. Use gam.check() for standard diagnostics - Additive Model
# This provides k-checks (are basis dimensions adequate?) and residual plots.
gam.check(m_additive)
```




```
Method: fREML   Optimizer: perf chol
$grad
[1] -2.203159e-07 -1.194440e-05 -5.990452e-08 -1.250850e-08  3.399407e-07
[6]  3.575220e-09

$hess
              [,1]          [,2]          [,3]          [,4]          [,5]
[1,]  1.409644e+00 -1.254825e-06 -3.264924e-02 -1.341775e-02  7.719297e-02
[2,] -1.254825e-06  1.194438e-05 -1.393144e-07 -1.814586e-06  1.365336e-06
[3,] -3.264924e-02 -1.393144e-07  1.316937e-01 -4.605060e-03 -4.015984e-02
[4,] -1.341775e-02 -1.814586e-06 -4.605060e-03  4.476653e-01  6.971215e-02
[5,]  7.719297e-02  1.365336e-06 -4.015984e-02  6.971215e-02  4.417628e+01
[6,]  5.439477e-03  1.378624e-07 -2.822755e-04 -9.333281e-04  2.574958e-01
              [,6]
[1,]  5.439477e-03
[2,]  1.378624e-07
```

```
[3,] -2.822755e-04
[4,] -9.333281e-04
[5,]  2.574958e-01
[6,]  1.368696e-01

Model rank =  200 / 200

Basis dimension (k) checking results. Low p-value (k-index<1) may
indicate that k is too low, especially if edf is close to k'.

                              k'    edf k-index p-value
s(temperature)            24.000 11.644    0.75  <2e-16 ***
s(wind_mean)              24.000  1.000    0.77  <2e-16 ***
s(butterflies_direct_sun) 24.000  1.819    0.40  <2e-16 ***
s(time_above_threshold)   24.000  2.327    0.28  <2e-16 ***
s(day_id)                 99.000 95.244      NA      NA
s(Observer)                4.000  0.883      NA      NA
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# gam.check() for interaction model
gam.check(m_int_temp_wind)
```



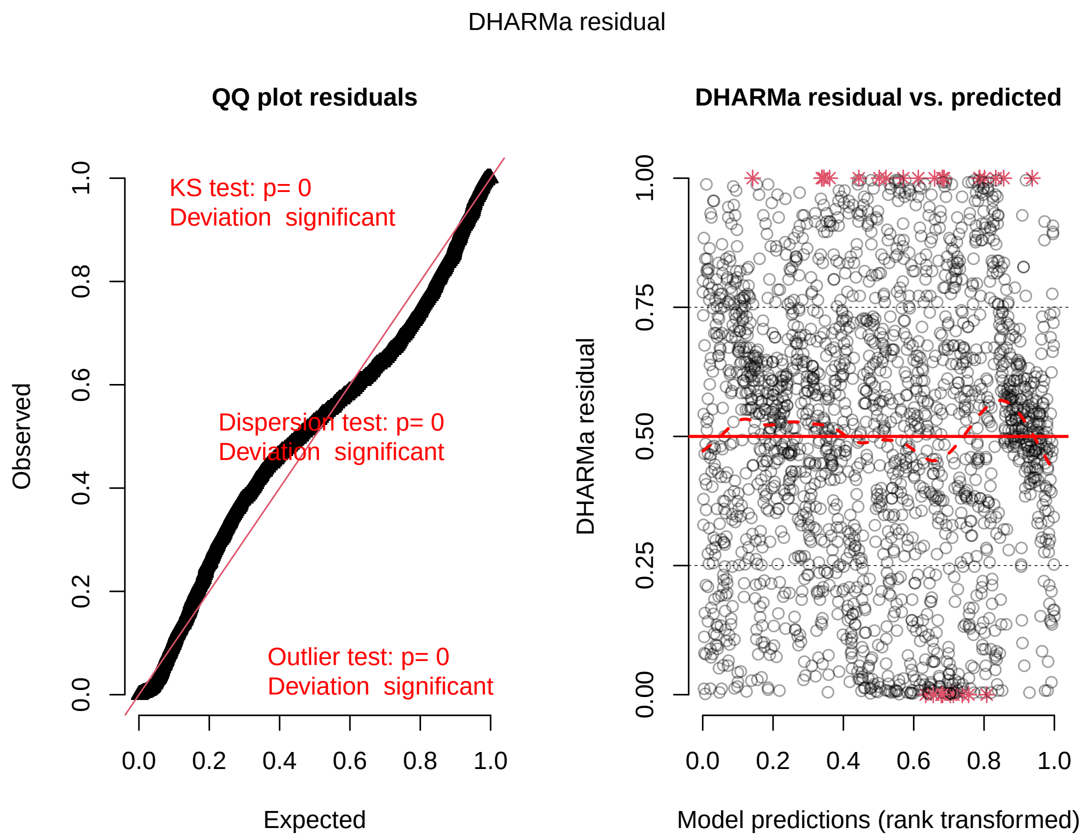

```
Method: fREML   Optimizer: perf chol
$grad
[1] -3.073437e-08 -1.418299e-05  1.264180e-08 -6.558273e-09 -6.572156e-09
[6] -3.030263e-09 -6.605788e-07 -5.819831e-09

$hess
              [,1]          [,2]          [,3]          [,4]          [,5]
[1,]  1.402411e+00 -2.180554e-07 -3.355815e-02 -1.459239e-02  6.322646e-02
[2,] -2.180554e-07  1.418276e-05 -1.669192e-09 -7.459535e-07 -1.152722e-08
[3,] -3.355815e-02 -1.669192e-09  1.406051e-01 -4.129179e-03 -1.701181e-03
[4,] -1.459239e-02 -7.459535e-07 -4.129179e-03  4.554739e-01  5.964487e-03
[5,]  6.322646e-02 -1.152722e-08 -1.701181e-03  5.964487e-03  1.975294e-01
[6,]  2.225232e-02 -1.479048e-06 -6.636009e-04  1.216904e-02  2.976199e-02
[7,]  7.010139e-02  4.395933e-07 -4.399687e-02  7.191106e-02  2.644437e-02
[8,]  5.101200e-03  3.197216e-08 -2.057891e-04 -1.095167e-03  2.546720e-03
              [,6]          [,7]          [,8]
```

```
[1,]  2.225232e-02  7.010139e-02  5.101200e-03
[2,] -1.479048e-06  4.395933e-07  3.197216e-08
[3,] -6.636009e-04 -4.399687e-02 -2.057891e-04
[4,]  1.216904e-02  7.191106e-02 -1.095167e-03
[5,]  2.976199e-02  2.644437e-02  2.546720e-03
[6,]  8.507487e-02  1.638532e-02  2.405023e-04
[7,]  1.638532e-02  4.421931e+01  2.489754e-01
[8,]  2.405023e-04  2.489754e-01  1.183717e-01

Model rank =  281 / 281

Basis dimension (k) checking results. Low p-value (k-index<1) may
indicate that k is too low, especially if edf is close to k'.

                              k'    edf k-index p-value
s(temperature)            24.000 11.833    0.76  <2e-16 ***
s(wind_mean)              24.000  1.000    0.77  <2e-16 ***
s(butterflies_direct_sun) 24.000  1.968    0.40  <2e-16 ***
s(time_above_threshold)   24.000  2.262    0.28  <2e-16 ***
ti(temperature,wind_mean) 81.000  3.218    0.90  <2e-16 ***
s(day_id)                 99.000 95.306      NA      NA
s(Observer)                4.000  0.822      NA      NA
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

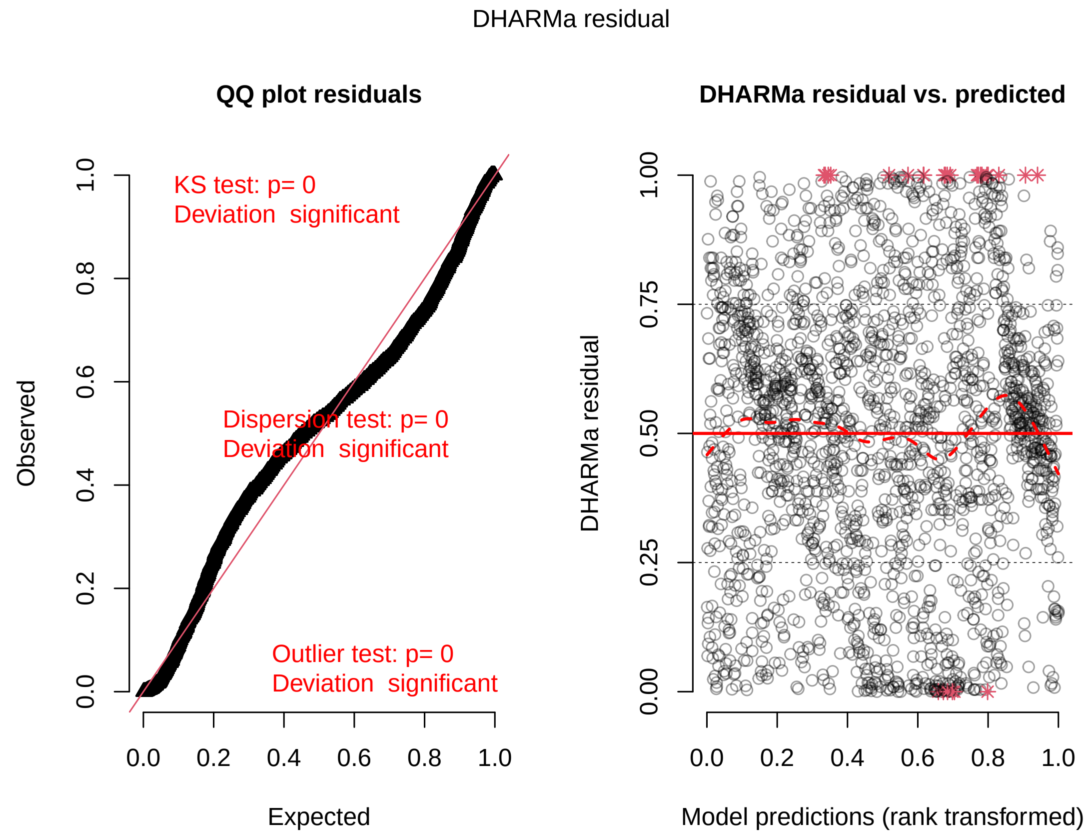
```r
# 3. Use DHARMa for more advanced residual diagnostics - Additive Model
# This simulates residuals from the fitted model and compares them to the observed residuals
sim_res_additive <- simulateResiduals(fittedModel = m_additive, n = 250)
```

```
Registered S3 methods overwritten by 'mgcViz':
  method       from
  +.gg         ggplot2
  simulate.gam gratia
```

```r
plot(sim_res_additive)
```




```r
# DHARMa diagnostics for interaction model
sim_res_int <- simulateResiduals(fittedModel = m_int_temp_wind, n = 250)
plot(sim_res_int)
```

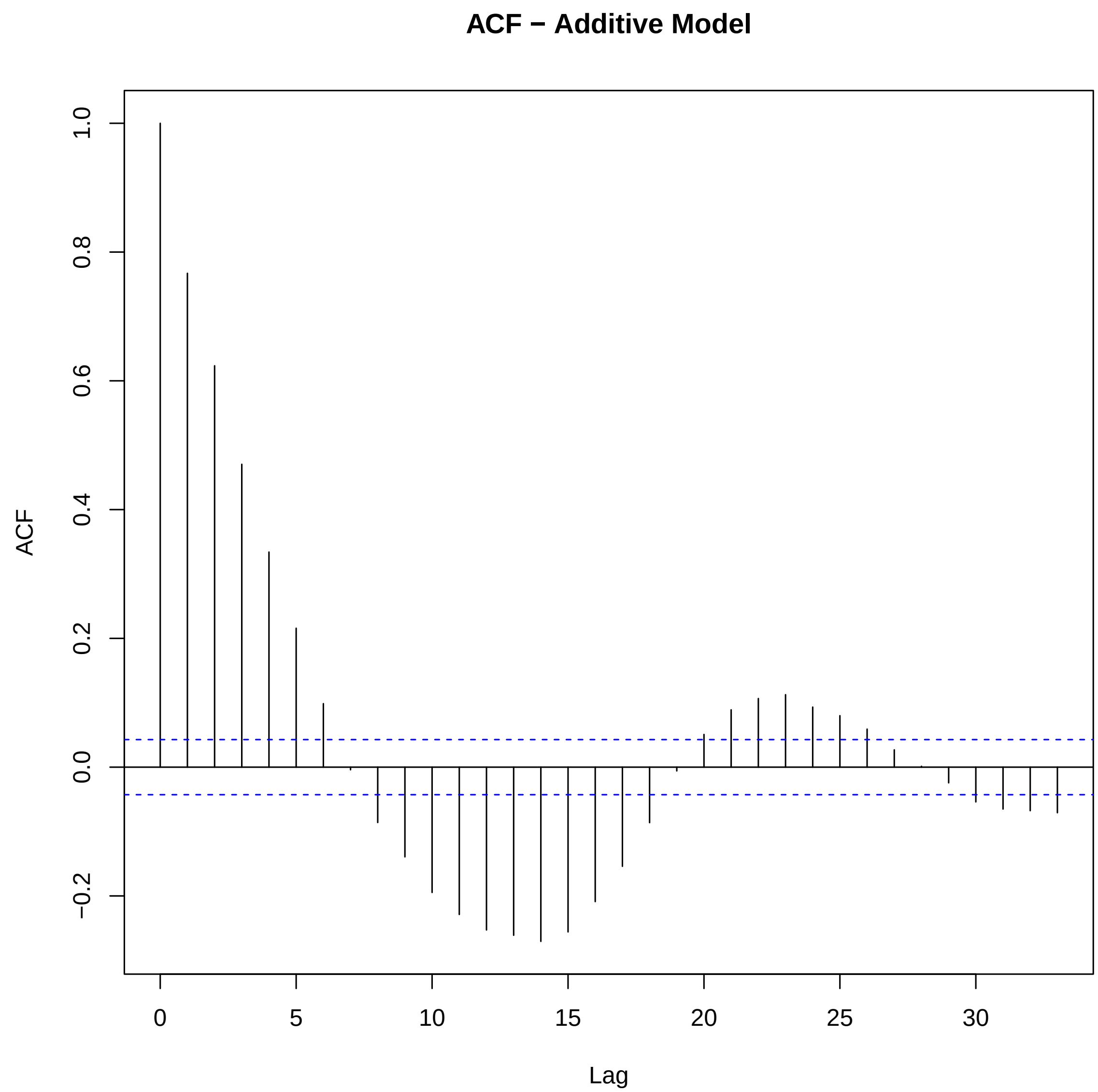



```
# 4. Check for temporal autocorrelation using ACF plots
# This is crucial for time series data to ensure residuals are not autocorrelated

library(gridExtra)
```

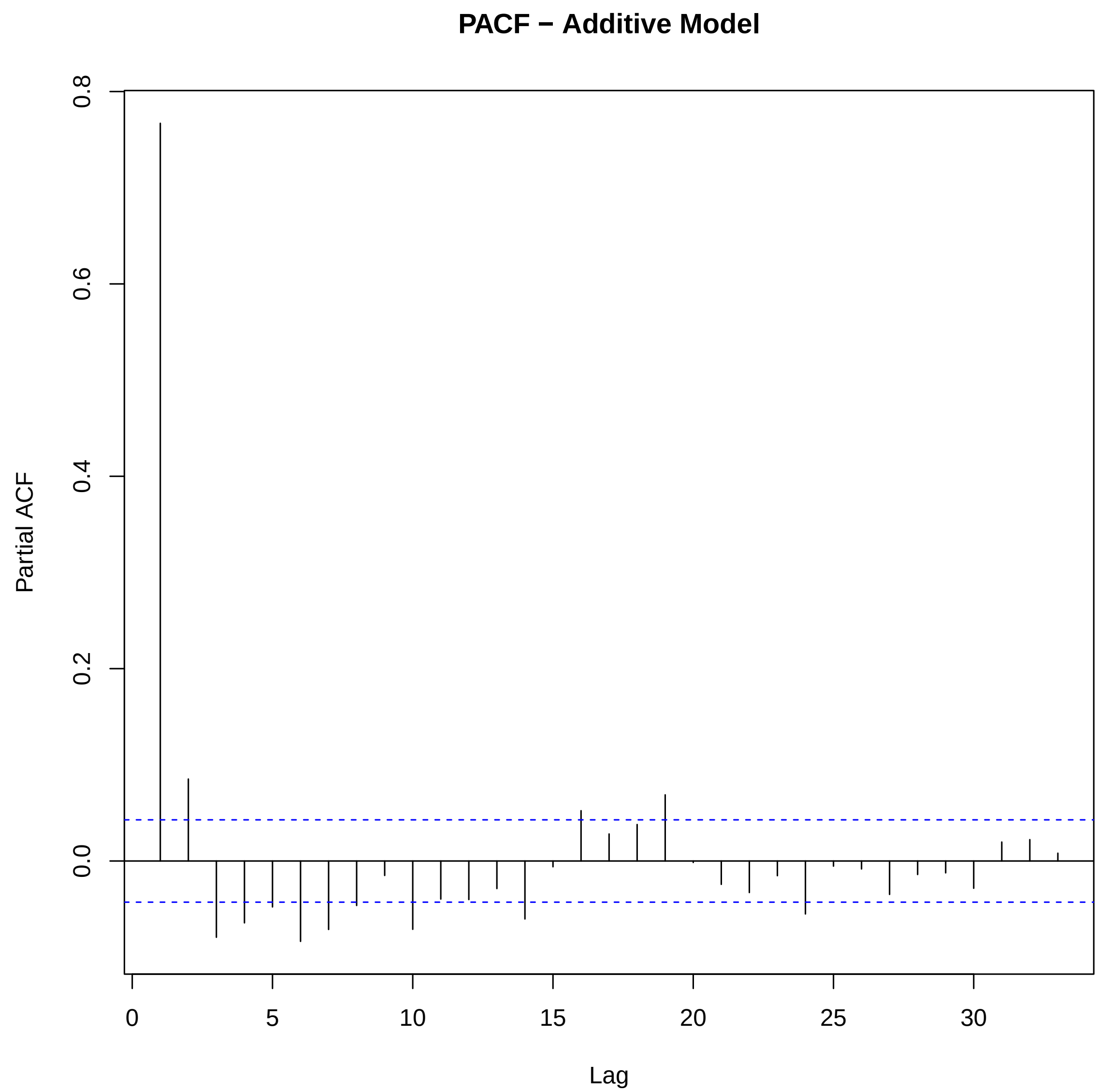
```
Attaching package: 'gridExtra'

The following object is masked from 'package:dplyr':

    combine
```

```
library(grid)

# Create ACF plots using base R but arrange with grid
# ACF for additive model
```

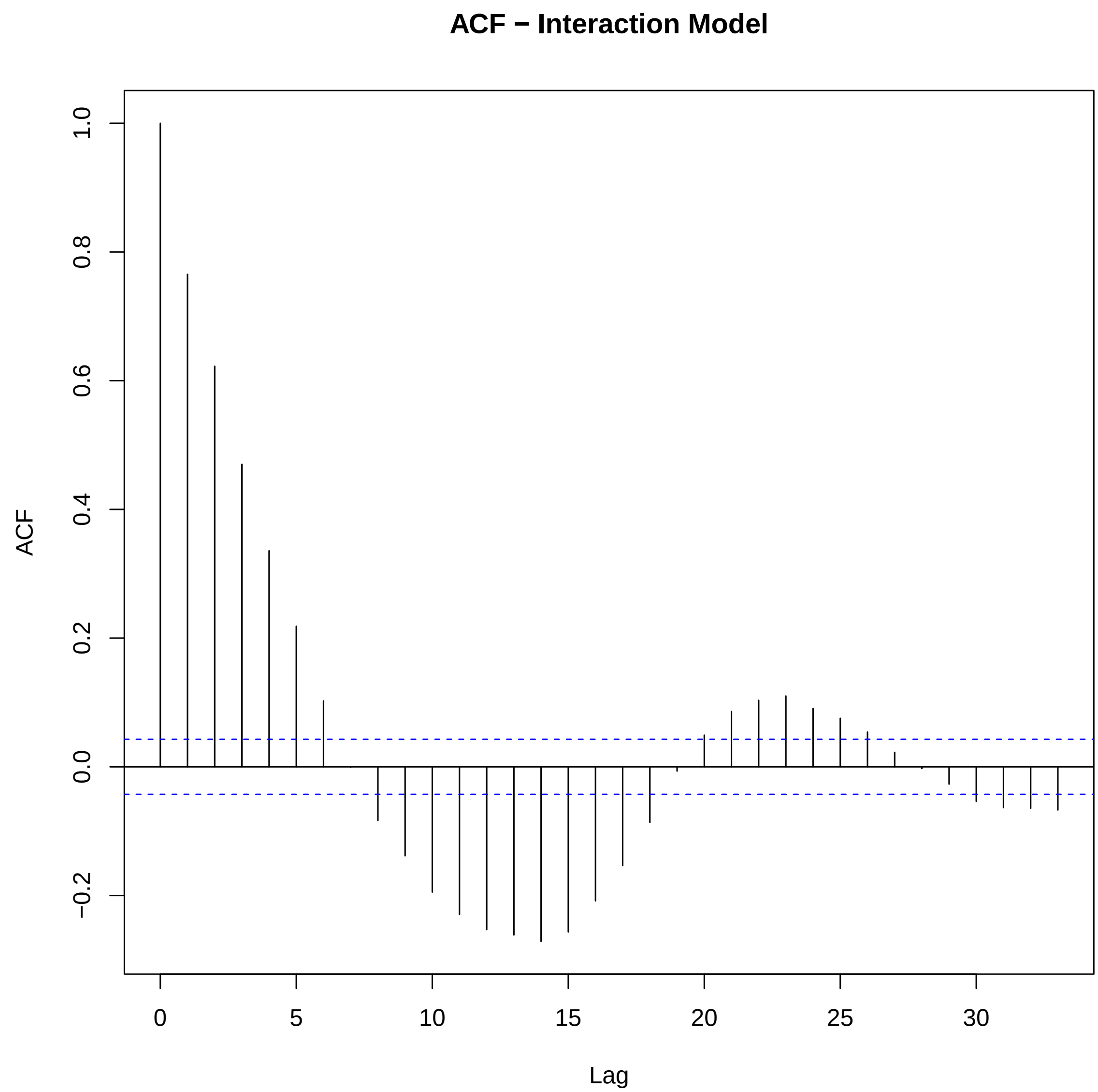
```r
acf_add <- acf(residuals(m_additive), plot = FALSE, main = "ACF - Additive Model")
pacf_add <- pacf(residuals(m_additive), plot = FALSE, main = "PACF - Additive Model")

# ACF for interaction model
acf_int <- acf(residuals(m_int_temp_wind), plot = FALSE, main = "ACF - Interaction Model")
pacf_int <- pacf(residuals(m_int_temp_wind), plot = FALSE, main = "PACF - Interaction Model")

# Plot ACF results
plot(acf_add, main = "ACF - Additive Model")
```



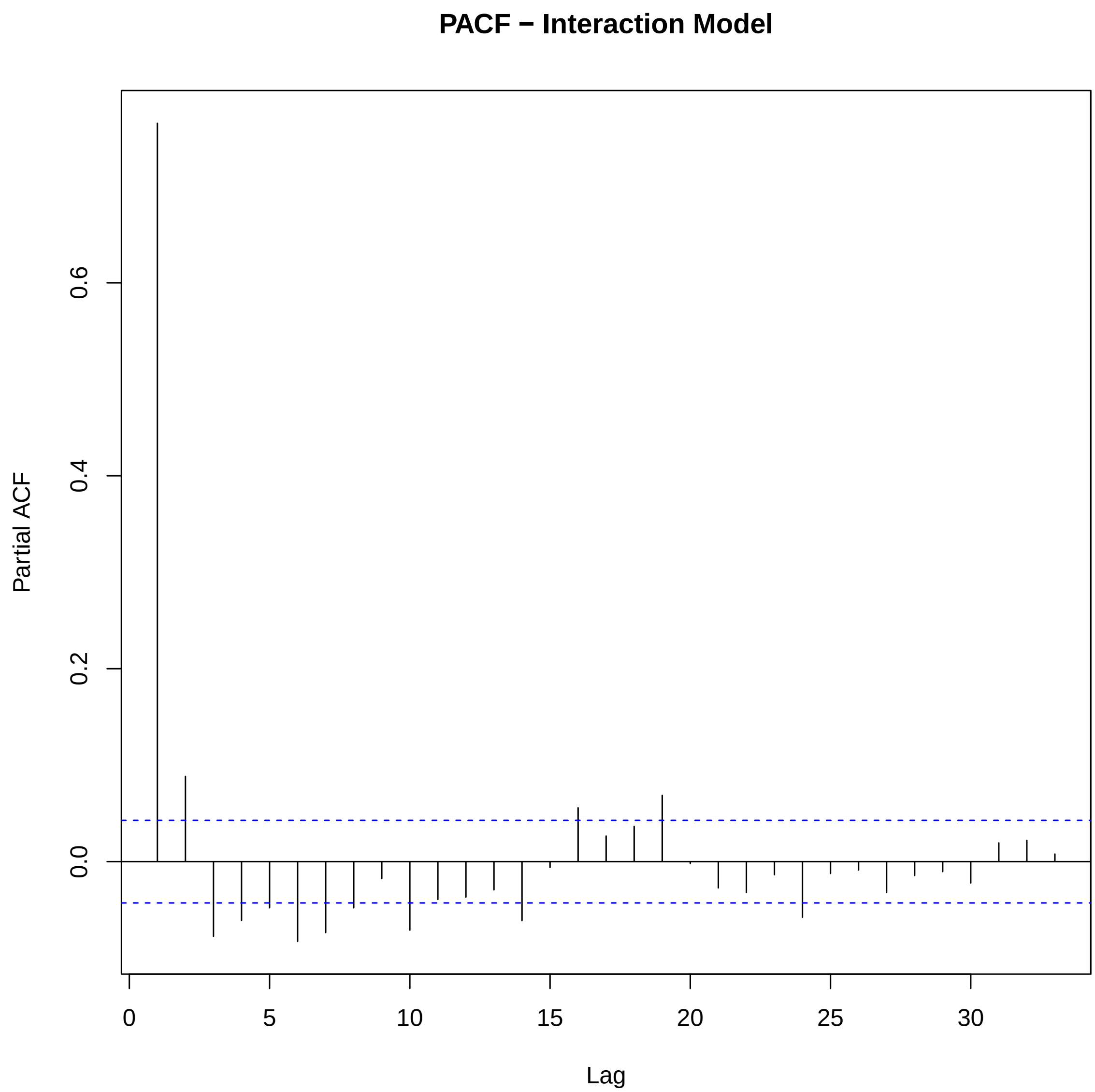

```
plot(pacf_add, main = "PACF - Additive Model")
```




```r
plot(acf_int, main = "ACF - Interaction Model")
```




```
plot(pacf_int, main = "PACF - Interaction Model")
```

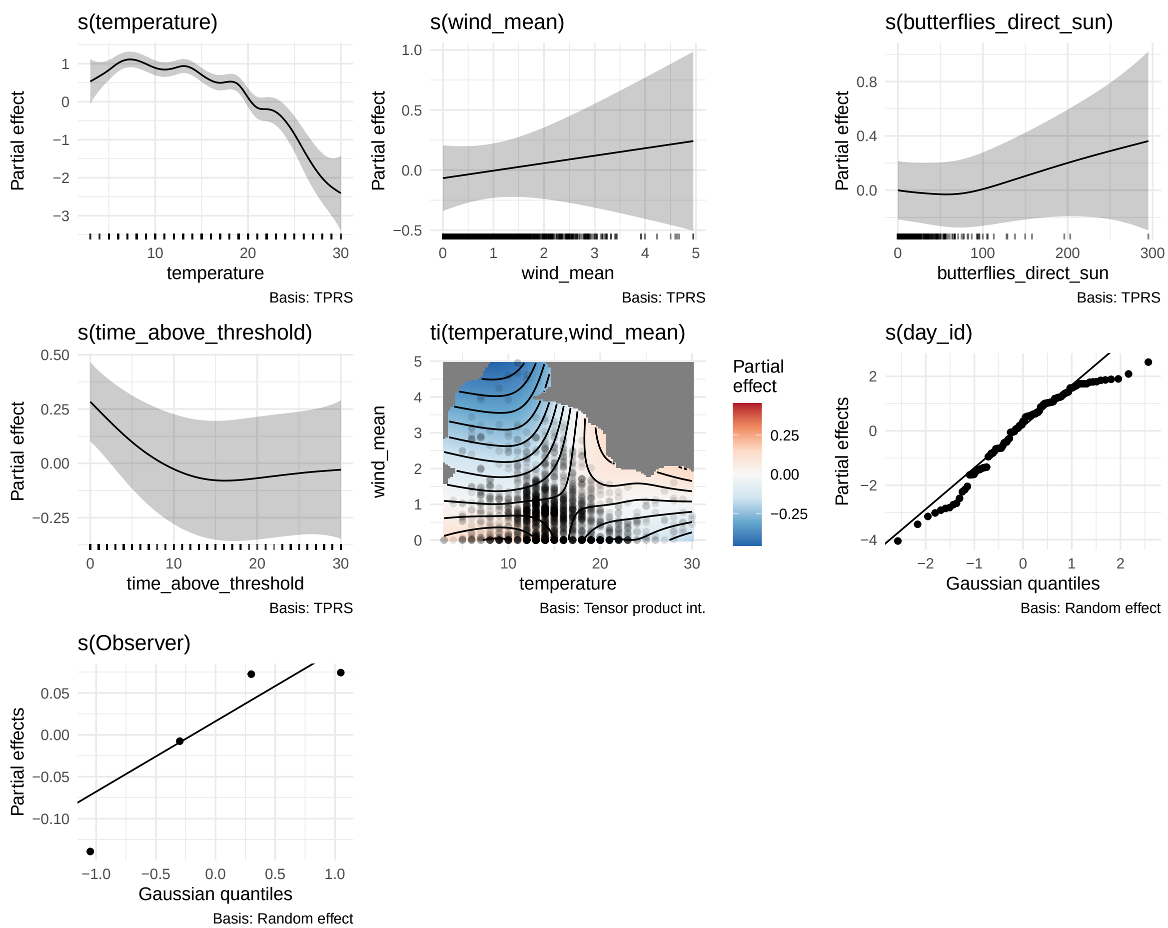



```
# 5. Compare models using AIC
# Create a list of models
model_list <- list(
  null = m_null,
  temp = m_temp,
  wind = m_wind,
  sun = m_sun,
  time = m_time,
  additive = m_additive,
```

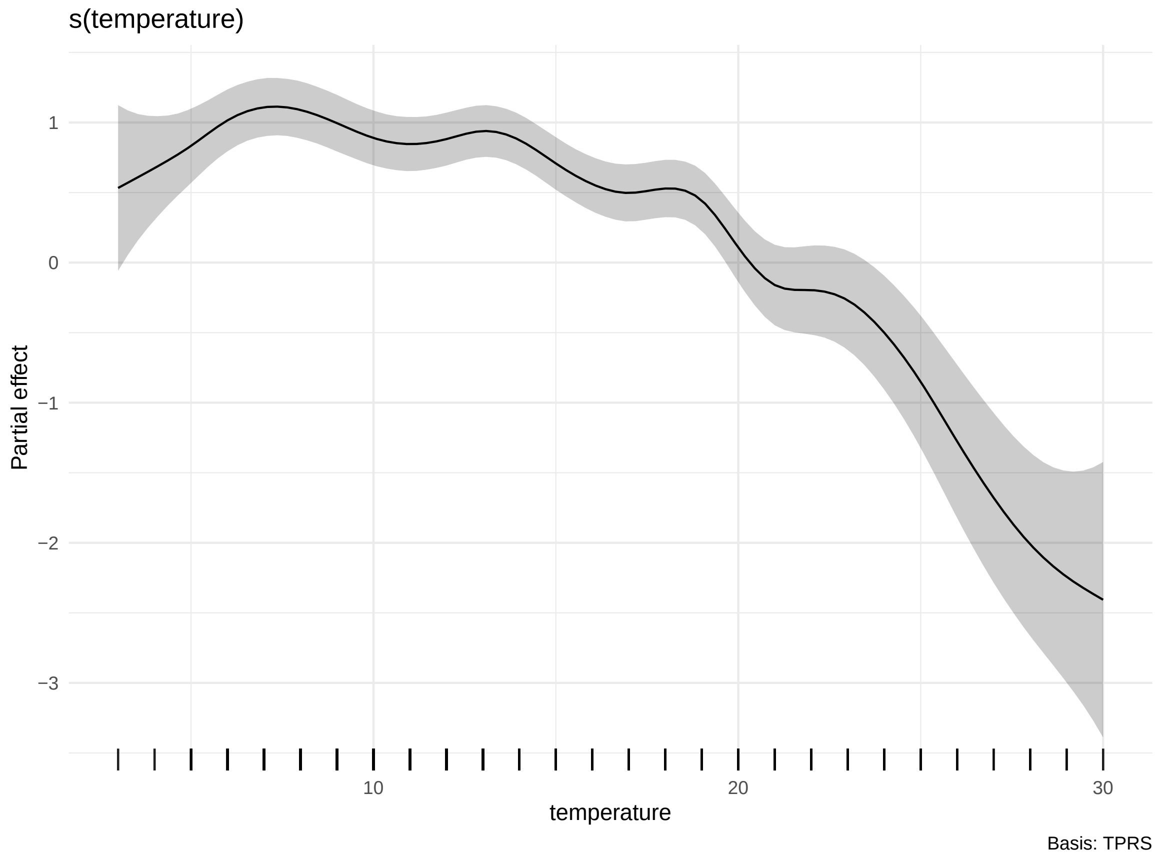
```r
  interaction = m_int_temp_wind
)

# Get AIC for each model
aic_values <- sapply(model_list, AIC)

# Create a summary table
aic_table <- tibble(
  model = names(aic_values),
  AIC = aic_values
) %>%
  arrange(AIC)

print(aic_table)
```

```
# A tibble: 7 x 2
  model          AIC
  <chr>        <dbl>
1 interaction 18090.
2 additive    18092.
3 temp        18107.
4 sun         18378.
5 time        18431.
6 wind        18438.
7 null        18438.
```

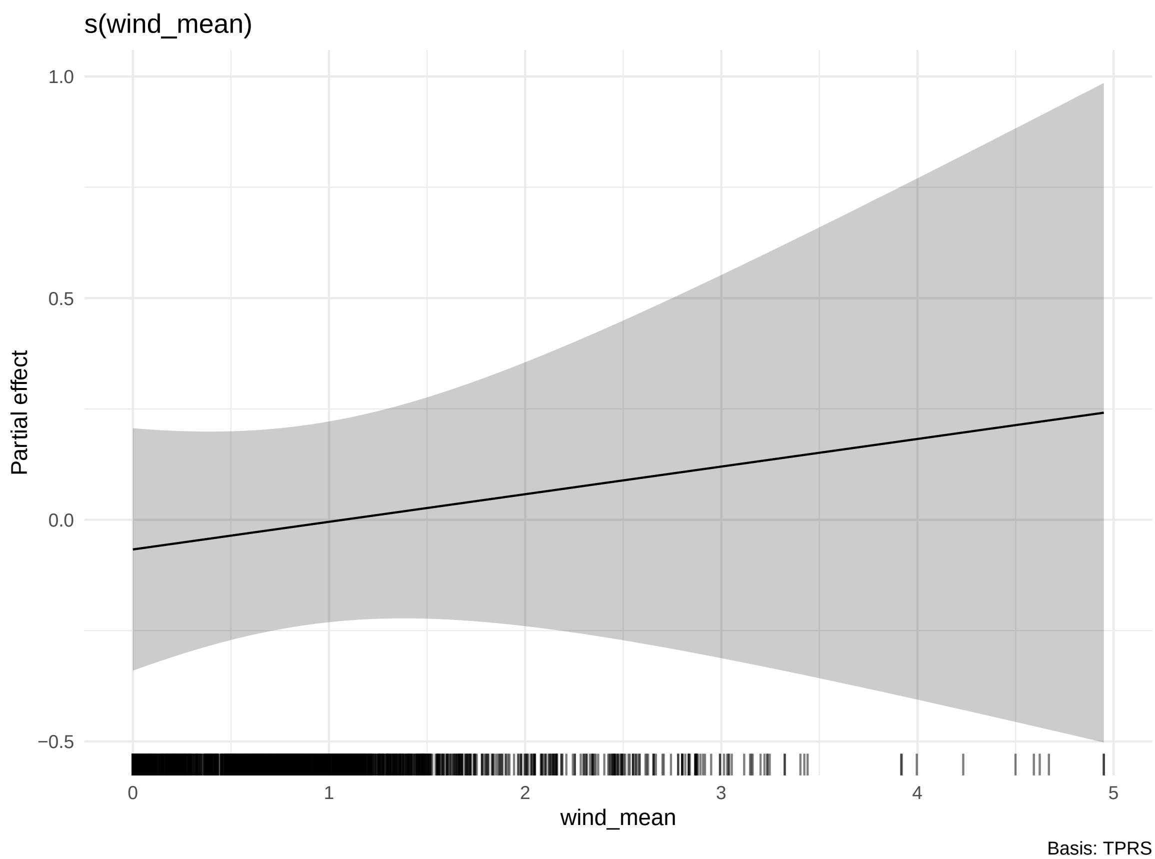
```r
# --- Plotting the Best Model ---

# The model with the lowest AIC is `m_int_temp_wind`.
# Let's visualize the effects from this model.

# 1. Plot all smooth terms (main effects and interactions) together
# `gratia::draw()` is excellent for this. `scales = "free"` allows each plot
# to have its own y-axis scale.
draw(m_int_temp_wind, scales = "free")
```



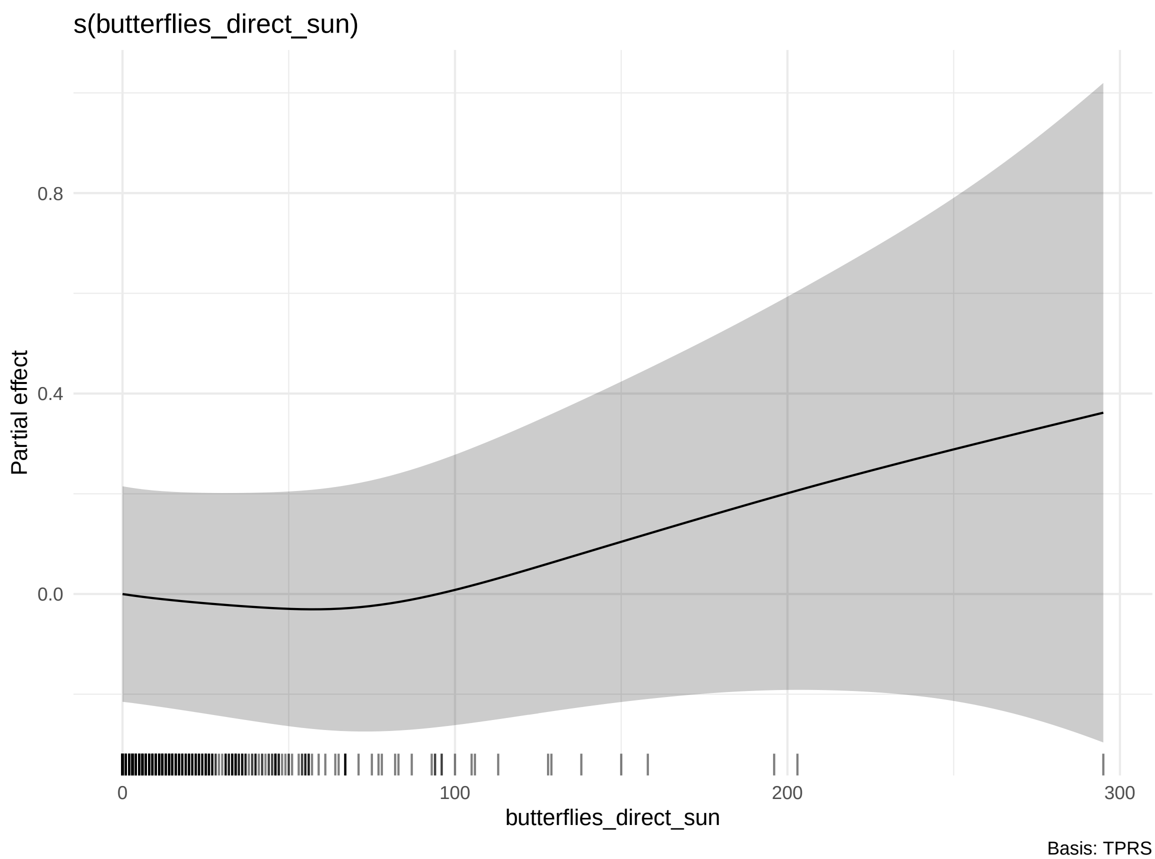

```
# 2. Plot individual effects for more detail

# Main effect of Temperature
draw(m_int_temp_wind, select = "s(temperature)")
```



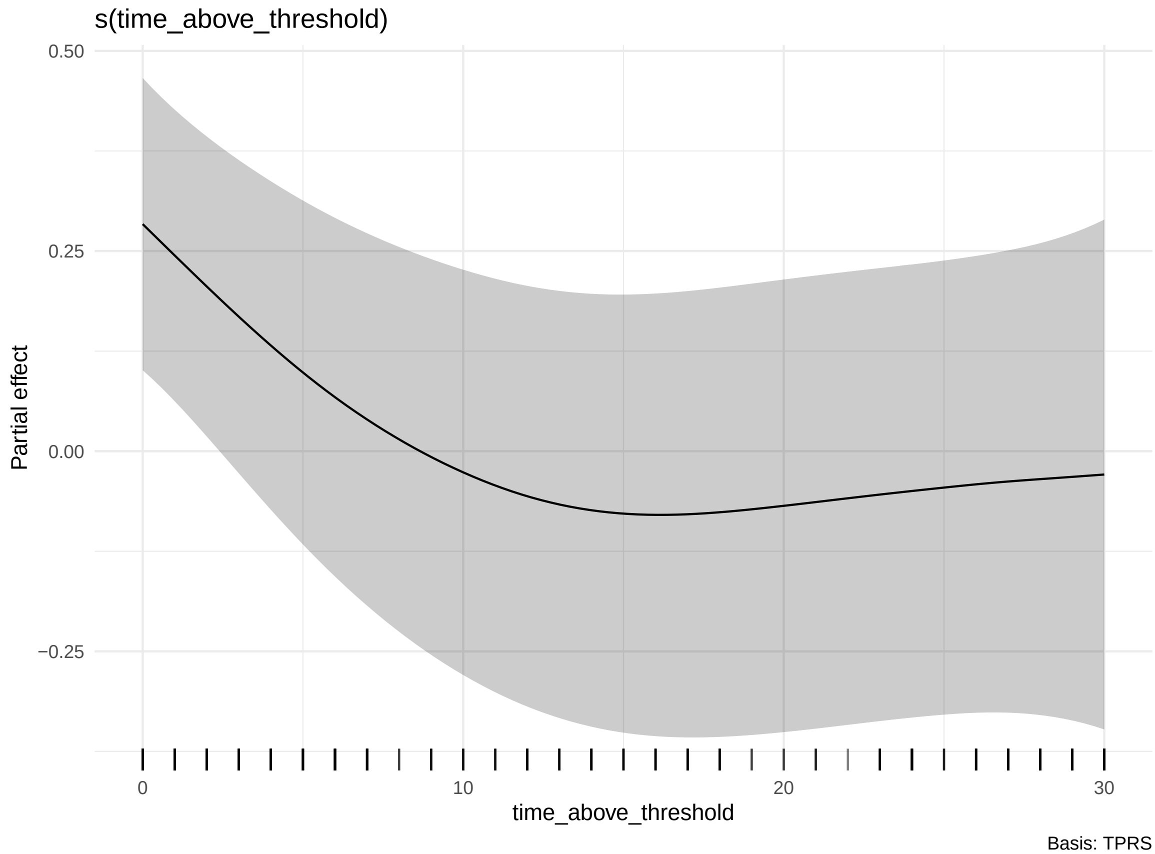

```r
# Main effect of Wind
draw(m_int_temp_wind, select = "s(wind_mean)")
```



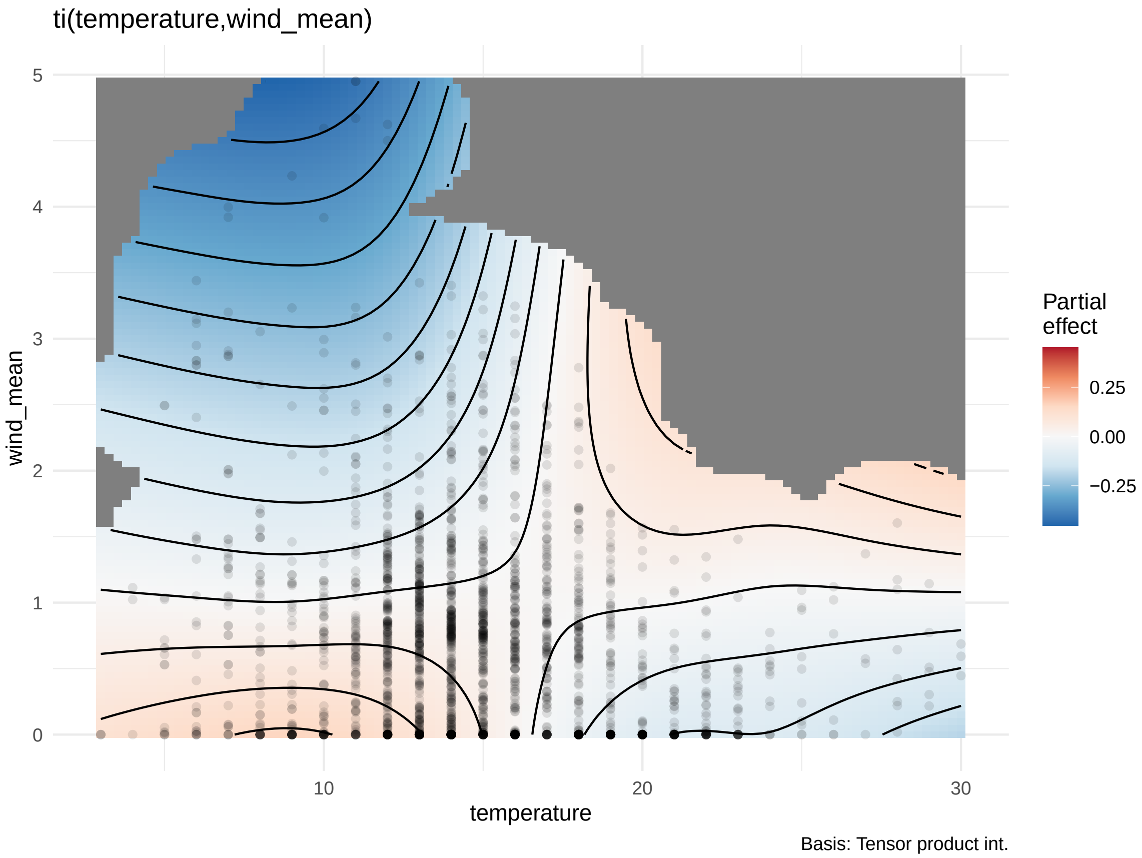

```r
# Main effect of Sun
draw(m_int_temp_wind, select = "s(butterflies_direct_sun)")
```




```r
# Main effect of Time Above Threshold
draw(m_int_temp_wind, select = "s(time_above_threshold)")
```

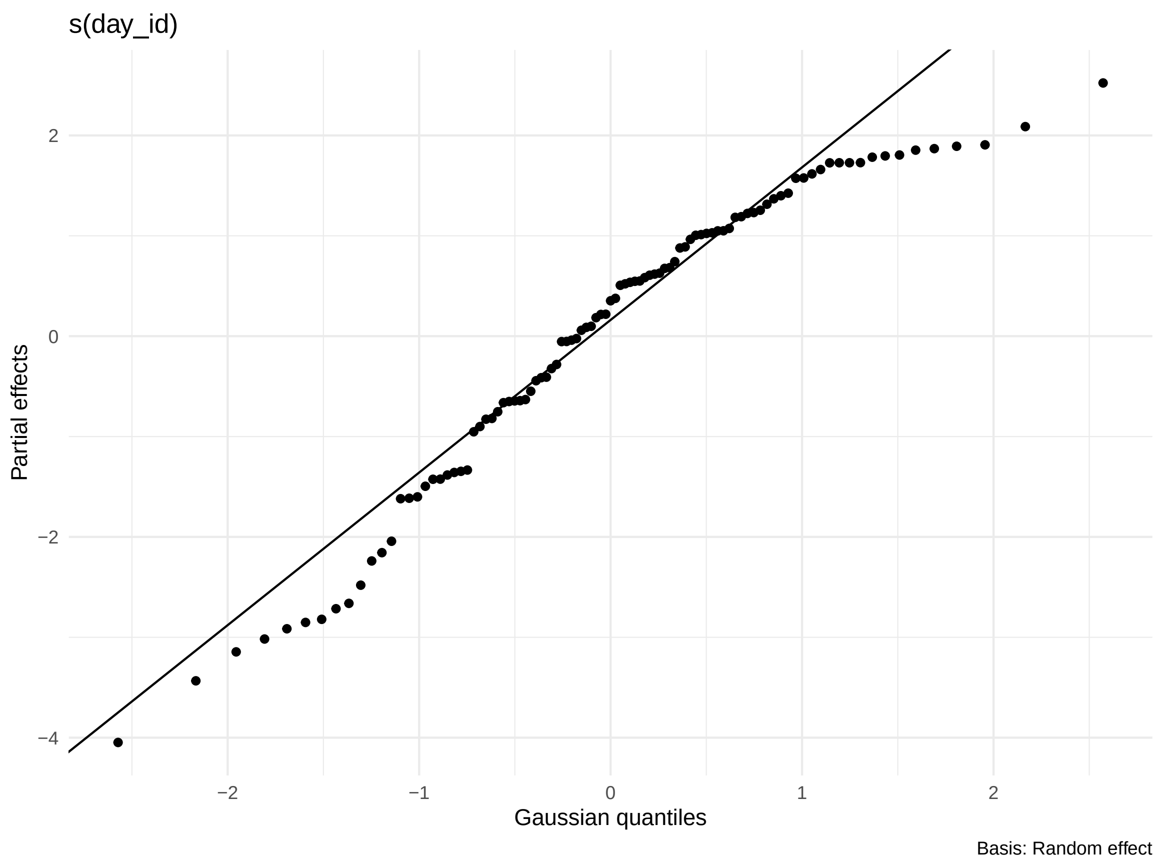



```r
# 3. Visualize the interaction between Temperature and Wind

# Option A: A 2D heatmap of the interaction surface using gratia
draw(m_int_temp_wind, select = "ti(temperature,wind_mean)")
```

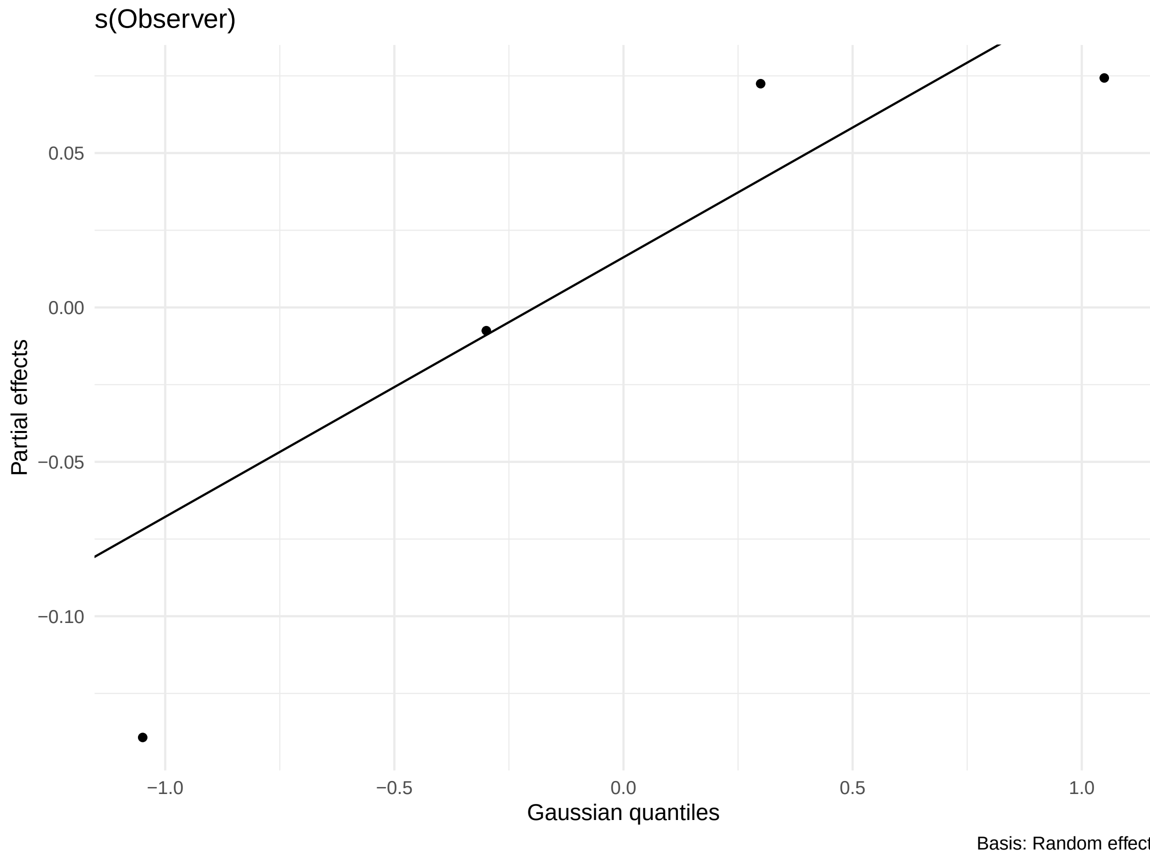



```
# Option B: Using ggeffects to plot conditional effects.
# This shows the effect of temperature at different levels of wind speed.
# It can sometimes be easier to interpret.
library(ggeffects)
# Note: ggpredict has issues with logical AR_start, so we use typical values
#ggpredict(m_int_temp_wind, terms = c("temperature", "wind_mean"),
#          typical = "mean") %>% plot()
```

```
# 4. Visualize the random effects
# This can help understand the variation among days and observers.
draw(m_int_temp_wind, select = "s(day_id)")
```




```
draw(m_int_temp_wind, select = "s(Observer)")
```


s(Observer)
Partial effects
0.05
0.00
−0.05
−0.10
−1.0
−0.5
0.0
0.5
1.0
Gaussian quantiles
Basis: Random effect