Leafhopper distributed lag model

Eric R. Scott

2020-01-25

Table of Contents

# TODO

* for these new hypotheses tests to work, I need to fix the penalization for the covariates at whatever they are in the full model.
* I might be able to do this with the sp parameter, or possibly by setting k and using fx = TRUE

*Last compiled: 2020-12-23*

library(tidyverse)  
library(here)  
library(conflicted)  
library(mgcv)  
library(dlnm)  
library(bbmle)  
  
conflict\_prefer("filter", "dplyr")  
conflict\_prefer("here", "here")  
conflict\_prefer("lag", "dplyr")  
conflict\_prefer("select", "dplyr")

# Methods

This analysis is based on Gasparrini A, Scheipl F, Armstrong B, Kenward MG (2017) A penalized framework for distributed lag non-linear models. Biometrics 73:938–948. doi: 10.1111/biom.12645. Specifically, it is based on ex1\_01.int.R provided in their code. This “internal” method for distributed lag models takes full(?) advantage of built in features of gams.

The goal is to model the effects of temperature and precipitation on leafhoppers as having a non-linear, lagged effect.

# Hypotheses

To achieve this goal, I’ll test three hypotheses for each weather variable (min, mean, and max daily tempearture, and precipitation)

1. Is *smooth* function necessary?
   * HA: Full model including crossbasis smooth for weather variable over 15 days lag.
   * H0: replace crossbasis term with mean of temperature in past 15 days
2. Is there an effect of temperature?
   * HA: Same as HA from #1 above
   * H0: Remove the weather variable term entirely
3. Is there a single lagged effect?
   * HA: Use the weather variable at a lag with maximum effect suggested by the DLNM (e.g. mean temperature 10 days ago) as a parameteric (not smoothed) term.
   * H0: Same as H0 from #2 above

# Read in Data

hoppers <- read\_rds(here("data", "cleaned", "hopper-weather.rds"))  
  
head(hoppers)

## # A tibble: 6 x 20  
## date\_time date day day\_post field interharvest\_id plant  
## <dttm> <date> <dbl> <dbl> <fct> <fct> <fct>  
## 1 2017-06-05 06:00:00 2017-06-05 1 0 A A1 A1   
## 2 2017-06-05 06:00:00 2017-06-05 1 0 A A1 A2   
## 3 2017-06-05 06:00:00 2017-06-05 1 0 A A1 A3   
## 4 2017-06-05 06:00:00 2017-06-05 1 0 A A1 A4   
## 5 2017-06-05 06:00:00 2017-06-05 1 0 A A1 A5   
## 6 2017-06-05 06:00:00 2017-06-05 1 0 A A1 A6   
## # … with 83 more variables: counter <fct>, hoppers <dbl>, n\_leaves <dbl>,  
## # temp\_at\_count <dbl>, temp\_mean <dbl>, temp\_min <dbl>, temp\_max <dbl>,  
## # precip <dbl>, Q\_precip[,"1"] <dbl>, [,"2"] <dbl>, [,"3"] <dbl>,  
## # [,"4"] <dbl>, [,"5"] <dbl>, [,"6"] <dbl>, [,"7"] <dbl>, [,"8"] <dbl>,  
## # [,"9"] <dbl>, [,"10"] <dbl>, [,"11"] <dbl>, [,"12"] <dbl>, [,"13"] <dbl>,  
## # [,"14"] <dbl>, [,"15"] <dbl>, Q\_temp\_mean[,"1"] <dbl>, [,"2"] <dbl>,  
## # [,"3"] <dbl>, [,"4"] <dbl>, [,"5"] <dbl>, [,"6"] <dbl>, [,"7"] <dbl>,  
## # [,"8"] <dbl>, [,"9"] <dbl>, [,"10"] <dbl>, [,"11"] <dbl>, [,"12"] <dbl>,  
## # [,"13"] <dbl>, [,"14"] <dbl>, [,"15"] <dbl>, Q\_temp\_min[,"1"] <dbl>,  
## # [,"2"] <dbl>, [,"3"] <dbl>, [,"4"] <dbl>, [,"5"] <dbl>, [,"6"] <dbl>,  
## # [,"7"] <dbl>, [,"8"] <dbl>, [,"9"] <dbl>, [,"10"] <dbl>, [,"11"] <dbl>,  
## # [,"12"] <dbl>, [,"13"] <dbl>, [,"14"] <dbl>, [,"15"] <dbl>,  
## # Q\_temp\_max[,"1"] <dbl>, [,"2"] <dbl>, [,"3"] <dbl>, [,"4"] <dbl>,  
## # [,"5"] <dbl>, [,"6"] <dbl>, [,"7"] <dbl>, [,"8"] <dbl>, [,"9"] <dbl>,  
## # [,"10"] <dbl>, [,"11"] <dbl>, [,"12"] <dbl>, [,"13"] <dbl>, [,"14"] <dbl>,  
## # [,"15"] <dbl>, L[,1] <int>, [,2] <int>, [,3] <int>, [,4] <int>, [,5] <int>,  
## # [,6] <int>, [,7] <int>, [,8] <int>, [,9] <int>, [,10] <int>, [,11] <int>,  
## # [,12] <int>, [,13] <int>, [,14] <int>, [,15] <int>

# str(hoppers)

Remove any rows with missing values

hoppers <-  
 hoppers %>%  
 rowwise() %>%  
 filter(!is.na(sum(Q\_temp\_max)))

#calculate mean of past 15 days  
hoppers <-  
 hoppers %>%  
 mutate(mean\_temp\_mean = mean(Q\_temp\_mean),  
 mean\_temp\_min = mean(Q\_temp\_min),  
 mean\_temp\_max = mean(Q\_temp\_max),  
 mean\_precip = mean(Q\_precip)) %>%   
 #and the mean of days 9–12  
 mutate(pick\_temp\_mean = mean(Q\_temp\_mean[ , 9:12]),  
 pick\_temp\_min = mean(Q\_temp\_min[ , 9:12]),  
 pick\_temp\_max = mean(Q\_temp\_max[ , 9:12]),  
 pick\_precip = mean(Q\_temp\_max[ , 9:12])) %>%   
 ungroup()

# Build the full models

Temp and precip are highly concurve, so cannot include all weather variables as predictors in the same model.

m\_hopper\_temp\_mean <-   
 gam(hoppers ~  
 s(Q\_temp\_mean, L,  
 bs = "cb",  
 k = c(10, 14),  
 xt = list(bs = "cr")  
 ) +  
 s(day\_post, bs = "cr") +  
 s(counter, bs = "re") + #no difference with random vs. fixed  
 # counter +  
 s(interharvest\_id, bs = "re") +  
 s(plant, bs = "re"),  
 family = poisson(),  
 offset = log(n\_leaves),  
 data = hoppers,  
 gamma = 1.2,  
 method = "REML")

## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'  
  
## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'  
  
## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'

m\_hopper\_temp\_min <-   
 gam(hoppers ~  
 s(Q\_temp\_min, L,  
 bs = "cb",  
 k = c(10, 14),  
 xt = list(bs = "cr")  
 ) +  
 s(day\_post, bs = "cr") +  
 s(counter, bs = "re") + #no difference with random vs. fixed  
 # counter +  
 s(interharvest\_id, bs = "re") +  
 s(plant, bs = "re"),  
 family = poisson(),  
 offset = log(n\_leaves),  
 data = hoppers,  
 gamma = 1.2,  
 method = "REML")

## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'  
  
## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'  
  
## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'

m\_hopper\_temp\_max <-   
 gam(hoppers ~  
 s(Q\_temp\_max, L,  
 bs = "cb",  
 k = c(10, 14),  
 xt = list(bs = "cr")  
 ) +  
 s(day\_post, bs = "cr") +  
 s(counter, bs = "re") + #no difference with random vs. fixed  
 # counter +  
 s(interharvest\_id, bs = "re") +  
 s(plant, bs = "re"),  
 family = poisson(),  
 offset = log(n\_leaves),  
 data = hoppers,  
 gamma = 1.2,  
 method = "REML")

## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'  
  
## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'  
  
## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'

m\_hopper\_precip <-   
 gam(hoppers ~  
 s(Q\_precip, L,  
 bs = "cb",  
 k = c(10, 14),  
 xt = list(bs = "cr")  
 ) +  
 s(day\_post, bs = "cr") +  
 s(counter, bs = "re") + #no difference with random vs. fixed  
 # counter +  
 s(interharvest\_id, bs = "re") +  
 s(plant, bs = "re"),  
 family = poisson(),  
 offset = log(n\_leaves),  
 data = hoppers,  
 gamma = 1.2,  
 method = "REML")

## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'  
  
## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'  
  
## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'

AICtab(m\_hopper\_temp\_mean, m\_hopper\_temp\_min, m\_hopper\_temp\_max, m\_hopper\_precip)

## dAIC df   
## m\_hopper\_temp\_min 0.0 38.8  
## m\_hopper\_temp\_mean 6.8 41.1  
## m\_hopper\_precip 7.5 43.4  
## m\_hopper\_temp\_max 9.1 39.7

Minimum temperature model “wins”

# Save models

write\_rds(m\_hopper\_precip, here("analysis", "models", "m\_hopper\_precip.rds"))  
write\_rds(m\_hopper\_temp\_max, here("analysis", "models", "m\_hopper\_temp\_max.rds"))  
write\_rds(m\_hopper\_temp\_mean, here("analysis", "models", "m\_hopper\_temp\_mean.rds"))  
write\_rds(m\_hopper\_temp\_min, here("analysis", "models", "m\_hopper\_temp\_min.rds"))

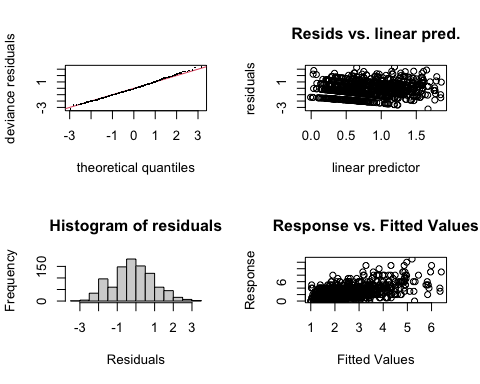
# Mean Temperature

## Inspect the GAM

#check for overdispersion  
sum(residuals(m\_hopper\_temp\_mean, type = "pearson")^2) / df.residual(m\_hopper\_temp\_mean)

## [1] 1.169463

gam.check(m\_hopper\_temp\_mean)



##   
## Method: REML Optimizer: outer newton  
## full convergence after 7 iterations.  
## Gradient range [-0.0007421781,0.0004188801]  
## (score 1453.49 & scale 1).  
## Hessian positive definite, eigenvalue range [0.0005558872,4.976186].  
## Model rank = 165 / 165   
##   
## 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(Q\_temp\_mean,L) 126.00 12.60 NA NA  
## s(day\_post) 9.00 1.01 0.98 0.5  
## s(counter) 3.00 1.80 NA NA  
## s(interharvest\_id) 6.00 4.03 NA NA  
## s(plant) 20.00 13.61 NA NA

Not overdispersed, appropriate numbers of knots used, residuals meet assumptions (mostly).

## Marginal Hypothesis test

P-values for covariates. P-values for crossbasis obtained through specific LRT below.

anova.gam(m\_hopper\_temp\_mean)

## Warning: partial match of 'val' to 'values'

##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## hoppers ~ s(Q\_temp\_mean, L, bs = "cb", k = c(10, 14), xt = list(bs = "cr")) +   
## s(day\_post, bs = "cr") + s(counter, bs = "re") + s(interharvest\_id,   
## bs = "re") + s(plant, bs = "re")  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value  
## s(Q\_temp\_mean,L) 12.600 16.026 59.406 0.000004548  
## s(day\_post) 1.006 1.012 4.524 0.0343  
## s(counter) 1.799 2.000 35.815 < 2e-16  
## s(interharvest\_id) 4.032 5.000 233.247 < 2e-16  
## s(plant) 13.614 19.000 82.829 0.000000666

## H1: Is smooth necessary?

sp\_hopper\_temp\_mean <- m\_hopper\_temp\_mean$sp  
  
m\_hopper\_temp\_mean1 <-   
 gam(hoppers ~  
 mean\_temp\_mean +  
 s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_mean["s(day\_post)"]) +  
 s(counter, bs = "re", sp = sp\_hopper\_temp\_mean["s(counter)"]) +   
 s(interharvest\_id, bs = "re",  
 sp = sp\_hopper\_temp\_mean["s(interharvest\_id)"]) +  
 s(plant, bs = "re", sp = sp\_hopper\_temp\_mean["s(plant)"]),  
 family = poisson(),  
 offset = log(n\_leaves),  
 data = hoppers,  
 gamma = 1.2,  
 method = "REML")

## Warning in attr(pterms[tind[j]], "term.label"): partial match of 'term.label' to  
## 'term.labels'

## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'

## Warning: partial match of 'w' to 'weights'

h1\_hopper\_temp\_mean <- anova.gam(m\_hopper\_temp\_mean1, m\_hopper\_temp\_mean, test = "Chisq")

## Warning: partial match of 'val' to 'values'

h1\_hopper\_temp\_mean

## Analysis of Deviance Table  
##   
## Model 1: hoppers ~ mean\_temp\_mean + s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_mean["s(day\_post)"]) +   
## s(counter, bs = "re", sp = sp\_hopper\_temp\_mean["s(counter)"]) +   
## s(interharvest\_id, bs = "re", sp = sp\_hopper\_temp\_mean["s(interharvest\_id)"]) +   
## s(plant, bs = "re", sp = sp\_hopper\_temp\_mean["s(plant)"])  
## Model 2: hoppers ~ s(Q\_temp\_mean, L, bs = "cb", k = c(10, 14), xt = list(bs = "cr")) +   
## s(day\_post, bs = "cr") + s(counter, bs = "re") + s(interharvest\_id,   
## bs = "re") + s(plant, bs = "re")  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)   
## 1 900.24 1207.2   
## 2 878.55 1150.3 21.695 56.847 0.000055 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Yes, smooth is significantly better than just a mean of the past 15 days.

## H2: Is there an effect of daily mean temp?

m\_hopper\_temp\_mean\_null <-   
 gam(hoppers ~  
 s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_mean["s(day\_post)"]) +  
 s(counter, bs = "re", sp = sp\_hopper\_temp\_mean["s(counter)"]) +   
 s(interharvest\_id, bs = "re",  
 sp = sp\_hopper\_temp\_mean["s(interharvest\_id)"]) +  
 s(plant, bs = "re", sp = sp\_hopper\_temp\_mean["s(plant)"]),  
 family = poisson(),  
 offset = log(n\_leaves),  
 data = hoppers,  
 gamma = 1.2,  
 method = "REML")

## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'

## Warning: partial match of 'w' to 'weights'

h2\_hopper\_temp\_mean <- anova.gam(m\_hopper\_temp\_mean\_null, m\_hopper\_temp\_mean, test = "Chisq")

## Warning: partial match of 'val' to 'values'

h2\_hopper\_temp\_mean

## Analysis of Deviance Table  
##   
## Model 1: hoppers ~ s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_mean["s(day\_post)"]) +   
## s(counter, bs = "re", sp = sp\_hopper\_temp\_mean["s(counter)"]) +   
## s(interharvest\_id, bs = "re", sp = sp\_hopper\_temp\_mean["s(interharvest\_id)"]) +   
## s(plant, bs = "re", sp = sp\_hopper\_temp\_mean["s(plant)"])  
## Model 2: hoppers ~ s(Q\_temp\_mean, L, bs = "cb", k = c(10, 14), xt = list(bs = "cr")) +   
## s(day\_post, bs = "cr") + s(counter, bs = "re") + s(interharvest\_id,   
## bs = "re") + s(plant, bs = "re")  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)   
## 1 901.20 1228.9   
## 2 878.55 1150.3 22.647 78.592 0.00000004251 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#larger residual DF means smaller estimated degrees of freedom

Yes, there is a significant effect of lagged mean daily temperature on leafhoppers

## H3: Is there an effect of a specific lag?

According to plots, all models had a steep slope (strong effect) between 9 and 12 days of lag. I’ll use the mean of those days as a predictor.

m\_hopper\_temp\_mean3 <-   
 gam(hoppers ~  
 pick\_temp\_mean +  
 s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_mean["s(day\_post)"]) +  
 s(counter, bs = "re", sp = sp\_hopper\_temp\_mean["s(counter)"]) +   
 s(interharvest\_id, bs = "re",  
 sp = sp\_hopper\_temp\_mean["s(interharvest\_id)"]) +  
 s(plant, bs = "re", sp = sp\_hopper\_temp\_mean["s(plant)"]),  
 family = poisson(),  
 offset = log(n\_leaves),  
 data = hoppers,  
 gamma = 1.2,  
 method = "REML")

## Warning in attr(pterms[tind[j]], "term.label"): partial match of 'term.label' to  
## 'term.labels'

## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'

## Warning: partial match of 'w' to 'weights'

h3\_hopper\_temp\_mean <- anova.gam(m\_hopper\_temp\_mean\_null, m\_hopper\_temp\_mean3, test = "Chisq")  
h3\_hopper\_temp\_mean

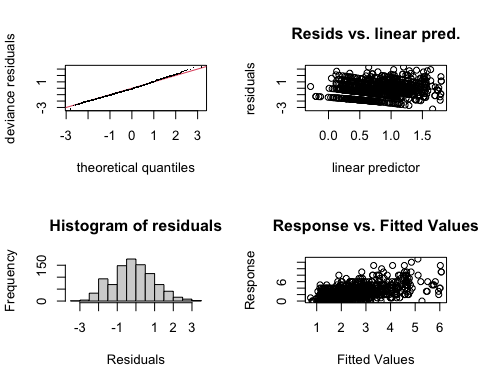
## Analysis of Deviance Table  
##   
## Model 1: hoppers ~ s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_mean["s(day\_post)"]) +   
## s(counter, bs = "re", sp = sp\_hopper\_temp\_mean["s(counter)"]) +   
## s(interharvest\_id, bs = "re", sp = sp\_hopper\_temp\_mean["s(interharvest\_id)"]) +   
## s(plant, bs = "re", sp = sp\_hopper\_temp\_mean["s(plant)"])  
## Model 2: hoppers ~ pick\_temp\_mean + s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_mean["s(day\_post)"]) +   
## s(counter, bs = "re", sp = sp\_hopper\_temp\_mean["s(counter)"]) +   
## s(interharvest\_id, bs = "re", sp = sp\_hopper\_temp\_mean["s(interharvest\_id)"]) +   
## s(plant, bs = "re", sp = sp\_hopper\_temp\_mean["s(plant)"])  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)   
## 1 901.2 1228.9   
## 2 900.2 1208.6 0.99195 20.323 0.000006428 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Significant effect of mean temp 10 days ago

# Minimum temperature

## Inspect the GAM

gam.check(m\_hopper\_temp\_min)



##   
## Method: REML Optimizer: outer newton  
## full convergence after 9 iterations.  
## Gradient range [-0.0004254752,0.0001679812]  
## (score 1449.283 & scale 1).  
## Hessian positive definite, eigenvalue range [0.0004251673,4.952216].  
## Model rank = 165 / 165   
##   
## 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(Q\_temp\_min,L) 126.00 11.18 NA NA  
## s(day\_post) 9.00 1.00 0.98 0.57  
## s(counter) 3.00 1.78 NA NA  
## s(interharvest\_id) 6.00 4.19 NA NA  
## s(plant) 20.00 13.63 NA NA

#check for overdispersion  
sum(residuals(m\_hopper\_temp\_min, type = "pearson")^2) / df.residual(m\_hopper\_temp\_min)

## [1] 1.166184

## Marginal Hypothesis test

P-values for covariates. P-values for crossbasis obtained through specific LRT below.

anova.gam(m\_hopper\_temp\_min)

## Warning: partial match of 'val' to 'values'

##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## hoppers ~ s(Q\_temp\_min, L, bs = "cb", k = c(10, 14), xt = list(bs = "cr")) +   
## s(day\_post, bs = "cr") + s(counter, bs = "re") + s(interharvest\_id,   
## bs = "re") + s(plant, bs = "re")  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value  
## s(Q\_temp\_min,L) 11.177 13.915 64.872 <2e-16  
## s(day\_post) 1.001 1.002 0.003 0.968  
## s(counter) 1.781 2.000 33.488 <2e-16  
## s(interharvest\_id) 4.187 5.000 235.671 <2e-16  
## s(plant) 13.628 19.000 86.049 <2e-16

## H1: Is smooth necessary?

sp\_hopper\_temp\_min <- m\_hopper\_temp\_min$sp  
  
m\_hopper\_temp\_min1 <-   
 gam(hoppers ~  
 mean\_temp\_min +  
 s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_min["s(day\_post)"]) +  
 s(counter, bs = "re", sp = sp\_hopper\_temp\_min["s(counter)"]) +   
 s(interharvest\_id, bs = "re",  
 sp = sp\_hopper\_temp\_min["s(interharvest\_id)"]) +  
 s(plant, bs = "re", sp = sp\_hopper\_temp\_min["s(plant)"]),  
 family = poisson(),  
 offset = log(n\_leaves),  
 data = hoppers,  
 gamma = 1.2,  
 method = "REML")

## Warning in attr(pterms[tind[j]], "term.label"): partial match of 'term.label' to  
## 'term.labels'

## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'

## Warning: partial match of 'w' to 'weights'

h1\_hopper\_temp\_min <- anova.gam(m\_hopper\_temp\_min1, m\_hopper\_temp\_min, test = "Chisq")

## Warning: partial match of 'val' to 'values'

h1\_hopper\_temp\_min

## Analysis of Deviance Table  
##   
## Model 1: hoppers ~ mean\_temp\_min + s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_min["s(day\_post)"]) +   
## s(counter, bs = "re", sp = sp\_hopper\_temp\_min["s(counter)"]) +   
## s(interharvest\_id, bs = "re", sp = sp\_hopper\_temp\_min["s(interharvest\_id)"]) +   
## s(plant, bs = "re", sp = sp\_hopper\_temp\_min["s(plant)"])  
## Model 2: hoppers ~ s(Q\_temp\_min, L, bs = "cb", k = c(10, 14), xt = list(bs = "cr")) +   
## s(day\_post, bs = "cr") + s(counter, bs = "re") + s(interharvest\_id,   
## bs = "re") + s(plant, bs = "re")  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)   
## 1 900.22 1227.4   
## 2 881.54 1148.2 18.683 79.248 1.969e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Yes, the smooth is a significant improvment over just mean minimum temperature

## H2: Is there an effect of min temp?

m\_hopper\_temp\_min\_null <-   
 gam(hoppers ~  
 s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_min["s(day\_post)"]) +  
 s(counter, bs = "re", sp = sp\_hopper\_temp\_min["s(counter)"]) +   
 s(interharvest\_id, bs = "re",  
 sp = sp\_hopper\_temp\_min["s(interharvest\_id)"]) +  
 s(plant, bs = "re", sp = sp\_hopper\_temp\_min["s(plant)"]),  
 family = poisson(),  
 offset = log(n\_leaves),  
 data = hoppers,  
 gamma = 1.2,  
 method = "REML")

## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'

## Warning: partial match of 'w' to 'weights'

h2\_hopper\_temp\_min <- anova.gam(m\_hopper\_temp\_min\_null, m\_hopper\_temp\_min, test = "Chisq")

## Warning: partial match of 'val' to 'values'

h2\_hopper\_temp\_min

## Analysis of Deviance Table  
##   
## Model 1: hoppers ~ s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_min["s(day\_post)"]) +   
## s(counter, bs = "re", sp = sp\_hopper\_temp\_min["s(counter)"]) +   
## s(interharvest\_id, bs = "re", sp = sp\_hopper\_temp\_min["s(interharvest\_id)"]) +   
## s(plant, bs = "re", sp = sp\_hopper\_temp\_min["s(plant)"])  
## Model 2: hoppers ~ s(Q\_temp\_min, L, bs = "cb", k = c(10, 14), xt = list(bs = "cr")) +   
## s(day\_post, bs = "cr") + s(counter, bs = "re") + s(interharvest\_id,   
## bs = "re") + s(plant, bs = "re")  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)   
## 1 901.22 1229.0   
## 2 881.54 1148.2 19.674 80.83 2.226e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#larger residual DF means smaller estimated degrees of freedom

yes, highly significant effect of temperature

## H3: Is there an effect of a specific lag?

m\_hopper\_temp\_min3 <-   
 gam(hoppers ~  
 pick\_temp\_min +  
 s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_min["s(day\_post)"]) +  
 s(counter, bs = "re", sp = sp\_hopper\_temp\_min["s(counter)"]) +   
 s(interharvest\_id, bs = "re",  
 sp = sp\_hopper\_temp\_min["s(interharvest\_id)"]) +  
 s(plant, bs = "re", sp = sp\_hopper\_temp\_min["s(plant)"]),  
 family = poisson(),  
 offset = log(n\_leaves),  
 data = hoppers,  
 gamma = 1.2,  
 method = "REML")

## Warning in attr(pterms[tind[j]], "term.label"): partial match of 'term.label' to  
## 'term.labels'

## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'

## Warning: partial match of 'w' to 'weights'

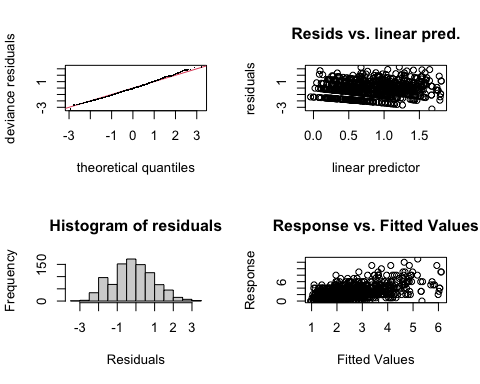
h3\_hopper\_temp\_min <- anova.gam(m\_hopper\_temp\_min\_null, m\_hopper\_temp\_min3, test = "Chisq")  
h3\_hopper\_temp\_min

## Analysis of Deviance Table  
##   
## Model 1: hoppers ~ s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_min["s(day\_post)"]) +   
## s(counter, bs = "re", sp = sp\_hopper\_temp\_min["s(counter)"]) +   
## s(interharvest\_id, bs = "re", sp = sp\_hopper\_temp\_min["s(interharvest\_id)"]) +   
## s(plant, bs = "re", sp = sp\_hopper\_temp\_min["s(plant)"])  
## Model 2: hoppers ~ pick\_temp\_min + s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_min["s(day\_post)"]) +   
## s(counter, bs = "re", sp = sp\_hopper\_temp\_min["s(counter)"]) +   
## s(interharvest\_id, bs = "re", sp = sp\_hopper\_temp\_min["s(interharvest\_id)"]) +   
## s(plant, bs = "re", sp = sp\_hopper\_temp\_min["s(plant)"])  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)  
## 1 901.22 1229.0   
## 2 900.22 1226.5 0.9981 2.4646 0.1161

# Maximum temperature

## Inspect the GAM

gam.check(m\_hopper\_temp\_max)



##   
## Method: REML Optimizer: outer newton  
## full convergence after 6 iterations.  
## Gradient range [-0.001735882,0.0003863561]  
## (score 1452.504 & scale 1).  
## Hessian positive definite, eigenvalue range [0.0100265,4.986818].  
## Model rank = 165 / 165   
##   
## 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(Q\_temp\_max,L) 126.00 11.04 NA NA  
## s(day\_post) 9.00 1.19 0.98 0.46  
## s(counter) 3.00 1.78 NA NA  
## s(interharvest\_id) 6.00 3.95 NA NA  
## s(plant) 20.00 13.64 NA NA

#check for overdispersion  
sum(residuals(m\_hopper\_temp\_max, type = "pearson")^2) / df.residual(m\_hopper\_temp\_max)

## [1] 1.172665

## Marginal Hypothesis test

P-values for covariates. P-values for crossbasis obtained through specific LRT below.

anova.gam(m\_hopper\_temp\_max)

## Warning: partial match of 'val' to 'values'

##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## hoppers ~ s(Q\_temp\_max, L, bs = "cb", k = c(10, 14), xt = list(bs = "cr")) +   
## s(day\_post, bs = "cr") + s(counter, bs = "re") + s(interharvest\_id,   
## bs = "re") + s(plant, bs = "re")  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value  
## s(Q\_temp\_max,L) 11.038 14.110 57.211 0.000003958  
## s(day\_post) 1.188 1.327 4.004 0.0903  
## s(counter) 1.776 2.000 32.089 0.000000291  
## s(interharvest\_id) 3.953 5.000 232.759 < 2e-16  
## s(plant) 13.641 19.000 83.668 0.000000671

## H1: Is smooth necessary?

sp\_hopper\_temp\_max <- m\_hopper\_temp\_max$sp  
m\_hopper\_temp\_max1 <-   
 gam(hoppers ~  
 mean\_temp\_max +  
 s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_max["s(day\_post)"]) +  
 s(counter, bs = "re", sp = sp\_hopper\_temp\_max["s(counter)"]) +  
 s(interharvest\_id, bs = "re",  
 sp = sp\_hopper\_temp\_max["s(interharvest\_id)"]) +  
 s(plant, bs = "re", sp = sp\_hopper\_temp\_max["s(plant)"]),  
 family = poisson(),  
 offset = log(n\_leaves),  
 data = hoppers,  
 gamma = 1.2,  
 method = "REML")

## Warning in attr(pterms[tind[j]], "term.label"): partial match of 'term.label' to  
## 'term.labels'

## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'

## Warning: partial match of 'w' to 'weights'

h1\_hopper\_temp\_max <- anova.gam(m\_hopper\_temp\_max1, m\_hopper\_temp\_max, test = "Chisq")

## Warning: partial match of 'val' to 'values'

h1\_hopper\_temp\_max

## Analysis of Deviance Table  
##   
## Model 1: hoppers ~ mean\_temp\_max + s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_max["s(day\_post)"]) +   
## s(counter, bs = "re", sp = sp\_hopper\_temp\_max["s(counter)"]) +   
## s(interharvest\_id, bs = "re", sp = sp\_hopper\_temp\_max["s(interharvest\_id)"]) +   
## s(plant, bs = "re", sp = sp\_hopper\_temp\_max["s(plant)"])  
## Model 2: hoppers ~ s(Q\_temp\_max, L, bs = "cb", k = c(10, 14), xt = list(bs = "cr")) +   
## s(day\_post, bs = "cr") + s(counter, bs = "re") + s(interharvest\_id,   
## bs = "re") + s(plant, bs = "re")  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)   
## 1 899.70 1202.9   
## 2 880.22 1155.5 19.474 47.389 0.0003995 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Smooth is marginally significant

## H2: Is there an effect of max temp?

m\_hopper\_temp\_max\_null <-   
 gam(hoppers ~  
 s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_max["s(day\_post)"]) +  
 s(counter, bs = "re", sp = sp\_hopper\_temp\_max["s(counter)"]) +   
 s(interharvest\_id, bs = "re",  
 sp = sp\_hopper\_temp\_max["s(interharvest\_id)"]) +  
 s(plant, bs = "re", sp = sp\_hopper\_temp\_max["s(plant)"]),  
 family = poisson(),  
 offset = log(n\_leaves),  
 data = hoppers,  
 gamma = 1.2,  
 method = "REML")

## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'

## Warning: partial match of 'w' to 'weights'

h2\_hopper\_temp\_max <- anova.gam(m\_hopper\_temp\_max\_null, m\_hopper\_temp\_max,  
 test = "Chisq")

## Warning: partial match of 'val' to 'values'

h2\_hopper\_temp\_max

## Analysis of Deviance Table  
##   
## Model 1: hoppers ~ s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_max["s(day\_post)"]) +   
## s(counter, bs = "re", sp = sp\_hopper\_temp\_max["s(counter)"]) +   
## s(interharvest\_id, bs = "re", sp = sp\_hopper\_temp\_max["s(interharvest\_id)"]) +   
## s(plant, bs = "re", sp = sp\_hopper\_temp\_max["s(plant)"])  
## Model 2: hoppers ~ s(Q\_temp\_max, L, bs = "cb", k = c(10, 14), xt = list(bs = "cr")) +   
## s(day\_post, bs = "cr") + s(counter, bs = "re") + s(interharvest\_id,   
## bs = "re") + s(plant, bs = "re")  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)   
## 1 900.58 1227.1   
## 2 880.22 1155.5 20.358 71.573 0.0000001282 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#larger residual DF means smaller estimated degrees of freedom

Significant effect of crossbasis

## H3: Is there an effect of a specific lag?

m\_hopper\_temp\_max3 <-   
 gam(hoppers ~  
 pick\_temp\_max +  
 s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_max["s(day\_post)"]) +  
 s(counter, bs = "re", sp = sp\_hopper\_temp\_max["s(counter)"]) +   
 s(interharvest\_id, bs = "re",  
 sp = sp\_hopper\_temp\_max["s(interharvest\_id)"]) +  
 s(plant, bs = "re", sp = sp\_hopper\_temp\_max["s(plant)"]),  
 family = poisson(),  
 offset = log(n\_leaves),  
 data = hoppers,  
 gamma = 1.2,  
 method = "REML")

## Warning in attr(pterms[tind[j]], "term.label"): partial match of 'term.label' to  
## 'term.labels'

## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'

## Warning: partial match of 'w' to 'weights'

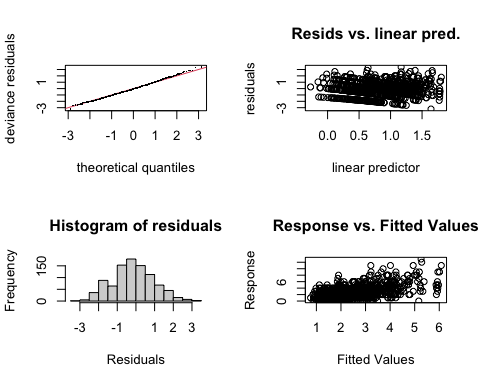
h3\_hopper\_temp\_max <- anova.gam(m\_hopper\_temp\_max\_null, m\_hopper\_temp\_max3, test = "Chisq")  
h3\_hopper\_temp\_max

## Analysis of Deviance Table  
##   
## Model 1: hoppers ~ s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_max["s(day\_post)"]) +   
## s(counter, bs = "re", sp = sp\_hopper\_temp\_max["s(counter)"]) +   
## s(interharvest\_id, bs = "re", sp = sp\_hopper\_temp\_max["s(interharvest\_id)"]) +   
## s(plant, bs = "re", sp = sp\_hopper\_temp\_max["s(plant)"])  
## Model 2: hoppers ~ pick\_temp\_max + s(day\_post, bs = "cr", sp = sp\_hopper\_temp\_max["s(day\_post)"]) +   
## s(counter, bs = "re", sp = sp\_hopper\_temp\_max["s(counter)"]) +   
## s(interharvest\_id, bs = "re", sp = sp\_hopper\_temp\_max["s(interharvest\_id)"]) +   
## s(plant, bs = "re", sp = sp\_hopper\_temp\_max["s(plant)"])  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)   
## 1 900.58 1227.1   
## 2 899.63 1206.3 0.95266 20.753 0.000004703 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Precip

## Inspect GAM

gam.check(m\_hopper\_precip)



##   
## Method: REML Optimizer: outer newton  
## full convergence after 8 iterations.  
## Gradient range [-0.001680635,0.0001684709]  
## (score 1449.65 & scale 1).  
## Hessian positive definite, eigenvalue range [0.00113261,4.953116].  
## Model rank = 165 / 165   
##   
## 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(Q\_precip,L) 126.00 13.44 NA NA  
## s(day\_post) 9.00 1.03 0.99 0.59  
## s(counter) 3.00 1.79 NA NA  
## s(interharvest\_id) 6.00 4.27 NA NA  
## s(plant) 20.00 13.63 NA NA

sum(residuals(m\_hopper\_precip, type = "pearson")^2) / df.residual(m\_hopper\_precip)

## [1] 1.168599

## Marginal hypothesis test

anova.gam(m\_hopper\_precip)

## Warning: partial match of 'val' to 'values'

##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## hoppers ~ s(Q\_precip, L, bs = "cb", k = c(10, 14), xt = list(bs = "cr")) +   
## s(day\_post, bs = "cr") + s(counter, bs = "re") + s(interharvest\_id,   
## bs = "re") + s(plant, bs = "re")  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value  
## s(Q\_precip,L) 13.441 17.745 63.000 0.00000132  
## s(day\_post) 1.029 1.053 3.534 0.0592  
## s(counter) 1.793 2.000 34.797 < 2e-16  
## s(interharvest\_id) 4.267 5.000 230.460 < 2e-16  
## s(plant) 13.627 19.000 83.477 < 2e-16

Marginally significant effect of day

## H1: Is smooth necessary?

sp\_hopper\_precip <- m\_hopper\_precip$sp  
  
m\_hopper\_precip1 <-   
 gam(hoppers ~  
 mean\_precip +  
 s(day\_post, bs = "cr", sp = sp\_hopper\_precip["s(day\_post)"]) +  
 s(counter, bs = "re", sp = sp\_hopper\_precip["s(counter)"]) +  
 s(interharvest\_id, bs = "re",  
 sp = sp\_hopper\_precip["s(interharvest\_id)"]) +  
 s(plant, bs = "re", sp = sp\_hopper\_precip["s(plant)"]),  
 family = poisson(),  
 offset = log(n\_leaves),  
 data = hoppers,  
 gamma = 1.2,  
 method = "REML")

## Warning in attr(pterms[tind[j]], "term.label"): partial match of 'term.label' to  
## 'term.labels'

## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'

## Warning: partial match of 'w' to 'weights'

h1\_hopper\_precip <- anova.gam(m\_hopper\_precip1, m\_hopper\_precip, test = "Chisq")

## Warning: partial match of 'val' to 'values'

h1\_hopper\_precip

## Analysis of Deviance Table  
##   
## Model 1: hoppers ~ mean\_precip + s(day\_post, bs = "cr", sp = sp\_hopper\_precip["s(day\_post)"]) +   
## s(counter, bs = "re", sp = sp\_hopper\_precip["s(counter)"]) +   
## s(interharvest\_id, bs = "re", sp = sp\_hopper\_precip["s(interharvest\_id)"]) +   
## s(plant, bs = "re", sp = sp\_hopper\_precip["s(plant)"])  
## Model 2: hoppers ~ s(Q\_precip, L, bs = "cb", k = c(10, 14), xt = list(bs = "cr")) +   
## s(day\_post, bs = "cr") + s(counter, bs = "re") + s(interharvest\_id,   
## bs = "re") + s(plant, bs = "re")  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)   
## 1 900.15 1204.3   
## 2 875.41 1146.5 24.736 57.806 0.0001841 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Yes, crossbasis smooth is significantly better than just mean over past 15 days.

## H2: Is there an effect of max temp?

m\_hopper\_precip\_null <-   
 gam(hoppers ~  
 s(day\_post, bs = "cr", sp = sp\_hopper\_precip["s(day\_post)"]) +  
 s(counter, bs = "re", sp = sp\_hopper\_precip["s(counter)"]) +   
 s(interharvest\_id, bs = "re",  
 sp = sp\_hopper\_precip["s(interharvest\_id)"]) +  
 s(plant, bs = "re", sp = sp\_hopper\_precip["s(plant)"]),  
 family = poisson(),  
 offset = log(n\_leaves),  
 data = hoppers,  
 gamma = 1.2,  
 method = "REML")

## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'

## Warning: partial match of 'w' to 'weights'

h2\_hopper\_precip <- anova.gam(m\_hopper\_precip\_null, m\_hopper\_precip, test = "Chisq")

## Warning: partial match of 'val' to 'values'

h2\_hopper\_precip

## Analysis of Deviance Table  
##   
## Model 1: hoppers ~ s(day\_post, bs = "cr", sp = sp\_hopper\_precip["s(day\_post)"]) +   
## s(counter, bs = "re", sp = sp\_hopper\_precip["s(counter)"]) +   
## s(interharvest\_id, bs = "re", sp = sp\_hopper\_precip["s(interharvest\_id)"]) +   
## s(plant, bs = "re", sp = sp\_hopper\_precip["s(plant)"])  
## Model 2: hoppers ~ s(Q\_precip, L, bs = "cb", k = c(10, 14), xt = list(bs = "cr")) +   
## s(day\_post, bs = "cr") + s(counter, bs = "re") + s(interharvest\_id,   
## bs = "re") + s(plant, bs = "re")  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)   
## 1 901.14 1228.8   
## 2 875.41 1146.5 25.726 82.297 0.00000007807 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#larger residual DF means smaller estimated degrees of freedom

Precipitation has a significant effect.

## H3: Is there an effect of a specific lag?

m\_hopper\_precip3 <-   
 gam(hoppers ~  
 pick\_precip +  
 s(day\_post, bs = "cr", sp = sp\_hopper\_precip["s(day\_post)"]) +  
 s(counter, bs = "re", sp = sp\_hopper\_precip["s(counter)"]) +  
 s(interharvest\_id, bs = "re",   
 sp = sp\_hopper\_precip["s(interharvest\_id)"]) +  
 s(plant, bs = "re", sp = sp\_hopper\_precip["s(plant)"]),  
 family = poisson(),  
 offset = log(n\_leaves),  
 data = hoppers,  
 gamma = 1.2,  
 method = "REML")

## Warning in attr(pterms[tind[j]], "term.label"): partial match of 'term.label' to  
## 'term.labels'

## Warning in seq.default(0, 1, length = nk): partial argument match of 'length' to  
## 'length.out'

## Warning: partial match of 'w' to 'weights'

h3\_hopper\_precip <- anova.gam(m\_hopper\_precip\_null, m\_hopper\_precip3, test = "Chisq")  
h3\_hopper\_precip

## Analysis of Deviance Table  
##   
## Model 1: hoppers ~ s(day\_post, bs = "cr", sp = sp\_hopper\_precip["s(day\_post)"]) +   
## s(counter, bs = "re", sp = sp\_hopper\_precip["s(counter)"]) +   
## s(interharvest\_id, bs = "re", sp = sp\_hopper\_precip["s(interharvest\_id)"]) +   
## s(plant, bs = "re", sp = sp\_hopper\_precip["s(plant)"])  
## Model 2: hoppers ~ pick\_precip + s(day\_post, bs = "cr", sp = sp\_hopper\_precip["s(day\_post)"]) +   
## s(counter, bs = "re", sp = sp\_hopper\_precip["s(counter)"]) +   
## s(interharvest\_id, bs = "re", sp = sp\_hopper\_precip["s(interharvest\_id)"]) +   
## s(plant, bs = "re", sp = sp\_hopper\_precip["s(plant)"])  
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)   
## 1 901.14 1228.8   
## 2 900.15 1207.5 0.9868 21.338 0.000003739 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Table

Collect model statistics:

full\_models <- list(m\_hopper\_temp\_min,  
 m\_hopper\_temp\_mean,  
 m\_hopper\_temp\_max,  
 m\_hopper\_precip)  
h1 <- list(h1\_hopper\_temp\_min,  
 h1\_hopper\_temp\_mean,  
 h1\_hopper\_temp\_max,  
 h1\_hopper\_precip)  
  
h2 <- list(h2\_hopper\_temp\_min,  
 h2\_hopper\_temp\_mean,  
 h2\_hopper\_temp\_max,  
 h2\_hopper\_precip)  
h3 <- list(h3\_hopper\_temp\_min,  
 h3\_hopper\_temp\_mean,  
 h3\_hopper\_temp\_max,  
 h3\_hopper\_precip)  
  
df <-   
 tibble(weather\_var = c("Minimum daily temp.",  
 "Mean daily temp.",  
 "Maximum daily temp.",  
 "Total daily precip."),  
 r2 = map\_dbl(full\_models, ~summary(.x)$r.sq),  
 aic = map\_dbl(full\_models, AIC),  
 h1\_df = map\_dbl(h1, ~pull(.x, "Df")[2]),  
 h1\_x2 = map\_dbl(h1, ~pull(.x, "Deviance")[2]),  
 h1\_p = map\_dbl(h1, ~pull(.x, "Pr(>Chi)")[2]),  
 h2\_df = map\_dbl(h2, ~pull(.x, "Df")[2]),  
 h2\_x2 = map\_dbl(h2, ~pull(.x, "Deviance")[2]),  
 h2\_p = map\_dbl(h2, ~pull(.x, "Pr(>Chi)")[2]),  
 h3\_df = map\_dbl(h3, ~pull(.x, "Df")[2]),  
 h3\_x2 = map\_dbl(h3, ~pull(.x, "Deviance")[2]),  
 h3\_p = map\_dbl(h3, ~pull(.x, "Pr(>Chi)")[2]),  
 )

## Warning: partial match of 'val' to 'values'  
  
## Warning: partial match of 'val' to 'values'  
  
## Warning: partial match of 'val' to 'values'  
  
## Warning: partial match of 'val' to 'values'

df

## # A tibble: 4 x 12  
## weather\_var r2 aic h1\_df h1\_x2 h1\_p h2\_df h2\_x2 h2\_p h3\_df h3\_x2  
## <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 Minimum da… 0.284 3426. 18.7 79.2 1.97e-9 19.7 80.8 2.23e-9 0.998 2.46  
## 2 Mean daily… 0.288 3432. 21.7 56.8 5.50e-5 22.6 78.6 4.25e-8 0.992 20.3   
## 3 Maximum da… 0.282 3435. 19.5 47.4 4.00e-4 20.4 71.6 1.28e-7 0.953 20.8   
## 4 Total dail… 0.293 3433. 24.7 57.8 1.84e-4 25.7 82.3 7.81e-8 0.987 21.3   
## # … with 1 more variable: h3\_p <dbl>

library(flextable)  
get\_flextable\_defaults()

## $font.family  
## [1] "Helvetica"  
##   
## $font.size  
## [1] 11  
##   
## $font.color  
## [1] "black"  
##   
## $text.align  
## [1] "left"  
##   
## $padding.bottom  
## [1] 3  
##   
## $padding.top  
## [1] 3  
##   
## $padding.left  
## [1] 3  
##   
## $padding.right  
## [1] 3  
##   
## $border.color  
## [1] "black"  
##   
## $background.color  
## [1] "transparent"  
##   
## $table.layout  
## [1] "fixed"  
##   
## $decimal.mark  
## [1] "."  
##   
## $big.mark  
## [1] ","  
##   
## $digits  
## [1] 2  
##   
## $na\_str  
## [1] ""  
##   
## $fmt\_date  
## [1] "%Y-%m-%d"  
##   
## $fmt\_datetime  
## [1] "%Y-%m-%d %H:%M:%S"  
##   
## $fonts\_ignore  
## [1] FALSE  
##   
## $theme\_fun  
## [1] "theme\_booktabs"

set\_flextable\_defaults(font.family = "Times New Roman",  
 font.size = 9)

df %>%   
 mutate(across(ends\_with("\_p"), ~gtsummary::style\_pvalue(.x))) %>%   
 flextable(col\_keys = c("weather\_var", "r2", "aic", "sep\_1", "h1\_df", "h1\_x2", "h1\_p", "sep\_2", "h2\_df", "h2\_x2", "h2\_p", "sep\_3", "h3\_df", "h3\_x2", "h3\_p")) %>%   
 colformat\_num(j = 3, big.mark = "") %>%   
 colformat\_num(j = c(4, 7, 10), digits = 3) %>%   
 colformat\_num(j = c(5, 8, 11), digits = 4) %>%   
 add\_header\_row(c("", "DLNM", "", "H1", "", "H2", "", "H3"), colwidths = c(1, 2, 1, 3, 1, 3, 1, 3), top = TRUE) %>%   
 set\_header\_labels(  
 weather\_var = "",  
 r2 = "R^2^",  
 aic = "AIC",  
 h1\_df = "df",  
 h1\_x2 = "X^2^",  
 h1\_p = "P",  
 h2\_df = "df",  
 h2\_x2 = "X^2^",  
 h2\_p = "P",  
 h3\_df = "df",  
 h3\_x2 = "X^2^",  
 h3\_p = "P"  
 ) %>%   
 merge\_v(part = "header") %>%   
 autofit() %>%  
 theme\_booktabs() %>%   
 align(align = "center", part = "header") %>%   
 align(i = 2, align = "right", part = "header") %>%   
 align(align = "right", part = "body")

|  | DLNM | |  | H1 | | |  | H2 | | |  | H3 | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| R^2^ | AIC | df | X^2^ | P | df | X^2^ | P | df | X^2^ | P |
| Minimum daily temp. | 0.28 | 3426 |  | 18.68 | 79 | <0.001 |  | 20 | 80.8 | <0.001 |  | 1.00 | 2.5 | 0.12 |
| Mean daily temp. | 0.29 | 3432 |  | 21.69 | 57 | <0.001 |  | 23 | 78.6 | <0.001 |  | 0.99 | 20.3 | <0.001 |
| Maximum daily temp. | 0.28 | 3435 |  | 19.47 | 47 | <0.001 |  | 20 | 71.6 | <0.001 |  | 0.95 | 20.8 | <0.001 |
| Total daily precip. | 0.29 | 3433 |  | 24.74 | 58 | <0.001 |  | 26 | 82.3 | <0.001 |  | 0.99 | 21.3 | <0.001 |

# Exploration of how edf works

If I constrain the sp parameter, does it constrain the resulting EDF? Mostly…

library(magrittr)  
m\_hopper\_temp\_mean$edf %>%  
 enframe() %>%  
 mutate(name = str\_remove(name, ".\\d+")) %>%  
 group\_by(name) %>%  
 summarize(edf = sum(value)) %T>%  
 print() %>%   
 summarize(total = sum(edf))

## `summarise()` ungrouping output (override with `.groups` argument)

## # A tibble: 6 x 2  
## name edf  
## <chr> <dbl>  
## 1 (Intercept) 1.   
## 2 s(counter) 1.80  
## 3 s(day\_post) 1.01  
## 4 s(interharvest\_id) 4.03  
## 5 s(plant) 13.6   
## 6 s(Q\_temp\_mean,L) 12.6

## # A tibble: 1 x 1  
## total  
## <dbl>  
## 1 34.1

m\_hopper\_temp\_mean1$edf %>%  
 enframe() %>%  
 mutate(name = str\_remove(name, ".\\d+")) %>%  
 group\_by(name) %>%  
 summarize(edf = sum(value)) %T>%  
 print() %>%   
 summarize(total = sum(edf))

## `summarise()` ungrouping output (override with `.groups` argument)

## # A tibble: 6 x 2  
## name edf  
## <chr> <dbl>  
## 1 (Intercept) 1.00  
## 2 mean\_temp\_mean 1.   
## 3 s(counter) 1.81  
## 4 s(day\_post) 1.01  
## 5 s(interharvest\_id) 4.44  
## 6 s(plant) 13.6

## # A tibble: 1 x 1  
## total  
## <dbl>  
## 1 22.9

m\_hopper\_temp\_mean\_null$edf %>%  
 enframe() %>%  
 mutate(name = str\_remove(name, ".\\d+")) %>%  
 group\_by(name) %>%  
 summarize(edf = sum(value)) %T>%  
 print() %>%   
 summarize(total = sum(edf))

## `summarise()` ungrouping output (override with `.groups` argument)

## # A tibble: 5 x 2  
## name edf  
## <chr> <dbl>  
## 1 (Intercept) 1.00  
## 2 s(counter) 1.81  
## 3 s(day\_post) 1.01  
## 4 s(interharvest\_id) 4.58  
## 5 s(plant) 13.6

## # A tibble: 1 x 1  
## total  
## <dbl>  
## 1 22.0

m\_hopper\_temp\_mean3$edf %>%  
 enframe() %>%  
 mutate(name = str\_remove(name, ".\\d+")) %>%  
 group\_by(name) %>%  
 summarize(edf = sum(value)) %T>%  
 print() %>%   
 summarize(total = sum(edf))

## `summarise()` ungrouping output (override with `.groups` argument)

## # A tibble: 6 x 2  
## name edf  
## <chr> <dbl>  
## 1 (Intercept) 1.00  
## 2 pick\_temp\_mean 1.   
## 3 s(counter) 1.81  
## 4 s(day\_post) 1.01  
## 5 s(interharvest\_id) 4.55  
## 6 s(plant) 13.6

## # A tibble: 1 x 1  
## total  
## <dbl>  
## 1 23.0