Canidate Models

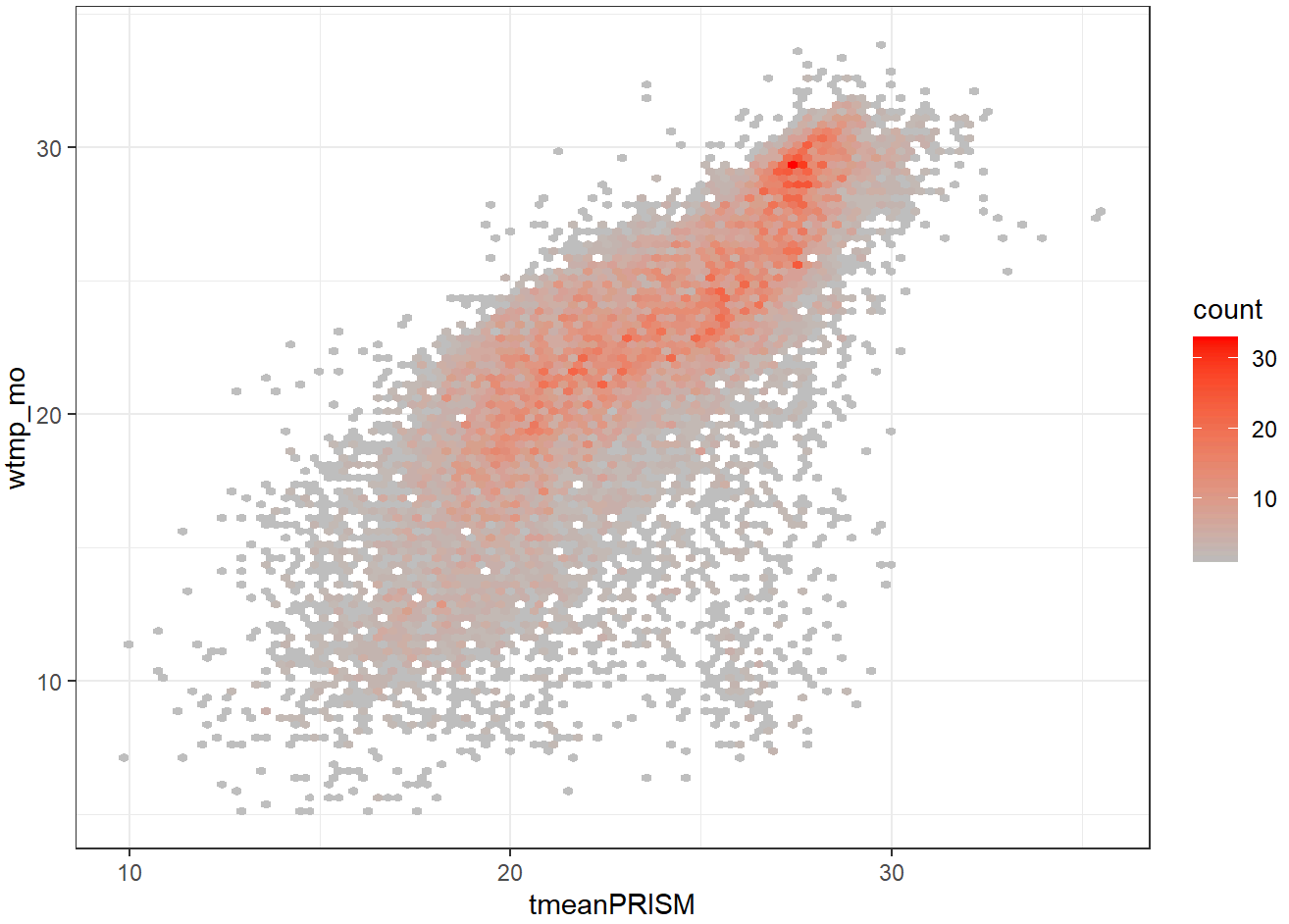
Libraries

```
library(data.table)
library(MASS)
library(lmerTest)
library(car)
library(tidyverse)
library(sf)
```

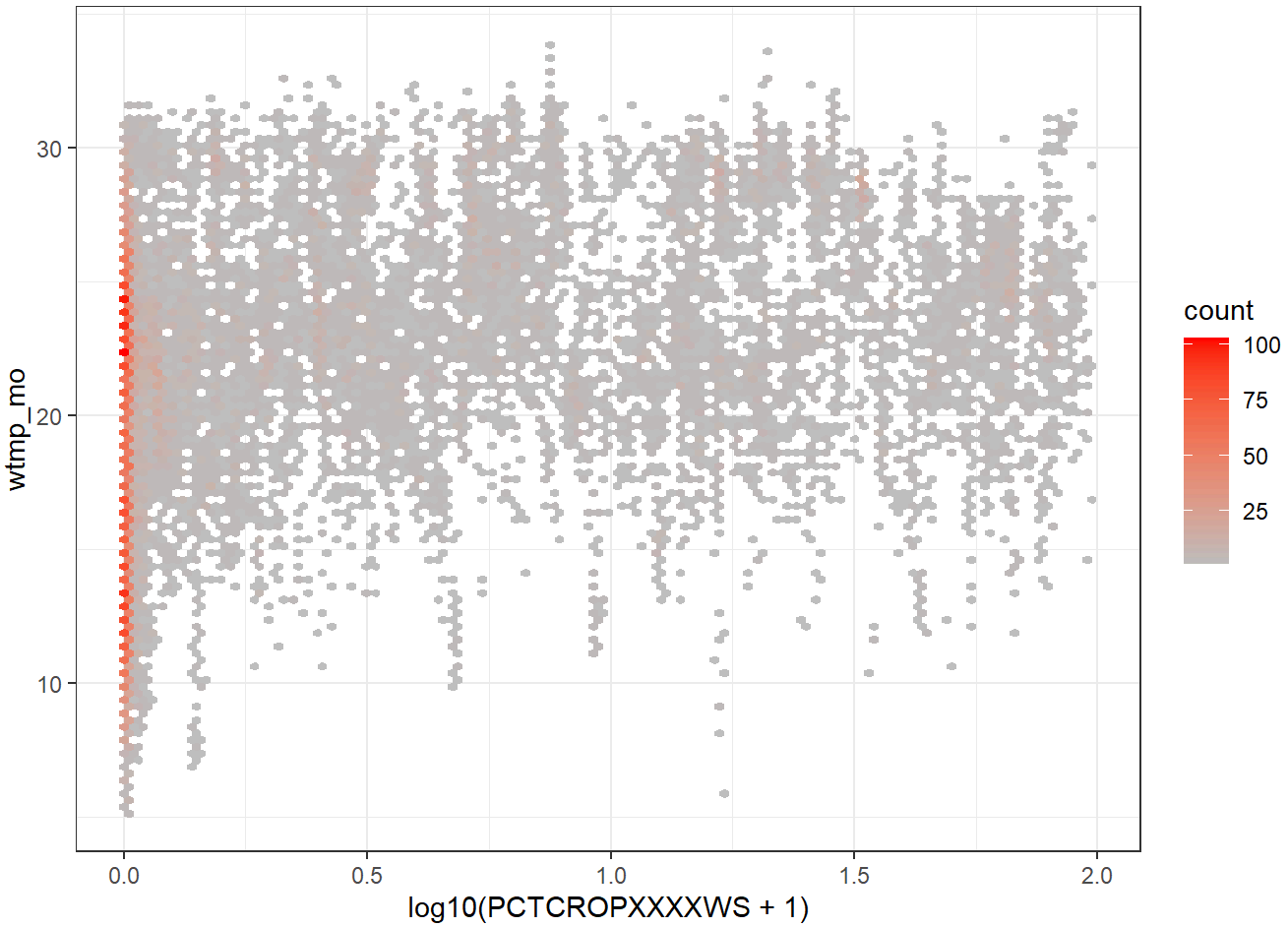
Initial Model Selection

- A fully-expressed formula is defined with multiple potential interactions with air temperature and month.
- Two selection procedures are then used to identify two candidate models. These procedures include:
- Backward selection (lmerTest::step) with a random intercept (COMID) included.
- Backward/Forward selection without a random intercept (COMID) included.
- Candidate models are then saved.

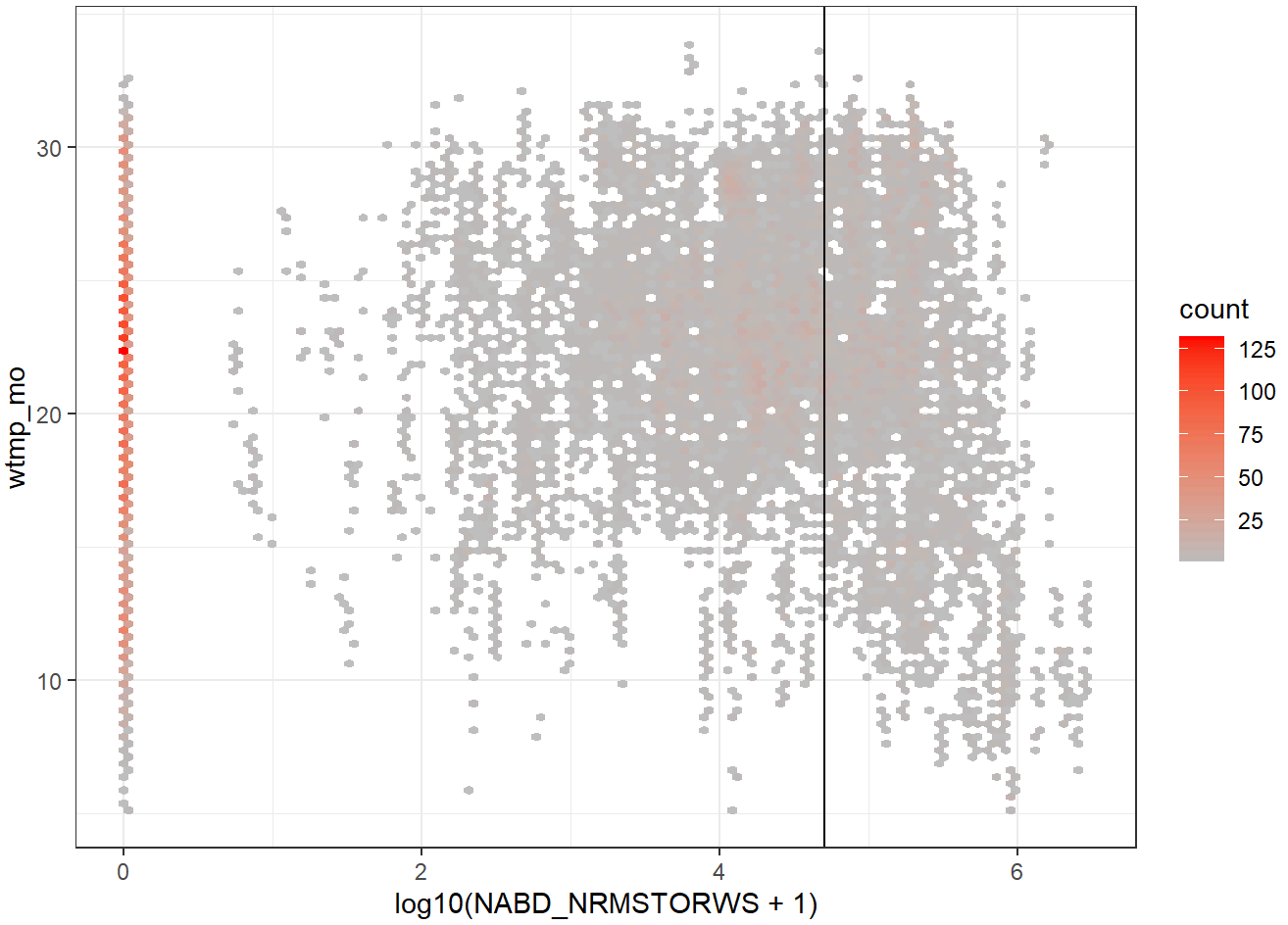
```
st <- read_rds('../data/summer_data.2024.08.08.rds') %>%
  dplyr::mutate(month = as.character(month),
                COMID = as.character(COMID)) %>%
  na.omit() %>%
  dplyr::ungroup() %>%
  st_drop_geometry() %>%
  dplyr::mutate(log10_dam = log10(NABD_NRMSTORWS + 1),
                large_dam = ifelse(log10_dam > 5, 1, 0) %>%
                  as.character(),
                no_dam = ifelse(log10_dam == 0, 1, 0) %>%
                  as.character())

ggplot(data = st,
       aes(x = tmeanPRISM,
           y = wtmp_mo)) +
  geom_hex(bins = 100) +
  scale_fill_gradient(low = "grey", high = "red") +
  theme_bw()
```



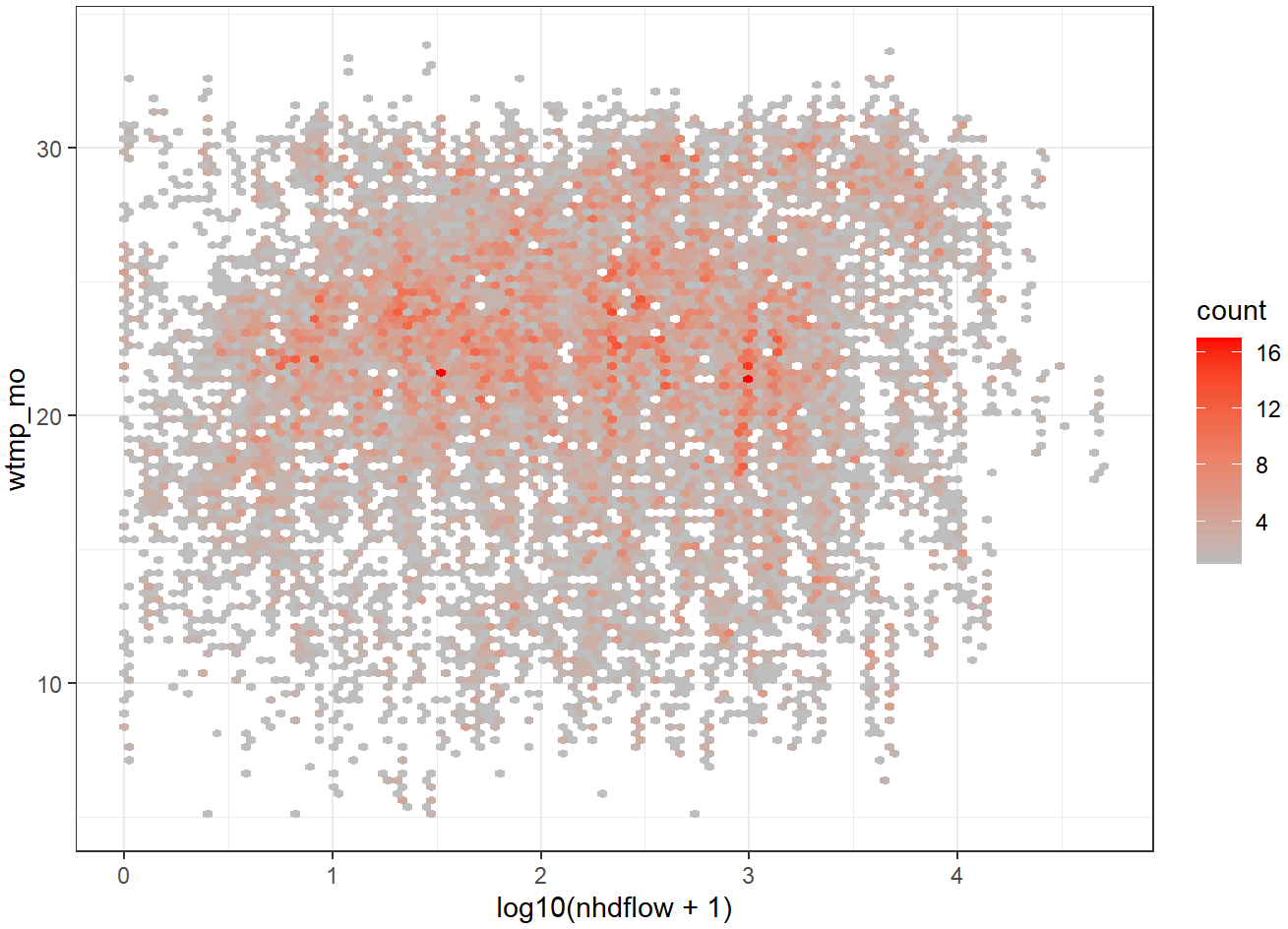

```
ggplot(data = st,
       aes(x = log10(PCTCROPXXXXWS+1),
           y = wtmp_mo)) +
  geom_hex(bins = 100) +
  scale_fill_gradient(low = "grey", high = "red") +
  theme_bw()
```




```
ggplot(data = st,
       aes(x = log10(NABD_NRMSTORWS + 1),
           y = wtmp_mo)) +
  geom_hex(bins = 100) +
  scale_fill_gradient(low = "grey", high = "red") +
  theme_bw() +
  geom_vline(xintercept = 4.7)
```




```
ggplot(data = st,
       aes(x = log10(nhdflow+1),
           y = wtmp_mo)) +
  geom_hex(bins = 100) +
  scale_fill_gradient(low = "grey", high = "red") +
  theme_bw()
```




```
formula <-
  wtmp_mo ~
  tmeanPRISM*month +
  tmeanPRISM*PCTOWXXXXWS*I(log10(NABD_NRMSTORWS+1))*large_dam*month +
  tmeanPRISM*ELEVCAT*month +
  tmeanPRISM*BFIWS*month +
  tmeanPRISM*I(log10(PCTCROPXXXXWS+1))*month +
  tmeanPRISM*WTDEPWS*month +
  tmeanPRISM*PCTFSTXXXXWSRP100*month +
  tmeanPRISM*PCTURBXXXXWS*month +
  tmeanPRISM*SANDWS*month +
  tmeanPRISM*WETINDEXWS*month +
  tmeanPRISM*I(log10(nhdflow+1))*month +
  tmeanPRISM*RUNOFFWS*month +
  tmeanPRISM*CAOWS*month +
  tmeanPRISM*BFIWS*month +
  tmeanPRISM*pptPRISM*month +
  (1|COMID)

initial.lmer <- lmerTest::lmer(formula,
                               data = st)

reduced <- lmerTest::step(initial.lmer)
```

```r
lmer_formula <-
  lmerTest::get_model(reduced) %>%
  formula() %>%
  update(. ~ .
         -(1 | COMID))

initial.lm <-
  lm(formula = update(formula, . ~ . - (1 | COMID)),
     data = st)

initial.stepaic <-
  MASS::stepAIC(initial.lm,
                direction = "both",
                trace = FALSE)

aic_formula <-
  formula(initial.stepaic) %>%
  update(. ~ .
         -(1 | COMID))

formula <-
  formula %>%
  update(. ~ .
         -(1 | COMID))

write_rds(lmer_formula, '../data/base_lmer_formula.rds')
write_rds(aic_formula, '../data/base_formula_stepaic.rds')
write_rds(formula, '../data/fullset_formula.rds')
```