Exercise 3. Data analysis

Sasha D. Hafner

07 March, 2024

Overview

In this exercise you will analyze volatilization of ammonia from field-applied manure to try to infer wheather rain has an effect.

1. Read and check data

Read in the data in the files `NH3_emis_rain_interval.csv` and `NH3_emis_rain_plot.csv` and merge by the field plot key `pmid`. Check the data. The relevant columns are

- `pmid`: field plot key
- `cta`: time after slurry application (h)
- `j__NH3`: ammonia volatilization rate in preceding interval (`cta[i-1]` to `cta[i]`) (kg N / h-ha)
- `rain_rate`: rainfall rate in preceding interval (mm/h)
- `air_temp`: air temperature (deg. C)
- `wind_2m`: wind speed (m/s)
- `app_method`: slurry application method

2. Single experimental unit

Plot the ammonia volatilization rate data from plot `pmid` = 2223. Do you see strong evidence of a rain effect? Focus on 50 h < `cta` < 100 h.

3. Multiple experimental units

Can you think of an approach for estimation and evaluation of an overall rain effect using data from all the plots? This is not simple. See how far you can get.

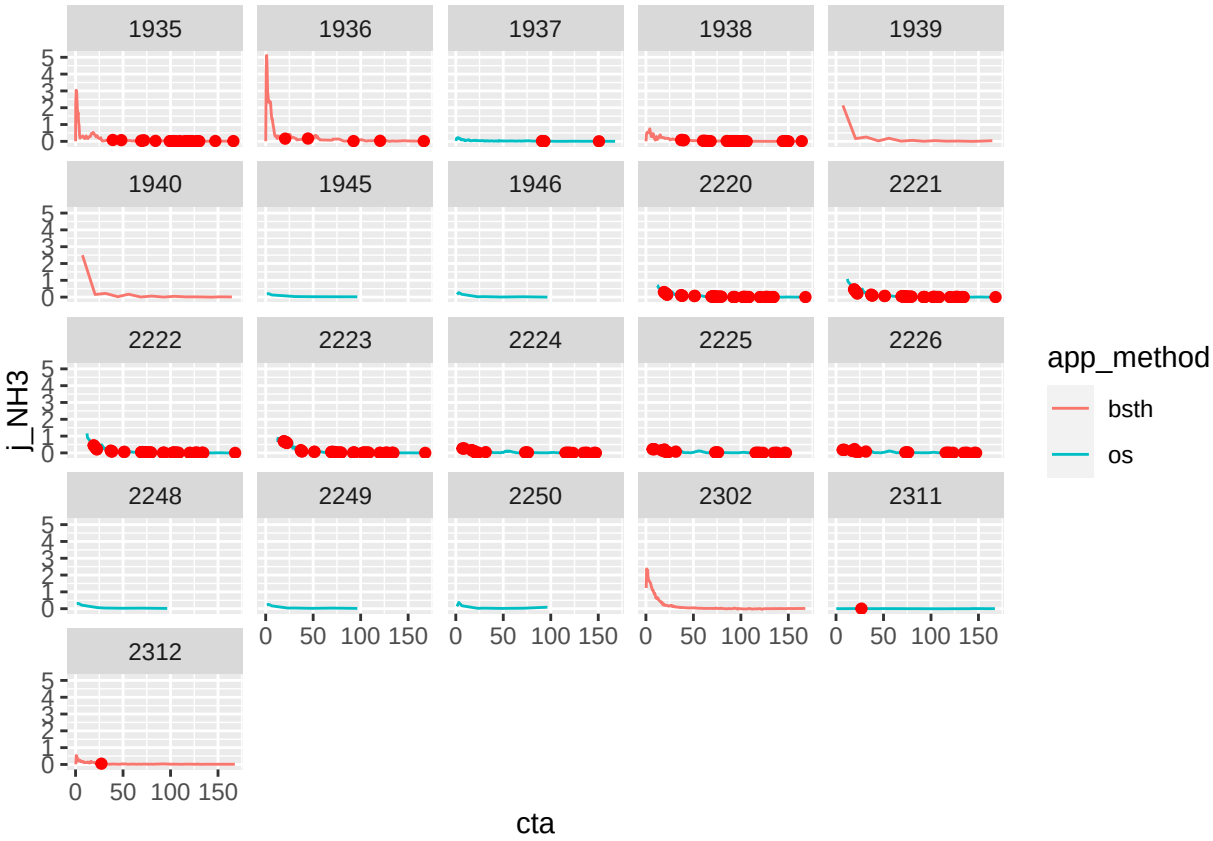
```
library(data.table)
library(ggplot2)
library(viridisLite)
```

```
idat <- fread('../data/NH3_emis_rain_interval.csv')
pdat <- fread('../data/NH3_emis_rain_plot.csv')
idat <- merge(idat, pdat)
```

```
ggplot(idat, aes(cta, j_NH3, colour = app_method, group = pmid)) +
  geom_line() +
  geom_point(data = idat[rain_rate > 0, ], colour = 'red') +
  facet_wrap(~ pmid) +
  xlim(0, 168)
```

```
## Warning: Removed 495 rows containing missing values (`geom_line()`).
```

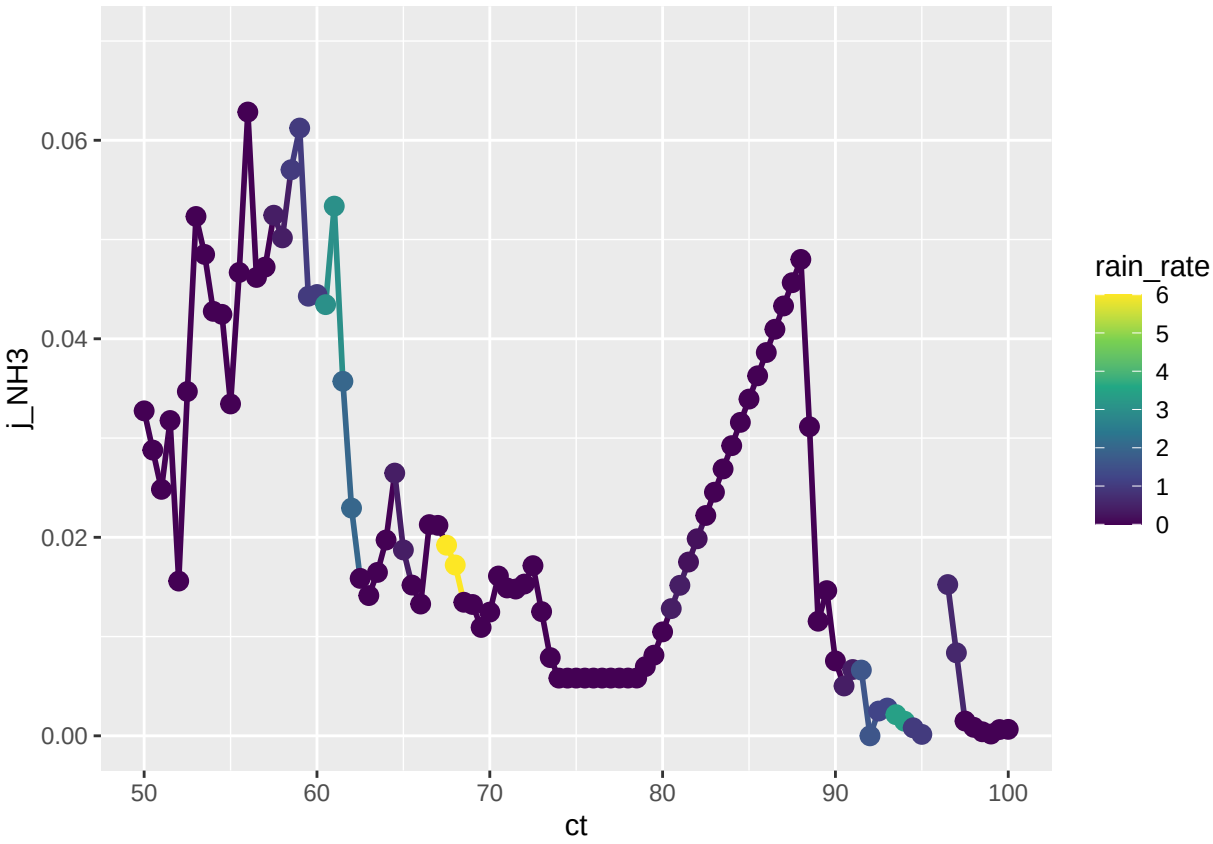
```
## Warning: Removed 25 rows containing missing values (`geom_point()`).
```




```
dd <- idat[pmid == 2223, ]
ggplot(dd, aes(ct, j_NH3, colour = rain_rate, group = pmid)) +
  geom_line(lwd = 1) +
  geom_point(size = 3) +
  xlim(50, 100) +
  ylim(0, 0.07) +
  scale_colour_viridis_c()
```

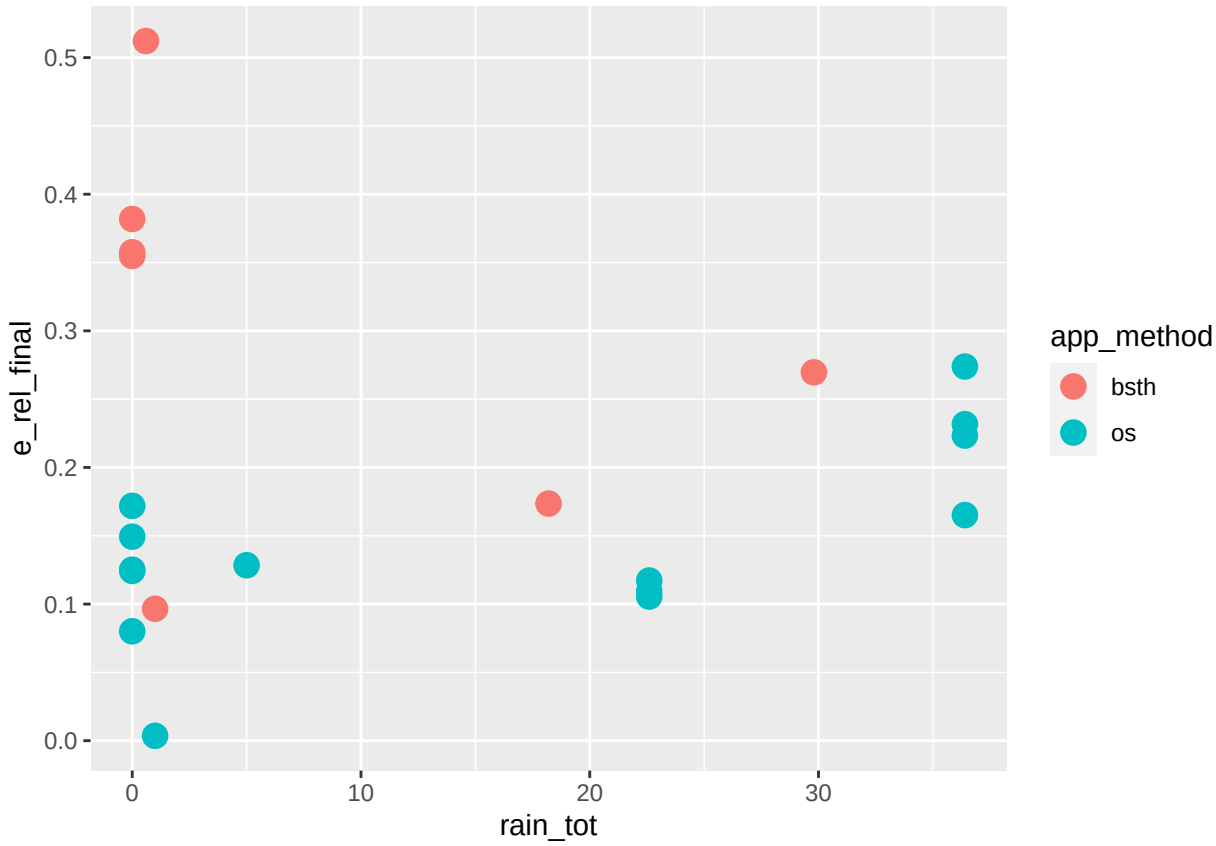
```
## Warning: Removed 220 rows containing missing values (`geom_line()`).
```

```
## Warning: Removed 222 rows containing missing values (`geom_point()`).
```



Look at total emission

```
ggplot(pdat, aes(rain_tot, e_rel_final, colour = app_method)) +
  geom_point(size = 4)
```




```
mod1 <- lm(e_rel_final ~ rain_tot * app_method, data = pdat)
summary(mod1)
```

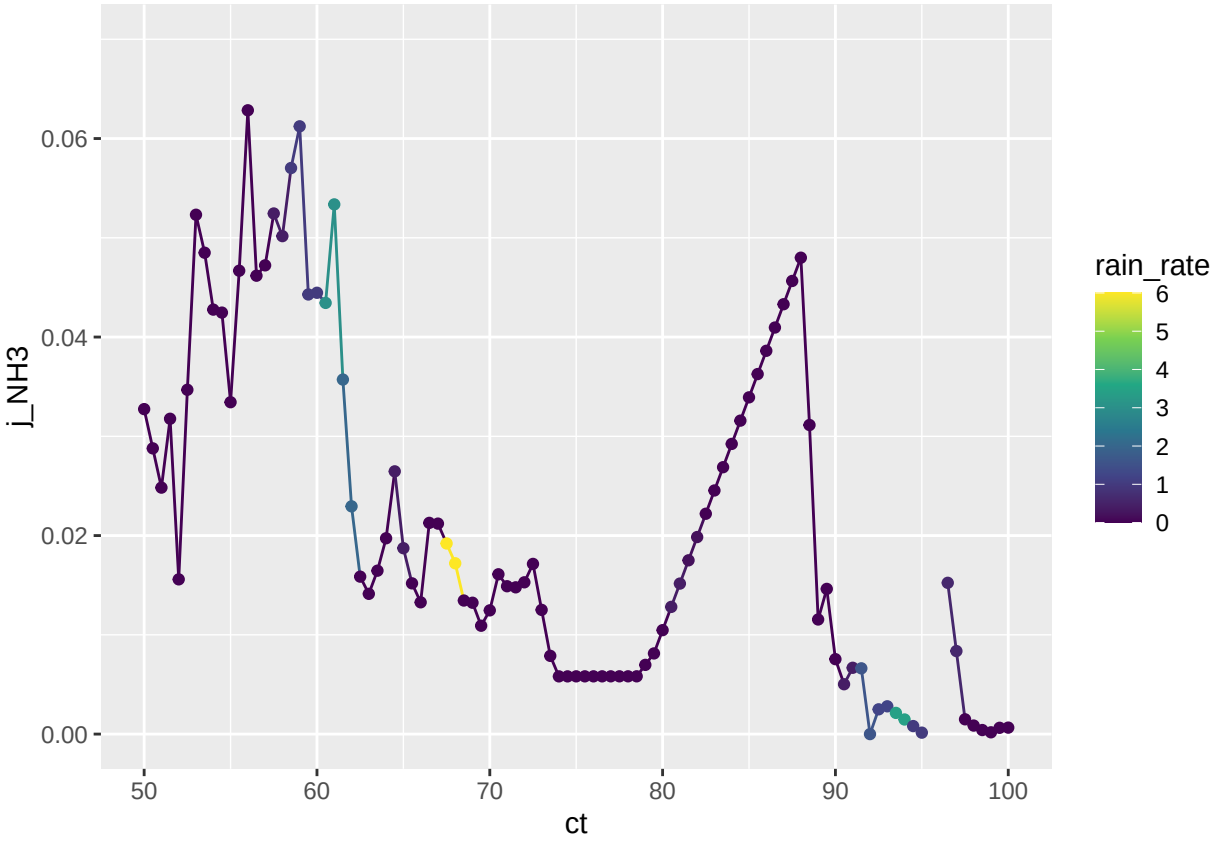
```
##
## Call:
## lm(formula = e_rel_final ~ rain_tot * app_method, data = pdat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.23531 -0.04488  0.02152  0.04587  0.17861
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)              0.336017   0.040138   8.371 1.96e-07 ***
## rain_tot                -0.004153   0.003040  -1.366 0.189686
## app_methodos            -0.234732   0.052549  -4.467 0.000339 ***
## rain_tot:app_methodos    0.006846   0.003404   2.011 0.060444 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08961 on 17 degrees of freedom
## Multiple R-squared:  0.5456, Adjusted R-squared:  0.4654
## F-statistic: 6.805 on 3 and 17 DF,  p-value: 0.003236
```

```
ggplot(dd, aes(ct, j_NH3, colour = rain_rate, group = pmid)) +
  geom_line() +
  geom_point() +
  xlim(50, 100) +
```

```
  ylim(0, 0.07) +
  scale_colour_viridis_c()
```

```
## Warning: Removed 220 rows containing missing values (`geom_line()`).
```

```
## Warning: Removed 222 rows containing missing values (`geom_point()`).
```




```
ggplot(idat, aes(cta, log10(j_NH3), colour = app_method, group = pmid)) +
  geom_line() +
  geom_point(data = idat[rain_rate > 0, ], colour = 'orange') +
  facet_wrap(~ pmid) +
  xlim(0, 168)
```

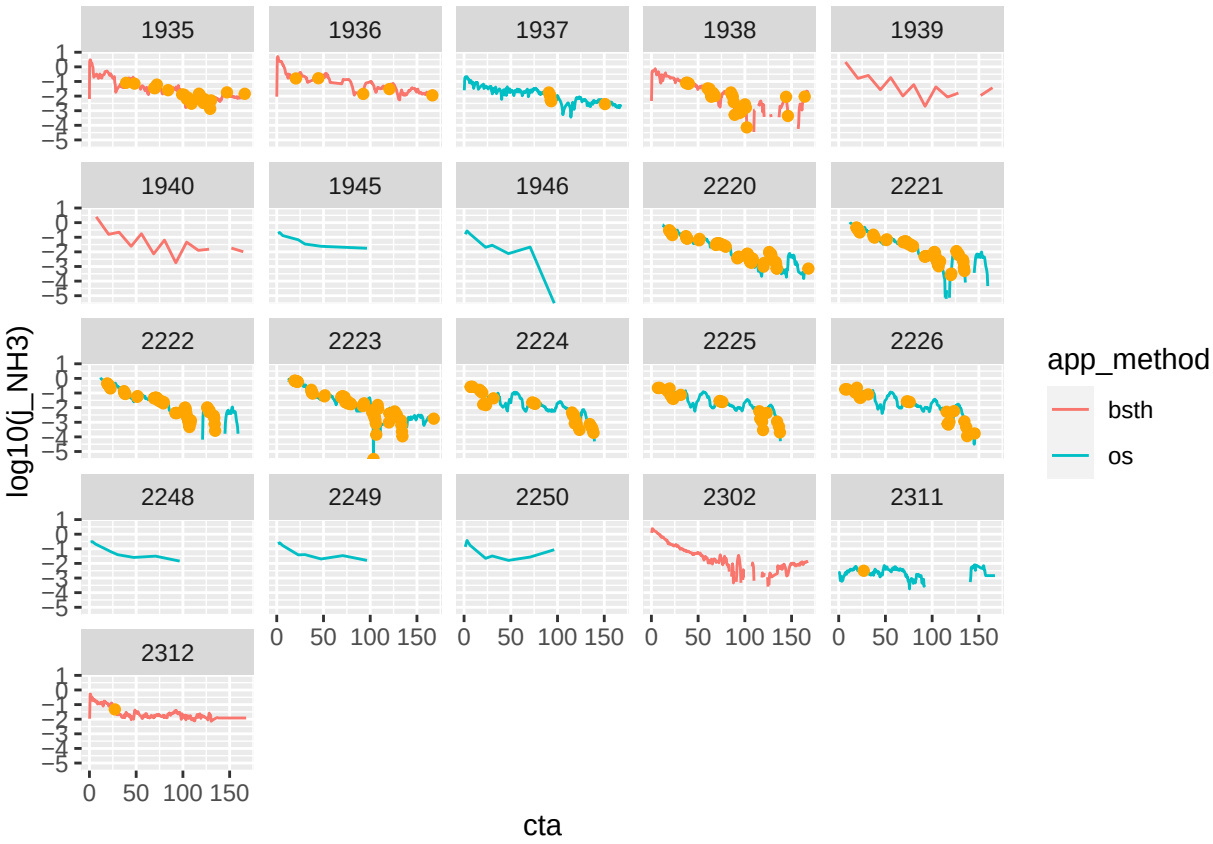
```
## Warning in FUN(X[[i]], ...): NaNs produced
```

```
## Warning in FUN(X[[i]], ...): NaNs produced
```

```
## Warning in FUN(X[[i]], ...): NaNs produced
```

```
## Warning: Removed 581 rows containing missing values (`geom_line()`).
```

```
## Warning: Removed 59 rows containing missing values (`geom_point()`).
```




```
idatsub <- idat[!is.na(j_NH3 + cta + rain + rain_rate + rain_cum + air_temp + wind_2m) & j_NH3 > 0, ]
mod1 <- lm(log10(j_NH3) ~ poly(cta, 3):app_method + (rain_cum + air_temp + wind_2m):factor(pmid) + app_m
summary(mod1)
```

```
##
## Call:
## lm(formula = log10(j_NH3) ~ poly(cta, 3):app_method + (rain_cum +
##     air_temp + wind_2m):factor(pmid) + app_method, data = idatsub)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.94570 -0.17174  0.00573  0.19493  1.72268
##
## Coefficients: (1 not defined because of singularities)
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    -1.737681   0.038356 -45.304  < 2e-16 ***
## app_methodos                   -0.737088   0.069817 -10.557  < 2e-16 ***
## poly(cta, 3)1:app_methodbsth  -18.794005   1.148120 -16.369  < 2e-16 ***
## poly(cta, 3)2:app_methodbsth   -2.508166   0.526028  -4.768 1.92e-06 ***
## poly(cta, 3)3:app_methodbsth   12.995775   0.532481  24.406  < 2e-16 ***
## poly(cta, 3)1:app_methodos    -17.836884   1.813934  -9.833  < 2e-16 ***
## poly(cta, 3)2:app_methodos     22.084612   1.308201  16.882  < 2e-16 ***
## poly(cta, 3)3:app_methodos      4.675644   1.675964   2.790 0.005297 **
## rain_cum:factor(pmid)1935      -0.012881   0.002514  -5.123 3.14e-07 ***
## rain_cum:factor(pmid)1936      -1.913692   0.142448 -13.434  < 2e-16 ***
## rain_cum:factor(pmid)1937      -0.062478   0.012179  -5.130 3.03e-07 ***
## rain_cum:factor(pmid)1938      -0.030866   0.004639  -6.654 3.21e-11 ***
```

```
## rain_cum:factor(pmid)2220     -0.039743   0.001957 -20.308  < 2e-16 ***
## rain_cum:factor(pmid)2221     -0.042377   0.002151 -19.699  < 2e-16 ***
## rain_cum:factor(pmid)2222     -0.035672   0.002249 -15.863  < 2e-16 ***
## rain_cum:factor(pmid)2223     -0.036985   0.001957 -18.901  < 2e-16 ***
## rain_cum:factor(pmid)2224     -0.049652   0.005446  -9.118  < 2e-16 ***
## rain_cum:factor(pmid)2225     -0.036950   0.005627  -6.566 5.77e-11 ***
## rain_cum:factor(pmid)2226     -0.039753   0.005600  -7.099 1.46e-12 ***
## rain_cum:factor(pmid)2302           NA         NA      NA       NA
## rain_cum:factor(pmid)2311      0.489762   0.059856   8.182 3.62e-16 ***
## rain_cum:factor(pmid)2312     -0.249559   0.052046  -4.795 1.68e-06 ***
## air_temp:factor(pmid)1935     -0.004965   0.003593  -1.382 0.167019
## air_temp:factor(pmid)1936      0.042282   0.004097  10.320  < 2e-16 ***
## air_temp:factor(pmid)1937      0.038181   0.007755   4.923 8.82e-07 ***
## air_temp:factor(pmid)1938      0.065584   0.011847   5.536 3.28e-08 ***
## air_temp:factor(pmid)2220      0.067900   0.006300  10.778  < 2e-16 ***
## air_temp:factor(pmid)2221      0.071691   0.006342  11.305  < 2e-16 ***
## air_temp:factor(pmid)2222      0.070688   0.006330  11.167  < 2e-16 ***
## air_temp:factor(pmid)2223      0.094956   0.006304  15.064  < 2e-16 ***
## air_temp:factor(pmid)2224      0.066014   0.008788   7.511 7.05e-14 ***
## air_temp:factor(pmid)2225      0.056311   0.008819   6.386 1.89e-10 ***
## air_temp:factor(pmid)2226      0.050892   0.008847   5.752 9.41e-09 ***
## air_temp:factor(pmid)2302      0.226144   0.014899  15.179  < 2e-16 ***
## air_temp:factor(pmid)2311     -0.078441   0.015425  -5.085 3.83e-07 ***
## air_temp:factor(pmid)2312     -0.045878   0.012442  -3.687 0.000229 ***
## wind_2m:factor(pmid)1935       0.161255   0.018770   8.591  < 2e-16 ***
## wind_2m:factor(pmid)1936       0.230236   0.016190  14.221  < 2e-16 ***
## wind_2m:factor(pmid)1937       0.239079   0.024496   9.760  < 2e-16 ***
## wind_2m:factor(pmid)1938       0.033164   0.019659   1.687 0.091684 .
## wind_2m:factor(pmid)2220       0.104324   0.021369   4.882 1.09e-06 ***
## wind_2m:factor(pmid)2221       0.118537   0.021727   5.456 5.15e-08 ***
## wind_2m:factor(pmid)2222       0.094063   0.021919   4.291 1.81e-05 ***
## wind_2m:factor(pmid)2223      -0.015017   0.021414  -0.701 0.483157
## wind_2m:factor(pmid)2224       0.191574   0.021863   8.763  < 2e-16 ***
## wind_2m:factor(pmid)2225       0.200575   0.022636   8.861  < 2e-16 ***
## wind_2m:factor(pmid)2226       0.221101   0.022767   9.711  < 2e-16 ***
## wind_2m:factor(pmid)2302      -0.109711   0.017931  -6.119 1.03e-09 ***
## wind_2m:factor(pmid)2311      -0.080203   0.036595  -2.192 0.028459 *
## wind_2m:factor(pmid)2312       0.192112   0.024086   7.976 1.92e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3736 on 4362 degrees of freedom
## Multiple R-squared:  0.7709, Adjusted R-squared:  0.7684
## F-statistic: 305.8 on 48 and 4362 DF,  p-value: < 2.2e-16
```

```
cc <- coef(mod1)
cr <- cc[grepl('rain', names(cc))]
ct <- cc[grepl('temp', names(cc))]
cr
```

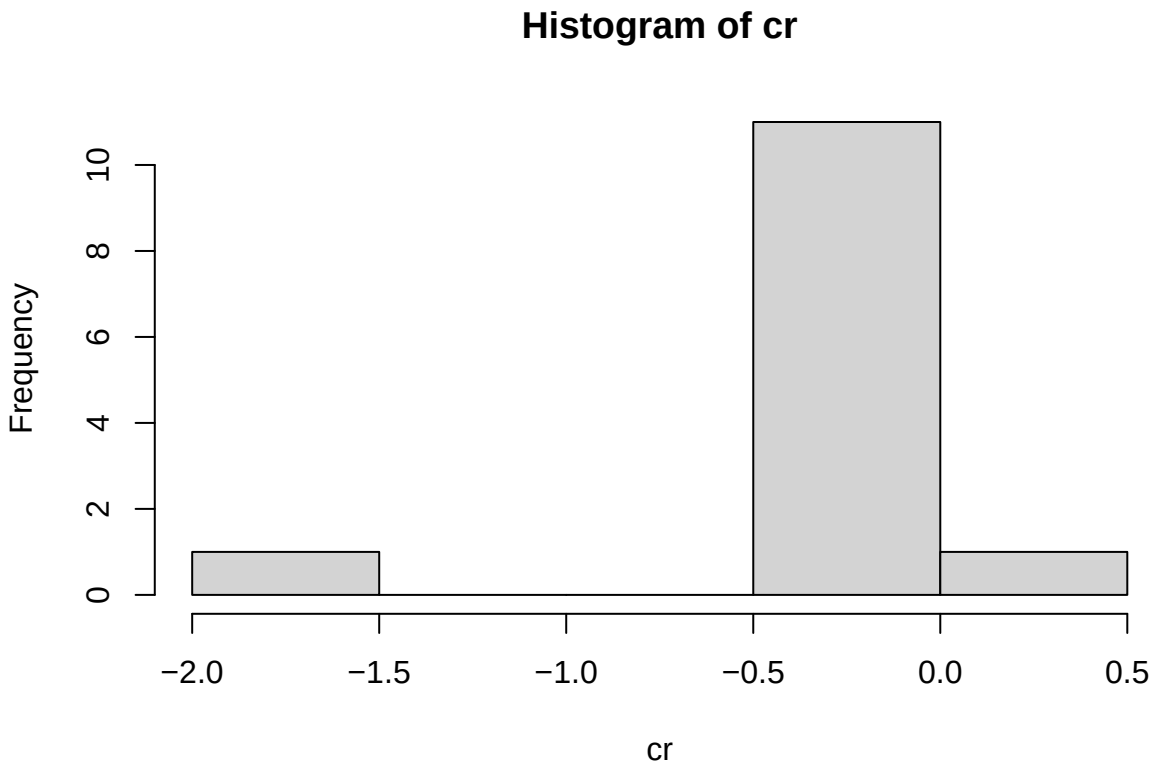
```
## rain_cum:factor(pmid)1935 rain_cum:factor(pmid)1936 rain_cum:factor(pmid)1937 rain_cum:factor(pmid)19
##               -0.01288061               -1.91369216               -0.06247797               -0.030865
## rain_cum:factor(pmid)2222 rain_cum:factor(pmid)2223 rain_cum:factor(pmid)2224 rain_cum:factor(pmid)22
##               -0.03567169               -0.03698524               -0.04965162               -0.036950
## rain_cum:factor(pmid)2311 rain_cum:factor(pmid)2312
```

```
##                0.48976151                -0.24955932
```

```
ct
```

```
## air_temp:factor(pmid)1935 air_temp:factor(pmid)1936 air_temp:factor(pmid)1937 air_temp:factor(pmid)19
##              -0.004965467                0.042282148                0.038181488                0.065583
## air_temp:factor(pmid)2222 air_temp:factor(pmid)2223 air_temp:factor(pmid)2224 air_temp:factor(pmid)22
##               0.070687751                0.094955685                0.066014331                0.056310
## air_temp:factor(pmid)2311 air_temp:factor(pmid)2312
##              -0.078441272               -0.045877936
```

```
idatsub[, j_pred := predict(mod1)]
hist(cr)
```

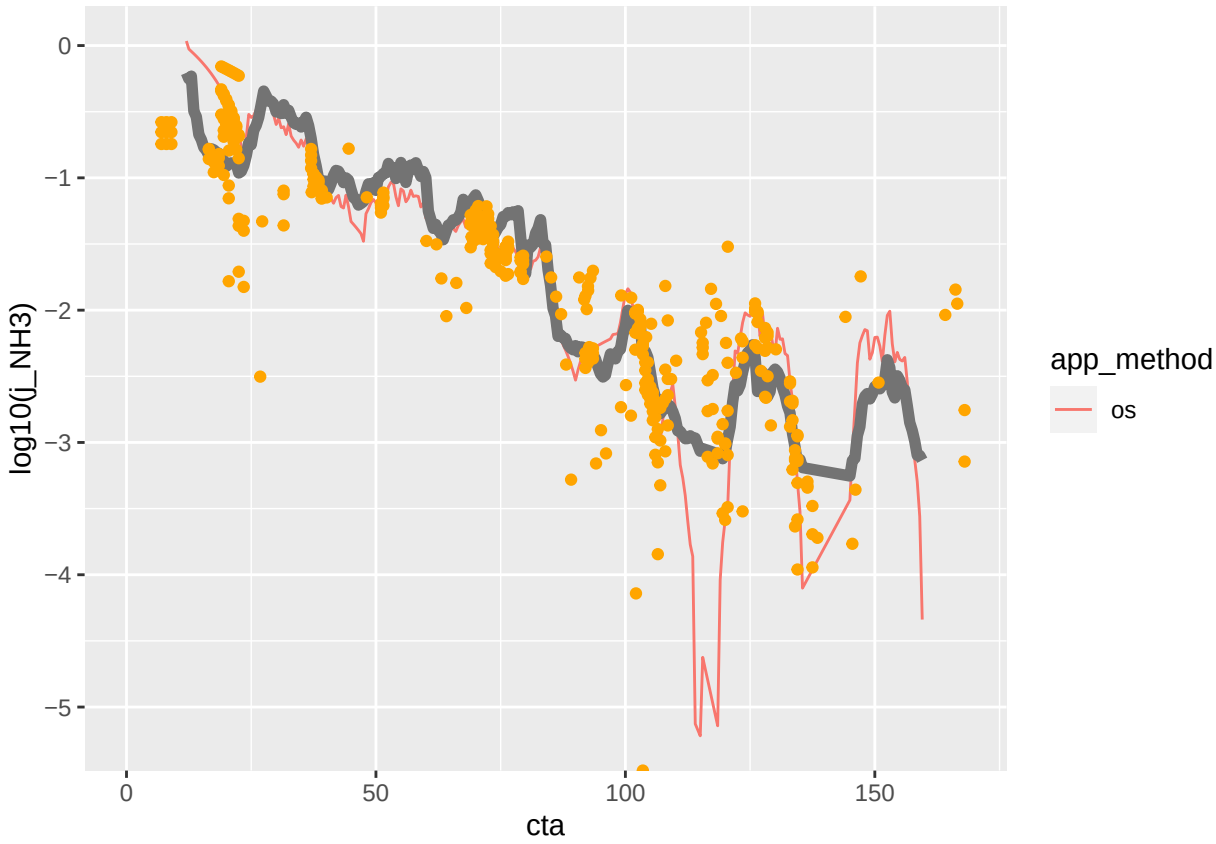



```
dd <- idatsub[pmid == 2221, ]
ggplot(dd, aes(cta, log10(j_NH3), colour = app_method, group = pmid)) +
  geom_line() +
  geom_line(aes(y = j_pred), colour = 'gray45', lwd = 2) +
  geom_point(data = idat[rain_rate > 0, ], colour = 'orange') +
  xlim(0, 168)
```

```
## Warning in FUN(X[[i]], ...): NaNs produced
```

```
## Warning: Removed 5 rows containing missing values (`geom_line()`).
## Removed 5 rows containing missing values (`geom_line()`).
```

```
## Warning: Removed 59 rows containing missing values (`geom_point()`).
```




```
t.test(cr)
```

```
##
##  One Sample t-test
##
## data:  cr
## t = -1.0358, df = 12, p-value = 0.3207
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.4919905  0.1749370
## sample estimates:
##  mean of x
## -0.1585267
```

```
t.test(ct)
```

```
##
##  One Sample t-test
##
## data:  ct
## t = 2.7545, df = 13, p-value = 0.0164
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.01111379 0.09193722
## sample estimates:
## mean of x
## 0.0515255
```

```r
binom.test(table(cr < 0))
```

```
##
##  Exact binomial test
##
## data:  table(cr < 0)
## number of successes = 1, number of trials = 13, p-value = 0.003418
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.001945628 0.360297435
## sample estimates:
## probability of success
##             0.07692308
```

```r
binom.test(table(ct < 0))
```

```
##
##  Exact binomial test
##
## data:  table(ct < 0)
## number of successes = 11, number of trials = 14, p-value = 0.05737
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.4920243 0.9534207
## sample estimates:
## probability of success
##              0.7857143
```