

# Stats to parse out DM and pH effects

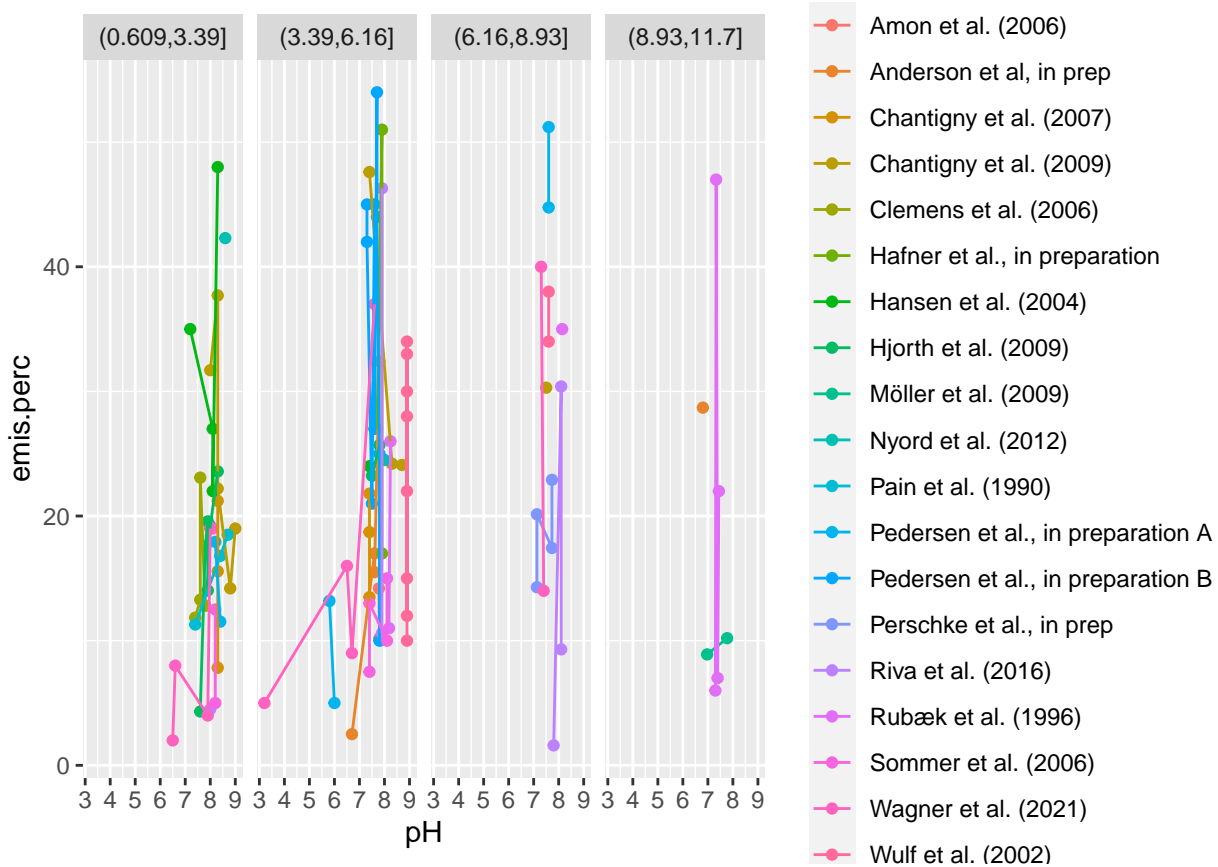
Sasha D. Hafner

30 January, 2023

## Plots

```
dat$DMgrp <- cut(dat$DM, 4)
dat$pHgrp <- cut(dat$pH, 4)
dd <- subset(dat, !is.na(pH) & !is.na(DM))
```

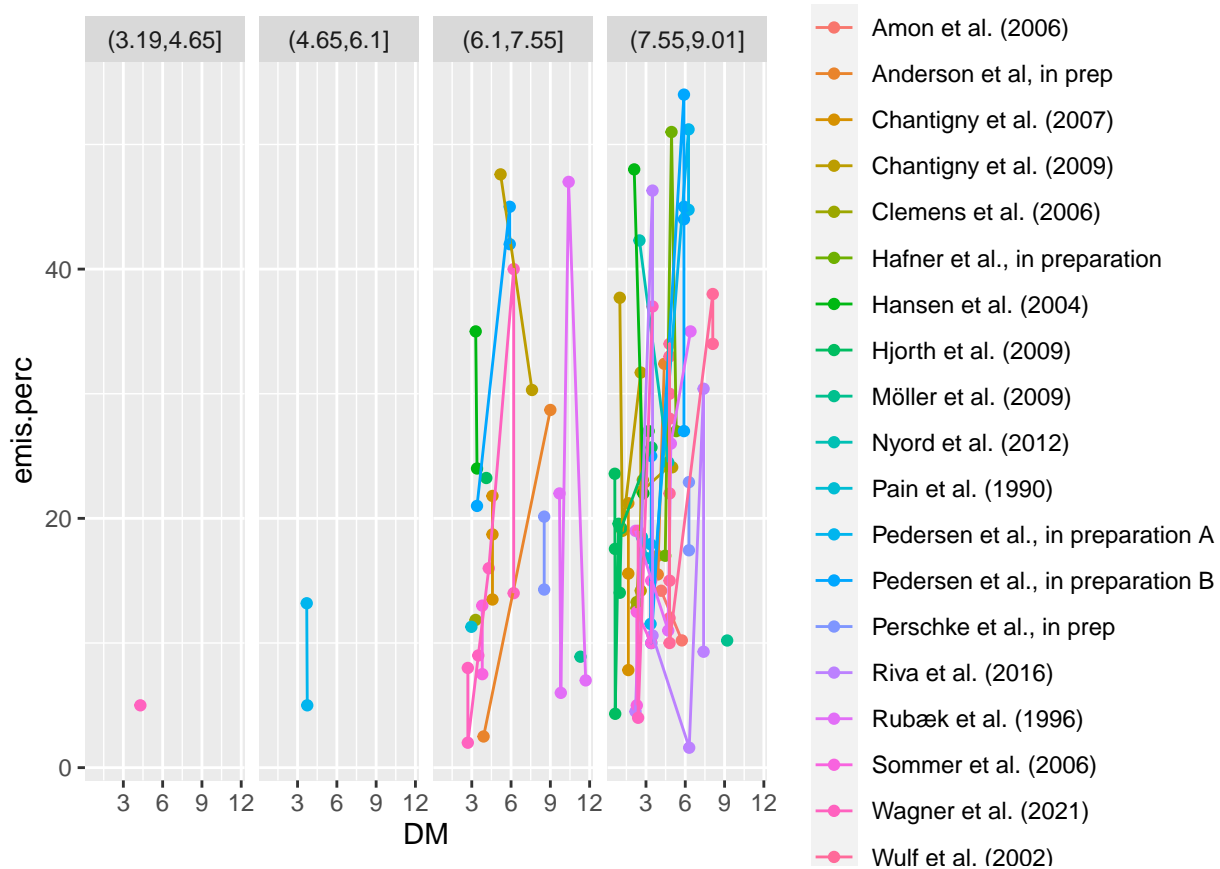
```
ggplot(dd, aes(pH, emis.perc, colour = source)) +
  geom_point() +
  geom_line() +
  facet_wrap(~ DMgrp, ncol = 4)
```



```
ggplot(dd, aes(DM, emis.perc, colour = source)) +
  geom_point() +
  geom_line() +
```

```
facet_wrap(~ pHgrp, ncol = 4)
```

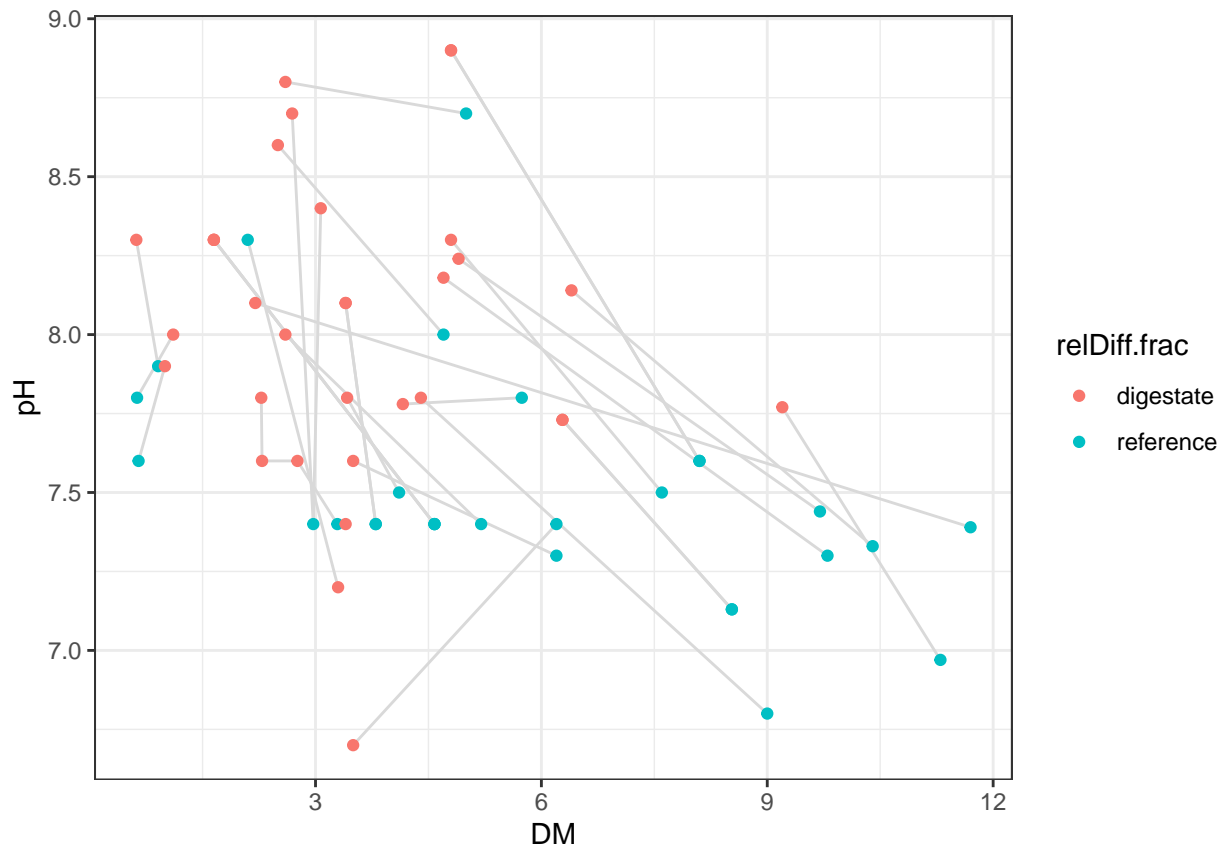
```
## `geom_line()`: Each group consists of only one observation.  
## i Do you need to adjust the group aesthetic?
```



Not pretty but there do seem to be correlations for both DM and pH.

How did digestion change DM and pH?

```
dd <- subset(dd, !is.na(relDiff.set))  
ggplot(dd, aes(DM, pH, colour = relDiff.frac, group = interaction(source, relDiff.set))) +  
  geom_line(colour = 'gray85') +  
  geom_point() +  
  theme_bw()
```



Generally increased pH and decreased DM, as expected in the absence of added high DM substrates.

## Models

Summary.

```
dfsummd(dat[, c('source', 'app.meth', 'DM', 'pH', 'relDiff.frac', 'emis.perc')])
```

```
##
## 122 rows and 6 columns
## 122 unique rows
##
##          source          app.meth      DM      pH
## Class          character      character numeric numeric
## Minimum      Amon et al. (2006) broadcast 0.62      3.2
## Maximum      Wulf et al. (2002) trailing shoe + harrowing 11.7      9
## Mean          <NA>            <NA>      4.39      7.76
## Unique (excl. NA)      21      6      71      36
## Missing values      0      7      2      14
## Sorted          FALSE          FALSE  FALSE  FALSE
##
##          relDiff.frac emis.perc
## Class          character      numeric
## Minimum      digestate      1.6
## Maximum      reference      60
## Mean          <NA>      22.7
## Unique (excl. NA)      2      98
## Missing values      46      0
```

```
## Sorted                FALSE    FALSE
##
```

Overall digestion effect (not expecting much)

```
m1 <- lm(emis.perc ~ source + app.meth + relDiff.frac, data = dat)
summary.aov(m1)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## source         14   4828    344.9     3.98 0.000137 ***
## app.meth         2   1324    662.1     7.64 0.001252 **
## relDiff.frac     1     63     63.2     0.73 0.397011
## Residuals       51   4420     86.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 53 observations deleted due to missingness
```

```
tail(coef(m1))
```

```
## sourceWagner et al. (2021)    sourceWulf et al. (2002)
##              12.798830              20.298830
## app.methopen slot injection    app.methtrailing hose
##              -13.750000              8.000000
##      app.methtrailing shoe    relDiff.fracreference
##              NA              1.933273
```

Nothing.

Separate DM and pH effects.

```
m2 <- lm(emis.perc ~ pH + DM + relDiff.frac + source + app.meth, data = dat)
summary.aov(m2)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## pH              1    67.9     67.9    0.914 0.34486
## DM              1   346.3    346.3    4.661 0.03692 *
## relDiff.frac    1     1.8     1.8    0.024 0.87754
## source          12 2495.8    208.0    2.799 0.00724 **
## app.meth         1   913.2    913.2   12.289 0.00114 **
## Residuals       40 2972.6     74.3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 65 observations deleted due to missingness
```

```
drop1(m2, test = 'F')
```

```
## Single term deletions
```

```
##
```

```
## Model:
```

```
## emis.perc ~ pH + DM + relDiff.frac + source + app.meth
##              Df Sum of Sq    RSS    AIC F value    Pr(>F)
## <none>                2972.6 259.38
## pH                    1      2.11 2974.7 257.43  0.0284 0.867126
## DM                    1      1.59 2974.2 257.42  0.0213 0.884604
## relDiff.frac          1      2.86 2975.4 257.44  0.0385 0.845358
## source                10    2404.38 5377.0 273.17  3.2354 0.003866 **
## app.meth               1      913.25 3885.8 272.66 12.2890 0.001139 **
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
m3 <- lm(emis.perc ~ pH + DM + source + app.meth, data = dat)
summary.aov(m3)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## pH              1     463    463.0    4.694 0.033286 *
## DM              1    1312   1312.4   13.305 0.000473 ***
## source         16    4701    293.8    2.979 0.000678 ***
## app.meth        5    3023    604.6    6.130 7.55e-05 ***
## Residuals      79    7792     98.6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 19 observations deleted due to missingness
```

```
drop1(m3, test = 'F')
```

```
## Single term deletions
##
## Model:
## emis.perc ~ pH + DM + source + app.meth
##              Df Sum of Sq    RSS    AIC F value    Pr(>F)
## <none>                7792.1 493.59
## pH              1      529.9  8322.0 498.37  5.3722 0.0230494 *
## DM              1      981.9  8774.1 503.82  9.9554 0.0022695 **
## source         16     4917.5 12709.6 511.98  3.1160 0.0004117 ***
## app.meth        5     3023.1 10815.2 517.36  6.1300 7.548e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cc <- coef(m3)[1:3]
cc
```

```
## (Intercept)          pH          DM
## -25.475142    4.205796    2.109639
```

```
confint(m3)[1:3, ]
```

```
##              2.5 %    97.5 %
## (Intercept) -63.6388860 12.688603
## pH           0.5939860  7.817605
## DM           0.7787866  3.440491
```

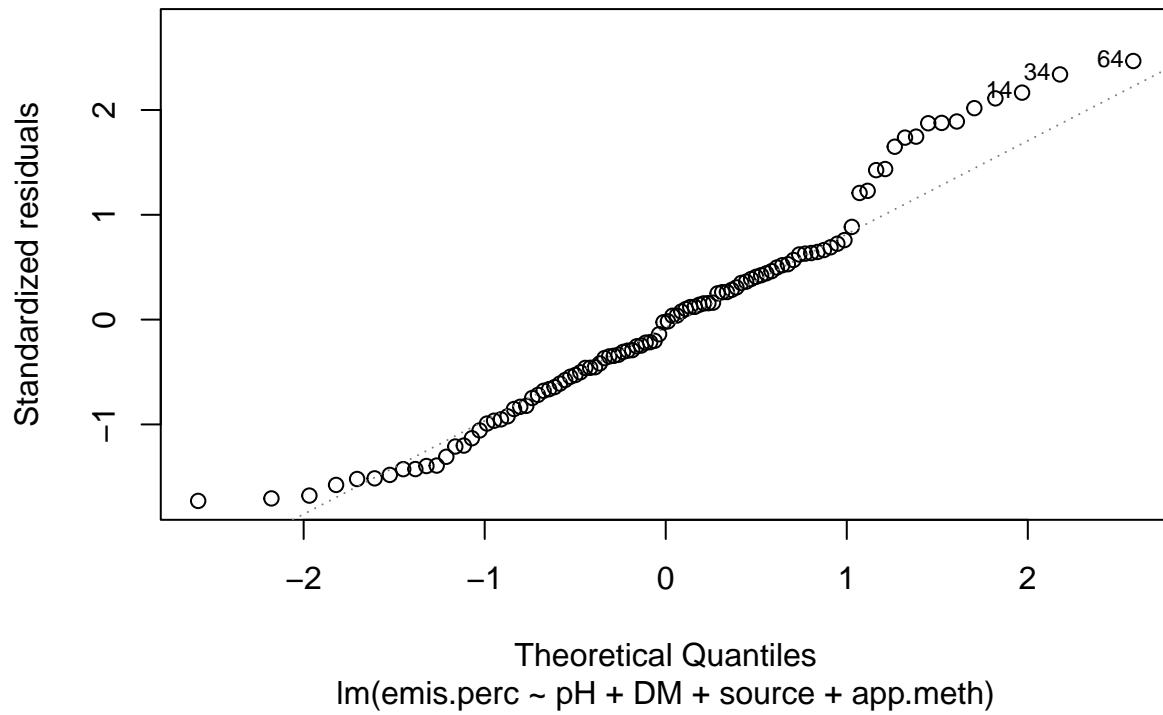
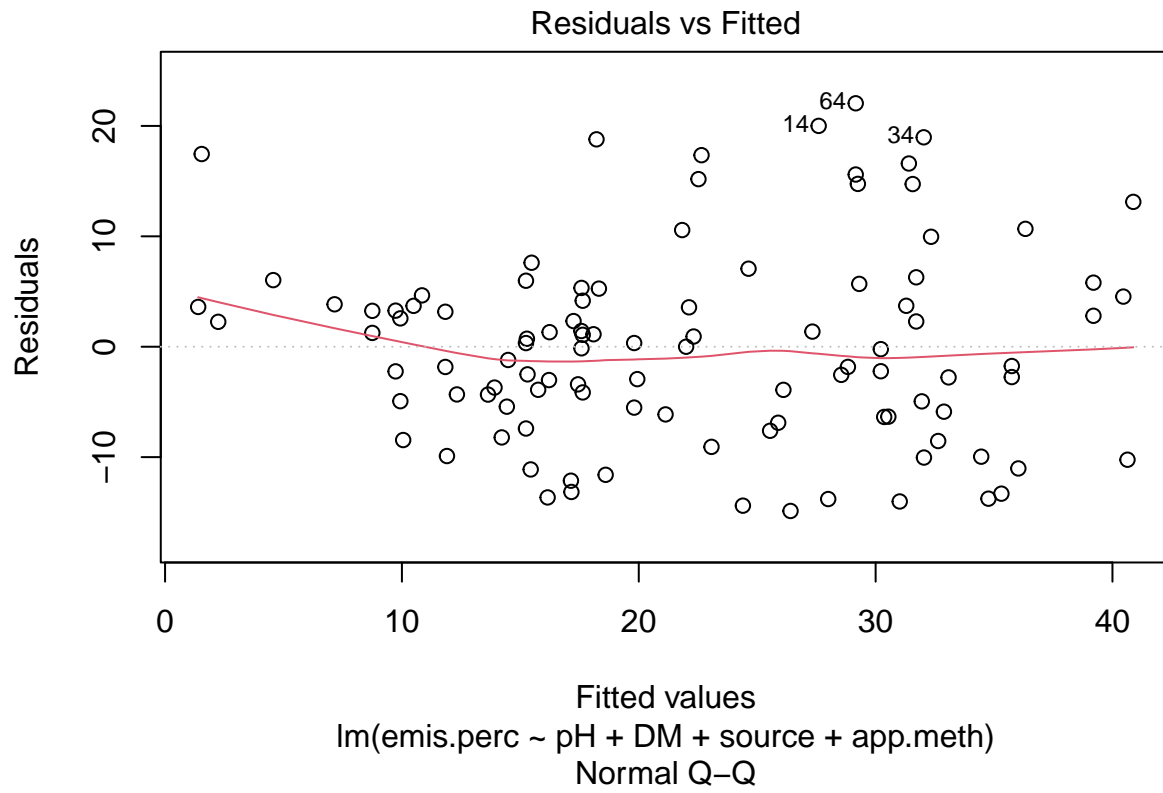
```
cc['pH'] / cc['DM']
```

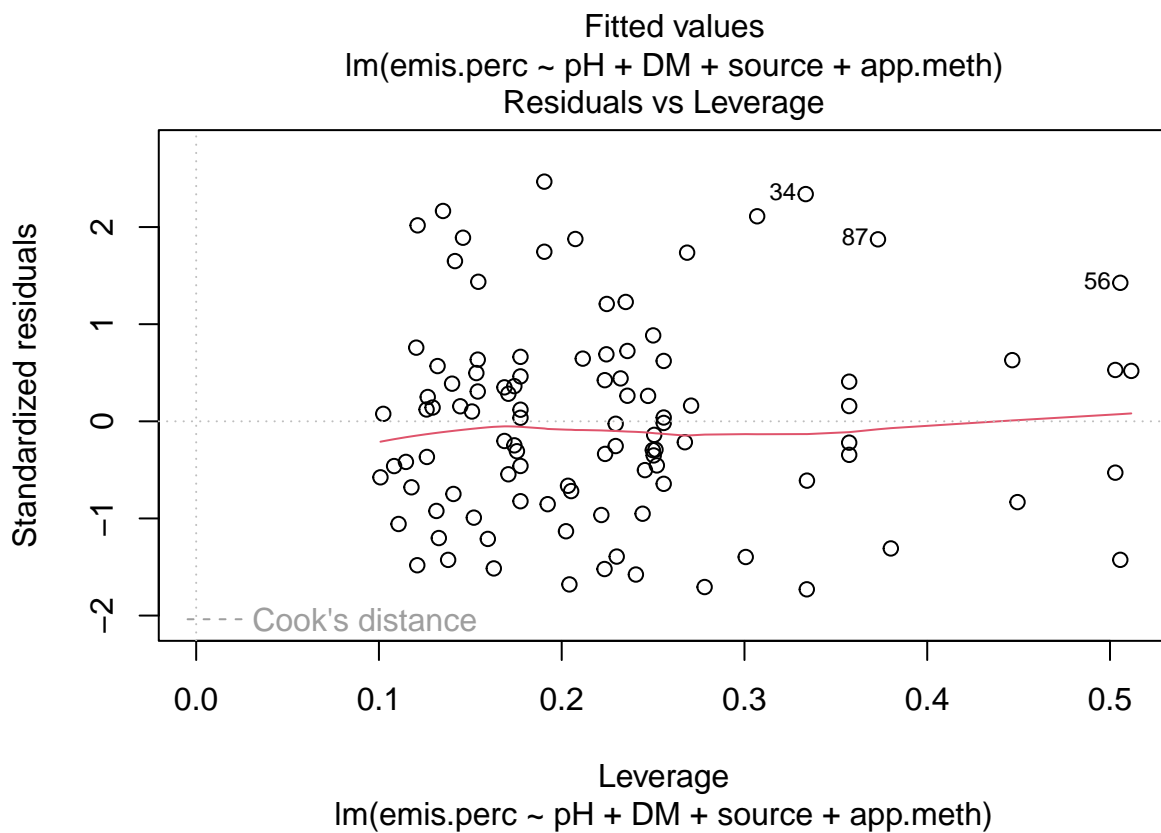
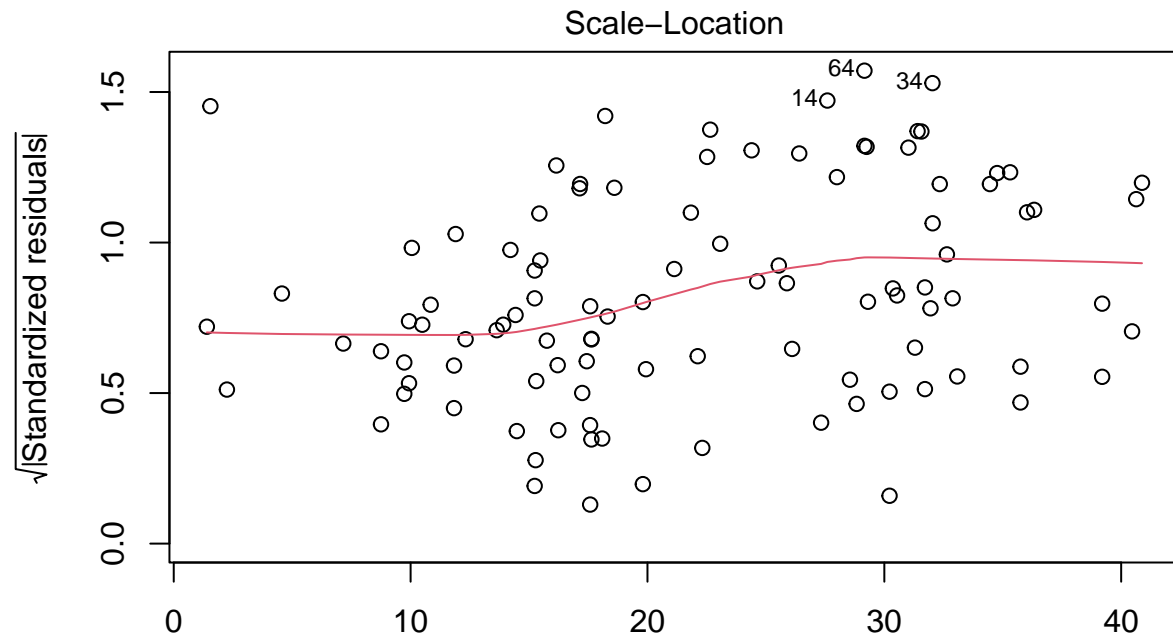
```
##          pH
## 1.993609
```

m3 is where it gets interesting. Clear positive effects of DM and pH, although the pH effect seems small (+4% emission (frac. of TAN) per unit pH change). Effect of pH is about 2x DM effect (1 unit pH change : 1% DM change).

```
plot(m3, ask = FALSE)
```

```
## Warning: not plotting observations with leverage one:
## 102
```





We might look at relative effect.

```
m4 <- lm(log10(emis.perc) ~ pH + DM + source + app.meth, data = dat)
summary.aov(m4)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## pH         1  0.726   0.7257  12.252 0.000767 ***
```

```
## DM          1  0.574  0.5744   9.698 0.002570 **
## source      16  2.545  0.1591   2.686 0.001971 **
## app.meth     5  1.689  0.3377   5.702 0.000153 ***
## Residuals   79  4.679  0.0592
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 19 observations deleted due to missingness
```

```
cc <- 100 * (10^coef(m4)[1:3] - 1)
cc
```

```
## (Intercept)      pH      DM
## -24.47896    38.27664   11.86561
```

```
100 * (10^confint(m4)[1:3, ] - 1)
```

```
##           2.5 %    97.5 %
## (Intercept) -91.232586 550.52562
## pH          12.782575  69.53354
## DM           3.772978  20.58934
```

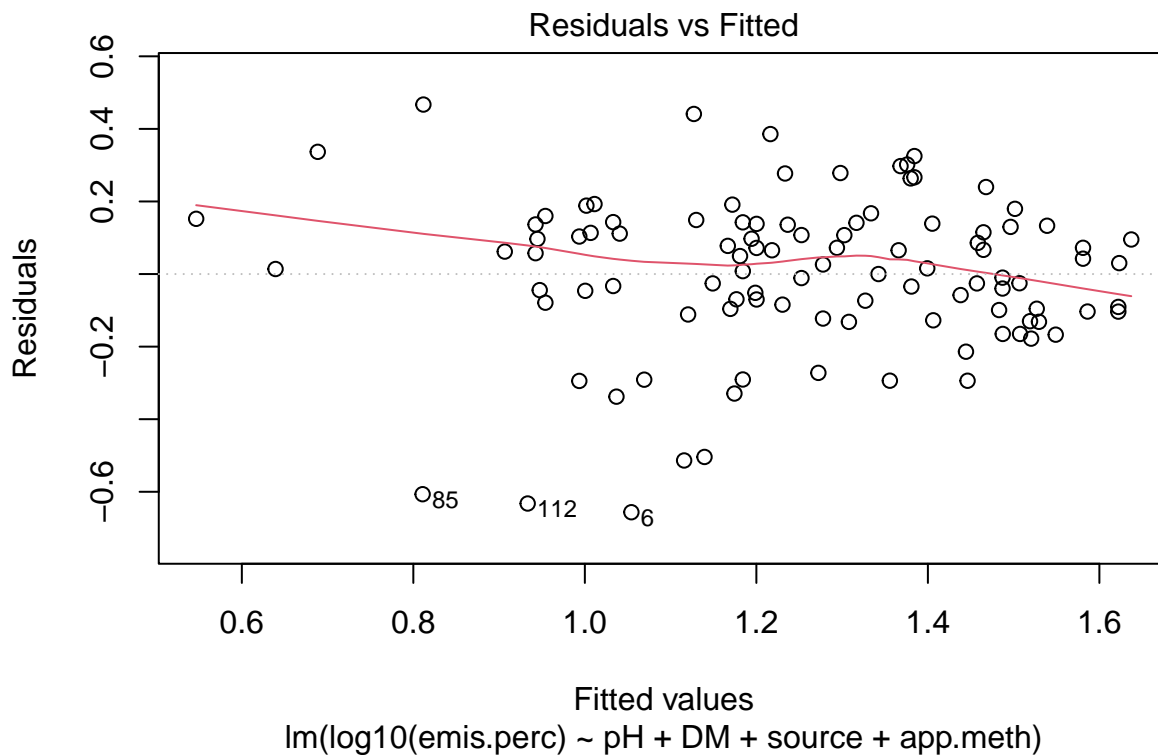
```
cc['pH'] / cc['DM']
```

```
##      pH
## 3.225847
```

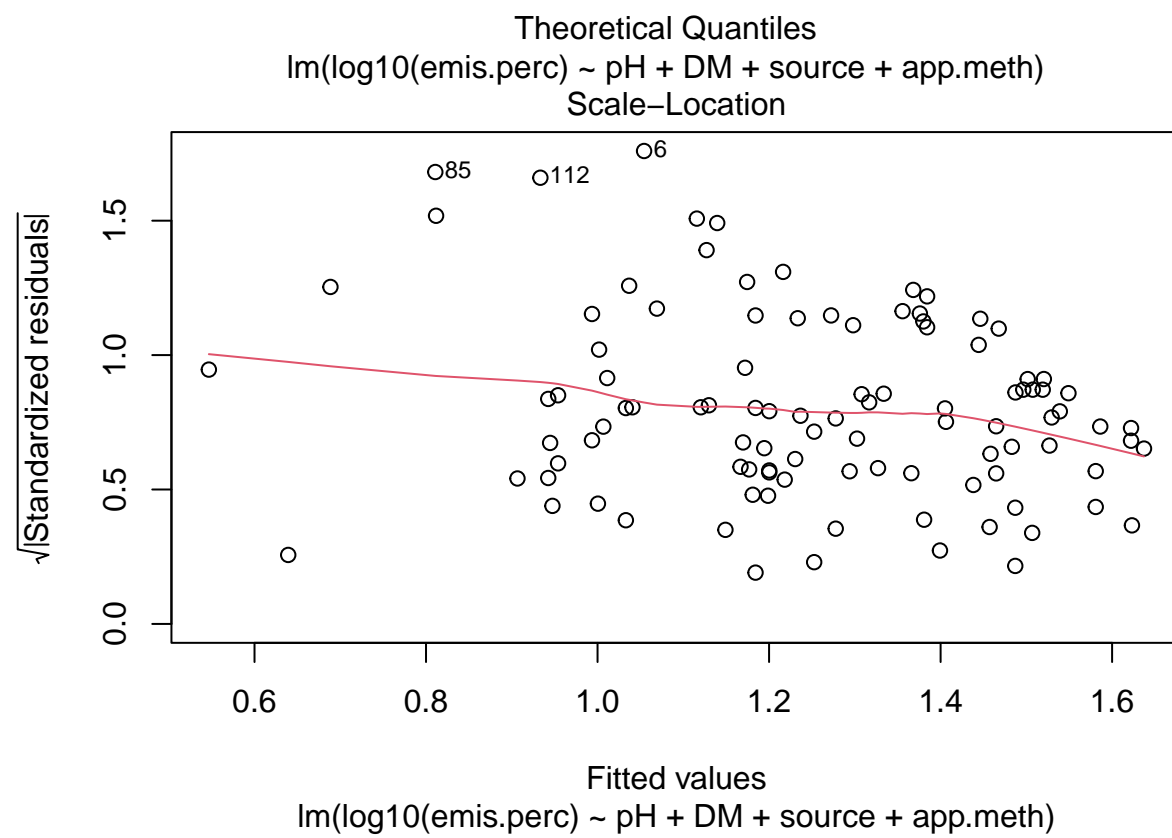
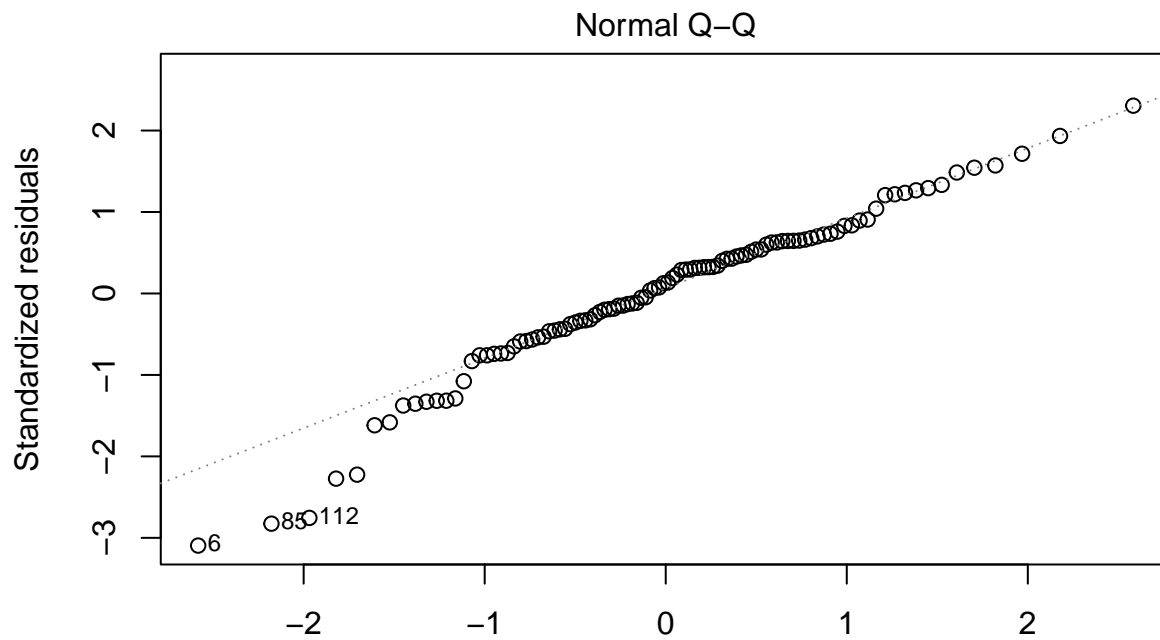
38% increase in emission with a 1 unit increase in pH, seems plausible. Effect of pH here 3x DM effect. Seems more plausible.

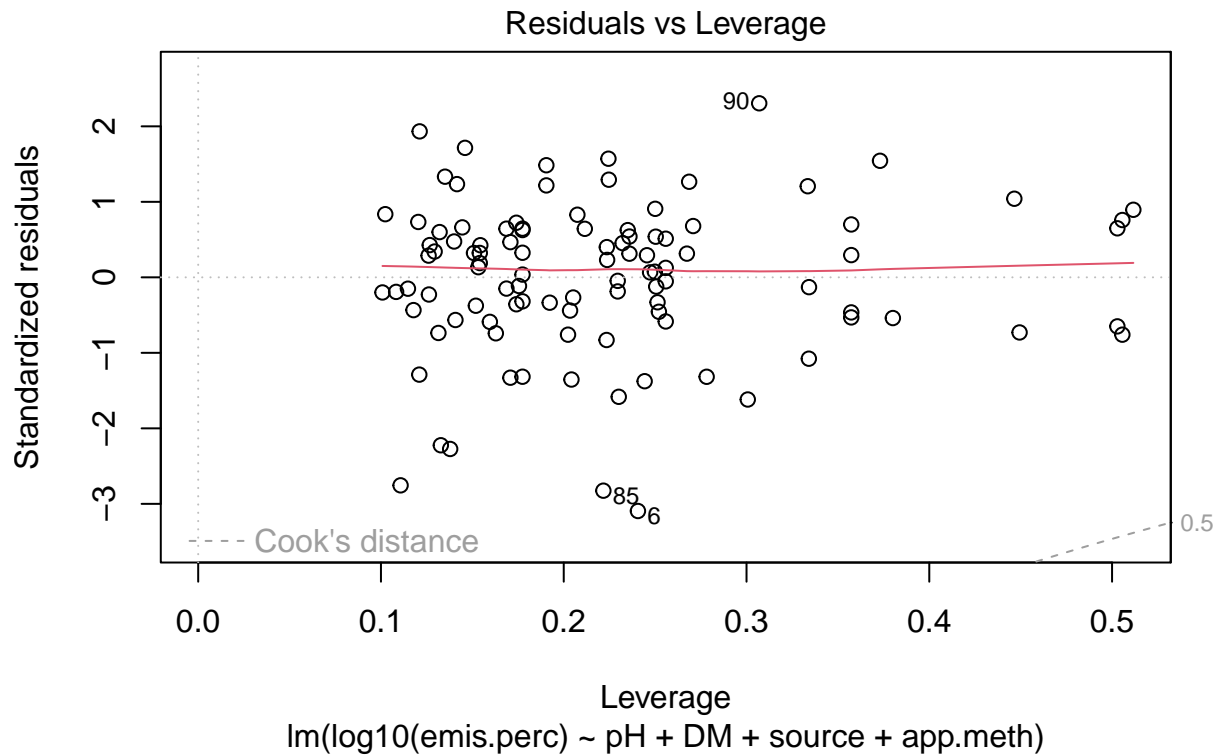
```
plot(m4, ask = FALSE)
```

```
## Warning: not plotting observations with leverage one:
## 102
```









## Mixed-effects model

As Johanna pointed out, `source` is really a random effect.

```
dfsumm(dd[, c('pH', 'DM', 'app.meth', 'source')])
```

```
##
## 62 rows and 4 columns
## 52 unique rows
##           pH      DM      app.meth      source
## Class      numeric numeric      character      character
## Minimum      6.7      0.62      broadcast Amon et al. (2006)
## Maximum      8.9      11.7      trailing shoe Wulf et al. (2002)
## Mean         7.79      4.63      <NA>      <NA>
## Unique (excl. NA) 28      46      4      15
## Missing values  0      0      5      0
## Sorted        FALSE    FALSE    FALSE    TRUE
##
```

```
dsub <- subset(dat, !is.na(pH) & !is.na(DM) & !is.na(app.meth))
m5 <- lmer(log10(emis.perc) ~ pH + DM + app.meth + (1|source), data = dsub)
```

```
summary(m5)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: log10(emis.perc) ~ pH + DM + app.meth + (1 | source)
## Data: dsub
##
## REML criterion at convergence: 35.5
##
```

```
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.69602 -0.37275  0.00472  0.58206  1.67428
##
## Random effects:
##   Groups   Name      Variance Std.Dev.
##   source   (Intercept) 0.01438  0.1199
##   Residual                0.05913  0.2432
## Number of obs: 103, groups: source, 17
##
## Fixed effects:
##                                     Estimate Std. Error t value
## (Intercept)                       -0.02950    0.35422  -0.083
## pH                                0.15097    0.03999   3.775
## DM                                0.04846    0.01355   3.575
## app.methclosed slot injection     -0.63499    0.13515  -4.699
## app.methopen slot injection       -0.35685    0.14158  -2.520
## app.methtrailing hose             -0.06463    0.09715  -0.665
## app.methtrailing shoe             -0.05203    0.12593  -0.413
## app.methtrailing shoe + harrowing -0.20991    0.26252  -0.800
##
## Correlation of Fixed Effects:
##      (Intr) pH      DM      app.mthcsi app.mthpsi app.mh app.ms
## pH      -0.961
## DM      -0.405  0.281
## app.mthclsi -0.089 -0.017 -0.034
## app.mthpnsi -0.210  0.088 -0.079  0.280
## app.mthtrlh -0.375  0.194 -0.040  0.412      0.575
## app.mthtrls -0.281  0.134  0.018  0.317      0.539      0.608
## app.mthts+h -0.024 -0.038 -0.015  0.184      0.160      0.240  0.187
```

```
ccm <- 100 * (10^fixef(m5)[1:3] - 1)
ccm
```

```
## (Intercept)      pH      DM
##   -6.566185   41.569348   11.805280
```

```
100 * (10^confint(m5)[4:5, ] - 1)
```

```
## Computing profile confidence intervals ...
```

```
##      2.5 %   97.5 %
## pH 19.334246 69.27394
## DM  5.430987 18.50035
```

```
ccm['pH'] / ccm['DM']
```

```
##      pH
## 3.521251
```

Large  $t$  statistics and similar coefficients as classical linear model `m4`. Compare:

```
cc
```

```
## (Intercept)      pH      DM
##   -24.47896   38.27664   11.86561
```

```
ccm
```

```
## (Intercept)      pH      DM
```

```
## -6.566185 41.569348 11.805280
```

But, let's try likelihood ratio test (I think that's what `anova()` is doing here).

```
m0 <- lmer(log10(emis.perc) ~ app.meth + (1|source), data = dsub)
m0pH <- lmer(log10(emis.perc) ~ DM + app.meth + (1|source), data = dsub)
m0DM <- lmer(log10(emis.perc) ~ pH + app.meth + (1|source), data = dsub)
anova(m5, m0pH)
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: dsub
```

```
## Models:
```

```
## m0pH: log10(emis.perc) ~ DM + app.meth + (1 | source)
```

```
## m5: log10(emis.perc) ~ pH + DM + app.meth + (1 | source)
```

```
##      npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
```

```
## m0pH    9 39.919 63.631 -10.9594  21.9189
```

```
## m5     10 27.293 53.641  -3.6466   7.2932 14.626  1 0.0001311 ***
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(m5, m0DM)
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: dsub
```

```
## Models:
```

```
## m0DM: log10(emis.perc) ~ pH + app.meth + (1 | source)
```

```
## m5: log10(emis.perc) ~ pH + DM + app.meth + (1 | source)
```

```
##      npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
```

```
## m0DM    9 38.512 62.225 -10.2561  20.5123
```

```
## m5     10 27.293 53.641  -3.6466   7.2932 13.219  1 0.0002771 ***
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(m5, m0)
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: dsub
```

```
## Models:
```

```
## m0: log10(emis.perc) ~ app.meth + (1 | source)
```

```
## m5: log10(emis.perc) ~ pH + DM + app.meth + (1 | source)
```

```
##      npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
```

```
## m0      8 44.082 65.160 -14.0409  28.0817
```

```
## m5     10 27.293 53.641  -3.6466   7.2932 20.788  2 3.061e-05 ***
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Look at raw and digestate separately.

```
m4r <- lm(log10(emis.perc) ~ pH + DM + app.meth, data = dsub, subset = relDiff.frac == 'reference')
summary(m4r)
```

```
##
```

```
## Call:
```

```
## lm(formula = log10(emis.perc) ~ pH + DM + app.meth, data = dsub,
```

```
##      subset = relDiff.frac == "reference")
```

```
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.41478 -0.15717  0.00526  0.14443  0.39722
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -0.60305     1.11683   -0.540  0.594897
## pH              0.22291     0.13906    1.603  0.123868
## DM              0.05360     0.01919    2.794  0.010878 *
## app.methopen slot injection -0.79886     0.20876   -3.827  0.000983 ***
## app.methtrailing hose      -0.07566     0.10865   -0.696  0.493867
## app.methtrailing shoe      -0.14913     0.25432   -0.586  0.563853
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2287 on 21 degrees of freedom
## (46 observations deleted due to missingness)
## Multiple R-squared:  0.4666, Adjusted R-squared:  0.3396
## F-statistic: 3.674 on 5 and 21 DF,  p-value: 0.01519

drop1(m4r, test = 'F')

## Single term deletions
##
## Model:
## log10(emis.perc) ~ pH + DM + app.meth
##      Df Sum of Sq  RSS   AIC F value    Pr(>F)
## <none>                 1.0987 -74.446
## pH          1   0.13444 1.2331 -73.329   2.5696  0.123868
## DM          1   0.40839 1.5071 -67.913   7.8058  0.010878 *
## app.meth    3   0.82686 1.9256 -65.297   5.2681  0.007231 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

m4d <- lm(log10(emis.perc) ~ pH + DM + app.meth, data = dsub, subset = relDiff.frac == 'digestate')
summary(m4d)

##
## Call:
## lm(formula = log10(emis.perc) ~ pH + DM + app.meth, data = dsub,
##     subset = relDiff.frac == "digestate")
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.35417 -0.13536  0.00286  0.11649  0.28672
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.44423     0.68138   0.652   0.521
## pH              0.08905     0.08113    1.098   0.283
## DM              0.02219     0.02451    0.905   0.374
## app.methopen slot injection -0.08560     0.15933   -0.537   0.596
## app.methtrailing hose      0.11317     0.10046    1.127   0.271
## app.methtrailing shoe      0.01676     0.14610    0.115   0.910
##
```

```
## Residual standard error: 0.1919 on 24 degrees of freedom
## (46 observations deleted due to missingness)
## Multiple R-squared: 0.193, Adjusted R-squared: 0.02487
## F-statistic: 1.148 on 5 and 24 DF, p-value: 0.3629
```

```
drop1(m4d, test = 'F')
```

```
## Single term deletions
##
## Model:
## log10(emis.perc) ~ pH + DM + app.meth
##           Df Sum of Sq      RSS      AIC F value Pr(>F)
## <none>                0.88427 -93.726
## pH          1  0.044386  0.92866 -94.256  1.2047 0.2833
## DM          1  0.030200  0.91447 -94.718  0.8197 0.3743
## app.meth    3  0.108323  0.99259 -96.259  0.9800 0.4186
```

## Predictions

Might be interesting to see how predictions here stack up against current DK EFs (based on ALFAM2).

```
dat$si <- factor(as.integer(factor(dat$source)))
m6 <- lm(log10(emis.perc) ~ pH + DM + si + app.meth, data = dat)
coef(m6)
```

```
##           (Intercept)                pH
##          -0.12193204             0.14074882
##                DM                si2
##          0.04869660             0.17827981
##                si3                si4
##          0.05756898             0.20300552
##                si5                si6
##          0.40397579             0.37177705
##                si7                si8
##          0.48801250             0.29512484
##                si11               si14
##          0.42117908             0.26718720
##                si15               si16
##          0.70218532             0.11573208
##                si17               si18
##          0.20769066             0.25782185
##                si19               si20
##          -0.01567678             0.14392388
##                si21  app.methclosed slot injection
##          0.25760194             -0.67956909
##  app.methopen slot injection  app.methtrailing hose
##          -0.57136678             -0.13501678
##  app.methtrailing shoe app.methtrailing shoe + harrowing
##          -0.31414078             -0.27965539
```

```
preddat <- data.frame(scenario = c('ref', 'dig', 'dig+'), si = factor(19),
                      app.meth = 'trailing hose',
                      DM = c(6.5, 5.1, 6.5), pH = c(7, 7.9, 7.9))
preddat$emis <- 10^predict(m6, newdata = preddat)
```

```
knitr::kable(preddat)
```

scenario	si	app.meth	DM	pH	emis
ref	19	trailing hose	6.5	7.0	10.69409
dig	19	trailing hose	5.1	7.9	12.23609
dig+	19	trailing hose	6.5	7.9	14.31588

## Different approach

```
ddw <- subset(dw, !is.na(emis.perc.ref + emis.perc.dig + DM.ref + DM.dig + pH.ref + pH.dig))
ddw$demis <- ddw$emis.perc.dig - ddw$emis.perc.ref
ddw$dpH <- ddw$pH.dig - ddw$pH.ref
ddw$dDM <- ddw$DM.dig - ddw$DM.ref
```

```
m7 <- lm(log10(emis.perc.dig) ~ log10(emis.perc.ref) + dpH + dDM, data = ddw)
summary(m7)
```

```
##
## Call:
## lm(formula = log10(emis.perc.dig) ~ log10(emis.perc.ref) + dpH +
##      dDM, data = ddw)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.34574 -0.05195  0.00894  0.06739  0.25809
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.59102     0.11477   5.150 1.84e-05 ***
## log10(emis.perc.ref) 0.54173     0.08754   6.188 1.10e-06 ***
## dpH             0.03294     0.04951   0.665  0.511
## dDM            -0.00223     0.01228  -0.182  0.857
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1324 on 28 degrees of freedom
## Multiple R-squared:  0.5817, Adjusted R-squared:  0.5369
## F-statistic: 12.98 on 3 and 28 DF,  p-value: 1.699e-05
```

```
m8 <- lm(demis ~ dpH + dDM, data = ddw)
summary(m8)
```

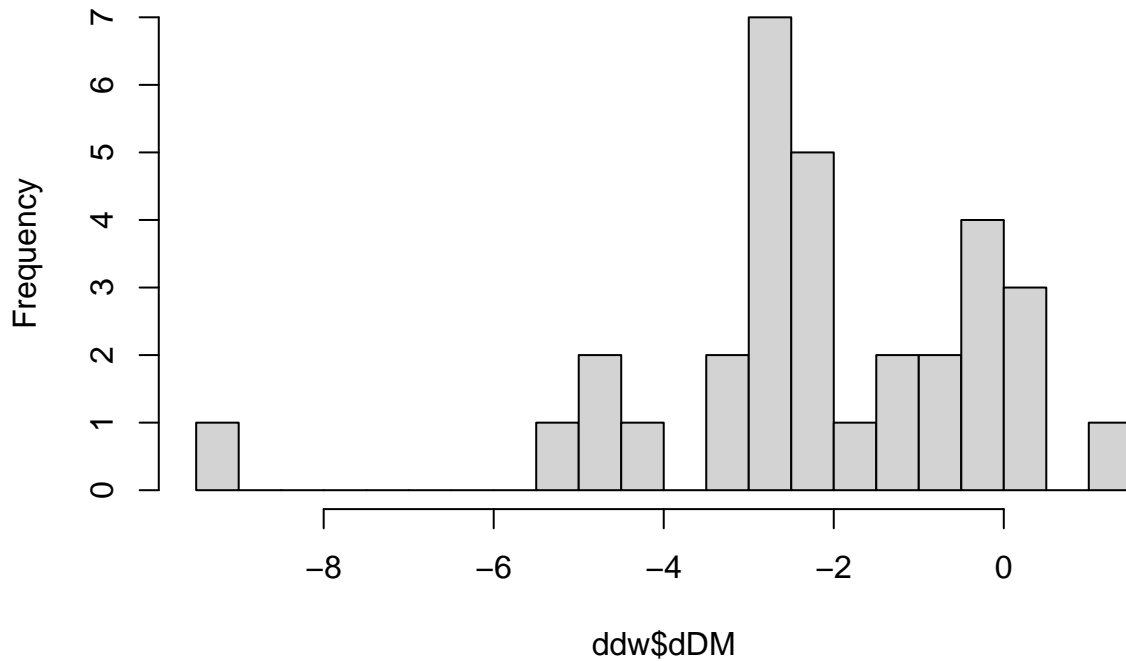
```
##
## Call:
## lm(formula = demis ~ dpH + dDM, data = ddw)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16.374  -4.769   1.514   3.550  17.253
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)  -0.5760    2.2367   -0.258    0.799
## dpH          2.5331    2.8854    0.878    0.387
## dDM          0.1806    0.7157    0.252    0.803
##
## Residual standard error: 7.728 on 29 degrees of freedom
## Multiple R-squared:  0.02609,    Adjusted R-squared:  -0.04108
## F-statistic: 0.3885 on 2 and 29 DF,  p-value: 0.6816
```

Wow. Nothing going on.

```
hist(ddw$dDM, breaks = 20)
```

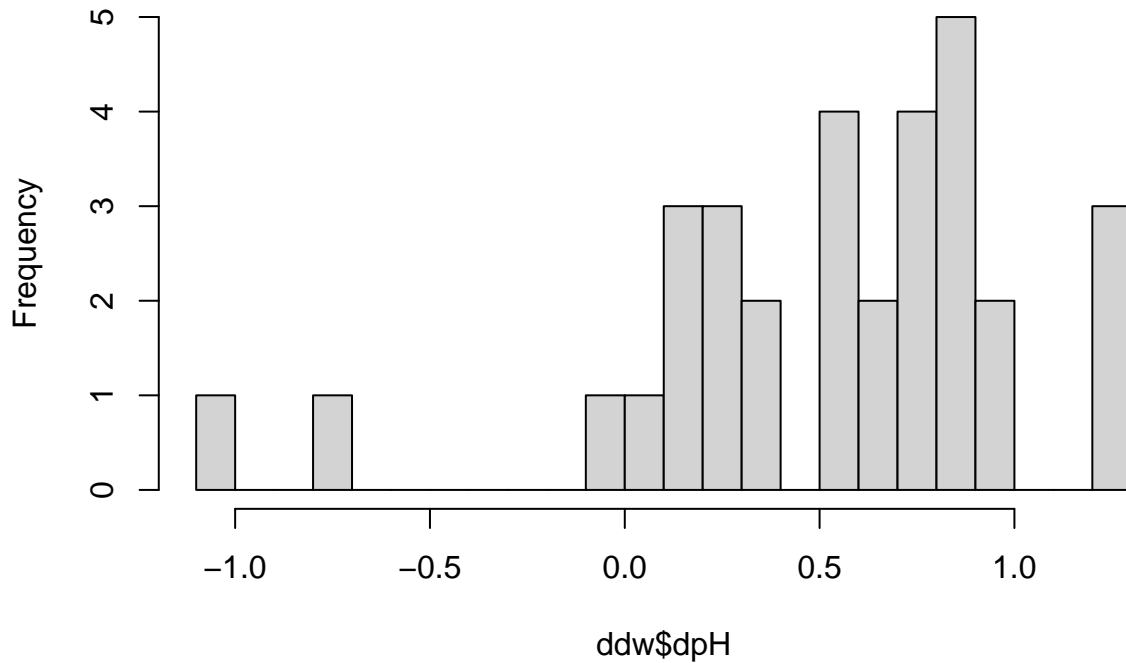
## Histogram of ddw\$dDM



```
hist(ddw$dpH, breaks = 20)
```



## Histogram of ddw\$d pH



Include interaction.

```
m9 <- lm(log10(emis.perc.dig) ~ log10(emis.perc.ref) + dpH*dDM, data = ddw)
summary(m9)
```

```
##
## Call:
## lm(formula = log10(emis.perc.dig) ~ log10(emis.perc.ref) + dpH *
##     dDM, data = ddw)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.35702 -0.06029  0.00276  0.06269  0.27676
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.654115   0.120589   5.424 9.76e-06 ***
## log10(emis.perc.ref) 0.502291   0.090015   5.580 6.44e-06 ***
## dpH             -0.006731   0.055663  -0.121   0.905
## dDM              0.021523   0.020275   1.062   0.298
## dpH:dDM         -0.037187   0.025535  -1.456   0.157
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1298 on 27 degrees of freedom
## Multiple R-squared:  0.6122, Adjusted R-squared:  0.5547
## F-statistic: 10.65 on 4 and 27 DF,  p-value: 2.592e-05
```