

Stats to parse out DM and pH effects

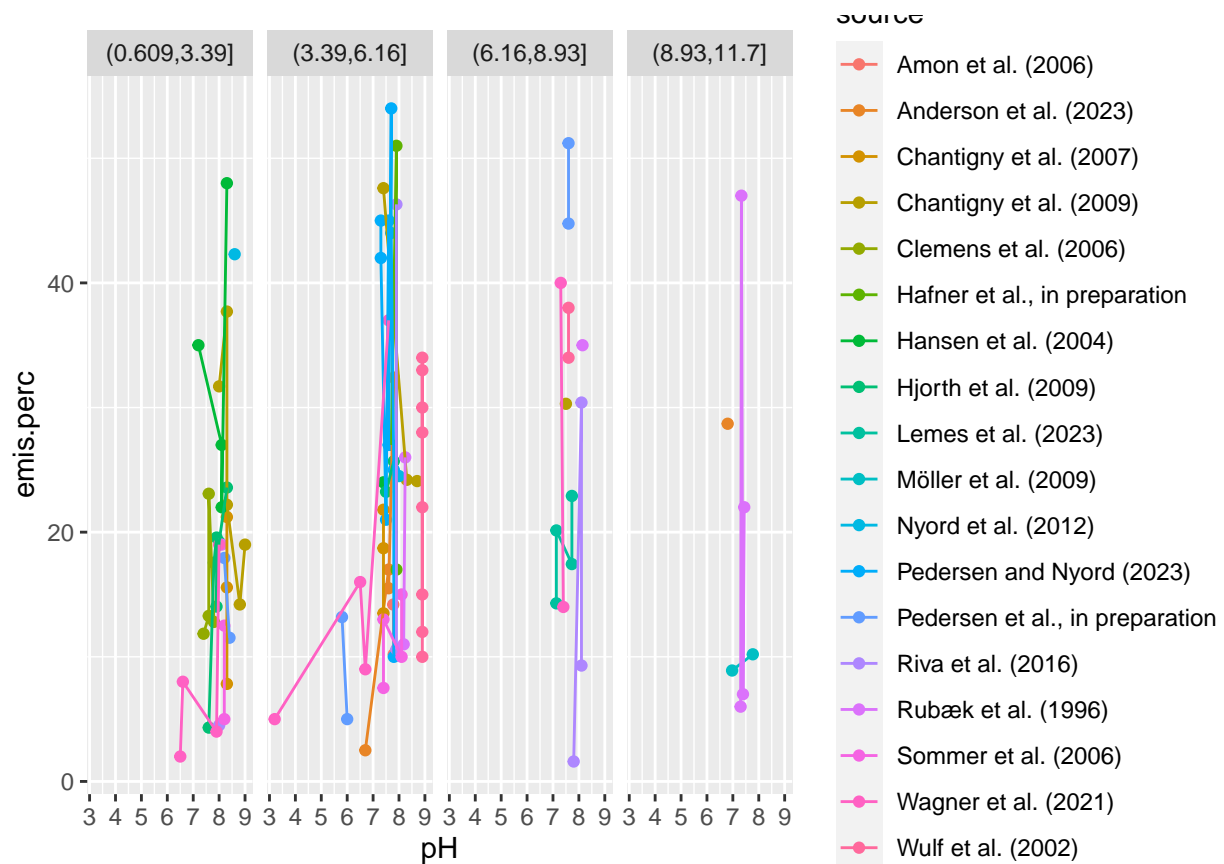
Sasha D. Hafner

25 oktober, 2023

Plots

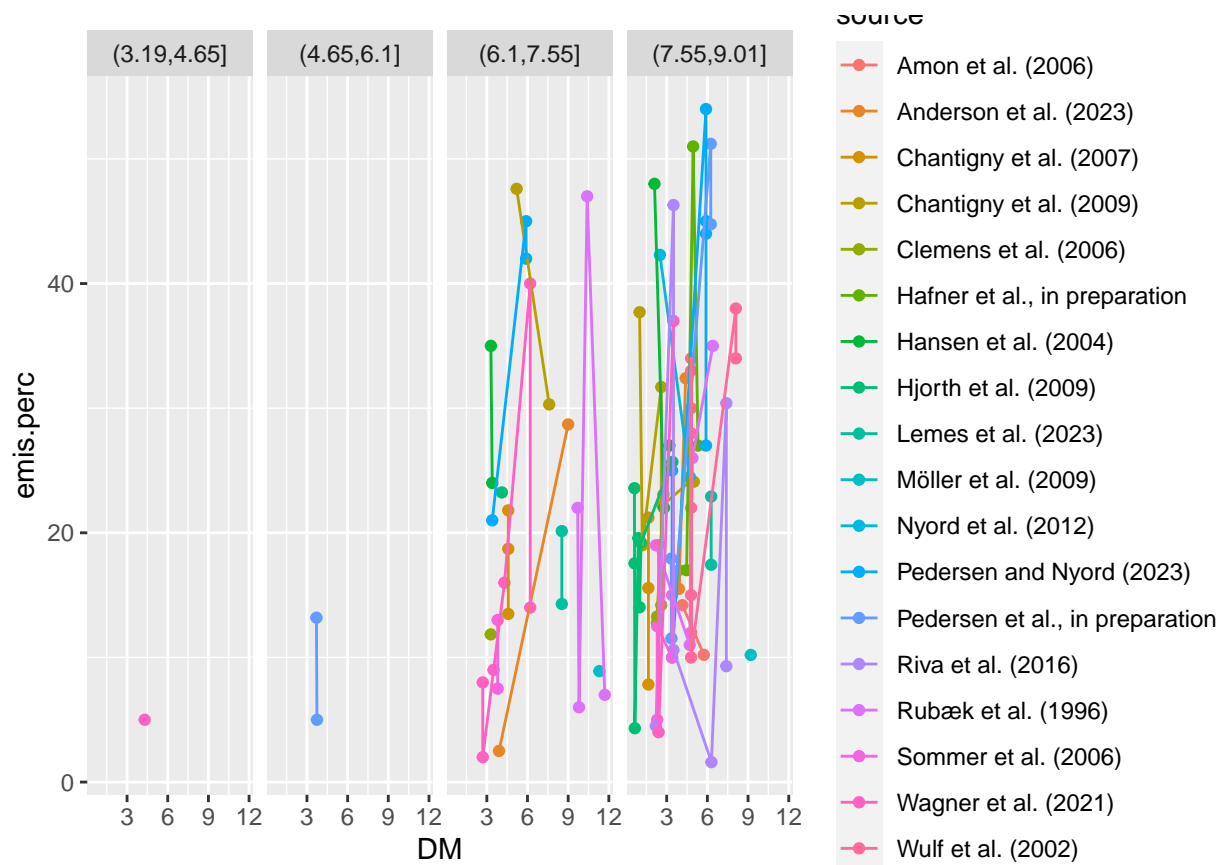
```
dat$DMgrp <- cut(dat$DM, 4)
dat$pHgrp <- cut(dat$pH, 4)
dd <- subset(dat, !is.na(pH) & !is.na(DM))
```

```
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)
```

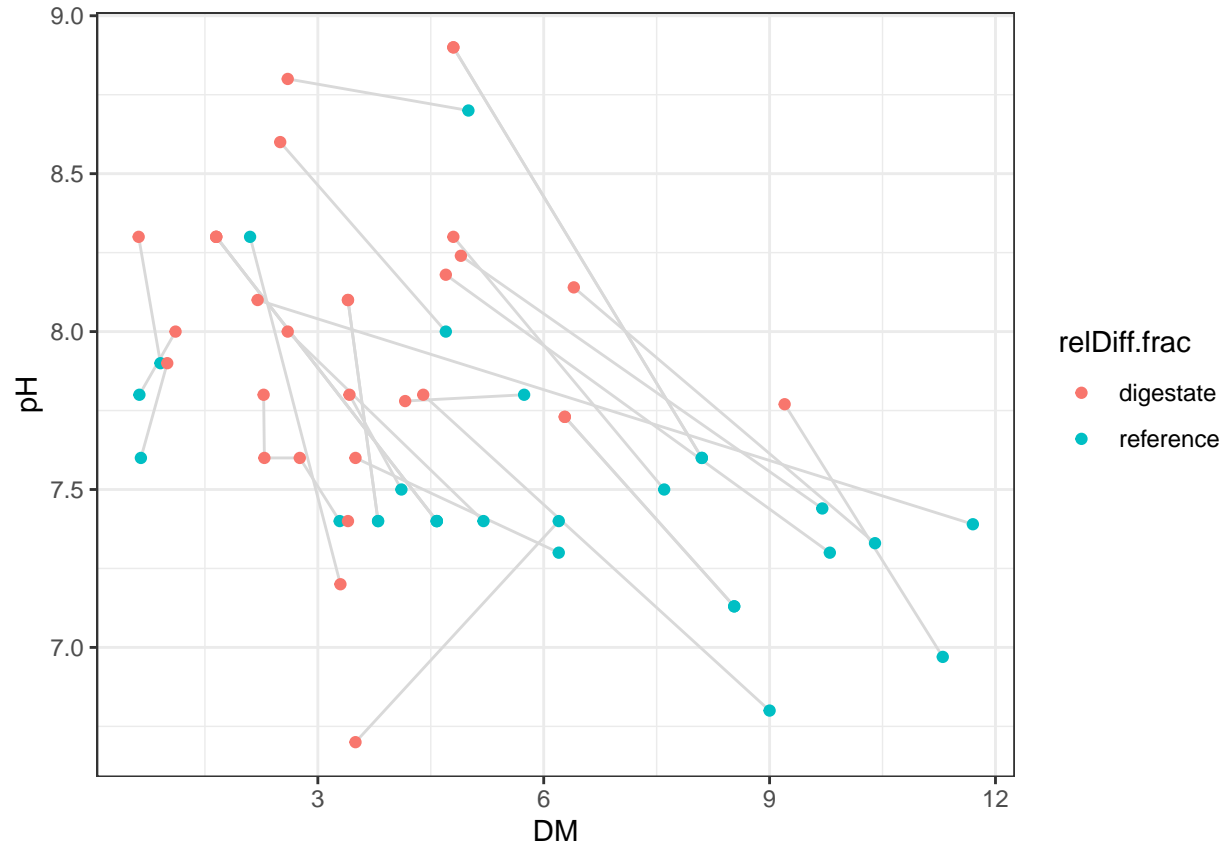
'geom_line()': Each group consists of only one observation.
 ## i Do you need to adjust the group aesthetic?



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.

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

```
##
## 119 rows and 6 columns
## 119 unique rows
##
##           source           app.meth      DM      pH relDiff.frac emis.pe
## Class           character      character numeric numeric      character  numer
## Minimum      Amon et al. (2006) broadcast    0.62    3.2    digestate    1
## Maximum      Wulf et al. (2002) trailing shoe + harrowing 11.7    9      reference    0
## Mean          <NA>              <NA>    4.43    7.75    <NA>    22
## Unique (excl. NA)      20              6      68    36      2      9
## Missing values          0              4      2    14      46
## Sorted          FALSE              FALSE FALSE FALSE      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
## 50 observations deleted due to missingness
```

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

```
## sourceWagner et al. (2021)    sourceWulf et al. (2002) app.methopen slot injection    app.methtr
##              12.798830              20.298830              -13.750000
##      relDiff.fracreference
##              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
## 62 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
## 16 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
```

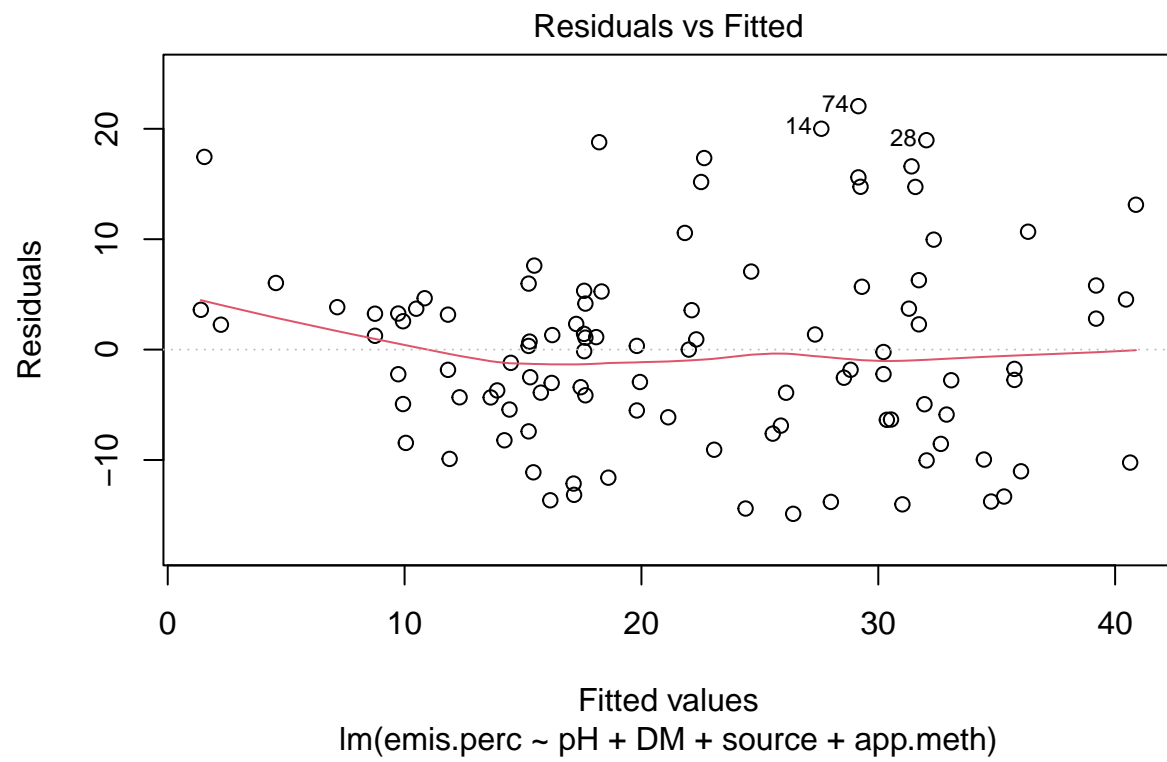
```
cc['pH'] / cc['DM']
```

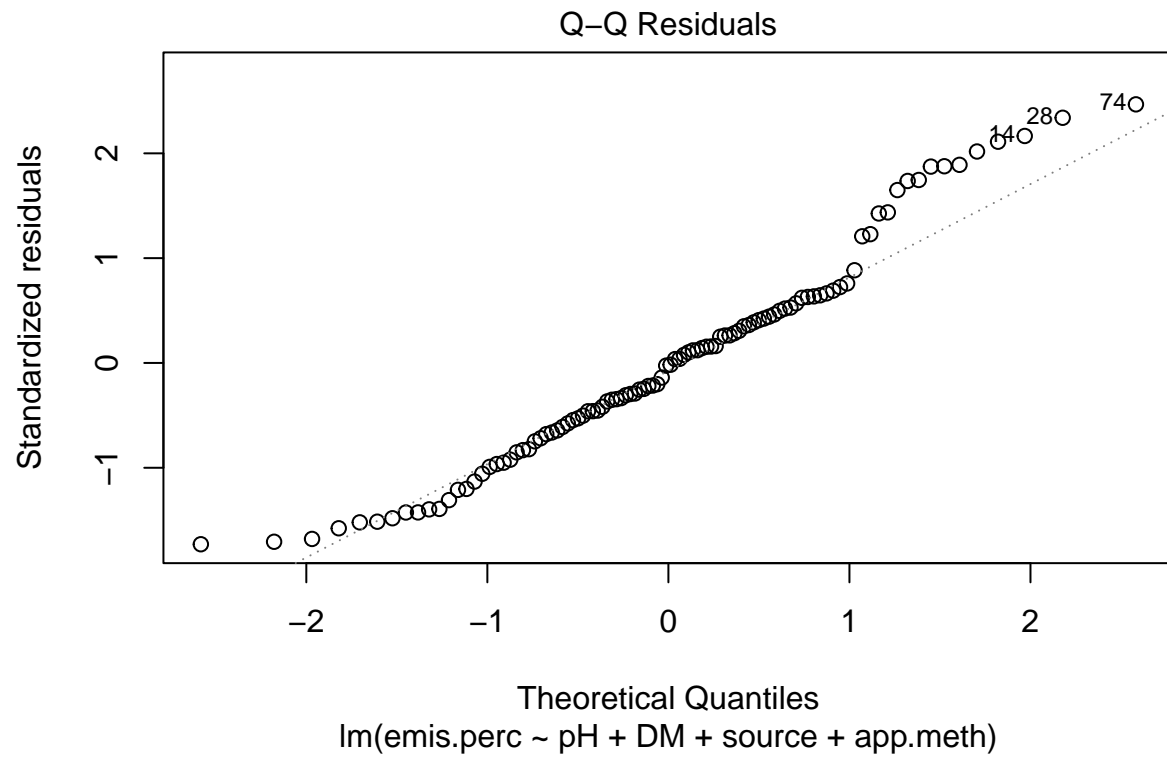
```
##          pH
## 1.993609
```

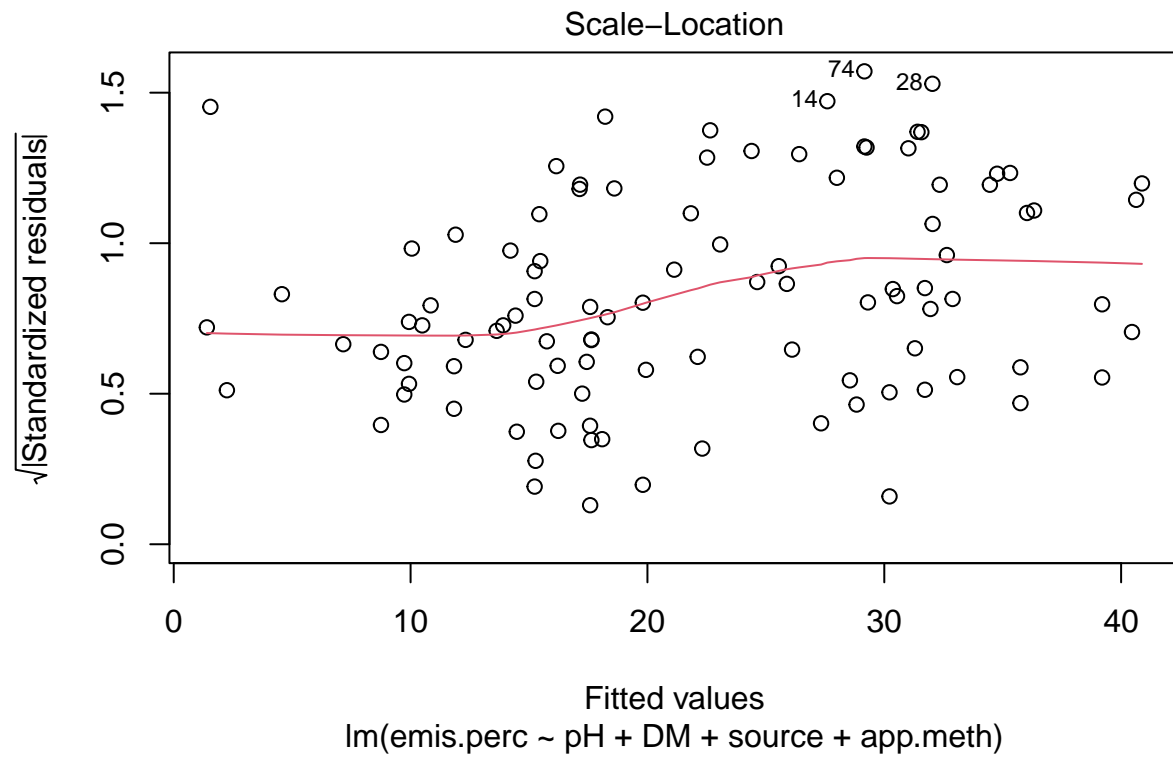
m3 is where it gets interesting. Clear positive effects of DM and pH, although the pH effect seems small (+4% emission (frac. of TAN) per unit pH change). Effect of pH is about 2x DM effect (1 unit pH change : 1% DM change).

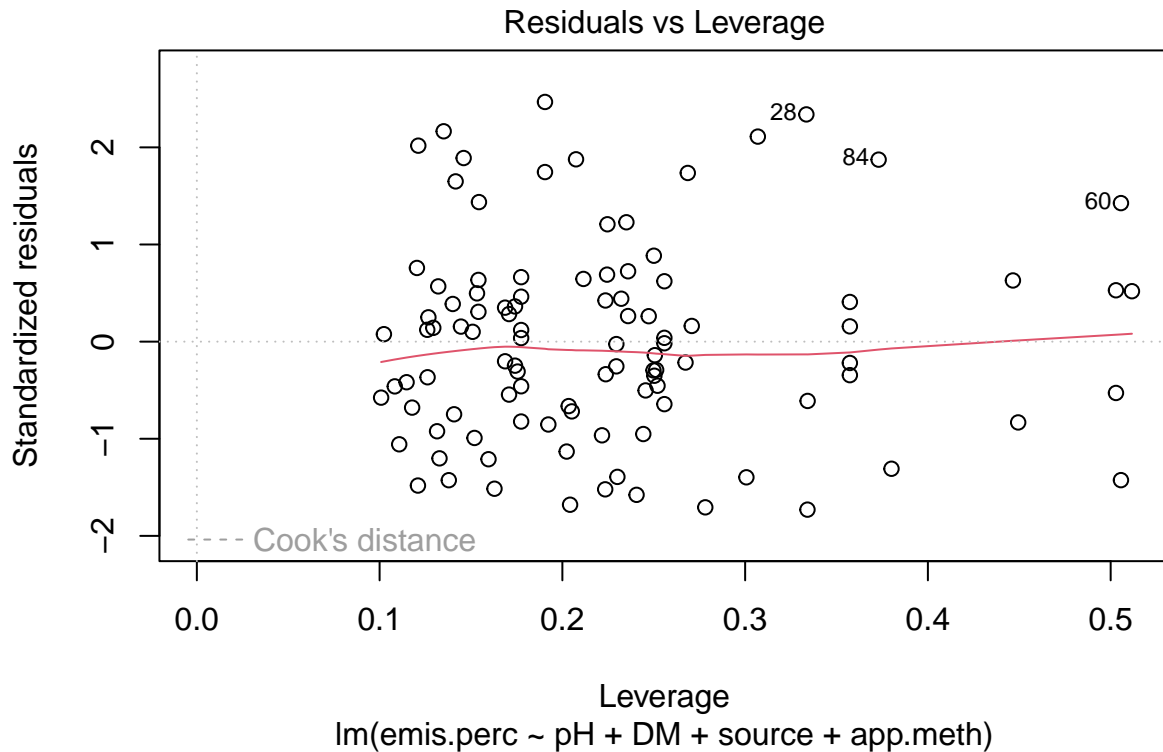
```
plot(m3, ask = FALSE)
```

```
## 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
## 16 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
```

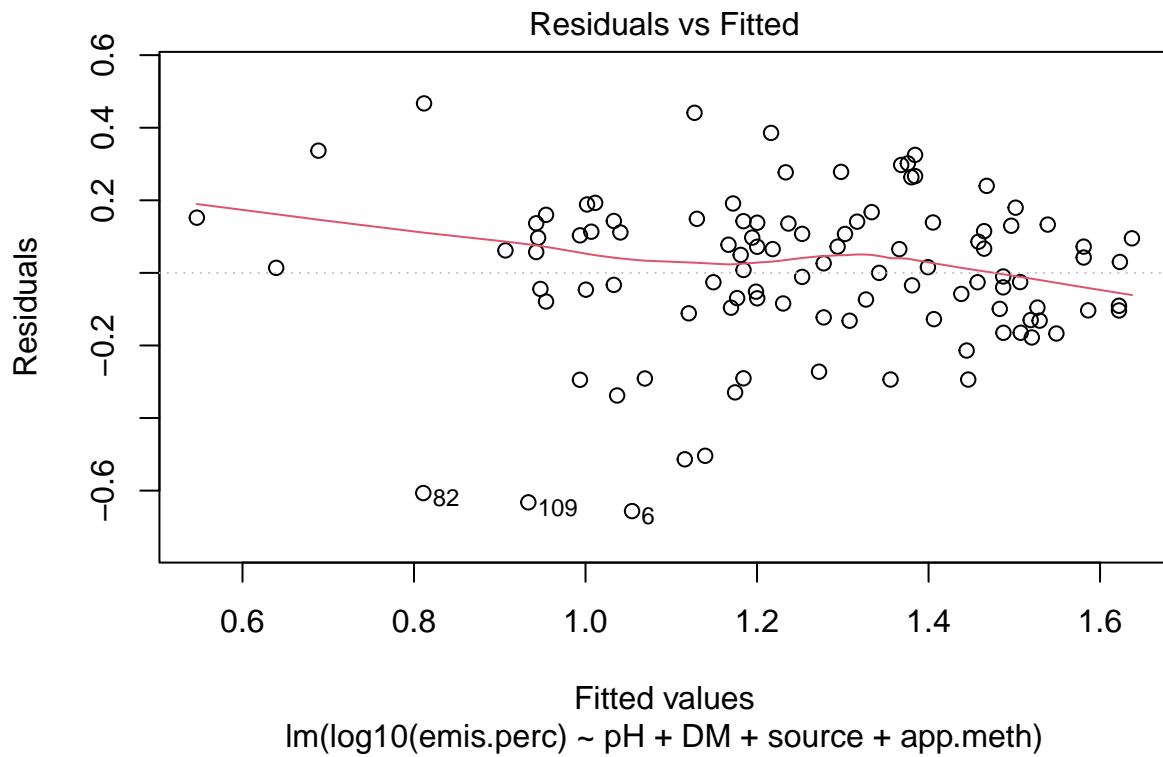
```
cc['pH'] / cc['DM']
```

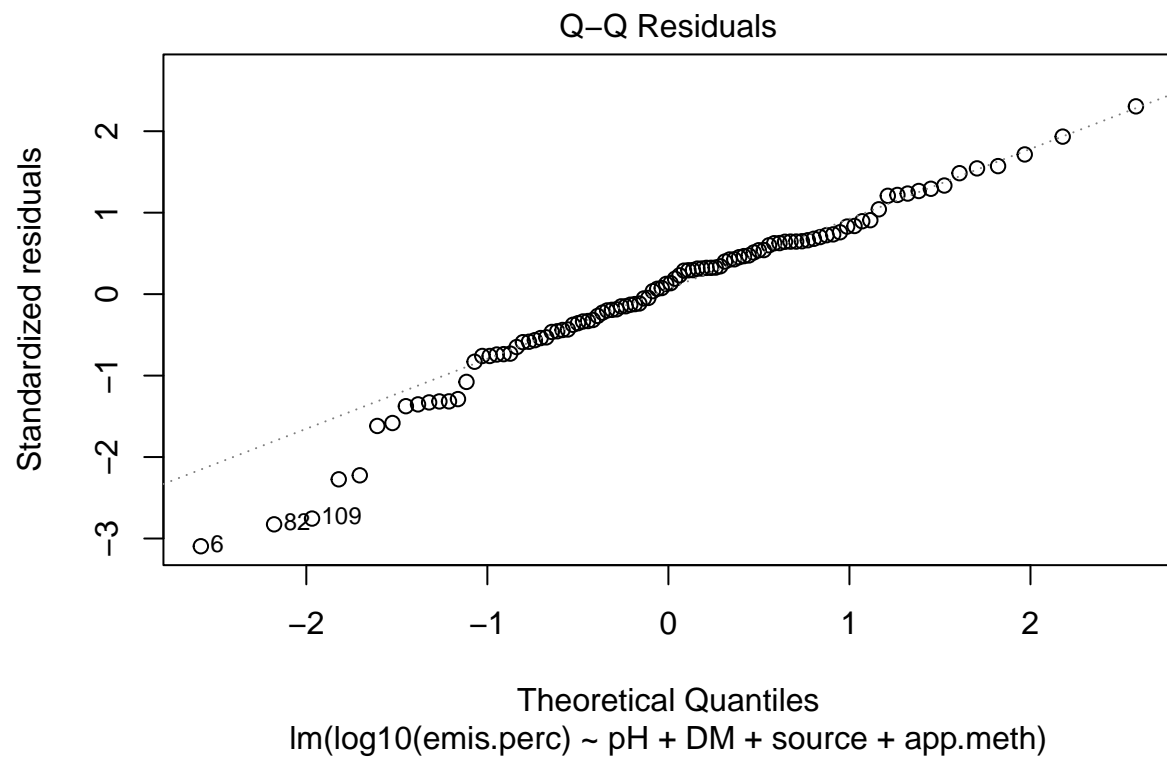
```
##      pH
## 3.225847
```

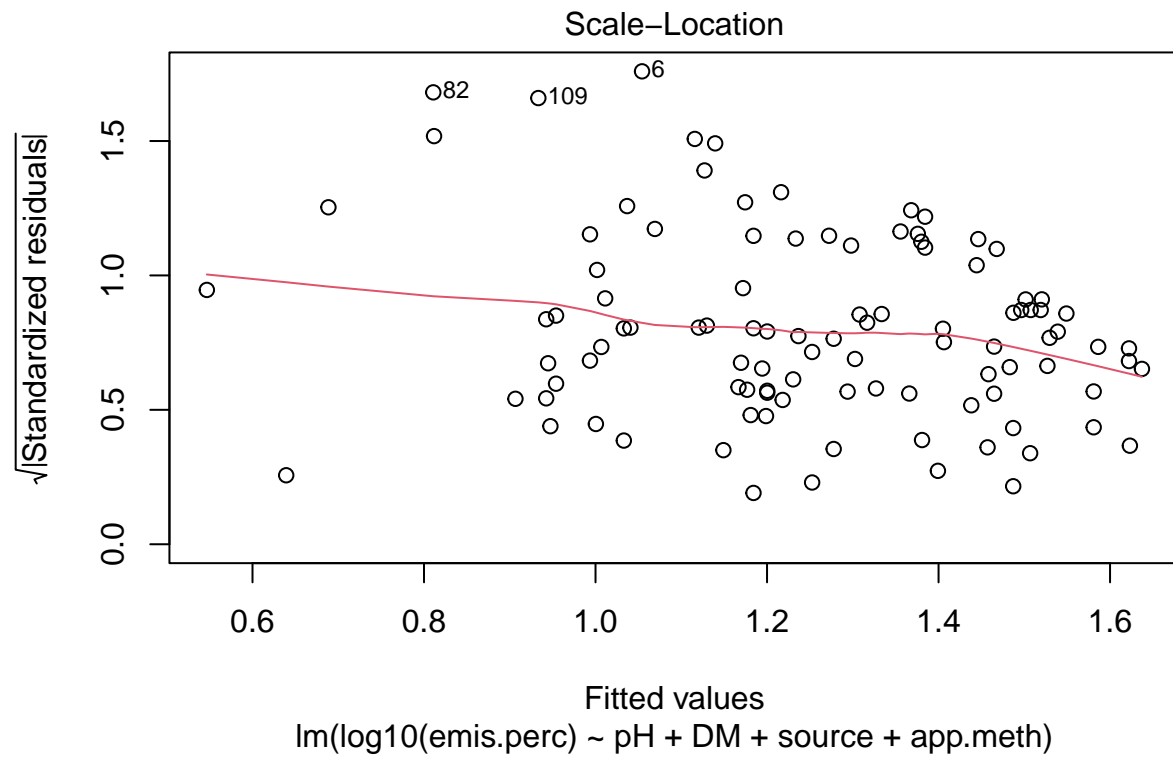
38% increase in emission with a 1 unit increase in pH, seems plausible. Effect of pH here 3x DM effect. Seems more plausible.

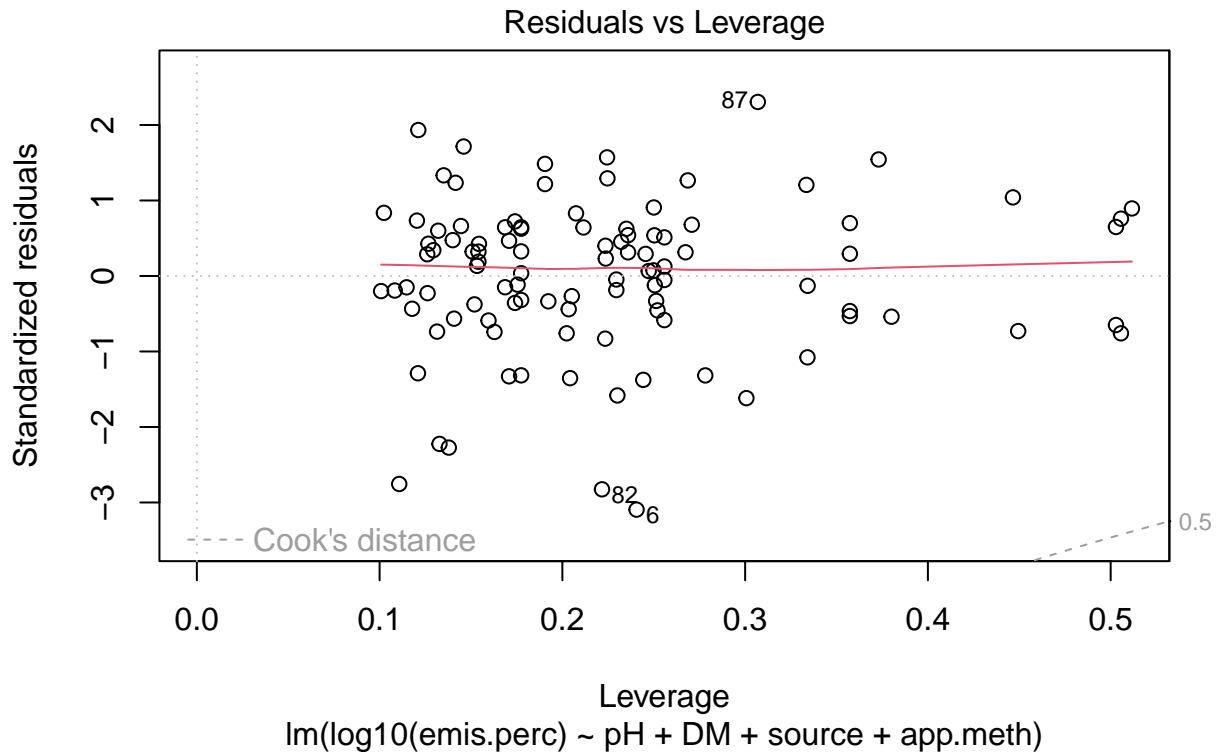
```
plot(m4, ask = FALSE)
```

```
## 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')])
```

```
##
## 59 rows and 4 columns
## 49 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      7.77      4.72      <NA>      <NA>
## Unique (excl. NA)      27      43      4      14
## Missing values      0      0      2      0
## Sorted      FALSE      FALSE      FALSE      TRUE
##
```

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

And there are multiple observations from each source.

```
table(dsub$source)
```

```
##
##           Amon et al. (2006)           Anderson et al. (2023)           Chantigny et al. (2007)
##                               2                               5                               6
##           Clemens et al. (2006)   Hafner et al., in preparation           Hansen et al. (2004)
##                               4                               3                               5
##           Lemes et al. (2023)           Neerackal et al. (2015)           Nyord et al. (2012)
##                               4                               8                               2
##           Pedersen et al. (2021) Pedersen et al., in preparation           Riva et al. (2016)
##                               4                               6                               6
##           Sommer et al. (2006)           Wagner et al. (2021)           Wulf et al. (2002)
##                               6                               10                              10
```

```
summary(m5)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: log10(emis.perc) ~ relDiff.frac + app.meth + (1 | source)
## Data: dsub
##
## REML criterion at convergence: -2.5
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.2292 -0.4857  0.0590  0.6099  1.6787
##
## Random effects:
## Groups   Name                Variance Std.Dev.
## source  (Intercept)  0.02839  0.1685
## Residual                    0.03507  0.1873
## Number of obs: 69, groups: source, 15
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)      1.303680   0.101085  12.897
## relDiff.fracreference -0.002955   0.045499  -0.065
## app.methopen slot injection -0.285128   0.107897  -2.643
## app.methtrailing hose      0.087273   0.108850   0.802
## app.methtrailing shoe     -0.136058   0.216767  -0.628
##
## Correlation of Fixed Effects:
##              (Intr) rLDff. app.si app.mh
## rLDff.frcrf -0.224
## app.mthpnsi -0.529  0.002
## app.mthtrlh -0.828  0.005  0.568
## app.mthtrls -0.455  0.052  0.247  0.386
```

```
ccm <- 100 * (10^fixef(m5) - 1)
ccm
```

```
##              (Intercept)      relDiff.fracreference app.methopen slot injection      app.methtrailing
##      1912.2427561              -0.6781026              -48.1353250
```

```
100 * (10^confint(m5) - 1)
```

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

```
##                2.5 %    97.5 %
## .sig01          19.34601  76.96722
## .sigma          42.42857  68.09626
## (Intercept)     1183.98821 3039.32502
## relDiff.fracreference -19.19513  21.85180
## app.methopen slot injection -67.74775 -16.80177
## app.methtrailing hose -24.99057  98.81704
## app.methtrailing shoe -71.20792  89.58999
```

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

```
##
## 59 rows and 4 columns
## 49 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          7.77      4.72      <NA>          <NA>
## Unique (excl. NA) 27      43          4          14
## Missing values    0      0          2          0
## Sorted          FALSE    FALSE      FALSE      TRUE
##
```

```
dsub <- subset(dat, !is.na(pH) & !is.na(DM) & !is.na(app.meth))
m6 <- lmer(log10(emis.perc) ~ pH + DM + app.meth + (1|source), data = dsub)
```

```
summary(m6)
```

```
## 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(m6) - 1)
ccm
```

```
## (Intercept) pH DM
## -6.566185 41.569348 11.80528
## app.methopen slot injection app.methtrailing hose app.methtrailing shoe
## -56.030607 -13.827757 -11.29042
```

```
100 * (10^confint(m6)[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 11.80528
## app.methopen slot injection app.methtrailing hose app.methtrailing shoe
## -56.030607 -13.827757 -11.29042
```


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)
mODM <- lmer(log10(emis.perc) ~ pH + app.meth + (1|source), data = dsub)
anova(m6, m0pH)
```

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

```
## Data: dsub
## Models:
## m0pH: log10(emis.perc) ~ DM + app.meth + (1 | source)
## m6: 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
## m6      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(m6, mODM)
```

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

```
## Data: dsub
## Models:
## mODM: log10(emis.perc) ~ pH + app.meth + (1 | source)
## m6: log10(emis.perc) ~ pH + DM + app.meth + (1 | source)
##      npar    AIC    BIC   logLik deviance  Chisq Df Pr(>Chisq)
## mODM     9 38.512 62.225 -10.2561  20.5123
## m6      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(m6, m0)
```

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

```
## Data: dsub
## Models:
## m0: log10(emis.perc) ~ app.meth + (1 | source)
## m6: 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
## m6      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
```

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)))
m7 <- lm(log10(emis.perc) ~ pH + DM + si + app.meth, data = dat)
coef(m7)
```

```
##              (Intercept)              pH              DM
##      -0.12193204          0.14074882          0.04869666
##              si3              si4              si5
##      0.05756898          0.20300552          0.40397579
##              si7              si8              si9
##      0.48801250          0.29512484          0.11573203
##              si13             si15             si16
##      0.70218532          0.26718720          0.20769066
##              si18             si19             si20
##      -0.01567678          0.14392388          0.25760194
##      app.methopen slot injection      app.methtrailing hose      app.methtrailing shoe
##      -0.57136678          -0.13501678          -0.31414079
```

```
preddat <- data.frame(scenario = c('ref', 'dig', 'DM', 'pH'), si = factor(19),
                      app.meth = 'trailing hose',
                      DM = c(6.5, 5.1, 5.1, 6.5), pH = c(7, 7.9, 7, 7.9))
preddat$emis <- 10^predict(m7, newdata = preddat)
preddat$rcDM <- 100 * ((preddat$emis - preddat$emis[1]) / preddat$emis[1]) / (preddat$DM - preddat$DM[1])
preddat$rcpH <- 100 * ((preddat$emis - preddat$emis[1]) / preddat$emis[1]) / (preddat$pH - preddat$pH[1])
```

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

scenario	si	app.meth	DM	pH	emis	rcDM	rcpH
ref	19	trailing hose	6.5	7.0	15.44346	NaN	NaN
dig	19	trailing hose	5.1	7.9	17.67028	-10.29939	16.02127
DM	19	trailing hose	5.1	7.0	13.19986	10.37702	-Inf
pH	19	trailing hose	6.5	7.9	20.67372	Inf	37.63014

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
```

```
m8 <- lm(log10(emis.perc.dig) ~ log10(emis.perc.ref) + dph + dDM, data = ddw)
summary(m8)
```

```
##
## Call:
## lm(formula = log10(emis.perc.dig) ~ log10(emis.perc.ref) + dph +
##      dDM, data = ddw)
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.33780 -0.05069  0.00758  0.07053  0.26165
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.57702    0.12075   4.779 6.03e-05 ***
## log10(emis.perc.ref) 0.55047    0.09142   6.021 2.33e-06 ***
## dpH            0.01750    0.05700   0.307   0.761
## dDM           -0.00567    0.01393  -0.407   0.687
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1364 on 26 degrees of freedom
## Multiple R-squared:  0.5862, Adjusted R-squared:  0.5385
## F-statistic: 12.28 on 3 and 26 DF,  p-value: 3.42e-05
```

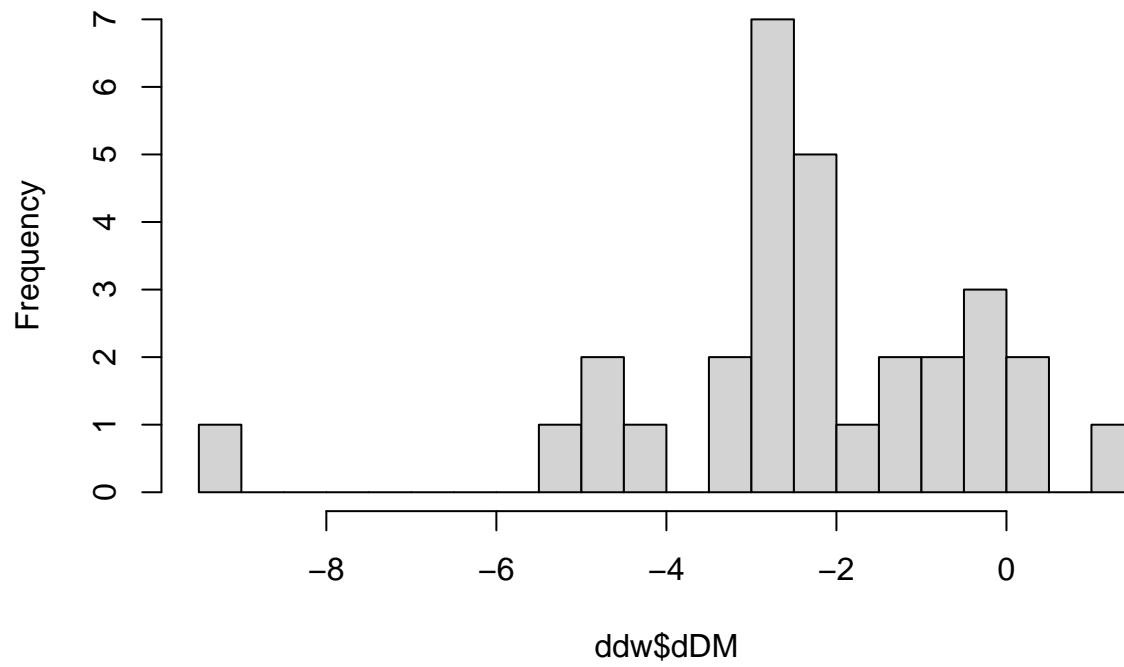
```
m9 <- lm(demis ~ dpH + dDM, data = ddw)
summary(m9)
```

```
##
## Call:
## lm(formula = demis ~ dpH + dDM, data = ddw)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16.083  -4.786   1.936   3.240  17.662
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.8593     2.3139  -0.371   0.713
## dpH           1.2468     3.3006   0.378   0.709
## dDM          -0.1133     0.8066  -0.140   0.889
##
## Residual standard error: 7.9 on 27 degrees of freedom
## Multiple R-squared:  0.0103, Adjusted R-squared:  -0.06301
## F-statistic: 0.1405 on 2 and 27 DF,  p-value: 0.8695
```

Wow. Nothing going on.

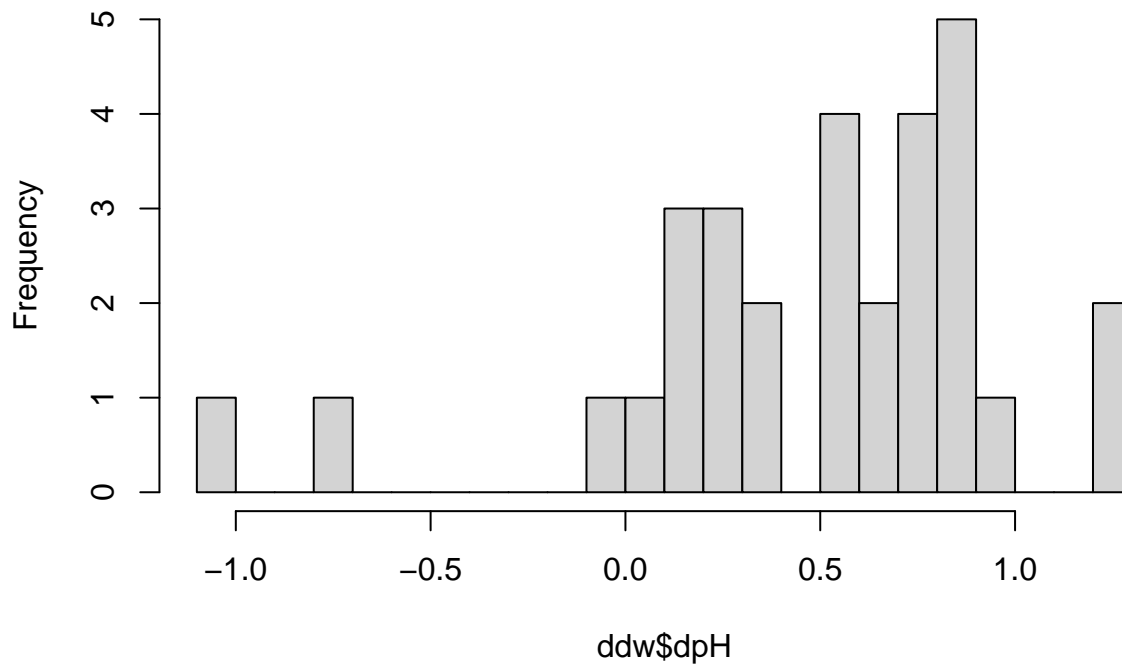
```
hist(ddw$dDM, breaks = 20)
```

Histogram of ddw\$dDM



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

Histogram of ddw\$dpH



Include interaction.

```
m10 <- lm(log10(emis.perc.dig) ~ log10(emis.perc.ref) + dpH*dDM, data = ddw)
summary(m10)
```

```
##
## Call:
## lm(formula = log10(emis.perc.dig) ~ log10(emis.perc.ref) + dpH *
##     dDM, data = ddw)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.34652 -0.06666 -0.00179  0.06010  0.28448
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.64078    0.12425   5.157 2.48e-05 ***
## log10(emis.perc.ref) 0.51054    0.09248   5.521 9.75e-06 ***
## dpH            -0.03484    0.06465  -0.539   0.595
## dDM             0.01932    0.02087   0.926   0.363
## dpH:dDM        -0.04188    0.02660  -1.574   0.128
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1327 on 25 degrees of freedom
## Multiple R-squared:  0.6235, Adjusted R-squared:  0.5633
```

```
## F-statistic: 10.35 on 4 and 25 DF, p-value: 4.371e-05
```

Better to use mixed-effects model.

```
m11 <- lmer(demis ~ dpH + dDM + (1|source), data = ddw)
summary(m11)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: demis ~ dpH + dDM + (1 | source)
## Data: ddw
##
## REML criterion at convergence: 191.9
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.34194 -0.41053  0.03202  0.35965  1.62213
##
## Random effects:
## Groups Name Variance Std.Dev.
## source (Intercept) 49.36 7.026
## Residual 25.34 5.034
## Number of obs: 30, groups: source, 14
##
## Fixed effects:
## Estimate Std. Error t value
## (Intercept) -3.3834 2.9617 -1.142
## dpH 0.3909 3.3565 0.116
## dDM -1.5697 0.8808 -1.782
##
## Correlation of Fixed Effects:
## (Intr) dpH
## dpH -0.322
## dDM 0.470 0.334
```

```
confint(m11)
```

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

```
##      2.5 %      97.5 %
## .sig01 2.226036 10.9701800
## .sigma 3.526253 7.5739994
## (Intercept) -9.353064 2.3565819
## dpH -6.020312 6.9034389
## dDM -3.342465 0.5623619
```