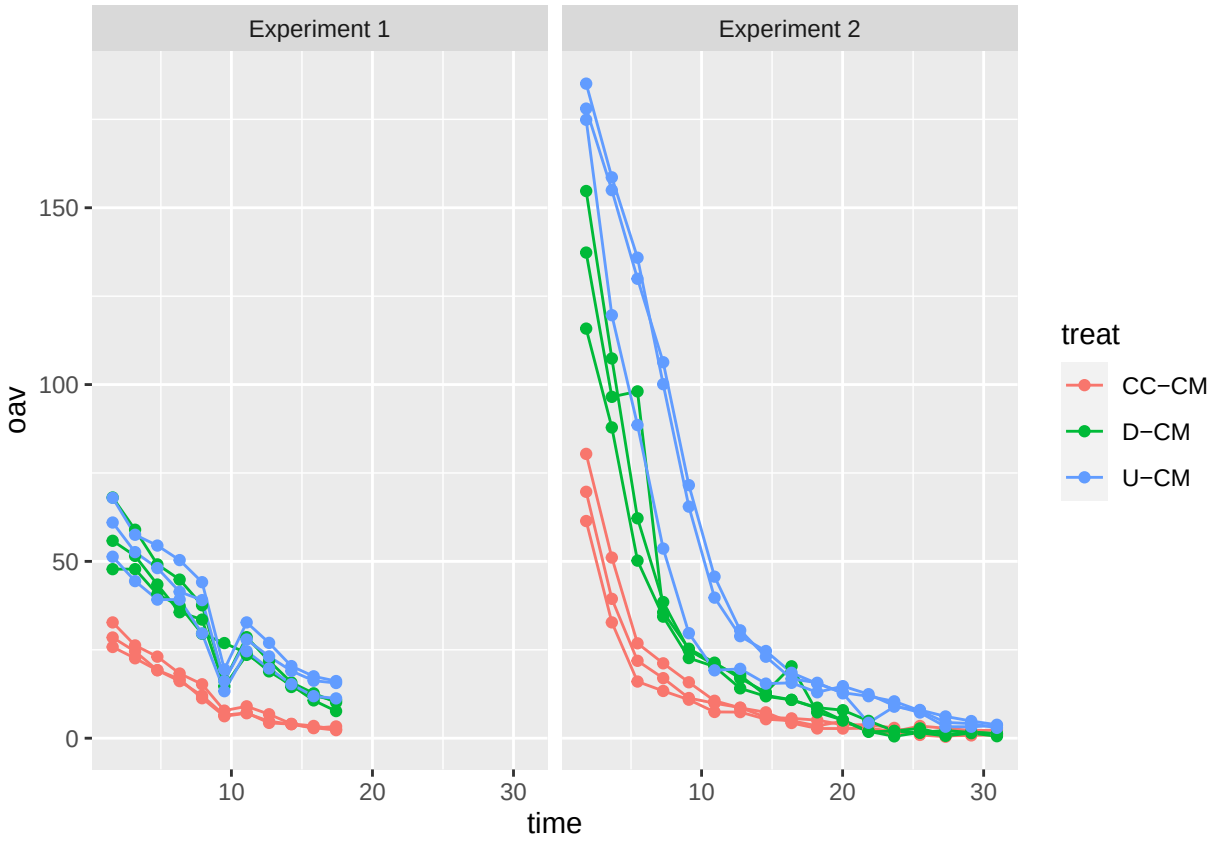
Data analysis for odor from digestate experiments

Sasha D. Hafner

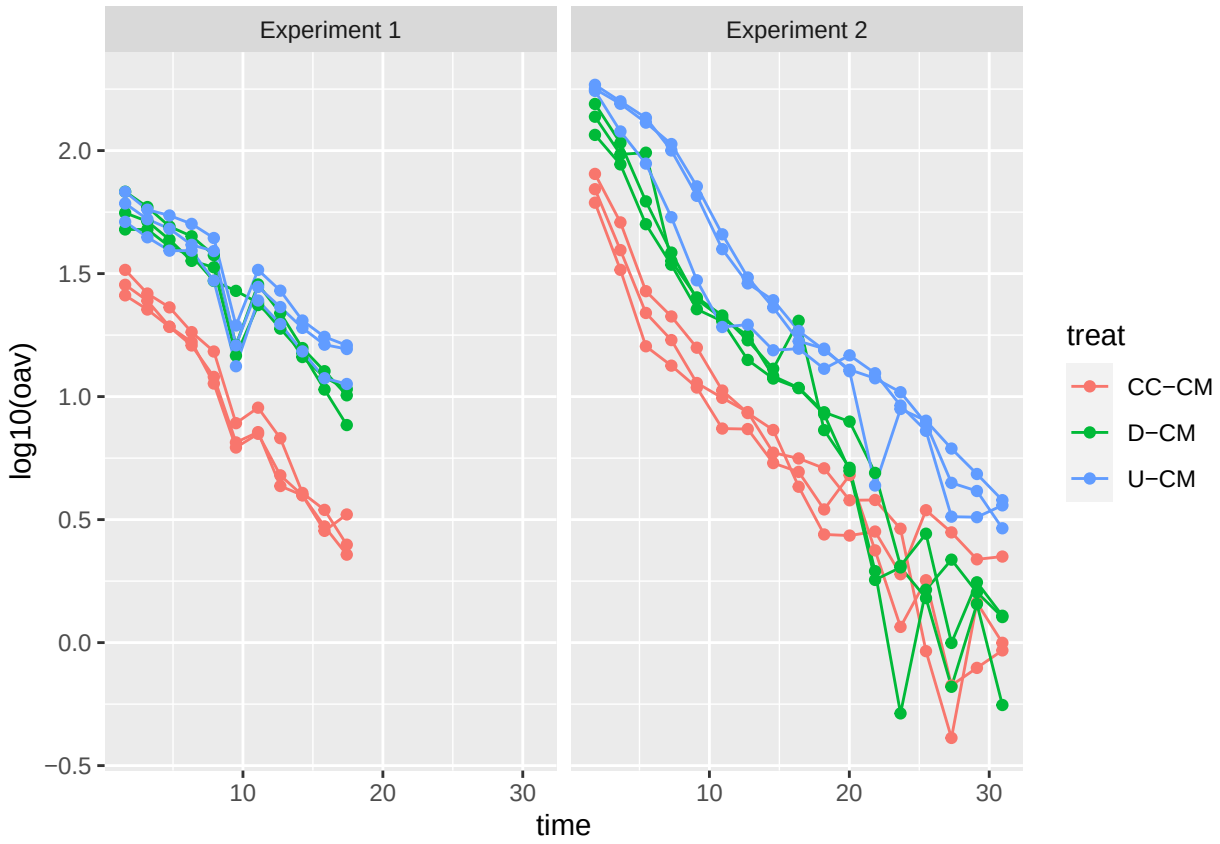
21 December, 2022

Plots

```
ggplot(dat, aes(time, oav, colour = treat, group = interaction(rep, treat))) +
  geom_line() +
  geom_point() +
  facet_wrap(~ exper)
```



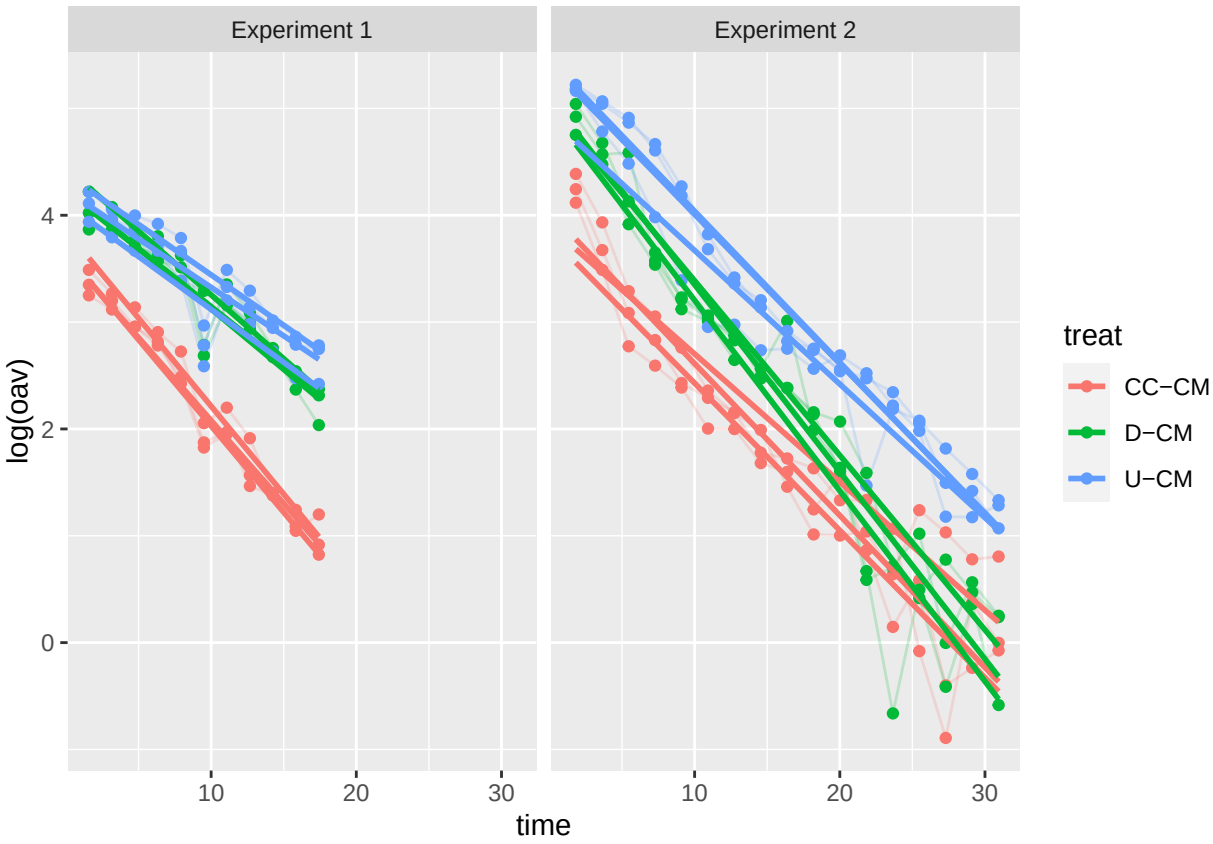

```
ggplot(dat, aes(time, log10(oav), colour = treat, group = interaction(rep, treat))) +
  geom_line() +
  geom_point() +
  facet_wrap(~ exper)
```



Looks linear enough.

```
ggplot(dat, aes(time, log(oav), colour = treat, group = interaction(rep, treat))) +
  geom_line(alpha = 0.2) +
  geom_point() +
  geom_smooth(method = lm, se = FALSE) +
  facet_wrap(~ exper)
```

```
## `geom_smooth()` using formula 'y ~ x'
```



Stats

So, our question will be (referring to last plot above) "are there differences in slope and initial value?". I had expected to use the intercept and slope terms in the analysis, but the least-squares lines don't fit very well in all cases at the start. And anyway, the initial measurements are our best estimate of initial OAV.

Set reference to untreated cattle manure.

```
dat$treat <- factor(dat$treat, levels = c('U-CM', 'CC-CM', 'D-CM'))
dat$exper <- factor(dat$exper)
```

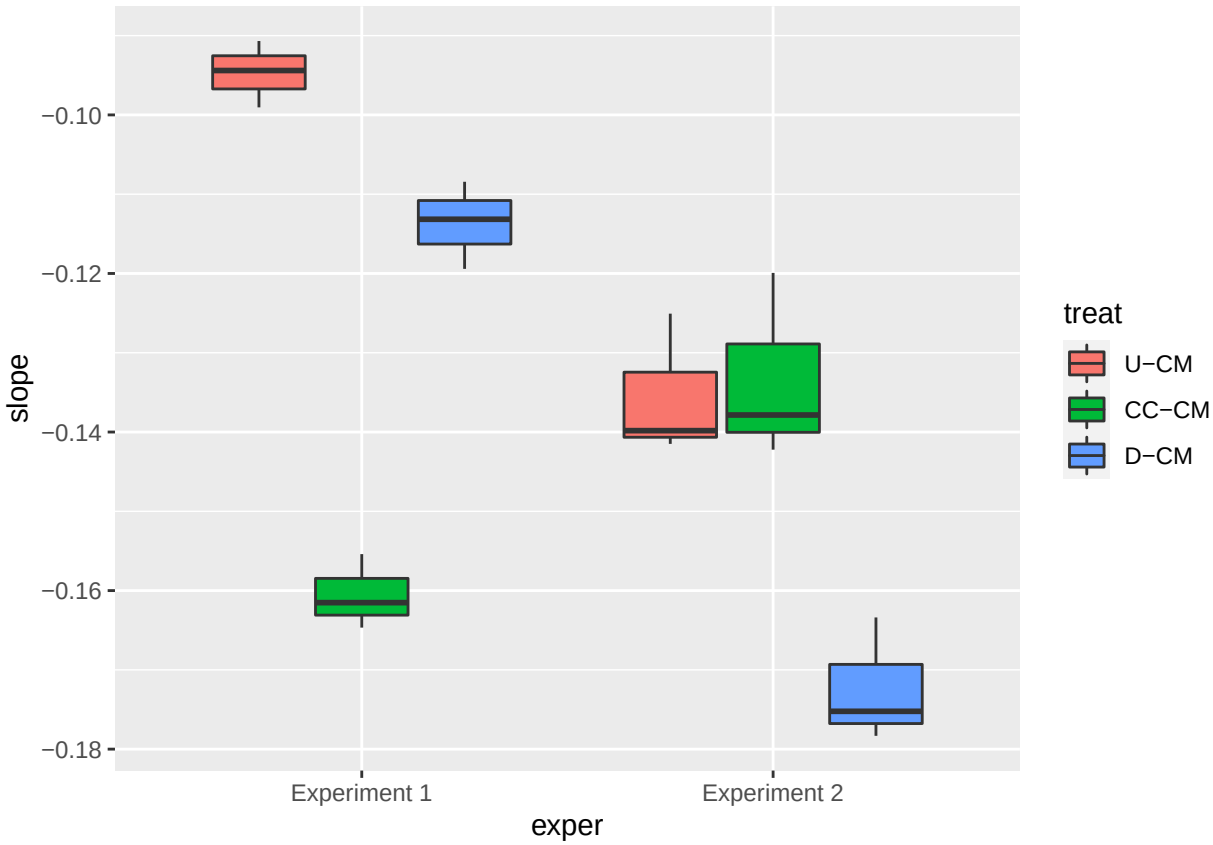
Unit of analysis will be wind tunnel plot.

First fit linear model to each wind tunnel to get the slopes.

```
lmods <- dat[, .(int = coef(lm(log(oav) ~ time))[1],
                 slope = coef(lm(log(oav) ~ time))[2]),
             by = .(exper, treat, rep)]
```

Take a look at slopes.

```
ggplot(lmods, aes(exper, slope, fill = treat)) +
  geom_boxplot()
```



Clear differences for D-CM in both experiments, but no difference for CC-CM in experiment 2.

Now analysis.

Look at slope.

```
modslope1 <- lm(slope ~ treat * exper, data = lmods)
summary.aov(modslope1)
```

```
##             Df   Sum Sq   Mean Sq F value   Pr(>F)
## treat        2 0.003617 0.0018086   30.69 1.91e-05 ***
## exper        1 0.002607 0.0026067   44.23 2.36e-05 ***
## treat:exper  2 0.006153 0.0030763   52.20 1.20e-06 ***
## Residuals   12 0.000707 0.0000589
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(modslope1)
```

```
##
## Call:
## lm(formula = slope ~ treat * exper, data = lmods)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.008889 -0.004479 -0.001953  0.004855  0.013407
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)                  -0.094707   0.004432 -21.369 6.43e-11 ***
## treatCC-CM                   -0.065823   0.006268 -10.502 2.11e-07 ***
## treatD-CM                    -0.018963   0.006268  -3.025   0.0106 *
## experExperiment 2            -0.040756   0.006268  -6.502 2.93e-05 ***
## treatCC-CM:experExperiment 2  0.067950   0.008864   7.666 5.80e-06 ***
## treatD-CM:experExperiment 2  -0.017886   0.008864  -2.018   0.0665 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.007677 on 12 degrees of freedom
## Multiple R-squared:  0.946,  Adjusted R-squared:  0.9234
## F-statistic:    42 on 5 and 12 DF,  p-value: 3.405e-07
```

Interactions complicated. Let's look by experiment. First experiment 1.

```
modexp1 <- aov(slope ~ treat, data = lmods, subset = exper == 'Experiment 1')
summary(modexp1)
```

```
##             Df   Sum Sq  Mean Sq F value   Pr(>F)
## treat        2 0.006888 0.003444   147.3 7.96e-06 ***
## Residuals    6 0.000140 0.000023
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary.lm(modexp1)
```

```
##
## Call:
## aov(formula = slope ~ treat, data = lmods, subset = exper ==
##     "Experiment 1")
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0057472 -0.0041397  0.0003088  0.0040298  0.0052367
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.094707   0.002792 -33.921 4.37e-08 ***
## treatCC-CM  -0.065823   0.003948 -16.671 2.97e-06 ***
## treatD-CM   -0.018963   0.003948  -4.803  0.00299 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.004836 on 6 degrees of freedom
## Multiple R-squared:   0.98,  Adjusted R-squared:  0.9734
## F-statistic: 147.3 on 2 and 6 DF,  p-value: 7.956e-06
```

```
coef(modexp1)
```

```
## (Intercept)  treatCC-CM   treatD-CM
## -0.09470724 -0.06582257 -0.01896326
```

```
confint(modexp1)
```

```
##                    2.5 %       97.5 %
## (Intercept) -0.10153891 -0.087875564
## treatCC-CM  -0.07548402 -0.056161130
```

```
## treatD-CM   -0.02862471 -0.009301821
```

```
model.tables(modexp1, type = 'means')
```

```
## Tables of means
## Grand mean
##
## -0.1229692
##
##  treat
## treat
##     U-CM    CC-CM     D-CM
## -0.09471 -0.16053 -0.11367
```

Use this model in paper. Both D-CM and CC-CM have lower slope than reference in experiment 1. Results are a first-order constant with units of 1/hr that describes the drop in flux over time. Note that `log()` (not `log10()`) used above gives this result.

Experiment 2 next.

```
modexp2 <- aov(slope ~ treat, data = lmods, subset = exper == 'Experiment 2')
summary(modexp2)
```

```
##             Df    Sum Sq   Mean Sq F value  Pr(>F)
## treat        2 0.0028816 0.0014408   15.25 0.00444 **
## Residuals    6 0.0005668 0.0000945
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary.lm(modexp2)
```

```
##
## Call:
## aov(formula = slope ~ treat, data = lmods, subset = exper ==
##     "Experiment 2")
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.008889 -0.006008 -0.004363  0.008924  0.013407
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.135463   0.005612 -24.139 3.32e-07 ***
## treatCC-CM   0.002128   0.007936   0.268  0.79760
## treatD-CM   -0.036850   0.007936  -4.643  0.00353 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.00972 on 6 degrees of freedom
## Multiple R-squared:  0.8356, Adjusted R-squared:  0.7808
## F-statistic: 15.25 on 2 and 6 DF,  p-value: 0.004441
```

```
coef(modexp2)
```

```
##  (Intercept)   treatCC-CM    treatD-CM
## -0.135463171  0.002127718 -0.036849566
```

```r
confint(modexp2)
```

```
##                   2.5 %      97.5 %
## (Intercept) -0.14919457 -0.12173177
## treatCC-CM  -0.01729141  0.02154685
## treatD-CM   -0.05626870 -0.01743043
```

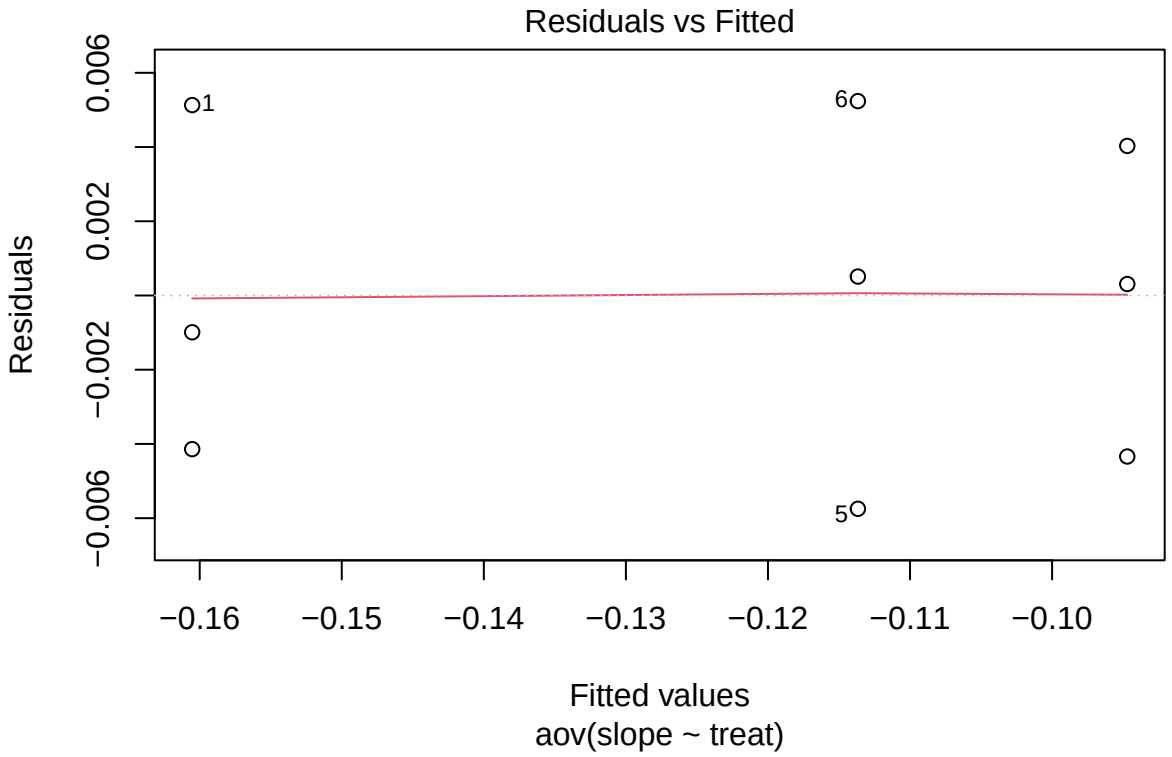
```r
model.tables(modexp2, type = 'means')
```

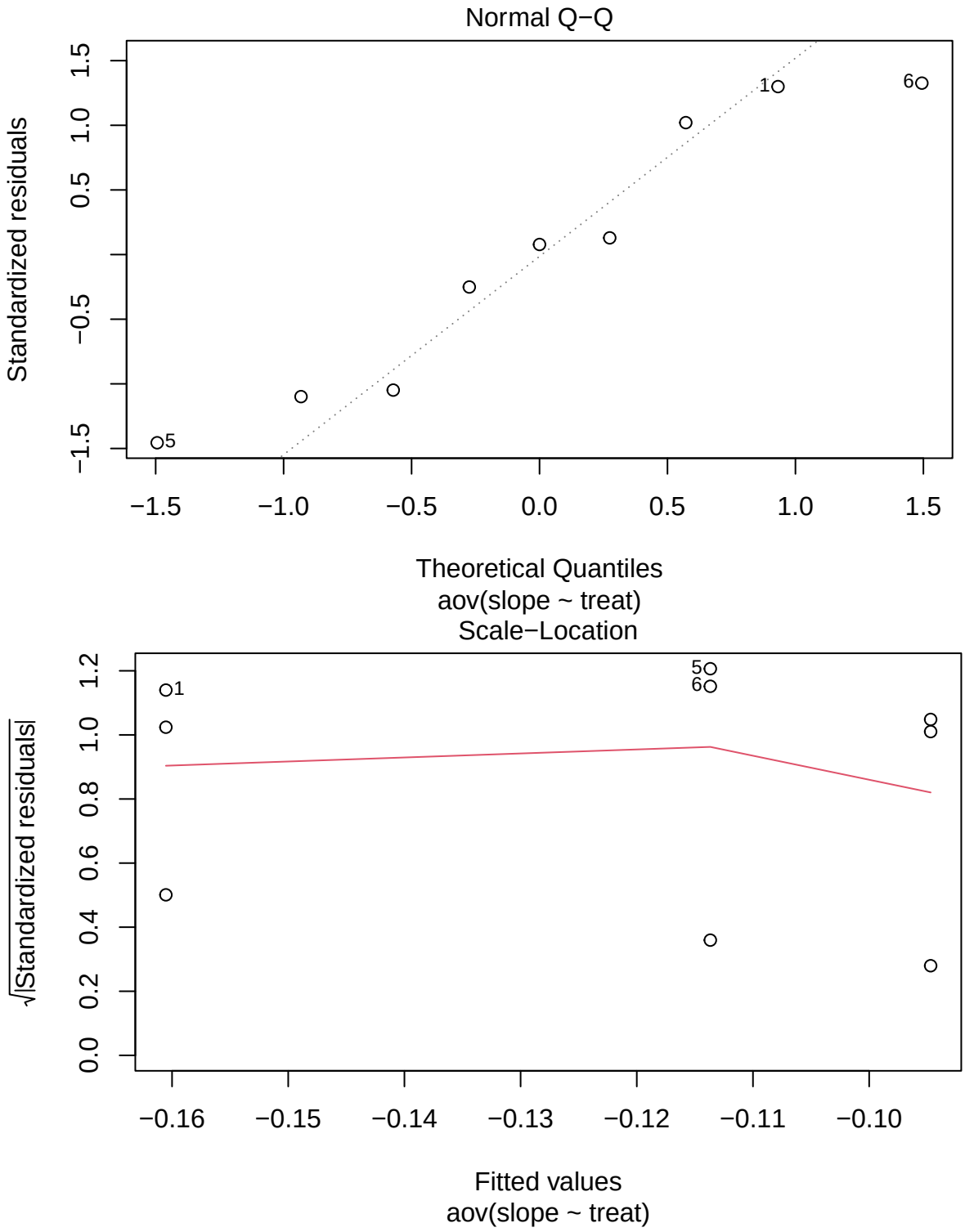
```
## Tables of means
## Grand mean
##
## -0.1470371
##
##  treat
## treat
##     U-CM    CC-CM     D-CM
## -0.13546 -0.13334 -0.17231
```

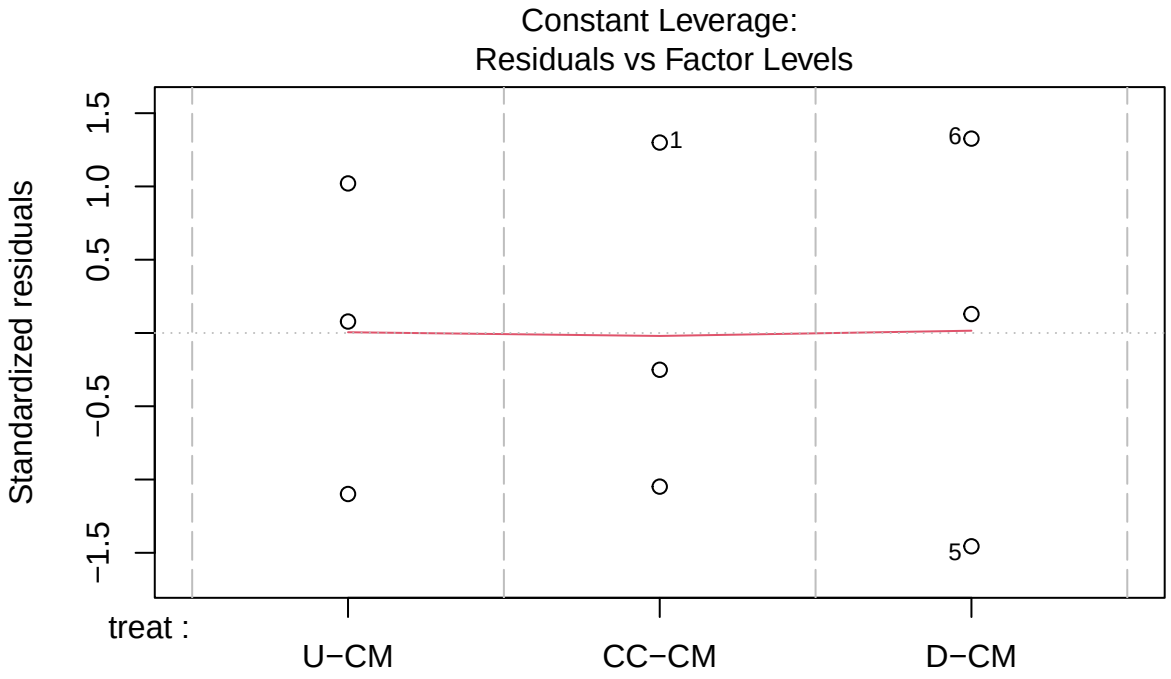
Here only D-CM is lower, and the effect seems larger than in experiment 1. **Use this model in paper for experiment 2.**

Diagnostic plots.

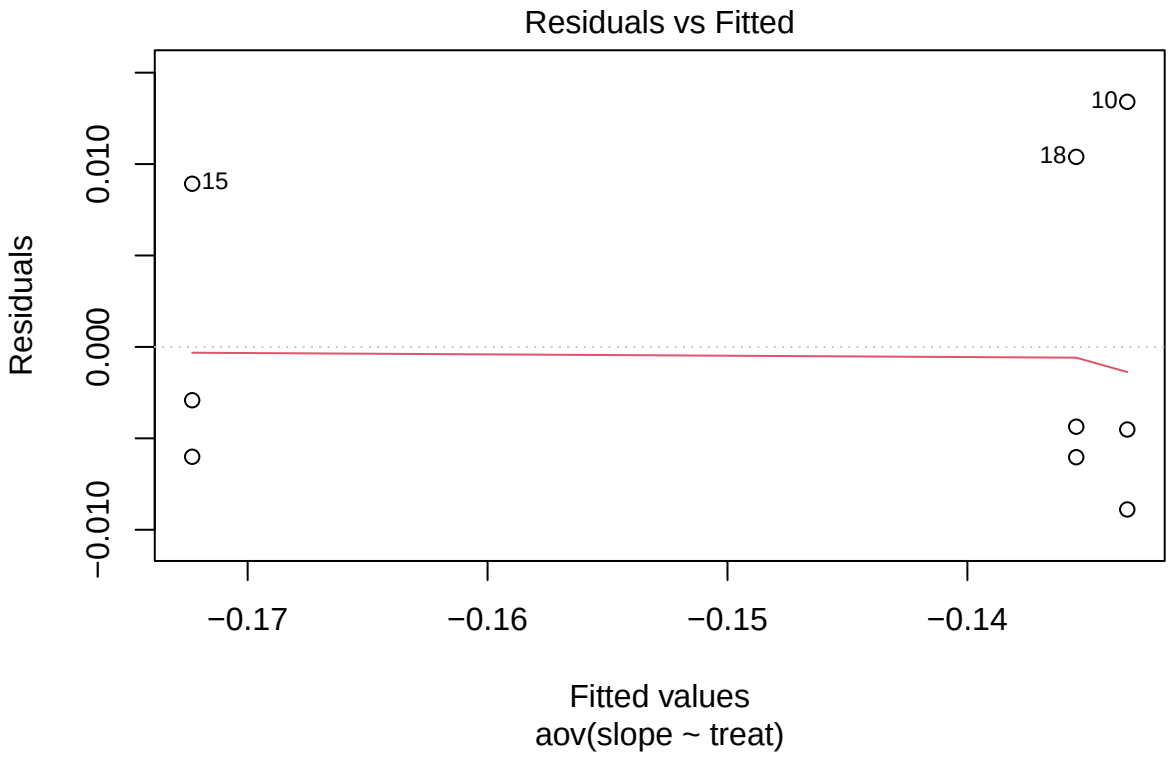
```r
plot(modexp1, ask = FALSE)
```

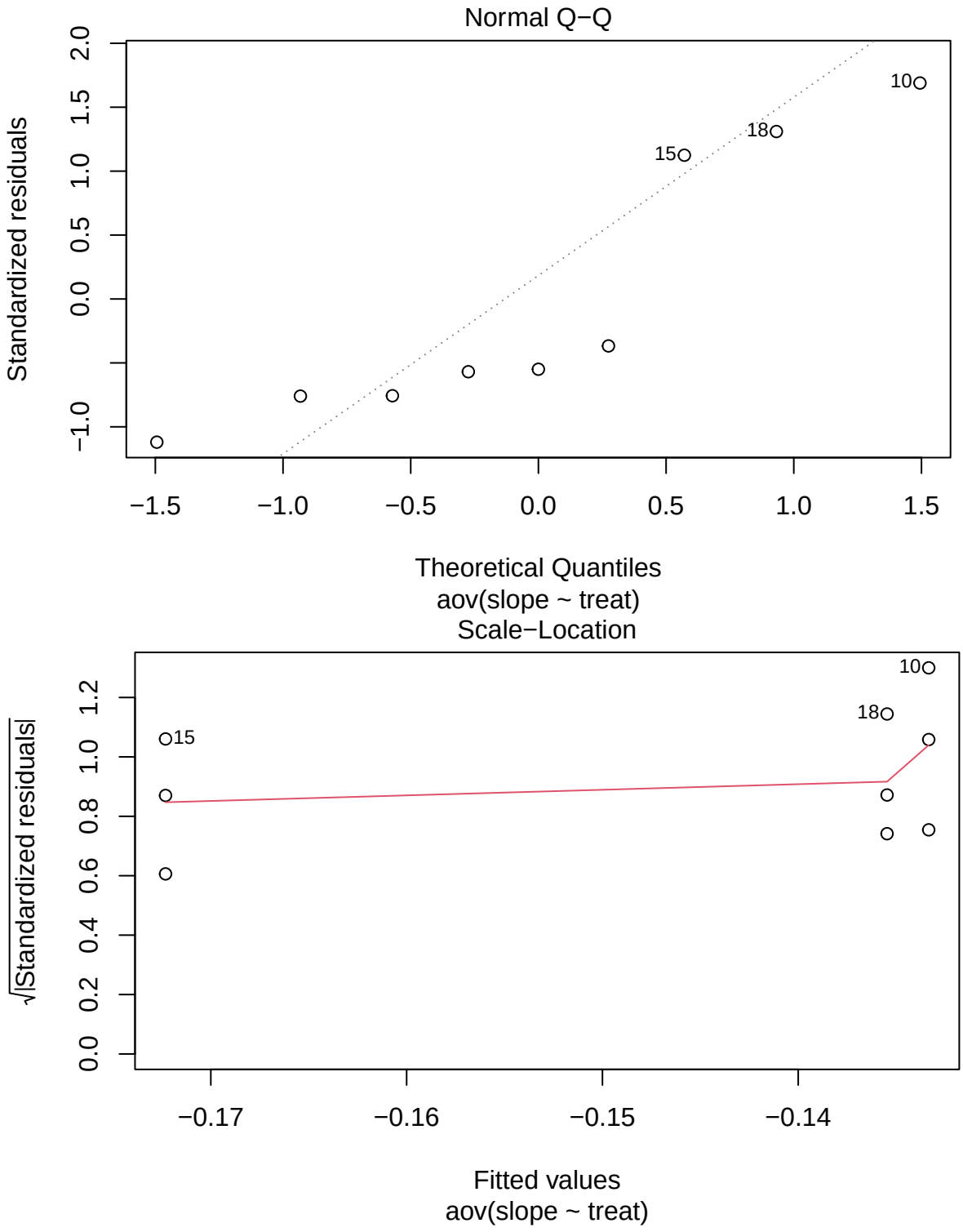


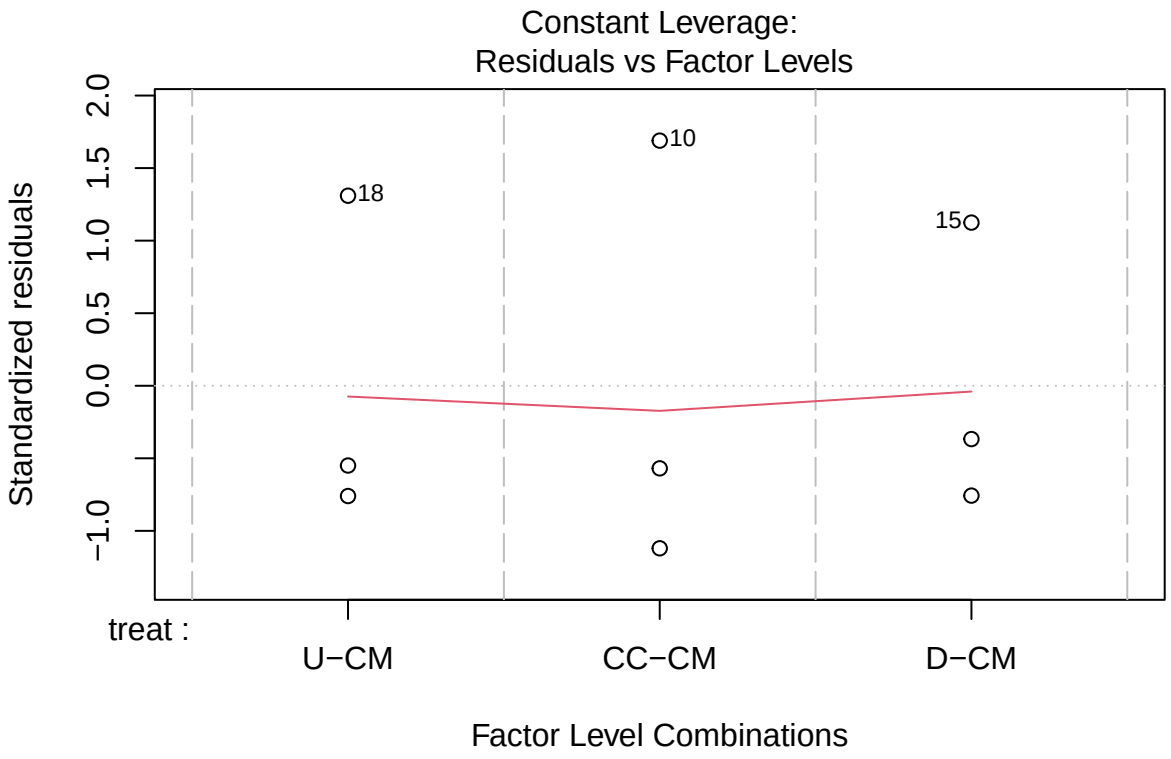

Normal Q−Q
Standardized residuals
Theoretical Quantiles
aov(slope ~ treat)
Scale−Location
√|Standardized residuals|
Fitted values
aov(slope ~ treat)




```
plot(modexp2, ask = FALSE)
```




Normal Q−Q
Standardized residuals
2.0
1.5
1.0
0.5
0.0
−1.0
10
18
15
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
Theoretical Quantiles
aov(slope ~ treat)
Scale−Location
√|Standardized residuals|
1.2
1.0
0.8
0.6
0.4
0.2
0.0
10
18
15
−0.17
−0.16
−0.15
−0.14
Fitted values
aov(slope ~ treat)



Not terrible.

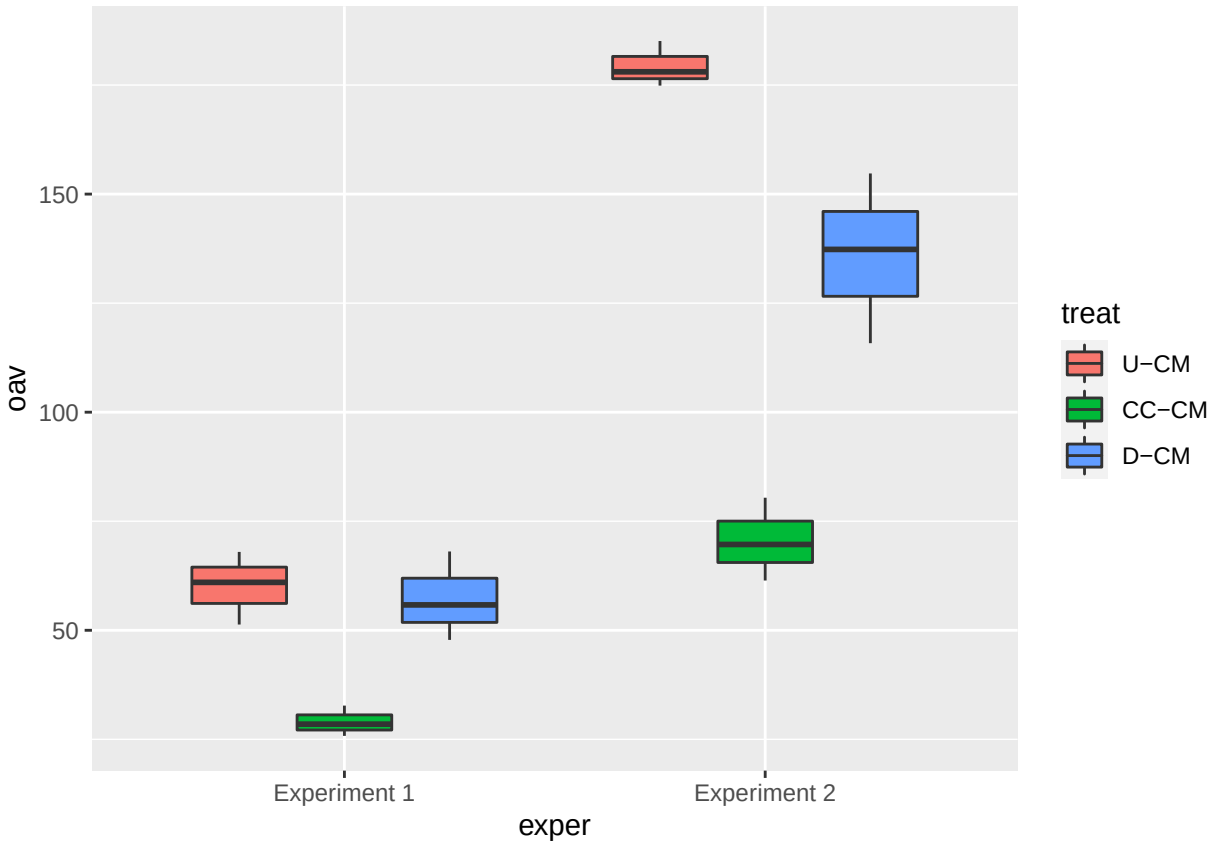
Now look at first measurement period in lieu of intercept.

Add interval number to data.

```
dat <- dat[, int := as.integer(factor(time)), by = exper]
```

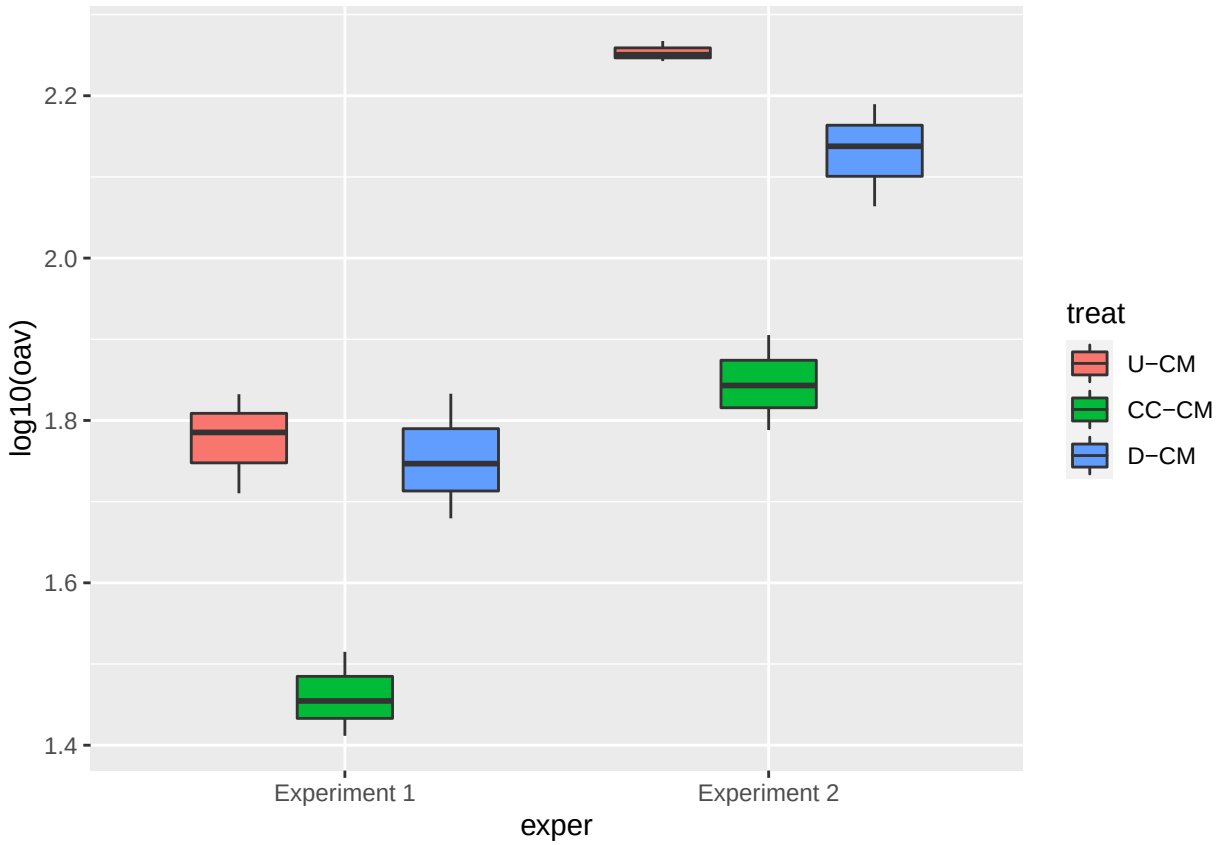
And take a look at the initial OAV values.

```
dd <- subset(dat, int == 1)
ggplot(dd, aes(exper, oav, fill = treat)) +
  geom_boxplot()
```



Check transformed values.

```
ggplot(dd, aes(exper, log10(oav), fill = treat)) +
  geom_boxplot()
```



See CC-CM lower in both experiments, D-CM lower in experiment 2 only.

```
modinit1 <- aov(log10(oav) ~ treat * exper, data = dat, subset = int == 1)
summary(modinit1)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## treat        2 0.4393  0.2196   65.92 3.37e-07 ***
## exper        1 0.7689  0.7689  230.77 3.36e-09 ***
## treat:exper  2 0.0093  0.0047    1.40    0.284
## Residuals   12 0.0400  0.0033
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Drop interaction.

```
modinit2 <- aov(log10(oav) ~ treat + exper, data = dat, subset = int == 1)
summary(modinit2)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## treat        2 0.4393  0.2196   62.36 1.07e-07 ***
## exper        1 0.7689  0.7689  218.30 6.22e-10 ***
## Residuals   14 0.0493  0.0035
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary.lm(modinit2)
```

```
##
## Call:
## aov(formula = log10(oav) ~ treat + exper, data = dat, subset = int ==
```

```
##     1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.09777 -0.03174  0.01001  0.03816  0.09791
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.80808    0.02798  64.627  < 2e-16 ***
## treatCC-CM        -0.36182    0.03426 -10.560 4.74e-08 ***
## treatD-CM         -0.07304    0.03426  -2.132   0.0512 .
## experExperiment 2  0.41336    0.02798  14.775 6.22e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05935 on 14 degrees of freedom
## Multiple R-squared:  0.9608, Adjusted R-squared:  0.9524
## F-statistic: 114.3 on 3 and 14 DF,  p-value: 4.403e-10
```

```
100 * (1 - 10^coef(modinit2))
```

```
##       (Intercept)        treatCC-CM         treatD-CM experExperiment 2
##       -6328.03822          56.53087          15.47974        -159.03353
```

```
100 * (1 - 10^confint(modinit2))
```

```
##                         2.5 %      97.5 %
## (Intercept)       -5498.52660 -7280.45530
## treatCC-CM           63.29793    48.51610
## treatD-CM            28.63745    -0.10396
## experExperiment 2  -125.60633  -197.41350
```

```
100 * (10^coef(modinit2))
```

```
##       (Intercept)        treatCC-CM         treatD-CM experExperiment 2
##        6428.03822          43.46913          84.52026         259.03353
```

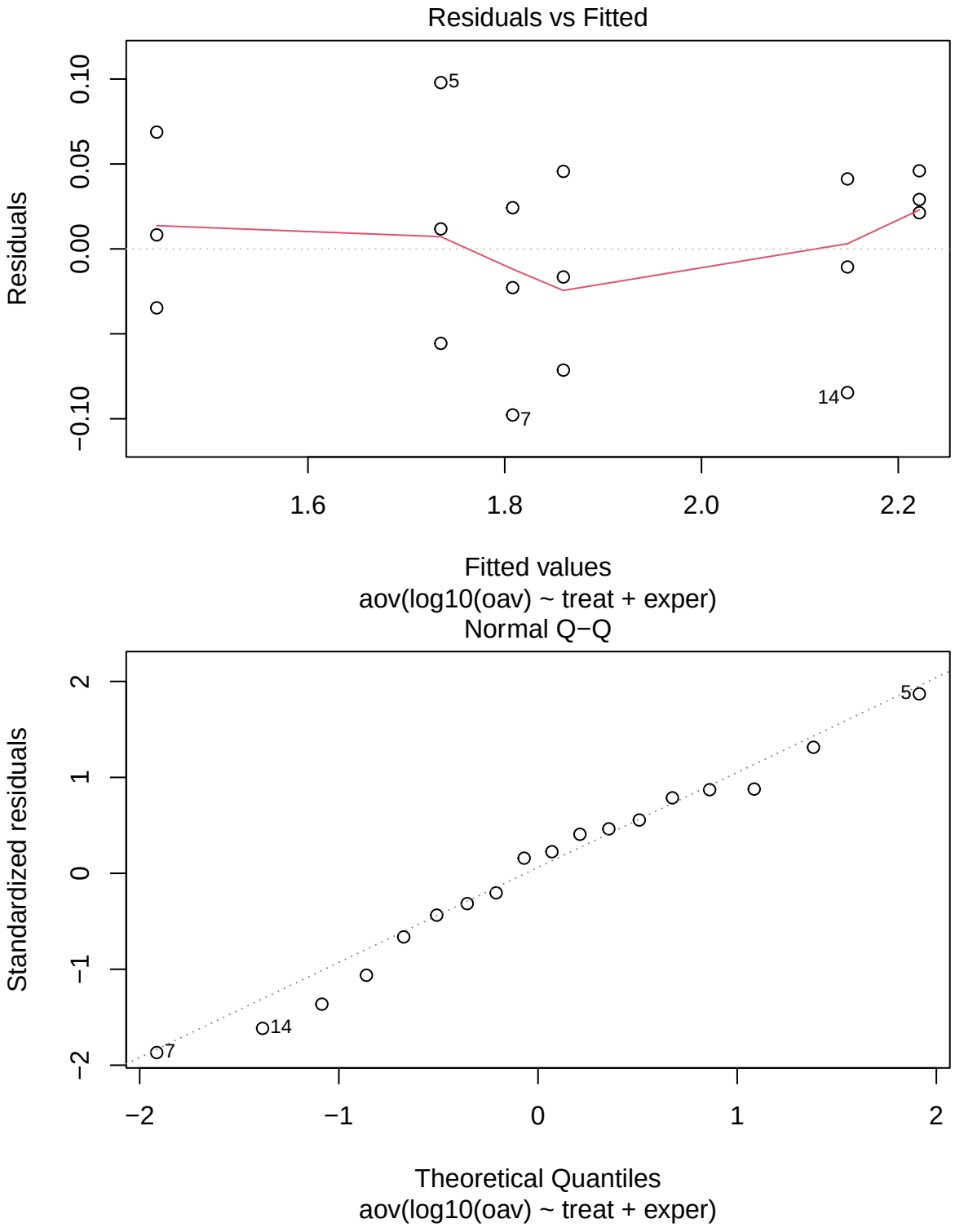
```
100 * (10^confint(modinit2))
```

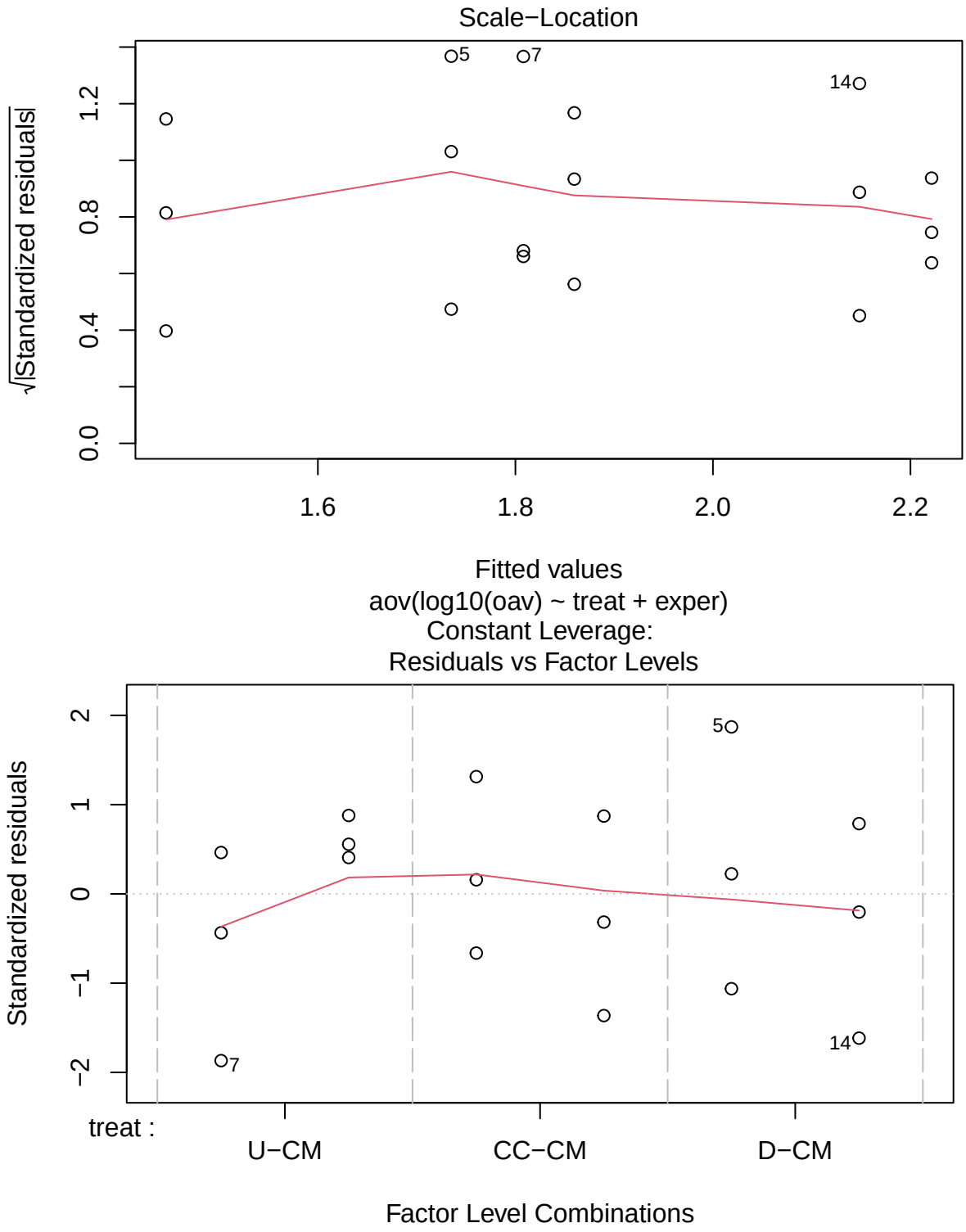
```
##                        2.5 %    97.5 %
## (Intercept)       5598.52660 7380.4553
## treatCC-CM          36.70207   51.4839
## treatD-CM           71.36255  100.1040
## experExperiment 2  225.60633  297.4135
```

Use this model in paper. Back-transformed coef and confint results give % reduction relative to reference. CC-CM clearly lower, D-CM not.

Check diagnostic plots.

```
plot(modinit2, ask = FALSE)
```


Residuals vs Fitted
Residuals
0.10
0.05
0.00
−0.10
5
7
14
1.6
1.8
2.0
2.2
Fitted values
aov(log10(oav) ~ treat + exper)
Normal Q−Q
Standardized residuals
2
1
0
−1
−2
5
14
7
−2
−1
0
1
2
Theoretical Quantiles
aov(log10(oav) ~ treat + exper)



These look good.