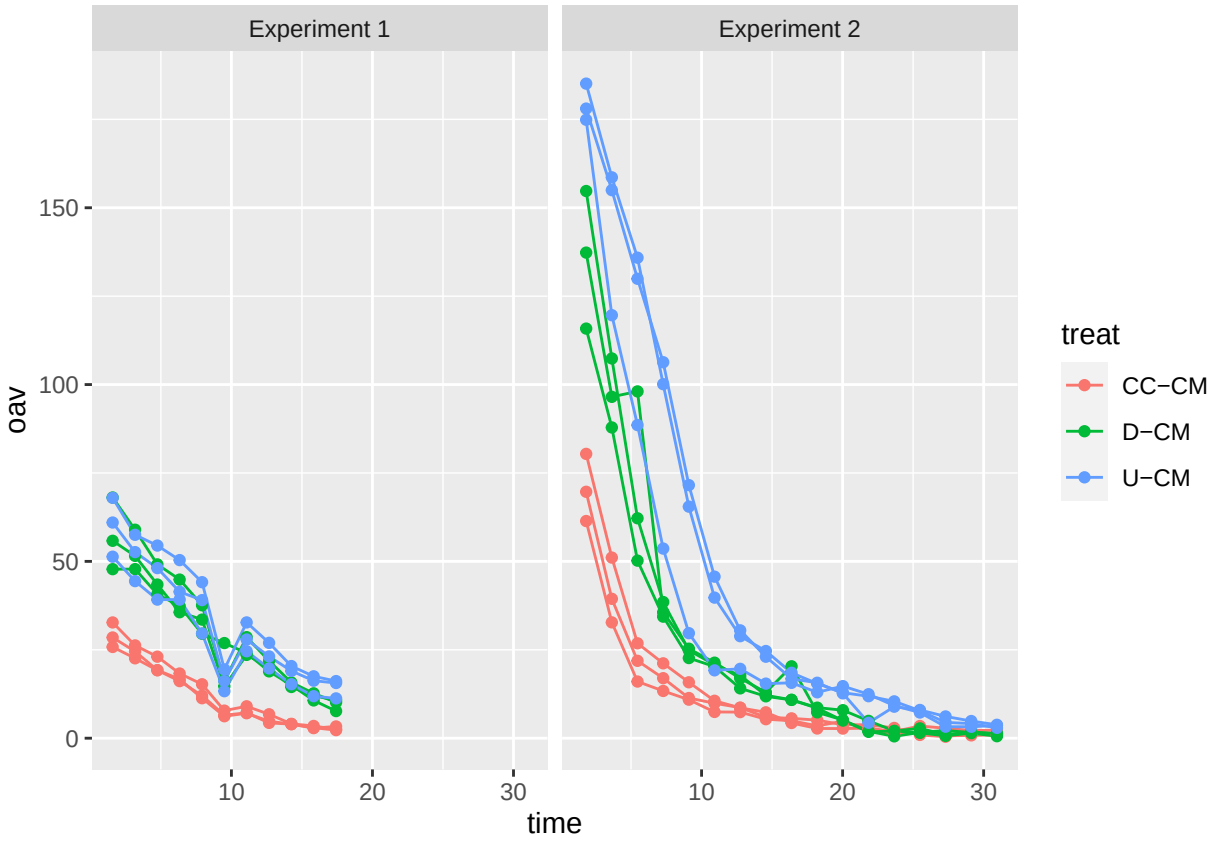
Data analysis for odor from digestate experiments

Sasha D. Hafner

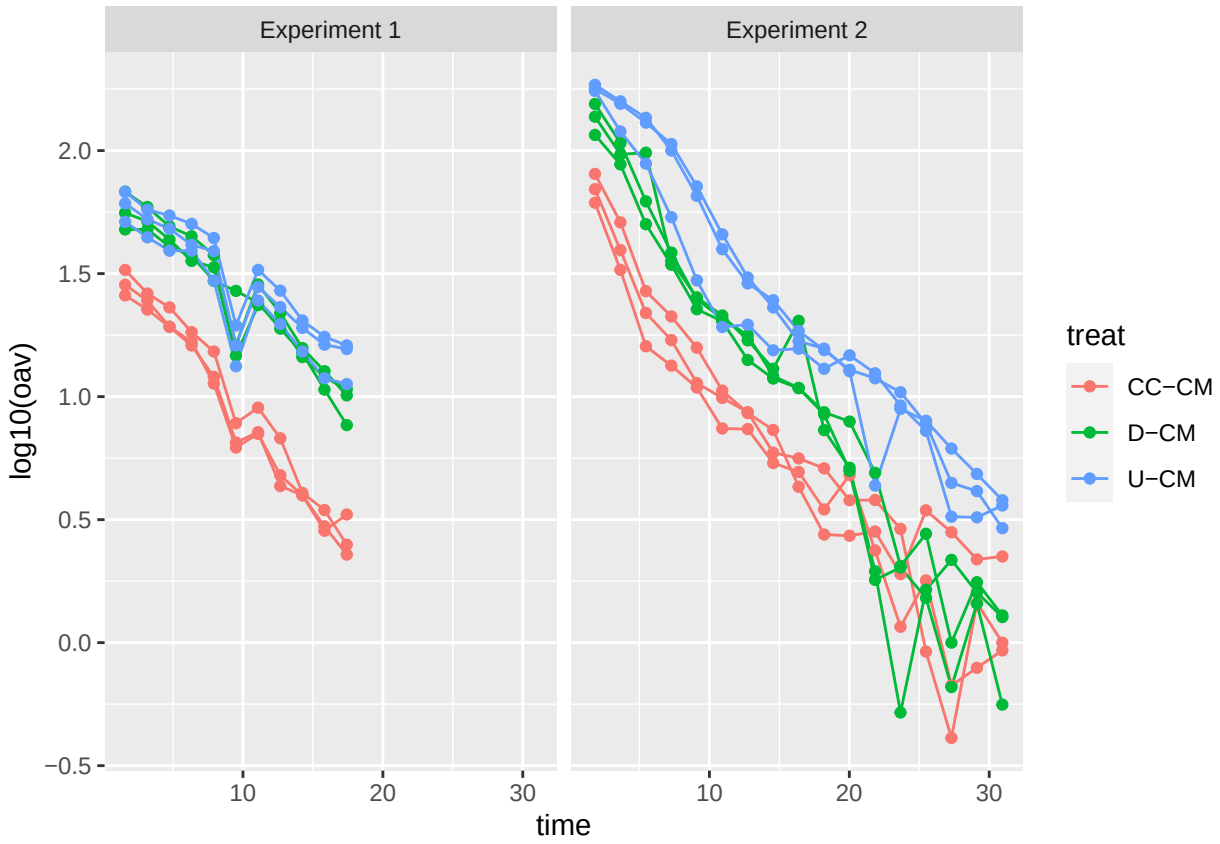
01 December, 2022

Plots

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



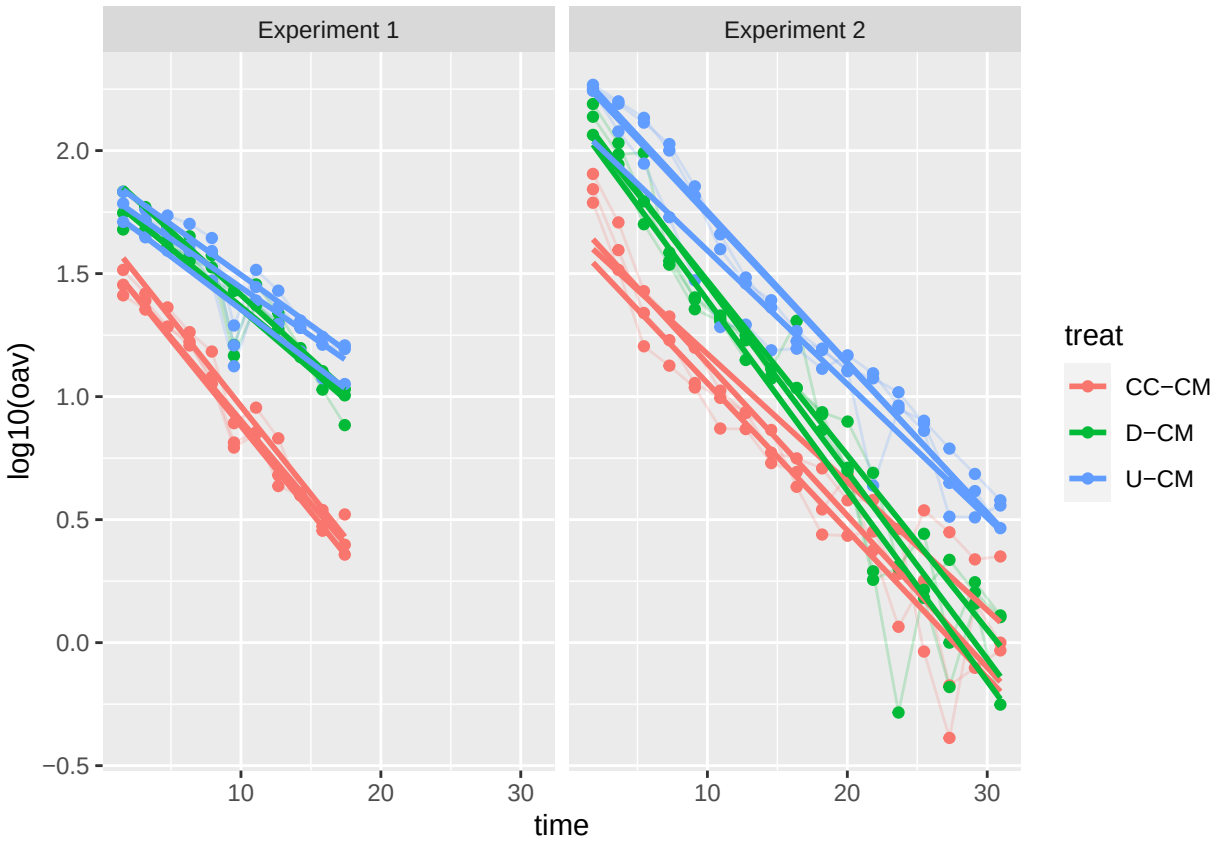

```
ggplot(dat, aes(time, log10(oav), colour = treat, group = interaction(rep, treat))) +
  geom_line() +
  geom_point() +
  facet_wrap(~ exper)
```



Looks linear enough.

```
ggplot(dat, aes(time, log10(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)
```

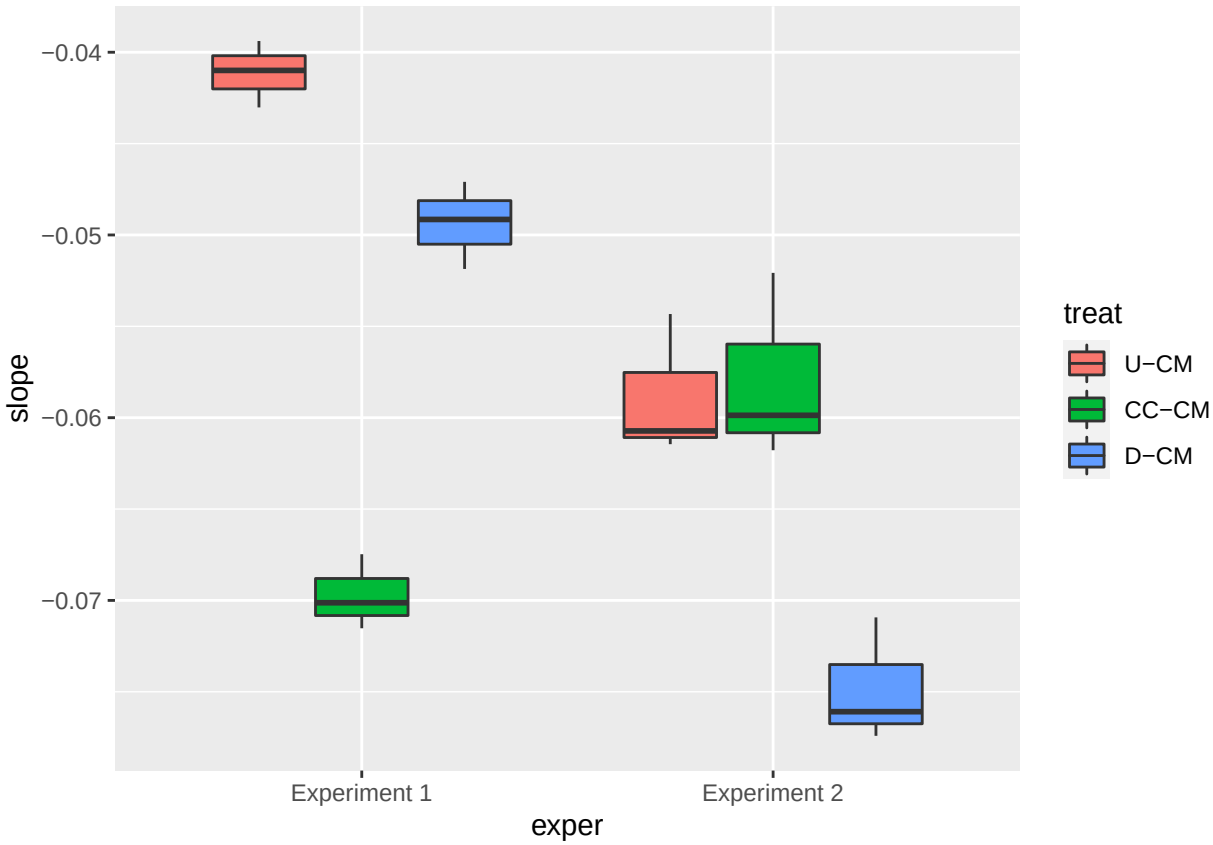
Unit of analysis will be wind tunnel plot.

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

```
lmods <- dat[, .(int = coef(lm(log10(oav) ~ time))[1],
                 slope = coef(lm(log10(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 0.0006813 0.0003406   30.60 1.94e-05 ***
## exper        1 0.0004912 0.0004912   44.12 2.39e-05 ***
## treat:exper  2 0.0011587 0.0005794   52.04 1.22e-06 ***
## Residuals   12 0.0001336 0.0000111
## ---
## 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
## -0.0038720 -0.0019442 -0.0008501  0.0021141  0.0058339
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)                  -0.041133   0.001926 -21.352 6.49e-11 ***
## treatCC-CM                   -0.028576   0.002724 -10.489 2.13e-07 ***
## treatD-CM                    -0.008234   0.002724  -3.022   0.0106 *
## experExperiment 2            -0.017700   0.002724  -6.497 2.95e-05 ***
## treatCC-CM:experExperiment 2  0.029502   0.003853   7.657 5.87e-06 ***
## treatD-CM:experExperiment 2  -0.007743   0.003853  -2.010   0.0675 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.003337 on 12 degrees of freedom
## Multiple R-squared:  0.9458, Adjusted R-squared:  0.9232
## F-statistic: 41.88 on 5 and 12 DF,  p-value: 3.463e-07
```

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

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

```
##             Df    Sum Sq   Mean Sq F value   Pr(>F)
## treat        2 0.0012982 0.0006491   146.8 8.04e-06 ***
## Residuals    6 0.0000265 0.0000044
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(modexp1)
```

```
##
## Call:
## lm(formula = slope ~ treat, data = lmods, subset = exper == "Experiment 1")
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0024915 -0.0018166  0.0001387  0.0017475  0.0022745
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.041133   0.001214 -33.877  4.4e-08 ***
## treatCC-CM  -0.028576   0.001717 -16.642  3.0e-06 ***
## treatD-CM   -0.008234   0.001717  -4.796  0.00301 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.002103 on 6 degrees of freedom
## Multiple R-squared:   0.98,  Adjusted R-squared:  0.9733
## F-statistic: 146.8 on 2 and 6 DF,  p-value: 8.037e-06
```

Both D-CM and CC-CM have lower slope than reference in experiment 1. **Use these results in paper.**

Experiment 2 now

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

```
##             Df    Sum Sq   Mean Sq F value  Pr(>F)
## treat        2 0.0005418 2.709e-04   15.18 0.00449 **
## Residuals    6 0.0001071 1.784e-05
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

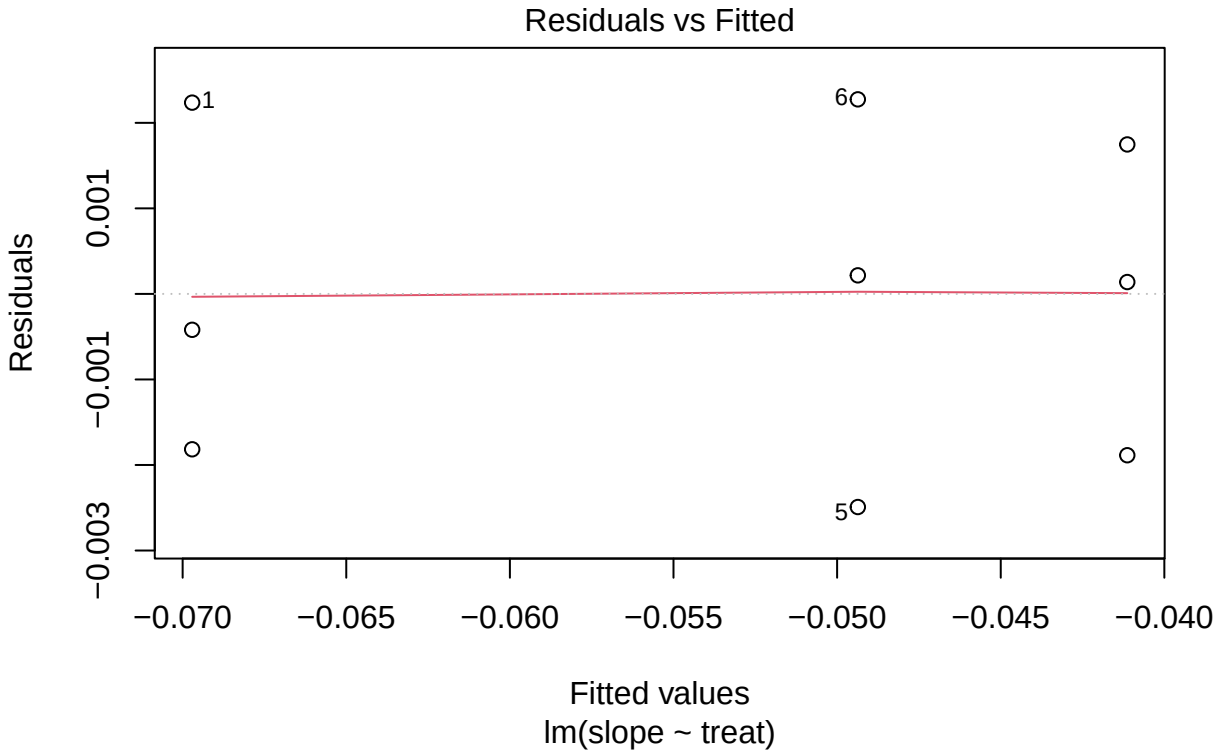
```
summary(modexp2)
```

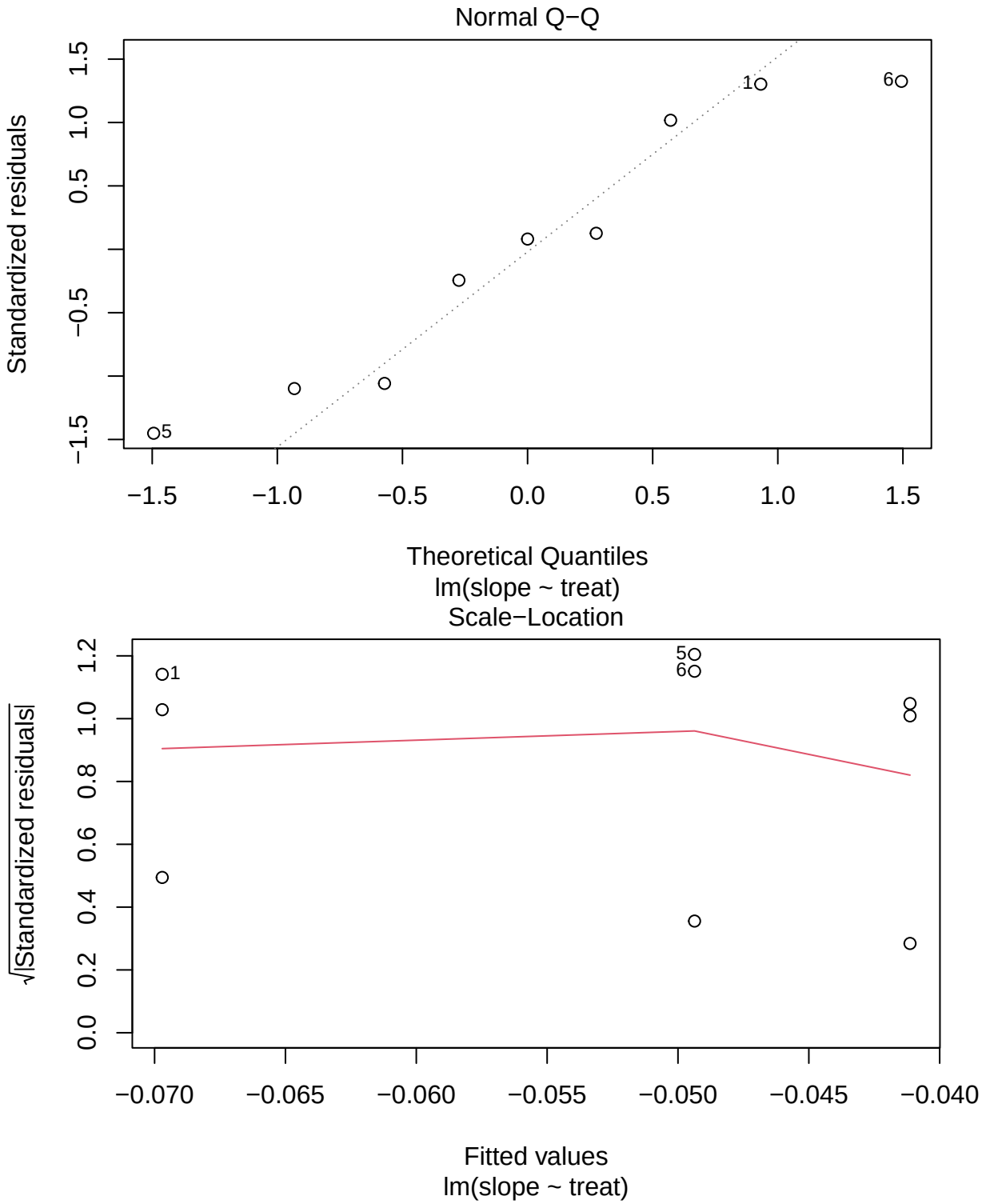
```
##
## Call:
## lm(formula = slope ~ treat, data = lmods, subset = exper == "Experiment 2")
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.003872 -0.002601 -0.001891  0.003881  0.005834
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.0588330  0.0024389 -24.123 3.33e-07 ***
## treatCC-CM   0.0009254  0.0034491   0.268  0.79745
## treatD-CM   -0.0159773  0.0034491  -4.632  0.00357 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.004224 on 6 degrees of freedom
## Multiple R-squared:  0.835,  Adjusted R-squared:   0.78
## F-statistic: 15.18 on 2 and 6 DF,  p-value: 0.004492
```

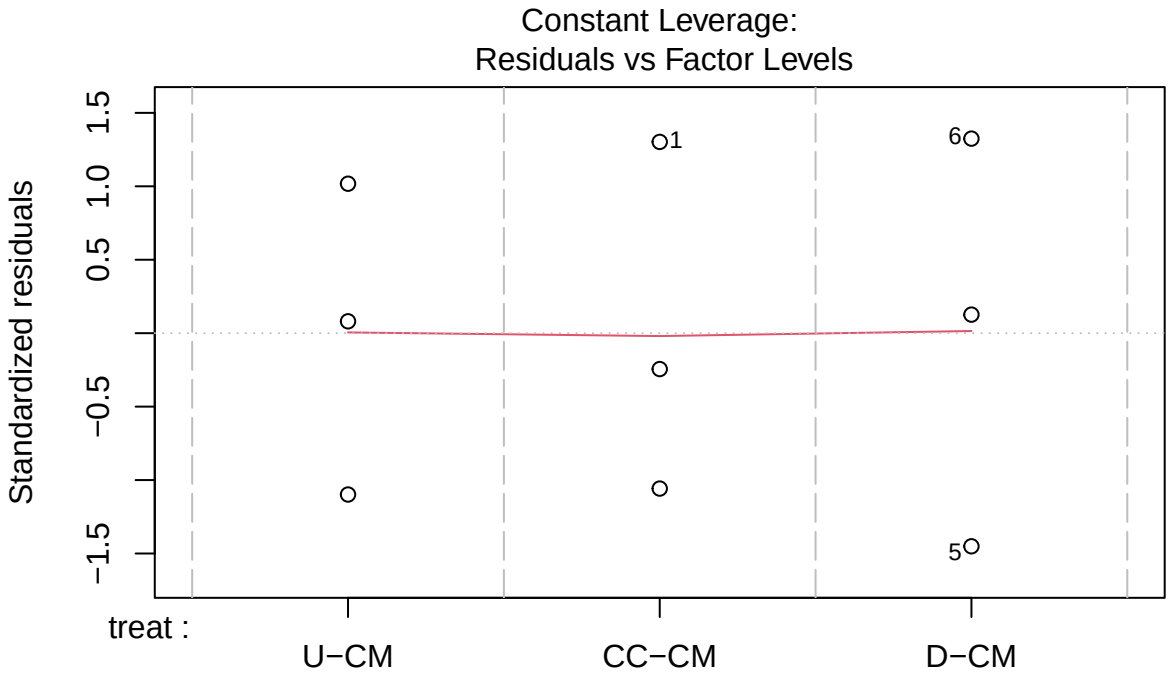
Here only D-CM is lower, and the effect seems larger than in experiment 1. **Use these results in paper.**

Diagnostic plots.

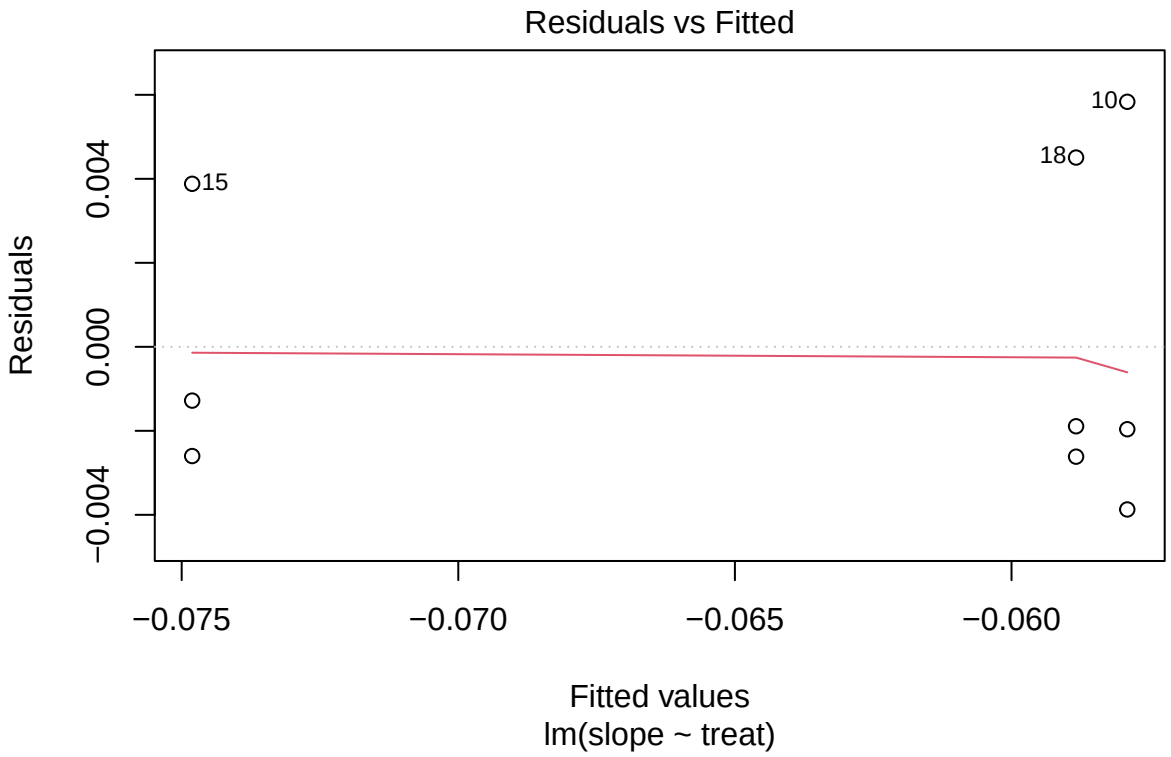
```
plot(modexp1, ask = FALSE)
```

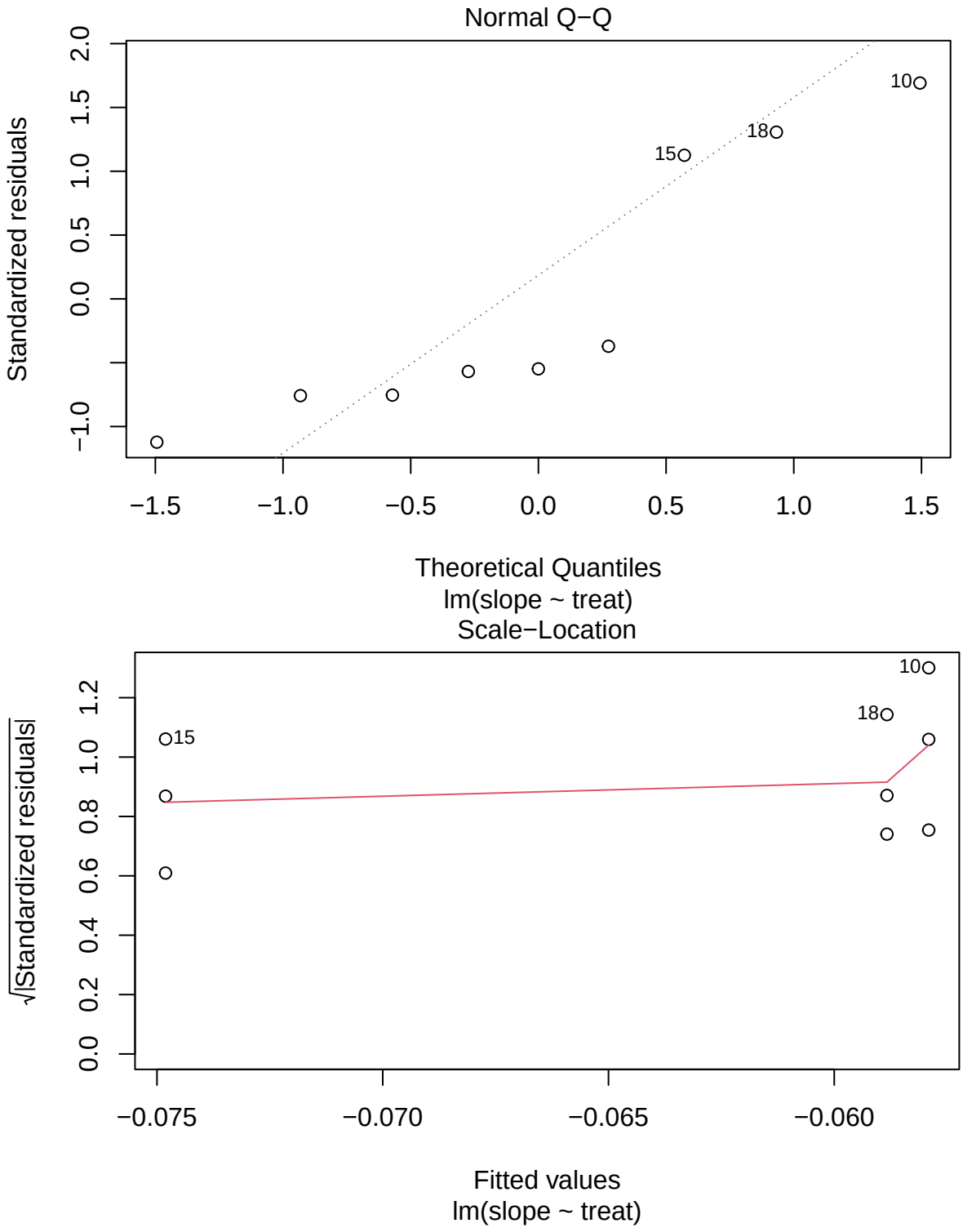


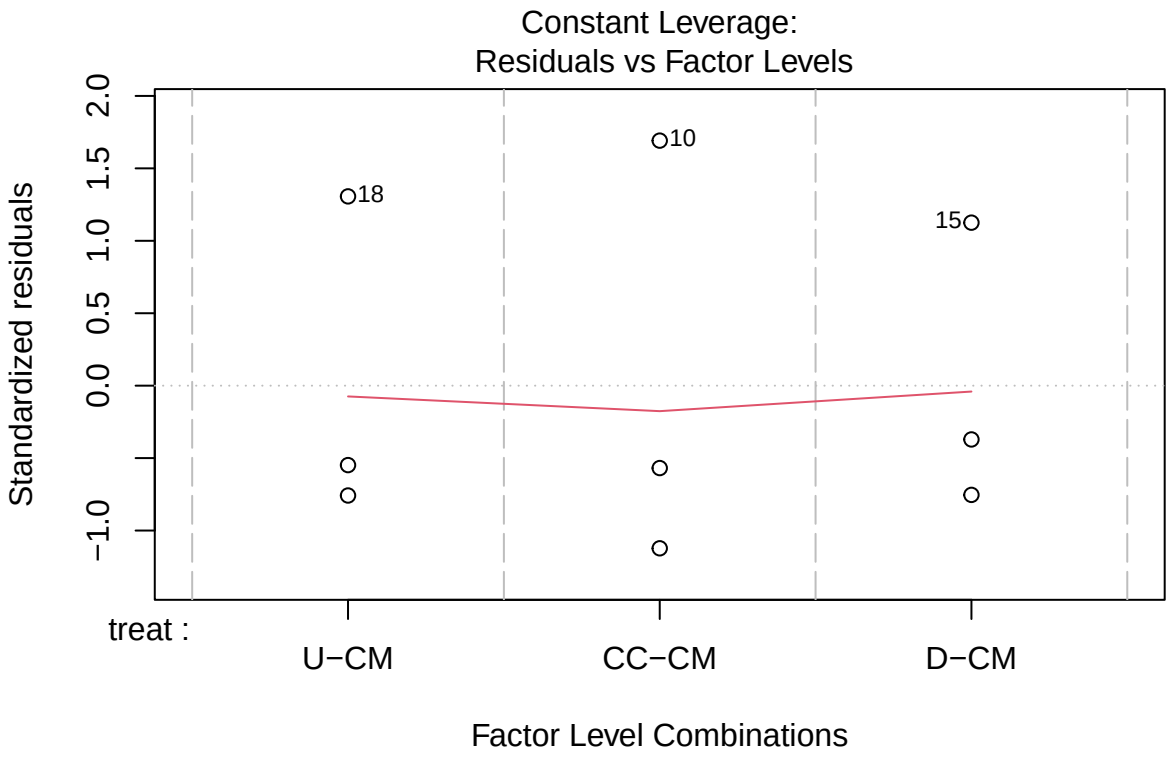

Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(slope ~ treat)
Scale−Location
√|Standardized residuals|
Fitted values
lm(slope ~ treat)




```
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
lm(slope ~ treat)
Scale−Location
√|Standardized residuals|
1.2
1.0
0.8
0.6
0.4
0.2
0.0
10
18
15
−0.075
−0.070
−0.065
−0.060
Fitted values
lm(slope ~ treat)



Not terrible.

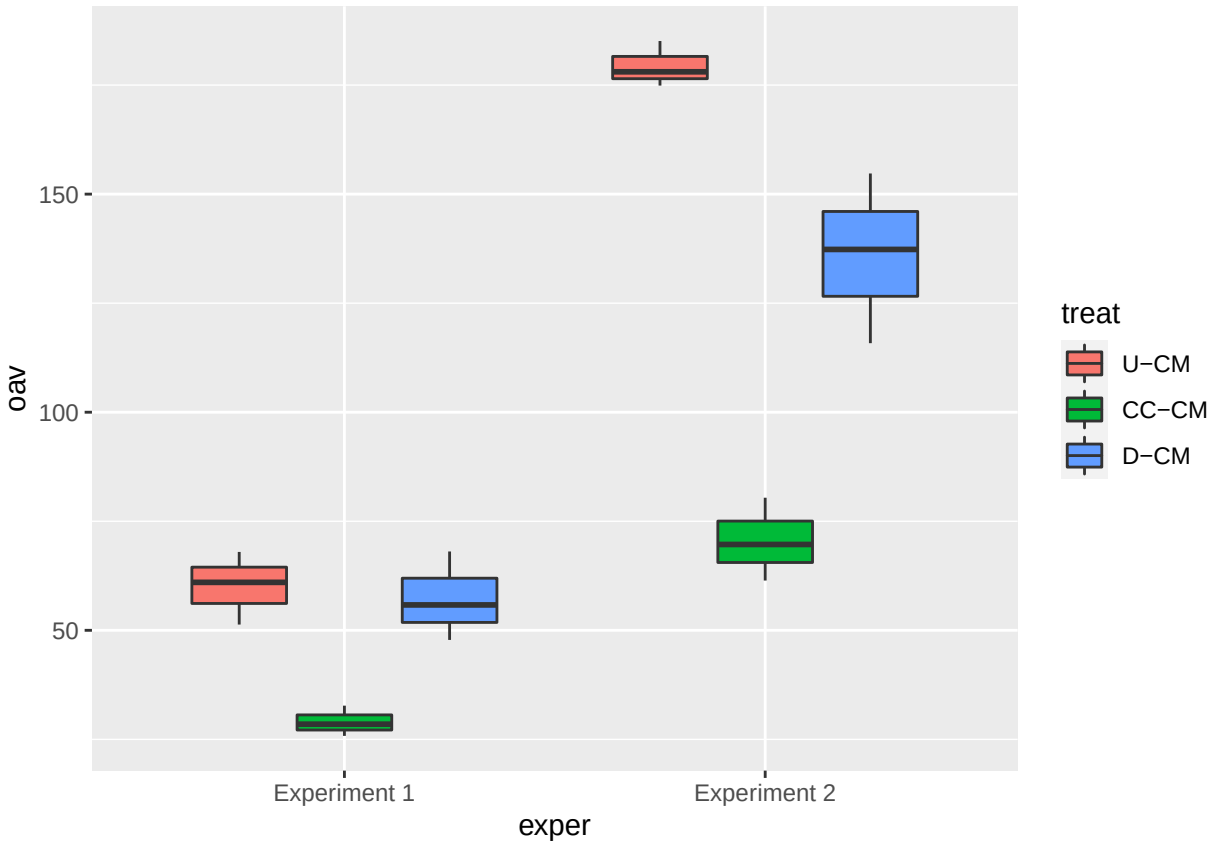
Now look at first measurement period in lieu of intercept.

Add interval number to data.

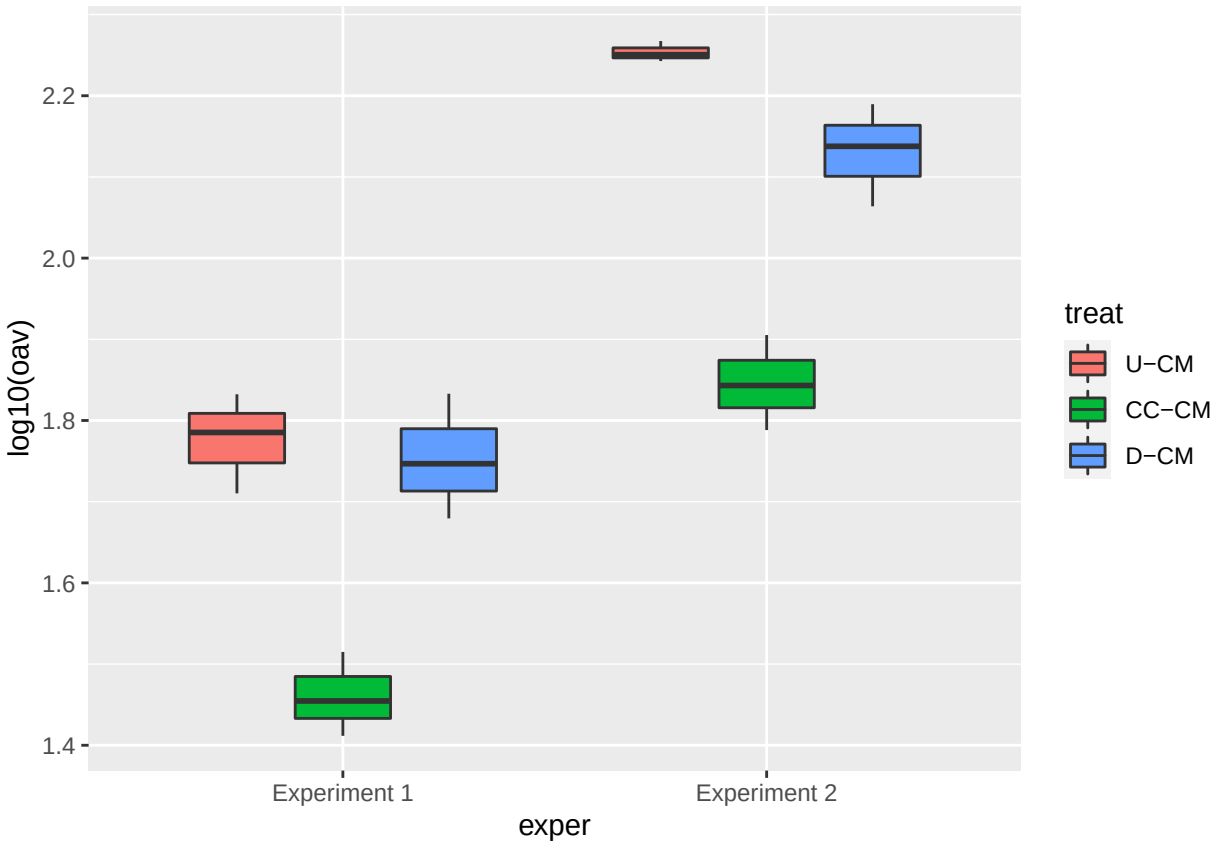
```
dat <- dat[, int := as.integer(factor(time)), by = exper]
```

And take a look at the initial OAV values.

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




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

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## treat        2 0.4392  0.2196  65.935 3.37e-07 ***
## exper        1 0.7689  0.7689 230.841 3.36e-09 ***
## treat:exper  2 0.0093  0.0047   1.401    0.284
## Residuals   12 0.0400  0.0033
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(modinit1)
```

```
##
## Call:
## lm(formula = log10(oav) ~ treat * exper, data = dat, subset = int ==
##     1)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.073630 -0.039265 -0.002732  0.044396  0.079898
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  1.77595    0.03332  53.298 1.25e-15 ***
## treatCC-CM                  -0.31559    0.04712  -6.697 2.21e-05 ***
```

```
## treatD-CM                    -0.02290    0.04712  -0.486    0.636
## experExperiment 2             0.47760    0.04712  10.135 3.10e-07 ***
## treatCC-CM:experExperiment 2 -0.09243    0.06664  -1.387    0.191
## treatD-CM:experExperiment 2  -0.10028    0.06664  -1.505    0.158
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05771 on 12 degrees of freedom
## Multiple R-squared:  0.9682, Adjusted R-squared:  0.955
## F-statistic:  73.1 on 5 and 12 DF,  p-value: 1.451e-08
```

Drop interaction.

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

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## treat        2 0.4392  0.2196   62.37 1.07e-07 ***
## exper        1 0.7689  0.7689  218.34 6.21e-10 ***
## Residuals   14 0.0493  0.0035
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

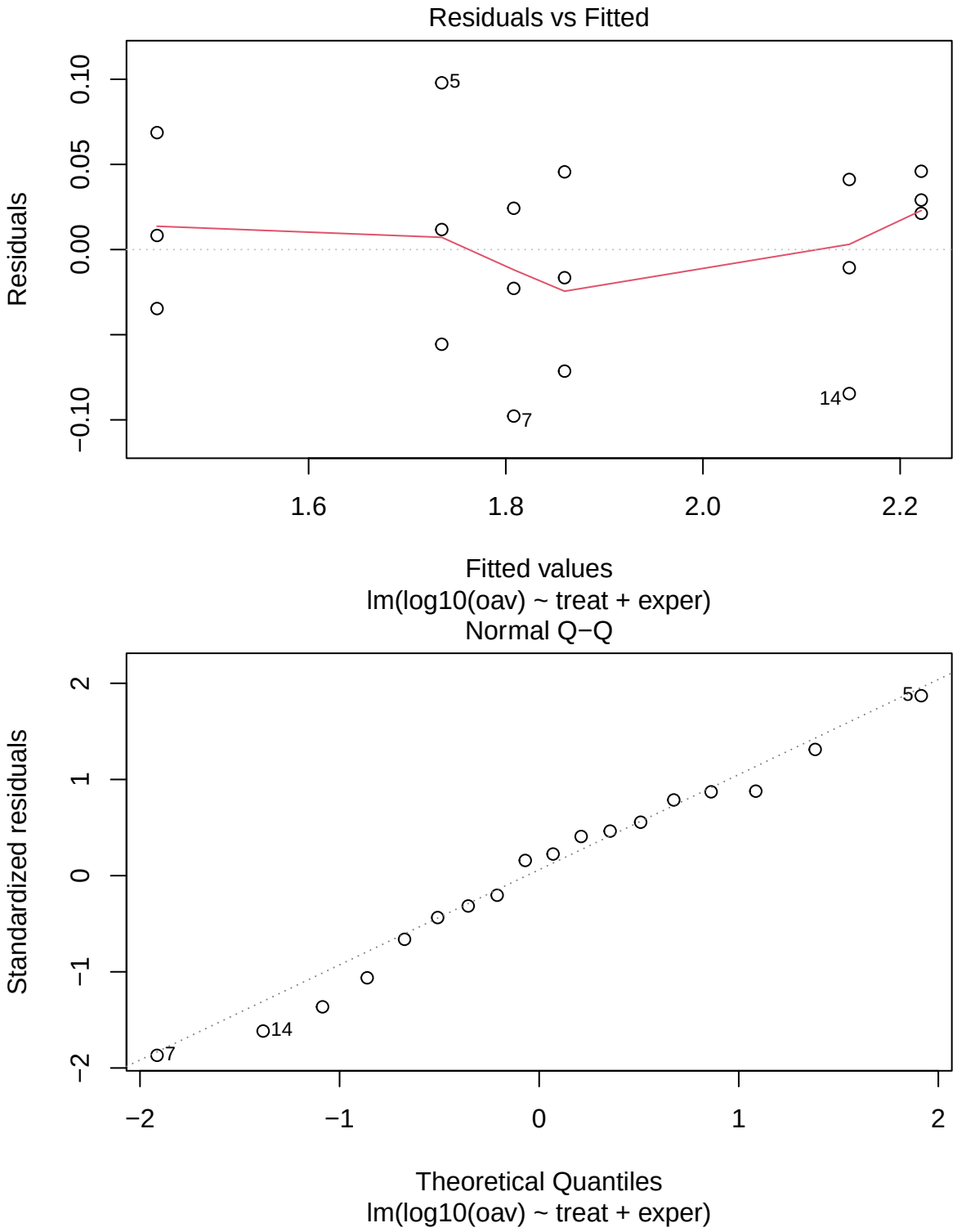
```
summary(modinit2)
```

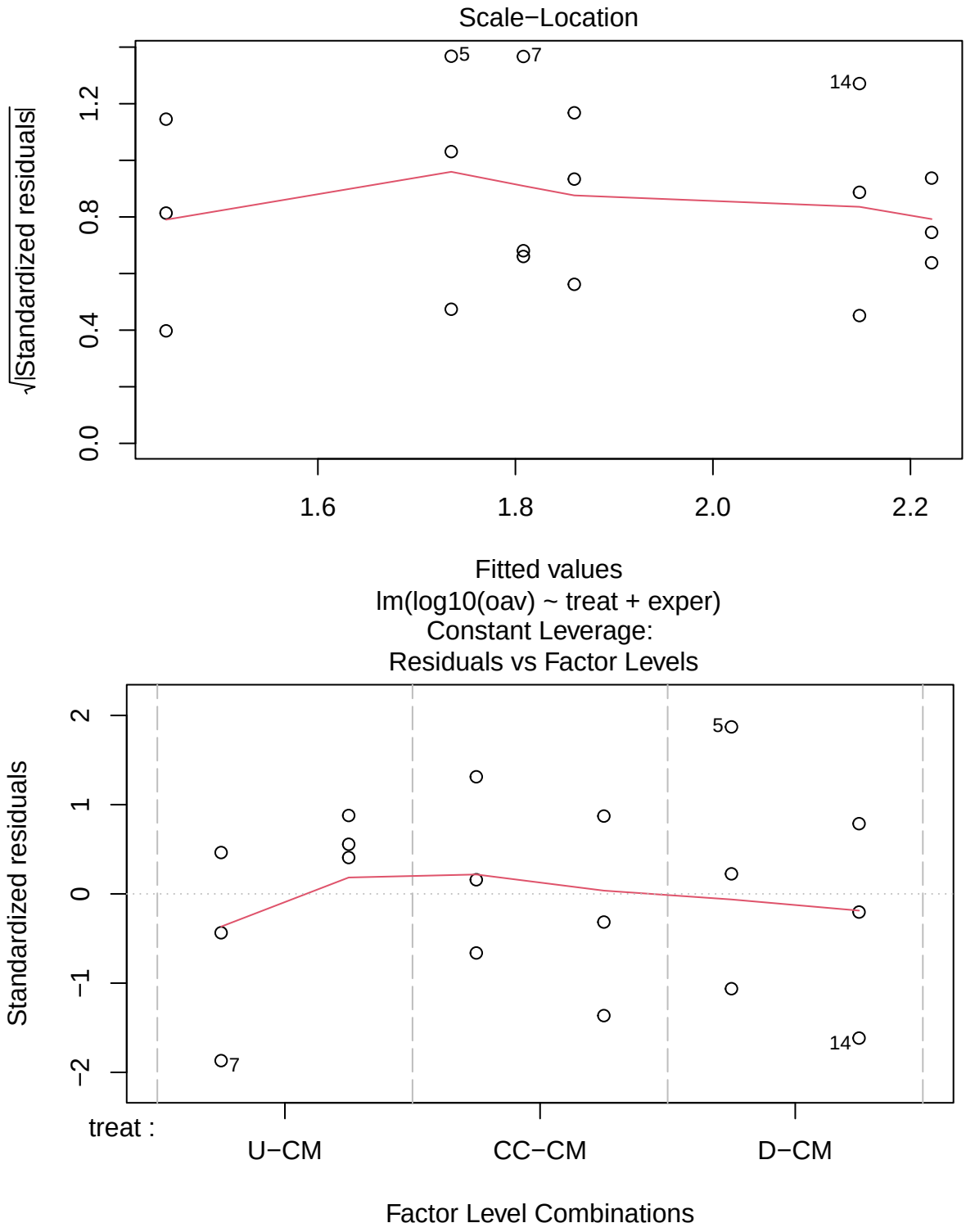
```
##
## Call:
## lm(formula = log10(oav) ~ treat + exper, data = dat, subset = int ==
##     1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.09779 -0.03169  0.01001  0.03815  0.09792
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.80807    0.02797  64.633  < 2e-16 ***
## treatCC-CM        -0.36180    0.03426 -10.560 4.74e-08 ***
## treatD-CM         -0.07304    0.03426  -2.132   0.0512 .
## experExperiment 2  0.41336    0.02797  14.776 6.21e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05934 on 14 degrees of freedom
## Multiple R-squared:  0.9608, Adjusted R-squared:  0.9524
## F-statistic: 114.4 on 3 and 14 DF,  p-value: 4.398e-10
```

CC-CM clearly lower, D-CM not. **Use these results in paper.**

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
lm(log10(oav) ~ treat + exper)
Normal Q−Q
Standardized residuals
2
1
0
−1
−2
5
14
7
−2
−1
0
1
2
Theoretical Quantiles
lm(log10(oav) ~ treat + exper)



These look good.