

Analysis

An analysis using LiDAR data to detect moose browsing effects, with ground truthing

```
library(readr)
library(ggplot2)
library(glmmTMB)
```

Get compiled dataset (see compile.R)

```
dat <- read_csv("../data/compiledDataset.csv")
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   locality_and_treatment = col_character(),
##   LocalityCode = col_character(),
##   LocalityName = col_character(),
##   Treatment = col_character(),
##   resolution_m = col_character(),
##   region = col_character()
## )
```

```
## See spec(...) for full column specifications.
```

```
head(dat)
```

```
## # A tibble: 6 x 27
##   locality_and_tr~ LocalityCode LocalityName Treatment Longitude Latitude
##   <chr>            <chr>         <chr>         <chr>         <dbl>    <dbl>
## 1 bratsberg_b     BRB          Bratsberg     B             10.5     63.4
## 2 bratsberg_ub    BRUB         Bratsberg     UB            10.5     63.4
## 3 didrik_holmsen_b DHB          Didrik Holm~ B             11.4     59.9
## 4 didrik_holmsen_~ DHUB         Didrik Holm~ UB            11.4     59.9
## 5 drangedal1_b    1DRB         Drangedal1    B             9.15     59.1
## 6 drangedal1_ub   1DRUB        Drangedal1    UB            9.15     59.1
## # ... with 21 more variables: Clear.cut <dbl>, Year.initiated <dbl>,
## #   LiDAR.data.from.year <dbl>, plot_density_m2 <dbl>, resolution_m <chr>,
## #   region <chr>, Moose2015 <dbl>, Reddeer2015 <dbl>, Roedeer2015 <dbl>,
## #   YrsSinceExclosure <dbl>, field_mean <dbl>, field_median <dbl>, mn <dbl>,
## #   md <dbl>, sd <dbl>, min <dbl>, max <dbl>, first_qu.25. <dbl>,
## #   third_qu.75. <dbl>, mad <dbl>, prod <dbl>
```

A quick data check

```
table(dat$Treatment, dat$Clear.cut)
```

```
##
##      2000 2002 2003 2004 2005 2006 2007 2008 2009
## B      1   4   3   7   8   4  10   7   1
## UB     1   4   3   7   8   4  10   7   1
```

```
table(dat$Year.initiated, dat$LiDAR.data.from.year)
```

```
##
##      2010 2011 2013 2015 2016 2017 2018 2019
## 2007     0   0   0   0   0   0   0   6
## 2008     4   4   0  18   2   2   0   0
## 2009     0   0   0   0   4  24   0   0
## 2010     0   0   0   0   4   4   4   0
## 2011     0   0   2   0   4   0   0   8
```

```
table(dat$plot_density_m2, dat$resolution_m)
```

```
##
##      0,25 0,5
## 2      2  54
## 5     32   2
```

Something odd there...

```
table(dat$region, dat$Treatment)
```

```
##
##           B UB
## Hedmark   16 16
## Telemark  14 14
## Trondelag 15 15
```

```
table(dat$LocalityName, dat$Treatment)
```

```
##
##           B UB
## Bratsberg      1  1
## Didrik Holmsen  1  1
## Drangedal1     1  1
## Drangedal3     1  1
## Drangedal4     1  1
## Eidskog        1  1
## Fet 3          1  1
## Fritsoe1       1  1
## Fritsoe2       1  1
```

```
## Furesdal 1 1
## Halvard Pramhus 1 1
## Hi_tydal 1 1
## Kongsvinger 1 1
## Kongsvinger 2 1 1
## Kviteseid1 1 1
## Kviteseid2 1 1
## Kviteseid3 1 1
## Maarud 1 1 1
## Maarud 2 1 1
## Maarud 3 1 1
## Malvik 1 1
## namdalseid_1kub 1 1
## Nes 1 1 1
## Nes 2 1 1
## Nome_Cappelen1 1 1
## Nome_Cappelen2 1 1
## Notodden3 1 1
## Notodden5 1 1
## Notodden6 1 1
## Nsb_Verdal 1 1
## Selbu_Flub 1 1
## Selbu_k1 1 1
## Selbu_S1 1 1
## Singsaas 1 1
## Sl_Tydal 1 1
## Soerum 1 1
## Stangeskovene Aurskog 1 1
## Stangeskovene Eidskog 1 1
## steinkjer_1BBb 1 1
## steinkjer_2BBb 1 1
## Stig Dahlen 1 1
## Sub_Namdalseid 1 1
## Truls Holm 1 1
## verdal_1vb 1 1
## verdal_2VB 1 1
```

Looks good.

Canopy growth per year

Lets first compute canopy growth per year since exclosure

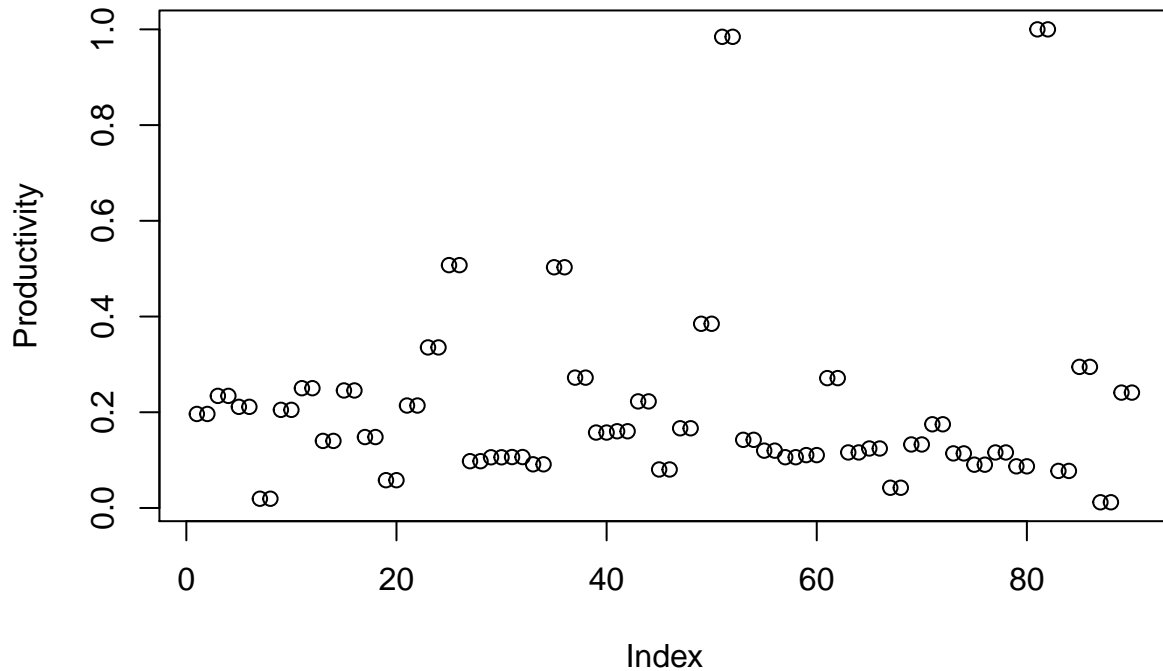
```
dat$canopygrowth <- dat$md/dat$YrsSinceExclosure
summary(dat$canopygrowth)
```

```
##      Min.   1st Qu.   Median     Mean  3rd Qu.    Max.
## 0.001214 0.022698 0.045375 0.096275 0.165922 0.420857
```

The numbers are in meters I'm pretty sure

One of the first things to decide on is what to do with productivity, as there are two outliers:

```
plot(dat$prod, ylab="Productivity")
```



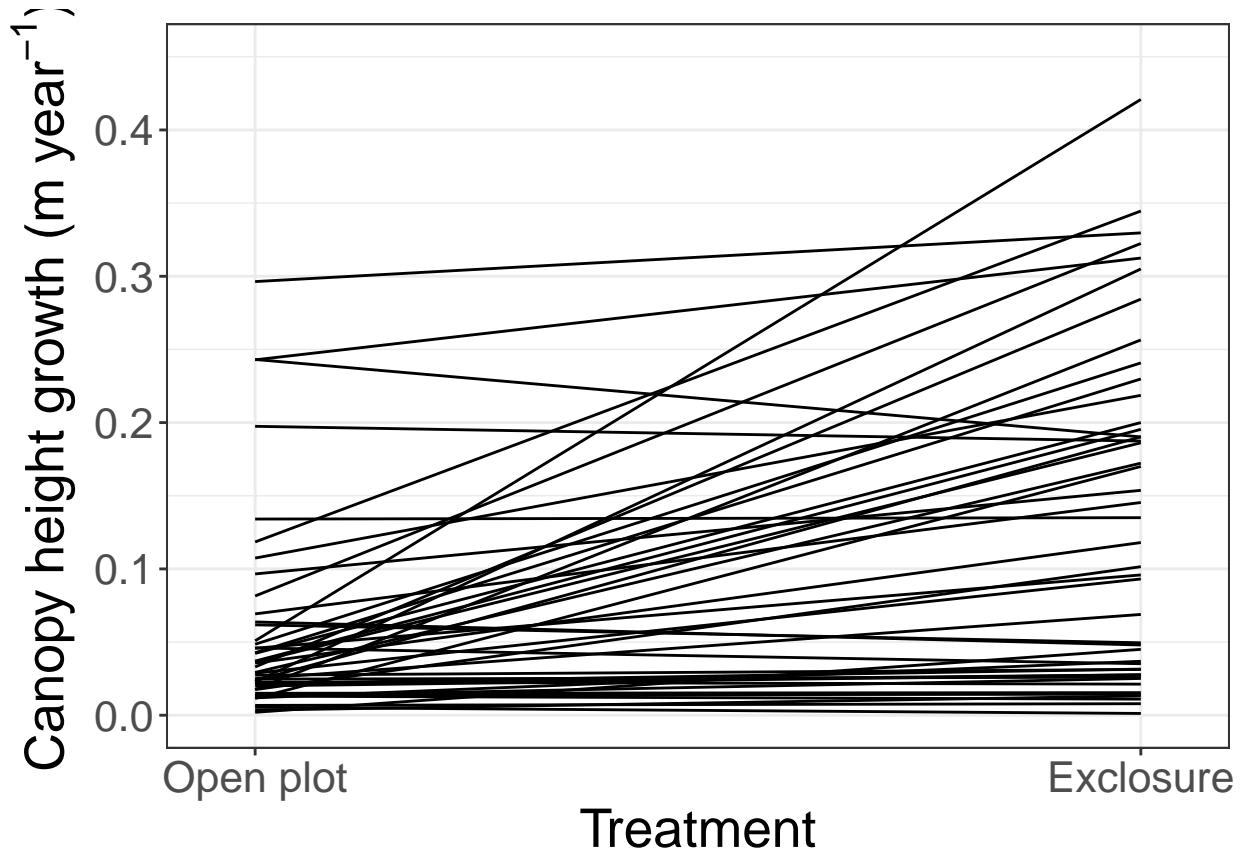
```
dat[dat$prod>0.6,c("LocalityName", "region", "prod")]
```

```
## # A tibble: 4 x 3
##   LocalityName region  prod
##   <chr>         <chr>  <dbl>
## 1 Nes 2        Hedmark 0.984
## 2 Nes 2        Hedmark 0.984
## 3 Stig Dahlen  Hedmark 1
## 4 Stig Dahlen  Hedmark 1
```

These are two sites in Hedmark that probably really are very productive, although there could have been serious sampling error due to chance. But they are legitimate, and we should be careful not to drop them too willingly. We can use the log of the productivity if residuals are acting strange.

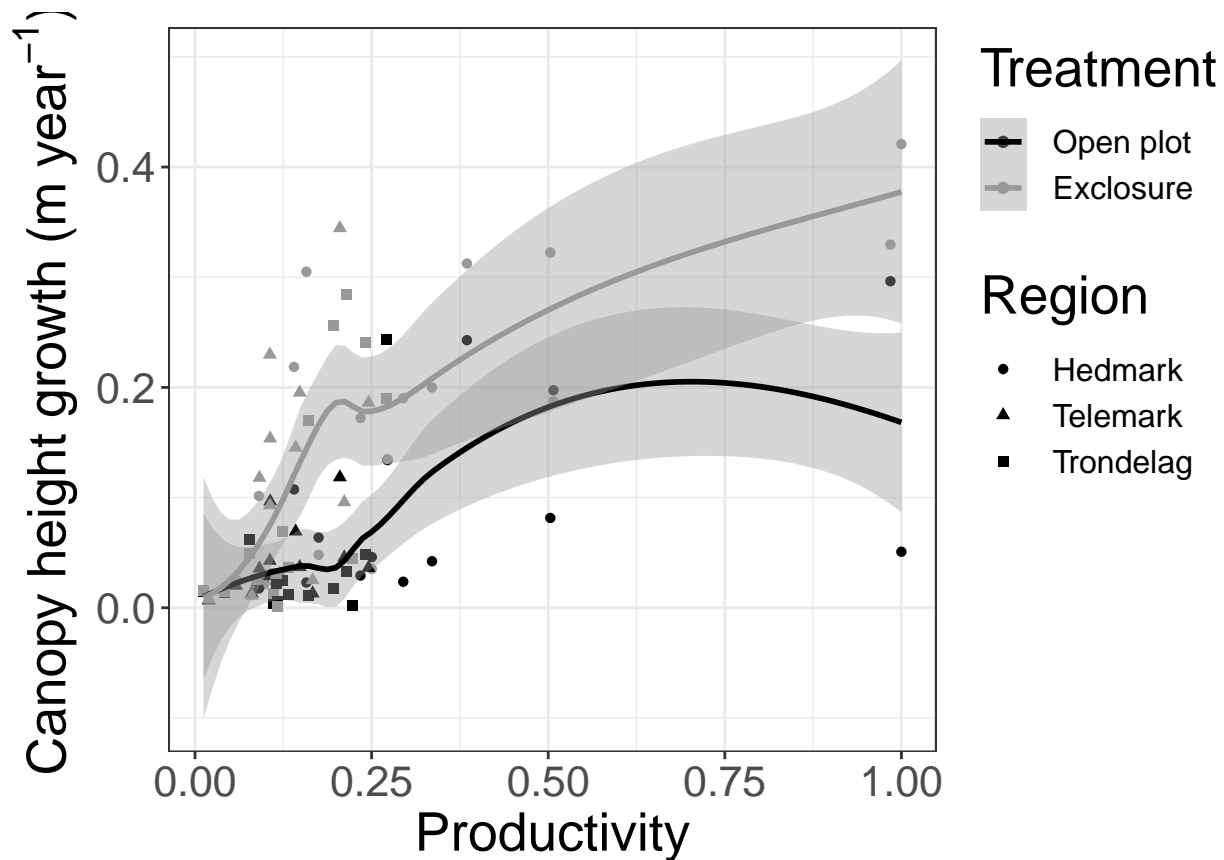
```
dat$Treatment <- as.factor(dat$Treatment)
levels(dat$Treatment) <- c('Open plot', 'Exclosure')
(chg_treat <- ggplot(dat, aes(x=Treatment, y=canopygrowth, group=LocalityName))+
  geom_line()+
  labs(y=expression(paste('Canopy height growth (m year'-1, '))), x='Treatment')+
  scale_linetype_manual(breaks = c("Exclosure", "Open plot"),
    labels = c("Open plots", "Exclosures"), values=c(1,2))+
  scale_x_discrete(limits = c('Open plot', 'Exclosure'),
```

```
breaks = c('Open plot', 'Exclosure'), expand = c(0.1,0))+
theme_bw()+
theme(text = element_text(size = 20))+
ylim(0, 0.45))
```



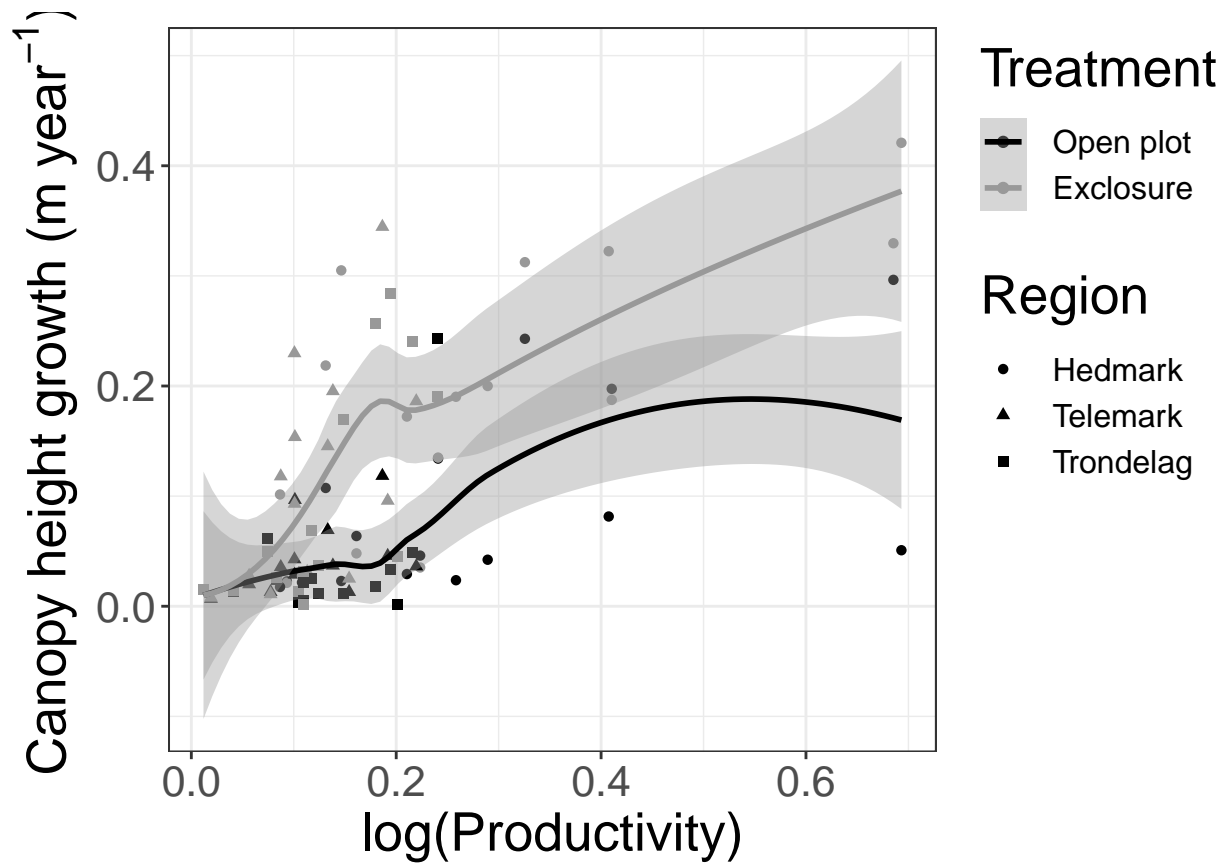
Lets use loess as it makes less assumptions about the shape of the relationship:

```
(chg_prod <- ggplot(data = dat,
                    aes(x = prod, y = canopygrowth))+
  geom_point(aes(colour= Treatment, shape=region))+
  geom_smooth(aes(colour= Treatment),
              method = "loess", formula = 'y ~ x')+
  labs(y=expression(paste('Canopy height growth (m year'-1', '))), x='Productivity')+
  theme_bw()+
  scale_color_manual(values = c("gray0", "gray60"))+
  labs(colour="Treatment", shape="Region")+
  theme(text = element_text(size = 20))+
  #ylim(0, 0.4)+
  theme(legend.position = 'right',
        legend.justification = c("left", "top"),
        legend.box.just = "left",
        #legend.margin = margin(5, 5, 5, 5),
        legend.text = element_text(size=12))
)
```



The highly productive sites are affecting this picture quite a bit I'd imagine. Let's try with a log-transformation:

```
dat$prod_1 <- log(dat$prod+1)
(chg_prod <- ggplot(data = dat,
                    aes(x = prod_1, y = canopygrowth))+
  geom_point(aes(colour= Treatment, shape=region))+
  geom_smooth(aes(colour= Treatment),
              method = "loess", formula = 'y ~ x')+
  labs(y=expression(paste('Canopy height growth (m year'-1', '))), x='log(Productivity)')+
  theme_bw()+
  scale_color_manual(values = c("gray0", "gray60"))+
  labs(colour="Treatment", shape="Region")+
  theme(text = element_text(size = 20))+
  #ylim(0, 0.4)+
  theme(legend.position = 'right',
        legend.justification = c("left", "top"),
        legend.box.just = "left",
        #legend.margin = margin(5, 5, 5, 5),
        legend.text = element_text(size=12))
)
```

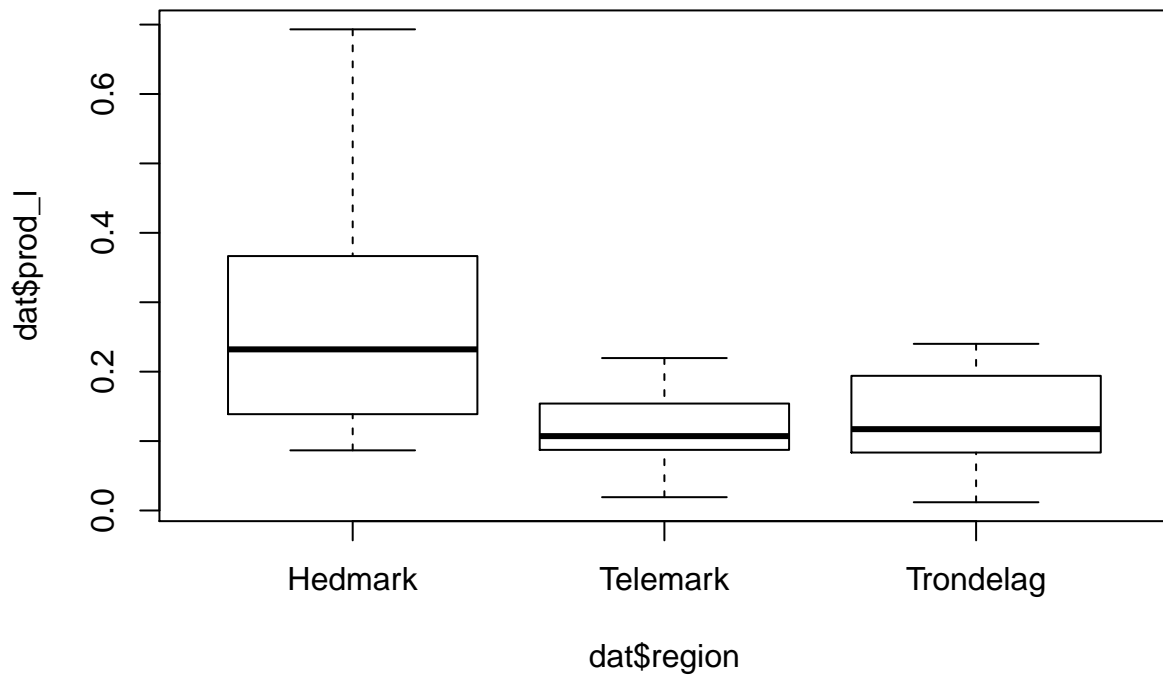


That's perhaps a little better. There's a relationship there it seems like, but perhaps not linear. Also note that all the really productive sites are in Hedmark:

```
table(round(dat$prod, 1), dat$region)
```

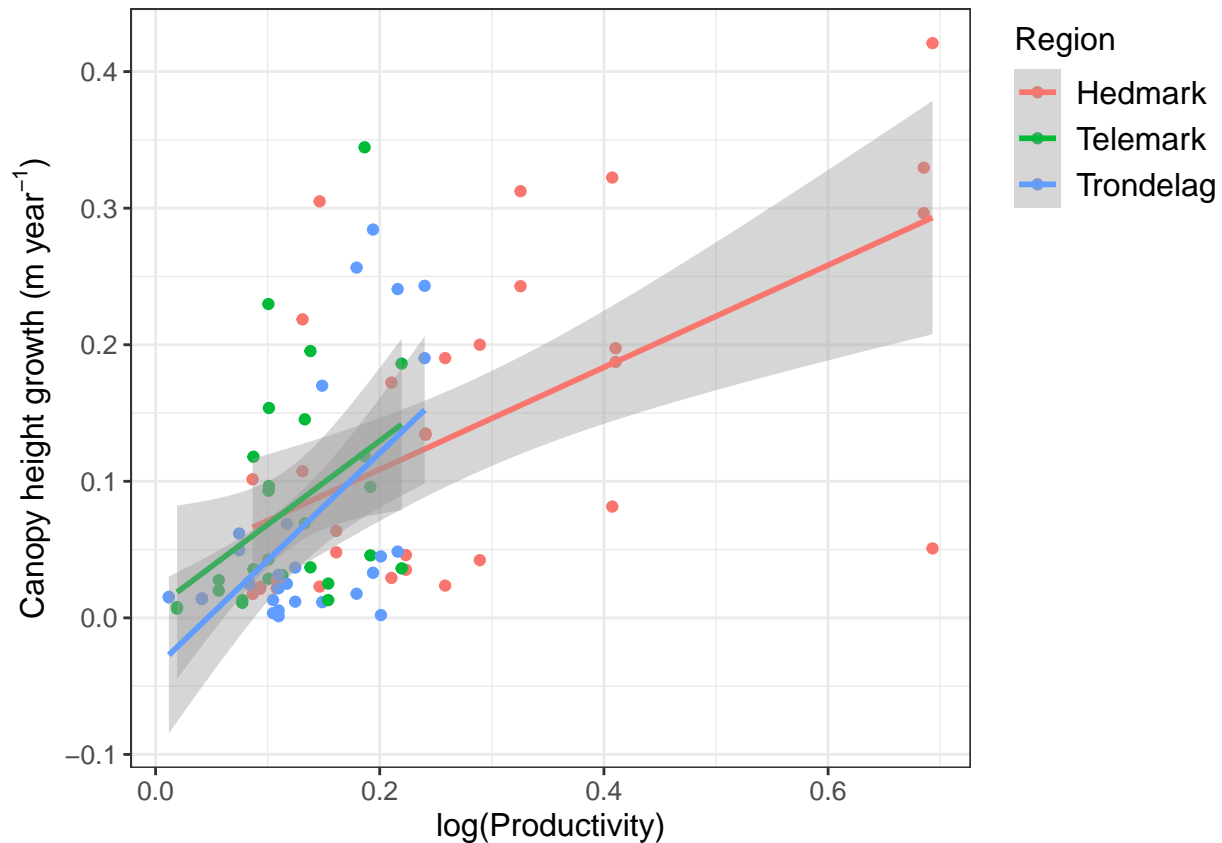
```
##
##      Hedmark Telemark Trondelag
## 0         0         2         4
## 0.1       8        18        14
## 0.2       6         8        10
## 0.3       8         0         2
## 0.4       2         0         0
## 0.5       4         0         0
## 1         4         0         0
```

```
boxplot(dat$prod_1~dat$region)
```



Lets investigate the effect of region a bit more.

```
ggplot(data = dat,
  aes(x = prod_l, y = canopygrowth, colour = region))+
  geom_point()+
  geom_smooth(method = "lm")+
  labs(y=expression(paste('Canopy height growth (m year-1, ')')), x='log(Productivity)')+
  theme_bw()+
  #scale_color_manual(values = c("gray0", "gray50", "grey100"))+
  labs(colour="Region")+
  theme(text = element_text(size = 12))+
  #ylim(0, 0.4)+
  theme(legend.position = 'right',
        legend.justification = c("left", "top"),
        legend.box.just = "left",
        #legend.margin = margin(5, 5, 5, 5),
        legend.text = element_text(size=12))
```

This plot tells us there is no justification for a random slope model. Also, the two productivity outliers appear less like outliers when only looking at the Hedmark data points.

Let's get some summary stats.

```
m <- tapply(dat$canopygrowth, dat$Treatment, FUN = mean)
# And standard error:
source('se.R')
s <- tapply(dat$canopygrowth, dat$Treatment, FUN = se)
(sumStats <- data.frame("mean canopy growth per year" = round(m, 3),
                        "standard error of the mean" = round(s, 3)))
```

```
##          mean.canopy.growth.per.year standard.error.of.the.mean
## Open plot                0.056                0.010
## Exclosure                0.136                0.017
```

We need a model to describe the interaction between treatment and productivity. The response is continuous and both positive and negative values can occur, so a gaussian family would be appropriate.

```
library(glmTMB) # using this package even though I dont need all its finesse...
dat$region <- as.factor(dat$region)
mod1 <- glmTMB(canopygrowth ~ Treatment * prod_1 + (1|region),
               data = dat, family = gaussian)
```

```
mod2 <- glmmTMB(canopygrowth ~ Treatment * prod_1 ,
  data = dat, family = gaussian)
```

```
AIC(mod1, mod2)
```

```
##      df      AIC
## mod1  6 -215.4912
## mod2  5 -217.4912
```

Model 2 is slightly better, suggesting we drop the random intercept. What does ANOVA say about this?

```
anova(mod1, mod2)
```

```
## Data: dat
## Models:
## mod2: canopygrowth ~ Treatment * prod_1, zi=~0, disp=~1
## mod1: canopygrowth ~ Treatment * prod_1 + (1 | region), zi=~0, disp=~1
##      Df      AIC      BIC logLik deviance Chisq Chi Df Pr(>Chisq)
## mod2  5 -217.49 -204.99 113.75 -227.49
## mod1  6 -215.49 -200.49 113.75 -227.49      0      1      1
```

The models explain exactly the same, so we can drop the random intercept. The only reason for keeping it is if we want to argue it's part of the design, which it is. However, I don't think that's more important than having a parsimonious model.

```
summary(mod2)
```

```
## Family: gaussian ( identity )
## Formula:          canopygrowth ~ Treatment * prod_1
## Data: dat
##
##      AIC      BIC    logLik deviance df.resid
##   -217.5   -205.0    113.7   -227.5      85
##
##
## Dispersion estimate for gaussian family (sigma^2): 0.00467
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    0.004464   0.016613   0.269  0.78815
## TreatmentExclosure 0.030452   0.023495   1.296  0.19493
## prod_1          0.287174   0.072826   3.943 8.04e-05 ***
## TreatmentExclosure:prod_1 0.275917   0.102992   2.679  0.00738 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

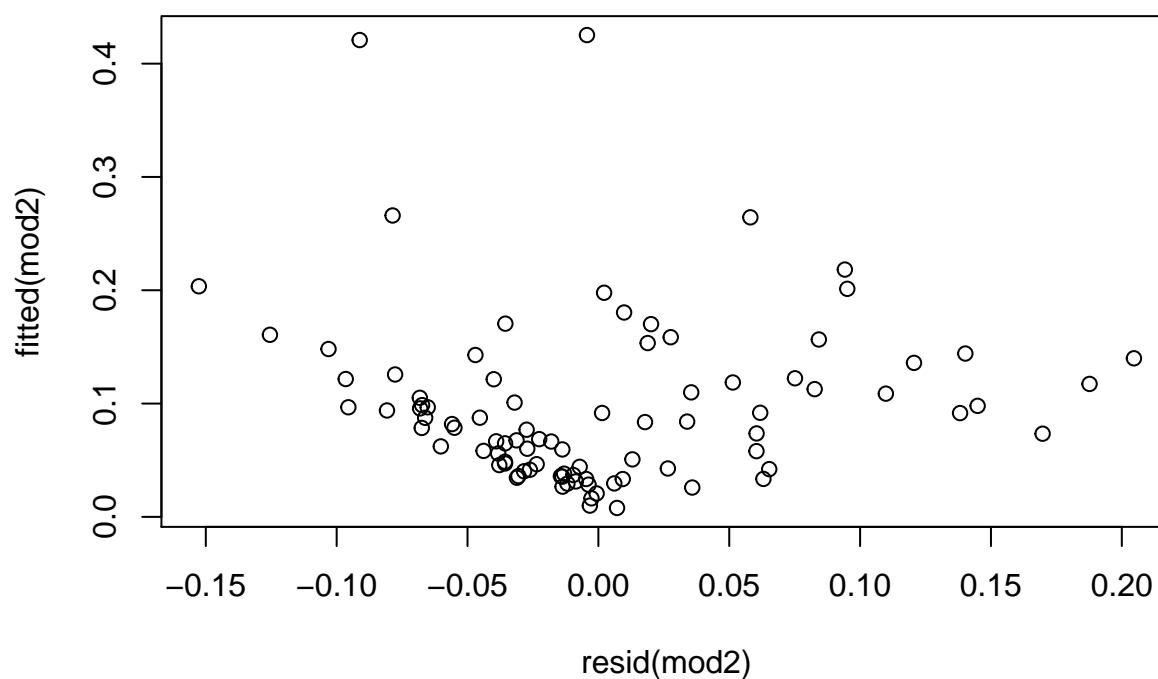
Significant interaction term, so we're keeping this full model. We can also see what the random effect would've give us:

```
ranef(mod1)
```

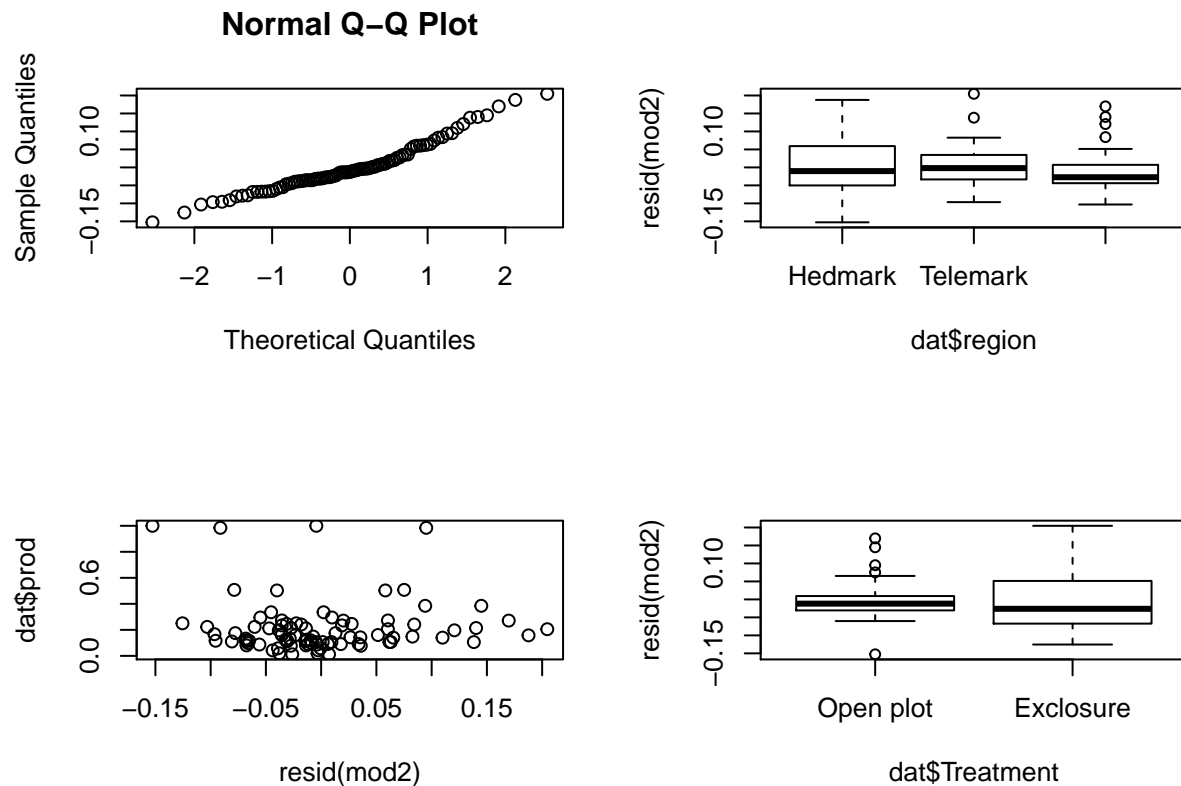
```
## $region  
##      (Intercept)  
## Hedmark  -1.809701e-12  
## Telemark   1.271089e-10  
## Trondelag -1.252993e-10
```

, and it's not very much.

```
plot(resid(mod2), fitted(mod2))
```

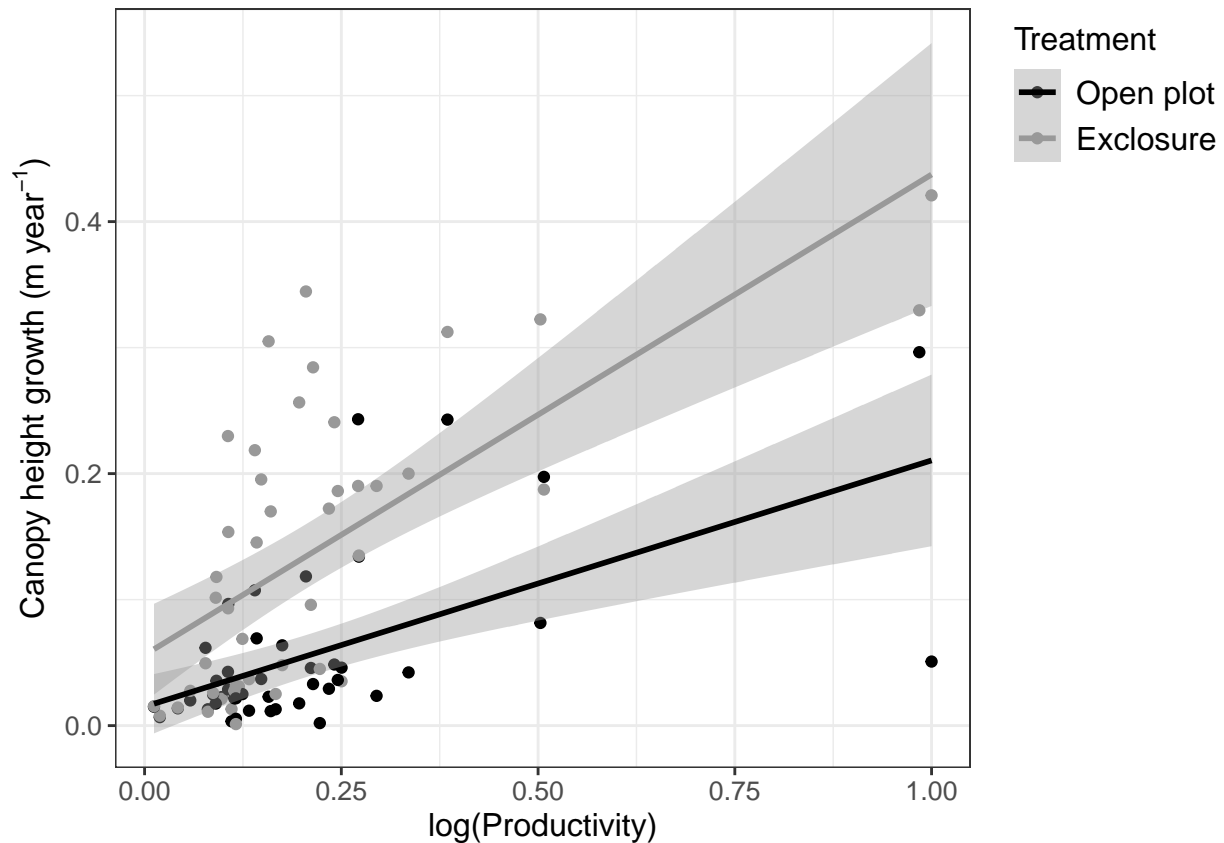


```
par(mfrow=c(2,2))  
qqnorm(resid(mod2))  
plot(resid(mod2) ~ dat$region)  
plot(resid(mod2), dat$prod)  
plot(resid(mod2) ~ dat$Treatment)
```



Looks like everything is in order. The residuals against productivity at least show a centering around zero.

```
(intPlot <- ggplot(data = dat,
  aes(x = prod, y = canopygrowth, colour= Treatment))+
  geom_point()+
  geom_smooth(method = "lm")+
  labs(y=expression(paste('Canopy height growth (m year-1, ')')), x='log(Productivity)')+
  theme_bw()+
  scale_color_manual(values = c("gray0", "gray60"))+
  labs(colour="Treatment")+
  theme(text = element_text(size = 12))+
  #ylim(0, 0.4)+
  theme(legend.position = 'right',
    legend.justification = c("left", "top"),
    legend.box.just = "left",
    #legend.margin = margin(5, 5, 5, 5),
    legend.text = element_text(size=12))
)
```



So here we're modelling a linear function of productivity. We could try adding a quadratic term.

```
dat$prod_12 <- dat$prod_1*dat$prod_1
mod3 <- glmmTMB(canopygrowth ~ Treatment * prod_1 + prod_12,
  data = dat, family = gaussian)
```

```
AIC(mod2, mod3)
```

```
##      df      AIC
## mod2  5 -217.4912
## mod3  6 -219.5023
```

This gave a slight improvement.

```
anova(mod2, mod3)
```

```
## Data: dat
## Models:
## mod2: canopygrowth ~ Treatment * prod_1, zi=~0, disp=~1
## mod3: canopygrowth ~ Treatment * prod_1 + prod_12, zi=~0, disp=~1
##      Df      AIC      BIC logLik deviance  Chisq Chi Df Pr(>Chisq)
## mod2  5 -217.49 -204.99 113.75  -227.49
## mod3  6 -219.50 -204.50 115.75  -231.50  4.0111      1    0.0452 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

But the numbers tell us we should keep it. Finally, we should do a sensitivity analysis by removing the outliers and comparing the results that those models give us.

```
dat2 <- dat[dat$prod < 0.7,]
dim(dat2)
```

```
## [1] 86 30
```

That deleted the two most productive sites.

```
mod_sens1 <- glmmTMB(canopygrowth ~ Treatment * prod + (1|region),
  data = dat2, family = gaussian)
mod_sens2 <- glmmTMB(canopygrowth ~ Treatment * prod,
  data = dat2, family = gaussian)
AIC(mod_sens1, mod_sens2)
```

```
##           df      AIC
## mod_sens1  6 -211.5329
## mod_sens2  5 -213.5329
```

```
mod_sens3 <- glmmTMB(canopygrowth ~ Treatment * prod + I(prod^2),
  data = dat2, family = gaussian)
AIC(mod_sens2, mod_sens3)
```

```
##           df      AIC
## mod_sens2  5 -213.5329
## mod_sens3  6 -212.8157
```

```
summary(mod_sens2)
```

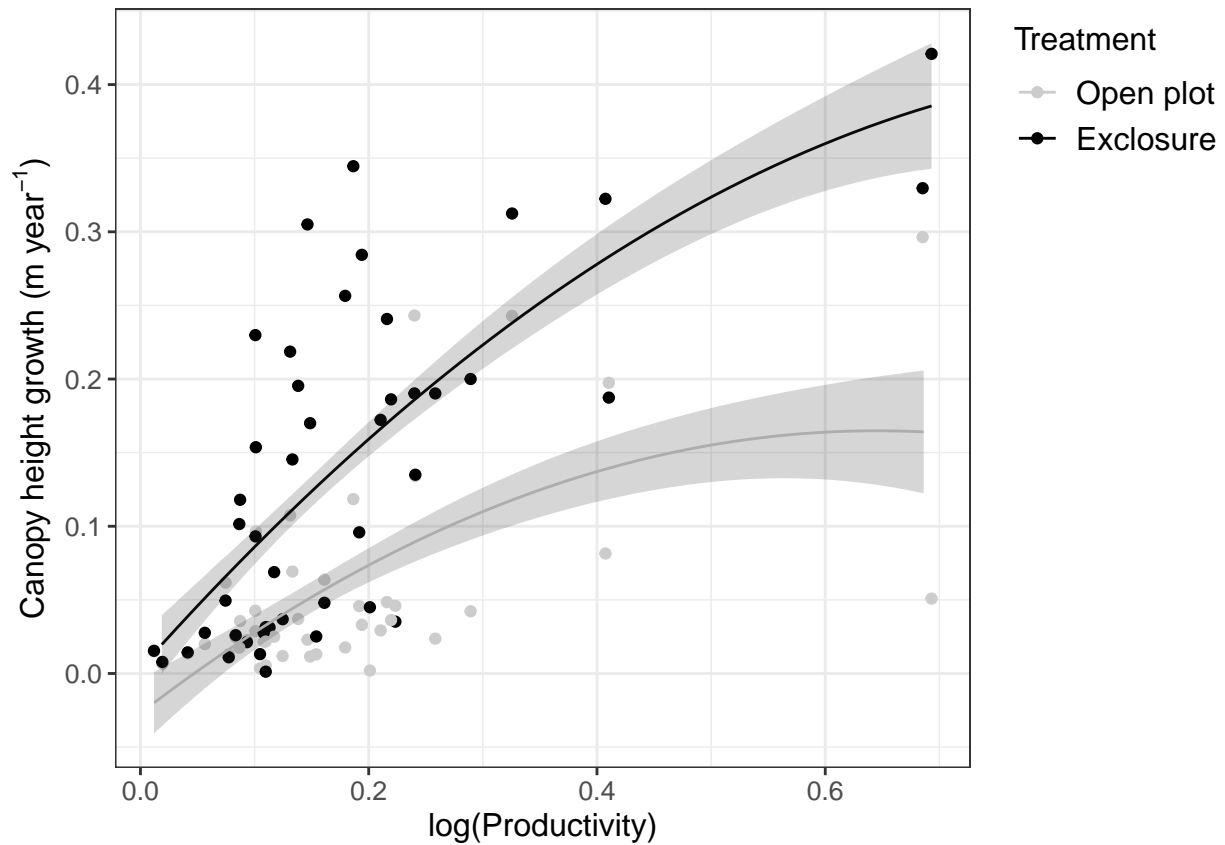
```
## Family: gaussian ( identity )
## Formula:          canopygrowth ~ Treatment * prod
## Data: dat2
##
##           AIC      BIC   logLik deviance df.resid
##    -213.5    -201.3    111.8   -223.5         81
##
##
## Dispersion estimate for gaussian family (sigma^2): 0.00435
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.003533   0.018914  -0.187  0.851835
## TreatmentExclosure  0.029838   0.026749   1.115  0.264644
## prod           0.311650   0.091980   3.388  0.000703 ***
## TreatmentExclosure:prod 0.256527   0.130079   1.972  0.048601 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The results are qualitatively the same.

Plot

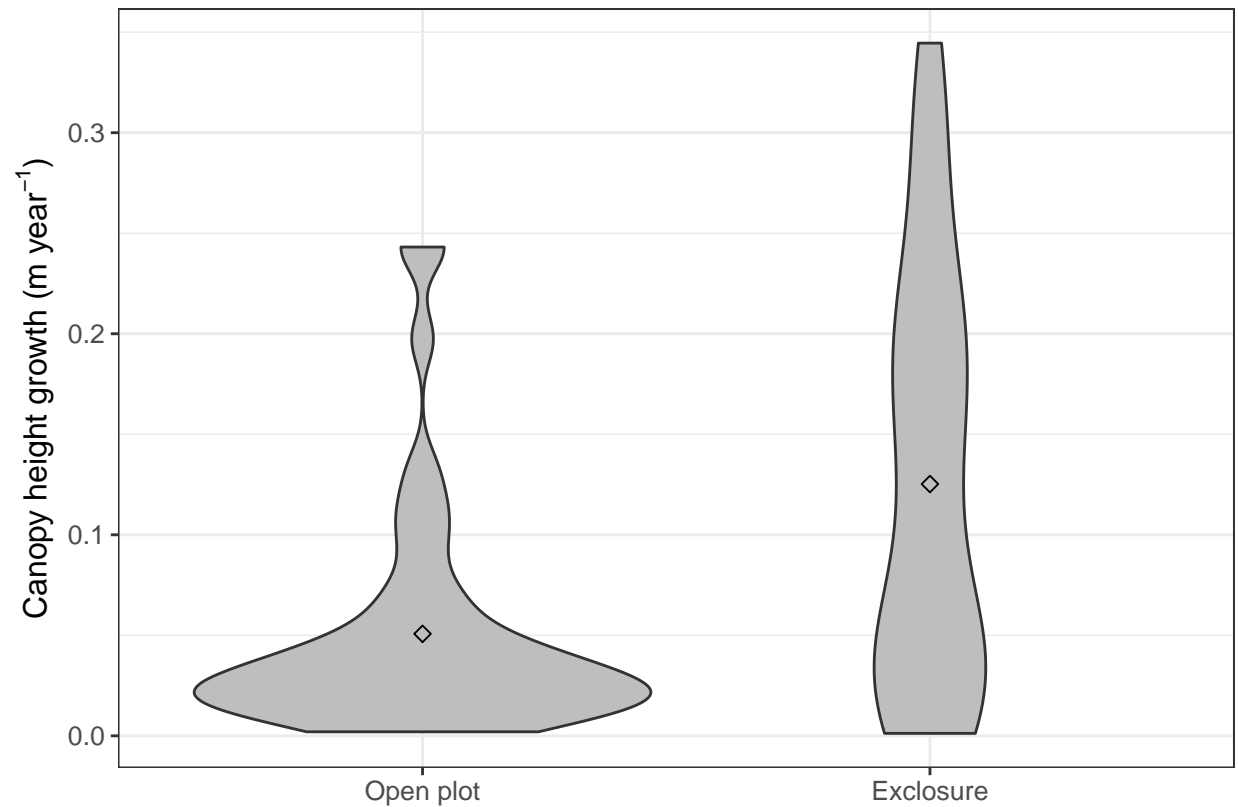
```
tr <- rep(c("Open plot", "Exclosure"), times=50, each=1)
prod_1 <- seq(from = min(dat$prod_1), to=max(dat$prod_1), length.out = 100)
prod_12 <- prod_1^2
datX <- data.frame(Treatment = rep(c("Open plot", "Exclosure"), times=50, each=1),
                  prod_1 = seq(from = min(dat$prod_1), to=max(dat$prod_1), length.out = 100))
datX$prod_12 <- datX$prod_1*datX$prod_1
pred <- predict(mod3, list(Treatment=tr, prod_1=prod_1, prod_12, prod_12), se.fit = TRUE)
pred2 <- data.frame(Treatment = tr,
                  prod_1 = prod_1,
                  pred = pred$fit,
                  se = pred$se.fit)

(predPlot <- ggplot()+
  geom_point(data = dat, aes(x = prod_1, y = canopygrowth, colour= Treatment))+
  geom_line(data=pred2, aes(x = prod_1, y=pred, colour = Treatment))+
  geom_ribbon(data=pred2, aes(x = prod_1,
                            ymin=pred-se,
                            ymax=pred+se,
                            group = Treatment),
            alpha=0.2,
            linetype="blank")+
  labs(y=expression(paste('Canopy height growth (m year-1, ')'), x='log(Productivity)')+
  theme_bw()+
  scale_color_manual(values = c("gray80", "black"))+
  labs(colour="Treatment")+
  theme(text = element_text(size = 12))+
  theme(legend.position = 'right',
        legend.justification = c("left", "top"),
        legend.box.just = "left",
        #legend.margin = margin(5, 5, 5, 5),
        legend.text = element_text(size=12))
)
```



We can also make these nice violin plots, even though the main plot should show the interaction effect like the plot above does.

```
(viol <- ggplot(data = dat2, aes(x = Treatment, y = canopygrowth)) +
  geom_violin(fill = "grey") +
  theme_bw() +
  theme(text = element_text(size = 12)) +
  labs(y = expression(paste('Canopy height growth (m year-1, ')'), x = '')) +
  stat_summary(fun.y = mean, geom = "point", shape = 23, size = 2))
```

MAD

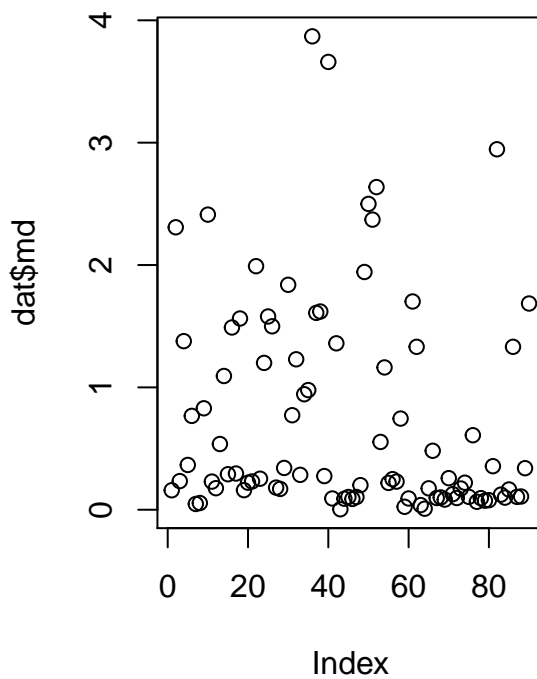
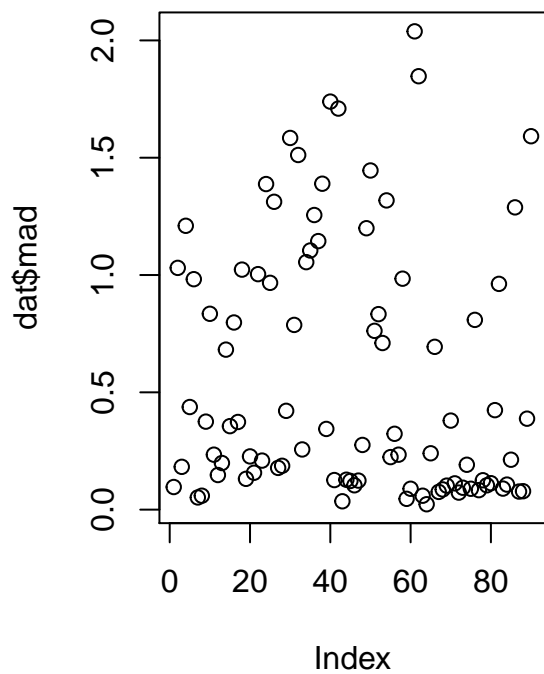
I think and foremost it's the relative MAD we should focus on. That's the equivalent of CV (Coefficient of Variation).

```
dat$rMAD <- dat$mad/dat$md
summary(dat$rMAD)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.3160  0.7399  1.0207  1.0869  1.2554  8.8956
```

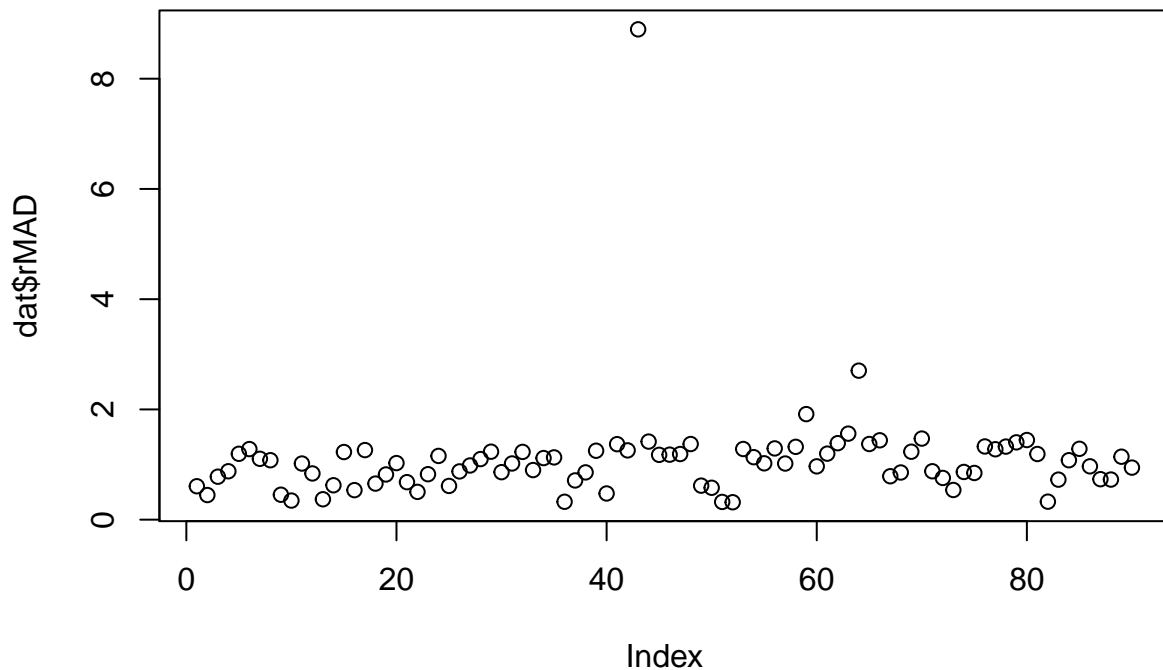
Looks sqewed.

```
par(mfrow=c(1,2))
plot(dat$mad)
plot(dat$md)
```



The two factors are well-behaved at first glance.

```
plot(dat$rMAD)
```



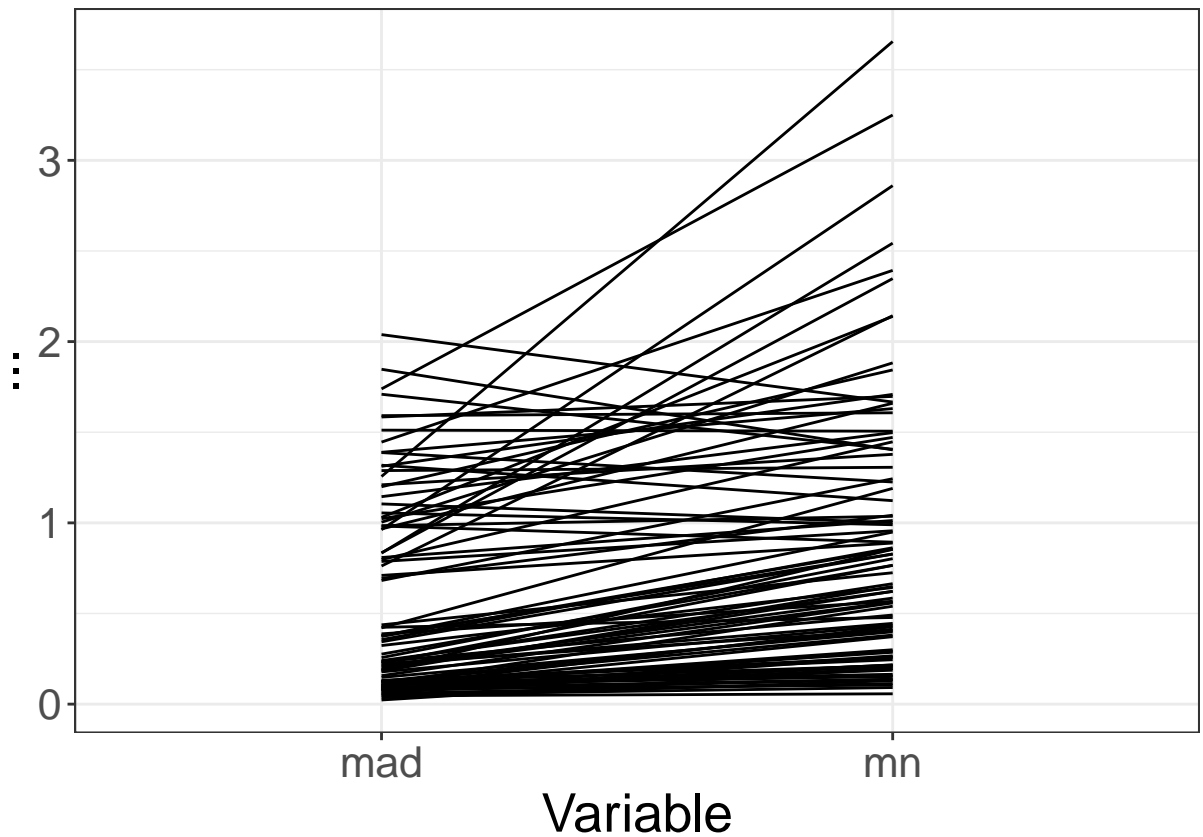
But this one is not.

```
dat[dat$rMAD>4, c("LocalityName", "Treatment", "md", "mn", "sd", "mad", "rMAD")]
```

```
## # A tibble: 1 x 7
##   LocalityName Treatment      md      mn      sd      mad rMAD
##   <chr>          <fct>    <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 namdalseid_1kub Open plot 0.00400 0.129 0.688 0.0356 8.90
```

It's the median that is causing trouble. It's very low, 0.004m.

```
library(reshape2)
myMelt <- melt(data = dat, id.vars = "LocalityCode", measure.vars = c("mad", "mn"))
ggplot(myMelt, aes(x=variable, y=value, group=LocalityCode))+
  geom_line()+
  labs(y='...', x='Variable')+
  theme_bw()+
  theme(text = element_text(size = 20))
```



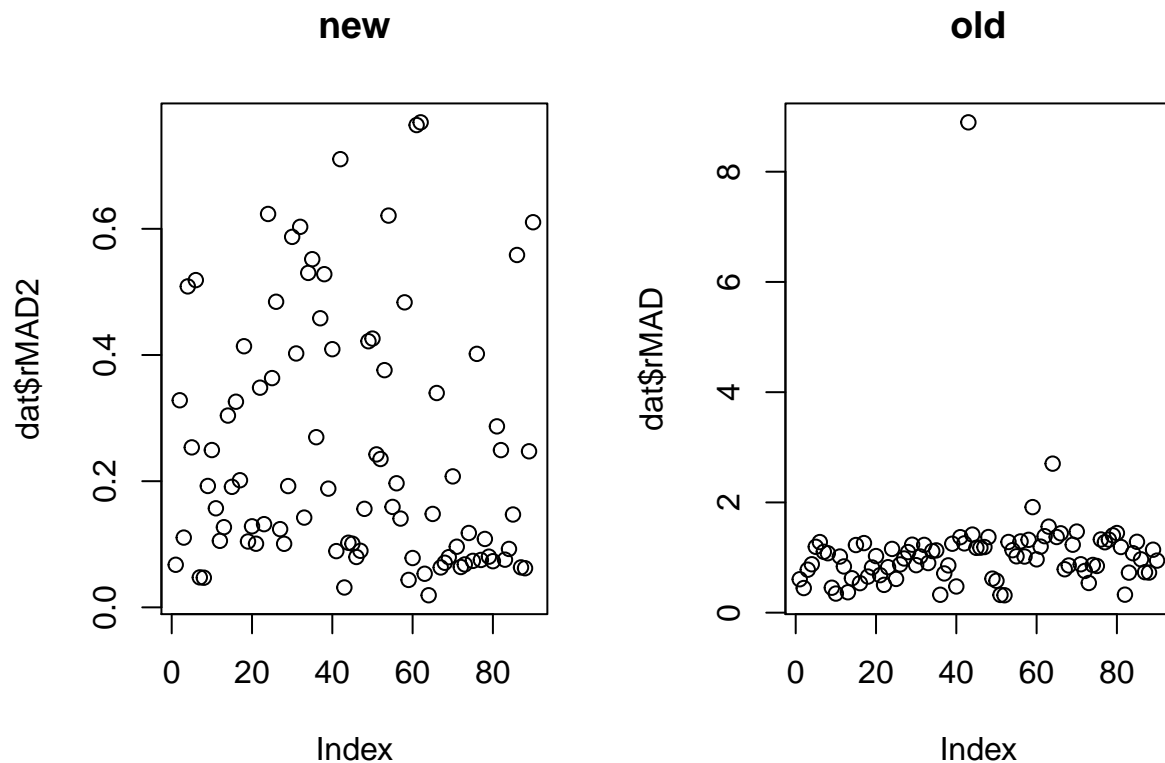
This doesn't look strange. It is simply the low median value that increases rMAD asymptotically. It's of course the lowest median in the data set, as you can see here:

```
min(dat$md)
```

```
## [1] 0.004
```

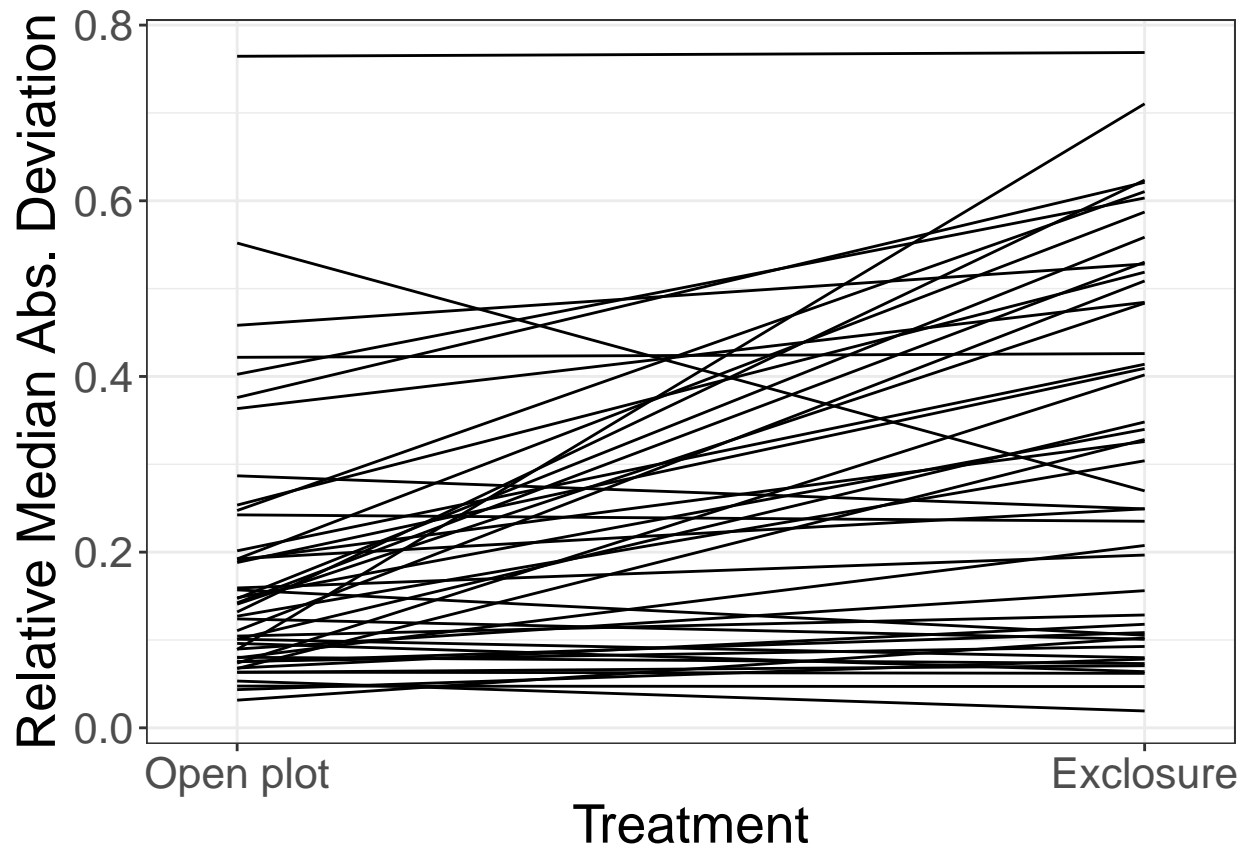
Let's remake rMAD with a slight moderation:

```
dat$rMAD2 <- dat$mad/(dat$mn+1)
par(mfrow=c(1,2))
plot(dat$rMAD2, main = "new")
plot(dat$rMAD, main = "old")
```

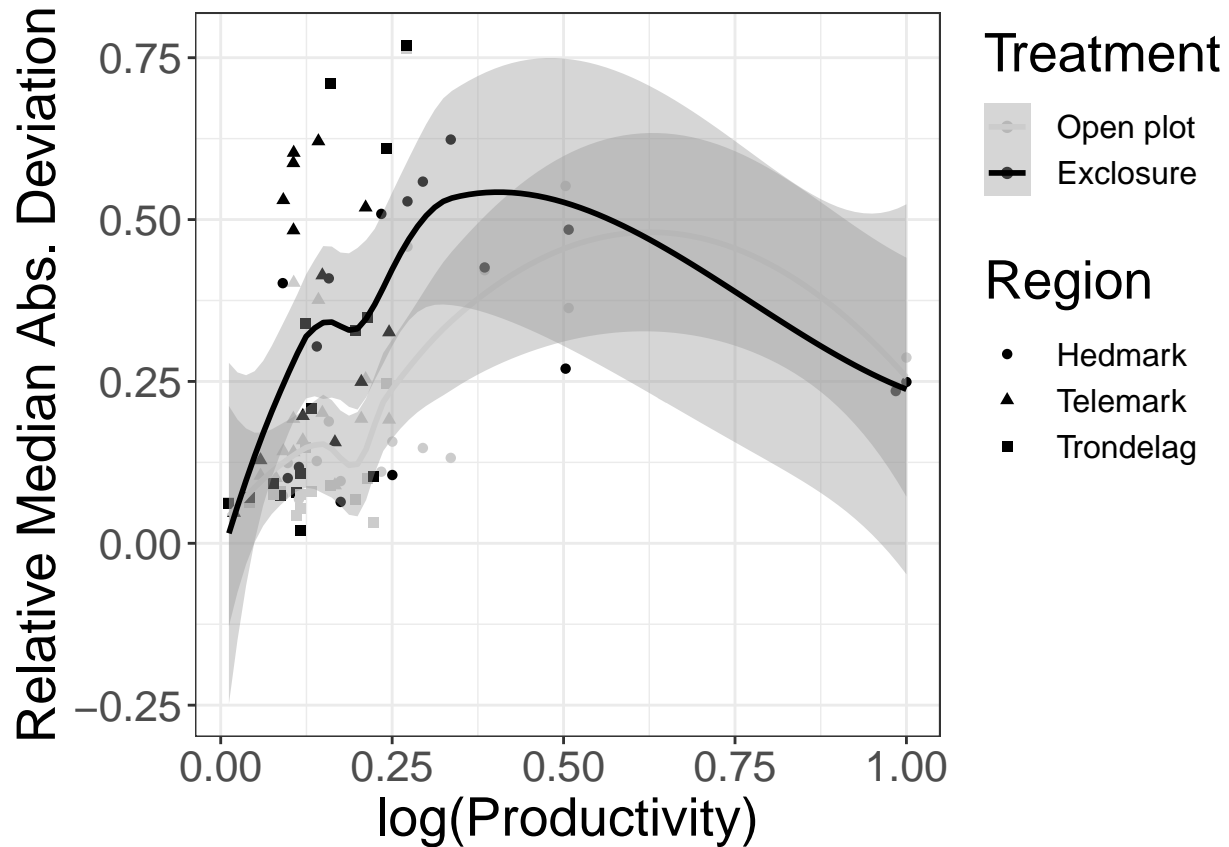


I don't see how this can affect the interpretation of rMAD. Let's use it.

```
ggplot(dat, aes(x=Treatment, y=rMAD2, group=LocalityName))+
  geom_line()+
  labs(y='Relative Median Abs. Deviation', x='Treatment')+
  scale_linetype_manual(breaks = c("Exclosure", "Open plot"),
                        labels = c("Open plots", "Exclosures"), values=c(1,2))+
  scale_x_discrete(limits = c('Open plot', 'Exclosure'),
                   breaks = c('Open plot', 'Exclosure'), expand = c(0.1,0))+
  theme_bw()+
  theme(text = element_text(size = 20))
```

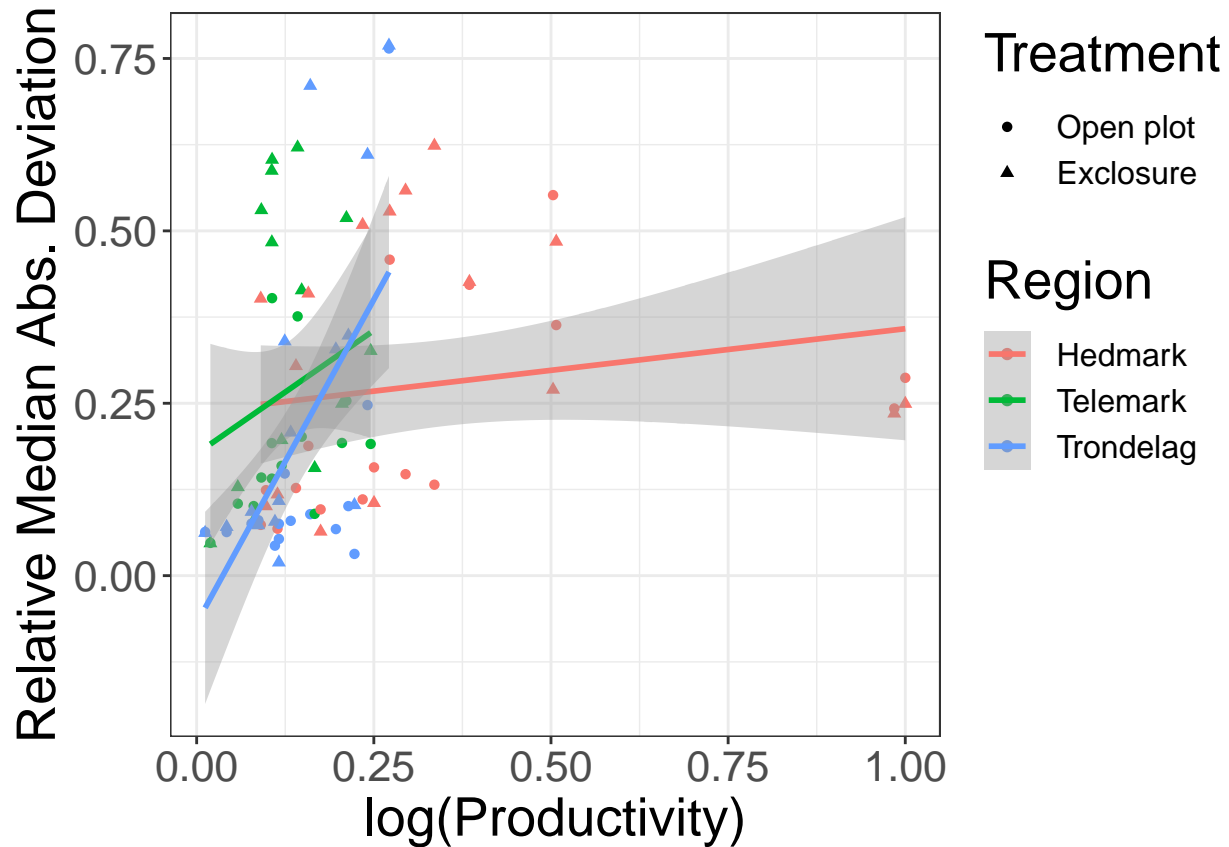


```
ggplot(data = dat,
       aes(x = prod, y = rMAD2))+
  geom_point(aes(colour= Treatment, shape=region))+
  geom_smooth(aes(colour= Treatment),
             method = "loess", formula = 'y ~ x')+
  labs(y='Relative Median Abs. Deviation', x='log(Productivity)')+
  theme_bw()+
  scale_color_manual(values = c("gray80", "black"))+
  labs(colour="Treatment", shape="Region")+
  theme(text = element_text(size = 20))+
  #ylim(0, 0.4)+
  theme(legend.position = 'right',
        legend.justification = c("left", "top"),
        legend.box.just = "left",
        #legend.margin = margin(5, 5, 5, 5),
        legend.text = element_text(size=12))
```



Probably no interaction with productivity, but the 'exclosure line' appear to lie higher. Let's see if there is an effect of region

```
ggplot(data = dat,
       aes(x = prod, y = rMAD2))+
  geom_point(aes(colour=region, shape=Treatment))+
  geom_smooth(aes(colour= region),
             method = "lm", formula = 'y ~ x')+
  labs(y='Relative Median Abs. Deviation', x='log(Productivity)')+
  theme_bw()+
  #scale_color_manual(values = c("gray80", "black", "grey20"))+
  labs(colour="Region", shape="Treatment")+
  theme(text = element_text(size = 20))+
  #ylim(0, 0.4)+
  theme(legend.position = 'right',
        legend.justification = c("left", "top"),
        legend.box.just = "left",
        #legend.margin = margin(5, 5, 5, 5),
        legend.text = element_text(size=12))
```



Looks like we need a random slope for this one.

```
modmad <- glmmTMB(data = dat,
  rMAD2~Treatment*prod_l+prod_l2+(prod_l|region),
  family = gaussian)
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; singular convergence
## (7). See vignette('troubleshooting')
```

```
summary(modmad)
```

```
## Family: gaussian ( identity )
## Formula:      rMAD2 ~ Treatment * prod_l + prod_l2 + (prod_l | region)
## Data: dat
##
##      AIC      BIC    logLik deviance df.resid
##    -60.3    -37.8     39.1    -78.3       81
##
## Random effects:
##
## Conditional model:
##   Groups   Name      Variance Std.Dev. Corr
##   region  (Intercept) 0.003101 0.05569
##           prod_l      0.026468 0.16269  -1.00
##   Residual              0.023795 0.15426
```



```
## Number of obs: 90, groups: region, 3
##
## Dispersion estimate for gaussian family (sigma^2): 0.0238
##
## Conditional model:
##
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.07869   0.06385  -1.232  0.21779
## TreatmentExclosure    0.16853   0.05301   3.179  0.00148 **
## prod_l         2.14520   0.42564   5.040 4.66e-07 ***
## prod_l2        -2.51124   0.58296  -4.308 1.65e-05 ***
## TreatmentExclosure:prod_l -0.18024   0.23236  -0.776  0.43792
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Note perfect correlation between random effect. Continue from here...

Compare LiDAR and field data

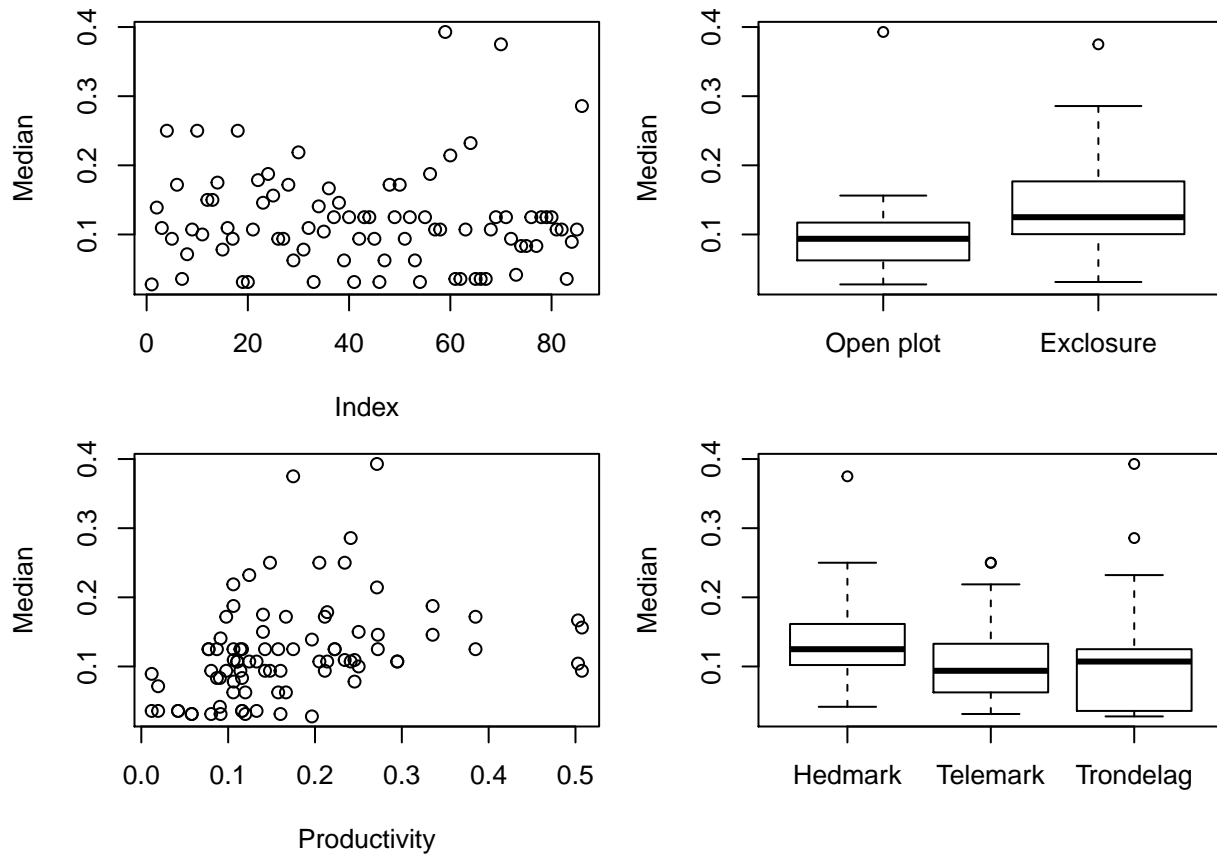
First, let's calculate canopy growth as we did for the LiDAR data

```
dat2$canopygrowth_f <- dat2$field_median/dat2$YrsSinceExclosure
```

```
summary(dat2$canopygrowth_f)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.02778 0.07943 0.10714 0.11947 0.14583 0.39286
```

```
par(mfrow=c(2,2), mar=c(4,4,1,1))
plot(dat2$canopygrowth_f, ylab="Median")
plot(dat2$canopygrowth_f~dat2$Treatment, xlab="", ylab="Median")
plot(dat2$canopygrowth_f~dat2$prod, xlab="Productivity", ylab="Median")
plot(dat2$canopygrowth_f~dat2$region, xlab="", ylab="Median")
```



Appears the annual median tree growth of a site (median of the median of a circle) is around 0.1, or precisely:

```
median(dat2$canopygrowth_f)
```

```
## [1] 0.1071429
```

Also, it is greater in exclosures and increases with productivity, and is highest in Hedmark. Very similar to the LiDAR data.

```
m2 <- tapply(dat2$canopygrowth_f, dat2$Treatment, FUN = mean)
s2 <- tapply(dat2$canopygrowth_f, dat2$Treatment, FUN = se)
(sumStats <- data.frame("mean canopy growth per year" = round(m2, 3),
                        "standard error of the mean" = round(s2, 3)))
```

```
##          mean.canopy.growth.per.year standard.error.of.the.mean
## Open plot                0.094                0.009
## Exclosure                0.145                0.011
```

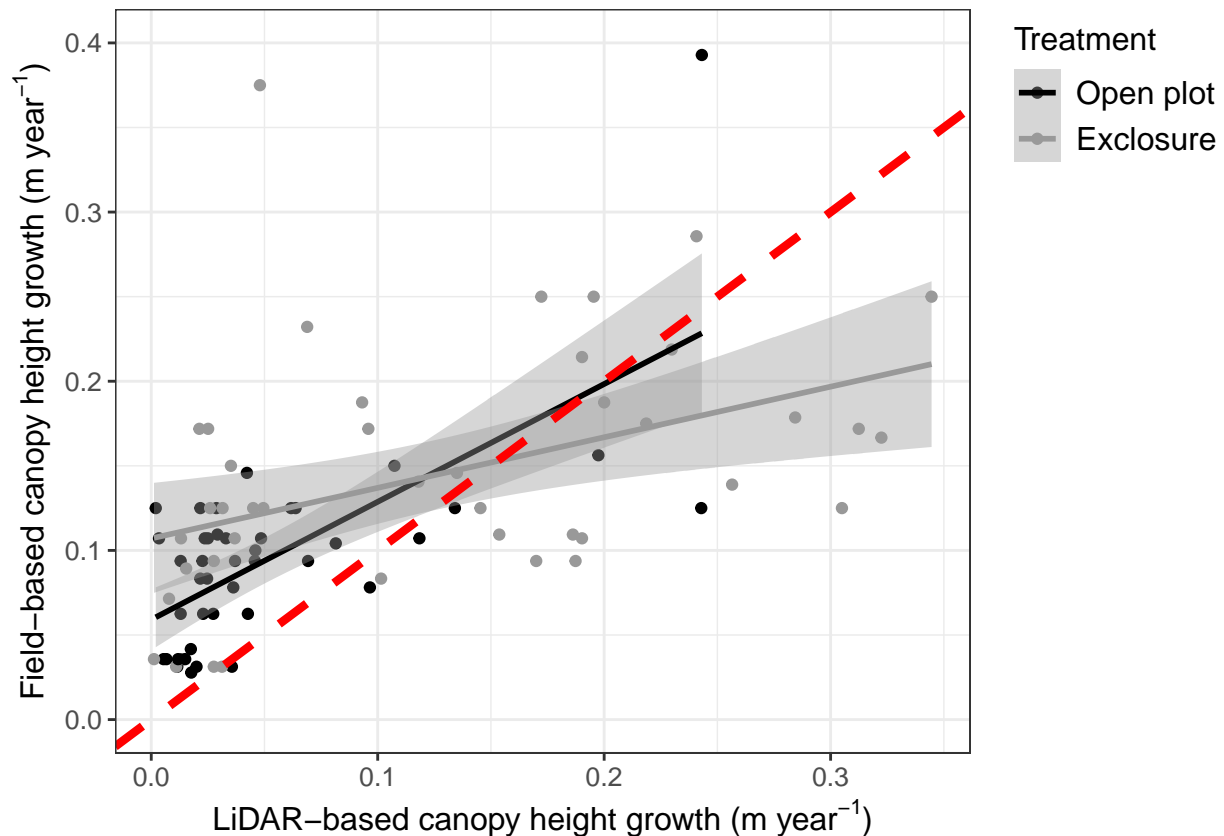
We don't want to analyse the field data alone, or by itself. That we have done in other papers and with more data. Here we are only interested in 'ground-truthing' the LiDAR data. So..

```
(corPlot <- ggplot(data = dat2,
                   aes(x = canopygrowth, y = canopygrowth_f, colour= Treatment))+
  geom_point()+
```

```

    geom_smooth(method = "lm")+
    labs(y=expression(paste('Field-based canopy height growth (m year-1', '))), x=expression(paste('L
    theme_bw()+
    scale_color_manual(values = c("gray0", "gray60"))+
    labs(colour="Treatment")+
    theme(text = element_text(size = 12))+
    ylim(0, 0.4)+
    theme(legend.position = 'right',
          legend.justification = c("left", "top"),
          legend.box.just = "left",
          #legend.margin = margin(5, 5, 5, 5),
          legend.text = element_text(size=12))+
    geom_abline(intercept = 0, slope = 1, color="red",
                linetype="dashed", size=1.5)
)

```



Perhaps the correlation strength differs between treatment, but not very much.

```

ggplot(data = dat2,
        aes(x = canopygrowth, y = canopygrowth_f, colour= region))+
    geom_point()+
    geom_smooth(method = "lm")+
    labs(y=expression(paste('Field-based canopy height growth (m year-1', '))), x=expression(paste('L
    theme_bw()+

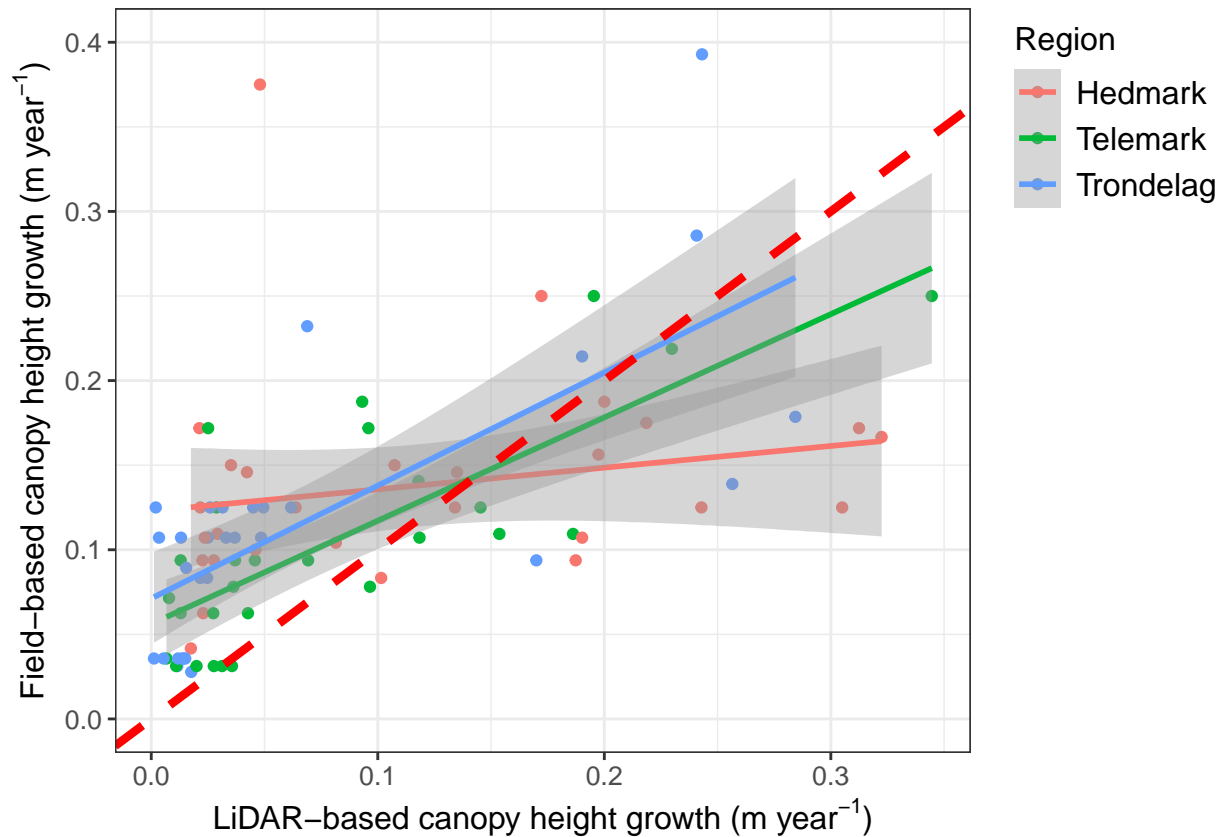
    labs(colour="Region")+

```

```

theme(text = element_text(size = 12))+
ylim(0, 0.4)+
theme(legend.position = 'right',
      legend.justification = c("left", "top"),
      legend.box.just = "left",
      #legend.margin = margin(5, 5, 5, 5),
      legend.text = element_text(size=12))+
geom_abline(intercept = 0, slope = 1, color="red",
            linetype="dashed", size=1.5)

```



Also it differs a bit between region.

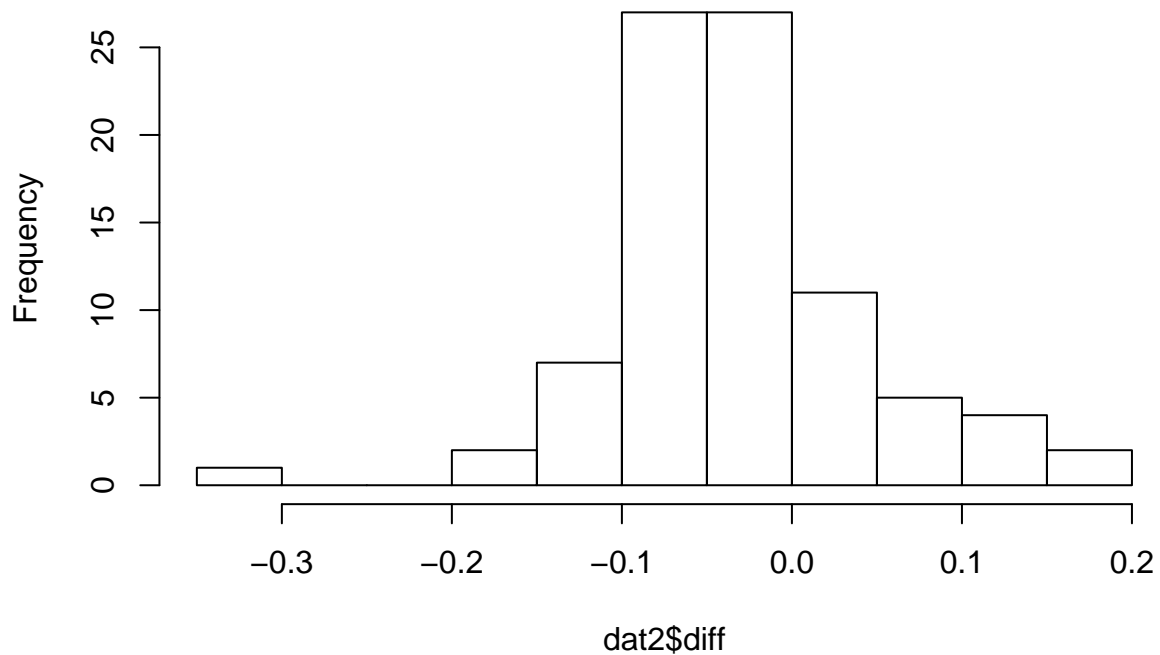
Also, we can look at the effect that experimental duration has had on the correlation:

```

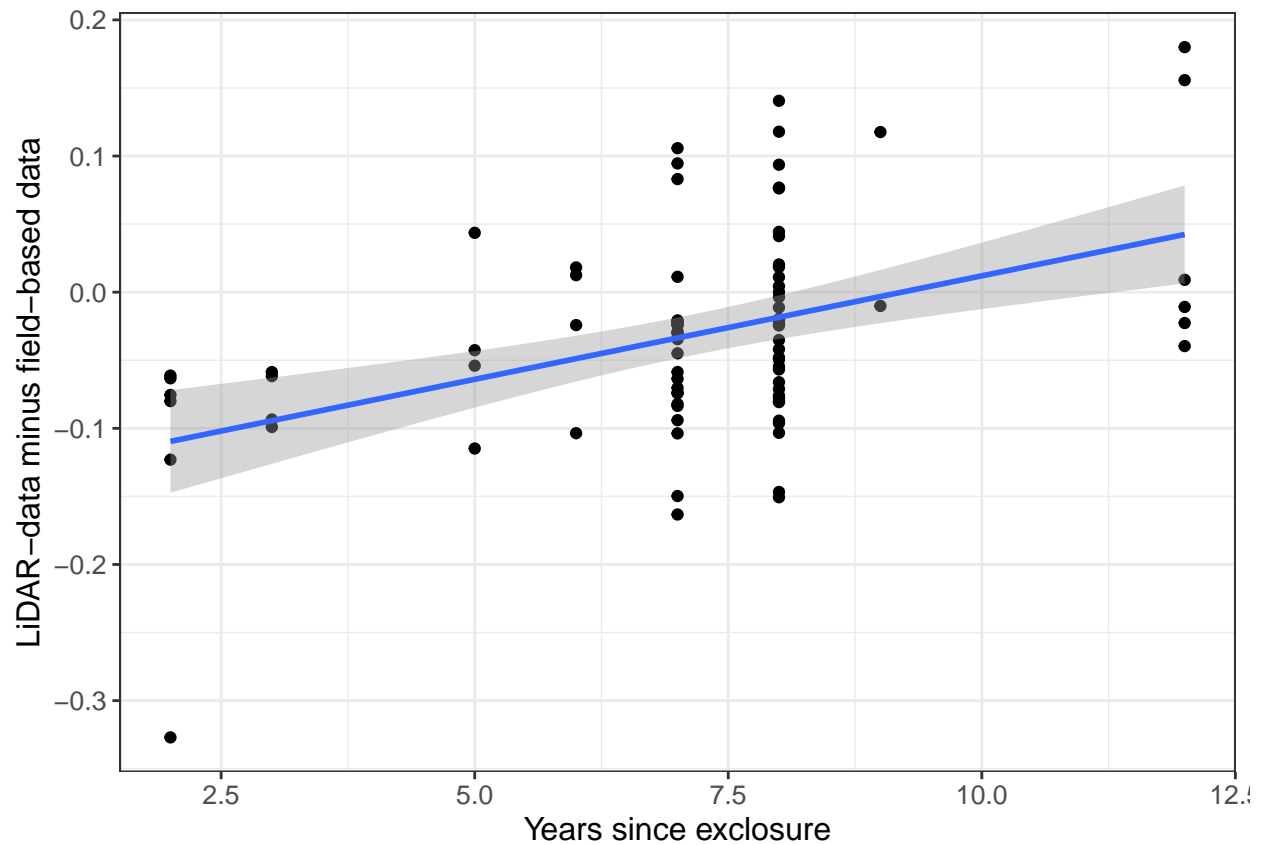
dat2$diff <- dat2$canopygrowth - dat2$canopygrowth_f
hist(dat2$diff)

```

Histogram of dat2\$diff



```
ggplot(data = dat2,  
       aes(x = YrsSinceExclosure, y = diff))+  
  geom_point()+  
  geom_smooth(method = "lm")+  
  labs(y="LiDAR-data minus field-based data",  
       x="Years since exclosure")+  
  theme_bw()+  
  theme(text = element_text(size = 12))+  
  #ylim(0, 0.4)+  
  theme(legend.position = 'right',  
        legend.justification = c("left", "top"),  
        legend.box.just = "left",  
        #legend.margin = margin(5, 5, 5, 5),  
        legend.text = element_text(size=12))+  
  geom_abline(intercept = 0, slope = 1, color="red",  
              linetype="dashed", size=1.5)
```



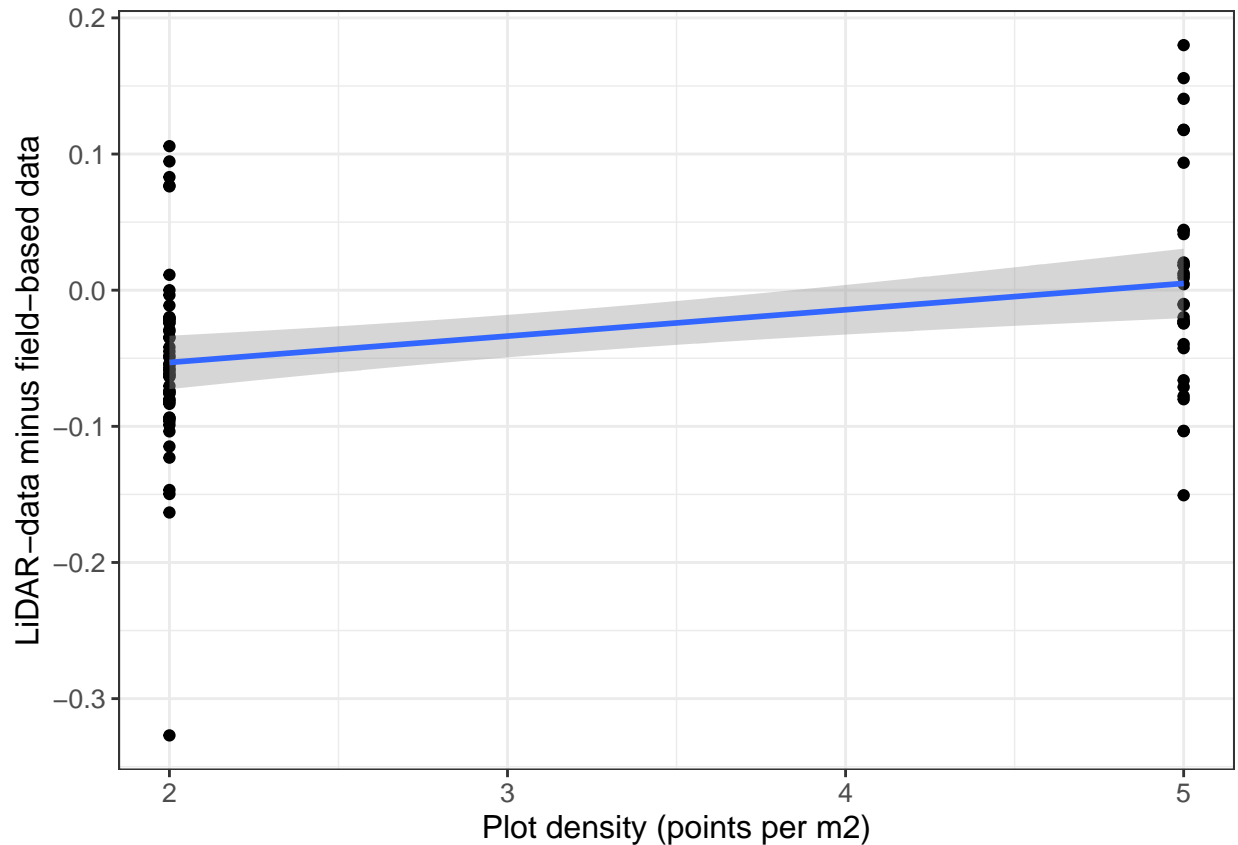
It looks like older difference between LiDAR and field-based data.

I would like to see if the correlation depends on LiDAR resolution, but then we should figure out what this is about:

```
table(dat2$plot_density_m2, dat2$resolution_m)
```

```
##
##      0,25 0,5
##      2    2  52
##      5    30   2
```

It's not clear which is derived from which, but I think its plot density that is the original.



No big difference.

Note that we are not seeing if the slope is close to 1 or not. We assume it will be. It's correlation strength that matters.

```
cor(dat2$canopygrowth, dat2$canopygrowth_f)
```

```
## [1] 0.5693676
```

And it's not very high. We can use a more sophisticated model to see where the field-based data is better or worse at predicting LiDAR results.

```
moddiff <- glmmTMB(canopygrowth~canopygrowth_f
  +region+canopygrowth_f:region
  +YrsSinceExclosure+YrsSinceExclosure:canopygrowth_f
  +plot_density_m2+plot_density_m2:canopygrowth_f
  +Treatment+Treatment:canopygrowth_f,
  family=gaussian,
  data=dat2)
summary(moddiff)
```

```
## Family: gaussian ( identity )
## Formula:
## canopygrowth ~ canopygrowth_f + region + canopygrowth_f:region +
## YrsSinceExclosure + YrsSinceExclosure:canopygrowth_f + plot_density_m2 +
## plot_density_m2:canopygrowth_f + Treatment + Treatment:canopygrowth_f
```

```
## Data: dat2
##
##      AIC      BIC    logLik deviance df.resid
##   -216.8   -184.9    121.4   -242.8      73
##
##
## Dispersion estimate for gaussian family (sigma^2): 0.00348
##
## Conditional model:
##
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    0.028696   0.067070   0.428   0.669
## canopygrowth_f -0.501168   0.437962  -1.144   0.252
## regionTelemark -0.021372   0.050695  -0.422   0.673
## regionTrondelag -0.027782   0.052147  -0.533   0.594
## YrsSinceExclosure -0.004944   0.010008  -0.494   0.621
## plot_density_m2  0.004487   0.014719   0.305   0.760
## TreatmentExclosure 0.021574   0.029502   0.731   0.465
## canopygrowth_f:regionTelemark 0.166680   0.400241   0.416   0.677
## canopygrowth_f:regionTrondelag 0.245282   0.404599   0.606   0.544
## canopygrowth_f:YrsSinceExclosure 0.109693   0.076821   1.428   0.153
## canopygrowth_f:plot_density_m2 0.077693   0.121632   0.639   0.523
## canopygrowth_f:TreatmentExclosure 0.141096   0.233017   0.606   0.545
```

Nothing significant. Applying dredge:

```
library(MuMIn)
dredge(moddiff, beta="none", rank = "AICc")
```

```
## Fixed terms are "cond((Int))" and "disp((Int))"
```

```
## Global model call: glmmTMB(formula = canopygrowth ~ canopygrowth_f + region + canopygrowth_f:region +
##   YrsSinceExclosure + YrsSinceExclosure:canopygrowth_f + plot_density_m2 +
##   plot_density_m2:canopygrowth_f + Treatment + Treatment:canopygrowth_f,
##   data = dat2, family = gaussian, ziformula = ~0, dispformula = ~1)
## ---
## Model selection table
##      cnd((Int)) dsp((Int)) cnd(cnp_f) cnd(plt_dns_m2) cnd(rgn) cnd(Trt)
## 284 0.0023810      +    -0.25930    0.0132300      +
## 412 0.0120600      +    -0.34500    0.0132000      +
## 316 0.0069280      +    -0.29720    0.0097240      +
## 282 0.0143100      +    -0.23430      +
## 444 0.0157800      +    -0.37570    0.0099600      +
## 288 0.0019960      +    -0.25810    0.0131900      +
## 276 0.0154100      +    -0.21890    0.0126900      +
## 308 0.0258100      +    -0.31440    0.0041080      +
## 410 0.0248900      +    -0.32840      +
## 286 0.0490300      +    -0.32760      +
## 416 0.0127300      +    -0.34760    0.0130200      +
## 28  -0.1254000      +     0.66090    0.0133100      +
## 320 0.0042720      +    -0.29230    0.0098010      +
## 60  -0.0913100      +     0.39650    0.0003423      +
## 274 0.0263900      +    -0.19630      +
## 414 0.0623100      +    -0.44350      +
```


## 156	-0.1305000	+	0.71430	0.0133400		+
## 280	0.0104300	+	-0.20820	0.0132600	+	
## 448	0.0141100	+	-0.37280	0.0099740	+	+
## 352	0.0023030	+	-0.25990	0.0132000	+	+
## 188	-0.0963300	+	0.44790	0.0004675		+
## 26	-0.1139000	+	0.69000			+
## 312	0.0156200	+	-0.29840	0.0046700	+	
## 278	0.0577500	+	-0.27790		+	
## 32	-0.1396000	+	0.66930	0.0150800	+	+
## 128	-0.0150200	+	-0.22430	-0.0066400	+	+
## 52	-0.0729000	+	0.38360	-0.0053750		
## 350	0.0488400	+	-0.32790		+	+
## 64	-0.1080000	+	0.38900	0.0019180	+	+
## 384	0.0089170	+	-0.33540	0.0059390	+	+
## 480	0.0178600	+	-0.38730	0.0130500	+	+
## 20	-0.1200000	+	0.76830	0.0127300		
## 12	-0.0691400	+	0.60820	0.0200400		+
## 154	-0.1186000	+	0.73910			+
## 96	-0.0886500	+	0.36830	0.0148200	+	+
## 160	-0.1447000	+	0.72150	0.0151700	+	+
## 256	0.0042250	+	-0.38600	-0.0081080	+	+
## 44	-0.0347000	+	0.34000	0.0068610		+
## 478	0.0669800	+	-0.48090		+	+
## 344	0.0079830	+	-0.19860	0.0131300	+	
## 30	-0.0963700	+	0.66200		+	+
## 192	-0.1130000	+	0.43950	0.0020680	+	+
## 512	0.0287000	+	-0.50120	0.0044870	+	+
## 112	0.0487800	+	-0.39710	-0.0034510	+	+
## 140	-0.0742200	+	0.66110	0.0200700		+
## 120	-0.0043950	+	-0.22100	-0.0111400	+	
## 56	-0.0953200	+	0.37400	-0.0030300	+	
## 18	-0.1092000	+	0.79260			
## 342	0.0542400	+	-0.26640		+	
## 224	-0.0867100	+	0.35180	0.0148000	+	+
## 172	-0.0396500	+	0.39070	0.0069860		+
## 94	-0.0449900	+	0.35430		+	+
## 376	0.0188000	+	-0.32810	0.0009479	+	
## 24	-0.1391000	+	0.77930	0.0152800	+	
## 158	-0.1003000	+	0.70400		+	+
## 240	0.0721600	+	-0.60100	-0.0054010	+	+
## 16	-0.0847000	+	0.62560	0.0219500	+	+
## 80	-0.0275100	+	0.24610	0.0206500	+	+
## 48	-0.0524300	+	0.33570	0.0082540	+	+
## 36	-0.0119100	+	0.32310	0.0010820		
## 222	-0.0412500	+	0.32180		+	+
## 104	0.0633800	+	-0.40290	-0.0080030	+	
## 4	-0.0598100	+	0.71960	0.0198400		
## 88	-0.0948500	+	0.52380	0.0149600	+	
## 22	-0.0953200	+	0.77290		+	
## 144	-0.0910700	+	0.68790	0.0220200	+	+
## 208	-0.0227600	+	0.20430	0.0205900	+	+
## 176	-0.0586600	+	0.39640	0.0084080	+	+
## 40	-0.0367100	+	0.31750	0.0032860	+	
## 86	-0.0508100	+	0.51060		+	

## 8	-0.0813200	+	0.73960	0.0225100	+	
## 72	-0.0292200	+	0.40180	0.0212600	+	
## 78	0.0733100	+	0.16150		+	+
## 10	-0.0083050	+	0.62530			+
## 14	0.0147400	+	0.59190		+	+
## 206	0.0818300	+	0.08194		+	+
## 138	-0.0125000	+	0.66970			+
## 142	0.0098830	+	0.64200		+	+
## 2	0.0002539	+	0.73440			
## 70	0.0746500	+	0.32050		+	
## 6	0.0209200	+	0.70950		+	
## 27	-0.0478900	+		0.0168800		+
## 11	-0.0150500	+		0.0211100		+
## 31	-0.0347200	+		0.0139000	+	+
## 15	-0.0004760	+		0.0188500	+	+
## 25	-0.0288100	+				+
## 29	0.0040740	+			+	+
## 13	0.0816700	+			+	+
## 9	0.0507400	+				+
## 3	0.0222000	+		0.0211100		
## 19	-0.0106400	+		0.0168800		
## 17	0.0084420	+				
## 23	0.0025320	+		0.0139000	+	
## 7	0.0367800	+		0.0188500	+	
## 21	0.0413300	+			+	
## 5	0.1189000	+			+	
## 1	0.0879900	+				
##	cnd(YSE) cnd(cnp_f:plt_dns_m2) cnd(cnp_f:rqn) cnd(cnp_f:Trt)					
## 284	-0.008476					
## 412	-0.009152				+	
## 316	-0.007803	0.02881				
## 282	-0.004758					
## 444	-0.008496	0.02662			+	
## 288	-0.008367					
## 276	-0.009101					
## 308	-0.007411	0.07081				
## 410	-0.005509				+	
## 286	-0.007004					
## 416	-0.009045				+	
## 28	0.010120					
## 320	-0.007587	0.03016				
## 60	0.010090	0.10650				
## 274	-0.005506					
## 414	-0.007912				+	
## 156	0.010120				+	
## 280	-0.008948					
## 448	-0.008296	0.02726			+	
## 352	-0.008277			+		
## 188	0.010090	0.10570			+	
## 26	0.013940					
## 312	-0.006928	0.07650				
## 278	-0.007579					
## 32	0.010450					
## 128	0.008615	0.18660		+		

## 52	0.010600	0.14930		
## 350	-0.007214		+	
## 64	0.010360	0.11460		
## 384	-0.004749	0.06617	+	
## 480	-0.008953		+	+
## 20	0.010720			
## 12				
## 154	0.013950			+
## 96	0.009280		+	
## 160	0.010410			+
## 256	0.008477	0.19830	+	+
## 44		0.10810		
## 478	-0.008046		+	+
## 344	-0.009819		+	
## 30	0.013500			
## 192	0.010330	0.11400		+
## 512	-0.004944	0.07769	+	+
## 112		0.20550	+	
## 140				+
## 120	0.009058	0.22670	+	
## 56	0.010770	0.15910		
## 18	0.014360			
## 342	-0.008757		+	
## 224	0.009267		+	+
## 172		0.10730		+
## 94	0.012250		+	
## 376	-0.003828	0.11100	+	
## 24	0.010960			
## 158	0.013480			+
## 240		0.22020	+	+
## 16				
## 80			+	
## 48		0.11870		
## 36		0.15410		
## 222	0.012220		+	+
## 104		0.24870	+	
## 4				
## 88	0.009977		+	
## 22	0.014060			
## 144				+
## 208			+	+
## 176		0.11800		+
## 40		0.16590		
## 86	0.012980		+	
## 8				
## 72			+	
## 78			+	
## 10				
## 14				
## 206			+	+
## 138				+
## 142				+
## 2				
## 70			+	

```

## 6
## 27 0.006445
## 11
## 31 0.007302
## 15
## 25 0.011140
## 29 0.010150
## 13
## 9
## 3
## 19 0.006445
## 17 0.011140
## 23 0.007302
## 7
## 21 0.010150
## 5
## 1
##      cnd(cnp_f:YSE) df  logLik   AICc delta weight
## 284      0.1402  7 121.087 -226.7  0.00  0.436
## 412      0.1453  8 121.176 -224.5  2.26  0.141
## 316      0.1351  8 121.147 -224.4  2.31  0.137
## 282      0.1408  6 117.585 -222.1  4.63  0.043
## 444      0.1403  9 121.227 -222.1  4.65  0.043
## 288      0.1400  9 121.089 -221.8  4.93  0.037
## 276      0.1491  6 117.381 -221.7  5.04  0.035
## 308      0.1359  7 117.730 -220.0  6.71  0.015
## 410      0.1464  7 117.684 -219.9  6.81  0.015
## 286      0.1495  8 118.812 -219.8  6.98  0.013
## 416      0.1453 10 121.182 -219.4  7.31  0.011
## 28      6 116.223 -219.4  7.35  0.011
## 320      0.1340 10 121.152 -219.4  7.37  0.011
## 60      7 117.052 -218.7  8.07  0.008
## 274      0.1493  5 114.402 -218.1  8.68  0.006
## 414      0.1563  9 118.963 -217.6  9.18  0.004
## 156      7 116.312 -217.2  9.55  0.004
## 280      0.1479  8 117.392 -216.9  9.82  0.003
## 448      0.1396 11 121.233 -216.9  9.84  0.003
## 352      0.1391 11 121.091 -216.6 10.12  0.003
## 188      8 117.130 -216.4 10.35  0.002
## 26      5 113.046 -215.3 11.40  0.001
## 312      0.1322  9 117.783 -215.2 11.54  0.001
## 278      0.1575  7 115.275 -215.1 11.62  0.001
## 32      8 116.409 -214.9 11.79  0.001
## 128     11 120.246 -214.9 11.81  0.001
## 52      6 113.889 -214.7 12.02  0.001
## 350      0.1514 10 118.814 -214.7 12.04  0.001
## 64      9 117.352 -214.3 12.40  0.001
## 384      0.1093 12 121.242 -214.2 12.53  0.001
## 480      0.1438 12 121.221 -214.2 12.57  0.001
## 20      5 112.323 -213.9 12.84  0.001
## 12      5 112.018 -213.3 13.45  0.001
## 154      6 113.116 -213.2 13.57  0.000
## 96     10 117.977 -213.0 13.72  0.000
## 160      9 116.492 -212.6 14.12  0.000

```

## 256		12	120.417	-212.6	14.18	0.000
## 44		6	112.792	-212.5	14.22	0.000
## 478	0.1569	11	118.994	-212.4	14.32	0.000
## 344	0.1562	10	117.489	-212.0	14.69	0.000
## 30		7	113.722	-212.0	14.73	0.000
## 192		10	117.427	-211.9	14.82	0.000
## 512	0.1097	13	121.425	-211.8	14.94	0.000
## 112		10	117.282	-211.6	15.11	0.000
## 140		6	112.097	-211.1	15.61	0.000
## 120		10	117.029	-211.1	15.61	0.000
## 56		8	114.349	-210.8	15.91	0.000
## 18		4	109.649	-210.8	15.93	0.000
## 342	0.1684	9	115.413	-210.5	16.28	0.000
## 224		11	117.980	-210.4	16.35	0.000
## 172		7	112.860	-210.3	16.45	0.000
## 94		9	115.292	-210.2	16.52	0.000
## 376	0.1054	11	117.888	-210.2	16.53	0.000
## 24		7	112.587	-209.7	17.00	0.000
## 158		8	113.773	-209.7	17.06	0.000
## 240		11	117.546	-209.5	17.21	0.000
## 16		7	112.298	-209.2	17.58	0.000
## 80		9	114.700	-209.0	17.71	0.000
## 48		8	113.219	-208.6	18.17	0.000
## 36		5	109.513	-208.3	18.46	0.000
## 222		10	115.303	-207.7	19.06	0.000
## 104		9	113.983	-207.6	19.14	0.000
## 4		4	108.005	-207.5	19.22	0.000
## 88		9	113.850	-207.3	19.41	0.000
## 22		6	110.060	-207.1	19.68	0.000
## 144		8	112.406	-206.9	19.80	0.000
## 208		10	114.718	-206.5	20.23	0.000
## 176		9	113.316	-206.3	20.47	0.000
## 40		7	110.176	-204.9	21.82	0.000
## 86		8	111.359	-204.8	21.89	0.000
## 8		6	108.436	-203.8	22.93	0.000
## 72		8	110.400	-202.9	23.81	0.000
## 78		8	109.344	-200.8	25.92	0.000
## 10		4	104.362	-200.2	26.51	0.000
## 14		6	106.547	-200.0	26.71	0.000
## 206		9	109.403	-198.4	28.30	0.000
## 138		5	104.409	-198.1	28.67	0.000
## 142		7	106.608	-197.8	28.96	0.000
## 2		3	101.104	-195.9	30.82	0.000
## 70		7	105.244	-195.1	31.69	0.000
## 6		5	102.886	-195.0	31.72	0.000
## 27		5	99.725	-188.7	38.04	0.000
## 11		4	98.486	-188.5	38.26	0.000
## 31		7	100.058	-184.7	42.06	0.000
## 15		6	98.612	-184.2	42.58	0.000
## 25		4	96.213	-183.9	42.80	0.000
## 29		6	98.475	-183.9	42.85	0.000
## 13		5	95.407	-180.1	46.67	0.000
## 9		3	92.174	-178.1	48.68	0.000
## 3		3	89.439	-172.6	54.15	0.000

```
## 19          4  90.440 -172.4 54.35  0.000
## 17          3  87.588 -168.9 57.85  0.000
## 23          6  90.708 -168.4 58.38  0.000
## 7           5  89.542 -168.3 58.40  0.000
## 21          5  89.431 -168.1 58.63  0.000
## 5           4  86.928 -165.4 61.38  0.000
## 1           2  84.256 -164.4 62.37  0.000
## Models ranked by AICc(x)
```

And the best model is...:

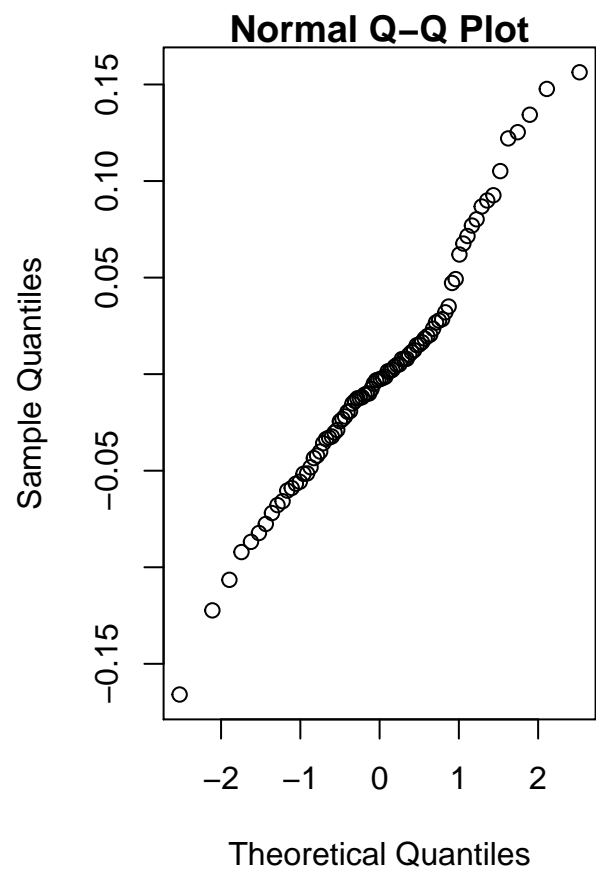
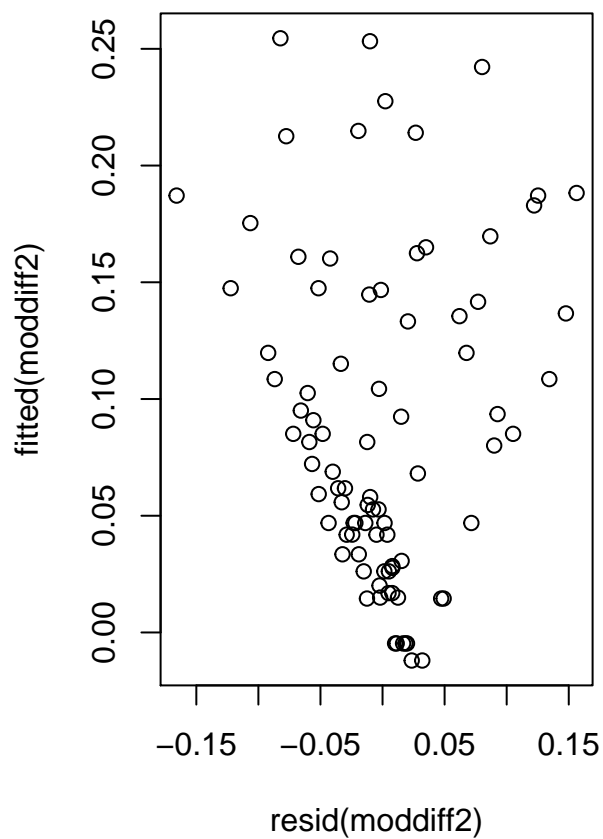
```
moddiff2 <- glmmTMB(canopygrowth~canopygrowth_f
                    +YrsSinceExclosure+YrsSinceExclosure:canopygrowth_f
                    +plot_density_m2
                    +Treatment,
                    family=gaussian,
                    data=dat2)
summary(moddiff2)
```

```
## Family: gaussian ( identity )
## Formula:
## canopygrowth ~ canopygrowth_f + YrsSinceExclosure + YrsSinceExclosure:canopygrowth_f +
##      plot_density_m2 + Treatment
## Data: dat2
##
##      AIC      BIC    logLik deviance df.resid
##   -228.2   -211.0     121.1   -242.2        79
##
##
## Dispersion estimate for gaussian family (sigma^2): 0.0035
##
## Conditional model:
##
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      0.002381   0.047428   0.050   0.95997
## canopygrowth_f    -0.259305   0.303197  -0.855   0.39242
## YrsSinceExclosure -0.008476   0.006628  -1.279   0.20100
## plot_density_m2    0.013229   0.004898   2.701   0.00691 **
## TreatmentExclosure 0.038179   0.013724   2.782   0.00540 **
## canopygrowth_f:YrsSinceExclosure 0.140224   0.043695   3.209   0.00133 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

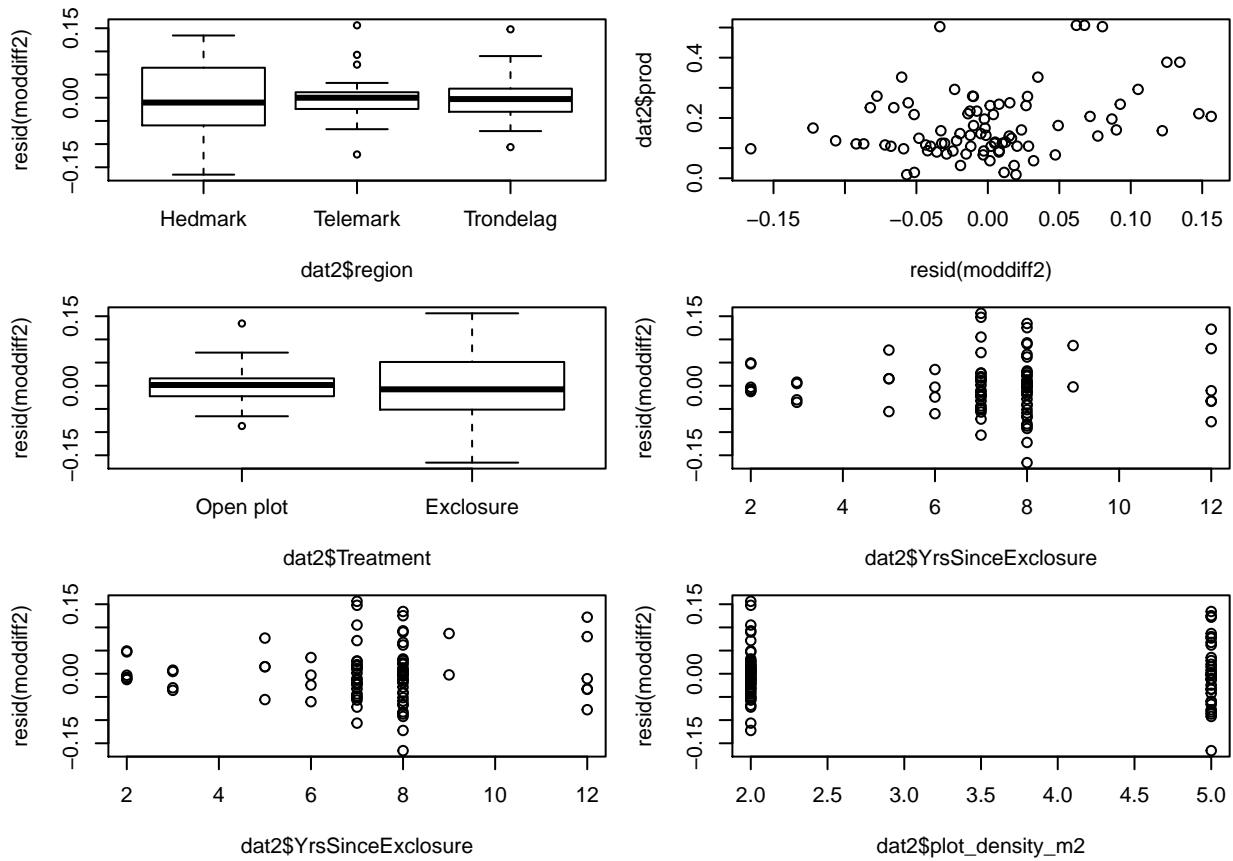
Interpretation: Field-based data have a medium correlation to LiDAR data (0.5ish). The correlation is dependent on Years since exclosure (probably this could be exchanged with tree height?) and LiDAR data is more likely to underestimate canopy growth at short experimental durations (see figure above). Similarly, lower LiDAR resolution induces underestimates in the same way. Field-based data is also not able to explain the variation induced by the fencing treatment. This could be a three way interaction as well, including years since exclusion. Not sure if it should be included here actually.

Validation

```
par(mfrow=c(1,2), mar = c(4,4,1,1))
plot(resid(moddiff2), fitted(moddiff2))
qqnorm(resid(moddiff2))
```



```
par(mfrow=c(3,2), mar = c(4,4,1,1))
plot(resid(moddiff2)~ dat2$region)
plot(resid(moddiff2), dat2$prod)
plot(resid(moddiff2)~ dat2$Treatment)
plot(resid(moddiff2)~ dat2$YrsSinceExclosure)
plot(resid(moddiff2)~ dat2$YrsSinceExclosure)
plot(resid(moddiff2)~ dat2$plot_density_m2)
```



Looks Ok

We also want to analyse MAD, and correlations between Lidar and field based treatment-differences.