

**Data initialisation** -> all valid data output as DBHdata

Original data -> 474 rows, 15 columns

Only extract DBH-related data for DBH modelling -> first 9 rows

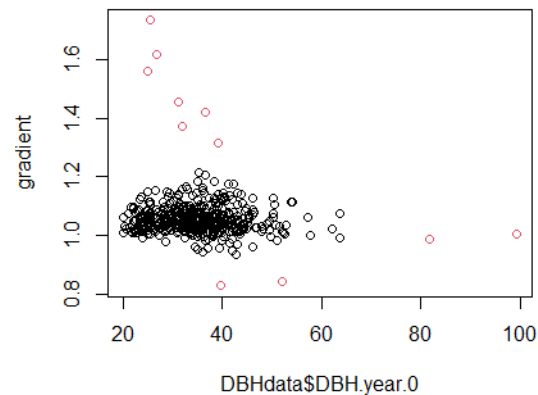
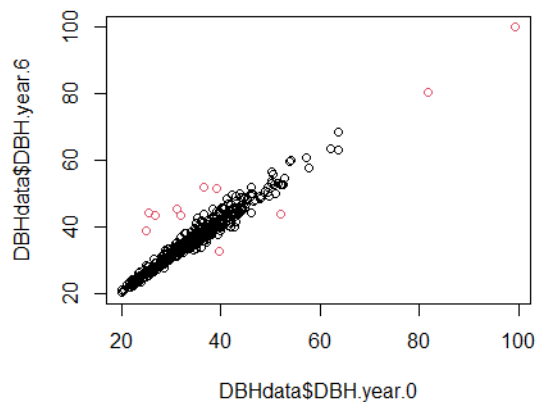
Assume “No” where there’s blank in “Large”

Remove rows with NA in “DBH.year.0” or NA in “DBH.year.6” -> 11 instances

**Further data preparation for models** -> DBHdatarm, DBHpositiverm

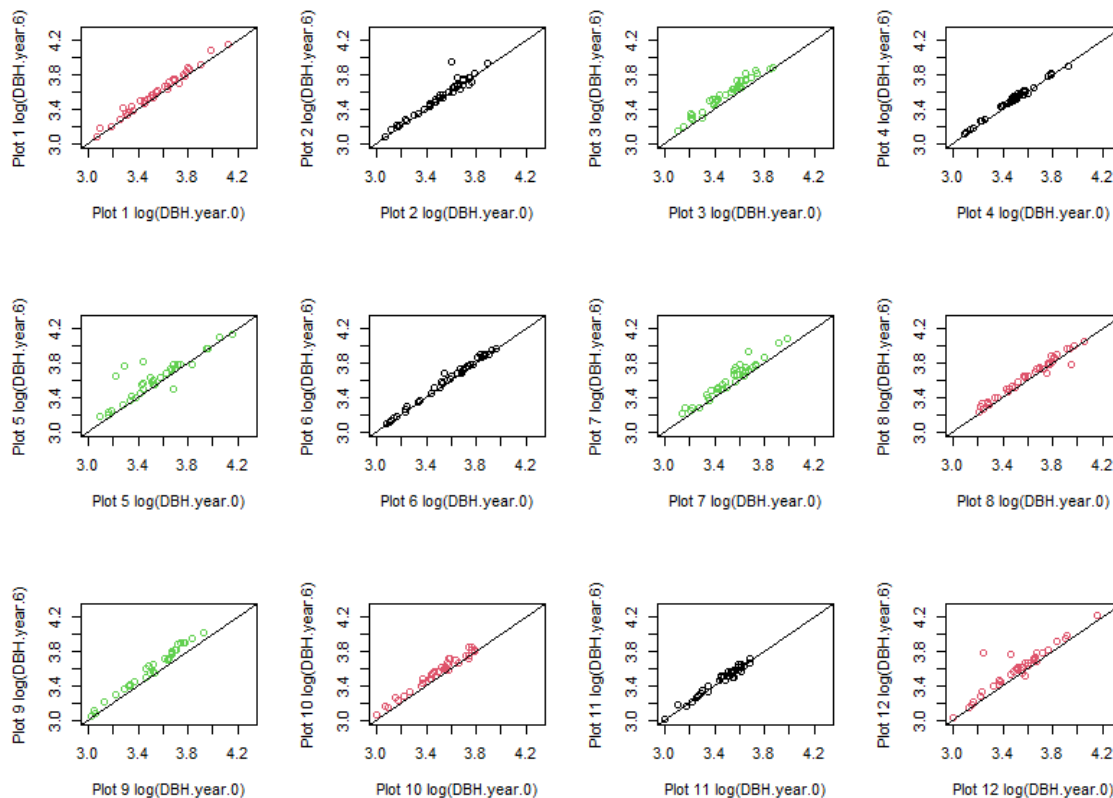
“positive” stands for positive tree growth, i.e. further remove rows where DBH.year.6 <= DBH.year.0 (tree growth <= 0)

“rm” stands for remove outliers (in the scatter plot, points outside the main group)



##	Tree.Number	Plot	Treatment	Large	Alive.at.year.6	DBH.year.0	DBH.year.6	Class.year.0	Class.year.6
## 20	58	8	T1: Gap	Yes	Yes	52.00000	43.9000	6	5
## 95	72	12	T1: Gap	Yes	Yes	31.85000	43.7000	4	5
## 96	71	12	T1: Gap	Yes	Yes	25.50000	44.2000	3	5
## 141	478	5	T2: Radial	No	Yes	31.20000	45.4000	4	5
## 144	172	5	T2: Radial	Yes	Yes	39.70000	32.9000	4	4
## 148	174	5	T2: Radial	Yes	Yes	99.30000	100.0000	NA	NA
## 149	176	5	T2: Radial	Yes	Yes	81.80000	80.7000	NA	NA
## 156	184	5	T2: Radial	Yes	Yes	26.80000	43.4000	3	5
## 157	185	5	T2: Radial	Yes	Yes	24.90000	38.9000	3	4
## 198	225	2	Control	Yes	Yes	36.60564	52.0000	4	6
## 412	208	7	T2: Radial	Yes	Yes	39.20000	51.5662	4	6

## random effect



Random effect is not significant. From the 12 scatter plots, it seems that there's no random intercept or random slope.

But it's still necessary to keep the random intercept term as a reference to human errors.

**Models:** random intercept mixed effect model using DBH.year.6 as response variable. Two version: Large VS InGap

### random intercept using InGap for log(DBHdatarm\$DBH.year.6)

Compared to the suggested model:

$\log(\text{DBH.year.6}) \sim \log(\text{DBH.year.0}) + \text{Treatment} + \text{Large} + \log(\text{DBH.year.0}) : \text{InGap} + \text{Treatment} : \text{InGap} + (1 | \text{Plot})$ ,

we've removed "Treatment:InGap" and replace "Large" with "InGap" (because "Large" is insignificant when there's "log(DBH.year.0):InGap")

**Question:** Whether "Treatment:InGap" is useless? (It's NA when Control or T1)

```

mixed2.0 = lmer(log(DBH.year.6)~log(DBH.year.0)+Treatment+InGap+log(DBH.year.
0):InGap+(1|Plot), data = DBHdatarm)
summary(mixed2.0)

## Linear mixed model fit by REML ['lmerMod']
## Formula:
## log(DBH.year.6) ~ log(DBH.year.0) + Treatment + InGap + log(DBH.year.0):In
Gap +
##      (1 | Plot)
##      Data: DBHdatarm
##
## REML criterion at convergence: -1778
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.6211 -0.6010 -0.0148  0.5975  3.2852
##
## Random effects:
##      Groups      Name              Variance Std.Dev.
##      Plot      (Intercept) 8.593e-05 0.00927
##      Residual              1.006e-03 0.03172
## Number of obs: 452, groups: Plot, 12
##
## Fixed effects:
##
##              Estimate Std. Error t value
## (Intercept)      0.072639   0.027788   2.614
## log(DBH.year.0)      0.986722   0.007783 126.785
## TreatmentT1: Gap      0.028683   0.007475   3.837
## TreatmentT2: Radial    0.039231   0.007962   4.927
## InGapTRUE            0.452284   0.107566   4.205
## log(DBH.year.0):InGapTRUE -0.114046   0.029285  -3.894
##
## Correlation of Fixed Effects:
##              (Intr) lg(DBH..0) TrT1:G TrT2:R IGTRUE
## lg(DBH.y.0)  -0.982
## TrtmntT1:Gp -0.095 -0.040
## TrtmntT2:Rd -0.248  0.125   0.463
## InGapTRUE    -0.237  0.241  -0.010  0.017
## l(DBH..0):I  0.258 -0.263   0.010 -0.035 -0.998

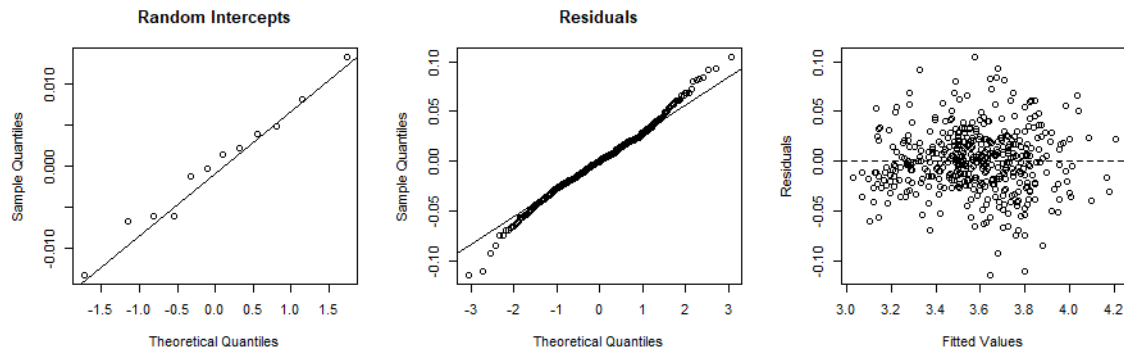
VarCorr(mixed2.0)

##      Groups      Name              Std.Dev.
##      Plot      (Intercept) 0.0092699
##      Residual              0.0317162

y.hat <- fitted(mixed2.0) # Fitted values
int.hat <- ranef(mixed2.0)[[1]][[1]] # Predicted intercepts
res.hat <- residuals(mixed2.0) # Estimated residuals
tss = sum((log(DBHdatarm$DBH.year.6)-mean(log(DBHdatarm$DBH.year.6)))^2) #TSS

```

```
rss = sum((log(DBHdatarm$DBH.year.6)-y.hat)^2) #RSS
(tss-rss)/tss #R^2 = 0.9801595
## [1] 0.9801595
```



**Reason:** high  $R^2$  value, small residual, simpler model, easy to interpret

**Achievement of this model:** explain priority1

(Thinning will increase tree growth in both treatments relative to the control.

Thinning will have an increased effect on the rate of tree growth in T2 as compared to T1.)

**Problem with this model:**

1. Cannot explain priority2, because “Large” is not included.

(Thinning will increase tree growth in the largest tree class size both treatments relative to the control.

Thinning will have an increased effect on the rate of tree growth in T2 as compared to T1 in the largest tree class size.)

-> use scatter plot of tree.growth VS DBH.year.0 to show trend

2. Cannot explain priority3.1 (larger DBH.year.0 results in larger tree growth). INSTEAD, we can only say: larger DBH.year.0 results in larger DBH.year.6.

(The greater the tree class size the greater the relative increase in the DBH after thinning across both treatments.)

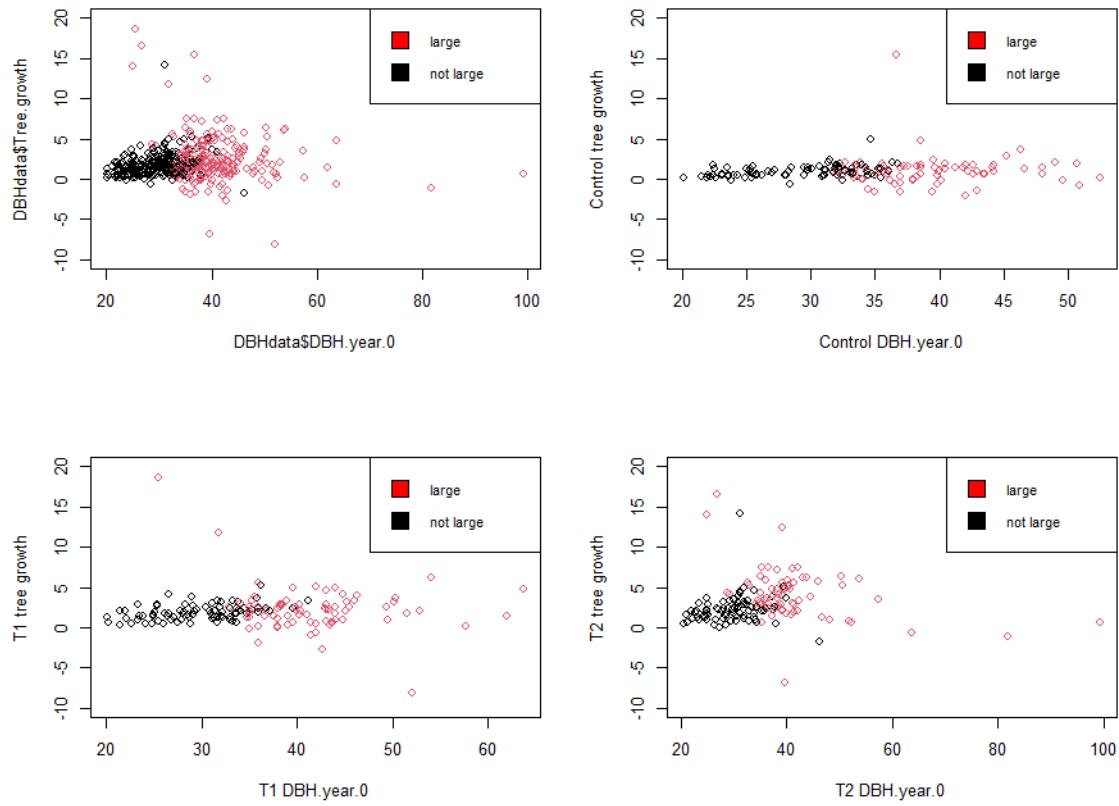
-> use scatter plot of tree.growth VS DBH.year.0 to show trend

3. Coefficients for “InGap” and “log(DBH.year.0):InGap” form constrains for DBH.year.0. Not sure how to interpret the constrains.

Constrains more strict to less strict: InGap>T2notInGap, InGap>T1, InGap>Control; model using DBHpositiverm, model using DBHdatarm

-> guess: few data points over constrains, so model is not accurate for large DBH.year.0

## scatter plot of tree.growth VS DBH.year.0



## Visualise constrain according to mixed2.0

