

## Appendix S2: Statistical analyses

Reb Bryant      Shan Kothari      Jeannine Cavender-Bares      Stephanie J. Curran  
Jake J. Grossman      Sarah E. Hobbie      Charlotte Nash      Grace C. Neumiller  
Craig R. See

### Setting up the workspace

This document is meant to provide a narrative summary to the statistical analyses accompanying the manuscript “Independent effects of tree diversity on aboveground and soil carbon pools after six years of experimental afforestation” by Bryant et al. (in review, as of 2023-12-01).

Our first step is to load a number of packages. Although this paper was a collaborative work in which multiple co-authors were working in their own R environments, all results have been checked under the package numbers listed in *R v.4.2.1*.

```
library(lme4) ## 1.1.32
library(lmerTest) ## 3.1.3
library(MuMIn) ## 1.47.5
library(car) ## 3.1.2
library(performance) ## 0.10.5
library(piecewiseSEM) ## 2.3.0
library(MASS) ## 7.3.57
library(ggplot2) ## 3.4.2
```

Next, we go to the working directory where we saved the archived data file and read it in.

```
# setwd(".")
carbon_seq<-read.csv("Cseq.csv")
```

### Relationships among aboveground, soil, and fine root carbon

One of the first results presented in the paper is the relationship between soil C in 2013 (pre-treatment) and 2019.

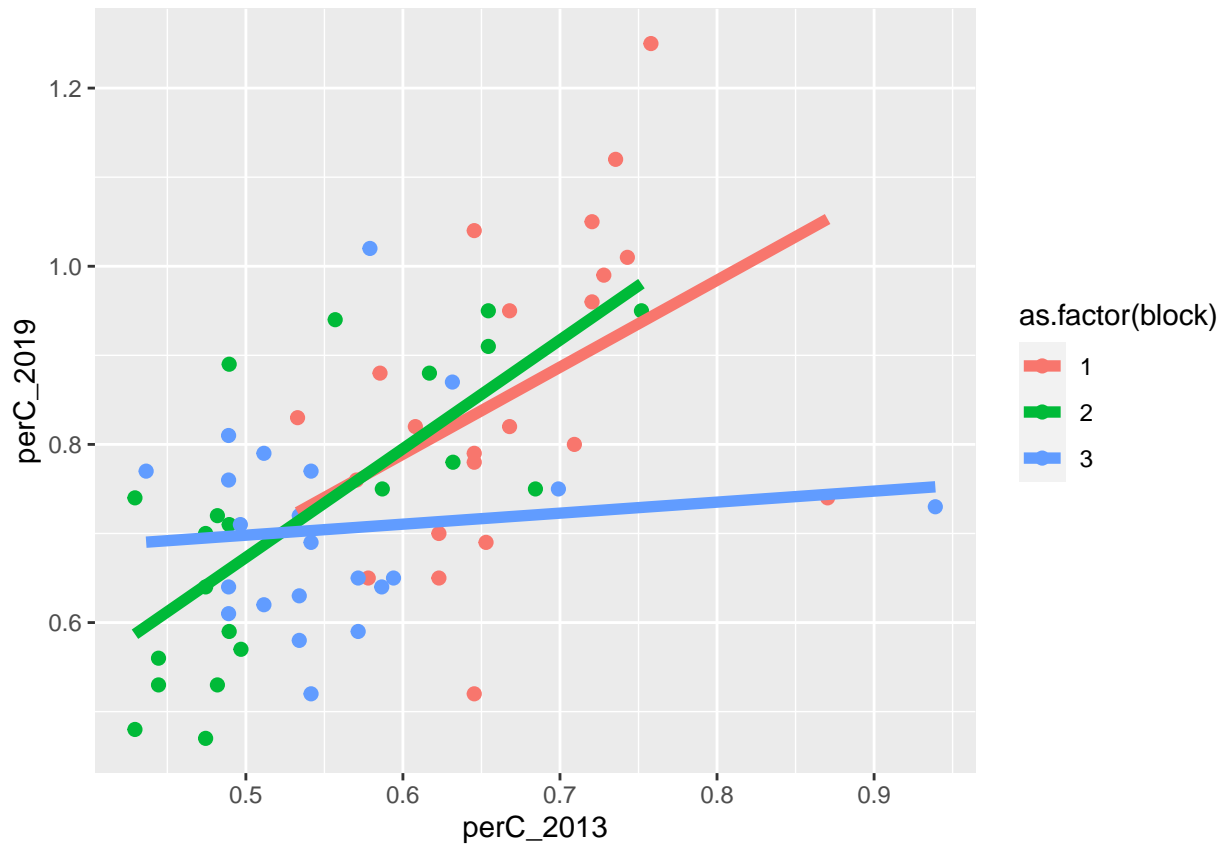
```
## here we use initial and final percent C, although the results are similar
## using initial and final soil C pools (which accounts for bulk density variation)

## one can add a block random intercept here with lmer, adding +(1/block)
## however, it doesn't change parameter values much
summary(lm(perC_2019~perC_2013,data=carbon_seq))
```

```
##
## Call:
## lm(formula = perC_2019 ~ perC_2013, data = carbon_seq)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.32357 -0.09997 -0.00773  0.09573  0.34801
##
```

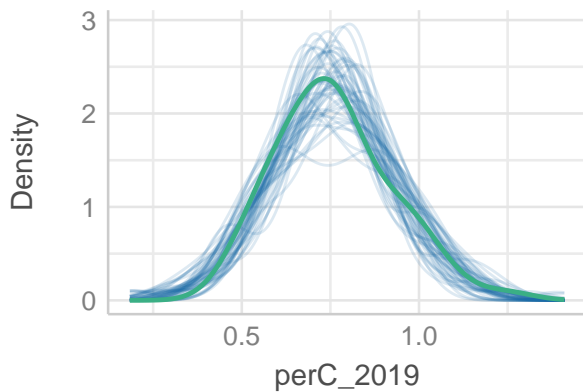
```
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.26754    0.09585   2.791  0.00691 **
## perC_2013    0.83710    0.16032   5.221  2.06e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1375 on 64 degrees of freedom
## (74 observations deleted due to missingness)
## Multiple R-squared:  0.2987, Adjusted R-squared:  0.2878
## F-statistic: 27.26 on 1 and 64 DF,  p-value: 2.058e-06

## in general, we plot separate regression lines for each block just to be
## able to check the generality within/among blocks more easily
## naturally, these do not correspond to the 'overall' best-fit line
ggplot(data=carbon_seq,aes(x=perC_2013,y=perC_2019,color=as.factor(block)))+
  geom_point(size=2)+geom_smooth(method="lm",se=F,linewidth=2)
```



### Posterior Predictive Check

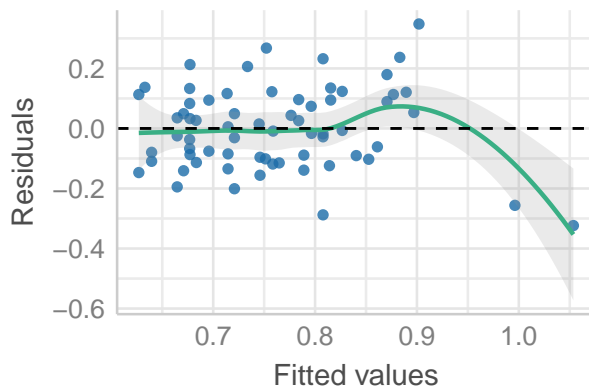
Model-predicted lines should resemble observed data



— Observed data — Model-predicted data

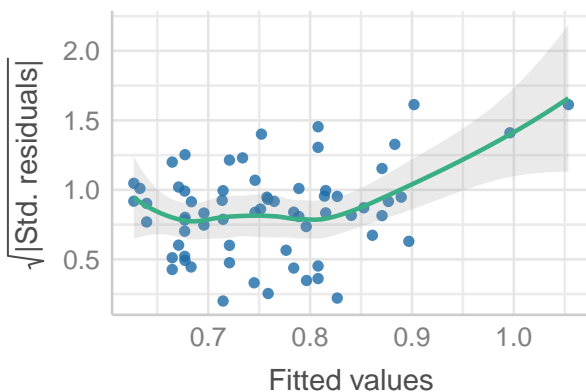
### Linearity

Reference line should be flat and horizontal



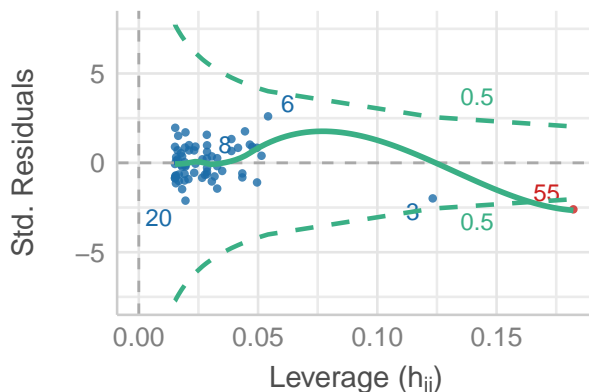
### Homogeneity of Variance

Reference line should be flat and horizontal



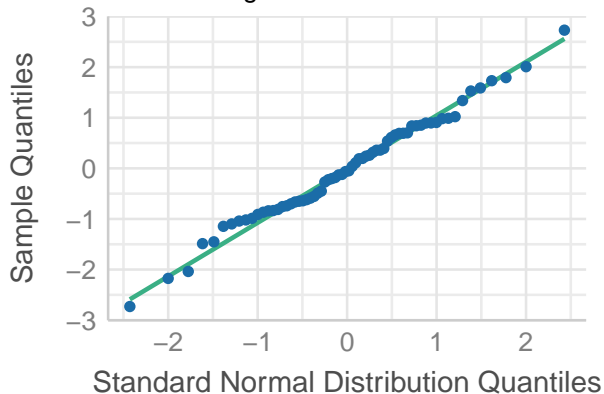
### Influential Observations

Points should be inside the contour lines



### Normality of Residuals

Dots should fall along the line



Despite a couple influential observations that declined in soil percent C, it's clear that pre-treatment and 2019 soil C are positively correlated. If we want to understand the drivers of carbon *sequestration*, we have good reason to consider the change in soil C pools rather than simply the endpoint in 2019. We might first want to check whether initial soil C determines how much C accumulates.

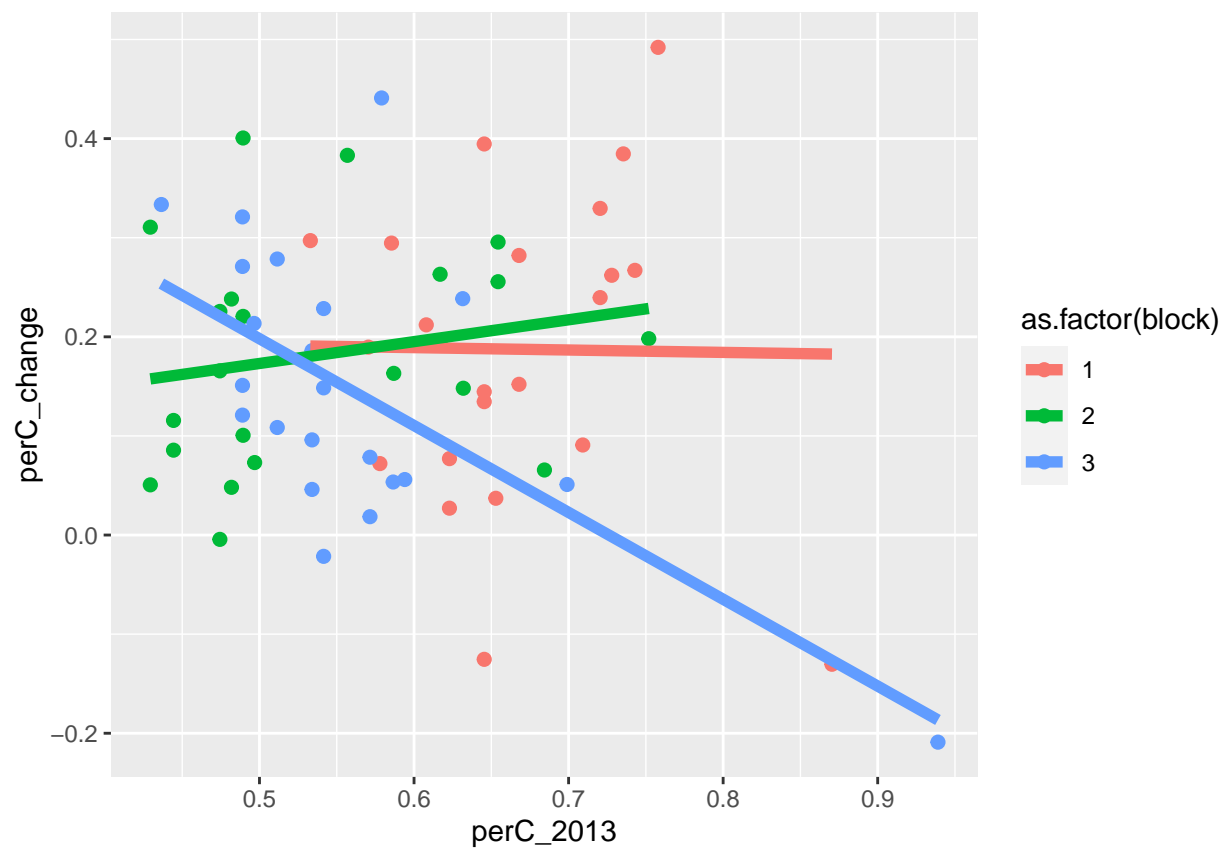
```

## again these results are robust to alternate ways of analyzing the data
## including considering the total pools and adding block random effects
carbon_seq$perC_change<-carbon_seq$perC_2019-carbon_seq$perC_2013
summary(lm(perC_change~perC_2013,data=carbon_seq))

##
## Call:
## lm(formula = perC_change ~ perC_2013, data = carbon_seq)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.32357 -0.09997 -0.00773  0.09573  0.34801
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.26754    0.09585   2.791  0.00691 **
## perC_2013    -0.16290    0.16032  -1.016  0.31340
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1375 on 64 degrees of freedom
## (74 observations deleted due to missingness)
## Multiple R-squared:  0.01588,    Adjusted R-squared:  0.0004994
## F-statistic: 1.032 on 1 and 64 DF,  p-value: 0.3134

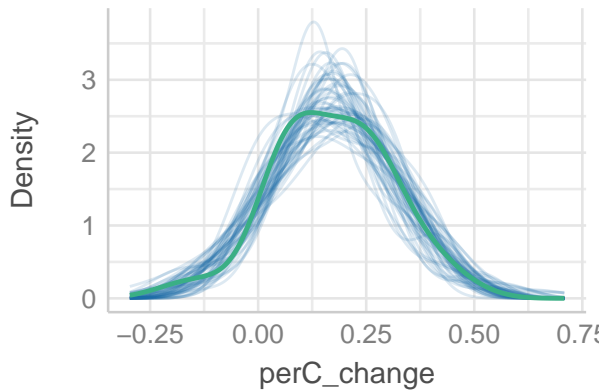
ggplot(data=carbon_seq,aes(x=perC_2013,y=perC_change,color=as.factor(block)))+
  geom_point(size=2)+geom_smooth(method="lm",se=F,linewidth=2)

```



### Posterior Predictive Check

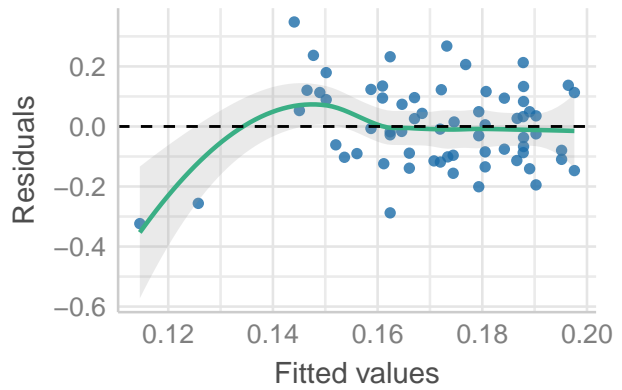
Model-predicted lines should resemble observed c



— Observed data — Model-predicted data

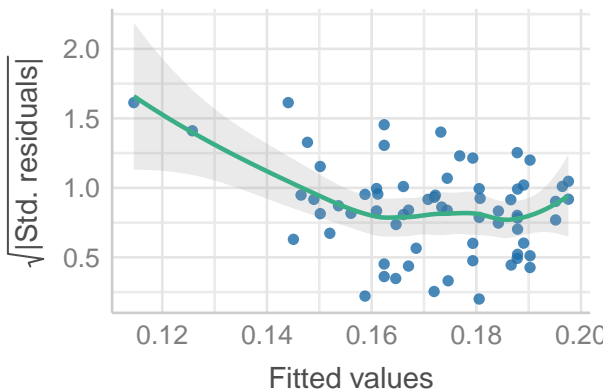
### Linearity

Reference line should be flat and horizontal



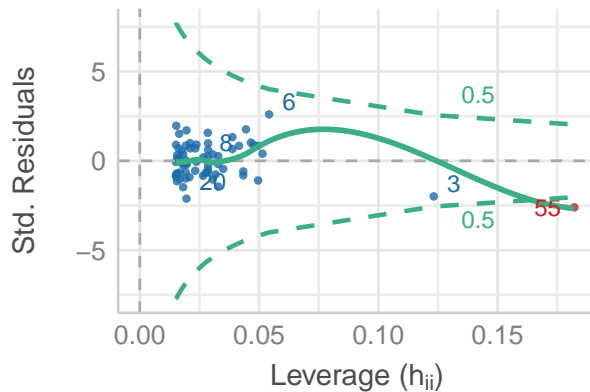
### Homogeneity of Variance

Reference line should be flat and horizontal



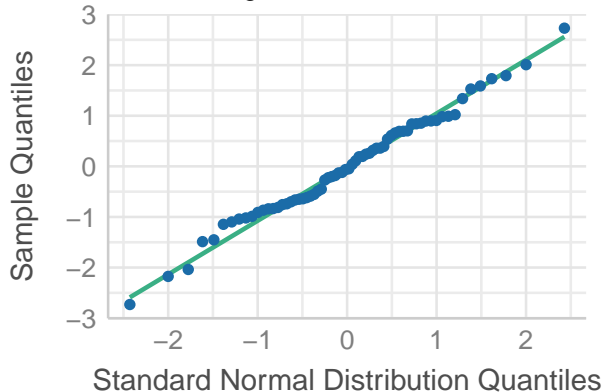
### Influential Observations

Points should be inside the contour lines



### Normality of Residuals

Dots should fall along the line



We see the same influential observations as before, but there remains no evidence for a relationship between initial soil C and change in soil C. We can assume that change in soil C is mainly driven by our treatments rather than be legacies of pre-existing soil conditions.

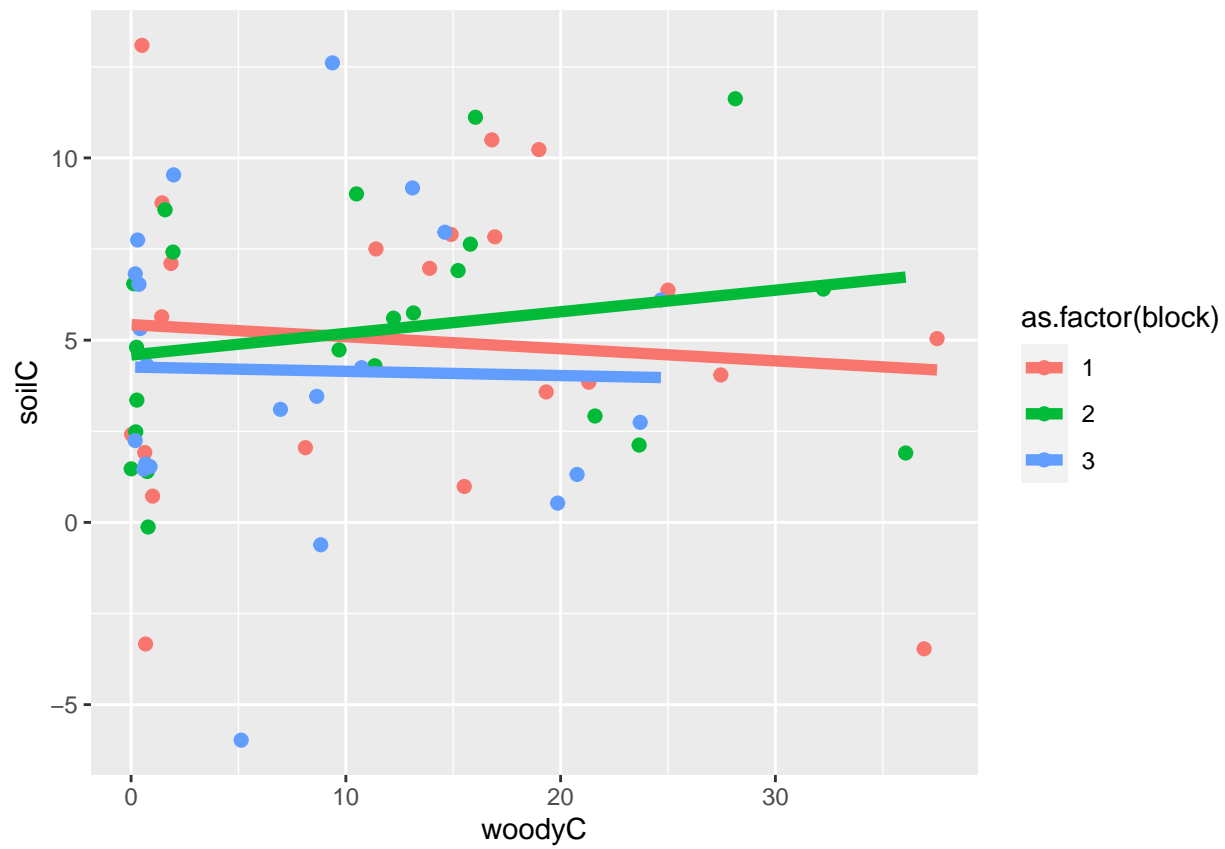
One of the key questions of the manuscript is: does change in soil C correlate with aboveground (woody) C?

```
## here we'll switch to using pools so that aboveground and soil are  
## in the same units
```

```
## note that a block random intercept results in a singular fit  
summary(lm(soilC~woodyC,data=carbon_seq))
```

```
##  
## Call:  
## lm(formula = soilC ~ woodyC, data = carbon_seq)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max   
## -10.704  -2.742  -0.100   2.720   8.424   
##  
## Coefficients:  
##              Estimate Std. Error t value Pr(>|t|)      
## (Intercept)  4.65738    0.67854   6.864 3.22e-09 ***  
## woodyC       0.01400    0.04479   0.313  0.756      
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 3.848 on 64 degrees of freedom  
##    (74 observations deleted due to missingness)  
## Multiple R-squared:  0.001524,    Adjusted R-squared:  -0.01408   
## F-statistic: 0.09767 on 1 and 64 DF,  p-value: 0.7557
```

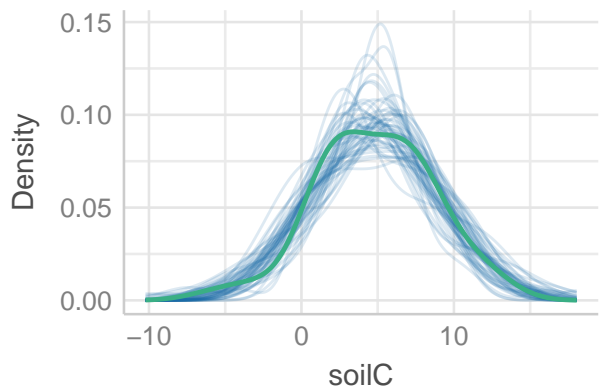
```
ggplot(data=carbon_seq,aes(x=woodyC,y=soilC,color=as.factor(block)))+  
  geom_point(size=2)+geom_smooth(method="lm",se=F,linewidth=2)
```





### Posterior Predictive Check

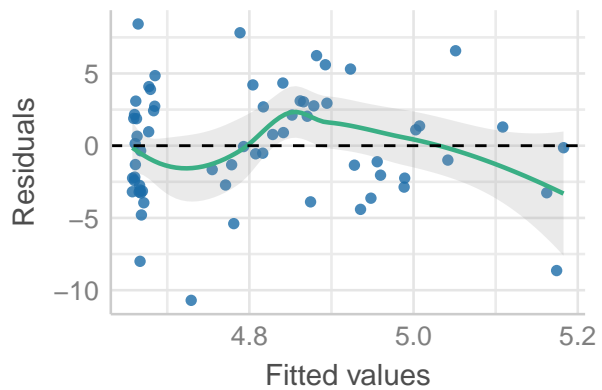
Model-predicted lines should resemble observed data



— Observed data — Model-predicted data

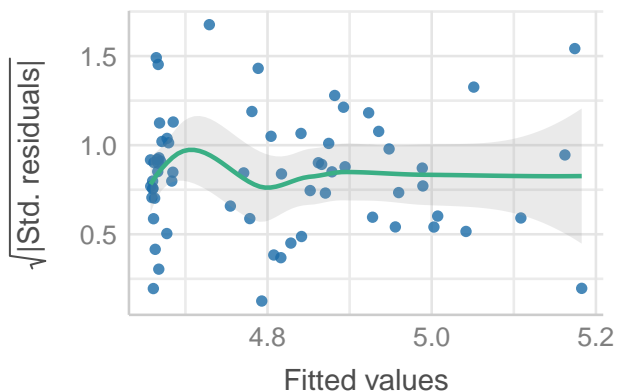
### Linearity

Reference line should be flat and horizontal



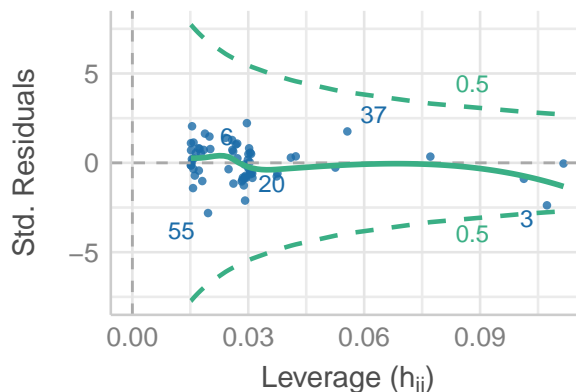
### Homogeneity of Variance

Reference line should be flat and horizontal



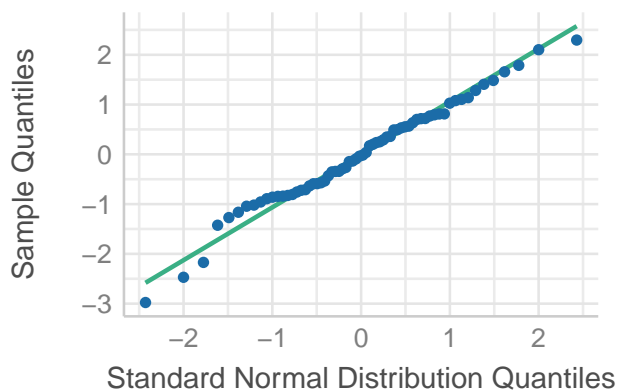
### Influential Observations

Points should be inside the contour lines



### Normality of Residuals

Dots should fall along the line



It is clear that there is no relationship between the two variables. However, as we report in the paper, if we subset the data to monoculture plots only, we see a significant interaction between species ID and aboveground C on soil C.

```
carbon_seq_mono<-subset(carbon_seq,species_richness==1)
## summary(aov(soilC~woodyC*species,data=carbon_seq_mono))
# ggplot(data=carbon_seq,aes(x=woodyC,y=soilC,color=as.factor(block)))+
#   geom_point(size=2)+geom_smooth(method="lm",se=F,linewidth=2)
```

Lastly, we look at the third pool we measured: fine root C. First of all, how does fine root C relate to soil C accumulation?

```
## note that a block random intercept results in a singular fit
summary(lm(soilC~rootC,data=carbon_seq))
```

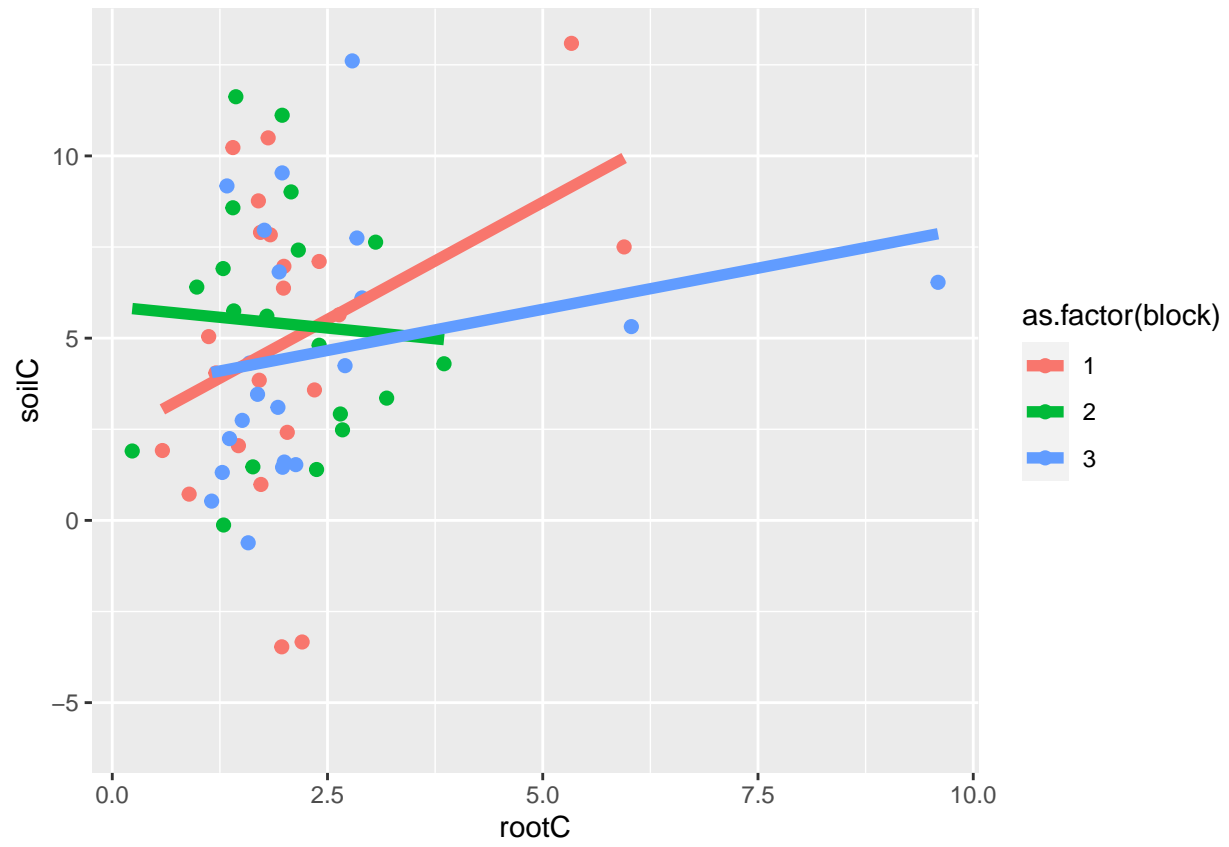
```
##
## Call:
## lm(formula = soilC ~ rootC, data = carbon_seq)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-8.3431	-2.5301	-0.3742	2.4080	7.2747

```
##
## Coefficients:
```

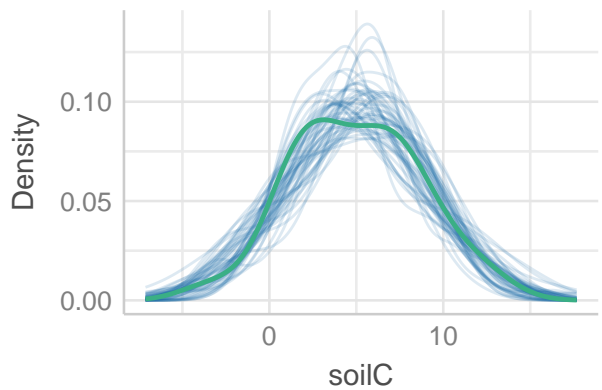
	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	3.7722	0.8464	4.457	3.7e-05 ***
rootC	0.5598	0.3245	1.725	0.0897 .

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.609 on 60 degrees of freedom
## (78 observations deleted due to missingness)
## Multiple R-squared:  0.04726,    Adjusted R-squared:  0.03138
## F-statistic: 2.976 on 1 and 60 DF,  p-value: 0.08965
ggplot(data=carbon_seq,aes(x=rootC,y=soilC,color=as.factor(block)))+
  geom_point(size=2)+geom_smooth(method="lm",se=F,linewidth=2)
```



### Posterior Predictive Check

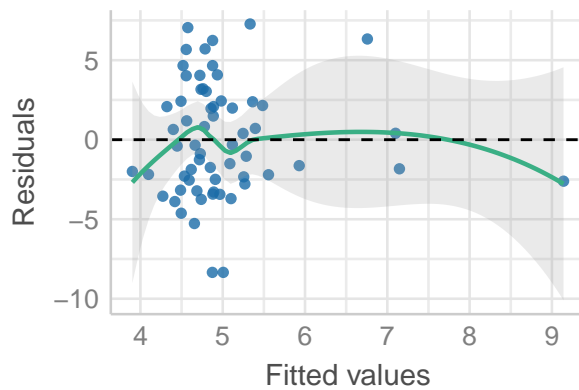
Model-predicted lines should resemble observed data



— Observed data — Model-predicted data

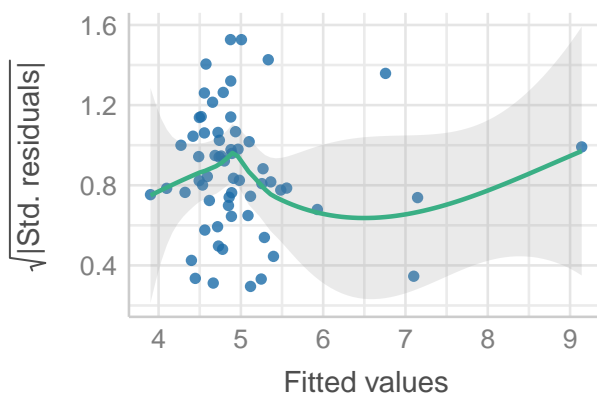
### Linearity

Reference line should be flat and horizontal



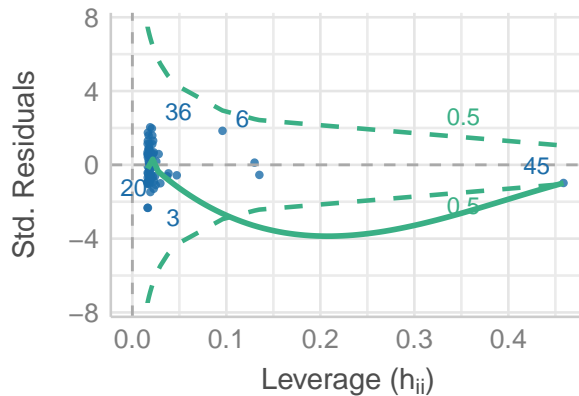
### Homogeneity of Variance

Reference line should be flat and horizontal



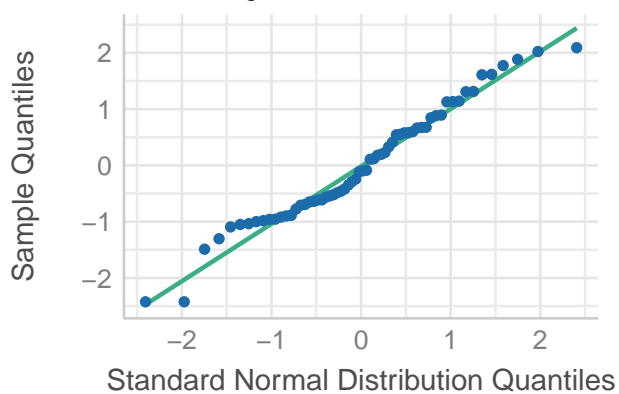
### Influential Observations

Points should be inside the contour lines



### Normality of Residuals

Dots should fall along the line

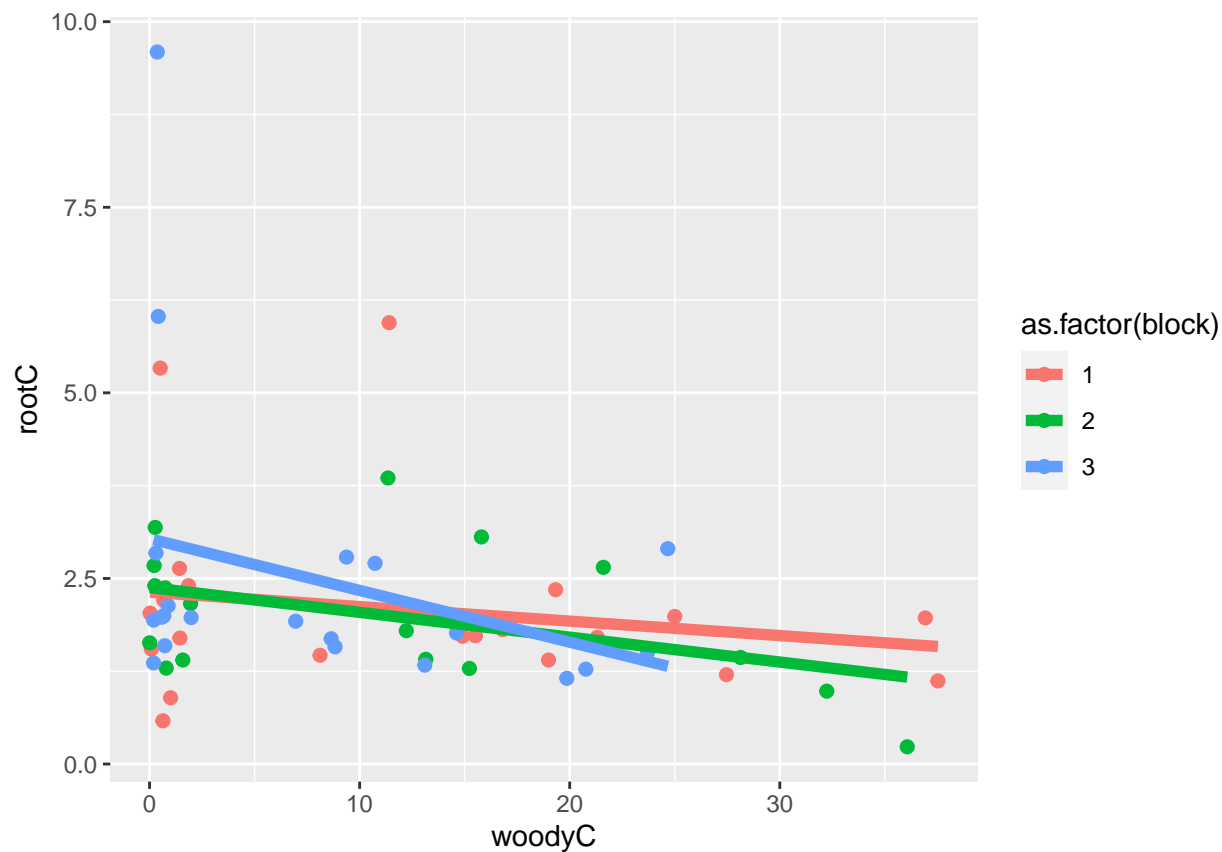


There is fairly weak evidence for a positive relationship between fine root C and soil C accumulation. Next, we check the relationship with aboveground C.

```
## note that a block random intercept results in a singular fit
summary(lm(rootC~woodyC,data=carbon_seq))
```

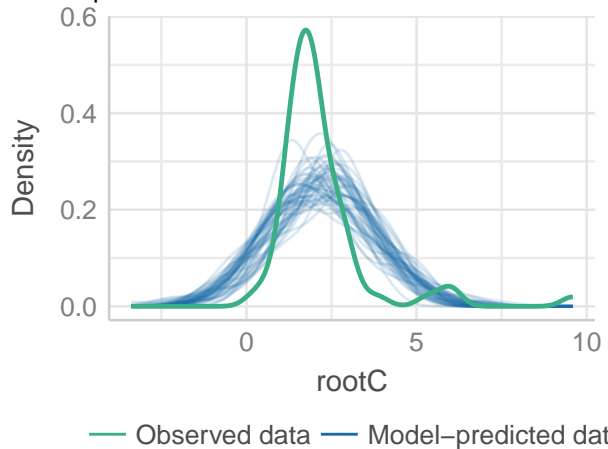
```
##
## Call:
## lm(formula = rootC ~ woodyC, data = carbon_seq)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9760 -0.6554 -0.3069  0.1031  7.0232
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.58183    0.24467  10.552 2.19e-15 ***
## woodyC       -0.03711    0.01615  -2.299  0.025 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.368 on 61 degrees of freedom
## (77 observations deleted due to missingness)
## Multiple R-squared:  0.07971,    Adjusted R-squared:  0.06463
## F-statistic: 5.284 on 1 and 61 DF,  p-value: 0.02497
```

```
ggplot(data=carbon_seq,aes(x=woodyC,y=rootC,color=as.factor(block)))+
  geom_point(size=2)+geom_smooth(method="lm",se=F,linewidth=2)
```



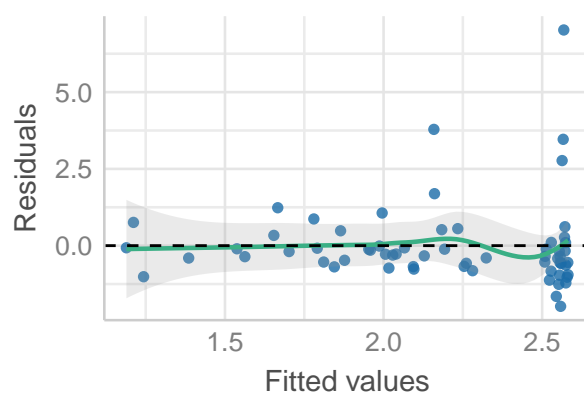
### Posterior Predictive Check

Model-predicted lines should resemble observed data



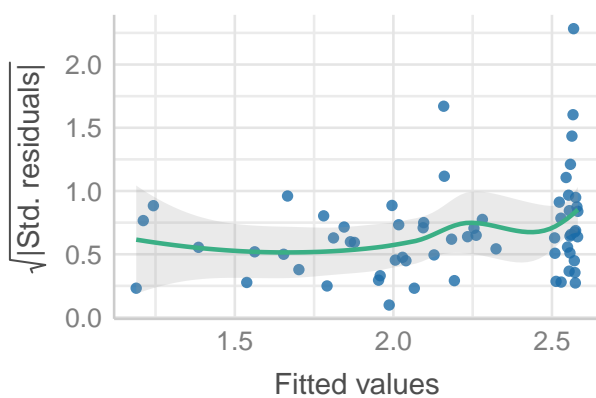
### Linearity

Reference line should be flat and horizontal



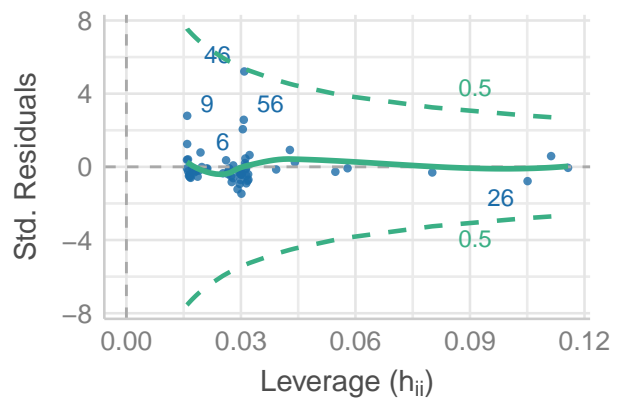
### Homogeneity of Variance

Reference line should be flat and horizontal



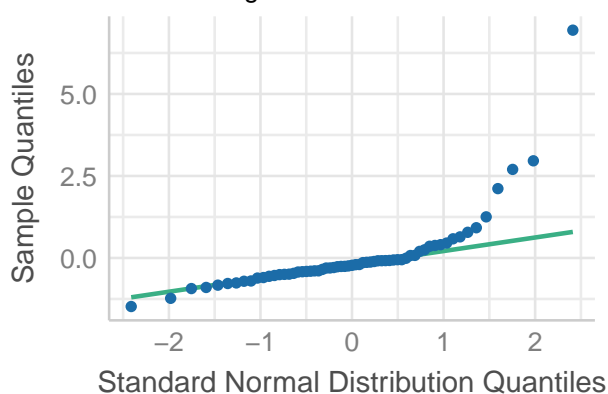
### Influential Observations

Points should be inside the contour lines



### Normality of Residuals

Dots should fall along the line

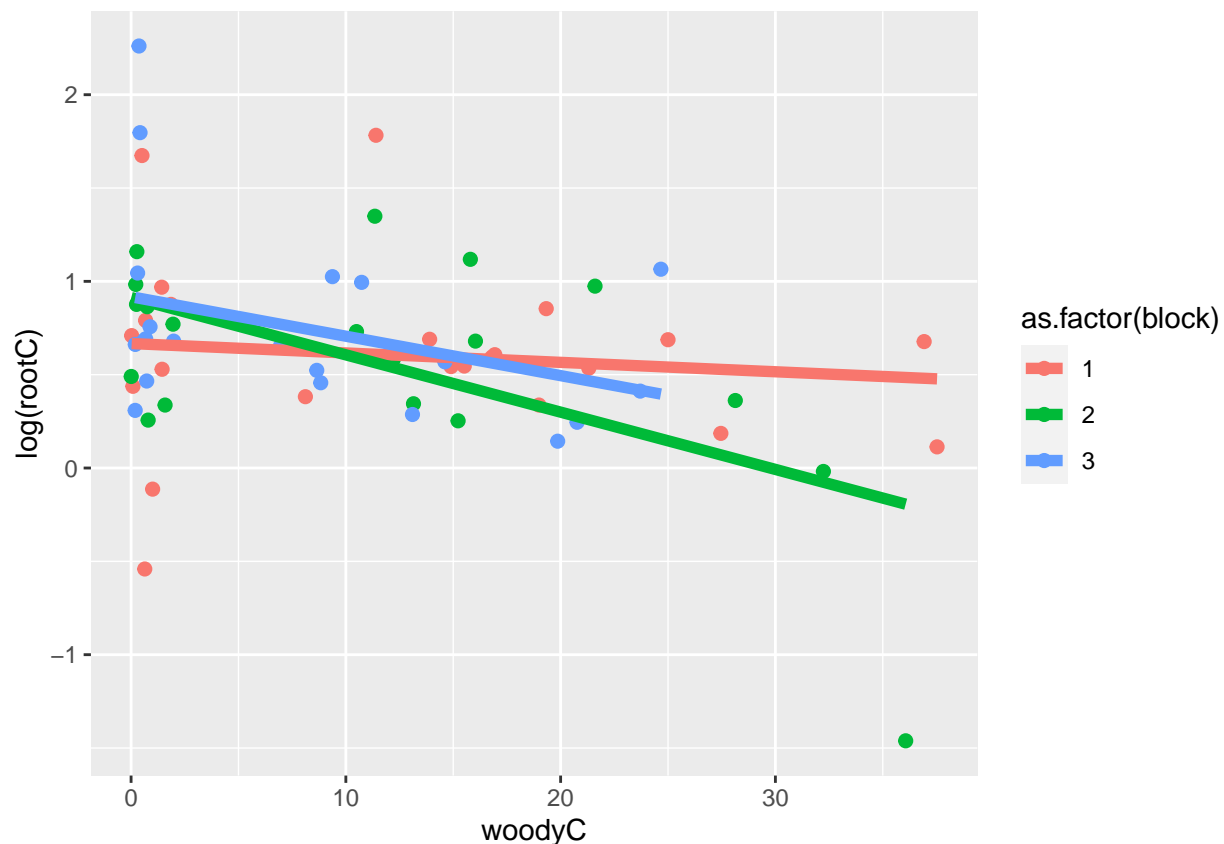


Due to the highly skewed distribution of root C, this model evidently violates the normality of residuals assumption of OLS. Although we have good reasons to want to keep root C in its original units, we could try the model with log-transformed root C as a robustness check:

```
## note that a block random intercept results in a singular fit
summary(lm(log(rootC)~woodyC,data=carbon_seq))
```

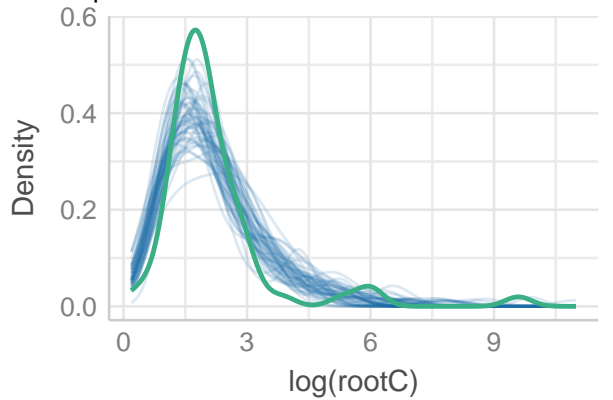
```
##
## Call:
## lm(formula = log(rootC) ~ woodyC, data = carbon_seq)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.64577 -0.23411 -0.02634  0.15896  1.43547
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.831995   0.089143   9.333 2.33e-13 ***
## woodyC       -0.017959   0.005882  -3.053  0.00335 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4985 on 61 degrees of freedom
## (77 observations deleted due to missingness)
## Multiple R-squared:  0.1325, Adjusted R-squared:  0.1183
## F-statistic: 9.321 on 1 and 61 DF,  p-value: 0.003354
```

```
ggplot(data=carbon_seq,aes(x=woodyC,y=log(rootC),color=as.factor(block)))+
  geom_point(size=2)+geom_smooth(method="lm",se=F,linewidth=2)
```



### Posterior Predictive Check

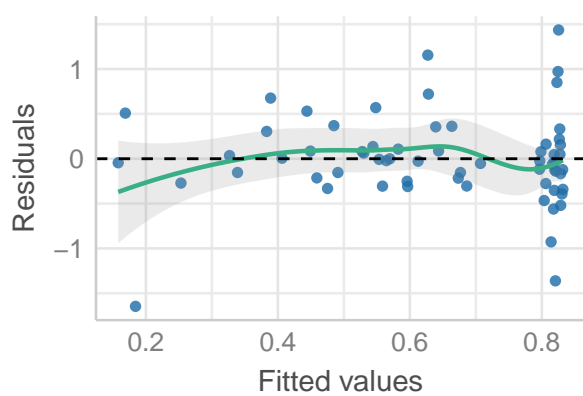
Model-predicted lines should resemble observed data



— Observed data — Model-predicted data

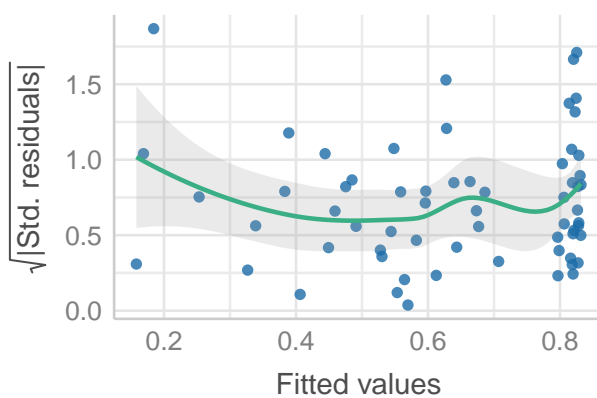
### Linearity

Reference line should be flat and horizontal



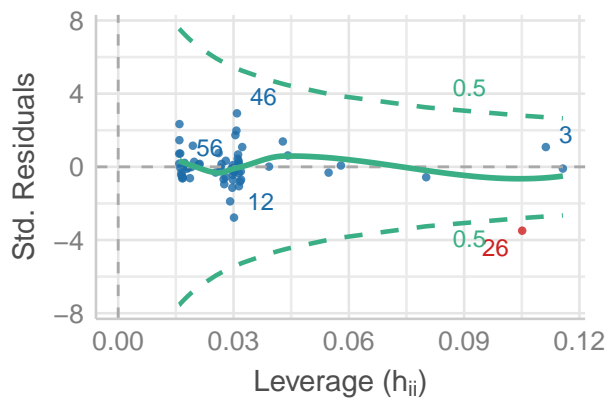
### Homogeneity of Variance

Reference line should be flat and horizontal



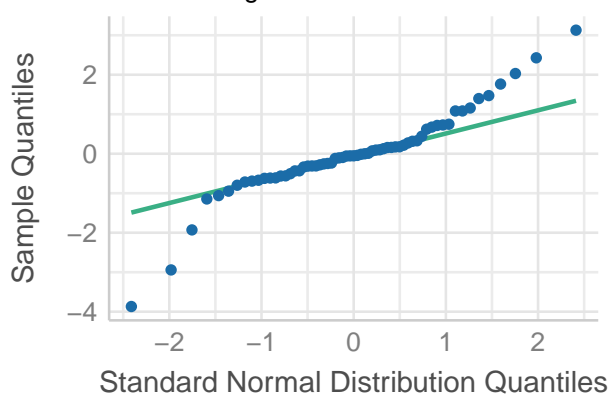
### Influential Observations

Points should be inside the contour lines



### Normality of Residuals

Dots should fall along the line



Although the functional form of the relationship now evidently differs, the coefficient is still negative and the model does somewhat better at meeting the assumptions.



## Changes in soil C in relation to other soil characteristics

A lot of recent work in ecosystem ecology links the accumulation of stable soil C to the formation of macroaggregates (250  $\mu\text{m}$  diameter and up). We could first ask: does the percentage of macroaggregates (in 2019) correlate with soil carbon in 2013 (pre-treatment) or 2019?

```
## note that these relationships would be significant in the absence  
## of the block random intercept
```

```
summary(lmer(soilC_2013~macro250+(1|block),data=carbon_seq))
```

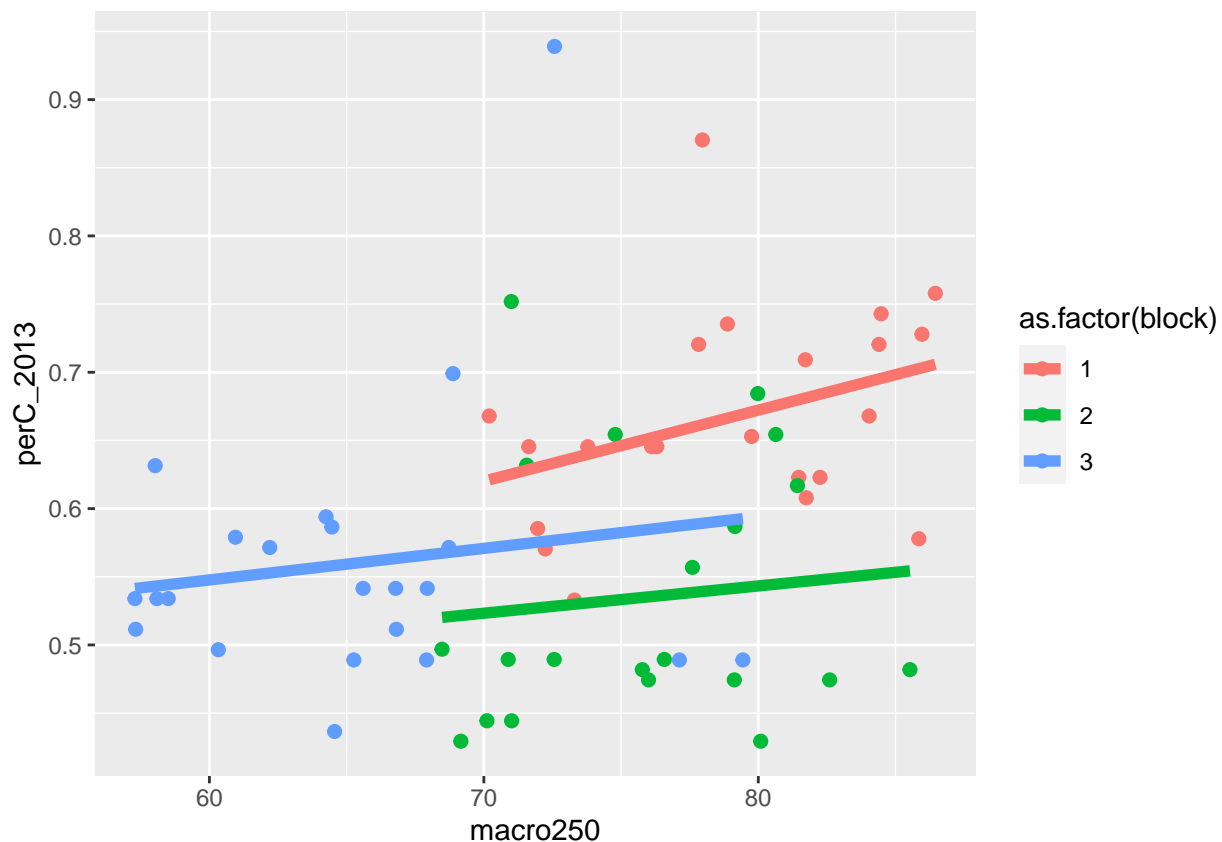
```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [  
## lmerModLmerTest]  
## Formula: soilC_2013 ~ macro250 + (1 | block)  
## Data: carbon_seq  
##  
## REML criterion at convergence: 312.1  
##  
## Scaled residuals:  
##      Min       1Q   Median       3Q      Max   
## -1.4538 -0.6916 -0.0738  0.4413  4.0100   
##  
## Random effects:  
## Groups   Name            Variance Std.Dev.   
## block    (Intercept)  0.8442     0.9188   
## Residual                    6.5752     2.5642   
## Number of obs: 65, groups:  block, 3  
##  
## Fixed effects:  
##              Estimate Std. Error    df t value Pr(>|t|)      
## (Intercept) 10.15177    3.85797 21.17731  2.631  0.0155 *   
## macro250     0.08554    0.05192 23.59887  1.647  0.1127   
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Correlation of Fixed Effects:  
##              (Intr)   
## macro250 -0.987
```

```
summary(lmer(perC_2013~macro250+(1|block),data=carbon_seq))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [  
## lmerModLmerTest]  
## Formula: perC_2013 ~ macro250 + (1 | block)  
## Data: carbon_seq  
##  
## REML criterion at convergence: -107.2  
##  
## Scaled residuals:  
##      Min       1Q   Median       3Q      Max   
## -1.3874 -0.6890 -0.1123  0.4286  3.9079   
##  
## Random effects:  
## Groups   Name            Variance Std.Dev.   
## block    (Intercept)  0.003521 0.05934   
## Residual                    0.008226 0.09070   
## Number of obs: 65, groups:  block, 3
```

```
##
## Fixed effects:
##           Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)  0.341085   0.151179 40.449103   2.256   0.0295 *
## macro250     0.003357   0.002002 52.194462   1.677   0.0995 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr)
## macro250 -0.971
```

```
ggplot(data=carbon_seq,aes(x=macro250,y=perC_2013,color=as.factor(block)))+
  geom_point(size=2)+geom_smooth(method="lm",se=F,linewidth=2)
```



```
## note that these relationships would be significant/marginal in the absence
## of the block random intercept
summary(lmer(soilC_2019~macro250+(1|block),data=carbon_seq))
```

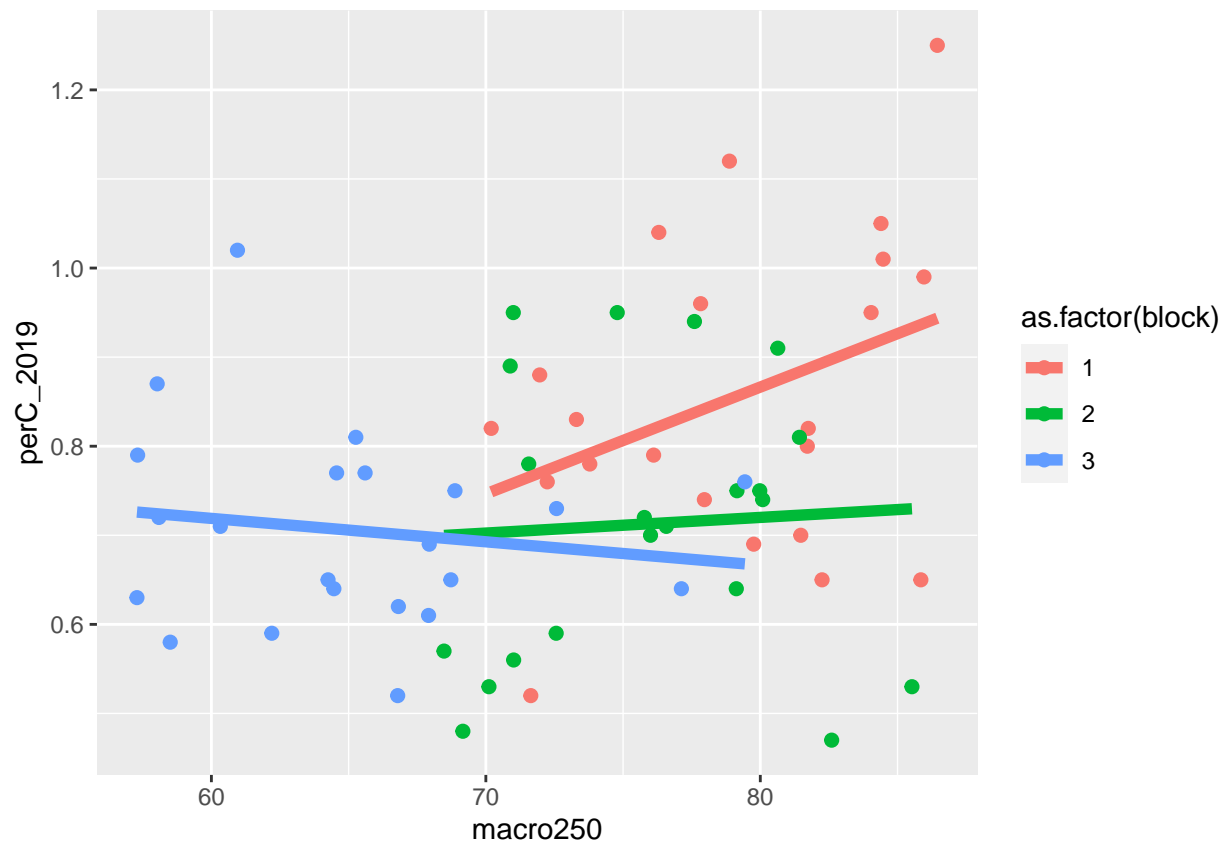
```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: soilC_2019 ~ macro250 + (1 | block)
## Data: carbon_seq
##
## REML criterion at convergence: 371.9
##
## Scaled residuals:
```

```

##      Min      1Q  Median      3Q      Max
## -2.0382 -0.7215 -0.1137  0.6404  2.4799
##
## Random effects:
##   Groups   Name                Variance Std.Dev.
##   block    (Intercept)  0.1458  0.3818
##   Residual                    17.5092  4.1844
## Number of obs: 65, groups:  block, 3
##
## Fixed effects:
##              Estimate Std. Error      df t value Pr(>|t|)
## (Intercept) 12.61086    5.01076  5.50575   2.517   0.049 *
## macro250     0.11715    0.06791  5.64541   1.725   0.138
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr)
## macro250 -0.994
summary(lmer(perC_2019~macro250+(1|block),data=carbon_seq))

## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: perC_2019 ~ macro250 + (1 | block)
##   Data: carbon_seq
##
## REML criterion at convergence: -44.8
##
## Scaled residuals:
##      Min      1Q  Median      3Q      Max
## -1.91549 -0.67435 -0.06883  0.59706  2.52338
##
## Random effects:
##   Groups   Name                Variance Std.Dev.
##   block    (Intercept)  0.003851  0.06206
##   Residual                    0.022650  0.15050
## Number of obs: 65, groups:  block, 3
##
## Fixed effects:
##              Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)  0.451005    0.232642 24.681756   1.939   0.0641 .
## macro250     0.004183    0.003124 28.437092   1.339   0.1912
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr)
## macro250 -0.985
ggplot(data=carbon_seq,aes(x=macro250,y=perC_2019,color=as.factor(block)))+
  geom_point(size=2)+geom_smooth(method="lm",se=F,linewidth=2)

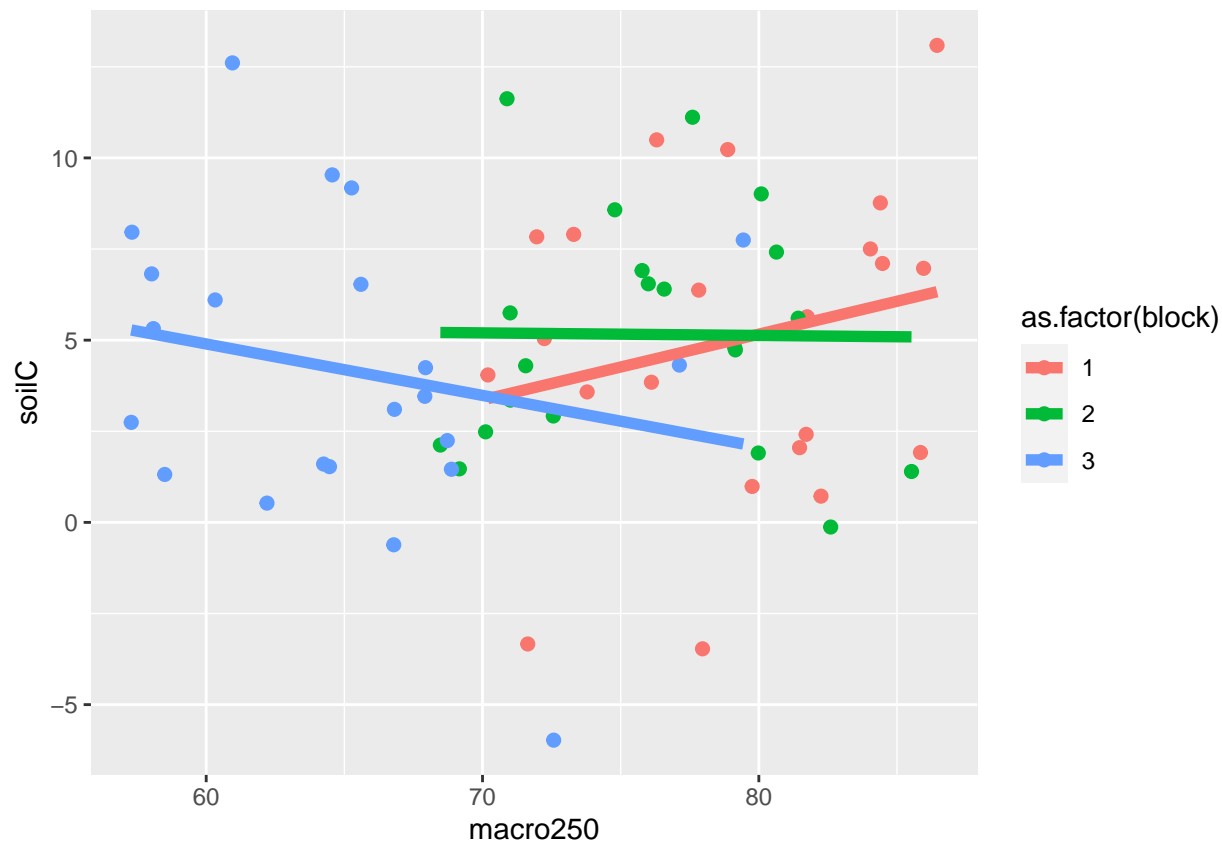
```



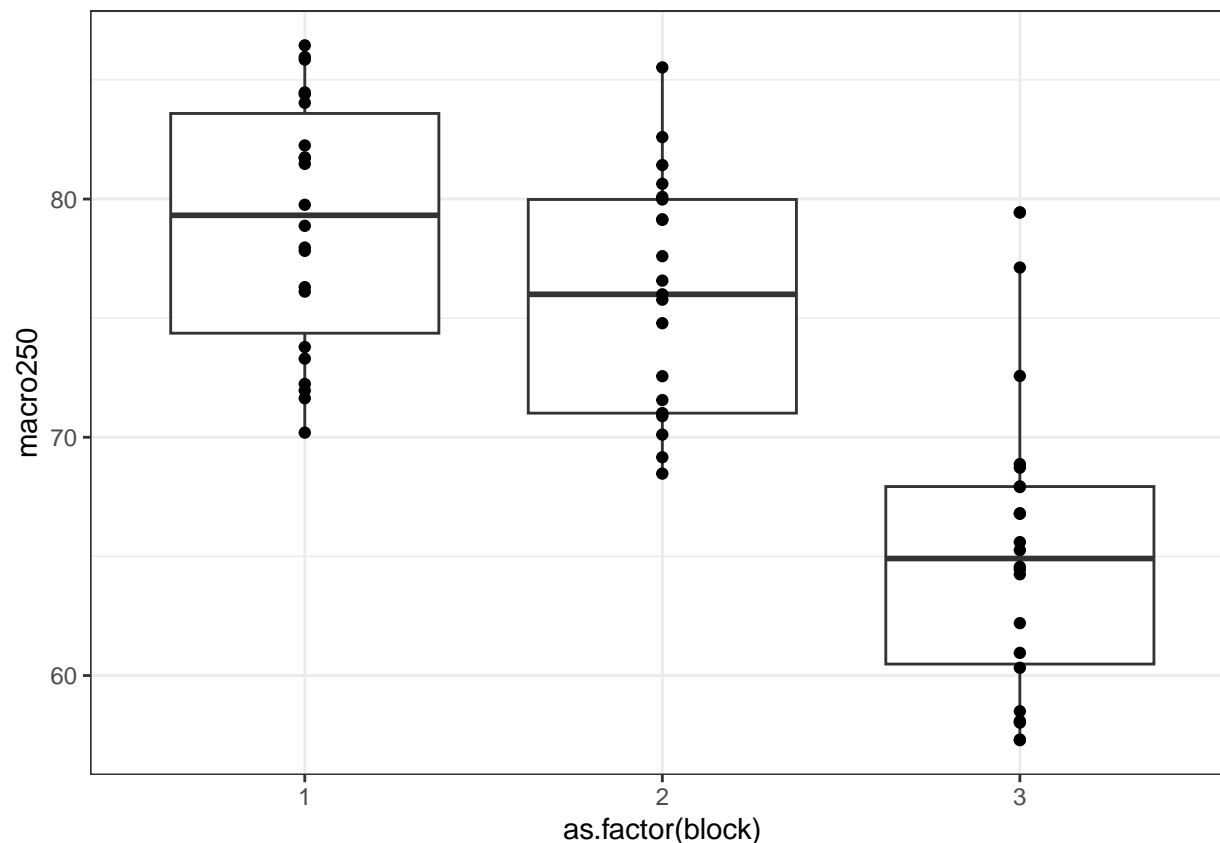
```
## when looking at change in soil C, the block effect reaches a singular fit
summary(lm(soilC~macro250,data=carbon_seq))
```

```
##
## Call:
## lm(formula = soilC ~ macro250, data = carbon_seq)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.7145  -2.8438  -0.2396   2.3915   8.2794
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.16897    4.40823   0.492   0.624
## macro250     0.03542    0.05977   0.593   0.556
##
## Residual standard error: 3.854 on 63 degrees of freedom
## (75 observations deleted due to missingness)
## Multiple R-squared:  0.005543,    Adjusted R-squared:  -0.01024
## F-statistic: 0.3511 on 1 and 63 DF,  p-value: 0.5556
```

```
ggplot(data=carbon_seq,aes(x=macro250,y=soilC,color=as.factor(block)))+
  geom_point(size=2)+geom_smooth(method="lm",se=F,linewidth=2)
```



These plots offer scant evidence that macroaggregates have much relation to soil C. To the extent that there appears to be a relationship across the full dataset, it appears largely driven by the fact that Block 1 has higher macroaggregates than Block 3 in particular, and greater initial and final percent C than Blocks 2 and 3. (These differences in percent C are largely offset by lower bulk density when calculating the top 20 cm C pool, such that there remain no post-treatment block differences in soil C pools or soil.) This block-level variation might prompt us to look more closely at macroaggregates across blocks:



Indeed, Block 3 has a much lower macroaggregate fraction than the other two.

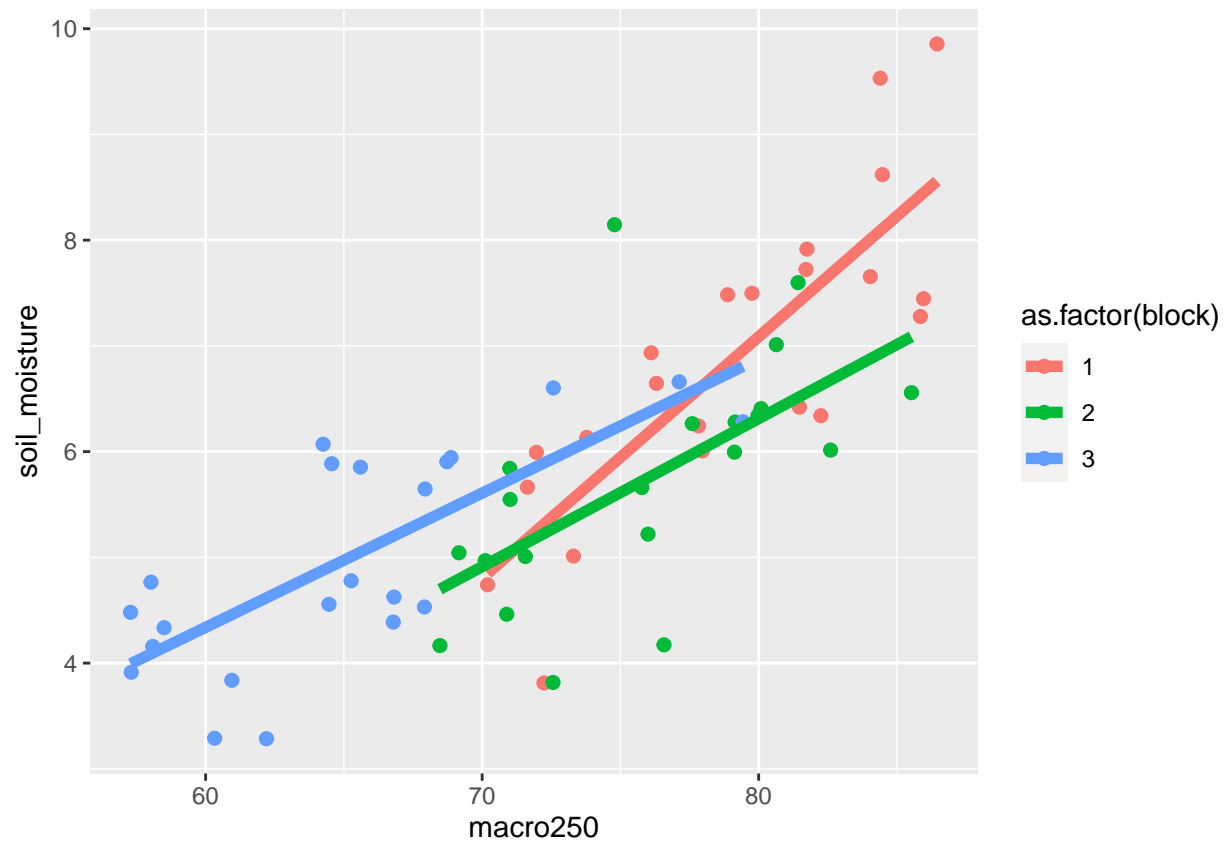
Macroaggregates are commonly associated with the ability to retain moisture, so we can test the relationship between macroaggregates and soil moisture.

```
summary(lmer(soil_moisture~macro250+(1|block),data=carbon_seq))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: soil_moisture ~ macro250 + (1 | block)
## Data: carbon_seq
##
## REML criterion at convergence: 174.4
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.32277 -0.71385 -0.03953  0.58157  2.94246
##
## Random effects:
## Groups   Name                Variance Std.Dev.
## block    (Intercept)  0.2087    0.4569
## Residual                    0.7259    0.8520
## Number of obs: 65, groups:  block, 3
##
## Fixed effects:
##              Estimate Std. Error    df t value Pr(>|t|)
## (Intercept) -5.54921    1.37804 33.53552  -4.027 0.000305 ***
```

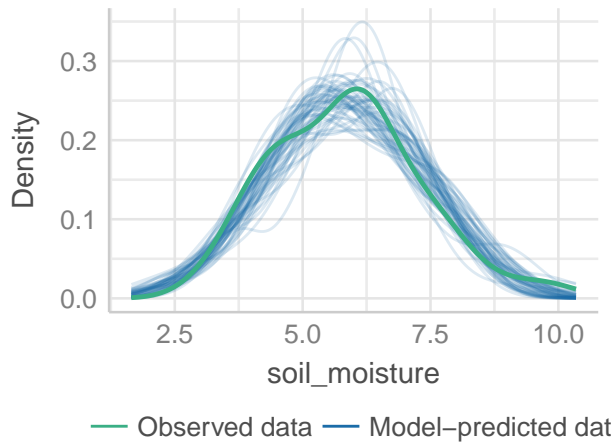
```
## macro250      0.15561      0.01838 41.56302      8.464 1.38e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr)
## macro250 -0.979
```

```
ggplot(data=carbon_seq,aes(x=macro250,y=soil_moisture,color=as.factor(block)))+
  geom_point(size=2)+geom_smooth(method="lm",se=F,linewidth=2)
```



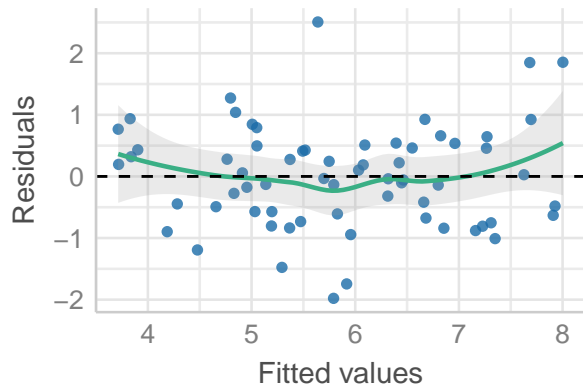
### Posterior Predictive Check

Model-predicted lines should resemble observed data



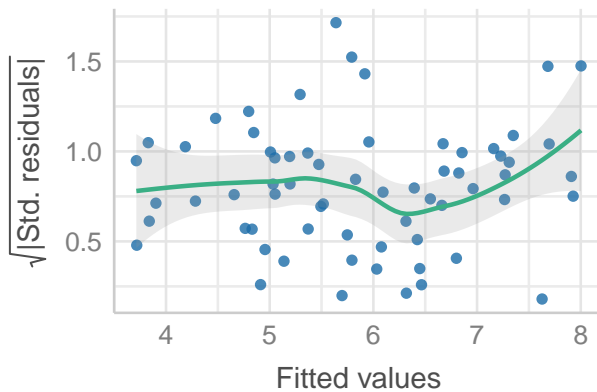
### Linearity

Reference line should be flat and horizontal



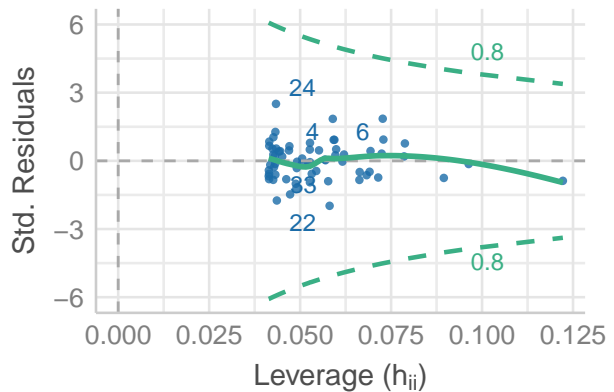
### Homogeneity of Variance

Reference line should be flat and horizontal



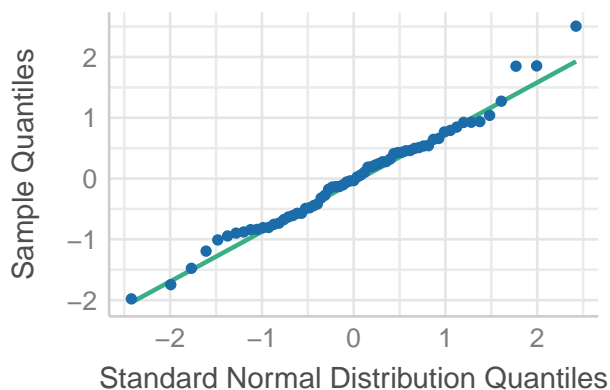
### Influential Observations

Points should be inside the contour lines



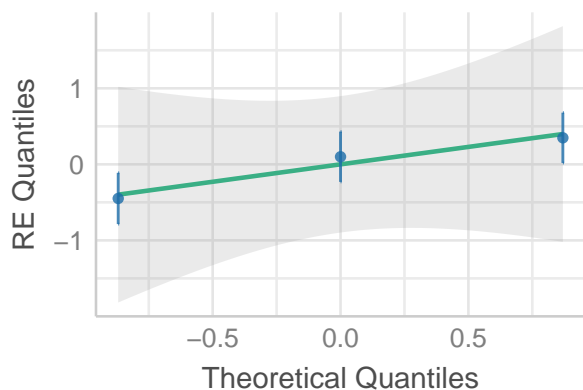
### Normality of Residuals

Dots should fall along the line



### Normality of Random Effects (block)

Dots should be plotted along the line



There is evidently a close relationship between soil moisture and macroaggregate fraction. What else is related to soil moisture? Although soil C and macroaggregates are not related, it turns out that soil C and soil moisture are, based on pools (or percent C) in both 2013 and 2019:



```
## this also yields positive relationships using initial and 2019 percent C
summary(lmer(soil_moisture~soilC_2013+(1|block),data=carbon_seq))
```

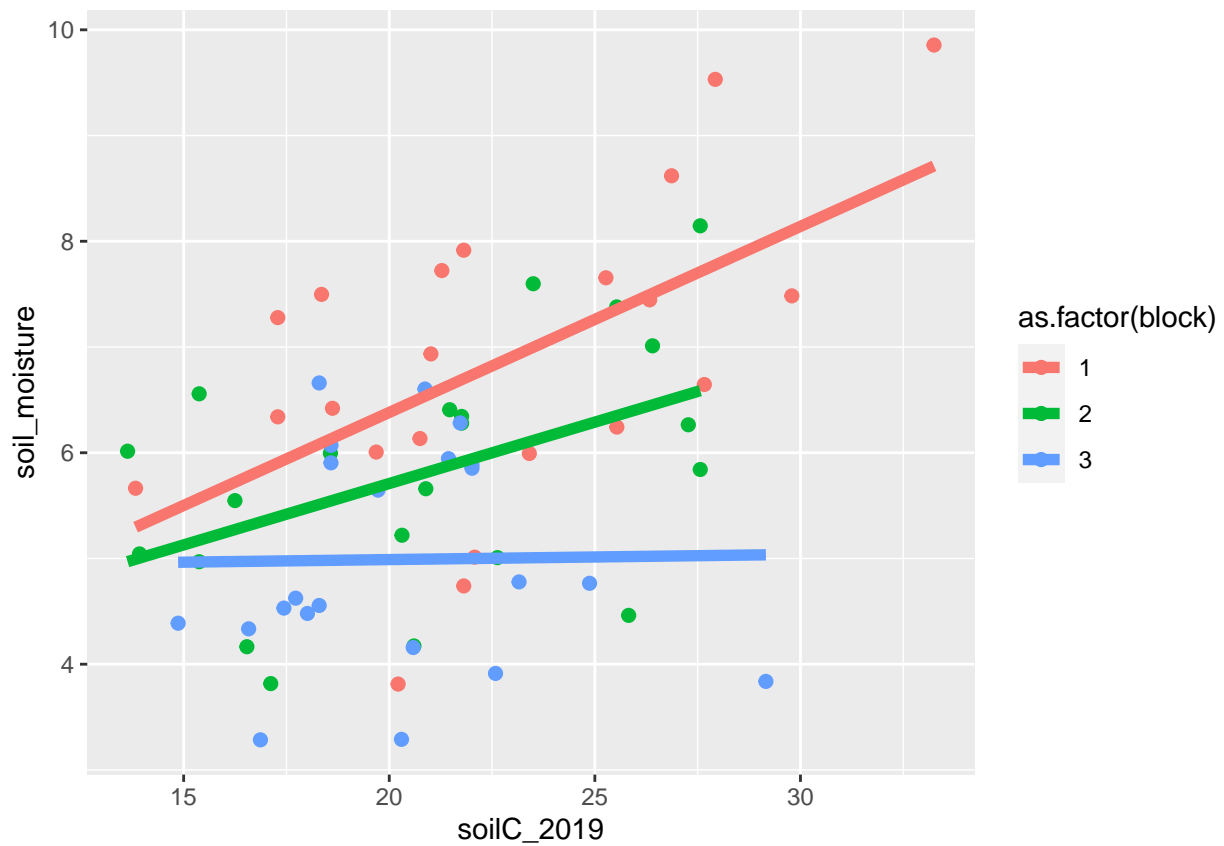
```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: soil_moisture ~ soilC_2013 + (1 | block)
## Data: carbon_seq
##
## REML criterion at convergence: 215.3
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.22554 -0.64940 -0.05845  0.67247  2.27265
##
## Random effects:
## Groups Name Variance Std.Dev.
## block (Intercept) 0.561 0.749
## Residual 1.276 1.130
## Number of obs: 67, groups: block, 3
##
## Fixed effects:
## Estimate Std. Error df t value Pr(>|t|)
## (Intercept) 2.72074 1.00632 30.59527 2.704 0.01109 *
## soilC_2013 0.19311 0.05461 64.23046 3.536 0.00076 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
## (Intr)
## soilC_2013 -0.892
```

```
summary(lmer(soil_moisture~soilC_2019+(1|block),data=carbon_seq))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: soil_moisture ~ soilC_2019 + (1 | block)
## Data: carbon_seq
##
## REML criterion at convergence: 213.1
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.35767 -0.69484  0.02931  0.81836  1.86267
##
## Random effects:
## Groups Name Variance Std.Dev.
## block (Intercept) 0.549 0.741
## Residual 1.276 1.130
## Number of obs: 66, groups: block, 3
##
## Fixed effects:
## Estimate Std. Error df t value Pr(>|t|)
## (Intercept) 3.2641 0.8480 20.1585 3.849 0.000989 ***
## soilC_2019 0.1234 0.0338 62.7304 3.651 0.000534 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr)
## soilC_2019 -0.848

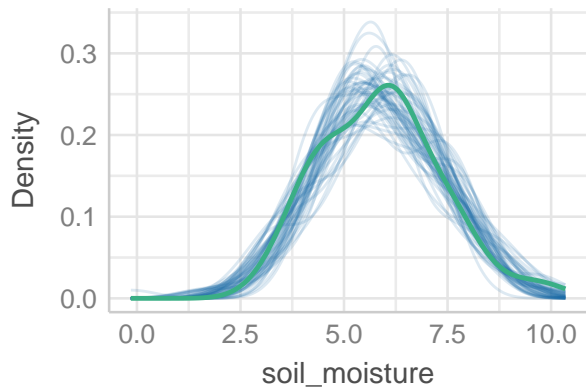
ggplot(data=carbon_seq,aes(x=soilC_2019,y=soil_moisture,color=as.factor(block)))+
  geom_point(size=2)+geom_smooth(method="lm",se=F,linewidth=2)
```



```
check_model(lmer(soil_moisture~soilC_2019+(1|block),data=carbon_seq))
```

### Posterior Predictive Check

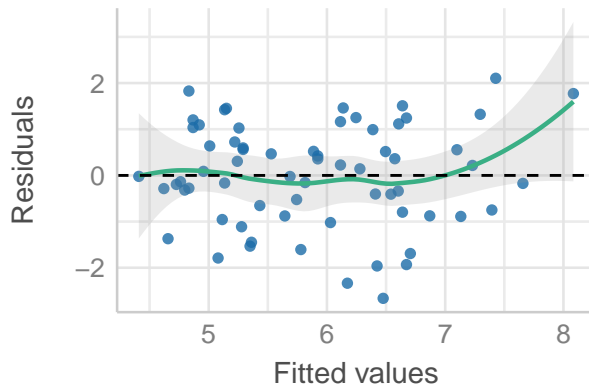
Model-predicted lines should resemble observed data



— Observed data — Model-predicted data

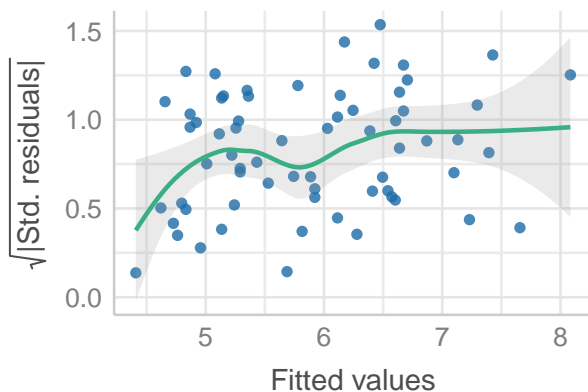
### Linearity

Reference line should be flat and horizontal



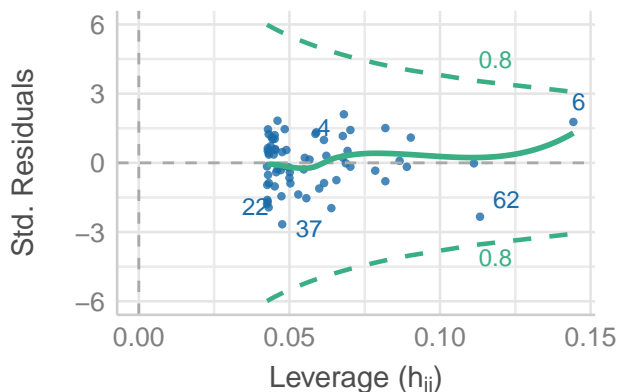
### Homogeneity of Variance

Reference line should be flat and horizontal



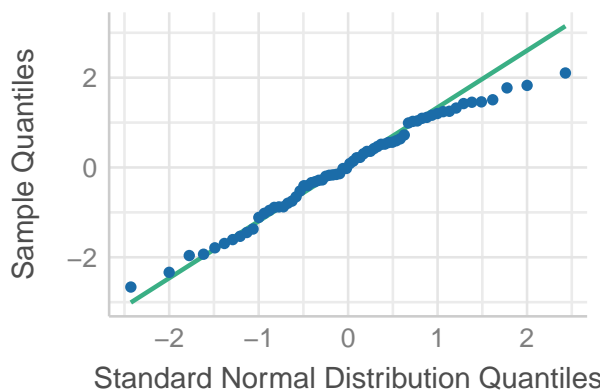
### Influential Observations

Points should be inside the contour lines



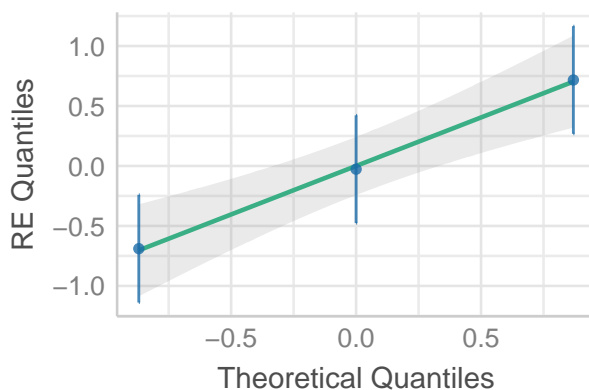
### Normality of Residuals

Dots should fall along the line



### Normality of Random Effects (block)

Dots should be plotted along the line



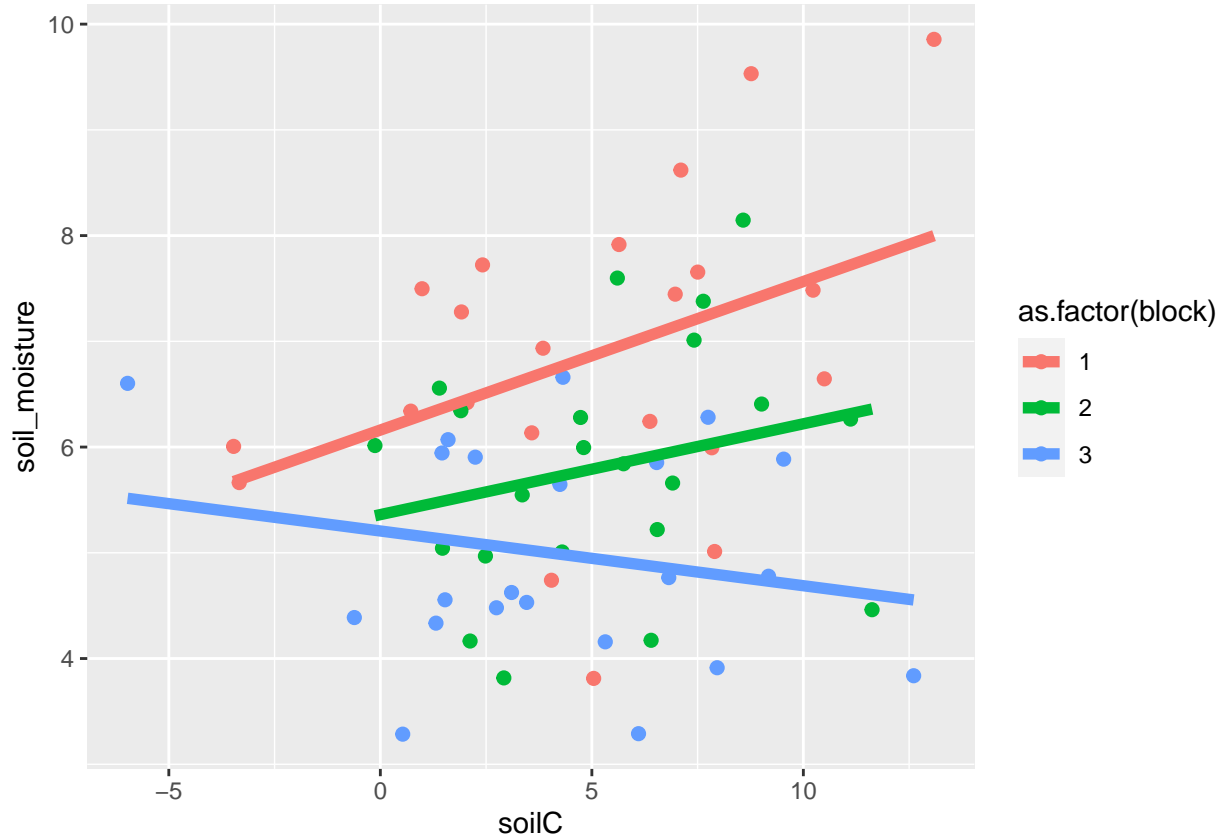
However, there is no relationship between soil moisture and soil C accumulation.

```
## this also yields positive relationships using initial and 2019 percent C
summary(lmer(soil_moisture~soilC+(1|block),data=carbon_seq))
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
```

```
## lmerModLmerTest]
## Formula: soil_moisture ~ soilC + (1 | block)
## Data: carbon_seq
##
## REML criterion at convergence: 222.9
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.4432 -0.5191 -0.1461  0.7314  2.1353
##
## Random effects:
## Groups Name Variance Std.Dev.
## block (Intercept) 0.7722  0.8787
## Residual 1.4829  1.2178
## Number of obs: 66, groups: block, 3
##
## Fixed effects:
##              Estimate Std. Error    df t value Pr(>|t|)
## (Intercept)  5.60863    0.56260  2.55339   9.969  0.00413 **
## soilC        0.05821    0.03981  62.15045   1.462  0.14871
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr)
## soilC -0.340
```

```
ggplot(data=carbon_seq,aes(x=soilC,y=soil_moisture,color=as.factor(block)))+
  geom_point(size=2)+geom_smooth(method="lm",se=F,linewidth=2)
```



We also measured pH, which is not correlated to soil moisture but is correlated with macroaggregates.

## Carbon accumulation and biodiversity facets

One of the core goals of our paper is to show how carbon storage in our experimental plots was influenced by the diversity of planted trees in the first six years of the experiment. We estimated carbon storage across three pools:

- aboveground wood
- soil (0-20 cm)–expressed as change from pre-treatment soil carbon
- fine roots (0-20 cm)

Besides these pools, we also calculated overyielding for mixture plots as the difference between the observed pools and monoculture-based expectations. For aboveground wood, we further partitioned overyielding into complementarity and selection effects following Loreau & Hector (2001).

We considered three facets of diversity:

- species richness
- Laliberté and Legendre (2010)’s functional dispersion (FDis), as reported by Grossman et al. (2017)
- Helmus et al. (2007)’s phylogenetic species variability (PSV), as reported by Grossman et al. (2017)

The latter two are metrics of functional and phylogenetic diversity that are designed not to be dependent on species richness. However, FDis and PSV are both 0 for monocultures–by default for FDis and as an analysis choice for PSV. As a result, they do end up somewhat correlated with each other and with species richness. As we will see, multicollinearity does not end up being a severe problem in most cases.

In keeping with the dominant paradigm in biodiversity-ecosystem function (BEF) research, we proposed that planted tree diversity would positively influence all three carbon pools. However, we did not pose hypotheses

about which specific facets of tree diversity would matter most. As a result, we used an AICc-based framework to test all models nested within a model that includes species richness, FDis, and PSV. We select the most parsimonious model with a  $\Delta AICc < 2$  from the model with the lowest AICc. (In effect, this is much like standard null hypothesis significance testing, where  $AICc < 2$  corresponds to  $p > 0.05$ .) When there are two or more equally parsimonious such models, we selected the one with the lower AICc.

## Aboveground carbon

We can start with aboveground woody carbon:

```
mwt<-lm(woodyC~species_richness+FDis+PSV,data=carbon_seq)
mwt_block<-lmer(woodyC~species_richness+FDis+PSV+(1|block),
               data=carbon_seq, REML=F)

options(na.action = "na.fail") # required for dredge to run
mwt_dredge <- dredge(mwt,beta = "none",evaluate = T, rank = AICc)
mwt_block_dredge <- dredge(mwt_block,beta = "none",evaluate = T,rank = AICc)
options(na.action = "na.omit")

## instead of selecting a 'best' model, one could instead perform model averaging
# summary(model.avg(mwt_dredge, subset = delta <= 2))
```

The `dredge` function tests all nested submodels of the specified model. Here we apply it to models both with and without random intercepts for block. There are three blocks in the FAB1 experiment. The block effect tends to either yield a singular fit or negligible changes in model parameters—the exceptions being in analyses of macroaggregates or soil moisture. As a result, we generally do not show the output of models with the block random effect.

We can examine our AIC table, select our ‘best’ model, and do some basic checks of model assumptions as follows:

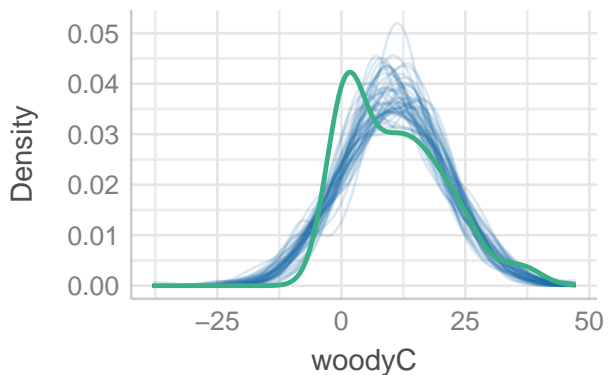
```
mwt_dredge ## we'll select the second model

## Global model call: lm(formula = woodyC ~ species_richness + FDis + PSV, data = carbon_seq)
## ---
## Model selection table
##      (Int)    FDs    PSV  spc_rch df   logLik   AICc delta weight
## 8  7.941 -5.549 20.870  0.62920  5 -499.261 1009.0  0.00  0.704
## 4  8.494 -4.251 19.140           4 -501.209 1010.7  1.74  0.294
## 3  6.563          9.453           3 -507.897 1022.0 13.00  0.001
## 7  6.662          9.560 -0.05419  4 -507.880 1024.1 15.09  0.000
## 2  8.863  1.263           3 -517.217 1040.6 31.64  0.000
## 1 10.690           2 -518.636 1041.4 32.39  0.000
## 5  9.826           0.32500  3 -518.053 1042.3 33.31  0.000
## 6  8.788  1.147           0.09113  4 -517.182 1042.7 33.69  0.000
## Models ranked by AICc(x)

mwt_model<-get.models(mwt_dredge, subset = 2)[[1]]
check_model(mwt_model)
```

### Posterior Predictive Check

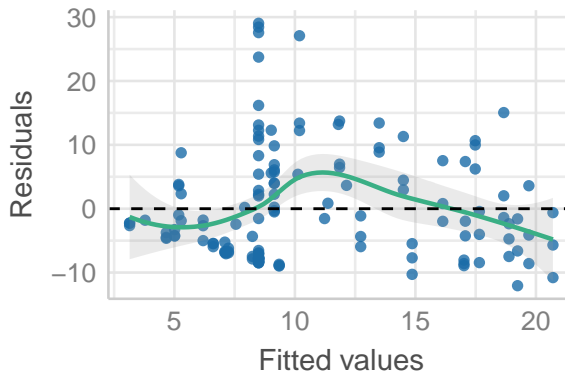
Model-predicted lines should resemble observed data



— Observed data — Model-predicted data

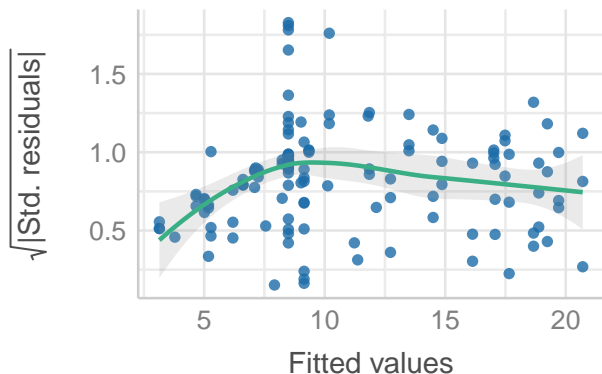
### Linearity

Reference line should be flat and horizontal



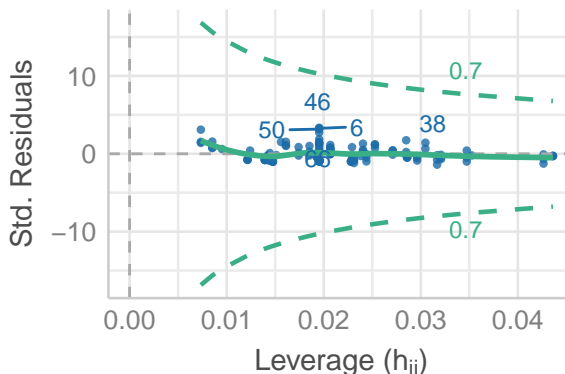
### Homogeneity of Variance

Reference line should be flat and horizontal



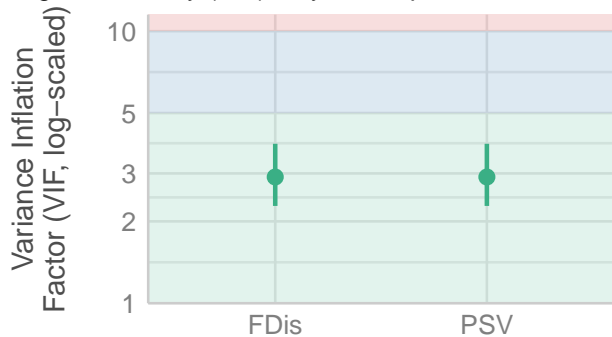
### Influential Observations

Points should be inside the contour lines



### Collinearity

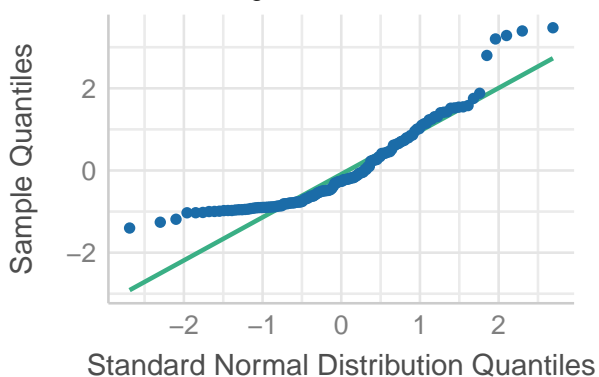
High collinearity (VIF) may inflate parameter uncertainty



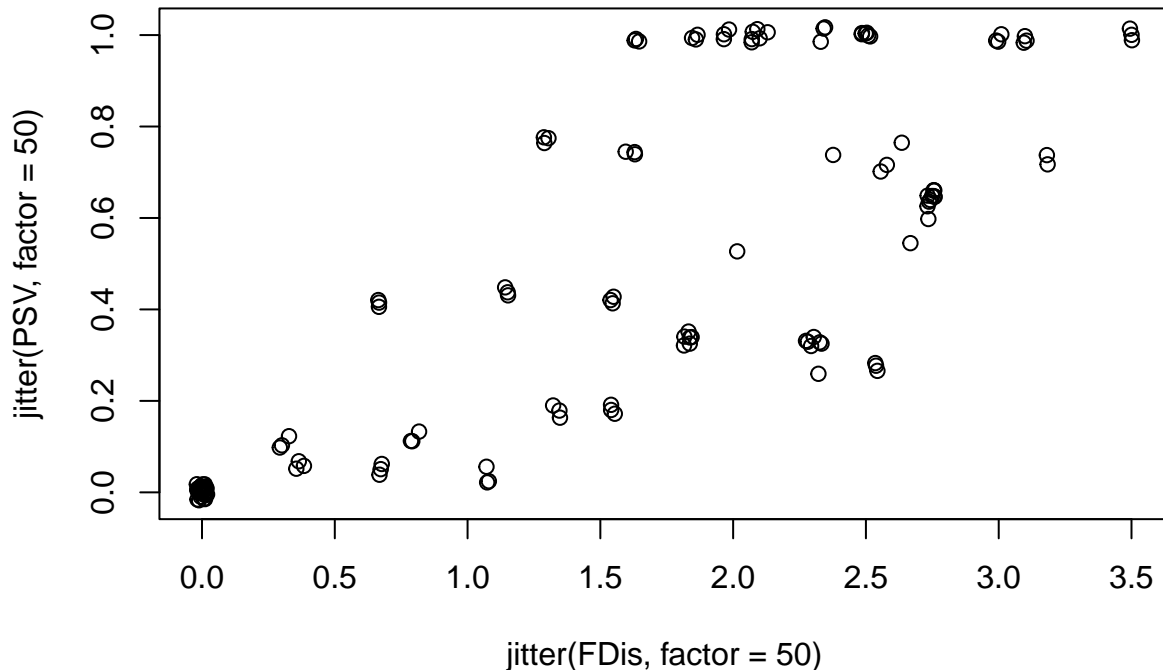
● Low (< 5)

### Normality of Residuals

Dots should fall along the line



A few things are apparent here. First of all, the two retained predictors (FDis and PSV) have opposite-sign coefficients. These predictors have a positive correlation with each other across the full dataset, as shown below:



Therefore, their “effects” partly offset each other. However, their variance inflation factor (VIF) is not so high that it would ordinarily raise concerns about multicollinearity.

Perhaps more troubling are the other checks: the posterior predictive check plot, the homoskedasticity plot, and the Q-Q (normality of residuals) plot. It is common for data from BEF experiments not to strictly meet these assumptions of ordinary least-squares regression because the variance is often greatest among monocultures. To assure ourselves that our results are robust, we can use robust regression through iterated re-weighted least squares to reduce the influence of high-leverage points.

```
mwt_robust<-rlm(woodyC~species_richness+FDis+PSV,data=carbon_seq)
```

```
options(na.action = "na.fail")
```

```
mwt_robust_dredge <- dredge(mwt_robust,beta = "none",evaluate = T,rank = AICc)
```

```
options(na.action = "na.omit")
```

```
mwt_robust_dredge
```

```
## Global model call: rlm(formula = woodyC ~ species_richness + FDis + PSV, data = carbon_seq)
```

```
## ---
```

```
## Model selection table
```

##	(Int)	FDs	PSV	spc_rch	df	logLik	AICc	delta	weight
## 8	5.734	-4.794	20.85	0.6500	5	-501.023	1012.5	0.00	0.571
## 4	6.656	-3.558	19.15		4	-502.384	1013.1	0.57	0.429
## 3	3.924		11.82		3	-510.742	1027.7	15.17	0.000
## 7	3.564		11.57	0.1425	4	-511.200	1030.7	18.20	0.000
## 1	10.380				2	-518.709	1041.5	29.01	0.000
## 2	7.361	1.794			3	-517.864	1041.9	29.41	0.000



```
## 5  9.306                0.3876  3 -518.165 1042.5 30.01  0.000
## 6  7.285  1.643          0.1125  4 -517.818 1043.9 31.44  0.000
## Models ranked by AICc(x)
```

Here, the ‘best’ model selected involves the same variables with similar (but not exactly the same) parameter values to when we were using OLS.

## Overyielding in aboveground carbon

Let’s repeat the same procedure for overyielding in aboveground carbon.

```
## for the dredge function to work, we need to remove plots with NAs
carbon_seq_woodyOY<-carbon_seq[which(!is.na(carbon_seq$woodyOY)),]
mw<-lm(woodyOY~species_richness+FDis+PSV,
      data=carbon_seq_woodyOY)
## as before, the posterior predictive check and Q-Q plot
## for the OLS model reveal violations of assumptions
mw_robust<-rlm(woodyOY~species_richness+FDis+PSV,data=carbon_seq_woodyOY)

options(na.action = "na.fail")
mw_dredge <- dredge(mw, beta = "none", evaluate = T, rank = AICc)
mw_robust_dredge <- dredge(mw_robust, beta = "none", evaluate = T, rank = AICc)
options(na.action = "na.omit")

mw_dredge ## we select the first model
```

```
## Global model call: lm(formula = woodyOY ~ species_richness + FDis + PSV, data = carbon_seq_woodyOY)
## ---
## Model selection table
##      (Int)    FDs    PSV spc_rch df   logLik  AICc delta weight
## 6 -2.1370 1.2200          0.2501  4 -269.594 547.6  0.00  0.450
## 8 -2.2040 0.9463 0.897  0.2729  5 -269.329 549.3  1.68  0.194
## 7 -1.4580          2.262  0.3643  4 -270.706 549.8  2.22  0.148
## 2 -2.0320 1.5830          3 -271.887 550.0  2.42  0.134
## 4 -2.0420 1.5400 0.159          4 -271.879 552.2  4.57  0.046
## 5 -0.1805          0.3808  3 -273.559 553.4  5.77  0.025
## 3 -0.3916          2.463          3 -276.153 558.5 10.95  0.002
## 1  1.0570          2 -279.205 562.5 14.94  0.000
## Models ranked by AICc(x)
```

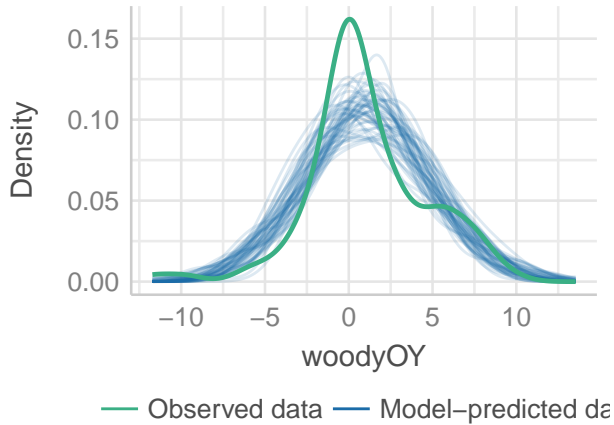
```
mw_robust_dredge ## we select the first model
```

```
## Global model call: rlm(formula = woodyOY ~ species_richness + FDis + PSV, data = carbon_seq_woodyOY)
## ---
## Model selection table
##      (Int)    FDs    PSV spc_rch df   logLik  AICc delta weight
## 6 -1.9010 1.0570          0.2476  4 -269.724 547.9  0.00  0.454
## 8 -1.9020 0.8079 0.7843  0.2619  5 -269.472 549.6  1.70  0.194
## 7 -1.2420          2.0290  0.3304  4 -270.797 550.0  2.15  0.155
## 2 -1.7270 1.3700          3 -272.086 550.4  2.56  0.126
## 4 -1.7240 1.3200 0.1664          4 -272.082 552.6  4.71  0.043
## 5 -0.2588          0.3588  3 -273.683 553.6  5.75  0.026
## 3 -0.3801          2.3410          3 -276.176 558.6 10.74  0.002
## 1  0.9372          2 -279.264 562.6 14.79  0.000
## Models ranked by AICc(x)
```

```
mw_model <- get.models(mw_dredge, subset = 1)[[1]]
check_model(mw_model)
```

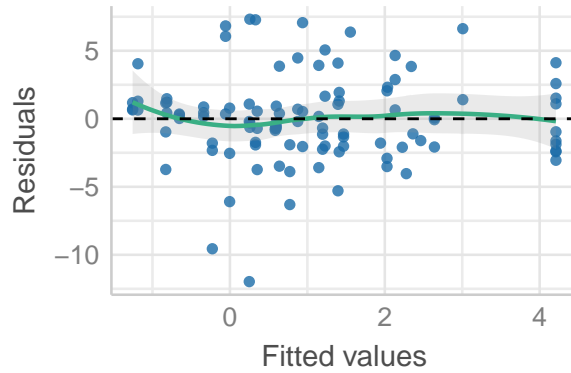
### Posterior Predictive Check

Model-predicted lines should resemble observed data



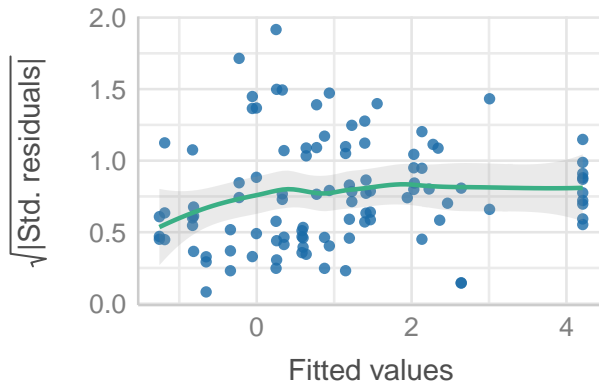
### Linearity

Reference line should be flat and horizontal



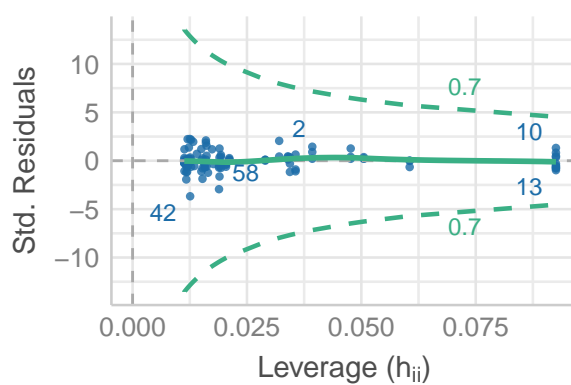
### Homogeneity of Variance

Reference line should be flat and horizontal



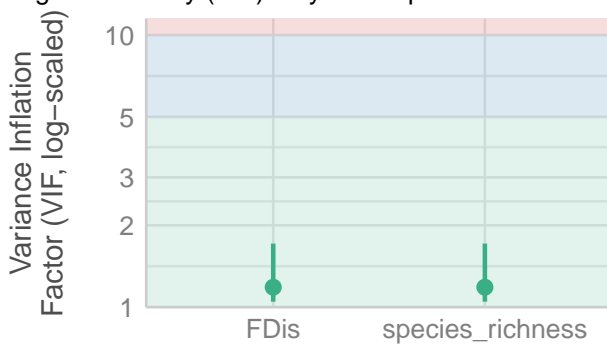
### Influential Observations

Points should be inside the contour lines



### Collinearity

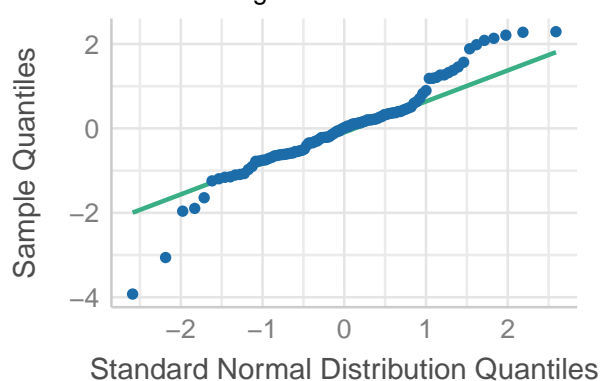
High collinearity (VIF) may inflate parameter uncertainty



● Low (< 5)

### Normality of Residuals

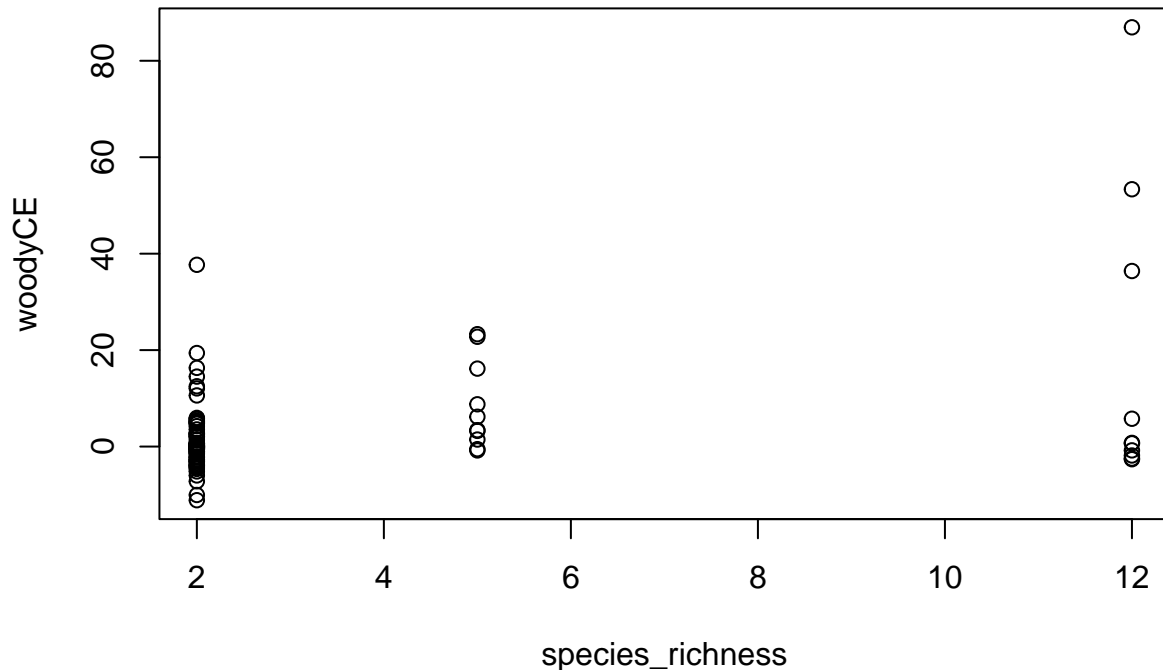
Dots should fall along the line



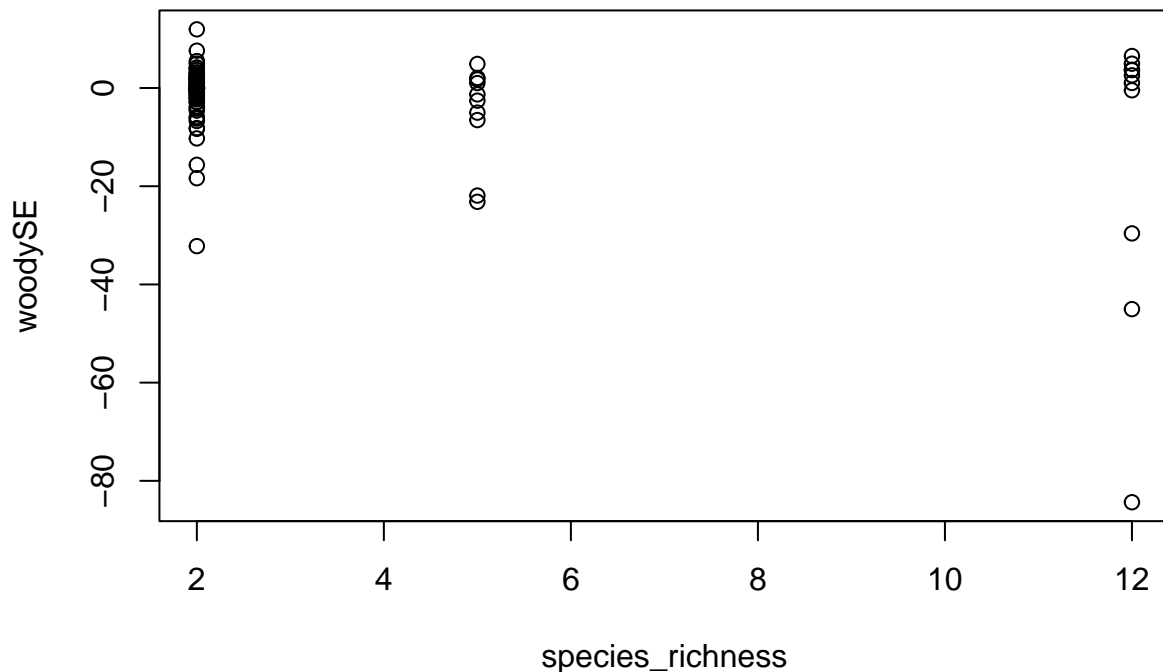
Robust regression tends to shrink parameter estimates slightly here, but yields a qualitatively very similar model.

Overyielding is often partitioned via Loreau & Hector (2001)'s method into complementarity and selection effects. Complementarity effects (CE) are based on the mean of species' *relative* yields and are often taken to indicate synergistic interactions like niche partitioning or facilitation. Selection (SE) effects describe how more productive species tend to show greater (or lesser) relative boosts in productivity than average in mixtures. Although CE and SE must sum to overyielding, it is possible for one to be greater than overyielding (or even total plot biomass) if the other is negative. (It is also possible for total overyielding to be negative, in which case it may be called underyielding.) Although CE and SE are expressed in the same units as overyielding (here, kg biomass per ha), it can be hard to make sense of extreme CE or SE values in these terms. For these reasons, we choose to focus on whether they are generally positive or negative than on their actual magnitudes. Starting with the Loreau & Hector (2001) paper itself, it is common for BEF researchers to expect positive CE and non-positive SE on average across plots. This would imply that (1) species tend to overperform in mixture and (2) it is not only the most productive species that overperform in mixture. It may help to look at the data quickly before proceeding.

```
plot(w WoodyCE~species_r richness,data=carbon_seq_woody0Y)
```



```
plot(w WoodySE~species_r richness,data=carbon_seq_woody0Y)
```



These plots reveal that there are a couple outliers with highly positive CE and negative SE. The most extreme are plots 14 and 50, both 12-species plots. Because of the odd distribution of CE and SE values, we'll conduct both a parametric t-test and a non-parametric Wilcoxon rank sum test to check whether CE or SE differ from 0. Most research does not concern itself with whether inferences about CE or SE are driven by extreme values, so the t-test is probably more consistent with how results are usually reported. However, the Wilcoxon test may be more appropriate given what we see in the plots above.

```
t.test(carbon_seq_woodyOY$woodyCE)
```

```
##
## One Sample t-test
##
## data: carbon_seq_woodyOY$woodyCE
## t = 2.8862, df = 103, p-value = 0.00475
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.087102 5.862464
## sample estimates:
## mean of x
##  3.474783
```

```
t.test(carbon_seq_woodyOY$woodySE)
```

```
##
## One Sample t-test
##
## data: carbon_seq_woodyOY$woodySE
## t = -2.1868, df = 103, p-value = 0.03102
```

```
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## -4.6108083 -0.2250558
## sample estimates:
## mean of x
## -2.417932
```

```
wilcox.test(carbon_seq_woodyOY$woodyCE)
```

```
##
## Wilcoxon signed rank test with continuity correction
##
## data: carbon_seq_woodyOY$woodyCE
## V = 3315, p-value = 0.05804
## alternative hypothesis: true location is not equal to 0
```

```
wilcox.test(carbon_seq_woodyOY$woodySE)
```

```
##
## Wilcoxon signed rank test with continuity correction
##
## data: carbon_seq_woodyOY$woodySE
## V = 2630, p-value = 0.747
## alternative hypothesis: true location is not equal to 0
```

There is some evidence that CE differs from 0, but not much evidence that SE does.

## Soil C accumulation and overyielding

We can now do the same for soil C accumulation:

```
carbon_seq_soilC<-carbon_seq[which(!is.na(carbon_seq$soilC)),]
mst<-lm(soilC~species_richness+FDis+PSV,data=carbon_seq_soilC)
```

```
options(na.action = "na.fail")
mst_dredge <- dredge(mst, beta = "none", evaluate = T, rank = AICc)
options(na.action = "na.omit")
```

```
mst_dredge ## we pick the first one
```

```
## Global model call: lm(formula = soilC ~ species_richness + FDis + PSV, data = carbon_seq_soilC)
## ---
```

```
## Model selection table
```

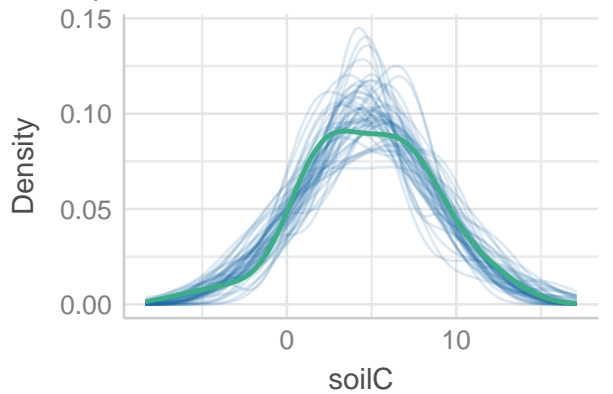
##	(Int)	FDis	PSV	spc_rch	df	logLik	AICc	delta	weight
## 5	3.881			0.2712	3	-178.975	364.3	0.00	0.356
## 3	4.099	2.5760			3	-180.022	366.4	2.09	0.125
## 7	3.839	0.5298	0.2406		4	-178.939	366.5	2.20	0.119
## 2	4.109	0.6608			3	-180.088	366.6	2.23	0.117
## 6	3.901	-0.1126	0.3000		4	-178.959	366.6	2.24	0.116
## 1	4.809				2	-181.616	367.4	3.09	0.076
## 8	3.823	-0.8537	2.7680	0.3294	5	-178.663	368.3	3.99	0.048
## 4	4.074	0.2645	1.6490		4	-179.983	368.6	4.28	0.042

```
## Models ranked by AICc(x)
```

```
mst_model<-get.models(mst_dredge, subset = 1)[[1]]
check_model(mst_model)
```

### Posterior Predictive Check

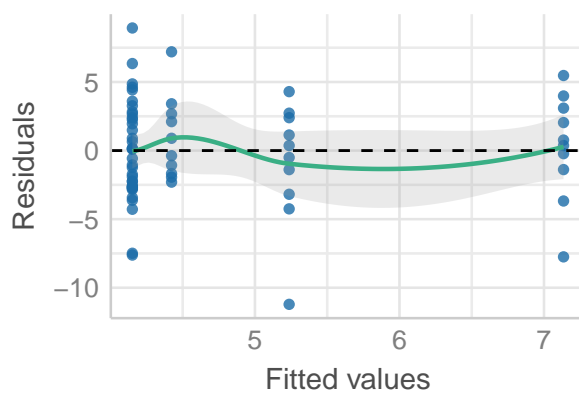
Model-predicted lines should resemble observed data



— Observed data — Model-predicted data

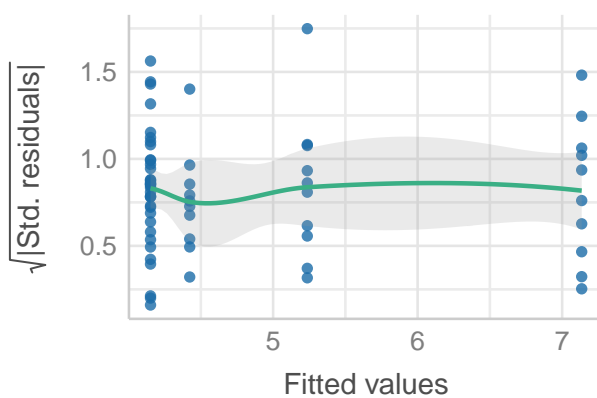
### Linearity

Reference line should be flat and horizontal



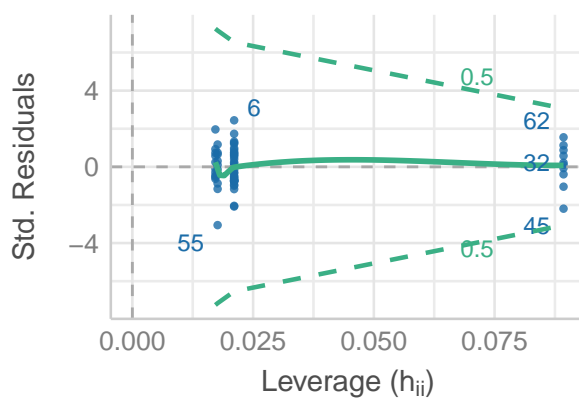
### Homogeneity of Variance

Reference line should be flat and horizontal



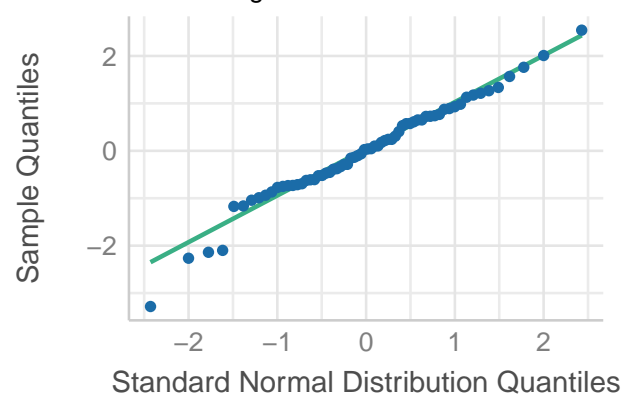
### Influential Observations

Points should be inside the contour lines



### Normality of Residuals

Dots should fall along the line



And for overyielding in soil C accumulation:

```
carbon_seq_soil0Y <- carbon_seq[which(!is.na(carbon_seq$soil0Y)),]
ms <- lm(soil0Y ~ species_richness + FDis + PSV, data = carbon_seq_soil0Y)
```

```
options(na.action = "na.fail")
ms_dredge <- dredge(ms, beta = "none", evaluate = T, rank = AICc)
options(na.action = "na.omit")
```

```
ms_dredge ## we pick the second one
```

```
## Global model call: lm(formula = soilOY ~ species_richness + FDis + PSV, data = carbon_seq_soilOY)
## ---
```

```
## Model selection table
```

	(Int)	FDs	PSV	spc_rch	df	logLik	AICc	delta	weight
## 5	-0.57840			0.3115	3	-81.153	169.2	0.00	0.325
## 1	1.39500				2	-82.987	170.4	1.19	0.179
## 7	2.12400		-4.879	0.3522	4	-80.519	170.6	1.41	0.161
## 2	-2.08600	1.4920			3	-82.271	171.5	2.24	0.106
## 6	0.05418	-0.3653		0.3462	4	-81.129	171.9	2.63	0.087
## 3	3.05200		-2.732		3	-82.805	172.5	3.30	0.062
## 4	-0.20830	1.6920	-3.863		4	-81.899	173.4	4.17	0.040
## 8	2.62100	-0.2960	-4.851	0.3800	5	-80.503	173.5	4.28	0.038

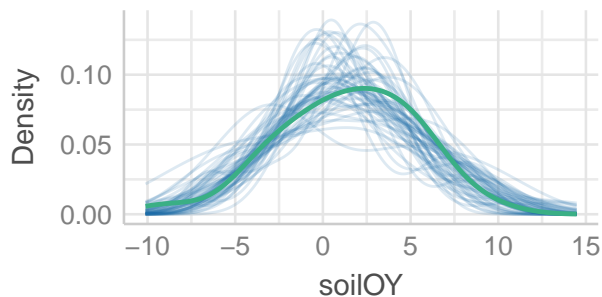
```
## Models ranked by AICc(x)
```

```
ms_model <- get.models(ms_dredge, subset = 2)[[1]]
```

```
check_model(ms_model)
```

### Posterior Predictive Check

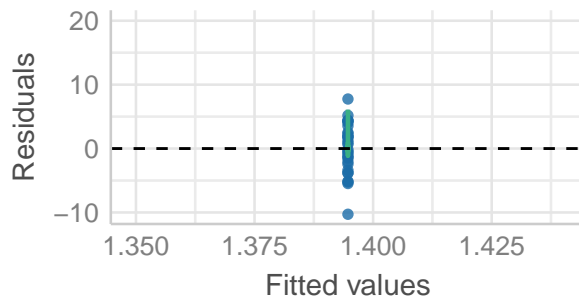
Model-predicted lines should resemble observed data



— Observed data — Model-predicted data

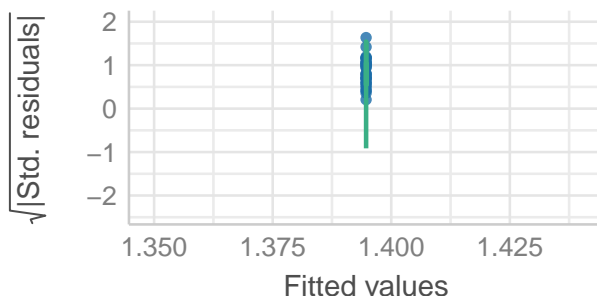
### Linearity

Reference line should be flat and horizontal



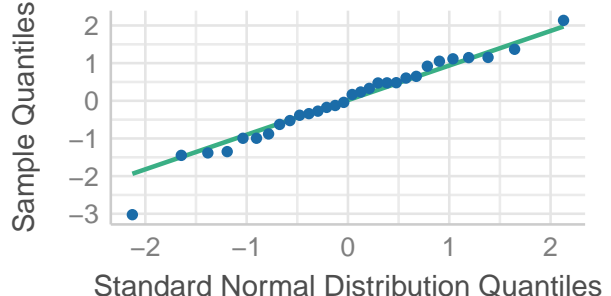
### Homogeneity of Variance

Reference line should be flat and horizontal



### Normality of Residuals

Dots should fall along the line



Both of these models seem to meet the assumptions quite well.

## Fine roots and overyielding

```
## total root C
carbon_seq_rootC<-carbon_seq[which(!is.na(carbon_seq$rootC)),]
mrt<-lm(rootC~species_richness+FDis+PSV,data=carbon_seq_rootC)

options(na.action = "na.fail")
mrt_dredge <- dredge(mrt, beta = "none", evaluate = T, rank = AICc)
options(na.action = "na.omit")

mrt_dredge ## we select the first model

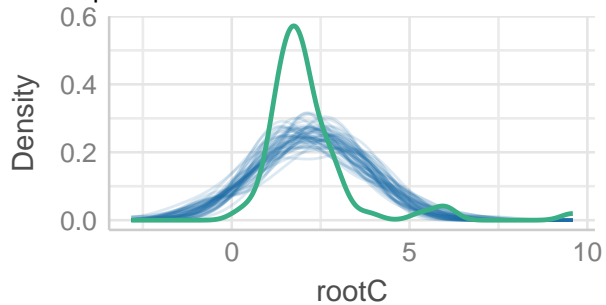
## Global model call: lm(formula = rootC ~ species_richness + FDis + PSV, data = carbon_seq_rootC)
## ---
## Model selection table
##      (Int)      FDs      PSV    spc_rch df    logLik  AICc delta weight
## 1 2.183
## 8 2.191 1.1200 -2.75600 -0.128900 5 -107.450 226.0 0.26 0.208
## 4 2.098 0.7036 -2.44400
## 2 2.056 0.1228
## 6 2.131 0.4149 -0.110000 4 -109.351 227.4 1.70 0.101
## 5 2.204 -0.006146 3 -110.736 227.9 2.19 0.080
## 3 2.165 0.06906 3 -110.738 227.9 2.19 0.079
## 7 2.186 0.26150 -0.021100 4 -110.685 230.1 4.37 0.027
## Models ranked by AICc(x)

mrt_model<-get.models(mrt_dredge, subset = 1)[[1]]
check_model(mrt_model)
```



### Posterior Predictive Check

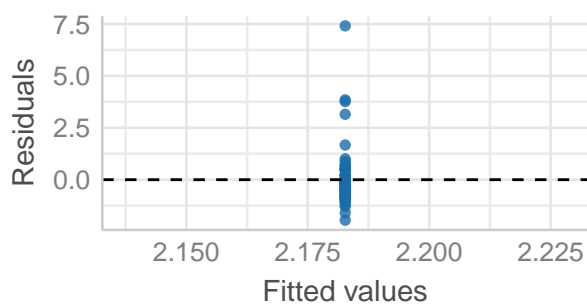
Model-predicted lines should resemble observed data



— Observed data — Model-predicted data

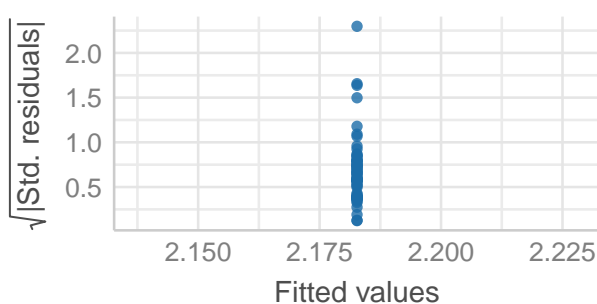
### Linearity

Reference line should be flat and horizontal



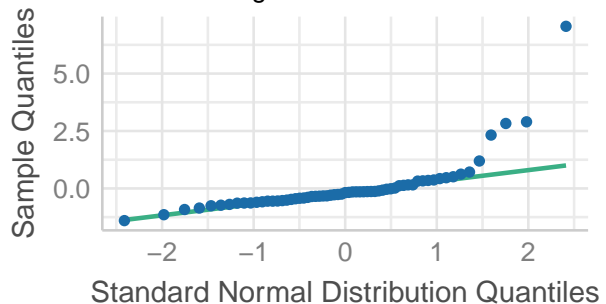
### Homogeneity of Variance

Reference line should be flat and horizontal



### Normality of Residuals

Dots should fall along the line



The selected model (with no predictors) does a poor job of meeting assumptions, but even changing out rootC for log(rootC), which meets the assumptions better, yields a ‘best’ model with no predictor variables. The same is true of robust regression. We move on to overyielding.

#### ## overyielding in root C

```
carbon_seq_rootOY<-carbon_seq[which(!is.na(carbon_seq$rootOY)),]
mr<-lm(rootOY~species_richness+FDis+PSV,data=carbon_seq_rootOY)
mr_robust<-rlm(rootOY~species_richness+FDis+PSV,data=carbon_seq_rootOY)

options(na.action = "na.fail")
mr_dredge <- dredge(mr, beta = "none", evaluate = T, rank = AICc)
mr_robust_dredge <- dredge(mr_robust, beta = "none", evaluate = T, rank = AICc)
options(na.action = "na.omit")

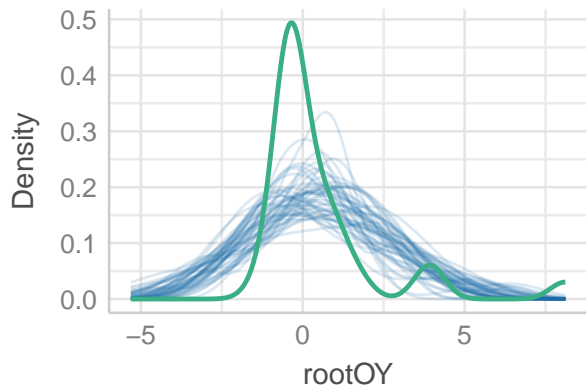
mr_dredge ## we select the first model
```

```
## Global model call: lm(formula = rootOY ~ species_richness + FDis + PSV, data = carbon_seq_rootOY)
## ---
## Model selection table
##      (Int)      FDis      PSV  spc_rch df  logLik  AICc  delta  weight
## 3 3.5240      -5.098      -0.03245  3 -54.654 116.3  0.00  0.506
## 7 3.5700      -4.823      -0.03245  4 -54.568 118.9  2.57  0.140
## 4 3.8510     -0.18630     -4.920      4 -54.605 118.9  2.64  0.135
## 1 0.5177
## 5 1.1020      -0.09082      3 -57.036 121.1  4.76  0.047
## 2 1.8690     -0.58290      3 -57.223 121.4  5.14  0.039
## 8 3.6110     -0.02526     -4.819     -0.03009  5 -54.568 121.9  5.56  0.031
```

```
## 6 1.3670 -0.15540      -0.07601  4 -57.021 123.8  7.47  0.012
## Models ranked by AICc(x)
mr_model<-get.models(mr_dredge, subset = 1)[[1]]
check_model(mr_model)
```

### Posterior Predictive Check

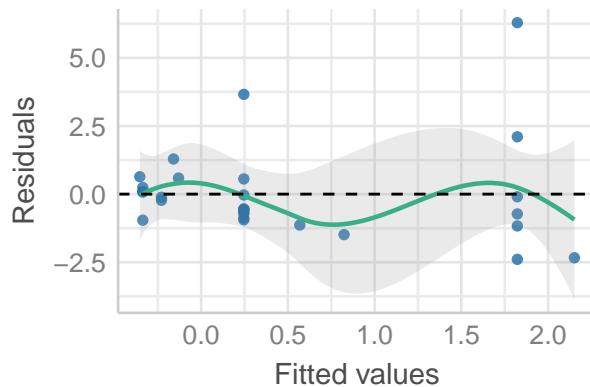
Model-predicted lines should resemble observed c



— Observed data — Model-predicted data

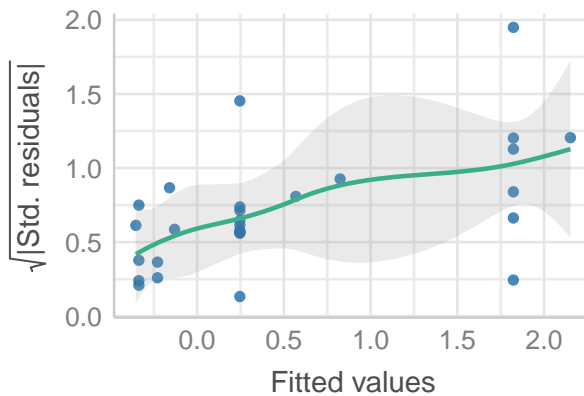
### Linearity

Reference line should be flat and horizontal



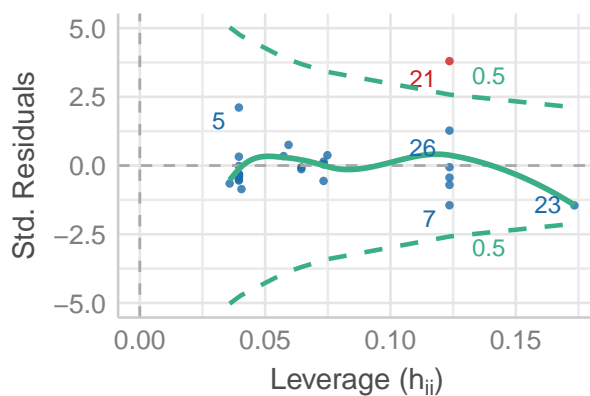
### Homogeneity of Variance

Reference line should be flat and horizontal



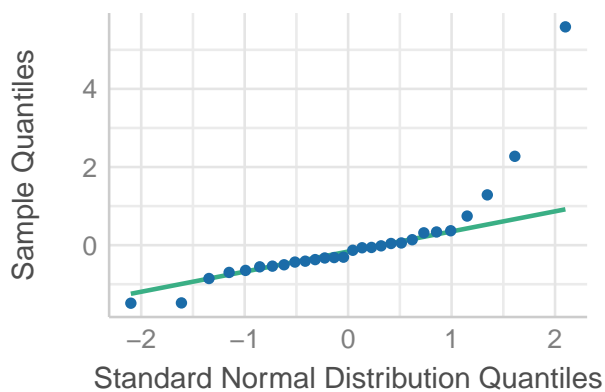
### Influential Observations

Points should be inside the contour lines



### Normality of Residuals

Dots should fall along the line



Evidently, this model also does a poor job of meeting assumptions. Thus, we show the robust regression results to buttress the claim that the same model is selected even when high-leverage points are not allowed to be as influential.

```
mr_robust<-rlm(rootOY~species_richness+FDis+PSV,data=carbon_seq_rootOY)

options(na.action = "na.fail")
mr_robust_dredge <- dredge(mr_robust, beta = "none", evaluate = T, rank = AICc)
options(na.action = "na.omit")

mr_robust_dredge

## Global model call: rlm(formula = rootOY ~ species_richness + FDis + PSV, data = carbon_seq_rootOY)
## ---
## Model selection table
##      (Int)      FDs  PSV  spc_rch df  logLik  AICc delta weight
## 3 1.6910          -2.632          3 -56.054 119.1  0.00  0.445
## 1 0.0654          2 -58.423 121.3  2.22  0.147
## 7 1.7780          -2.492 -0.02781 4 -55.955 121.6  2.54  0.125
## 4 1.8470 -0.079010 -2.585          4 -56.005 121.8  2.64  0.119
## 5 0.3619          -0.04437 3 -57.960 122.9  3.81  0.066
## 2 0.5933 -0.226700          3 -58.165 123.3  4.22  0.054
## 8 1.6320  0.099930 -2.543 -0.03692 5 -55.959 124.6  5.54  0.028
## 6 0.3433  0.009522          -0.04512 4 -57.967 125.7  6.57  0.017
## Models ranked by AICc(x)
```

PSV is still the only variable selected, but its parameter estimate shrinks quite a bit.

## Structural equation modeling

We use structural equation modeling (SEM) to tie together our hypotheses in a more integrative manner. Although we found no evidence elsewhere for a correlation between aboveground and soil C accumulation, SEMs allow us to directly test whether the influence of species richness on soil C accumulation is mediated by aboveground C.

We select six variables:

- species richness (exogenous)
- percent of planted trees associated with AM fungi (exogenous)
- percent of planted trees that are conifers (exogenous)
- aboveground C (endogenous)
- macroaggregates (endogenous)
- soil C accumulation (endogenous)

The first step is to drop rows with NAs in any of these variables, then z-standardize all of them so that regression coefficients in the SEM are standardized:

```
carbon_seq_sub<-carbon_seq[-which(is.na(carbon_seq$soilC) | is.na(carbon_seq$macro250)),]

carbon_seq_standard<-carbon_seq_sub
standard_cols<-c("species_richness","woodyC","soilC",
                 "macro250","percentAM","percentCon")
carbon_seq_standard[,standard_cols]<-scale(carbon_seq_standard[,standard_cols])
```

As explained in the main text and Appendix I, we proposed an initial model based on *a priori* hypotheses and earlier data analyses. The structure of this model is as follows:

```

localfit_orig_model<-psem(
  lmer(woodyC~species_richness+percentCon+(1|block),data=carbon_seq_standard),
  lmer(macro250~percentAM+percentCon+woodyC+(1|block),data=carbon_seq_standard),
  lmer(soilC~species_richness+percentAM+percentCon+woodyC+macro250+(1|block),
    data=carbon_seq_standard)
)

summary(localfit_orig_model)

```

```

##      |
##
## Structural Equation Model of localfit_orig_model
##
## Call:
##   woodyC ~ species_richness + percentCon
##   macro250 ~ percentAM + percentCon + woodyC
##   soilC ~ species_richness + percentAM + percentCon + woodyC + macro250
##
##      AIC
## 457.856
##
## ---
## Tests of directed separation:
##
##               Independ.Claim Test.Type      DF Crit.Value P.Value
## macro250 ~ species_richness + ...      coef 58.0493   -0.1347  0.8933
##               woodyC ~ percentAM + ...      coef 59.0416   -3.9468  0.0002 ***
##
## --
## Global goodness-of-fit:
##
## Chi-Squared = 6.883 with P-value = 0.032 and on 2 degrees of freedom
## Fisher's C = 17.132 with P-value = 0.002 and on 4 degrees of freedom
##
## ---
## Coefficients:
##
## Response      Predictor Estimate Std.Error    DF Crit.Value P.Value
## woodyC species_richness  0.1463    0.0599 60.1193    2.4410  0.0176
## woodyC      percentCon  0.8564    0.0600 60.1829   14.2796  0.0000
## macro250      percentAM -0.1520    0.0852 59.0404   -1.7827  0.0798
## macro250      percentCon -0.0673    0.1684 59.1552   -0.3996  0.6909
## macro250        woodyC -0.2397    0.1748 59.2409   -1.3710  0.1756
## soilC species_richness  0.2884    0.1252 59.0000    2.3031  0.0248
## soilC      percentAM  0.2972    0.1324 59.0000    2.2439  0.0286
## soilC      percentCon -0.0085    0.2725 59.0000   -0.0313  0.9751
## soilC        woodyC  0.0767    0.2740 59.0000    0.2799  0.7806
## soilC      macro250  0.1330    0.1249 59.0000    1.0648  0.2913
## Std.Estimate
##      0.1463  *
##      0.8564 ***
##     -0.1520
##     -0.0673

```

```
##      -0.2397
##      0.2884  *
##      0.2972  *
##      -0.0085
##      0.0767
##      0.1330
##
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05
##
## ---
## Individual R-squared:
##
## Response method Marginal Conditional
##      woodyC      none      0.75      0.77
##      macro250     none      0.07      0.73
##      soilC        none      0.16      0.16
```

By running `check_models()` on each subcomponent of the model, it becomes clear that percent conifer and aboveground C have moderately high VIF due to their high correlation with each other. We removed aboveground C from the model for macroaggregates and percent conifer from the model for soil C accumulation, although we caution that in any case it is a challenge to attribute anything definitely to one or the other variable.

Two other things are apparent. One is that the block random effect for soil C reaches a singular fit, so we remove it. The other is that the tests of directed separation show conditional dependence between percent AM and woody C. Since percent AM is exogenous, we add it as a predictor of woody C. We arrive at the following model:

```
localfit_full_model<-psem(
  lmer(woodyC~species_richness+percentCon+percentAM+(1|block),data=carbon_seq_standard),
  lmer(macro250~percentAM+percentCon+(1|block),data=carbon_seq_standard),
  lm(soilC~species_richness+percentAM+woodyC+macro250,
    data=carbon_seq_standard)
)

summary(localfit_full_model)
```

```
##      |
##
## Structural Equation Model of localfit_full_model
##
## Call:
##      woodyC ~ species_richness + percentCon + percentAM
##      macro250 ~ percentAM + percentCon
##      soilC ~ species_richness + percentAM + woodyC + macro250
##
##      AIC
##      430.913
##
## ---
## Tests of directed separation:
##
##      Independ.Claim Test.Type      DF Crit.Value P.Value
##      macro250 ~ species_richness + ...      coef 59.0081      -0.5527  0.5825
##      soilC ~ percentCon + ...      coef 59.0000      -0.0313  0.9751
```

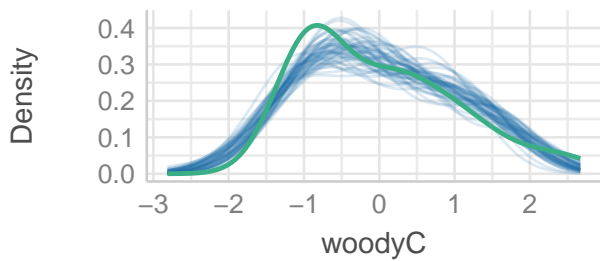
```
##          macro250 ~ woodyC + ...      coef 58.2852    -1.2467    0.2175
##
## --
## Global goodness-of-fit:
##
## Chi-Squared = NA with P-value = NA and on 3 degrees of freedom
## Fisher's C = 4.182 with P-value = 0.652 and on 6 degrees of freedom
##
## ---
## Coefficients:
##
##      Response      Predictor Estimate Std.Error    DF Crit.Value P.Value
##      woodyC species_richness   0.1361   0.0539  59.0976    2.5268  0.0142
##      woodyC   percentCon      0.8635   0.0539  59.1531   16.0316  0.0000
##      woodyC   percentAM     -0.2125   0.0538  59.0416   -3.9468  0.0002
##      macro250 percentAM     -0.0996   0.0769  60.0035   -1.2954  0.2001
##      macro250 percentCon     -0.2729   0.0770  60.0133   -3.5439  0.0008
##      soilC species_richness   0.2896   0.1190  60.0000    2.4339  0.0179
##      soilC   percentAM      0.2956   0.1212  60.0000    2.4380  0.0177
##      soilC      woodyC      0.0690   0.1217  60.0000    0.5671  0.5728
##      soilC   macro250      0.1340   0.1198  60.0000    1.1183  0.2679
## Std.Estimate
##      0.1361  *
##      0.8635 ***
##     -0.2125 ***
##     -0.0996
##     -0.2729 ***
##      0.2896  *
##      0.2956  *
##      0.0690
##      0.1340
##
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05
##
## ---
## Individual R-squared:
##
##      Response method Marginal Conditional
##      woodyC   none      0.79      0.82
##      macro250 none      0.07      0.71
##      soilC    none      0.17      NA
```

As the Fisher's  $C > 0.05$  shows, this model fits better than the one above. The in-depth interpretation of this model is in the main text. We can check each of the submodels of the full SEM to make sure that it means standard assumptions.

```
check_model(lmer(woodyC~species_richness+percentCon+percentAM+(1|block),data=carbon_seq_standard))
```

## Posterior Predictive Check

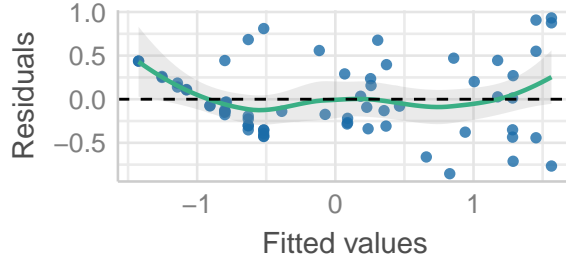
Model-predicted lines should resemble observed data



— Observed data — Model-predicted data

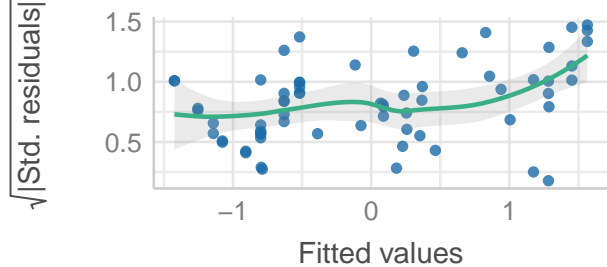
## Linearity

Reference line should be flat and horizontal



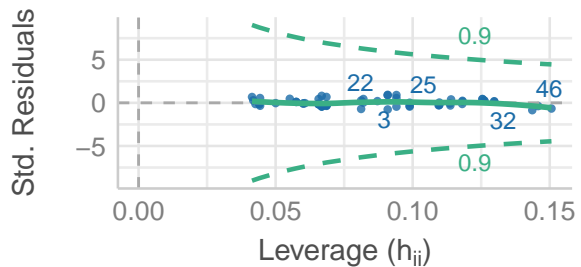
## Homogeneity of Variance

Reference line should be flat and horizontal



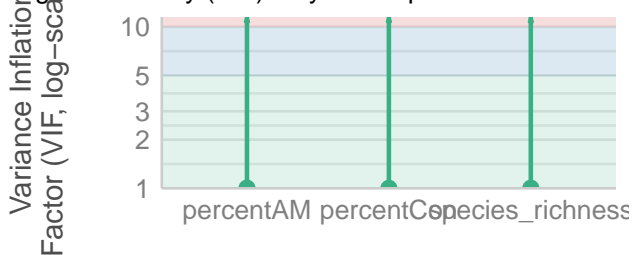
## Influential Observations

Points should be inside the contour lines



## Collinearity

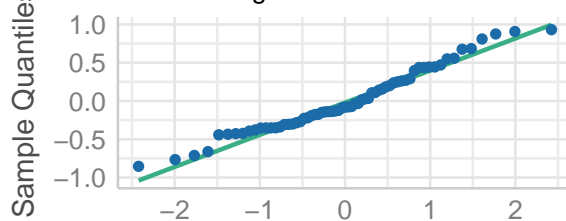
High collinearity (VIF) may inflate parameter uncertainty



● Low (< 5)

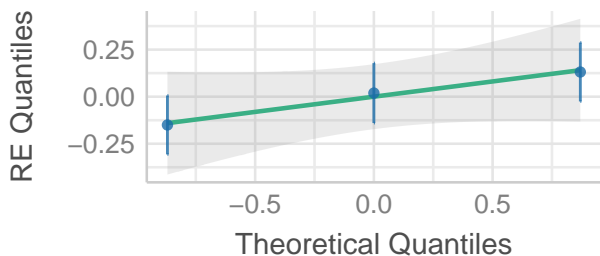
## Normality of Residuals

Dots should fall along the line



## Normality of Random Effects (block)

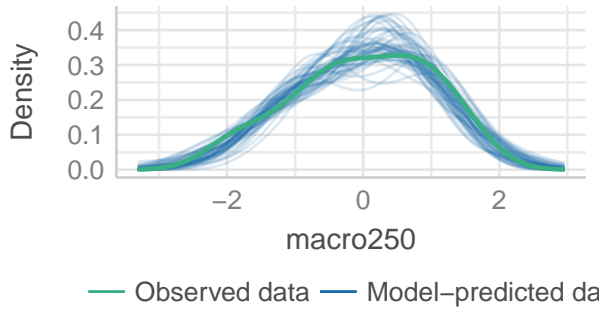
Dots should be plotted along the line



```
check_model(lmer(macro250~percentAM+percentCon+(1|block),data=carbon_seq_standard))
```

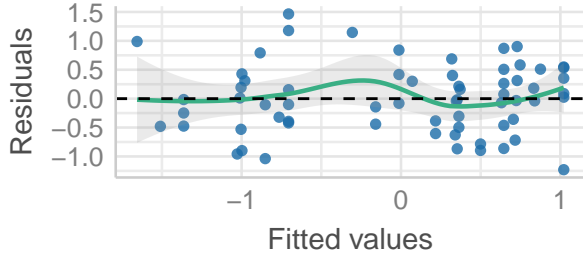
### Posterior Predictive Check

Model-predicted lines should resemble observed data



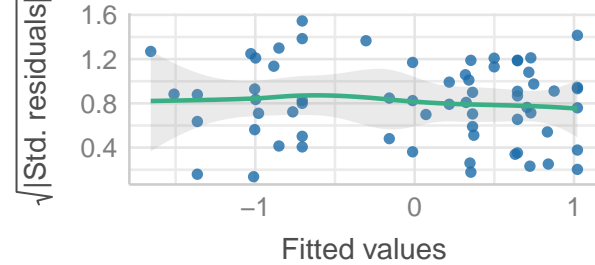
### Linearity

Reference line should be flat and horizontal



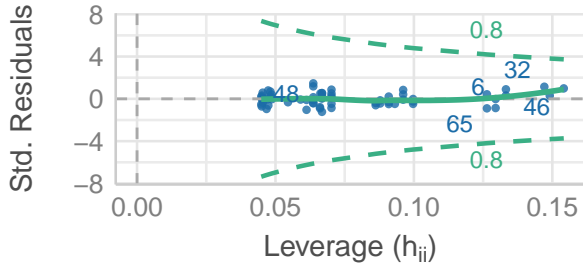
### Homogeneity of Variance

Reference line should be flat and horizontal



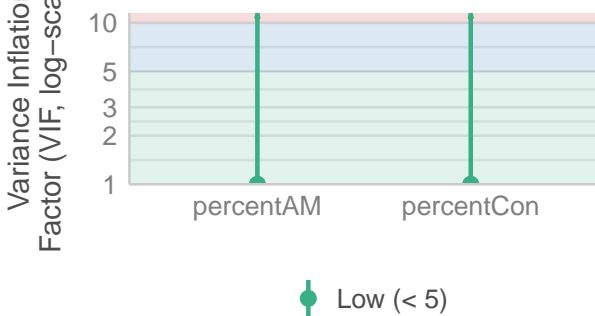
### Influential Observations

Points should be inside the contour lines



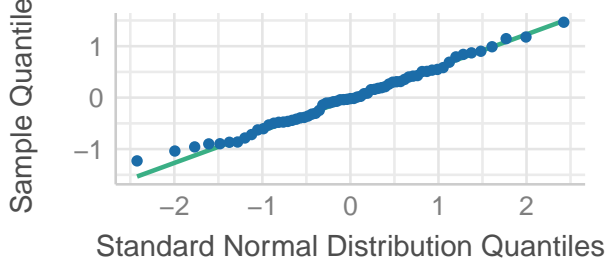
### Collinearity

High collinearity (VIF) may inflate parameter uncertainty



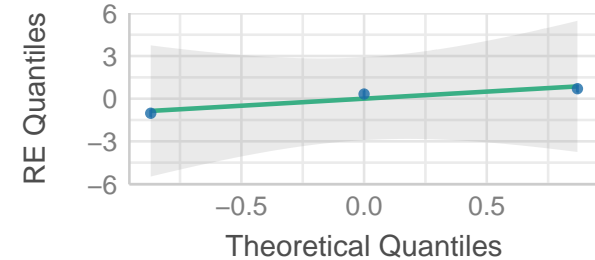
### Normality of Residuals

Dots should fall along the line



### Normality of Random Effects (block)

Dots should be plotted along the line

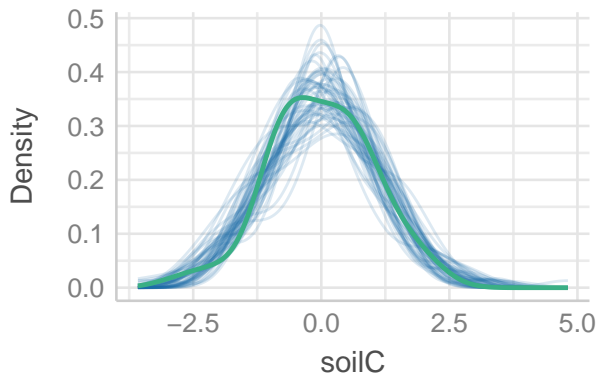


```
check_model(lm(soilC~species_richness+percentAM+woodyC+macro250,data=carbon_seq_standard))
```



### Posterior Predictive Check

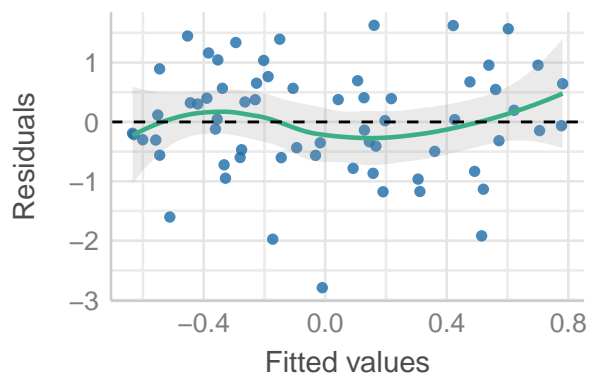
Model-predicted lines should resemble observed data



— Observed data — Model-predicted data

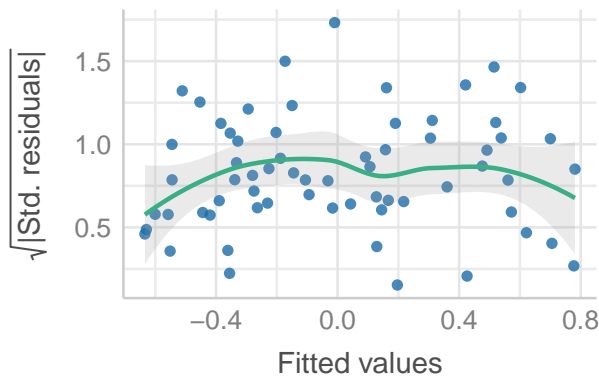
### Linearity

Reference line should be flat and horizontal



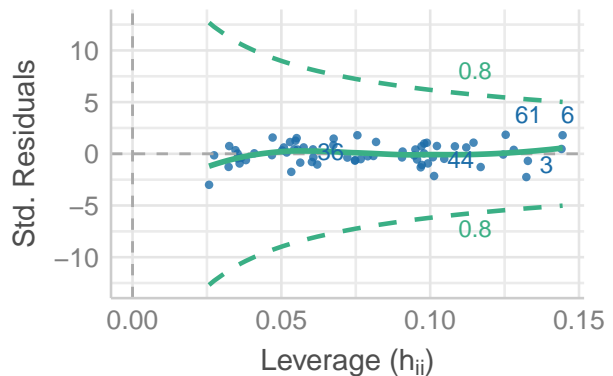
### Homogeneity of Variance

Reference line should be flat and horizontal



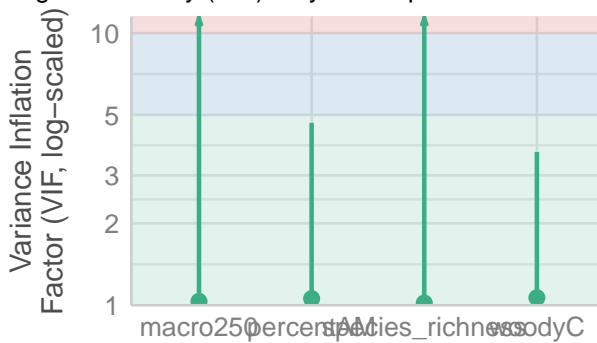
### Influential Observations

Points should be inside the contour lines



### Collinearity

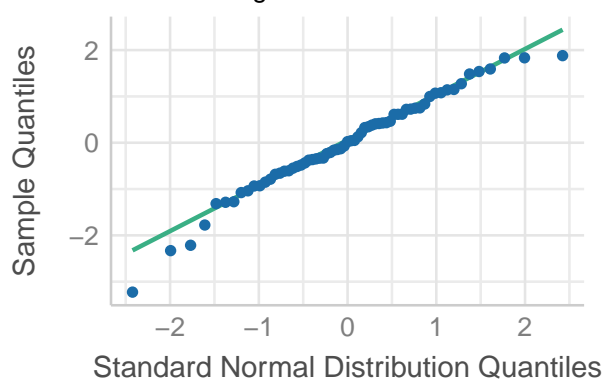
High collinearity (VIF) may inflate parameter uncertainty



● Low (< 5)

### Normality of Residuals

Dots should fall along the line



Each of these plots looks reasonably good. There is no particular need to fit a more minimal model that only retains significant terms, but if we did, it would look like this:

```
localfit_min_model <- psem(
  lmer(woodyC ~ species_richness + percentCon + percentAM + (1|block), data=carbon_seq_standard),
```

```
lmer(macro250~percentCon+(1|block),data=carbon_seq_standard),  
lm(soilC~species_richness+percentAM,  
  data=carbon_seq_standard)  
)
```