Curry Co. SOD Analysis

Hazel Daniels

2/5/2021

# R setup and libraries

# Reading and manipulating data

SODSites <- read\_xlsx("BasicData.xlsx", 1)  
SODSite.ord <- SODSites[-c(31),,drop=F] #removing HD-4331 as it skews results strongly  
  
# Subsetting data  
AllSites <- SODSites[-c(76, 78, 79, 82, 83, 84, 88, 97, 98, 99),,drop=F] #removing HD-4376, 4378, 4379, 4382, 4383, 4384, 4388, 4397, 4398, 4399 as these sites were measured before and after treatment (removing treated data). Removing these duplicate sites to a data subset for separate analysis.  
AllSites <- subset(AllSites, AllSites$Lineage != "U") #removing uninfested sites  
NA1 <- subset(AllSites, AllSites$Lineage == "NA1")  
  
AllSeeds <- read\_xlsx("BasicData.xlsx", 2)  
AllSeeds <- column\_to\_rownames(AllSeeds, var = "New Site")  
AllSeeds <- AllSeeds[-11] #removing CE as it skews results strongly  
AllSeeds <- AllSeeds[-c(31),,drop=F] #removing HD-4331 as it skews results strongly  
  
AllFire <- read\_xlsx("BasicData.xlsx", 3)

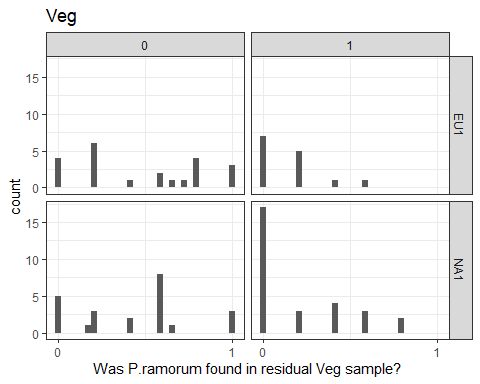
EU1wide <- read\_xlsx("BasicData.xlsx", 4)

EU1long <- read\_xlsx("BasicData.xlsx", 5)

# Visualizing raw data

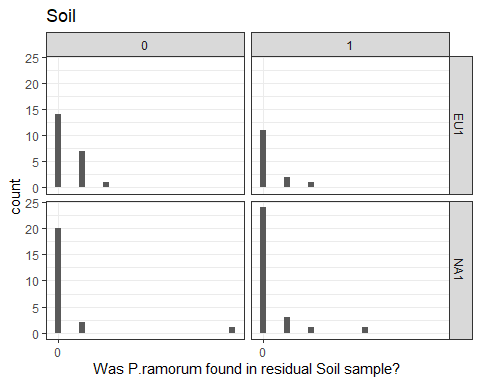
ggplot(AllSites, aes(x=x)) +  
 geom\_histogram( aes(x = VegProp, y = stat(count))) +  
 scale\_x\_continuous(name = "Was P.ramorum found in residual Veg sample?", breaks = seq(0, 1)) +  
 facet\_grid(Lineage ~ TrtBin) +  
 labs(title = "Veg") +  
 theme\_bw()

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



ggplot(AllSites, aes(x=x)) +  
 geom\_histogram( aes(x = SoilProp, y = stat(count))) +  
 scale\_x\_continuous(name = "Was P.ramorum found in residual Soil sample?", breaks = seq(0, 1)) +  
 facet\_grid(Lineage ~ TrtBin) +  
 labs(title = "Soil") +  
 theme\_bw()

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

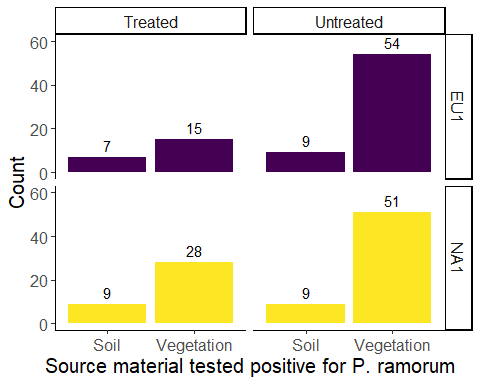


Trt <- c(rep("Untreated", 4), rep("Treated", 4))  
Lin <- c("EU1","EU1","NA1","NA1","EU1","EU1","NA1","NA1")  
Tis <- c("Vegetation","Soil","Vegetation","Soil","Vegetation","Soil","Vegetation","Soil")  
Count <- c(54, 9, 51, 9, 15, 7, 28, 9)  
df <- data.frame(Trt,Lin,Tis,Count)  
df

## Trt Lin Tis Count  
## 1 Untreated EU1 Vegetation 54  
## 2 Untreated EU1 Soil 9  
## 3 Untreated NA1 Vegetation 51  
## 4 Untreated NA1 Soil 9  
## 5 Treated EU1 Vegetation 15  
## 6 Treated EU1 Soil 7  
## 7 Treated NA1 Vegetation 28  
## 8 Treated NA1 Soil 9

ggplot(df,aes(x=Tis,y=Count,fill=Lin))+  
 geom\_bar(stat="identity",position="dodge")+  
 scale\_fill\_discrete(name="Lineage",labels=c("EU1", "NA1"))+  
 scale\_fill\_viridis\_d() +  
 theme\_classic()+  
 theme(legend.position="none", text = element\_text(size=15))+  
 xlab("Source material tested positive for P. ramorum")+  
 ylab("Count")+  
 ylim(0, 60)+  
 geom\_text(aes(label = Count), vjust = -0.5)+  
 facet\_grid(Lin~Trt)

## Scale for 'fill' is already present. Adding another scale for 'fill', which  
## will replace the existing scale.



#This visualization includes all 100 sites sampled, including uninfested sites, and EU1 sites that were sampled before and after treatment.

# Vegetation

## All Veg

V1 <- lm(VegProp ~ Lineage + TrtBin, data = AllSites)  
  
anova(V1)

## Analysis of Variance Table  
##   
## Response: VegProp  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Lineage 1 0.0398 0.03984 0.4419 0.508   
## TrtBin 1 1.6465 1.64646 18.2644 4.998e-05 \*\*\*  
## Residuals 85 7.6624 0.09015   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

summary(V1)

##   
## Call:  
## lm(formula = VegProp ~ Lineage + TrtBin, data = AllSites)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.45539 -0.17791 -0.05539 0.21665 0.62209   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.451828 0.056050 8.061 4.27e-12 \*\*\*  
## LineageNA1 0.003567 0.066013 0.054 0.957   
## TrtBin -0.277489 0.064930 -4.274 5.00e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3002 on 85 degrees of freedom  
## Multiple R-squared: 0.1804, Adjusted R-squared: 0.1611   
## F-statistic: 9.353 on 2 and 85 DF, p-value: 0.0002131

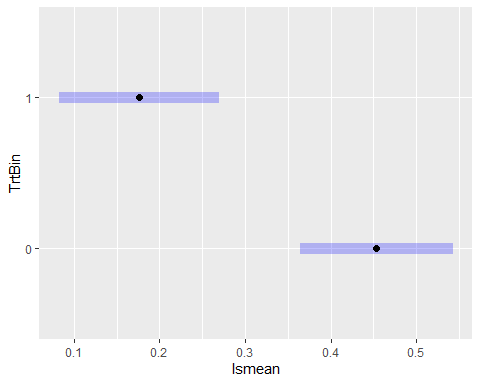
summ(V1)

## MODEL INFO:  
## Observations: 88  
## Dependent Variable: VegProp  
## Type: OLS linear regression   
##   
## MODEL FIT:  
## F(2,85) = 9.35, p = 0.00  
## R² = 0.18  
## Adj. R² = 0.16   
##   
## Standard errors: OLS  
## ------------------------------------------------  
## Est. S.E. t val. p  
## ----------------- ------- ------ -------- ------  
## (Intercept) 0.45 0.06 8.06 0.00  
## LineageNA1 0.00 0.07 0.05 0.96  
## TrtBin -0.28 0.06 -4.27 0.00  
## ------------------------------------------------

# Contrast statement  
V1.lsm <- lsmeans(V1, "TrtBin")  
V1.lsm

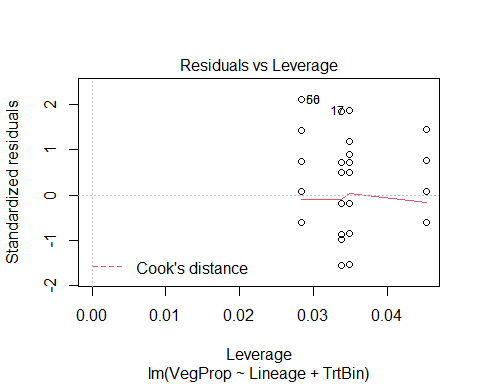
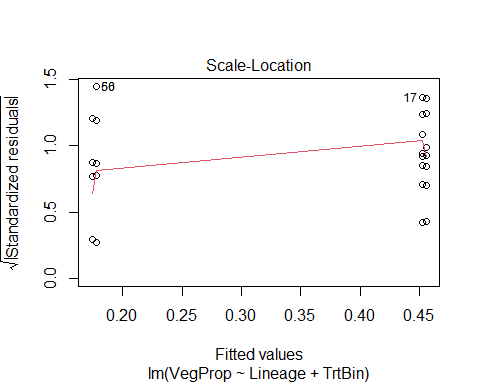
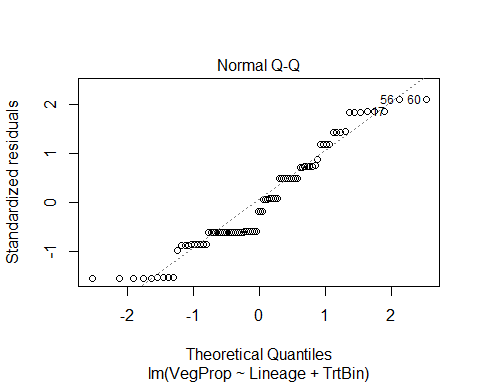
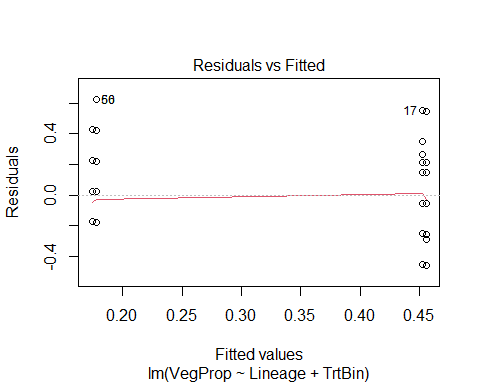
## TrtBin lsmean SE df lower.CL upper.CL  
## 0 0.454 0.0448 85 0.3646 0.543  
## 1 0.176 0.0472 85 0.0823 0.270  
##   
## Results are averaged over the levels of: Lineage   
## Confidence level used: 0.95

plot(V1.lsm)



## All Veg Assumptions

plot(V1)



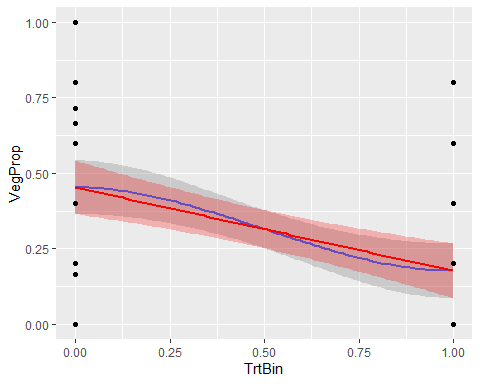
gvlma(V1)

##   
## Call:  
## lm(formula = VegProp ~ Lineage + TrtBin, data = AllSites)  
##   
## Coefficients:  
## (Intercept) LineageNA1 TrtBin   
## 0.451828 0.003567 -0.277489   
##   
##   
## ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS  
## USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:  
## Level of Significance = 0.05   
##   
## Call:  
## gvlma(x = V1)   
##   
## Value p-value Decision  
## Global Stat 5.9626 0.2020 Assumptions acceptable.  
## Skewness 2.3811 0.1228 Assumptions acceptable.  
## Kurtosis 2.1058 0.1467 Assumptions acceptable.  
## Link Function 0.4347 0.5097 Assumptions acceptable.  
## Heteroscedasticity 1.0410 0.3076 Assumptions acceptable.

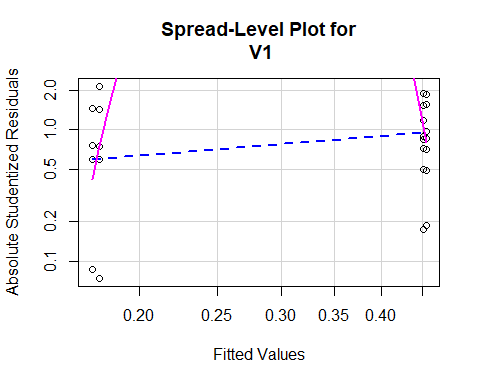
#linearity  
ggplot(AllSites, aes(TrtBin,VegProp)) +   
 stat\_smooth(method="loess") +   
 stat\_smooth(method="lm", color="red", fill="red", alpha=.25) +   
 geom\_point()

## `geom\_smooth()` using formula 'y ~ x'

## `geom\_smooth()` using formula 'y ~ x'

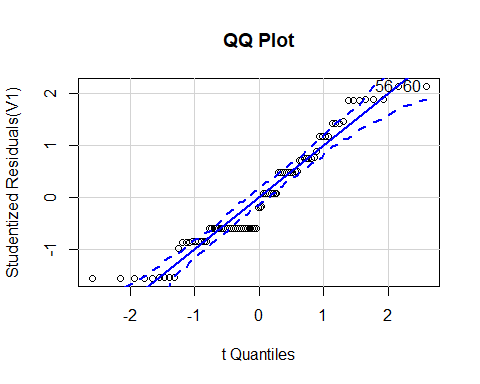


#constant variance  
spreadLevelPlot(V1)

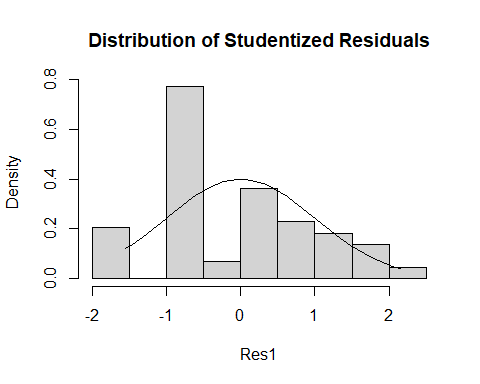


##   
## Suggested power transformation: 0.5043893

#normality of residuals  
Resq1 <- qqPlot(V1, main="QQ Plot")



Res1 <- studres(V1)  
hist(Res1, freq = FALSE, main="Distribution of Studentized Residuals")  
xfit1 <- seq(min(Res1), max(Res1), length = 40)  
yfit1 <- dnorm(xfit1)  
lines(xfit1, yfit1)



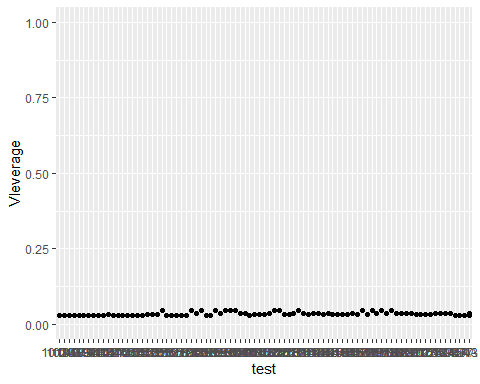
#multicollinearity  
vif(V1)

## Lineage TrtBin   
## 1.028348 1.028348

sqrt(vif(V1)) > 2

## Lineage TrtBin   
## FALSE FALSE

#outliers  
AllSites$Vleverage <- hatvalues(V1)  
ggplot(AllSites, aes(Plot, Vleverage)) + geom\_point() + ylim(0,1) + xlab("test")



## All Veg Correlations

#testing correlations of random effects which were not included in the model  
VC1 <- cor.test(AllSites$VegProp, AllSites$YrSinceTrt, method=c("pearson", "kendall", "spearman"))  
VC11 <- cor.test(AllSites$VegProp, AllSites$TotalTO, method=c("pearson", "kendall", "spearman"))  
VC111 <- cor.test(AllSites$VegProp, AllSites$Num\_Seedlings, method=c("pearson", "kendall", "spearman"))  
VC1111 <- cor.test(AllSites$VegProp, AllSites$SoilProp, method=c("pearson", "kendall", "spearman"))  
VC1

##   
## Pearson's product-moment correlation  
##   
## data: AllSites$VegProp and AllSites$YrSinceTrt  
## t = -1.6704, df = 86, p-value = 0.09848  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.37286767 0.03341372  
## sample estimates:  
## cor   
## -0.1772691

VC11

##   
## Pearson's product-moment correlation  
##   
## data: AllSites$VegProp and AllSites$TotalTO  
## t = 0.83633, df = 86, p-value = 0.4053  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.1219169 0.2937354  
## sample estimates:  
## cor   
## 0.08981887

VC111

##   
## Pearson's product-moment correlation  
##   
## data: AllSites$VegProp and AllSites$Num\_Seedlings  
## t = -1.1365, df = 86, p-value = 0.2589  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.32286202 0.09009204  
## sample estimates:  
## cor   
## -0.1216456

VC1111 #correlation between soil and veg!

##   
## Pearson's product-moment correlation  
##   
## data: AllSites$VegProp and AllSites$SoilProp  
## t = 2.931, df = 86, p-value = 0.004328  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.09811131 0.48047666  
## sample estimates:  
## cor   
## 0.301361

## NA1 Veg Wildfire

V2 <- lm(VegProp ~ RX, data = AllFire)  
  
anova(V2)

## Analysis of Variance Table  
##   
## Response: VegProp  
## Df Sum Sq Mean Sq F value Pr(>F)   
## RX 3 1.0706 0.35687 4.1361 0.01095 \*  
## Residuals 48 4.1416 0.08628   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

summary(V2)

##   
## Call:  
## lm(formula = VegProp ~ RX, data = AllFire)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.43623 -0.24000 -0.04444 0.16377 0.56377   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.43623 0.06125 7.122 4.78e-09 \*\*\*  
## RX1 -0.19623 0.09749 -2.013 0.0498 \*   
## RX2 -0.11623 0.14494 -0.802 0.4266   
## RX3 -0.39179 0.11549 -3.392 0.0014 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.2937 on 48 degrees of freedom  
## Multiple R-squared: 0.2054, Adjusted R-squared: 0.1557   
## F-statistic: 4.136 on 3 and 48 DF, p-value: 0.01095

summ(V2)

## MODEL INFO:  
## Observations: 52  
## Dependent Variable: VegProp  
## Type: OLS linear regression   
##   
## MODEL FIT:  
## F(3,48) = 4.14, p = 0.01  
## R² = 0.21  
## Adj. R² = 0.16   
##   
## Standard errors: OLS  
## ------------------------------------------------  
## Est. S.E. t val. p  
## ----------------- ------- ------ -------- ------  
## (Intercept) 0.44 0.06 7.12 0.00  
## RX1 -0.20 0.10 -2.01 0.05  
## RX2 -0.12 0.14 -0.80 0.43  
## RX3 -0.39 0.12 -3.39 0.00  
## ------------------------------------------------

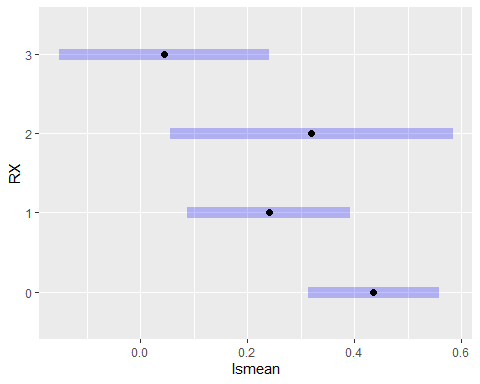
# Contrast statement  
levels(AllFire$RX)

## [1] "0" "1" "2" "3"

# 0 - no WF, no RX  
# 1 - no WF, yes RX  
# 2 - yes WF, no RX  
# 3 - yes WF, yes RX  
  
V2.lsm <- lsmeans(V2, ~RX, adjst="tukey")  
V2.lsm

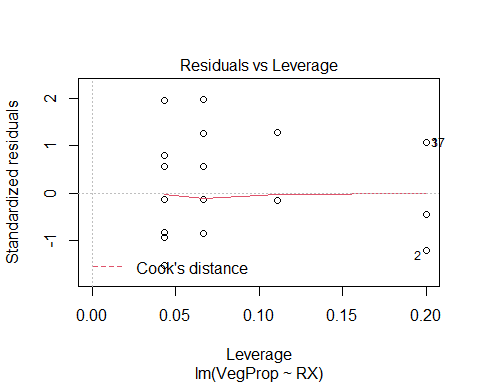
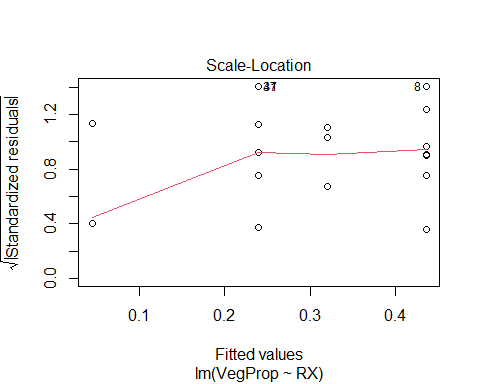
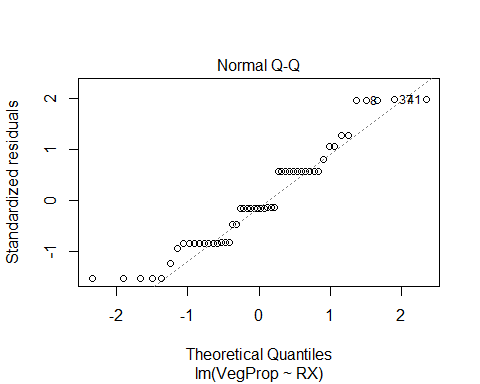
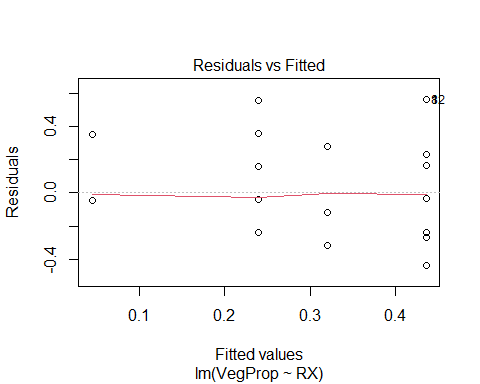
## RX lsmean SE df lower.CL upper.CL  
## 0 0.4362 0.0612 48 0.3131 0.559  
## 1 0.2400 0.0758 48 0.0875 0.392  
## 2 0.3200 0.1314 48 0.0559 0.584  
## 3 0.0444 0.0979 48 -0.1524 0.241  
##   
## Confidence level used: 0.95

plot(V2.lsm)



## NA1 Veg Wildfire Assumptions

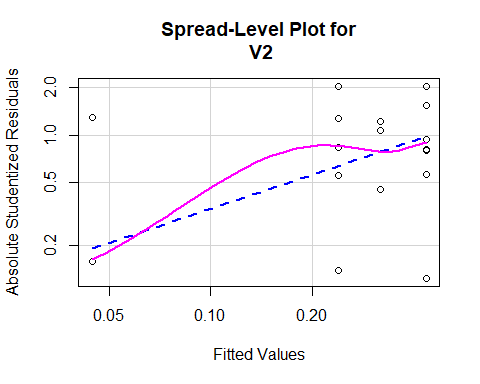
plot(V2)



gvlma(V2)

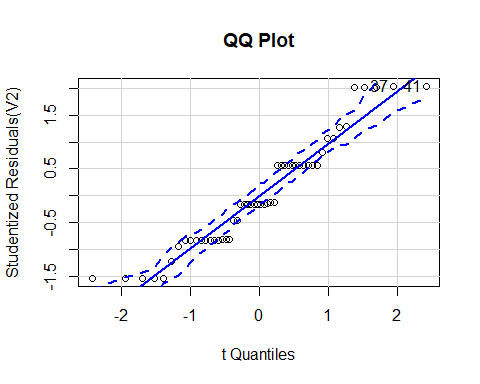
##   
## Call:  
## lm(formula = VegProp ~ RX, data = AllFire)  
##   
## Coefficients:  
## (Intercept) RX1 RX2 RX3   
## 0.4362 -0.1962 -0.1162 -0.3918   
##   
##   
## ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS  
## USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:  
## Level of Significance = 0.05   
##   
## Call:  
## gvlma(x = V2)   
##   
## Value p-value Decision  
## Global Stat 3.127e+00 0.5368 Assumptions acceptable.  
## Skewness 1.354e+00 0.2445 Assumptions acceptable.  
## Kurtosis 8.049e-01 0.3696 Assumptions acceptable.  
## Link Function -6.622e-16 1.0000 Assumptions acceptable.  
## Heteroscedasticity 9.674e-01 0.3253 Assumptions acceptable.

#constant variance  
spreadLevelPlot(V2)

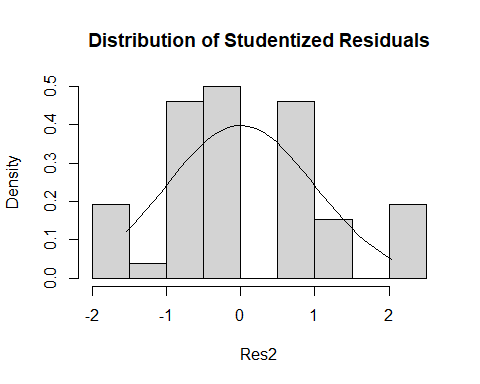


##   
## Suggested power transformation: 0.2873349

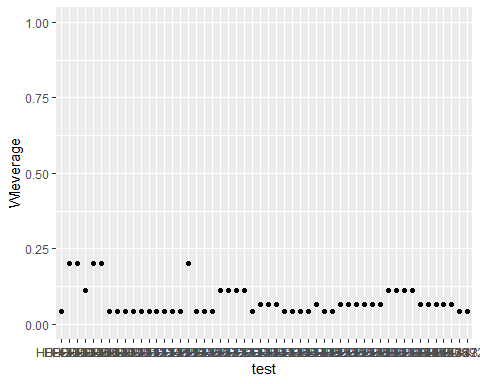
#normality of residuals  
Resq2 <- qqPlot(V2, main="QQ Plot")



Res2 <- studres(V2)  
hist(Res2, freq = FALSE,  
 main="Distribution of Studentized Residuals")  
xfit2 <- seq(min(Res2), max(Res2), length = 40)  
yfit2 <- dnorm(xfit2)  
lines(xfit2, yfit2)



#outliers  
AllFire$Wleverage <- hatvalues(V2)  
ggplot(AllFire, aes(`New Site`, Wleverage)) + geom\_point() + ylim(0,1) + xlab("test")



## NA1 Veg Wildfire Correlations

VC2 <- cor.test(AllFire$VegProp, AllFire$TotalTO, method=c("pearson", "kendall", "spearman"))  
VC22 <- cor.test(AllFire$VegProp, AllFire$SoilProp, method=c("pearson", "kendall", "spearman"))  
  
VC2

##   
## Pearson's product-moment correlation  
##   
## data: AllFire$VegProp and AllFire$TotalTO  
## t = 1.093, df = 50, p-value = 0.2796  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.1253661 0.4086259  
## sample estimates:  
## cor   
## 0.1527605

VC22 #correlation between soil and veg!

##   
## Pearson's product-moment correlation  
##   
## data: AllFire$VegProp and AllFire$SoilProp  
## t = 3.2198, df = 50, p-value = 0.002256  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.1595480 0.6174743  
## sample estimates:  
## cor   
## 0.4144049

## EU1 Repeated Measures

V3 <- lme(VegProp ~ TrtBin, random = ~1|Plot, data = EU1long)  
  
anova(V3)

## numDF denDF F-value p-value  
## (Intercept) 1 9 16.286969 0.0029  
## TrtBin 1 9 8.684625 0.0163

summary(V3)

## Linear mixed-effects model fit by REML  
## Data: EU1long   
## AIC BIC logLik  
## 24.68194 28.24342 -8.340968  
##   
## Random effects:  
## Formula: ~1 | Plot  
## (Intercept) Residual  
## StdDev: 0.08728716 0.3273545  
##   
## Fixed effects: VegProp ~ TrtBin   
## Value Std.Error DF t-value p-value  
## (Intercept) 0.5314286 0.1071354 9 4.960343 0.0008  
## TrtBin -0.4314286 0.1463974 9 -2.946969 0.0163  
## Correlation:   
## (Intr)  
## TrtBin -0.683  
##   
## Standardized Within-Group Residuals:  
## Min Q1 Med Q3 Max   
## -1.5033354 -0.3755643 -0.2234416 0.2972945 2.5651355   
##   
## Number of Observations: 20  
## Number of Groups: 10

DT <-array(c(1,8,0,1),  
 dim=c(2, 2),  
 dimnames=list(  
 Time2=c("ThenPos", "ThenNeg"),   
 Time1=c("FirstPos", "FirstNeg"))  
 )  
DT

## Time1  
## Time2 FirstPos FirstNeg  
## ThenPos 1 0  
## ThenNeg 8 1

mcnemar.test(DT, y=NULL, correct = FALSE)

##   
## McNemar's Chi-squared test  
##   
## data: DT  
## McNemar's chi-squared = 8, df = 1, p-value = 0.004678

chisq.test(EU1long$VegProp, EU1long$TrtBin)

## Warning in chisq.test(EU1long$VegProp, EU1long$TrtBin): Chi-squared  
## approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: EU1long$VegProp and EU1long$TrtBin  
## X-squared = 13.733, df = 5, p-value = 0.0174

fisher.test(EU1long$VegProp, EU1long$TrtBin)

##   
## Fisher's Exact Test for Count Data  
##   
## data: EU1long$VegProp and EU1long$TrtBin  
## p-value = 0.001093  
## alternative hypothesis: two.sided

## EU1 Repeated Measures Correlations

VC3 <- cor.test(EU1wide$VegPre, EU1wide$VegPost, method=c("pearson", "kendall", "spearman"))  
VC3

##   
## Pearson's product-moment correlation  
##   
## data: EU1wide$VegPre and EU1wide$VegPost  
## t = 0.18975, df = 8, p-value = 0.8542  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.5874476 0.6683936  
## sample estimates:  
## cor   
## 0.06693682

# Soil

## All Soil

S1 <- glm(SoilBin ~ Lineage + TrtBin, family = "binomial", data = AllSites)  
  
anova(S1)

## Analysis of Deviance Table  
##   
## Model: binomial, link: logit  
##   
## Response: SoilBin  
##   
## Terms added sequentially (first to last)  
##   
##   
## Df Deviance Resid. Df Resid. Dev  
## NULL 87 91.816  
## Lineage 1 2.85043 86 88.965  
## TrtBin 1 0.15705 85 88.808

summary(S1)

##   
## Call:  
## glm(formula = SoilBin ~ Lineage + TrtBin, family = "binomial",   
## data = AllSites)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.8831 -0.8076 -0.6092 -0.5524 1.9775   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -0.7404 0.4129 -1.793 0.073 .  
## LineageNA1 -0.8496 0.5347 -1.589 0.112   
## TrtBin -0.2127 0.5375 -0.396 0.692   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 91.816 on 87 degrees of freedom  
## Residual deviance: 88.808 on 85 degrees of freedom  
## AIC: 94.808  
##   
## Number of Fisher Scoring iterations: 4

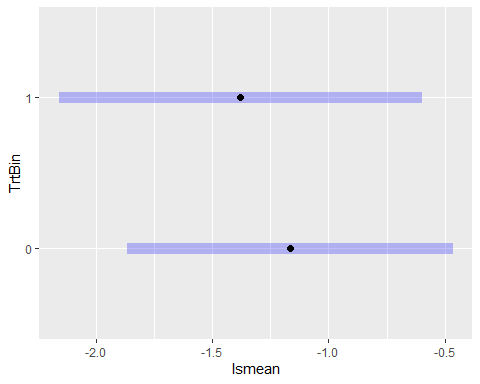
summ(S1)

## MODEL INFO:  
## Observations: 88  
## Dependent Variable: SoilBin  
## Type: Generalized linear model  
## Family: binomial   
## Link function: logit   
##   
## MODEL FIT:  
## <U+03C7>²(2) = 3.01, p = 0.22  
## Pseudo-R² (Cragg-Uhler) = 0.05  
## Pseudo-R² (McFadden) = 0.03  
## AIC = 94.81, BIC = 102.24   
##   
## Standard errors: MLE  
## ------------------------------------------------  
## Est. S.E. z val. p  
## ----------------- ------- ------ -------- ------  
## (Intercept) -0.74 0.41 -1.79 0.07  
## LineageNA1 -0.85 0.53 -1.59 0.11  
## TrtBin -0.21 0.54 -0.40 0.69  
## ------------------------------------------------

# Contrast statement  
S1.lsm <- lsmeans(S1, "TrtBin")  
S1.lsm

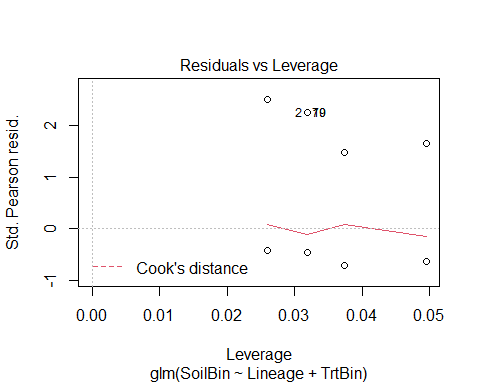
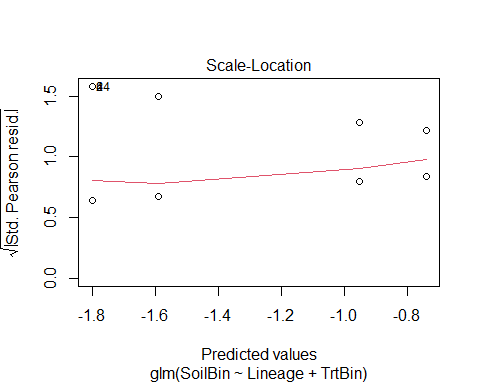
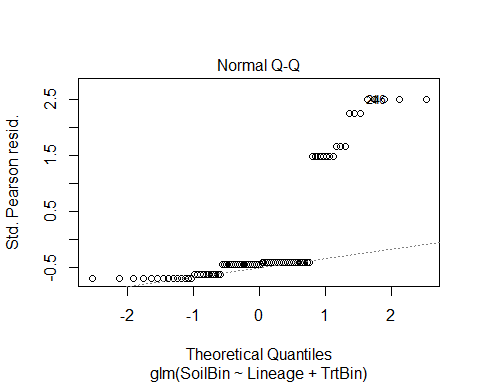
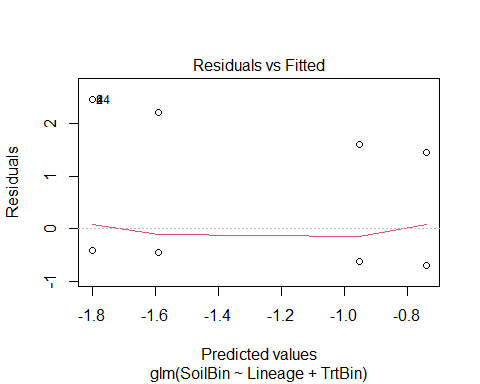
## TrtBin lsmean SE df asymp.LCL asymp.UCL  
## 0 -1.17 0.356 Inf -1.86 -0.467  
## 1 -1.38 0.398 Inf -2.16 -0.597  
##   
## Results are averaged over the levels of: Lineage   
## Results are given on the logit (not the response) scale.   
## Confidence level used: 0.95

plot(S1.lsm)



## All Soil Assumptions

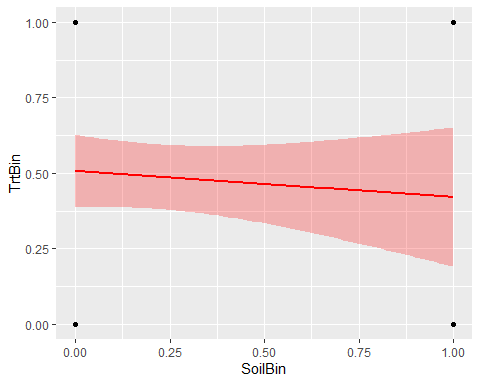
plot(S1)



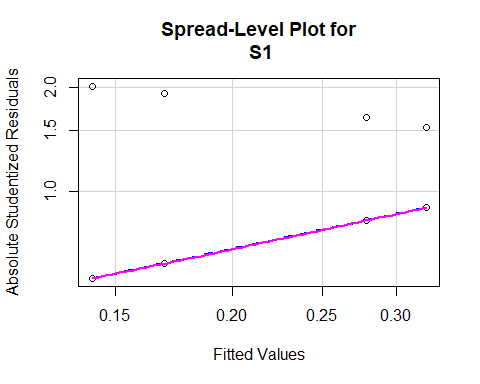
#linearity  
ggplot(AllSites, aes(SoilBin,TrtBin)) +   
 stat\_smooth(method="loess") +   
 stat\_smooth(method="lm", color="red", fill="red", alpha=.25) +   
 geom\_point()

## `geom\_smooth()` using formula 'y ~ x'

## `geom\_smooth()` using formula 'y ~ x'

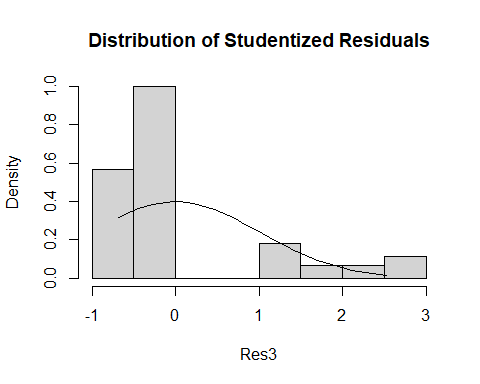


#constant varience  
spreadLevelPlot(S1)



##   
## Suggested power transformation: 0.4250135

#normality of residuals  
Res3 <- studres(S1)  
hist(Res3, freq = FALSE,  
 main="Distribution of Studentized Residuals")  
xfit3 <- seq(min(Res3), max(Res3), length = 40)  
yfit3 <- dnorm(xfit3)  
lines(xfit3, yfit3)



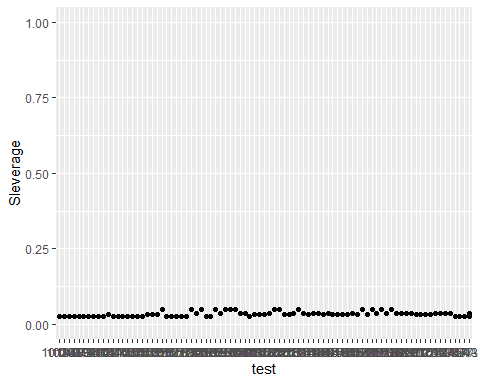
#assumption violated?  
  
#multicollinearity  
vif(S1)

## Lineage TrtBin   
## 1.024027 1.024027

sqrt(vif(S1)) > 2

## Lineage TrtBin   
## FALSE FALSE

#outliers  
AllSites$Sleverage <- hatvalues(S1)  
ggplot(AllSites, aes(Plot, Sleverage)) + geom\_point() + ylim(0,1) + xlab("test")



## All Soil Correlations

#testing correlations of random effects which were not included in the model  
SC1 <- cor.test(AllSites$SoilBin, AllSites$YrSinceTrt, method=c("pearson", "kendall", "spearman"))  
SC11 <- cor.test(AllSites$SoilBin, AllSites$TotalTO, method=c("pearson", "kendall", "spearman"))  
SC111 <- cor.test(AllSites$SoilBin, AllSites$Num\_Seedlings, method=c("pearson", "kendall", "spearman"))  
SC1111 <- cor.test(AllSites$SoilBin, AllSites$VegProp, method=c("pearson", "kendall", "spearman"))  
SC1

##   
## Pearson's product-moment correlation  
##   
## data: AllSites$SoilBin and AllSites$YrSinceTrt  
## t = -0.66715, df = 86, p-value = 0.5065  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.2770337 0.1397882  
## sample estimates:  
## cor   
## -0.07175484

SC11

##   
## Pearson's product-moment correlation  
##   
## data: AllSites$SoilBin and AllSites$TotalTO  
## t = -0.020186, df = 86, p-value = 0.9839  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.2115225 0.2073601  
## sample estimates:  
## cor   
## -0.002176686

SC111

##   
## Pearson's product-moment correlation  
##   
## data: AllSites$SoilBin and AllSites$Num\_Seedlings  
## t = -1.256, df = 86, p-value = 0.2125  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.33426265 0.07740132  
## sample estimates:  
## cor   
## -0.1342167

SC1111 #correlation between soil and veg!

##   
## Pearson's product-moment correlation  
##   
## data: AllSites$SoilBin and AllSites$VegProp  
## t = 2.6705, df = 86, p-value = 0.009055  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.0714214 0.4595330  
## sample estimates:  
## cor   
## 0.2767242

## NA1 Soil Wildfire

S2 <- glm(SoilBin ~ RX, data = AllFire)  
  
anova(S2)

## Analysis of Deviance Table  
##   
## Model: gaussian, link: identity  
##   
## Response: SoilBin  
##   
## Terms added sequentially (first to last)  
##   
##   
## Df Deviance Resid. Df Resid. Dev  
## NULL 51 6.7692  
## RX 3 1.2272 48 5.5420

summary(S2)

##   
## Call:  
## glm(formula = SoilBin ~ RX, data = AllFire)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.6000 -0.1333 -0.1304 0.0000 0.8696   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.130435 0.070852 1.841 0.07181 .   
## RX1 0.002899 0.112771 0.026 0.97960   
## RX2 0.469565 0.167666 2.801 0.00733 \*\*  
## RX3 -0.130435 0.133599 -0.976 0.33380   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for gaussian family taken to be 0.1154589)  
##   
## Null deviance: 6.7692 on 51 degrees of freedom  
## Residual deviance: 5.5420 on 48 degrees of freedom  
## AIC: 41.148  
##   
## Number of Fisher Scoring iterations: 2

summ(S2)

## MODEL INFO:  
## Observations: 52  
## Dependent Variable: SoilBin  
## Type: Linear regression   
##   
## MODEL FIT:  
## <U+03C7>²(3) = 1.23, p = 0.01  
## Pseudo-R² (Cragg-Uhler) = 0.33  
## Pseudo-R² (McFadden) = 0.25  
## AIC = 41.15, BIC = 50.90   
##   
## Standard errors: MLE  
## ------------------------------------------------  
## Est. S.E. t val. p  
## ----------------- ------- ------ -------- ------  
## (Intercept) 0.13 0.07 1.84 0.07  
## RX1 0.00 0.11 0.03 0.98  
## RX2 0.47 0.17 2.80 0.01  
## RX3 -0.13 0.13 -0.98 0.33  
## ------------------------------------------------  
##   
## Estimated dispersion parameter = 0.12

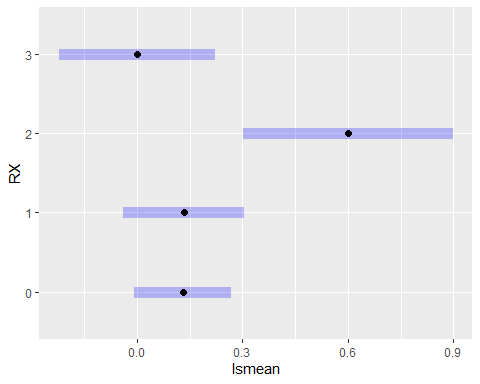
# Contrast statement  
levels(AllFire$RX)

## [1] "0" "1" "2" "3"

# 0 - no WF, no RX  
# 1 - no WF, yes RX  
# 2 - yes WF, no RX  
# 3 - yes WF, yes RX  
  
S2.lsm <- lsmeans(S2, ~RX, adjst="tukey")  
S2.lsm

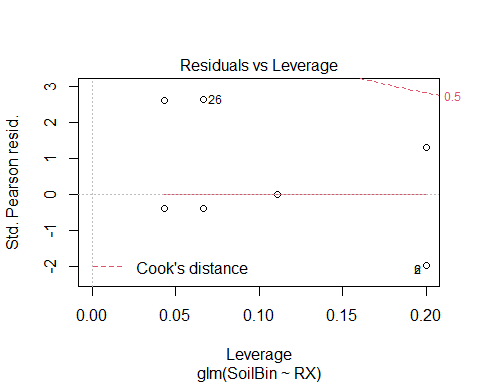
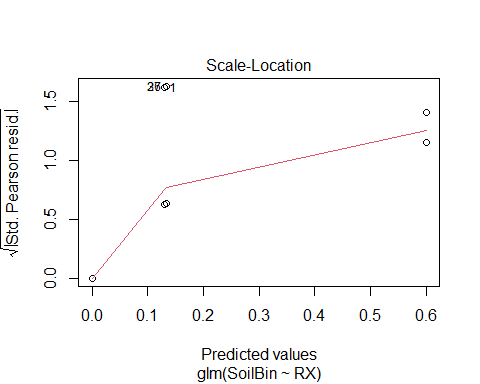
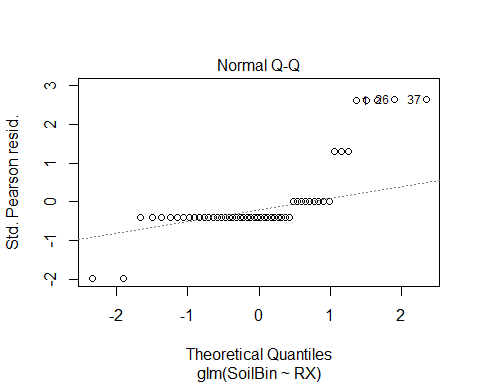
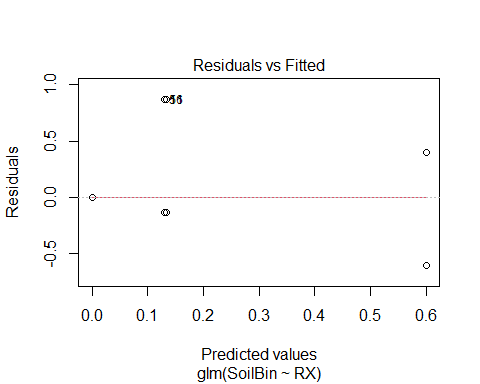
## RX lsmean SE df asymp.LCL asymp.UCL  
## 0 0.130 0.0709 Inf -0.00843 0.269  
## 1 0.133 0.0877 Inf -0.03862 0.305  
## 2 0.600 0.1520 Inf 0.30216 0.898  
## 3 0.000 0.1133 Inf -0.22199 0.222  
##   
## Confidence level used: 0.95

plot(S2.lsm)



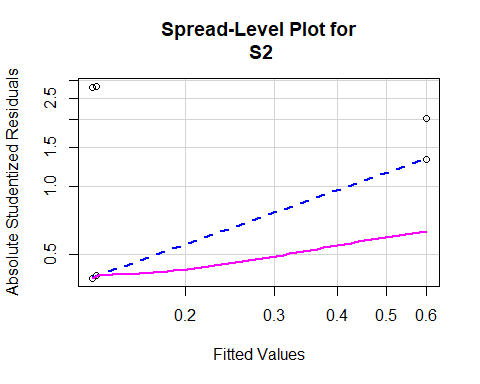
## NA1 Soil Wildfire Assumptions

plot(S2)



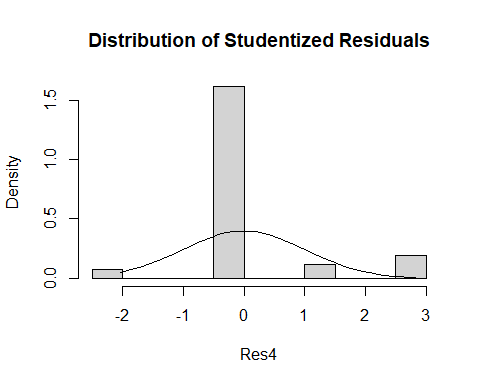
#constant variance  
spreadLevelPlot(S2)

## Warning in spreadLevelPlot.lm(S2):   
## 9 negative fitted values removed

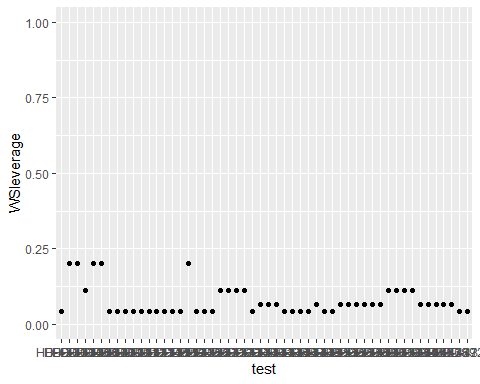


##   
## Suggested power transformation: 0.1937396

#normality of residuals  
Res4 <- studres(S2)  
hist(Res4, freq = FALSE,  
 main="Distribution of Studentized Residuals")  
xfit4 <- seq(min(Res4), max(Res4), length = 40)  
yfit4 <- dnorm(xfit4)  
lines(xfit4, yfit4)



#outliers  
AllFire$WSleverage <- hatvalues(S2)  
ggplot(AllFire, aes(`New Site`, WSleverage)) + geom\_point() + ylim(0,1) + xlab("test")



## NA1 Wildfire Correlations

SC2 <- cor.test(AllFire$VegProp, AllFire$TotalTO, method=c("pearson", "kendall", "spearman"))  
SC22 <- cor.test(AllFire$VegProp, AllFire$SoilProp, method=c("pearson", "kendall", "spearman"))  
  
SC2

##   
## Pearson's product-moment correlation  
##   
## data: AllFire$VegProp and AllFire$TotalTO  
## t = 1.093, df = 50, p-value = 0.2796  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.1253661 0.4086259  
## sample estimates:  
## cor   
## 0.1527605

SC22 #correlation between soil and veg!

##   
## Pearson's product-moment correlation  
##   
## data: AllFire$VegProp and AllFire$SoilProp  
## t = 3.2198, df = 50, p-value = 0.002256  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.1595480 0.6174743  
## sample estimates:  
## cor   
## 0.4144049

# Ordination of seedling data

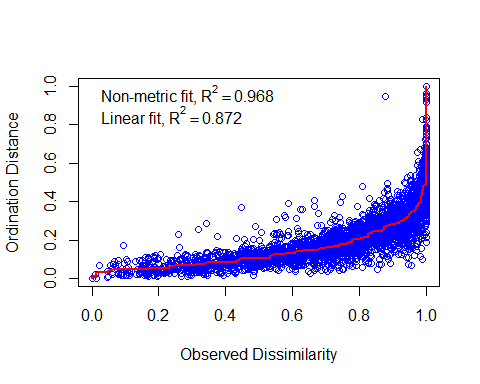
## NMDS

PCASeeds <- AllSeeds[5:35]  
DistSeed <- vegdist(PCASeeds, method = "bray")  
  
NMDS.scree <- function(x) { #where x is the name of the data frame variable  
 plot(rep(1, 10), replicate(10, metaMDS(x, autotransform = F, k = 1)$stress), xlim = c(1, 10),ylim = c(0, 0.30), xlab = "# of Dimensions", ylab = "Stress", main = "NMDS stress plot")  
 for (i in 1:10) {  
 points(rep(i + 1,10),replicate(10, metaMDS(x, autotransform = F, k = i + 1)$stress))  
 }  
}  
  
#NMDS.scree(DistSeed) chunk excluded to save space

NMDS1 <- metaMDS(DistSeed, k = 2, trymax = 100, trace = F)  
NMDS1

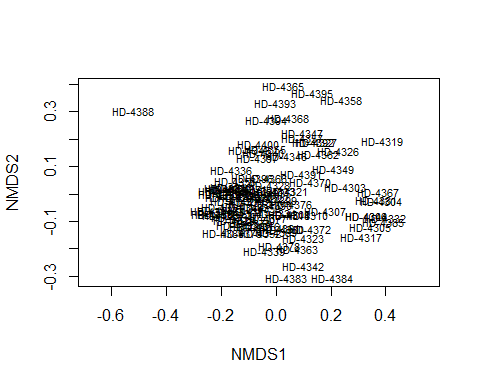
##   
## Call:  
## metaMDS(comm = DistSeed, k = 2, trymax = 100, trace = F)   
##   
## global Multidimensional Scaling using monoMDS  
##   
## Data: DistSeed   
## Distance: bray   
##   
## Dimensions: 2   
## Stress: 0.1791655   
## Stress type 1, weak ties  
## Two convergent solutions found after 20 tries  
## Scaling: centring, PC rotation   
## Species: scores missing

Stress <- stressplot(NMDS1)



Stressplot <- plot(NMDS1, type = "t")

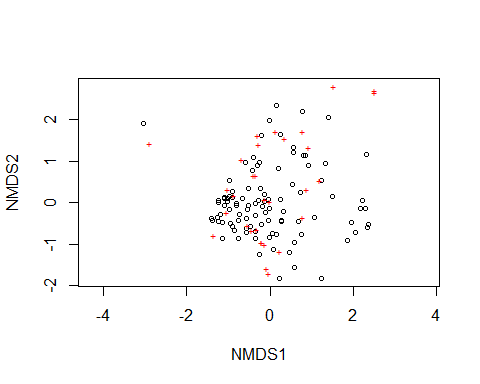
## species scores not available



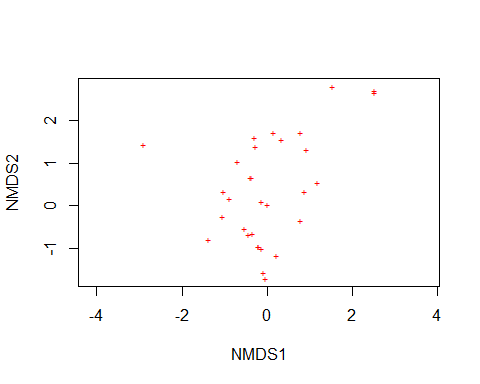
NMDS2 <- metaMDS(PCASeeds, k = 2, trymax = 100, trace = F)  
NMDS2

##   
## Call:  
## metaMDS(comm = PCASeeds, k = 2, trymax = 100, trace = F)   
##   
## global Multidimensional Scaling using monoMDS  
##   
## Data: wisconsin(sqrt(PCASeeds))   
## Distance: bray   
##   
## Dimensions: 2   
## Stress: 0.2403067   
## Stress type 1, weak ties  
## No convergent solutions - best solution after 100 tries  
## Scaling: centring, PC rotation, halfchange scaling   
## Species: expanded scores based on 'wisconsin(sqrt(PCASeeds))'

NMDS3 <- metaMDS(PCASeeds, k = 2, trymax = 100, trace = F, autotransform = FALSE, distance="bray")  
NMDS3plot <- plot(NMDS3)



NMDS3plotsp <- plot(NMDS3, display = "species")

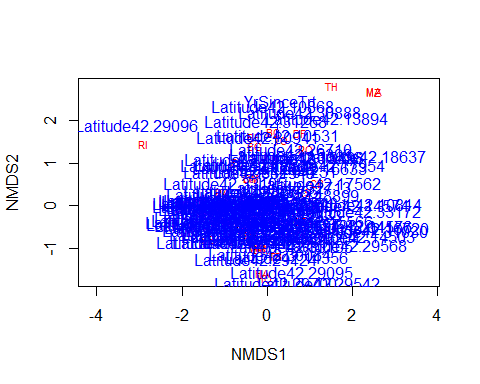


## Interpretation (including effects variables)

#Interpretation  
EFvars <- c("New Site", "Lineage", "Latitude", "Longitude", "VegBin", "SoilBin", "YrSinceTrt")  
EFSites <- SODSite.ord[EFvars]  
head(EFSites)

## # A tibble: 6 x 7  
## `New Site` Lineage Latitude Longitude VegBin SoilBin YrSinceTrt  
## <chr> <chr> <chr> <chr> <dbl> <dbl> <dbl>  
## 1 HD-3901 EU1 42.332505 -124.399209 1 1 0  
## 2 HD-4301 NA1 42.19457 -124.3507 1 1 0  
## 3 HD-4303 NA1 42.14574 -124.14850 0 0 1  
## 4 HD-4304 NA1 42.14574 -124.14850 1 1 1  
## 5 HD-4305 NA1 42.14503 -124.15112 0 0 1  
## 6 HD-4306 NA1 42.14556 -124.15621 1 1 1

ef <- envfit(NMDS3, EFSites, permutations = 999, na.rm = TRUE)  
  
efplot <- plot(NMDS3, type = "t", display = "species")  
efplot2 <- plot(ef, p.max = 0.05)

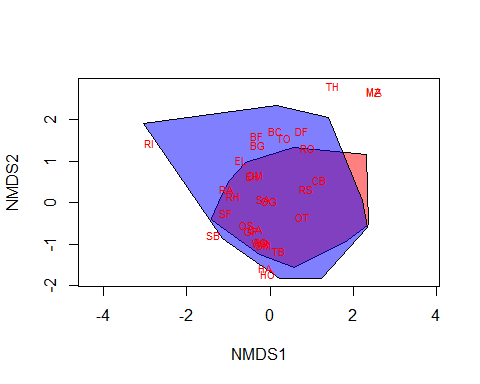


efplot2

## NULL

## Interpretation by treatment

group = c(rep("Untreated", 56), rep("Treated", 43))  
colors = c(rep("red", 56), rep("blue", 43))  
  
ordiplot(NMDS3, type = "n")  
for(i in unique(group)) {  
 ordihull(NMDS3$point[grep(i, group),], draw="polygon",  
 groups = group[group == i],col = colors[grep(i,group)],label=F) }  
  
orditorp(NMDS3, display = "species", col = "red", air = 0.01)



#orditorp(NMDS3, display = "sites", col = c(rep("red",12),  
# rep("blue", 12)), air = 0.01, cex = 1.25)

#MRPP or other analysis coming soon!