

Field Old and Young Leaf Water ANOVAs

Kenia E. Segura Aba

5/5/2020

This file was run in R version 3.5.3. The packages used are `tidyverse` version 1.3.0, `readr` version 1.3.1, `RRPP` version 0.4.2.9000, `mixOmics` version 6.6.2, and `labdsv` version 2.0-1. This file must be in the same directory as the Box sync folder in order to run. The following analysis of lipid metabolites was conducted using a split-split-plot analysis of variance (ANOVA) of Young and Old *P. virgatum* leaves using residual randomization permutation procedure (RRPP). Patterns in metabolite classification were visualized using `mixOmics` for principle component analysis (PCA) and partial least squares discriminant analysis (PLS-DA). Dufrene-Legendre indicator analysis was performed to identify specific metabolites indicative of plant response to water treatment and fungal treatment (`labdsv`).

1. Load necessary packages

```
library(tidyverse)
```

```
## -- Attaching packages -----
```

```
## v ggplot2 3.3.0      v purrr  0.3.4
## v tibble  3.0.1      v dplyr  0.8.5
## v tidyr   1.0.3      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.5.0
```

```
## -- Conflicts -----
```

```
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
library(readr)
```

```
library(RRPP)
```

```
library(mixOmics)
```

```
## Loading required package: MASS
```

```
##
```

```
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:dplyr':
```

```
##
```

```
##      select
```

```
## Loading required package: lattice
```

```
##
```

```
## Loaded mixOmics 6.12.0
```

```
## Thank you for using mixOmics!
```

```
## Tutorials: http://mixomics.org
```

```
## Bookdown vignette: https://mixomicsteam.github.io/Bookdown
```

```
## Questions, issues: Follow the prompts at http://mixomics.org/contact-us
```

```
## Cite us: citation('mixOmics')
```

```
##
## Attaching package: 'mixOmics'
## The following object is masked from 'package:purrr':
##
##      map
library(labdsv)

## Loading required package: mgcv
## Loading required package: nlme
##
## Attaching package: 'nlme'
## The following object is masked from 'package:dplyr':
##
##      collapse
## This is mgcv 1.8-31. For overview type 'help("mgcv-package")'.
## Registered S3 methods overwritten by 'labdsv':
##      method      from
##      summary.pca mixOmics
##      plot.pca     mixOmics
## This is labdsv 2.0-1
## convert existing ordinations with as.dsvord()
##
## Attaching package: 'labdsv'
## The following object is masked from 'package:mixOmics':
##
##      pca
## The following object is masked from 'package:stats':
##
##      density
```

Secondary Metabolites (Neg)

RRPP

2. Define dependent variable matrix and class matrix.

```
path <- "~/Box/Summer 2018 TX Endo Field Samples and Analysis/Statistics/Kenia_Thesis_Analysis/"
SM_neg <- read_tsv(paste(path,"XCMS Online Results/Pairwise Old vs Young/Pairwise_SM_Neg/XCMS.annotated

## Warning: Missing column names filled in: 'X1' [1]
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   name = col_character(),
##   updown = col_character(),
##   isotopes = col_character(),
##   adduct = col_logical()
## )
```

```
## See spec(...) for full column specifications.
# dependent variable: metabolite intensities
Y <- SM_neg[,c(9,23:108)] %>% data.frame(row.names=1) %>% t %>% data.frame()
scaled_Y <- scale(Y)

# class: sample factors
class <- read.csv(paste(path,"XCMS Online Results/pw_class.csv", sep=""), header = T, row.names = 1)

# define dependent variable matrices for each water treatment
high <- data.frame(cbind(Y,class)) %>% group_by(Water) %>% filter(Water=="High")
shigh <- scale(data.frame(high[,1:42]))
low <- data.frame(cbind(Y,class)) %>% group_by(Water) %>% filter(Water=="Low")
slow <- scale(low[,1:36])
```

3. Define and run multivariate regression models, then print out the results.

```
highLM <- lm.rrpp(shigh ~ Block * Fungus * Age, data = high, SS.type = "III", print.progress = F); summary
```

```
##
## Linear Model fit with lm.rrpp
##
## Number of observations: 40
## Number of dependent variables: 42
## Data space dimensions: 39
## Sums of Squares and Cross-products: Type III
## Number of permutations: 1000
##
## Full Model Analysis of Variance
##
##              Df Residual Df      SS Residual SS      Rsq      F
## Block * Fungus * Age  7          32 1232.285    405.7145 0.752311 13.8849
##              Z (from F)      Pr(>F)
## Block * Fungus * Age  7.048247 0.0005714286
##
##
## Redundancy Analysis (PCA on fitted values and residuals)
##
##              Trace Proportion Rank
## Fitted      31.59706  0.752311    7
## Residuals  10.40294  0.247689   32
## Total      42.00000  1.000000   39
##
## Eigenvalues
##
##              PC1      PC2      PC3      PC4      PC5      PC6      PC7
## Fitted      28.012547 1.683203 1.169873 0.293019 0.250454 0.112612 0.075355
## Residuals   2.524440 1.651228 1.151526 0.727367 0.567341 0.558263 0.440477
## Total      28.892001 2.951074 2.307701 1.745361 1.082585 0.602674 0.584975
##              PC8      PC9      PC10     PC11     PC12     PC13     PC14
## Fitted
## Residuals  0.409370 0.322615 0.289139 0.271689 0.239725 0.222207 0.168993
## Total      0.514725 0.424674 0.362244 0.351511 0.307229 0.274266 0.243421
##              PC15     PC16     PC17     PC18     PC19     PC20     PC21
## Fitted
```

```
## Residuals 0.147154 0.119771 0.107353 0.088959 0.082827 0.064484 0.049986
## Total 0.186895 0.164913 0.157192 0.143584 0.117081 0.109981 0.084363
## PC22 PC23 PC24 PC25 PC26 PC27 PC28
## Fitted
## Residuals 0.045113 0.036122 0.029325 0.026990 0.021241 0.014106 0.009124
## Total 0.075983 0.067298 0.051246 0.037312 0.036049 0.029930 0.024344
## PC29 PC30 PC31 PC32 PC33 PC34 PC35
## Fitted
## Residuals 0.007940 0.004725 0.002670 0.000663
## Total 0.018868 0.014910 0.010832 0.008293 0.006890 0.004194 0.002890
## PC36 PC37 PC38 PC39
## Fitted
## Residuals
## Total 0.001487 0.000674 0.000236 0.000113
```

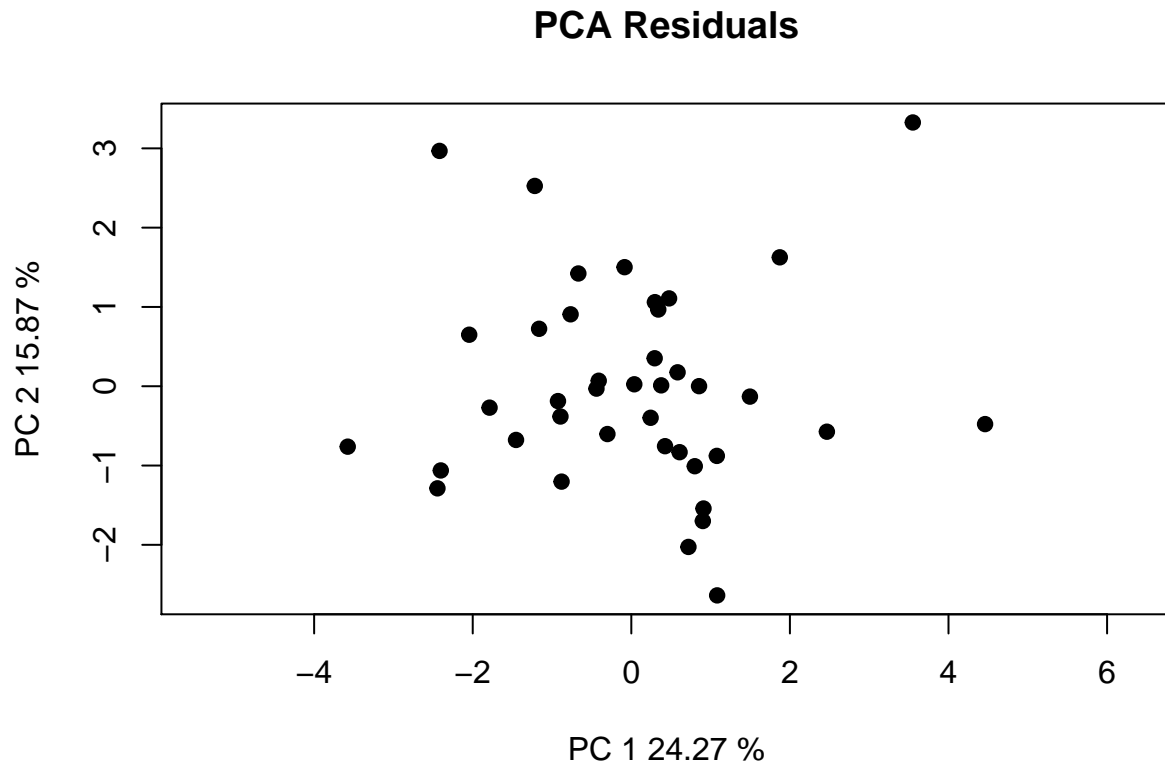
```
lowLM <- lm.rrpp(slow ~ Block * Fungus * Age, data = low, SS.type = "III", print.progress = F) ; summary
```

```
##
## Linear Model fit with lm.rrpp
##
## Number of observations: 46
## Number of dependent variables: 36
## Data space dimensions: 36
## Sums of Squares and Cross-products: Type III
## Number of permutations: 1000
##
## Full Model Analysis of Variance
##
## Df Residual Df SS Residual SS Rsq F
## Block * Fungus * Age 7 38 1376.215 243.7847 0.8495156 30.64542
## Z (from F) Pr(>F)
## Block * Fungus * Age 7.104153 0.0005714286
##
##
## Redundancy Analysis (PCA on fitted values and residuals)
##
## Trace Proportion Rank
## Fitted 30.58256 0.8495156 7
## Residuals 5.41744 0.1504844 36
## Total 36.00000 1.0000000 36
##
## Eigenvalues
##
## PC1 PC2 PC3 PC4 PC5 PC6 PC7
## Fitted 29.797871 0.372564 0.163081 0.127522 0.055156 0.040756 0.025613
## Residuals 1.745070 0.636750 0.531043 0.374811 0.329799 0.278353 0.246424
## Total 30.566142 1.314627 0.738552 0.585819 0.390087 0.365994 0.296074
## PC8 PC9 PC10 PC11 PC12 PC13 PC14
## Fitted
## Residuals 0.200664 0.170020 0.140925 0.123163 0.090950 0.089432 0.074957
## Total 0.266647 0.234843 0.186902 0.166514 0.138291 0.115327 0.099709
## PC15 PC16 PC17 PC18 PC19 PC20 PC21
## Fitted
## Residuals 0.065204 0.056145 0.043930 0.035723 0.033297 0.031437 0.027277
## Total 0.085500 0.074805 0.062057 0.052233 0.043223 0.037251 0.032847
```

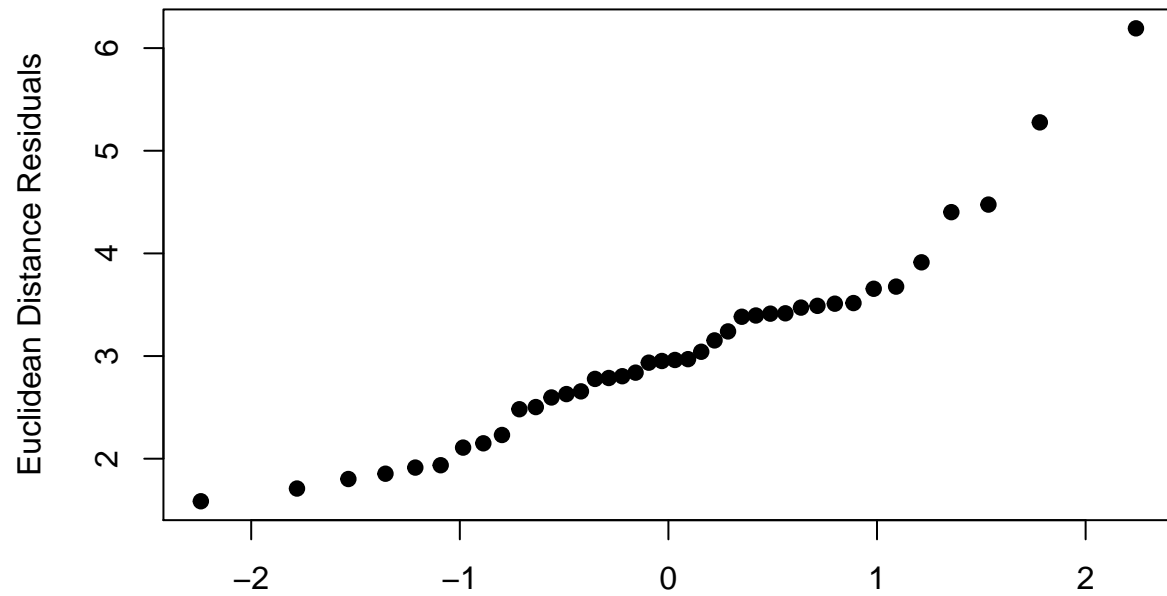
```
##          PC22      PC23      PC24      PC25      PC26      PC27      PC28
## Fitted
## Residuals 0.020364 0.019020 0.013295 0.012309 0.008317 0.007014 0.005325
## Total     0.027553 0.025772 0.020312 0.018253 0.014986 0.011428 0.009957
##          PC29      PC30      PC31      PC32      PC33      PC34      PC35
## Fitted
## Residuals 0.003226 0.001894 0.000515 0.000389 0.000281 0.000101 0.000013
## Total     0.006799 0.004269 0.003275 0.002460 0.000820 0.000463 0.000162
##          PC36
## Fitted
## Residuals 0.000003
## Total     0.000046
```

4. Examine RRPP plots to check for assumptions.

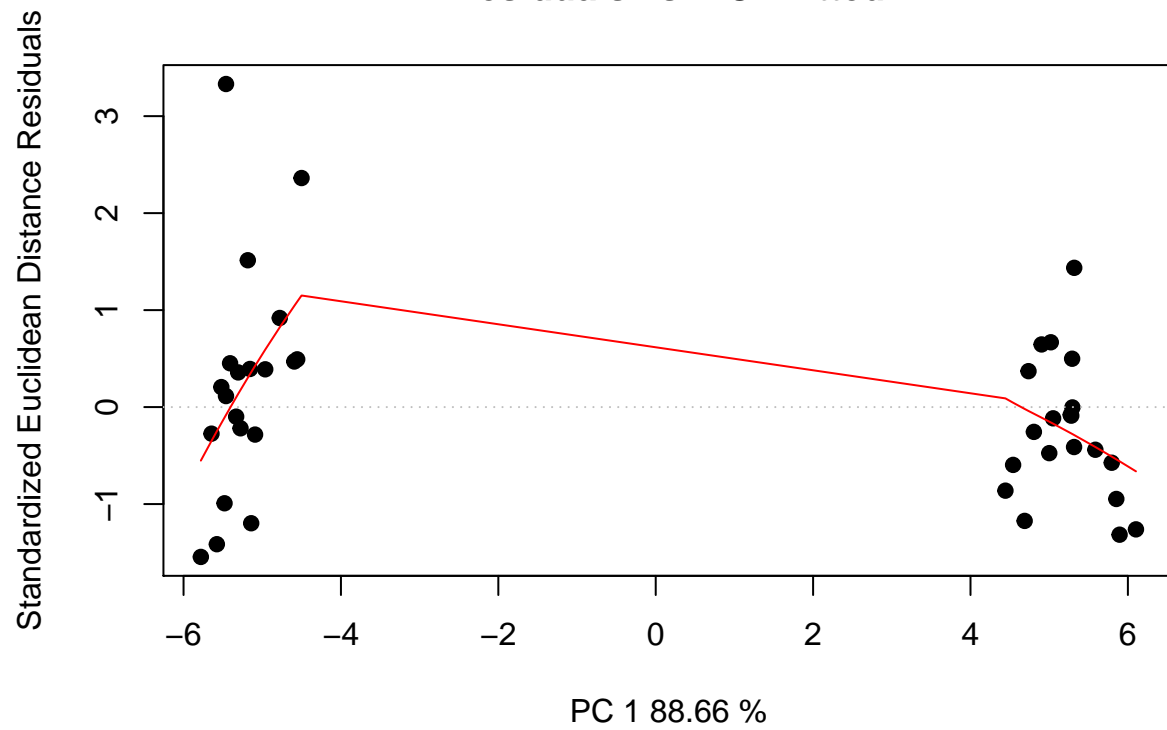
```
## High
# residuals vs fitted values (homoscedasticity check)
diagnostics <- plot(highLM, type = "diagnostics")
```

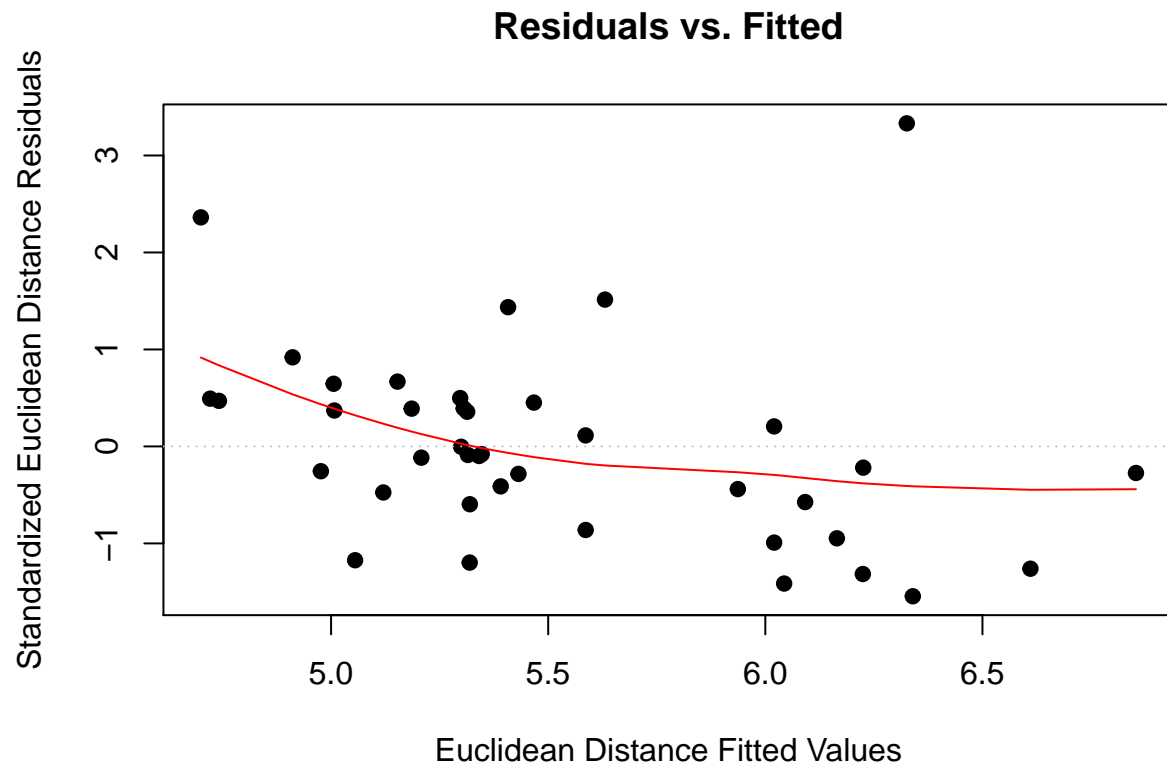


Q-Q plot

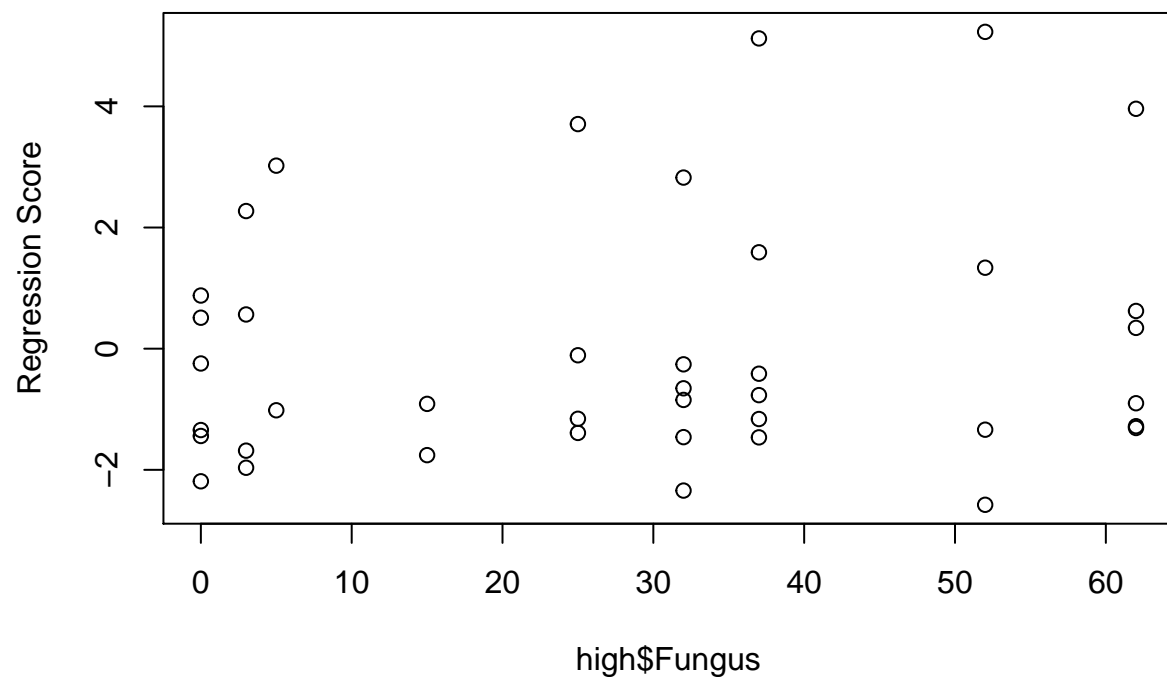


Theoretical Quantiles
Residuals vs. PC 1 fitted

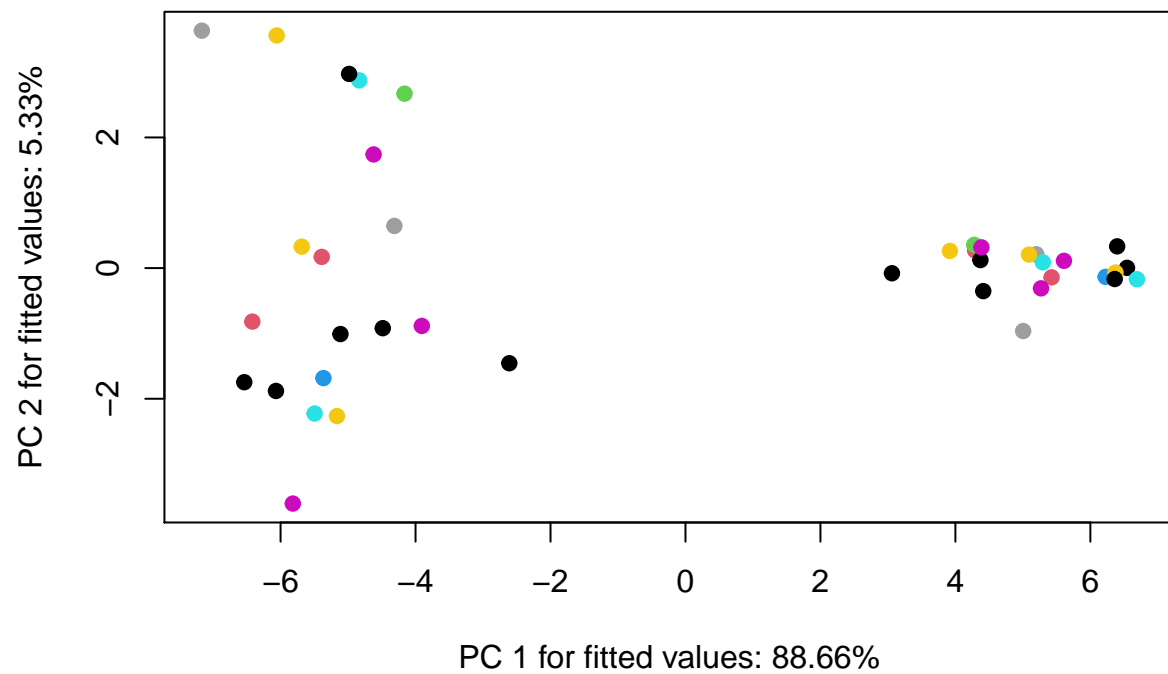




```
# linear regression plot
regression <- plot(highLM, type = "regression", predictor = high$Fungus, reg.type = "RegScore")
```

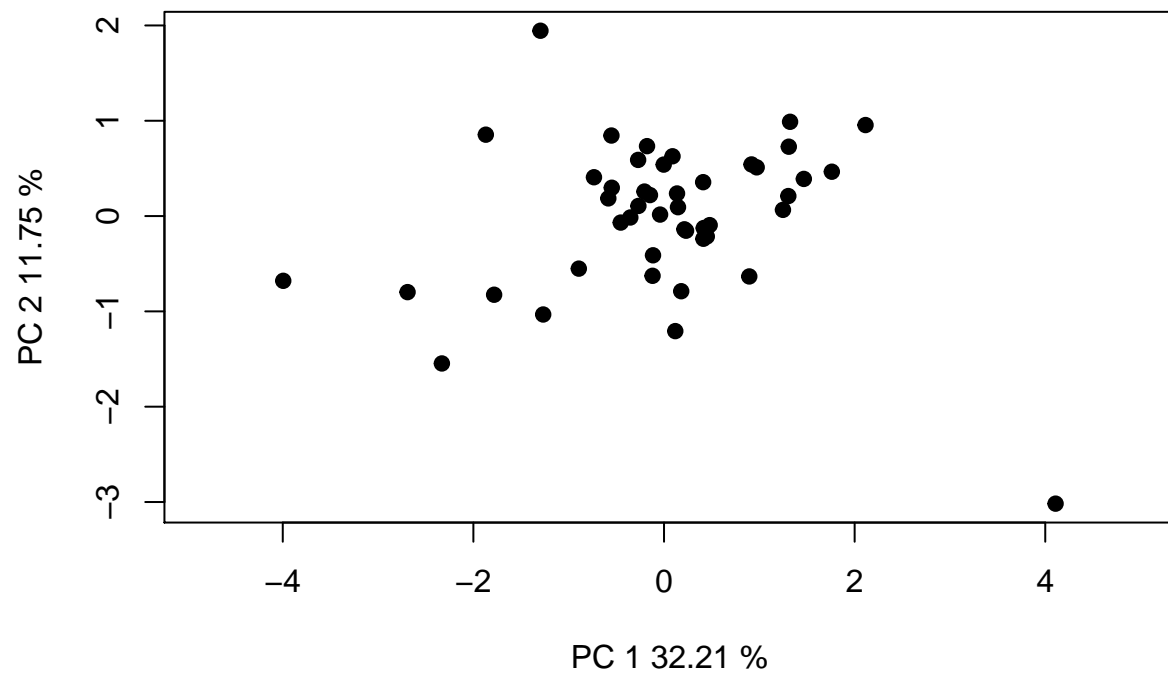


```
# pca plot
pcplot <- plot(highLM, type = "PC", pch = 19, col = interaction(high$Water, high$Fungus))
```

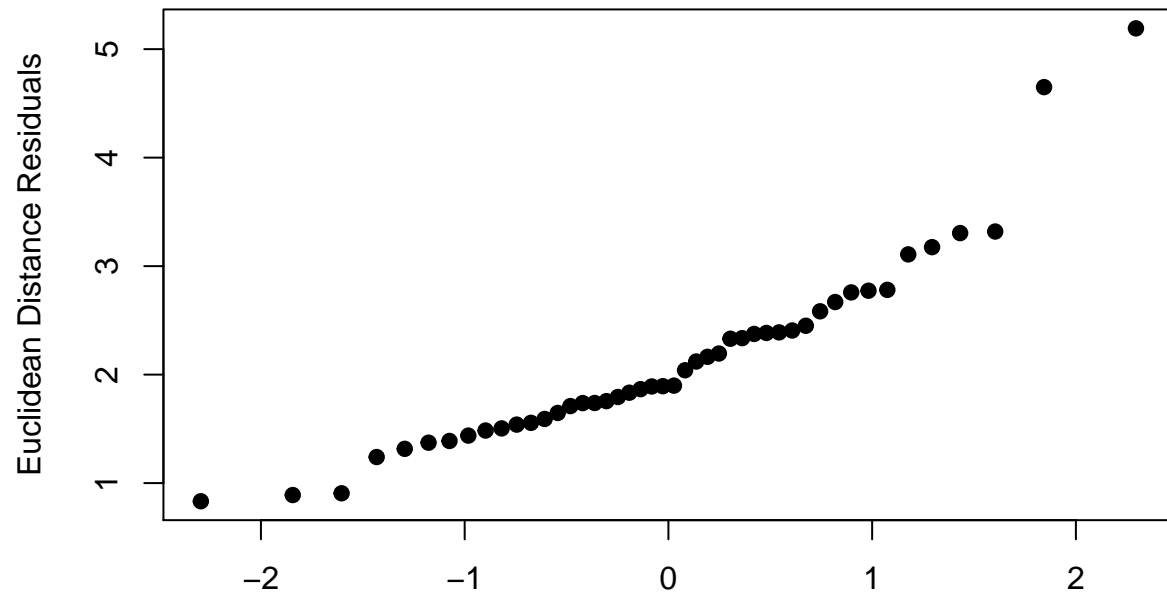


```
## Low
# residuals vs fitted values (homoscedasticity check)
diagnostics <- plot(lowLM, type = "diagnostics")
```

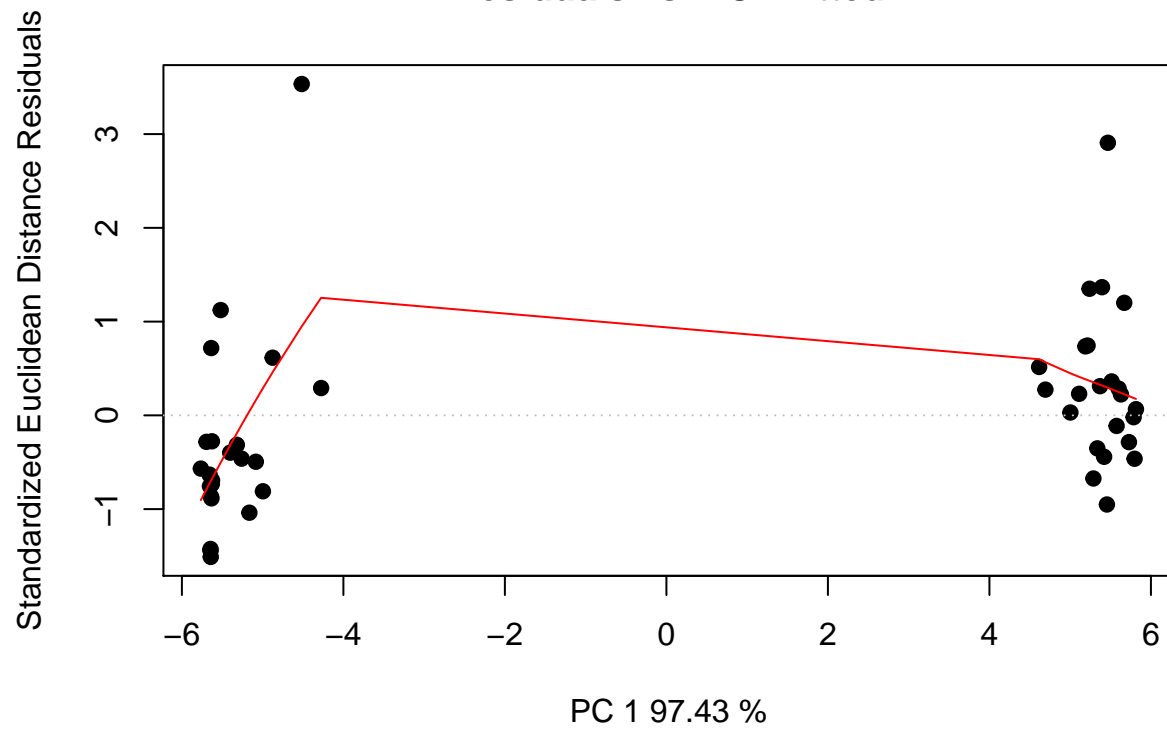
PCA Residuals

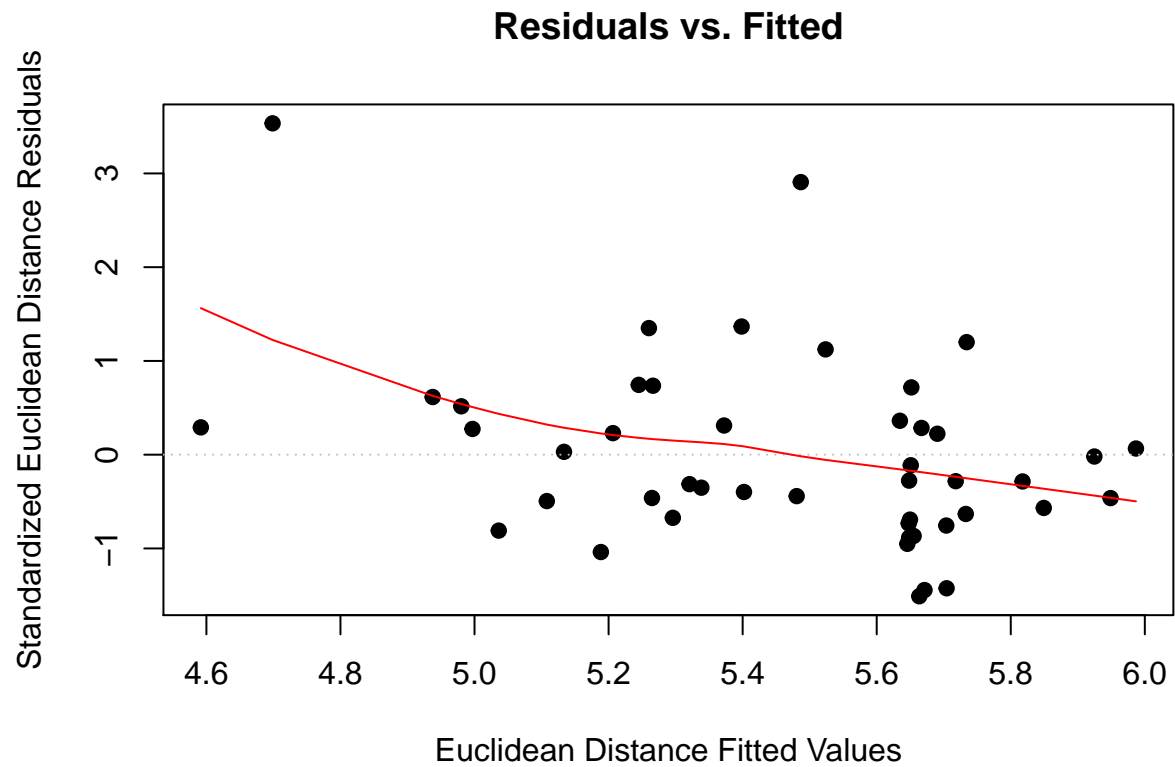


Q-Q plot

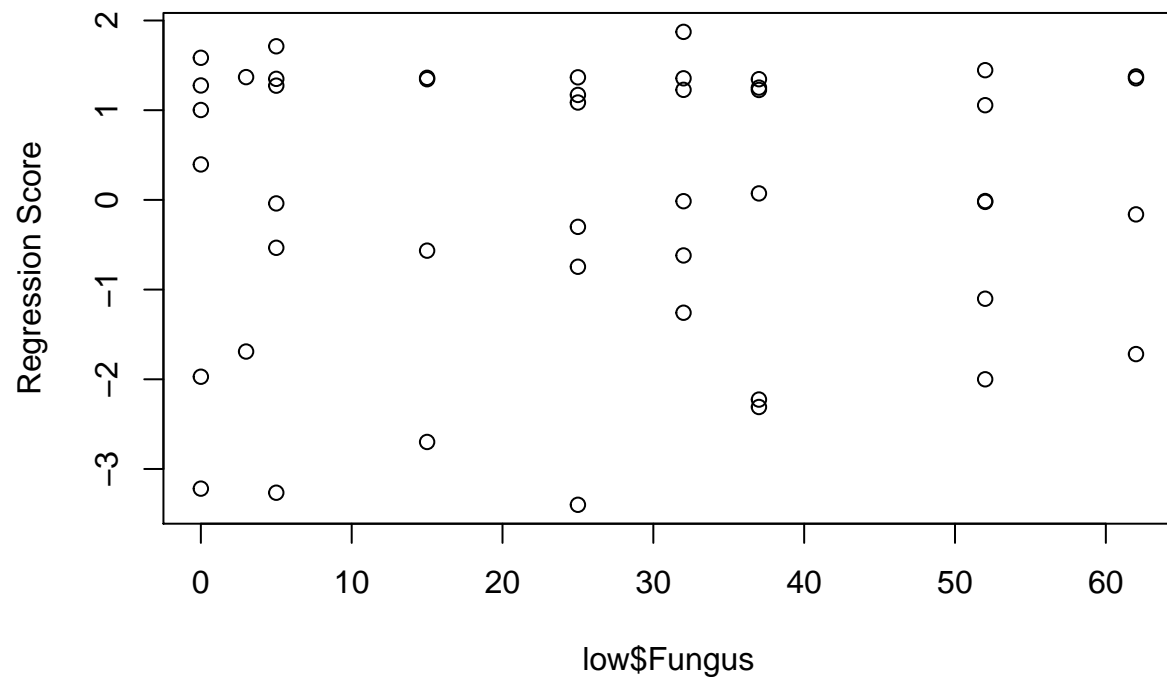


Theoretical Quantiles
Residuals vs. PC 1 fitted

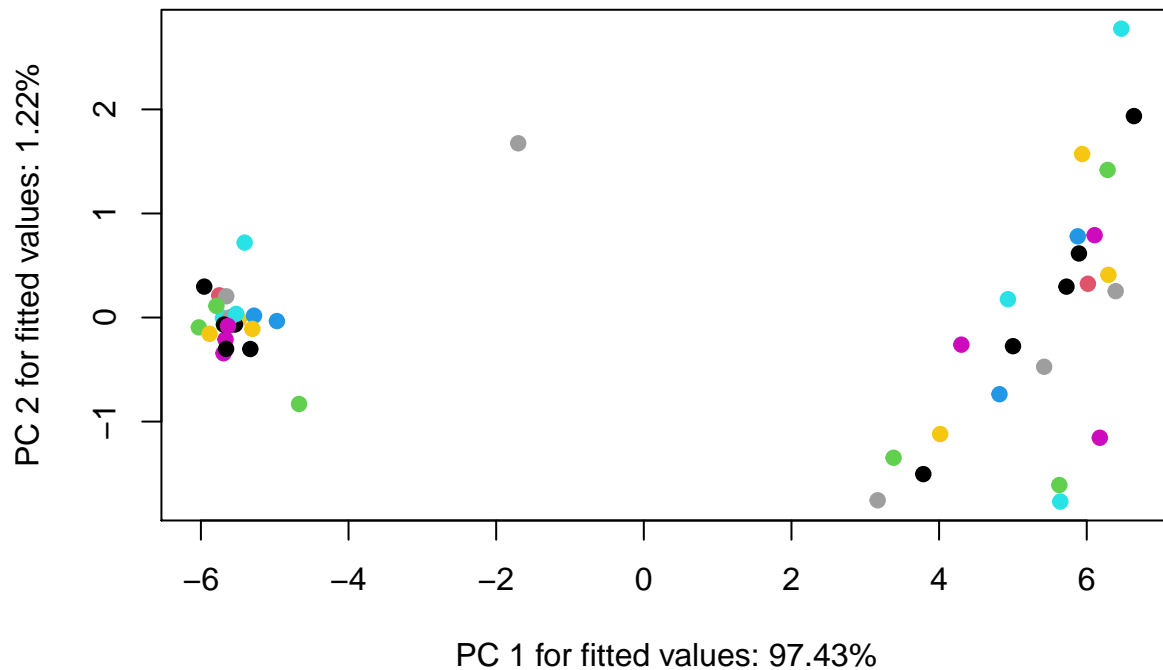




```
# linear regression plot
regression <- plot(lowLM, type = "regression", predictor = low$Fungus, reg.type = "RegScore")
```



```
# pca plot
pcplot <- plot(lowLM, type = "PC", pch = 19, col = interaction(low$Water, low$Fungus))
```



5. Perform an RRPP ANOVA and print results.

```
## High
highANOVA <- anova(highLM, effect.type = "F", error = c("Residuals", "Block:Fungus", "Block:Fungus:Age"))

##
## Analysis of Variance, using Residual Randomization
## Permutation procedure: Randomization of null model residuals
## Number of permutations: 1000
## Estimation method: Ordinary Least Squares
## Sums of Squares and Cross-products: Type III
## Effect sizes (Z) based on F distributions
##
##          Df      SS      MS      Rsq      F      Z Pr(>F)
## Block          1   12.52 12.517 0.007642 0.9873 0.13270 0.451
## Fungus          1   22.36 22.361 0.013652 1.4136 1.29729 0.090 .
## Age             1   50.86 50.862 0.031051 4.3302 3.15824 0.003 **
## Block:Fungus    1   15.82 15.819 0.009657 1.2477 0.69129 0.247
## Block:Age       1    6.97  6.966 0.004253 0.5931 -1.46098 0.933
## Fungus:Age      1   18.58 18.575 0.011340 1.5814 1.83643 0.023 *
## Block:Fungus:Age 1   11.75 11.746 0.007171 0.9264 0.01472 0.485
## Residuals      32 405.71 12.679 0.247689
## Total          39 1638.00
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Call: lm.rrpp(f1 = shigh ~ Block * Fungus * Age, SS.type = "III", data = high,
##   print.progress = F)

## Low
lowANOVA <- anova(lowLM, effect.type = "F", error = c("Residuals", "Block:Fungus", "Block:Fungus:Age",
```

```
##
## Analysis of Variance, using Residual Randomization
```

```
## Permutation procedure: Randomization of null model residuals
## Number of permutations: 1000
## Estimation method: Ordinary Least Squares
## Sums of Squares and Cross-products: Type III
## Effect sizes (Z) based on F distributions
##
##           Df      SS      MS      Rsq      F      Z Pr(>F)
## Block           1    15.87   15.869 0.009796   2.4737  1.9728  0.046 *
## Fungus           1     7.37    7.367 0.004548   0.9881  0.0114  0.514
## Age             1   105.28  105.279 0.064987  20.3426  4.6811  0.001 **
## Block:Fungus     1     7.46    7.456 0.004602   1.1622  0.4793  0.284
## Block:Age        1     8.15    8.146 0.005029   1.5741  1.0905  0.141
## Fungus:Age       1     4.59    4.592 0.002834   0.8873 -0.3714  0.654
## Block:Fungus:Age 1     5.18    5.175 0.003195   0.8067 -0.2995  0.604
## Residuals       38   243.78    6.415 0.150484
## Total           45  1620.00
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Call:  lm.rrpp(f1 = slow ~ Block * Fungus * Age, SS.type = "III", data = low,
##             print.progress = F)
```

6. Test lm.rrpp model coefficients. “d” is the amount of change in a variable for the coefficient indicated.

```
## High
# test model coefficients
highcoef <- coef(highLM, test = T) ; summary(highcoef)

##
## Linear Model fit with lm.rrpp
##
## Number of observations: 40
## Number of dependent variables: 42
## Data space dimensions: 39
## Sums of Squares and Cross-products: Type III
## Number of permutations: 1000
##
## Statistics (distances) of coefficients with 95 percent confidence intervals,
## effect sizes, and probabilities of exceeding observed values based on
## 1000 random permutations using RRPP
##
##           d.obs  UCL (95%)      Zd Pr(>d)
## (Intercept)    4.64544136  6.77057901 -1.7759641  0.977
## Block          1.68708791  2.13000672  0.4493312  0.292
## Fungus          0.12419106  0.11965559  2.0515760  0.034
## AgeYoung       9.02192022  5.84965839  5.2247895  0.001
## Block:Fungus    0.05356759  0.06069471  1.0602147  0.149
## Block:AgeYoung  1.77994302  2.96640098 -0.8301463  0.802
## Fungus:AgeYoung 0.16007499  0.16451790  1.7089160  0.059
## Block:Fungus:AgeYoung 0.06527915  0.08306624  0.3395585  0.327

## Low
# test model coefficients
lowcoef <- coef(lowLM, test = T) ; summary(lowcoef)

##
```

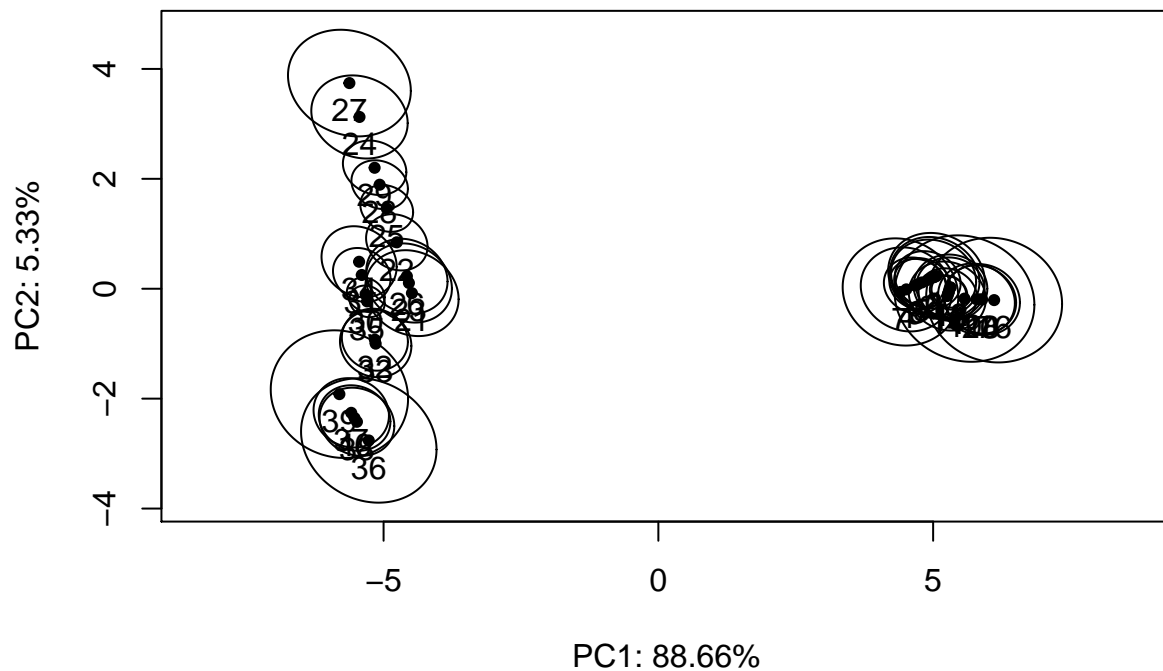
```
## Linear Model fit with lm.rrpp
##
## Number of observations: 46
## Number of dependent variables: 36
## Data space dimensions: 36
## Sums of Squares and Cross-products: Type III
## Number of permutations: 1000
##
## Statistics (distances) of coefficients with 95 percent confidence intervals,
## effect sizes, and probabilities of exceeding observed values based on
## 1000 random permutations using RRPP
##
##           d.obs  UCL (95%)      Zd Pr(>d)
## (Intercept)    7.05682799  8.53559314 -0.07468424  0.515
## Block          1.67705345  1.51479296  2.62012275  0.026
## Fungus         0.06854513  0.09016382  0.66918271  0.198
## AgeYoung      12.34545346  4.99371361  9.74319837  0.001
## Block:Fungus   0.03269716  0.04192239  0.66190146  0.203
## Block:AgeYoung 1.69926748  1.98153932  0.95211214  0.159
## Fungus:AgeYoung 0.07653146  0.11559563 -0.30619592  0.568
## Block:Fungus:AgeYoung 0.03852538  0.05618653 -0.08807986  0.491
```

Fungus has the largest effect on the model for young leaves. The standard is the mean for High water treatment. Block has the largest effect on the model for old leaves. Fungus coming in second.

7. Compute predicted values from the lm.rrpp model fit using bootstrapped residuals to generate confidence intervals (precision of group mean estimates).

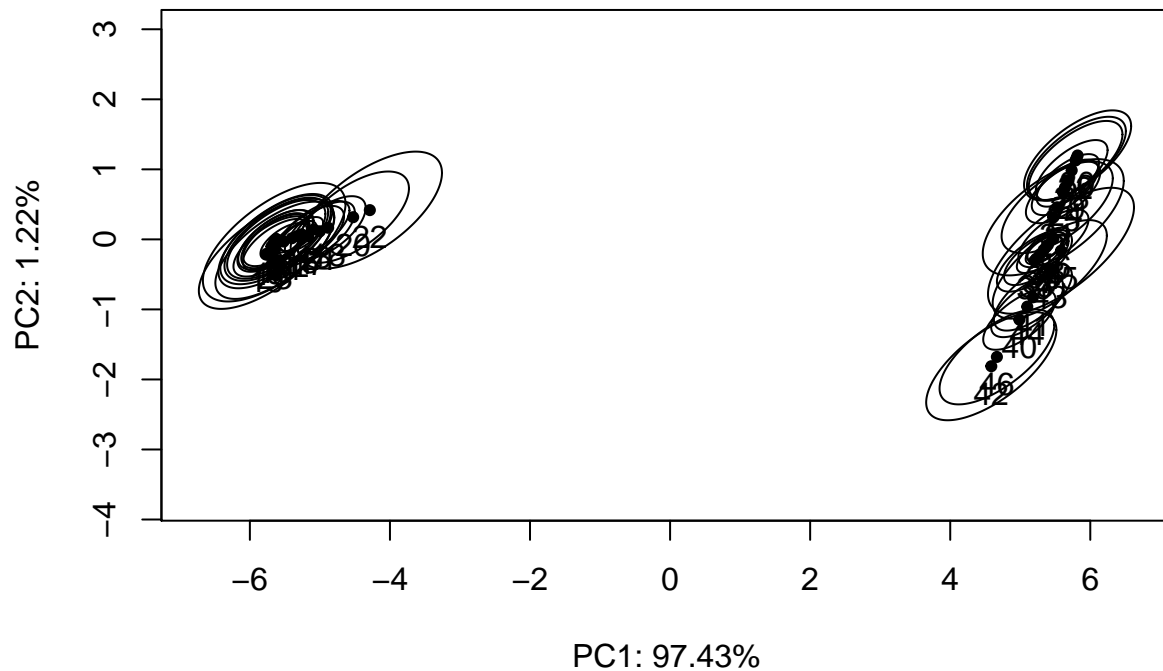
```
high_pred <- predict(highLM) ; plot(high_pred, PC = T, ellipse = T)
```

Among-prediction PC rotation; 95% confidence limits



```
low_pred <- predict(lowLM) ; plot(low_pred, PC = T, ellipse = T)
```

Among-prediction PC rotation; 95% confidence limits



8. Test pairwise differences between least squares means. Similar to `tukeyHSD` function in the `r stats` package. The `pairwise` function will generate tables with confidence intervals and p-values for the pairwise statistic, Euclidean distance between least-squares means.

```
# pairwise differences of fungus
highpw <- pairwise(highLM, groups = high$Fungus); summary(highpw, confidence = 0.95, stat.table = T)
```

```
##
## Pairwise comparisons
##
## Groups: 0 3 5 15 25 32 37 52 62
##
## RRPP: 1000 permutations
##
## LS means:
## Vectors hidden (use show.vectors = TRUE to view)
##
## Pairwise distances between means, plus statistics
##
```

| | d | UCL (95%) | Z | Pr > d |
|------|-----------|-----------|------------|--------|
| 0:3 | 0.6487654 | 1.0279415 | -1.7430693 | 0.974 |
| 0:5 | 1.2724804 | 1.9800080 | -1.6937319 | 0.970 |
| 0:15 | 1.4276935 | 2.0432325 | -1.3370105 | 0.917 |
| 0:25 | 0.6370393 | 1.1069173 | -1.8266206 | 0.974 |
| 0:32 | 0.8154103 | 1.4168541 | -1.8266206 | 0.974 |
| 0:37 | 0.9428182 | 1.6382375 | -1.8266206 | 0.974 |
| 0:52 | 1.9810164 | 2.9545804 | -1.0574236 | 0.848 |
| 0:62 | 1.5798575 | 2.7451548 | -1.8266206 | 0.974 |
| 3:5 | 0.6256567 | 0.9508451 | -1.6363622 | 0.962 |

```
## 3:15 0.8680976 1.1797757 -0.9520427 0.825
## 3:25 0.6638594 1.1522283 -1.7202650 0.973
## 3:32 0.7692220 1.3650807 -1.7242678 0.972
## 3:37 0.8593680 1.5353672 -1.7310537 0.975
## 3:52 1.7279186 2.6984536 -1.0380548 0.846
## 3:62 1.3998395 2.5367485 -1.7656427 0.975
## 5:15 0.5058106 0.7732454 -0.8909657 0.816
## 5:25 1.1234287 1.8037749 -1.7305529 0.971
## 5:32 1.1434796 1.8897376 -1.7421193 0.970
## 5:37 1.1742960 1.9731927 -1.7477790 0.970
## 5:52 1.6575038 2.7680614 -1.2025433 0.897
## 5:62 1.4910002 2.6489333 -1.7660378 0.974
## 15:25 1.1133770 1.5951205 -1.4329479 0.937
## 15:32 1.0776910 1.6025105 -1.5400689 0.950
## 15:37 1.0698253 1.6302964 -1.6256786 0.958
## 15:52 1.1929268 2.0309329 -1.3422607 0.920
## 15:62 1.2438048 2.1816207 -1.9216665 0.986
## 25:32 0.1783710 0.3099368 -1.8266206 0.974
## 25:37 0.3057789 0.5313203 -1.8266206 0.974
## 25:52 1.4376881 1.9906295 -0.7844551 0.787
## 25:62 0.9428182 1.6382375 -1.8266206 0.974
## 32:37 0.1274079 0.2213835 -1.8266206 0.974
## 32:52 1.3008838 1.7508842 -0.7024267 0.755
## 32:62 0.7644472 1.3283007 -1.8266206 0.974
## 37:52 1.2098146 1.5981244 -0.6541664 0.751
## 37:62 0.6370393 1.1069173 -1.8266206 0.974
## 52:62 0.8985529 1.3291502 -0.9532062 0.826
```

```
# pairwise differences of age
```

```
highpw2 <- pairwise(highLM, groups = high$Age) ; summary(highpw2, confidence = 0.95, stat.table = T)
```

```
##
## Pairwise comparisons
##
## Groups: Old Young
##
## RRPP: 1000 permutations
##
## LS means:
## Vectors hidden (use show.vectors = TRUE to view)
##
## Pairwise distances between means, plus statistics
##           d UCL (95%)           Z Pr > d
## Old:Young 10.42093 10.99091 -0.2300221 0.594
```

```
# pairwise differences of fungus
```

```
lowpw <- pairwise(lowLM, groups = low$Fungus); summary(lowpw, confidence = 0.95, stat.table = T)
```

```
##
## Pairwise comparisons
##
## Groups: 0 3 5 15 25 32 37 52 62
##
## RRPP: 1000 permutations
##
## LS means:
```

```

## Vectors hidden (use show.vectors = TRUE to view)
##
## Pairwise distances between means, plus statistics
##          d UCL (95%)          Z Pr > d
## 0:3    0.89759560 1.4916107 -0.9465119 0.825
## 0:5    0.07937776 0.1658745 -1.5036437 0.978
## 0:15   0.47953810 0.8214114 -1.0834297 0.875
## 0:25   0.39688878 0.8293727 -1.5036437 0.978
## 0:32   0.50801764 1.0615971 -1.5036437 0.978
## 0:37   0.58739539 1.2274716 -1.5036437 0.978
## 0:52   0.82552866 1.7250952 -1.5036437 0.978
## 0:62   0.98428417 2.0568443 -1.5036437 0.978
## 3:5    0.86229491 1.4390573 -0.9897130 0.840
## 3:15   0.49872366 0.8999841 -1.2197671 0.902
## 3:25   0.78807845 1.4033095 -1.3541174 0.934
## 3:32   0.79109557 1.4416361 -1.5273370 0.960
## 3:37   0.80271879 1.4935217 -1.6485043 0.968
## 3:52   0.88064926 1.6995109 -1.9275639 0.983
## 3:62   0.96229336 1.9058521 -2.0192244 0.989
## 5:15   0.42212520 0.6992318 -1.0524859 0.867
## 5:25   0.31751102 0.6634982 -1.5036437 0.978
## 5:32   0.42863988 0.8957225 -1.5036437 0.978
## 5:37   0.50801764 1.0615971 -1.5036437 0.978
## 5:52   0.74615090 1.5592207 -1.5036437 0.978
## 5:62   0.90490641 1.8909698 -1.5036437 0.978
## 15:25  0.31166945 0.5299654 -1.5901310 0.949
## 15:32  0.34118024 0.6201203 -1.8583908 0.975
## 15:37  0.38116592 0.7314209 -1.9233423 0.979
## 15:52  0.55457000 1.1272527 -1.8652459 0.986
## 15:62  0.69190302 1.4389922 -1.8038380 0.985
## 25:32  0.11112886 0.2322244 -1.5036437 0.978
## 25:37  0.19050661 0.3980989 -1.5036437 0.978
## 25:52  0.42863988 0.8957225 -1.5036437 0.978
## 25:62  0.58739539 1.2274716 -1.5036437 0.978
## 32:37  0.07937776 0.1658745 -1.5036437 0.978
## 32:52  0.31751102 0.6634982 -1.5036437 0.978
## 32:62  0.47626653 0.9952472 -1.5036437 0.978
## 37:52  0.23813327 0.4976236 -1.5036437 0.978
## 37:62  0.39688878 0.8293727 -1.5036437 0.978
## 52:62  0.15875551 0.3317491 -1.5036437 0.978

```