Tomato nutrient study correlations

Rachel Schattman and Becky Maden

November 7, 2018

## Load library

library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(ggpubr)

## Loading required package: ggplot2

## Loading required package: magrittr

library(ggplot2)  
library(car)

## Loading required package: carData

##   
## Attaching package: 'car'

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

library(gvlma)

## Start script

script\_path <- "C:/Users/rschattman/Documents/Research/Greenhouse\_tomatoes\_2018"  
in\_dir <- "C:/Users/rschattman/Documents/Research/Greenhouse\_tomatoes\_2018"  
out\_dir <- "C:/Users/rschattman/Documents/Research/Greenhouse\_tomatoes\_2018/output"  
data <- data.frame(read.csv("C:/Users/rschattman/Documents/Research/Greenhouse\_tomatoes\_2018/Nutrient\_levels\_yield\_NH\_VT\_MA\_11\_14\_18.csv"))

# Nitrate

## Create dataframe w/variables of interest for Nitrate only, omit rows with NA

tomatofun\_nitrate <- subset(data, select = c("Nitrate.N..ppm..SME", "Nitrate.N..ppm..MM", "TN.....Leaf.Analysis"))  
tomatofun\_nitrateNONA <- na.omit(tomatofun\_nitrate)

## Check distribution - all are non normal

shapiro.test(tomatofun\_nitrateNONA$Nitrate.N..ppm..SME)

##   
## Shapiro-Wilk normality test  
##   
## data: tomatofun\_nitrateNONA$Nitrate.N..ppm..SME  
## W = 0.86211, p-value = 9.017e-05

shapiro.test(tomatofun\_nitrateNONA$Nitrate.N..ppm..MM)

##   
## Shapiro-Wilk normality test  
##   
## data: tomatofun\_nitrateNONA$Nitrate.N..ppm..MM  
## W = 0.88707, p-value = 0.0004461

shapiro.test(tomatofun\_nitrateNONA$TN.....Leaf.Analysis)

##   
## Shapiro-Wilk normality test  
##   
## data: tomatofun\_nitrateNONA$TN.....Leaf.Analysis  
## W = 0.96102, p-value = 0.1421

## Correlation tests

cor(tomatofun\_nitrateNONA, method = "spearman")

## Nitrate.N..ppm..SME Nitrate.N..ppm..MM  
## Nitrate.N..ppm..SME 1.000000 0.9770220  
## Nitrate.N..ppm..MM 0.977022 1.0000000  
## TN.....Leaf.Analysis 0.269594 0.2836652  
## TN.....Leaf.Analysis  
## Nitrate.N..ppm..SME 0.2695940  
## Nitrate.N..ppm..MM 0.2836652  
## TN.....Leaf.Analysis 1.0000000

cor.test(x = tomatofun\_nitrateNONA$Nitrate.N..ppm..SME, y = tomatofun\_nitrateNONA$Nitrate.N..ppm..MM, alternative = "two.sided", method = "spearman", exact = NULL)

## Warning in cor.test.default(x =  
## tomatofun\_nitrateNONA$Nitrate.N..ppm..SME, : Cannot compute exact p-value  
## with ties

##   
## Spearman's rank correlation rho  
##   
## data: tomatofun\_nitrateNONA$Nitrate.N..ppm..SME and tomatofun\_nitrateNONA$Nitrate.N..ppm..MM  
## S = 326.06, p-value < 2.2e-16  
## alternative hypothesis: true rho is not equal to 0  
## sample estimates:  
## rho   
## 0.977022

cor.test(x = tomatofun\_nitrateNONA$Nitrate.N..ppm..SME, y = tomatofun\_nitrateNONA$TN.....Leaf.Analysis, alternative = "two.sided", method = "spearman", exact = NULL)

## Warning in cor.test.default(x =  
## tomatofun\_nitrateNONA$Nitrate.N..ppm..SME, : Cannot compute exact p-value  
## with ties

##   
## Spearman's rank correlation rho  
##   
## data: tomatofun\_nitrateNONA$Nitrate.N..ppm..SME and tomatofun\_nitrateNONA$TN.....Leaf.Analysis  
## S = 10364, p-value = 0.07677  
## alternative hypothesis: true rho is not equal to 0  
## sample estimates:  
## rho   
## 0.269594

cor.test(x = tomatofun\_nitrateNONA$Nitrate.N..ppm..MM, y = tomatofun\_nitrateNONA$TN.....Leaf.Analysis, alternative = "two.sided", method = "spearman", exact = NULL)

## Warning in cor.test.default(x = tomatofun\_nitrateNONA$Nitrate.N..ppm..MM, :  
## Cannot compute exact p-value with ties

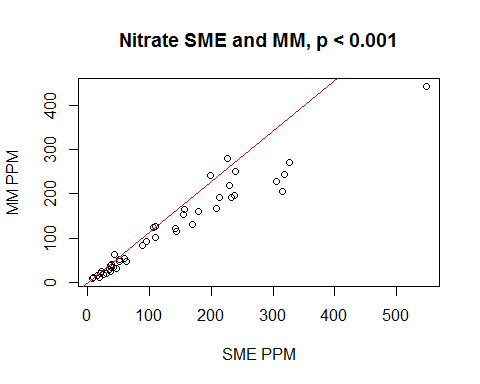
##   
## Spearman's rank correlation rho  
##   
## data: tomatofun\_nitrateNONA$Nitrate.N..ppm..MM and tomatofun\_nitrateNONA$TN.....Leaf.Analysis  
## S = 10165, p-value = 0.06204  
## alternative hypothesis: true rho is not equal to 0  
## sample estimates:  
## rho   
## 0.2836652

## plot SME and MM correlations

?plot

## starting httpd help server ... done

plot(tomatofun\_nitrateNONA$Nitrate.N..ppm..SME, tomatofun\_nitrateNONA$Nitrate.N..ppm..MM, type = "p", main = "Nitrate SME and MM, p < 0.001", xlab = "SME PPM", ylab = "MM PPM")  
abline(lm(tomatofun\_nitrateNONA$Nitrate.N..ppm..SME~tomatofun\_nitrateNONA$Nitrate.N..ppm..MM), col = "red") # regression line

 # Phosphorus ## Create dataframe w/variables of interest for P only, omit rows with NA

tomatofun\_P <- subset(data, select = c("Phosphorus..ppm..SME", "Phosphorus..lb.Ac..MM", "P.....Leaf.Analysis"))  
tomatofun\_P\_NONA <- na.omit(tomatofun\_P)

## Check distribution

shapiro.test(tomatofun\_P\_NONA$Phosphorus..ppm..SME)

##   
## Shapiro-Wilk normality test  
##   
## data: tomatofun\_P\_NONA$Phosphorus..ppm..SME  
## W = 0.68523, p-value = 5.792e-10

shapiro.test(tomatofun\_P\_NONA$Phosphorus..lb.Ac..MM)

##   
## Shapiro-Wilk normality test  
##   
## data: tomatofun\_P\_NONA$Phosphorus..lb.Ac..MM  
## W = 0.80092, p-value = 1.704e-07

shapiro.test(tomatofun\_P\_NONA$P.....Leaf.Analysis)

##   
## Shapiro-Wilk normality test  
##   
## data: tomatofun\_P\_NONA$P.....Leaf.Analysis  
## W = 0.95124, p-value = 0.01922

## Correlation tests

cor(tomatofun\_P\_NONA, method = "spearman")

## Phosphorus..ppm..SME Phosphorus..lb.Ac..MM  
## Phosphorus..ppm..SME 1.0000000 0.8404120  
## Phosphorus..lb.Ac..MM 0.8404120 1.0000000  
## P.....Leaf.Analysis 0.2051023 0.3496184  
## P.....Leaf.Analysis  
## Phosphorus..ppm..SME 0.2051023  
## Phosphorus..lb.Ac..MM 0.3496184  
## P.....Leaf.Analysis 1.0000000

cor.test(x = tomatofun\_P\_NONA$Phosphorus..ppm..SME, y = tomatofun\_P\_NONA$Phosphorus..lb.Ac..MM, alternative = "two.sided", method = "spearman", exact = NULL)

## Warning in cor.test.default(x = tomatofun\_P\_NONA$Phosphorus..ppm..SME, y =  
## tomatofun\_P\_NONA$Phosphorus..lb.Ac..MM, : Cannot compute exact p-value with  
## ties

##   
## Spearman's rank correlation rho  
##   
## data: tomatofun\_P\_NONA$Phosphorus..ppm..SME and tomatofun\_P\_NONA$Phosphorus..lb.Ac..MM  
## S = 5461.1, p-value < 2.2e-16  
## alternative hypothesis: true rho is not equal to 0  
## sample estimates:  
## rho   
## 0.840412

cor.test(x = tomatofun\_P\_NONA$Phosphorus..ppm..SME, y = tomatofun\_P\_NONA$P.....Leaf.Analysis, alternative = "two.sided", method = "spearman", exact = NULL)

## Warning in cor.test.default(x = tomatofun\_P\_NONA$Phosphorus..ppm..SME, y =  
## tomatofun\_P\_NONA$P.....Leaf.Analysis, : Cannot compute exact p-value with  
## ties

##   
## Spearman's rank correlation rho  
##   
## data: tomatofun\_P\_NONA$Phosphorus..ppm..SME and tomatofun\_P\_NONA$P.....Leaf.Analysis  
## S = 27201, p-value = 0.1192  
## alternative hypothesis: true rho is not equal to 0  
## sample estimates:  
## rho   
## 0.2051023

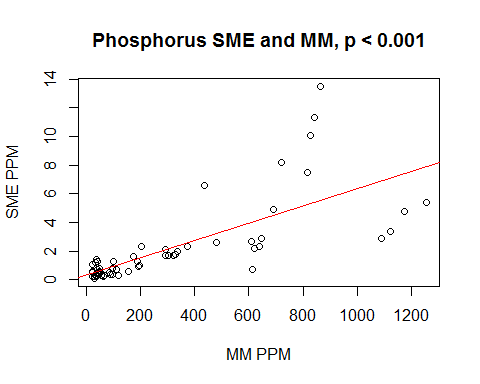
cor.test(x = tomatofun\_P\_NONA$Phosphorus..lb.Ac..MM, y = tomatofun\_P\_NONA$P.....Leaf.Analysis, alternative = "two.sided", method = "spearman", exact = NULL)

## Warning in cor.test.default(x = tomatofun\_P\_NONA$Phosphorus..lb.Ac..MM, :  
## Cannot compute exact p-value with ties

##   
## Spearman's rank correlation rho  
##   
## data: tomatofun\_P\_NONA$Phosphorus..lb.Ac..MM and tomatofun\_P\_NONA$P.....Leaf.Analysis  
## S = 22256, p-value = 0.006643  
## alternative hypothesis: true rho is not equal to 0  
## sample estimates:  
## rho   
## 0.3496184

## plot SME and MM correlations

plot(tomatofun\_P\_NONA$Phosphorus..lb.Ac..MM, tomatofun\_P\_NONA$Phosphorus..ppm..SME, type = "p", main = "Phosphorus SME and MM, p < 0.001", xlab = "MM PPM", ylab = "SME PPM")  
  
abline(lm(tomatofun\_P\_NONA$Phosphorus..ppm..SME~tomatofun\_P\_NONA$Phosphorus..lb.Ac..MM), col = "red") # regression line



# Potassium

## Create dataframe w/variables of interest for P only, omit rows with NA

head(data)

tomatofun\_K <- subset(data, select = c("Potassium..ppm..SME", "Potassium..lb.Ac..MM", "K.....Leaf.Analysis"))  
tomatofun\_K\_NONA <- na.omit(tomatofun\_K)

## Check distribution

shapiro.test(tomatofun\_K\_NONA$Potassium..ppm..SME)

##   
## Shapiro-Wilk normality test  
##   
## data: tomatofun\_K\_NONA$Potassium..ppm..SME  
## W = 0.83161, p-value = 1.075e-06

shapiro.test(tomatofun\_K\_NONA$Potassium..lb.Ac..MM)

##   
## Shapiro-Wilk normality test  
##   
## data: tomatofun\_K\_NONA$Potassium..lb.Ac..MM  
## W = 0.93469, p-value = 0.003474

shapiro.test(tomatofun\_K\_NONA$K.....Leaf.Analysis)

##   
## Shapiro-Wilk normality test  
##   
## data: tomatofun\_K\_NONA$K.....Leaf.Analysis  
## W = 0.96719, p-value = 0.1118

cor(tomatofun\_K\_NONA, method = "spearman")

## Potassium..ppm..SME Potassium..lb.Ac..MM  
## Potassium..ppm..SME 1.00000000 0.8702502  
## Potassium..lb.Ac..MM 0.87025023 1.0000000  
## K.....Leaf.Analysis -0.08586213 -0.1142807  
## K.....Leaf.Analysis  
## Potassium..ppm..SME -0.08586213  
## Potassium..lb.Ac..MM -0.11428071  
## K.....Leaf.Analysis 1.00000000

cor.test(x = tomatofun\_K\_NONA$Potassium..ppm..SME, y = tomatofun\_K\_NONA$Potassium..lb.Ac..MM, alternative = "two.sided", method = "spearman", exact = NULL)

## Warning in cor.test.default(x = tomatofun\_K\_NONA$Potassium..ppm..SME, y =  
## tomatofun\_K\_NONA$Potassium..lb.Ac..MM, : Cannot compute exact p-value with  
## ties

##   
## Spearman's rank correlation rho  
##   
## data: tomatofun\_K\_NONA$Potassium..ppm..SME and tomatofun\_K\_NONA$Potassium..lb.Ac..MM  
## S = 4440, p-value < 2.2e-16  
## alternative hypothesis: true rho is not equal to 0  
## sample estimates:  
## rho   
## 0.8702502

cor.test(x = tomatofun\_K\_NONA$Potassium..ppm..SME, y = tomatofun\_K\_NONA$K.....Leaf.Analysis, alternative = "two.sided", method = "spearman", exact = NULL)

## Warning in cor.test.default(x = tomatofun\_K\_NONA$Potassium..ppm..SME, y =  
## tomatofun\_K\_NONA$K.....Leaf.Analysis, : Cannot compute exact p-value with  
## ties

##   
## Spearman's rank correlation rho  
##   
## data: tomatofun\_K\_NONA$Potassium..ppm..SME and tomatofun\_K\_NONA$K.....Leaf.Analysis  
## S = 37158, p-value = 0.5179  
## alternative hypothesis: true rho is not equal to 0  
## sample estimates:  
## rho   
## -0.08586213

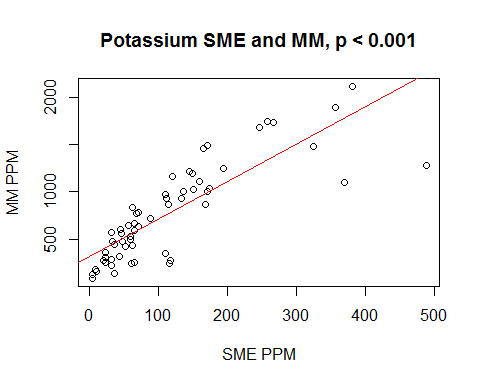
cor.test(x = tomatofun\_K\_NONA$Potassium..lb.Ac..MM, y = tomatofun\_K\_NONA$K.....Leaf.Analysis, alternative = "two.sided", method = "spearman", exact = NULL)

## Warning in cor.test.default(x = tomatofun\_K\_NONA$Potassium..lb.Ac..MM, y =  
## tomatofun\_K\_NONA$K.....Leaf.Analysis, : Cannot compute exact p-value with  
## ties

##   
## Spearman's rank correlation rho  
##   
## data: tomatofun\_K\_NONA$Potassium..lb.Ac..MM and tomatofun\_K\_NONA$K.....Leaf.Analysis  
## S = 38131, p-value = 0.3888  
## alternative hypothesis: true rho is not equal to 0  
## sample estimates:  
## rho   
## -0.1142807

## plot SME and MM correlations

plot(tomatofun\_K\_NONA$Potassium..ppm..SME, tomatofun\_K\_NONA$Potassium..lb.Ac..MM, type = "p", main = "Potassium SME and MM, p < 0.001", xlab = "SME PPM", ylab = "MM PPM")  
abline(lm(tomatofun\_K\_NONA$Potassium..lb.Ac..MM~tomatofun\_K\_NONA$Potassium..ppm..SME), col = "red") # regression line



# P SME and pH MM

## Create dataframe w/variables of interest, omit rows with NA

head(data)

tomatofun\_P\_pH <- subset(data, select = c("Phosphorus..ppm..SME", "Soil.pH"))  
tomatofun\_P\_pH\_NONA <- na.omit(tomatofun\_P\_pH)

cor(tomatofun\_P\_pH\_NONA, method = "spearman")

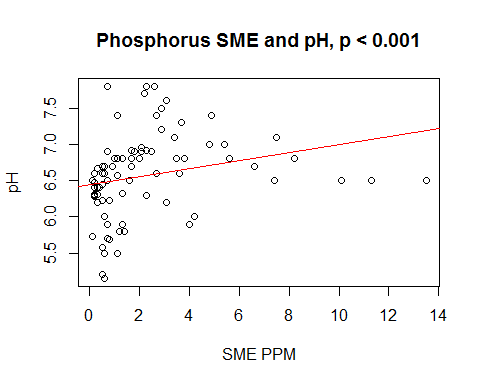
## Phosphorus..ppm..SME Soil.pH  
## Phosphorus..ppm..SME 1.0000000 0.4851736  
## Soil.pH 0.4851736 1.0000000

cor.test(x = tomatofun\_P\_pH\_NONA$Phosphorus..ppm..SME, y = tomatofun\_P\_pH\_NONA$Soil.pH, alternative = "two.sided", method = "spearman", exact = NULL)

## Warning in cor.test.default(x = tomatofun\_P\_pH\_NONA$Phosphorus..ppm..SME, :  
## Cannot compute exact p-value with ties

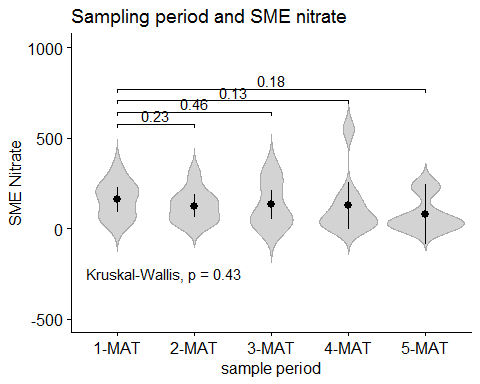
##   
## Spearman's rank correlation rho  
##   
## data: tomatofun\_P\_pH\_NONA$Phosphorus..ppm..SME and tomatofun\_P\_pH\_NONA$Soil.pH  
## S = 52687, p-value = 2.527e-06  
## alternative hypothesis: true rho is not equal to 0  
## sample estimates:  
## rho   
## 0.4851736

plot(tomatofun\_P\_pH\_NONA$Phosphorus..ppm..SME, tomatofun\_P\_pH\_NONA$Soil.pH, type = "p", main = "Phosphorus SME and pH, p < 0.001", xlab = "SME PPM", ylab = "pH")  
abline(lm(tomatofun\_P\_pH\_NONA$Soil.pH~tomatofun\_P\_pH\_NONA$Phosphorus..ppm..SME), col = "red") # regression line

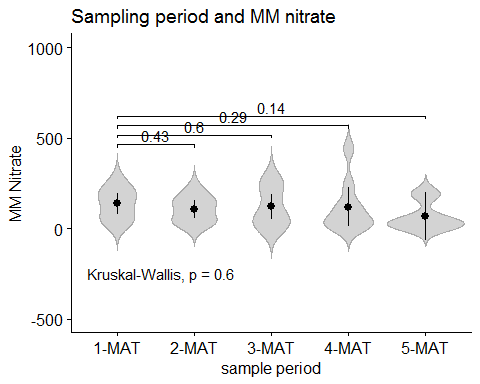


# Take home message: No significant difference between Nitrate levels (any test) by sample period.

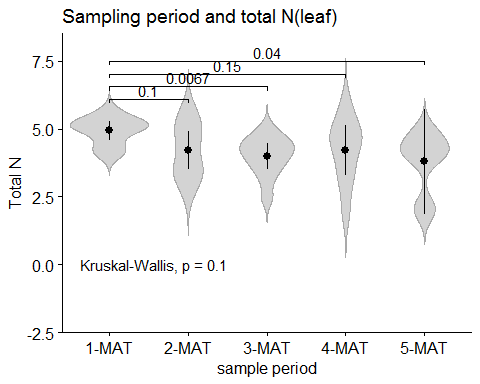
tomatofun\_yield\_N <- subset(data, select = c("Total.Yield.per.sq.ft", "Sample.Period", "Nitrate.N..ppm..SME", "Nitrate.N..ppm..MM", "TN.....Leaf.Analysis"))  
  
tomatofun\_yield\_N\_NONA <- na.omit(tomatofun\_yield\_N)  
  
#set comparisons for error plots  
my\_comparisons <- list(c("Pre-Plant", "1-MAT"), c("Pre-Plant", "2-MAT"), c("Pre-Plant", "3-MAT"), c("Pre-Plant", "4-MAT"), c("Pre-Plant", "5-MAT"))  
my\_comparisonsb <- list(c("1-MAT", "2-MAT"), c("1-MAT", "3-MAT"), c("1-MAT", "4-MAT"), c("1-MAT", "5-MAT"))  
  
#correct order of sample periods  
tomatofun\_yield\_N\_NONA$"Sample.Period" <- factor(tomatofun\_yield\_N\_NONA$"Sample.Period", levels = c('Pre-Plant',   
 '1-MAT',   
 '2-MAT',   
 '3-MAT',   
 '4-MAT',   
 '5-MAT'))  
  
ggerrorplot(tomatofun\_yield\_N\_NONA, x = "Sample.Period",   
 y = "Nitrate.N..ppm..SME",  
 combine = FALSE, merge = TRUE,  
 desc\_stat = "mean\_ci",   
 color = "black",  
 palette = "npg",  
 title = "Sampling period and SME nitrate",  
 add = "violin", add.params = list(color = "darkgray", fill="lightgray"),  
 ylim = c(-500, 1000),  
 legend = "bottom",  
 xlab = "sample period",  
 ylab = "SME Nitrate",  
 orientation = "vertical") +   
 stat\_compare\_means(comparisons = my\_comparisonsb) +  
 stat\_compare\_means(label.y = -250, label.x = )



ggerrorplot(tomatofun\_yield\_N\_NONA, x = "Sample.Period",   
 y = "Nitrate.N..ppm..MM",  
 combine = FALSE, merge = TRUE,  
 desc\_stat = "mean\_ci",   
 color = "black",  
 palette = "npg",  
 title = "Sampling period and MM nitrate",  
 add = "violin", add.params = list(color = "darkgray", fill="lightgray"),  
 ylim = c(-500, 1000),  
 legend = "bottom",  
 xlab = "sample period",  
 ylab = "MM Nitrate",  
 orientation = "vertical") +   
 stat\_compare\_means(comparisons = my\_comparisonsb) +  
 stat\_compare\_means(label.y = -250, label.x = )



ggerrorplot(tomatofun\_yield\_N\_NONA, x = "Sample.Period",   
 y = "TN.....Leaf.Analysis",  
 combine = FALSE, merge = TRUE,  
 desc\_stat = "mean\_ci",   
 color = "black",  
 palette = "npg",  
 title = "Sampling period and total N(leaf)",  
 add = "violin", add.params = list(color = "darkgray", fill="lightgray"),  
 ylim = c(-2, 8),  
 legend = "bottom",  
 xlab = "sample period",  
 ylab = "Total N",  
 orientation = "vertical") +   
 stat\_compare\_means(comparisons = my\_comparisonsb) +  
 stat\_compare\_means(label.y = -0, label.x = )



## Linear multiple regression models for all measures of N

# take away message: no models are significant. In otherwords, there is no evidence that testing N at a

# particular MAT has any relationship to yeild.

as.factor(tomatofun\_yield\_N\_NONA$"Total.Yield.per.sq.ft")

## [1] 2.79 2.79 2.79 2.79 3.52 3.52 3.52 3.52 3.52 2.85 2.85 2.85 2.85 2.85  
## [15] 1.88 1.88 1.88 1.92 1.92 1.92 5.11 5.11 5.11 5.11 2.94 2.94 2.94 2.94  
## [29] 2.94 5.11 5.11 5.11 5.11 2.7 2.7 2.7 2.7 2.59 2.59 2.59 1.43 1.43  
## [43] 1.43 1.43  
## Levels: 1.43 1.88 1.92 2.59 2.7 2.79 2.85 2.94 3.52 5.11

as.factor(tomatofun\_yield\_N\_NONA$"Sample.Period")

## [1] 1-MAT 2-MAT 3-MAT 4-MAT 1-MAT 2-MAT 3-MAT 4-MAT 5-MAT 1-MAT 2-MAT  
## [12] 3-MAT 4-MAT 5-MAT 2-MAT 3-MAT 4-MAT 1-MAT 2-MAT 3-MAT 1-MAT 3-MAT  
## [23] 4-MAT 5-MAT 1-MAT 2-MAT 3-MAT 4-MAT 5-MAT 1-MAT 2-MAT 3-MAT 4-MAT  
## [34] 1-MAT 2-MAT 3-MAT 4-MAT 1-MAT 2-MAT 3-MAT 1-MAT 2-MAT 3-MAT 4-MAT  
## Levels: Pre-Plant 1-MAT 2-MAT 3-MAT 4-MAT 5-MAT

as.factor(tomatofun\_yield\_N\_NONA$"Nitrate.N..ppm..SME")

## [1] 88.4 110 94.7 110 237 305 318 548 232 170 144 315 142 59.4  
## [15] 37 34.9 42.6 327 209 106 62.1 39.1 23 20.4 229 51.6 15.9 7.1   
## [29] 8.9 51.8 46.7 38.5 41.6 180 156 226 199 214 155 240 37 31   
## [43] 19.1 26.2  
## 42 Levels: 7.1 8.9 15.9 19.1 20.4 23 26.2 31 34.9 37 38.5 39.1 ... 548

as.factor(tomatofun\_yield\_N\_NONA$"Nitrate.N..ppm..MM")

## [1] 83 102 93 127 197 228 243 442 191 131 114 206 122 53 37 28 62   
## [18] 271 167 124 47 33 24 21 218 52 16 9 10 47 32 41 34 159  
## [35] 165 279 240 192 153 251 24 21 12 17   
## 41 Levels: 9 10 12 16 17 21 24 28 32 33 34 37 41 47 52 53 62 83 93 ... 442

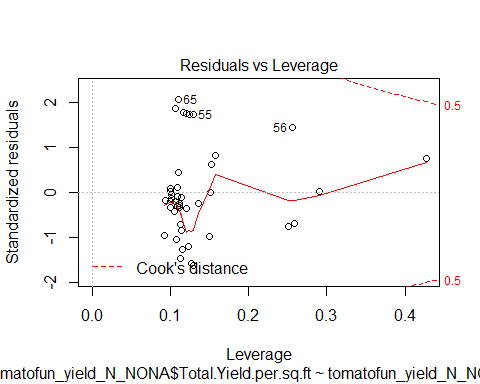
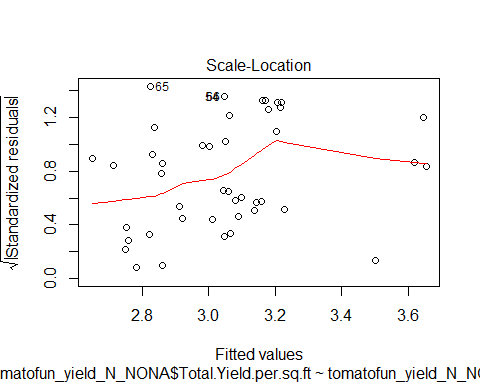
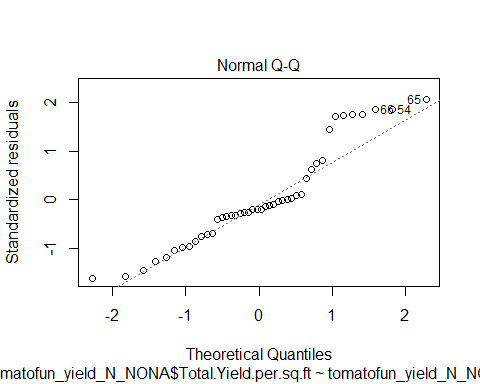
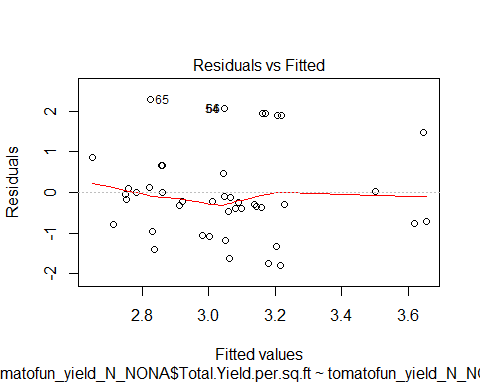
as.factor(tomatofun\_yield\_N\_NONA$"TN.....Leaf.Analysis")

## [1] 5.29 5.51 4.98 5.24 5.06 4.73 3.89 4.17 4.04 5.65 4.65 4.56 4.78 4.25  
## [15] 4.96 4.22 3.46 4.21 5.53 4.32 4.06 3.56 4.66 4.83 5.3 3.46 3.83 1.99  
## [29] 2.05 4.95 3.33 4.66 4.63 5.07 3.45 4.29 5.89 4.54 2.75 2.48 5.16 3.72  
## [43] 3.19 3.13  
## 42 Levels: 1.99 2.05 2.48 2.75 3.13 3.19 3.33 3.45 3.46 3.56 3.72 ... 5.89

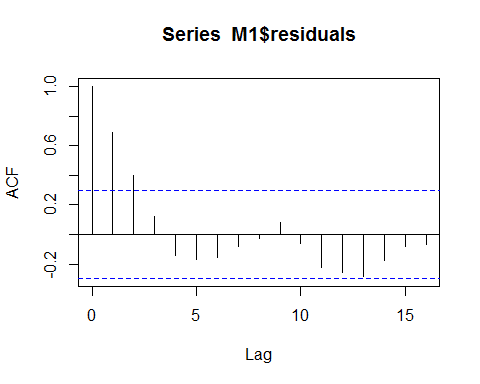
#run model  
M1 <- lm(tomatofun\_yield\_N\_NONA$"Total.Yield.per.sq.ft" ~ tomatofun\_yield\_N\_NONA$"Nitrate.N..ppm..SME" + tomatofun\_yield\_N\_NONA$"Sample.Period", data = tomatofun\_yield\_N\_NONA, method = "qr")  
  
#check assumptions  
mean(M1$residuals) #Mean is <1, which is good.

## [1] 5.830051e-18

plot(M1) #Residuals vs fitted indicate homoscedasticity. Will need to rerun model. Points are normally distributed.



acf(M1$residuals) #No autocorrelation of residuals.



lmtest::dwtest(M1) #No autocorrelation of residuals.

##   
## Durbin-Watson test  
##   
## data: M1  
## DW = 0.56018, p-value = 1.915e-08  
## alternative hypothesis: true autocorrelation is greater than 0

cor.test(tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..SME, M1$residuals) #X variables and residuals are not correlated.

##   
## Pearson's product-moment correlation  
##   
## data: tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..SME and M1$residuals  
## t = -4.4726e-17, df = 42, p-value = 1  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.2968804 0.2968804  
## sample estimates:  
## cor   
## -6.901411e-18

var(tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..SME) #Variability is positive.

## [1] 13515.8

vif(M1) #Low GIF (<4) means low chance of multi-collinearity.

## GVIF Df GVIF^(1/(2\*Df))  
## tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..SME 1.034017 1 1.016866  
## tomatofun\_yield\_N\_NONA$Sample.Period 1.034017 4 1.004190

gvlma::gvlma(M1) #Shortcut for checking assumptions.

##   
## Call:  
## lm(formula = tomatofun\_yield\_N\_NONA$Total.Yield.per.sq.ft ~ tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..SME +   
## tomatofun\_yield\_N\_NONA$Sample.Period, data = tomatofun\_yield\_N\_NONA,   
## method = "qr")  
##   
## Coefficients:  
## (Intercept)   
## 3.2046641   
## tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..SME   
## -0.0006807   
## tomatofun\_yield\_N\_NONA$Sample.Period2-MAT   
## -0.3468934   
## tomatofun\_yield\_N\_NONA$Sample.Period3-MAT   
## -0.1296509   
## tomatofun\_yield\_N\_NONA$Sample.Period4-MAT   
## 0.0293010   
## tomatofun\_yield\_N\_NONA$Sample.Period5-MAT   
## 0.4549130   
##   
##   
## ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS  
## USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:  
## Level of Significance = 0.05   
##   
## Call:  
## gvlma::gvlma(x = M1)   
##   
## Value p-value Decision  
## Global Stat 7.8813 0.09602 Assumptions acceptable.  
## Skewness 2.7081 0.09984 Assumptions acceptable.  
## Kurtosis 0.3095 0.57801 Assumptions acceptable.  
## Link Function 0.3612 0.54782 Assumptions acceptable.  
## Heteroscedasticity 4.5025 0.03384 Assumptions NOT satisfied!

summary(M1)

##   
## Call:  
## lm(formula = tomatofun\_yield\_N\_NONA$Total.Yield.per.sq.ft ~ tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..SME +   
## tomatofun\_yield\_N\_NONA$Sample.Period, data = tomatofun\_yield\_N\_NONA,   
## method = "qr")  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.7861 -0.7274 -0.2209 0.5223 2.2840   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) 3.2046641 0.4490854 7.136  
## tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..SME -0.0006807 0.0015713 -0.433  
## tomatofun\_yield\_N\_NONA$Sample.Period2-MAT -0.3468934 0.5296923 -0.655  
## tomatofun\_yield\_N\_NONA$Sample.Period3-MAT -0.1296509 0.5165868 -0.251  
## tomatofun\_yield\_N\_NONA$Sample.Period4-MAT 0.0293010 0.5437301 0.054  
## tomatofun\_yield\_N\_NONA$Sample.Period5-MAT 0.4549130 0.7080027 0.643  
## Pr(>|t|)   
## (Intercept) 1.61e-08 \*\*\*  
## tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..SME 0.667   
## tomatofun\_yield\_N\_NONA$Sample.Period2-MAT 0.516   
## tomatofun\_yield\_N\_NONA$Sample.Period3-MAT 0.803   
## tomatofun\_yield\_N\_NONA$Sample.Period4-MAT 0.957   
## tomatofun\_yield\_N\_NONA$Sample.Period5-MAT 0.524   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.178 on 38 degrees of freedom  
## Multiple R-squared: 0.04375, Adjusted R-squared: -0.08207   
## F-statistic: 0.3477 on 5 and 38 DF, p-value: 0.8805

# if the model had been successful, I would take the time to remove the outliers and re-run, but since the model is not   
# significant, I think it would be a waste of time.  
  
M2 <- lm(tomatofun\_yield\_N\_NONA$"Total.Yield.per.sq.ft" ~ tomatofun\_yield\_N\_NONA$"Nitrate.N..ppm..MM" + tomatofun\_yield\_N\_NONA$"Sample.Period", data = tomatofun\_yield\_N\_NONA, method = "qr")  
gvlma::gvlma(M2)

##   
## Call:  
## lm(formula = tomatofun\_yield\_N\_NONA$Total.Yield.per.sq.ft ~ tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..MM +   
## tomatofun\_yield\_N\_NONA$Sample.Period, data = tomatofun\_yield\_N\_NONA,   
## method = "qr")  
##   
## Coefficients:  
## (Intercept)   
## 3.263140   
## tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..MM   
## -0.001221   
## tomatofun\_yield\_N\_NONA$Sample.Period2-MAT   
## -0.359382   
## tomatofun\_yield\_N\_NONA$Sample.Period3-MAT   
## -0.130513   
## tomatofun\_yield\_N\_NONA$Sample.Period4-MAT   
## 0.030738   
## tomatofun\_yield\_N\_NONA$Sample.Period5-MAT   
## 0.425796   
##   
##   
## ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS  
## USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:  
## Level of Significance = 0.05   
##   
## Call:  
## gvlma::gvlma(x = M2)   
##   
## Value p-value Decision  
## Global Stat 7.3133 0.12023 Assumptions acceptable.  
## Skewness 2.0121 0.15605 Assumptions acceptable.  
## Kurtosis 0.3797 0.53777 Assumptions acceptable.  
## Link Function 0.3236 0.56944 Assumptions acceptable.  
## Heteroscedasticity 4.5978 0.03201 Assumptions NOT satisfied!

summary(M2)

##   
## Call:  
## lm(formula = tomatofun\_yield\_N\_NONA$Total.Yield.per.sq.ft ~ tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..MM +   
## tomatofun\_yield\_N\_NONA$Sample.Period, data = tomatofun\_yield\_N\_NONA,   
## method = "qr")  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.8431 -0.7461 -0.2011 0.5440 2.2453   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) 3.263140 0.449851 7.254  
## tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..MM -0.001221 0.001855 -0.658  
## tomatofun\_yield\_N\_NONA$Sample.Period2-MAT -0.359382 0.528029 -0.681  
## tomatofun\_yield\_N\_NONA$Sample.Period3-MAT -0.130513 0.513951 -0.254  
## tomatofun\_yield\_N\_NONA$Sample.Period4-MAT 0.030738 0.540464 0.057  
## tomatofun\_yield\_N\_NONA$Sample.Period5-MAT 0.425796 0.706087 0.603  
## Pr(>|t|)   
## (Intercept) 1.12e-08 \*\*\*  
## tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..MM 0.514   
## tomatofun\_yield\_N\_NONA$Sample.Period2-MAT 0.500   
## tomatofun\_yield\_N\_NONA$Sample.Period3-MAT 0.801   
## tomatofun\_yield\_N\_NONA$Sample.Period4-MAT 0.955   
## tomatofun\_yield\_N\_NONA$Sample.Period5-MAT 0.550   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.174 on 38 degrees of freedom  
## Multiple R-squared: 0.04986, Adjusted R-squared: -0.07516   
## F-statistic: 0.3988 on 5 and 38 DF, p-value: 0.8465

M3 <- lm(tomatofun\_yield\_N\_NONA$"Total.Yield.per.sq.ft" ~ tomatofun\_yield\_N\_NONA$TN.....Leaf.Analysis + tomatofun\_yield\_N\_NONA$"Sample.Period", data = tomatofun\_yield\_N\_NONA, method = "qr")  
gvlma::gvlma(M3)

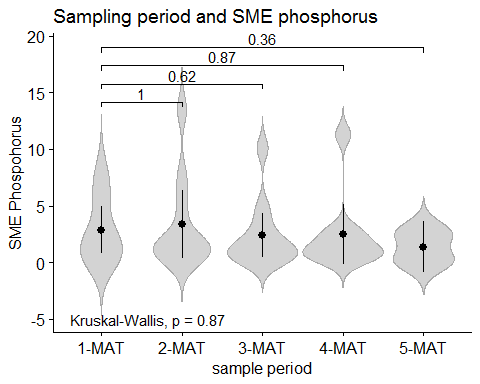
##   
## Call:  
## lm(formula = tomatofun\_yield\_N\_NONA$Total.Yield.per.sq.ft ~ tomatofun\_yield\_N\_NONA$TN.....Leaf.Analysis +   
## tomatofun\_yield\_N\_NONA$Sample.Period, data = tomatofun\_yield\_N\_NONA,   
## method = "qr")  
##   
## Coefficients:  
## (Intercept)   
## 2.56671   
## tomatofun\_yield\_N\_NONA$TN.....Leaf.Analysis   
## 0.10738   
## tomatofun\_yield\_N\_NONA$Sample.Period2-MAT   
## -0.24568   
## tomatofun\_yield\_N\_NONA$Sample.Period3-MAT   
## -0.01059   
## tomatofun\_yield\_N\_NONA$Sample.Period4-MAT   
## 0.12827   
## tomatofun\_yield\_N\_NONA$Sample.Period5-MAT   
## 0.63104   
##   
##   
## ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS  
## USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:  
## Level of Significance = 0.05   
##   
## Call:  
## gvlma::gvlma(x = M3)   
##   
## Value p-value Decision  
## Global Stat 10.0546 0.03952 Assumptions NOT satisfied!  
## Skewness 4.4622 0.03465 Assumptions NOT satisfied!  
## Kurtosis 0.1077 0.74283 Assumptions acceptable.  
## Link Function 1.4842 0.22312 Assumptions acceptable.  
## Heteroscedasticity 4.0006 0.04549 Assumptions NOT satisfied!

summary(M3)

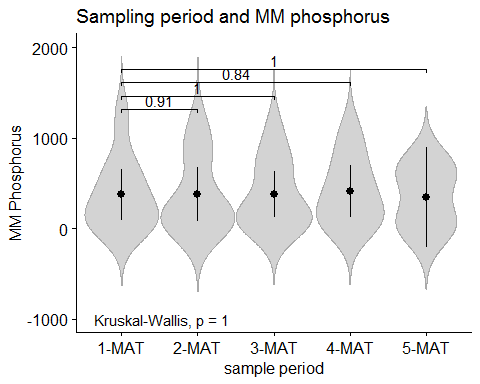
##   
## Call:  
## lm(formula = tomatofun\_yield\_N\_NONA$Total.Yield.per.sq.ft ~ tomatofun\_yield\_N\_NONA$TN.....Leaf.Analysis +   
## tomatofun\_yield\_N\_NONA$Sample.Period, data = tomatofun\_yield\_N\_NONA,   
## method = "qr")  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.6908 -0.6716 -0.2141 0.3854 2.4314   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) 2.56671 1.09352 2.347  
## tomatofun\_yield\_N\_NONA$TN.....Leaf.Analysis 0.10738 0.20861 0.515  
## tomatofun\_yield\_N\_NONA$Sample.Period2-MAT -0.24568 0.54730 -0.449  
## tomatofun\_yield\_N\_NONA$Sample.Period3-MAT -0.01059 0.54963 -0.019  
## tomatofun\_yield\_N\_NONA$Sample.Period4-MAT 0.12827 0.56075 0.229  
## tomatofun\_yield\_N\_NONA$Sample.Period5-MAT 0.63104 0.73547 0.858  
## Pr(>|t|)   
## (Intercept) 0.0242 \*  
## tomatofun\_yield\_N\_NONA$TN.....Leaf.Analysis 0.6097   
## tomatofun\_yield\_N\_NONA$Sample.Period2-MAT 0.6560   
## tomatofun\_yield\_N\_NONA$Sample.Period3-MAT 0.9847   
## tomatofun\_yield\_N\_NONA$Sample.Period4-MAT 0.8203   
## tomatofun\_yield\_N\_NONA$Sample.Period5-MAT 0.3963   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.177 on 38 degrees of freedom  
## Multiple R-squared: 0.04568, Adjusted R-squared: -0.07988   
## F-statistic: 0.3638 on 5 and 38 DF, p-value: 0.8701

head(data) # P

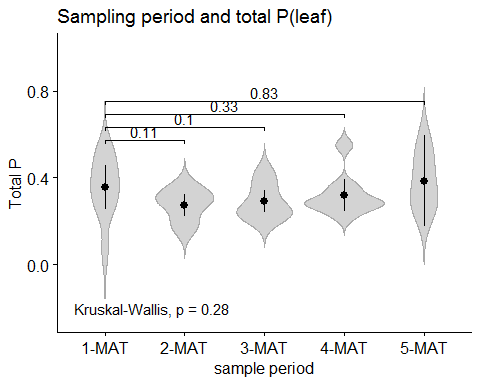
tomatofun\_yield\_P <- subset(data, select = c("Total.Yield.per.sq.ft", "Sample.Period", "Phosphorus..ppm..SME", "Phosphorus..lb.Ac..MM", "P.....Leaf.Analysis"))  
  
tomatofun\_yield\_P\_NONA <- na.omit(tomatofun\_yield\_P)  
  
#set comparisons for error plots  
my\_comparisons <- list(c("Pre-Plant", "1-MAT"), c("Pre-Plant", "2-MAT"), c("Pre-Plant", "3-MAT"), c("Pre-Plant", "4-MAT"), c("Pre-Plant", "5-MAT"))  
my\_comparisonsb <- list(c("1-MAT", "2-MAT"), c("1-MAT", "3-MAT"), c("1-MAT", "4-MAT"), c("1-MAT", "5-MAT"))  
  
#correct order of sample periods  
tomatofun\_yield\_P\_NONA$"Sample.Period" <- factor(tomatofun\_yield\_P\_NONA$"Sample.Period", levels = c('Pre-Plant',   
 '1-MAT',   
 '2-MAT',   
 '3-MAT',   
 '4-MAT',   
 '5-MAT'))  
  
ggerrorplot(tomatofun\_yield\_P\_NONA, x = "Sample.Period",   
 y = "Phosphorus..ppm..SME",  
 combine = FALSE, merge = TRUE,  
 desc\_stat = "mean\_ci",   
 color = "black",  
 palette = "npg",  
 title = "Sampling period and SME phosphorus",  
 add = "violin", add.params = list(color = "darkgray", fill="lightgray"),  
 ylim = c(-5, 19),  
 legend = "bottom",  
 xlab = "sample period",  
 ylab = "SME Phospohorus",  
 orientation = "vertical") +   
 stat\_compare\_means(comparisons = my\_comparisonsb) +  
 stat\_compare\_means(label.y = -5, label.x = )



ggerrorplot(tomatofun\_yield\_P\_NONA, x = "Sample.Period",   
 y = "Phosphorus..lb.Ac..MM",  
 combine = FALSE, merge = TRUE,  
 desc\_stat = "mean\_ci",   
 color = "black",  
 palette = "npg",  
 title = "Sampling period and MM phosphorus",  
 add = "violin", add.params = list(color = "darkgray", fill="lightgray"),  
 ylim = c(-1000, 2000),  
 legend = "bottom",  
 xlab = "sample period",  
 ylab = "MM Phosphorus",  
 orientation = "vertical") +   
 stat\_compare\_means(comparisons = my\_comparisonsb) +  
 stat\_compare\_means(label.y = -1000, label.x = )



ggerrorplot(tomatofun\_yield\_P\_NONA, x = "Sample.Period",   
 y = "P.....Leaf.Analysis",  
 combine = FALSE, merge = TRUE,  
 desc\_stat = "mean\_ci",   
 color = "black",  
 palette = "npg",  
 title = "Sampling period and total P(leaf)",  
 add = "violin", add.params = list(color = "darkgray", fill="lightgray"),  
 ylim = c(-.25, 1),  
 legend = "bottom",  
 xlab = "sample period",  
 ylab = "Total P",  
 orientation = "vertical") +   
 stat\_compare\_means(comparisons = my\_comparisonsb) +  
 stat\_compare\_means(label.y = -.2, label.x = )

 ## Linear multiple regression models for all measures of P # take away message: no models are significant. In otherwords, there is no evidence that testing P at a # particular MAT has any relationship to yeild.

M4 <- lm(tomatofun\_yield\_P\_NONA$"Total.Yield.per.sq.ft" ~ tomatofun\_yield\_P\_NONA$"Phosphorus..ppm..SME" + tomatofun\_yield\_P\_NONA$"Sample.Period", data = tomatofun\_yield\_P\_NONA, method = "qr")  
gvlma::gvlma(M4)

##   
## Call:  
## lm(formula = tomatofun\_yield\_P\_NONA$Total.Yield.per.sq.ft ~ tomatofun\_yield\_P\_NONA$Phosphorus..ppm..SME +   
## tomatofun\_yield\_P\_NONA$Sample.Period, data = tomatofun\_yield\_P\_NONA,   
## method = "qr")  
##   
## Coefficients:  
## (Intercept)   
## 2.96971   
## tomatofun\_yield\_P\_NONA$Phosphorus..ppm..SME   
## 0.04385   
## tomatofun\_yield\_P\_NONA$Sample.Period2-MAT   
## -0.34493   
## tomatofun\_yield\_P\_NONA$Sample.Period3-MAT   
## -0.08910   
## tomatofun\_yield\_P\_NONA$Sample.Period4-MAT   
## 0.06990   
## tomatofun\_yield\_P\_NONA$Sample.Period5-MAT   
## 0.57500   
##   
##   
## ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS  
## USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:  
## Level of Significance = 0.05   
##   
## Call:  
## gvlma::gvlma(x = M4)   
##   
## Value p-value Decision  
## Global Stat 9.7969 0.04399 Assumptions NOT satisfied!  
## Skewness 4.2923 0.03828 Assumptions NOT satisfied!  
## Kurtosis 0.2538 0.61442 Assumptions acceptable.  
## Link Function 0.8879 0.34605 Assumptions acceptable.  
## Heteroscedasticity 4.3628 0.03673 Assumptions NOT satisfied!

summary(M2)

##   
## Call:  
## lm(formula = tomatofun\_yield\_N\_NONA$Total.Yield.per.sq.ft ~ tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..MM +   
## tomatofun\_yield\_N\_NONA$Sample.Period, data = tomatofun\_yield\_N\_NONA,   
## method = "qr")  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.8431 -0.7461 -0.2011 0.5440 2.2453   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) 3.263140 0.449851 7.254  
## tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..MM -0.001221 0.001855 -0.658  
## tomatofun\_yield\_N\_NONA$Sample.Period2-MAT -0.359382 0.528029 -0.681  
## tomatofun\_yield\_N\_NONA$Sample.Period3-MAT -0.130513 0.513951 -0.254  
## tomatofun\_yield\_N\_NONA$Sample.Period4-MAT 0.030738 0.540464 0.057  
## tomatofun\_yield\_N\_NONA$Sample.Period5-MAT 0.425796 0.706087 0.603  
## Pr(>|t|)   
## (Intercept) 1.12e-08 \*\*\*  
## tomatofun\_yield\_N\_NONA$Nitrate.N..ppm..MM 0.514   
## tomatofun\_yield\_N\_NONA$Sample.Period2-MAT 0.500   
## tomatofun\_yield\_N\_NONA$Sample.Period3-MAT 0.801   
## tomatofun\_yield\_N\_NONA$Sample.Period4-MAT 0.955   
## tomatofun\_yield\_N\_NONA$Sample.Period5-MAT 0.550   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.174 on 38 degrees of freedom  
## Multiple R-squared: 0.04986, Adjusted R-squared: -0.07516   
## F-statistic: 0.3988 on 5 and 38 DF, p-value: 0.8465

M5 <- lm(tomatofun\_yield\_P\_NONA$"Total.Yield.per.sq.ft" ~ tomatofun\_yield\_P\_NONA$"Phosphorus..lb.Ac..MM" + tomatofun\_yield\_P\_NONA$"Sample.Period", data = tomatofun\_yield\_P\_NONA, method = "qr")  
gvlma::gvlma(M5)

##   
## Call:  
## lm(formula = tomatofun\_yield\_P\_NONA$Total.Yield.per.sq.ft ~ tomatofun\_yield\_P\_NONA$Phosphorus..lb.Ac..MM +   
## tomatofun\_yield\_P\_NONA$Sample.Period, data = tomatofun\_yield\_P\_NONA,   
## method = "qr")  
##   
## Coefficients:  
## (Intercept)   
## 2.379658   
## tomatofun\_yield\_P\_NONA$Phosphorus..lb.Ac..MM   
## 0.001919   
## tomatofun\_yield\_P\_NONA$Sample.Period2-MAT   
## -0.325629   
## tomatofun\_yield\_P\_NONA$Sample.Period3-MAT   
## -0.113530   
## tomatofun\_yield\_P\_NONA$Sample.Period4-MAT   
## -0.022123   
## tomatofun\_yield\_P\_NONA$Sample.Period5-MAT   
## 0.564152   
##   
##   
## ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS  
## USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:  
## Level of Significance = 0.05   
##   
## Call:  
## gvlma::gvlma(x = M5)   
##   
## Value p-value Decision  
## Global Stat 4.8305 0.3051 Assumptions acceptable.  
## Skewness 0.7755 0.3785 Assumptions acceptable.  
## Kurtosis 1.9220 0.1656 Assumptions acceptable.  
## Link Function 1.9180 0.1661 Assumptions acceptable.  
## Heteroscedasticity 0.2149 0.6429 Assumptions acceptable.

summary(M5) #model is significant, however the only coefficient that is sig is the Ph.MM measure (not any sampling period)

##   
## Call:  
## lm(formula = tomatofun\_yield\_P\_NONA$Total.Yield.per.sq.ft ~ tomatofun\_yield\_P\_NONA$Phosphorus..lb.Ac..MM +   
## tomatofun\_yield\_P\_NONA$Sample.Period, data = tomatofun\_yield\_P\_NONA,   
## method = "qr")  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.2733 -0.7170 -0.1273 0.6248 1.6503   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) 2.3796578 0.3183046 7.476  
## tomatofun\_yield\_P\_NONA$Phosphorus..lb.Ac..MM 0.0019190 0.0003721 5.158  
## tomatofun\_yield\_P\_NONA$Sample.Period2-MAT -0.3256290 0.4050385 -0.804  
## tomatofun\_yield\_P\_NONA$Sample.Period3-MAT -0.1135304 0.3957261 -0.287  
## tomatofun\_yield\_P\_NONA$Sample.Period4-MAT -0.0221228 0.4163838 -0.053  
## tomatofun\_yield\_P\_NONA$Sample.Period5-MAT 0.5641520 0.5359219 1.053  
## Pr(>|t|)   
## (Intercept) 5.63e-09 \*\*\*  
## tomatofun\_yield\_P\_NONA$Phosphorus..lb.Ac..MM 8.10e-06 \*\*\*  
## tomatofun\_yield\_P\_NONA$Sample.Period2-MAT 0.426   
## tomatofun\_yield\_P\_NONA$Sample.Period3-MAT 0.776   
## tomatofun\_yield\_P\_NONA$Sample.Period4-MAT 0.958   
## tomatofun\_yield\_P\_NONA$Sample.Period5-MAT 0.299   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.9057 on 38 degrees of freedom  
## Multiple R-squared: 0.4347, Adjusted R-squared: 0.3604   
## F-statistic: 5.845 on 5 and 38 DF, p-value: 0.0004242