Part I:

1) i) mossart = read.csv("Moss\_ArthropodSR.csv", skip = 4)

ii) mossart = mossart %>%

rename("rep"= Replicates, "tr" = Treatment, "A" = Area..cm2., "n" = Species.Richness)

iii) table(mossart$rep, mossart$tr, mossart$A)

The design is almost perfectly sampled with the exception of the very first block in the area 20 site.

iv) mossart$tr = as.factor(mossart$tr)

mossart$A = as.factor(mossart$A)

mossart$rep = as.factor(mossart$rep)

2) Fixed effect models:

mod1 = lm(n~tr\*A\*rep, data = mossart)

summary(mod1)

anova(mod1)

A fully interactive model is not possible, as R gives an error for perfect fit and values of 0 for the residuals. Model 1 is overfit.

mod2 = lm(n~tr\*A+rep, data = mossart)

summary(mod2)

anova(mod2)

Model 2 explains ~83% of the variation, 73% adjusted

mod3 = lm(n~tr+A+rep, data = mossart)

summary(mod3)

anova(mod3)

Model 3 also explains 83% variance, 75% adjusted. Slightly higher df though, so I’m going to opt for model 2.

3) Mixed-effect models:

mod4 = lme(n~tr\*A,random=~1|rep,data=mossart,control=list(opt="optim"))

summary(mod4) #Tau = 0.04608588, sigma = 3.007672

anova(mod4)

mod5 = lme(n~tr+A,random=~1|rep,data=mossart,control=list(opt="optim"))

summary(mod5) #Tau = 0.04766504, sigma = 2.942209

anova(mod5)

Tau values are larger than sigmas, good sign models aren’t completely inappropriate.

VarCorr(mod4)

vars = as.numeric(VarCorr(mod4)[,1])

corr = vars[1]/sum(vars)

Tau = vars[1]

Sig = vars[2]

Comparing variation among and within blocks.

mean.dat <- data.frame(fix.ests=predict(mod2), mix.ests=predict(mod4))

moss2 <- cbind(mossart, mean.dat)

moss2$mix.ests = as.numeric(moss2$mix.ests)

moss3 = moss2 %>%

group\_by(tr, A) %>%

mutate(pop\_mean = mean(n)) %>%

select(everything())

4) Rep, when controlled for, explains a substantial degree of the variation among the data left unexplained by the relationship between area and treatment type.

5) Mixed model compacts a substantial degree of the variation compared to both the raw data and the data from the fixed effects model. Note the relative spread of values in the graph below. Mixed model holds much tighter to the actual observed population means in the data.



Part II:

#Molly Jenkins

#Assignment 3

#ENEC563

#02/14/2017

# The data are contained in a comma separated file. If you open the file in text reader,

#you will notice that the first four rows contain descriptive information.

# 1\_1) Read the data into R with one line of code.

#To do this you will need to learn how to skip rows at the beginning of a file by examing the help for the read.csv file.

# 1\_2) Rename the columns something more easy to type.

# 1\_3) Generate a table that shows how many replicates there are for each treatment combination for each block

#and indicate if the experiment is balanced.

# 1\_4) You will be analyzing all predictor variables (Treatment, Size and Replicate) as categorical variables.

#Make sure you convert them all to categorical ones.

# 2) Analyze this experiment as a fixed effect model

# 2\_1) Find the best model which uses the two experimental treatments and the blocking variable as fixed effects.

#You will be analyzing all predictor variables (Treatment, Size and Replicate) as categorical variables.

# 2\_2) Explain why you cannot examine the three way interaction among all three variable.

# 3) Find the best mixed model which uses the two experimental treatments as fixed effects and the blocking variable as a random effect.

# Compare the amount of variation of the observations that is due to among block variation to the within block variation.

# 4) Give a one or two sentence description of what your final model from question 3 says.

# 5) Create a single graphic (the graphic may contain more than one panel) that summarizes the results of your analysis in question 3.

# It should show the average species richness of plots as predicted by your model for those treatments that had a significant effect on richness.

# It should clearly demonstrate how the individual treatment factors (not block) operate separately and in concert.

# It should show the actual data points

#setwd("C:/git/coursework/ENEC563")

library(dplyr)

library(nlme)

####Problem 1####

#1\_1)

mossart = read.csv("Moss\_ArthropodSR.csv", skip = 4)

#1\_2)

mossart = mossart %>%

rename("rep"= Replicates, "tr" = Treatment, "A" = Area..cm2., "n" = Species.Richness)

#1\_3

table(mossart$rep, mossart$tr, mossart$A)

#almost perfectly balanced save for the first block in the 20 Area patch

#1\_4

mossart$tr = as.factor(mossart$tr)

mossart$A = as.factor(mossart$A)

mossart$rep = as.factor(mossart$rep)

####Problem 2:Fixed effect mod####

mod1 = lm(n~tr\*A\*rep, data = mossart)

summary(mod1) #just can't happen! 0 vals for the residuals!

anova(mod1)

mod2 = lm(n~tr\*A+rep, data = mossart)

summary(mod2) #explains ~83% of the variation, 73% adjusted

anova(mod2)

mod3 = lm(n~tr+A+rep, data = mossart)

summary(mod3) #explains 83% var again, 75% adjusted

anova(mod3)

#####Problem 3: Mixed effect mod####

mod4 = lme(n~tr\*A,random=~1|rep,data=mossart,control=list(opt="optim"))

summary(mod4) #Tau = 0.04608588, sigma = 3.007672

anova(mod4)

mod5 = lme(n~tr+A,random=~1|rep,data=mossart,control=list(opt="optim"))

summary(mod5) #Tau = 0.04766504, sigma = 2.942209

anova(mod5)

VarCorr(mod4)

vars = as.numeric(VarCorr(mod4)[,1])

corr = vars[1]/sum(vars)

Tau = vars[1]

Sig = vars[2]

pop.mean = fixef(mod4)[3]

#in summary we can see our est for our intercept, the effect, the SE estimates, and p values

#we can also see StdDev and Intercept = Tau; Residual = sigma

#if squared would get sigma and tau squared

#if intercept higher than residual then indicative original model would be a source of huge error

#random = U\_0i -> random effect of pot -> is a normal with mean 0 and variance Tau^2

#within each pot there's a random mean, and we don't know what that is

#there's var in each plant in each pot, and each pot is also going to have random var

#can partition that variation and error into fixed (mean of 0, var of sigma) and random (mean of 0 and var of Tau)

mean.dat <- data.frame(fix.ests=predict(mod2), mix.ests=fitted(mod4)) #make df for plotting this, gen predictions based on plot and plant type

moss2 <- cbind(mossart, mean.dat) #just slap onto original df so real data there too

as.numeric(moss2$rep)

moss2$mix.ests = as.numeric(moss2$mix.ests)

moss2$pop.mean = fixef(mod4)[3]

####Problem 4####

##compare among block to within block by examining dfs & ratio of Tau to sig

#is there an interaction between treatment and area? looks like the interactive models are marginally better than the additive mods in both cases, but not by much at all

#compare mod2 and mod4? fixed vs random/mixed

####Problem 5####

library(ggplot2)

theme\_set(theme\_bw())

ggplot(moss2,aes(x=tr,y=n))+geom\_point(aes(color="Raw data"),size=1)+

labs(x="tr",y="richness",color="")+facet\_wrap(~A)+

geom\_point(aes(y=mix.ests,color="Conditional means"),shape="-", size = 4)+

geom\_point(aes(y=fix.ests,color="Fixed estimates"),shape="-", size = 4) +

geom\_line(aes(y=pop.mean,x=tr,color="Population Mean"),

linetype=2)

# It should show the average species richness of plots as predicted by your model for those treatments that had a significant effect on richness.

# It should clearly demonstrate how the individual treatment factors (not block) operate separately and in concert.

# It should show the actual data points