Potatoes Cold Climate Adaptation

Sina Mokhtar, Behzad FallahiFard

Summary

This report aims to investigate how plants adapt to cold climates by analyzing the effects of three experimental factors (`variety`, acclimatization `regime`, and `temperature`) on damage score for ion leakage.

The report employed Analysis of Variance (ANOVA) to examine the differences among means and determine the effect of each factor on the variable. Boxplots were generated to display differences in mean leak by factor `variety`, while an interaction plot was created to investigate the effect of each factor on the others.

The Analysis showed that the main effects of `variety`, `regime`, and `temperature` each has a statistically significant effect on ion leakage. Additionally, the two-way interaction between `variety` and `regime` has a significant effect on the response ion leakage, suggesting that the effect of `variety` on ion leakage rely on the acclimatization `regime`. The report concluded by discussing the importance of each main effect and interaction term in the model.

Introduction:

The study explores how plants adapt to cold climates by examining the effects of three experimental factors (`variety`, `regime`, and `temperature`) on ion leakage. The investigation was prompted by the observation that plants conditioned to cold climates experience less damage when exposed to cold temperatures. Two potato varieties (s1; s2) were studied, and each plant was assigned to one of two acclimatization regimes levels (C = plant was kept in a cold room; R = plant was kept at room `temperature`) and later subjected to one of two cold temperatures levels (-4 degrees; -8 degrees). The damage scores for photosynthesis and ion leakage were measured, but ion leakage was selected as the response variable for this project. The data is available **here**. Some of the 80 plants assigned initially to the treatment combinations were lost during the experiment. In subsequent sections, we will discuss the methodology used for data analysis, including the statistical models and hypothesis testing. We will then present the analysis results, including the experimental factors' effects on ion leakage. Finally, we will conclude with a discussion of the implications of the findings and suggestions for future research.

Exploration

For investigating the effect of a factor on a variable we can use Analysis of variance (ANOVA) which is a collection of statistical models and their associated estimation procedures used to analyze the differences among means. For following this procedure, we started exploration by making boxplots for the levels of each factor (Figure 1)

It is obvious the mean of leak for varieties 1 (s1) and 2 (s2) are different, showing difference in mean leak by factor `variety`. This seems also hold true for factors `regime` and `temperature`. In the next stage we would like to investigate if the effect of each factor depends on the other factors; The profile plot of sample means (i.e., the interaction plot) (Figure 2) shows a clear interaction between `variety` and `regime`.

the difference in the mean for `regime` R and `regime` C depends on the varieties, with larger `variety` differences for `regime` R than for the `regime` C so it seems there is at least one interaction. This indicates in addition to main effects it is better include interaction effects in the model to analyze the data and inference whether the main effects and interaction effects are statistically significant.
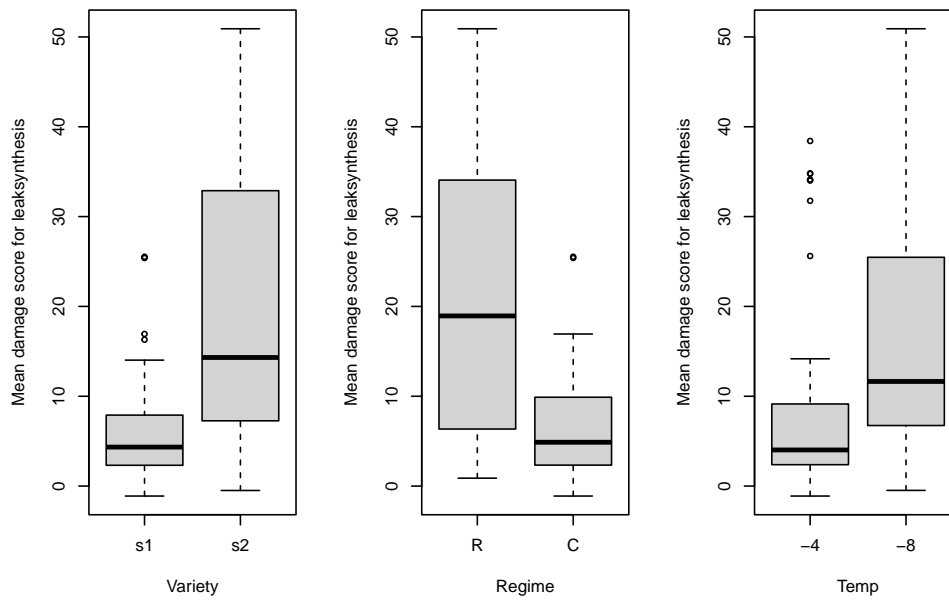
Figure 1: Boxplots of the mean damage score for leaksynthesis by variety, regime, and temperature in potato plants.
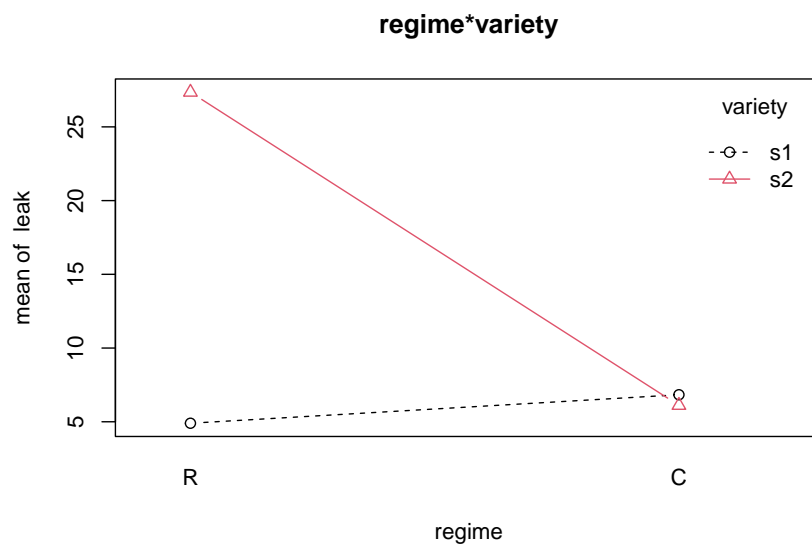

Figure 2: Interaction plot of the mean damage score for leaksynthesis in potato plants, showing the effect of variety and regime on leak values.

Full model

The unbalanced full model that we proceeded to fit is:

$$Leak_{ijkr} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkr}$$

Here, μ is the overall mean, α_i is the effect of level 1 of `variety`, β_j is the effect of level j of `regime`, and γ_k is the effect of level k of `temperature`. $(\alpha\beta)_{ij}$ represents the two-way interaction effect between `variety` and `regime`, $(\alpha\gamma)_{ik}$ represents the two-way interaction effect between `variety` and `temperature`, $(\beta\gamma)_{jk}$ represents the two-way interaction effect between `regime` and `temperature`, and $(\alpha\beta\gamma)_{ijk}$ represents the three-way interaction effect between `variety`, `regime`, and `temperature`. ϵ_{ijkr} represents the error term that cannot be controlled. The indices $i = 1, 2$, $j = 1, 2$, $k = 1, 2$, and $r = 1, 2, \ldots, n_{ijk}$ denote the levels of the `variety`, `regime`, `temperature`, and the number of replicates, respectively.

We checked the assumptions for our initial model and Significant level for all results is 0.05. The normality assumption was not met as seen from the normal QQ plot (Figure 4), and the Shapiro-Wilks Test ($p = 0.03256$). Also, from the plot of the residuals versus the fitted values (Figure 5), and the Breusch Pagan Test ($p = 2.307 \times 10^{-6}$), there is evidence to conclude that the constant variance assumption was not met. We do not know enough about how the data was collected to make a definitive statement about independence, and do not see any reason to question the assumption as it pertains to this data.

Remedial Measures

To address the violations of the model assumptions, a transformation on the response variable was attempted. The Box-Cox transformation suggested log_{10} transformation (Figure 6). We checked the assumptions for our log_{10} transformed model. By fitting the log_{10} transformation, no violation of the equal variance assumption occurred (Figure 8). Normality still appeared to be violated based on the normal QQ plot. Also, according to the Bonferoni Outlier Test of residuals ($p = 0.036851$) the observation number 25 is an outlier. To address the normality assumption violations of the log_{10} transformed model, again a transformation on the log_{10} response variable was attempted. The Box-Cox transformation applied on log_{10} transformed model suggested squared transformation (Figure 9). The two transformations indicate $[log_{10}(leak)]^2$. By fitting the $[log_{10}(leak)]^2$ transformation, no violation of normality (Figure 10) and the equal variance assumption (Figure 11) occurred, there is also no outliers (by this transformation we keep all the observations) and this transformation was applied for the next stage of model selection.

$$[\log_{10}(leak)]^2 = variety + regime + temp + variety : regime : temp + \\ variety : regime + variety : temp + temp : regime \tag{1}$$

The ANOVA table (Table 1) of the following model:

Table 1: Anova Table (Type III tests) for model (1)

Term	Sum_Sq	Df	F_value	P_value
(Intercept)	62.884	1	486.2824	< 2.2e-16 ***
variety	3.782	1	29.2463	9.472e-07 ***
regime	2.743	1	21.2151	1.927e-05 ***
temp	2.487	1	19.2345	4.253e-05 ***
variety:regime	3.830	1	29.6161	8.302e-07 ***
variety:temp	0.046	1	0.3586	0.55134
regime:temp	0.549	1	4.2472	0.04326 *
variety:regime:temp	0.045	1	0.3458	0.55851
Residuals	8.535	66	NA	NA

Reduced Model

The ANOVA table of the adequate transformed full model (Table 1) indicates the three-way interaction `variety:regime:temp` and the two way interaction `variety:temp` are not significant. At the beginning we removed the three-way interaction effect and consider the first reduced model:

Reduced model (R1):

$$[log_{10}(leak)]^2 = variety + regime + temp + variety : regime + variety : temp + temp : regime$$

Reduced model (R1) meets all the model assumptions. We compared model (1) with reduced model (R1). The p-value ($p = 0.5585$) indicates we do not reject the reduced model (R1). In the second step, the ANOVA table of reduced model (R1) indicates the two way interaction `variety:temp` is not significant so we removed it and considered the second reduced model:

Reduced model (R2):

$$[log_{10}(leak)]^2 = variety + regime + temp + variety : regime$$

Reduced model (R2) meets all the model assumptions. We compared model (R1) with model (R2). The p-value $p = 0.5871$ indicates we do not reject the reduced model (R2). In the third step, the ANOVA table of reduced model (R2) indicates the two way interaction `temp:regime` is not significant so we removed it and considered the third reduced model:

Reduced model (R3):

$$[log_{10}(leak)]^2 = variety + regime + temp + variety : regime$$

Reduced model (R3) meets all the model assumptions. We compared model (R2) with model (R3). The p-value $p = 0.05471$ indicates we do not reject the reduced model (R3). We decided to select the third reduced model (R3) as our final model since all its effects are significant and congruent with the assumptions very well.

Final Model

The final model we selected for this data is:

$$[log_{10}(leak)]^2 = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + \epsilon_{ijkr}$$

$i = 1, 2$
$j = 1, 2$

$k = 1, 2$
$r = 1, 2, \ldots, n_{ijk}$

The residuals of the model were plotted against fitted value (Fig. 12) and we took note of there is no outliers. We see the scatter is even overall. The Bonferoni Outlier Test of residuals ($p = 0.56156$) also confirms that. The BP test for equal variance results in a p-value of 0.4549, indicating that the residuals from this model do not show evidence of unequal variance, and confirmed our visual analysis of the plot. The diagnostics on the normality of the residuals points to no violation of this assumption. The normal QQ plot of the residuals (Fig. 11) shows that all values are following the straight line without any significant departure. The Shapiro-Wilks test for normality on these residuals results in a p-value of 0.6591, providing strong evidence that the residuals do follow the normal distribution. The ANOVA table for final mode is:

Table 2: Anova Table (Type III tests) for Final Model

Term	Sum_Sq	Df	F_value	P_value
(Intercept)	63.141	1	478.654	< 2.2e-16 ***
variety	3.734	1	28.309	1.214e-06 ***
regime	2.702	1	20.487	2.438e-05 ***
temp	3.309	1	25.088	4.035e-06 ***
variety:regime	3.884	1	29.445	8.032e-07 ***
Residuals	9.102	69	NA	NA

There is a statistically significant interaction between `variety` and `regime` (p-value = 8.03×10^{-7}), and significant main effects for `variety`, `regime`, and `temperature`. The following table shows the estimations for the fitted final model.

Table 3: Linear regression model summary for the relationship between $[log_{10}(leak)]^2$ and variety, regime, temp, and their interactions

Term	Estimate	Std.Error	t_value	P_value
(Intercept)	0.9965	0.0456	21.878	< 2e-16 ***
variety1	-0.2424	0.0456	-5.321	1.21e-06 ***
regime1	0.2062	0.0456	4.526	2.44e-05 ***
temp1	-0.2117	0.0423	-5.009	4.03e-06 ***
variety1:regime1	-0.2472	0.0456	-5.426	8.03e-07 ***

The sample mean ion leakage damage score for temperatures -4 and -8 are about 9.60 and 17.40 respectively indicating adaptability for `temperature` -4 is better than -8. The difference is statistically significant, according to very small p-value for `temp`. The significant interaction between `variety` and `regime` reinforces that varieties do differ in that the size of the differences in mean ion leakage between regimes depends on varieties. So we made a pairwise comparison for the levels of interaction `variety:regime` instead of main effects `variety` and `regime`.

Table 4: Table showing the mean values of leak for each combination
of variety and regime in the potato dataset.

variety	regime	leak
s1	R	4.901000
s2	R	27.350385
s1	C	6.833600
s2	C	6.113571

variety 2 has significantly worse adaptability to the cold weather when previously conditioned at room
temperature than other combinations of factor levels of variety and regime, explaining where the interaction
is coming from (Figure 3).
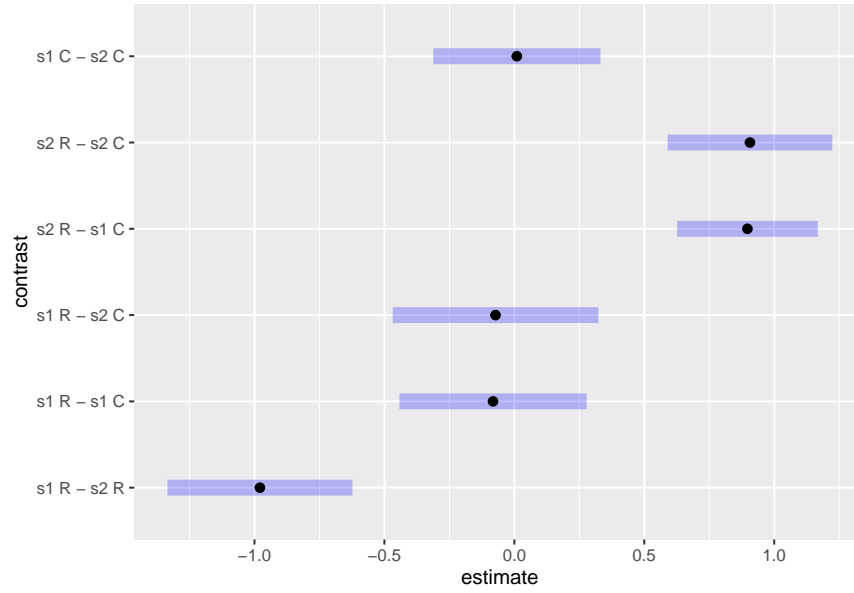
Figure 3: Contrasts plot for pairwise comparisons between variety:regime levels

Conclusion

This report discusses a study on how plants (potatoes) adapt to cold climates by analyzing the effects of
three experimental factors: variety, acclimatization regime, and temperature on ion leakage. The study
used two potato varieties, s1 and s2, and assigned each plant to one of two acclimatization regimes, C and
R, which were later subjected to one of two cold temperature levels, -4 degrees and -8 degrees. The report
employed Analysis of Variance (ANOVA) to examine the differences among means and determine the effect of
each factor on the variable. The analysis showed that the main effects of variety, regime, and temperature
were all significant, indicating that each has an independent effect on ion leakage. Additionally, the two-way
interaction between variety and regime was significant, suggesting that the effect of variety on ion leakage
may rely on the acclimatization regime levels.
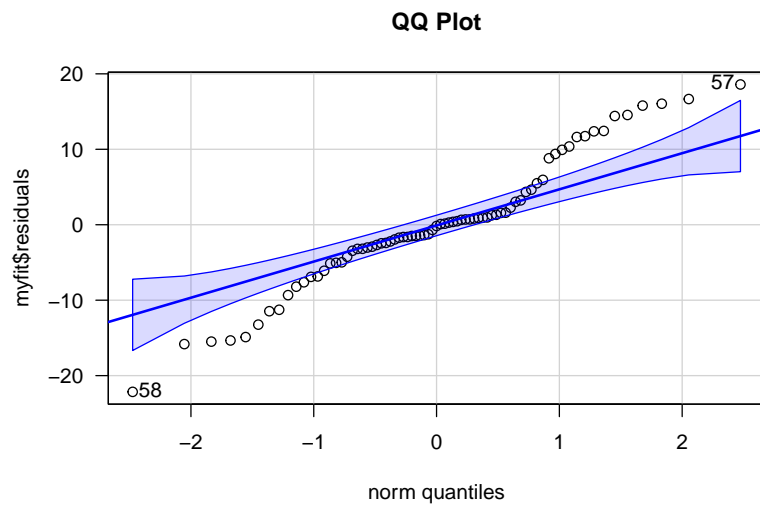
Appendix A
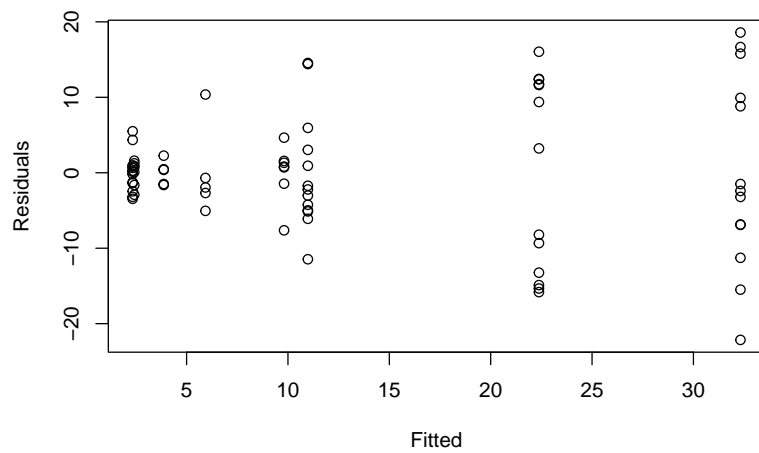

Figure 4: Normal QQ plot of residuals for full model

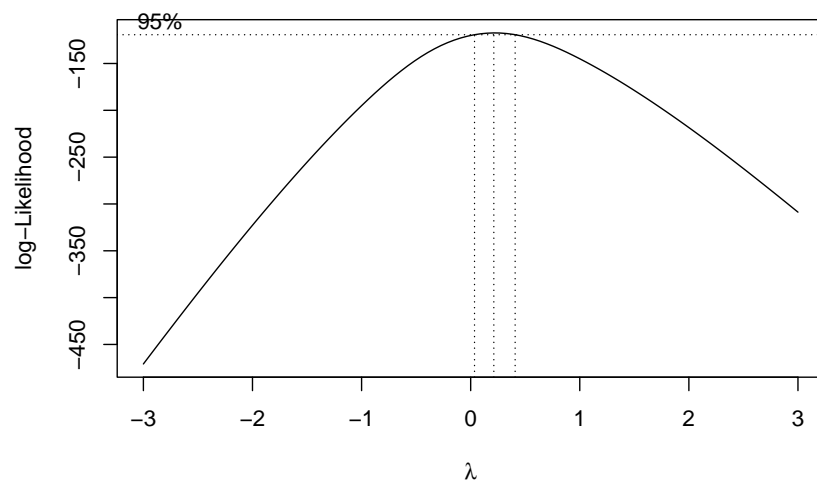Figure 5: The relationship between the fitted values and residuals of a full model

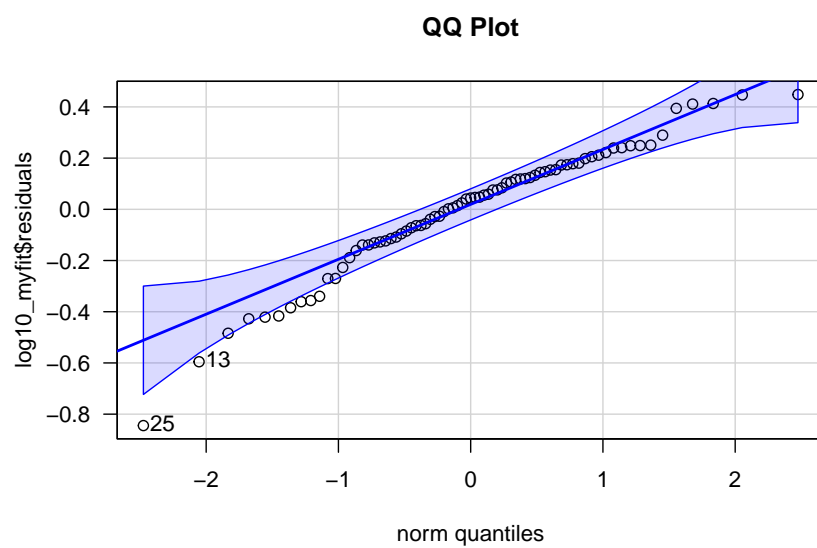Figure 6: The Box-Cox plot for full model

QQ Plot

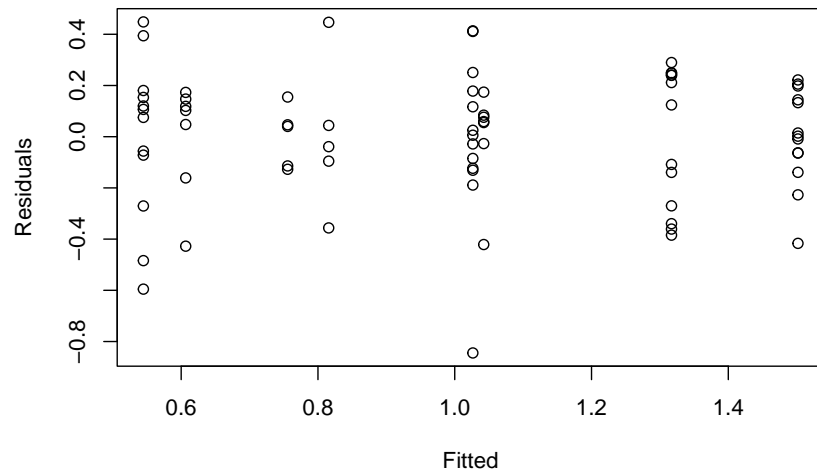Figure 7: Normal QQ plot of residuals for log transformed model

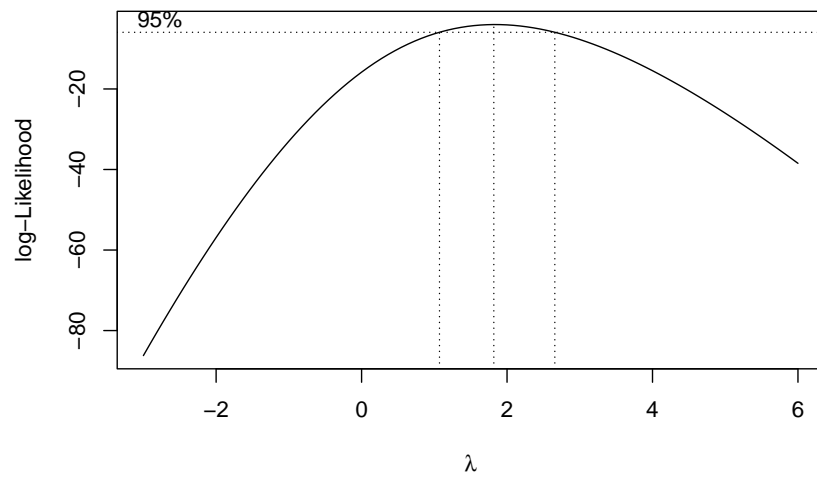Figure 8: The relationship between the fitted values and residuals of a log transformed model

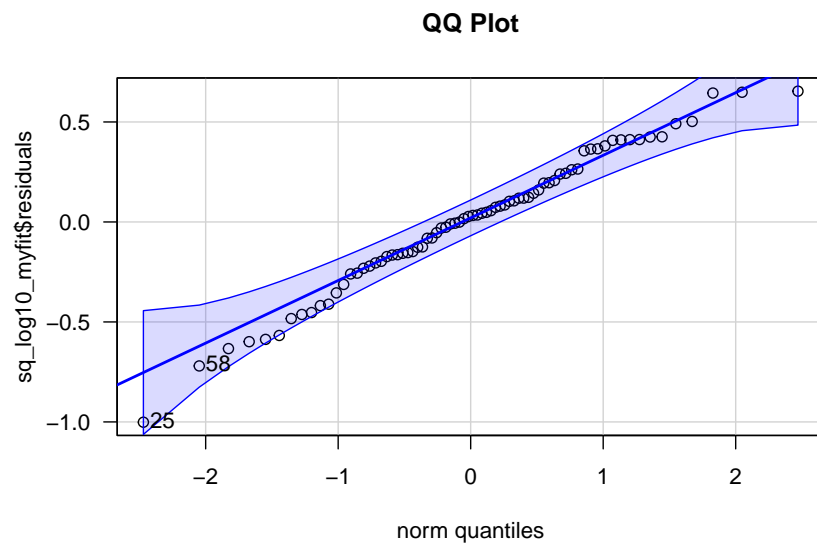Figure 9: The Box-Cox plot for log transformed model

QQ Plot

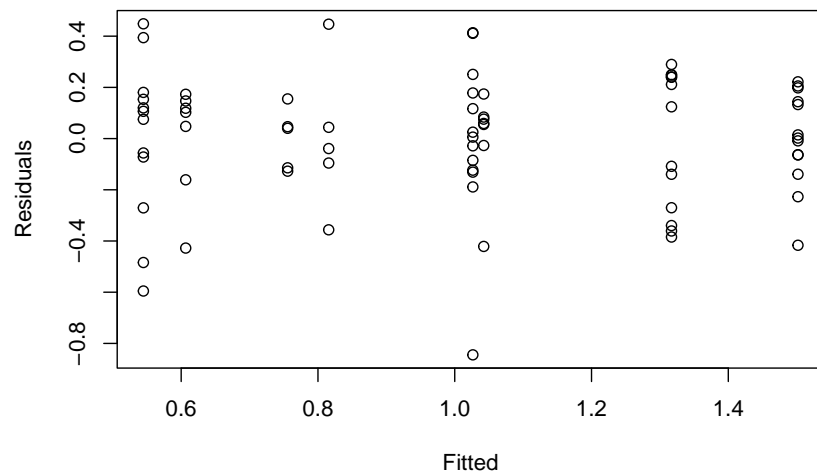Figure 10: Normal QQ plot of residuals for the model (1)

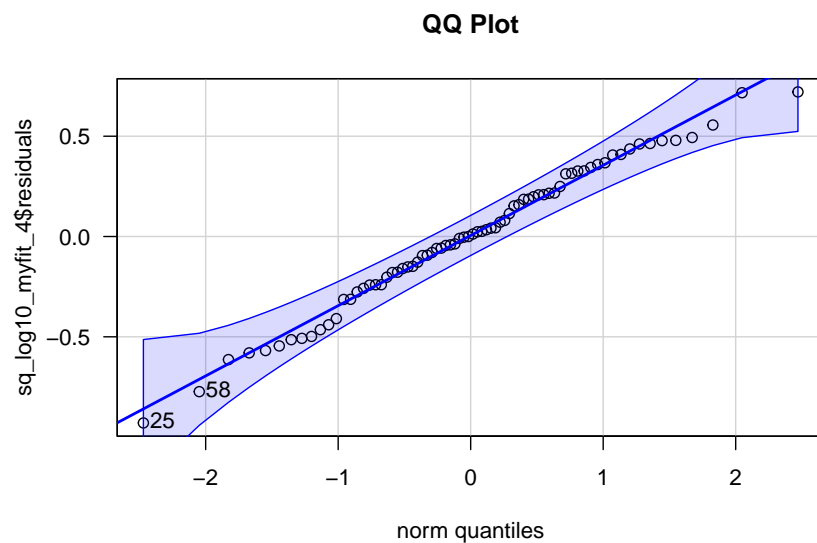Figure 11: The relationship between the fitted values and residuals of the model (1)

QQ Plot

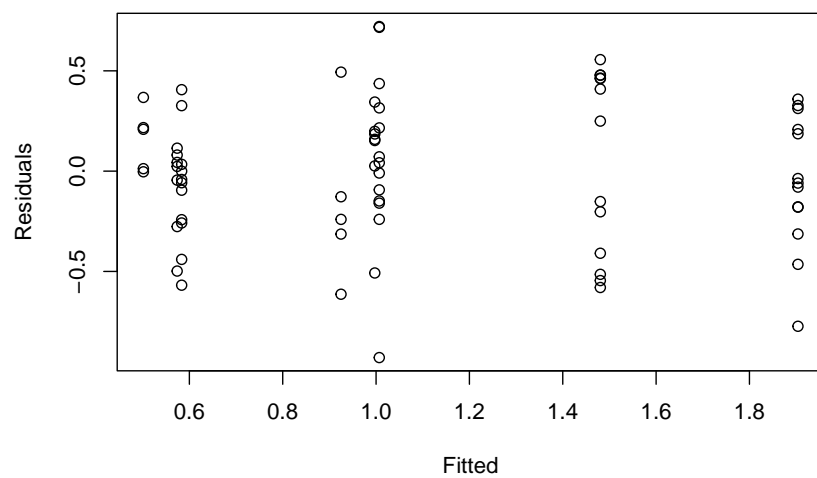Figure 12: Normal QQ plot of residuals for final model

Figure 13: The relationship between the fitted values and residuals for the final model

11

Apendix B

```r
library(lmtest)
library(car)

potato<-read.table("RawData/potato.txt",header=TRUE)
n<-nrow(potato)
n

potato$regime<-factor(potato$regime,label=c("R","C"))
potato$variety<-factor(potato$variety,label=c("s1","s2"))
potato$temp<-factor(potato$temp,label=c("-4","-8"))
attach(potato)
potato[1:10,]

tapply(leak,regime,mean)
tapply(leak,variety,mean)
tapply(leak,temp,mean)
aggregate(leak~variety+regime+temp, data=potato, mean)


#boxplot
par(mfrow=c(1,3))
boxplot(leak ~ variety,
data = potato,
xlab = "Variety",
ylab = "Mean damage score for leaksynthesis")
boxplot(leak ~ regime,
data = potato,
xlab = "Regime",
ylab = "Mean damage score for leaksynthesis")
boxplot(leak ~ temp,
data = potato,
xlab = "Temp",
ylab = "Mean damage score for leaksynthesis")


par(mfrow=c(1,1))
boxplot(leak ~ variety:regime:temp,
data = potato,
xlab = "V:R:T",
ylab = "Mean score of photo")


interaction.plot(regime,variety:temp,leak,type='b',
col=1:4, pch=1:4)


interaction.plot(regime,variety,leak,type='b',
col=1:2, pch=1:2,main="regime*variety")


potato2<-potato[temp==-4,]
```

```r
nrow(potato2)


interaction.plot(potato2$regime,potato2$variety,potato2$leak,type='b',
col=1:2, pch=1:2,main="regime*variety, temp=-4")


##fit full model
myfit<-lm(leak~variety*regime*temp,
contrasts = c(variety=contr.sum, regime=contr.sum,temp=contr.sum))
Anova(myfit,type=3)


qqPlot(myfit$residuals, las = 1, main="QQ Plot")
shapiro.test(myfit$resid)
plot(myfit$fitted,myfit$res,xlab="Fitted",ylab="Residuals")
leveneTest(myfit)
outlierTest(myfit)


min(leak) # -1.11
leak2<- leak+2 #make response positive
myfit2<-lm(leak2~variety*regime*temp,contrasts = c(variety=contr.sum,
regime=contr.sum,temp=contr.sum))
library(MASS)
par(mfrow=c(1,1))
BC <- boxcox(myfit2, lambda = seq(-3, 3, length = 10))

log10_myfit<-lm(log10(leak2)~variety*regime*temp,
contrasts = c(variety=contr.sum, regime=contr.sum,temp=contr.sum))
library(car)
Anova(log10_myfit,type=3)
qqPlot(log10_myfit$residuals, las = 1, main="QQ Plot")


shapiro.test(log10_myfit$resid)
plot(log10_myfit$fitted,log10_myfit$res,xlab="Fitted",ylab="Residuals")
leveneTest(log10_myfit) #can not be appy for reduced models


#perform Breusch-Pagan test
#bptest(log10_myfit)
outlierTest(log10_myfit)

min(log10(leak2)) # -0.05060999
logleak21 <- log10(leak2)+1
log10p1_myfit<-lm(logleak21 ~ variety*regime*temp,
contrasts = c(variety=contr.sum, regime=contr.sum,temp=contr.sum))
par(mfrow=c(1,1))
BC2<-boxcox(log10p1_myfit, lambda = seq(-3, 6, length = 20))

## the second transformation
#sq_log10_leak <- (log10(leak2)+1)^0.5 #+ (log10(leak2)+2)^0.5
```

```
sq_log10_leak <- (log10(leak2))^(1.5)
sq_log10_myfit<-lm(sq_log10_leak~variety*regime*temp,
contrasts = c(variety=contr.sum, regime=contr.sum,temp=contr.sum))
library(car)
Anova(sq_log10_myfit,type=3)

qqPlot(sq_log10_myfit$residuals, las = 1, main="QQ Plot")
shapiro.test(sq_log10_myfit$resid)
plot(sq_log10_myfit$fitted,sq_log10_myfit$res,xlab="Fitted",ylab="Residuals")

leveneTest(sq_log10_myfit)
#bptest(sq_log10_myfit)
outlierTest(sq_log10_myfit)

sq_log10_myfit_2 <- lm(sq_log10_leak ~ variety+regime+temp+variety:regime+variety:temp +temp:regime,
contrasts = c(variety=contr.sum, regime=contr.sum,temp=contr.sum))
anova(sq_log10_myfit,sq_log10_myfit_2)

Anova(sq_log10_myfit_2,type=3)
qqPlot(sq_log10_myfit_2$residuals, las = 1, main="QQ Plot")

shapiro.test(sq_log10_myfit_2$resid)
plot(sq_log10_myfit_2$fitted,sq_log10_myfit_2$res,xlab="Fitted",ylab="Residuals")

#leveneTest(log10_myfit) can not be appy for reduced models
#perform Breusch-Pagan test
bptest(sq_log10_myfit_2)

outlierTest(sq_log10_myfit_2)



sq_log10_myfit_3<- lm(sq_log10_leak~variety+regime+temp+variety:regime +temp:regime,
contrasts = c(variety=contr.sum, regime=contr.sum,temp=contr.sum))
anova(sq_log10_myfit_3,sq_log10_myfit_2)
Anova(sq_log10_myfit_3,type=3)
qqPlot(sq_log10_myfit_3$residuals, las = 1, main="QQ Plot")

shapiro.test(sq_log10_myfit_3$resid)

plot(sq_log10_myfit_3$fitted,sq_log10_myfit_3$res,xlab="Fitted",ylab="Residuals")
#leveneTest(sq_log10_myfit_2) can not be appy for reduced models
#perform Breusch-Pagan test
bptest(sq_log10_myfit_3)
outlierTest(sq_log10_myfit_3)



sq_log10_myfit_4<- lm(sq_log10_leak~variety+regime+temp+variety:regime,
contrasts = c(variety=contr.sum, regime=contr.sum,temp=contr.sum))
anova(sq_log10_myfit_4,sq_log10_myfit_3)

Anova(sq_log10_myfit_4,type=3)
```

```r
qqPlot(sq_log10_myfit_4$residuals, las = 1, main="QQ Plot")

shapiro.test(sq_log10_myfit_4$resid)

plot(sq_log10_myfit_4$fitted,sq_log10_myfit_4$res,xlab="Fitted",ylab="Residuals")

bptest(sq_log10_myfit_4)

outlierTest(sq_log10_myfit_4)

summary(sq_log10_myfit_4)
aggregate(leak ~ temp, data=potato, mean)
aggregate(leak ~ variety + regime, data=potato, mean)

library(lsmeans)
library(multcomp)

comp1<-lsmeans(sq_log10_myfit_4, pairwise ~ variety:regime, adjust="tukey")
plot(comp1$contrasts)
```