# Regression Analysis (Test 2)

Kevin Eng

Fall 2013

document created on: November 21, 2013

created from file: `report2.tex`

cover page automatically created with `CoverPage.sty`

(available at your favourite CTAN mirror)

## Introduction

Forest fires are uncontrolled fires that occur in vegetated regions. These fires have immediate effects on both the surrounding ecology and nearby human settlement. For forest fires that come close human settlements, fire suppression helps prevent large scale property damage and loss of life. However, because forest fires are also part of some ecosystems, knowing which fires can be left to burn out safely is important.

Many metrics have been developed in an attempt to predict forest fires. In this analysis we take a close look at four of the FWI metrics: FFMC, DMC, DC, and ISI and see if they can help predict the total area burned by a fire. We will try to determine if the interaction between rain and temperature play a hand as well. Lastly, we will also attempt to discover where fires occur in Montesinho Natural Park.

## Exploratory Data Analysis/ Initial Modeling

Table 1 list various summary statistics of the 517 recorded forest fires that occurred in Montesinho Natural Park between the years 2000 and 2003. No observation contained incomplete data. Month and days are listed as ordinal categorical variables. January and Monday are coded as 1.

| | Mean | Median | Range | IQR | Variance |
|---|---|---|---|---|---|
| X | 4.67 | 4.00 | 8.00 | 4.00 | 5.35 |
| Y | 4.30 | 4.00 | 7.00 | 1.00 | 1.51 |
| Month | 7.48 | 8.00 | 11.00 | 2.00 | 5.18 |
| Day | 4.26 | 5.00 | 6.00 | 4.00 | 4.30 |
| FFMC | 90.64 | 91.60 | 77.50 | 2.70 | 30.47 |
| DMC | 110.87 | 108.30 | 290.20 | 73.80 | 4101.95 |
| DC | 547.94 | 664.20 | 852.70 | 276.20 | 61536.84 |
| ISI | 9.02 | 8.40 | 56.10 | 4.30 | 20.79 |
| Temp | 18.89 | 19.30 | 31.10 | 7.30 | 33.72 |
| RH | 44.29 | 42.00 | 85.00 | 20.00 | 266.26 |
| Wind | 4.02 | 4.00 | 9.00 | 2.20 | 3.21 |
| Rain | 0.02 | 0.00 | 6.40 | 0.00 | 0.09 |
| Area | 12.85 | 0.52 | 1090.84 | 6.57 | 4052.06 |

Table 1: Summary statistics for recorded forest fires

The mean and median for area are quite far apart given that the IQR is only 6.57. Furthermore, the variance is huge. It dwarfs the IQR. This usually is indicative of either an extreme outlier or a very skewed distribution. Another eye catching detail is that the IQR for rain is 0. This means during most incidents of the 517 recorded forest fires, there was an absence of rain.

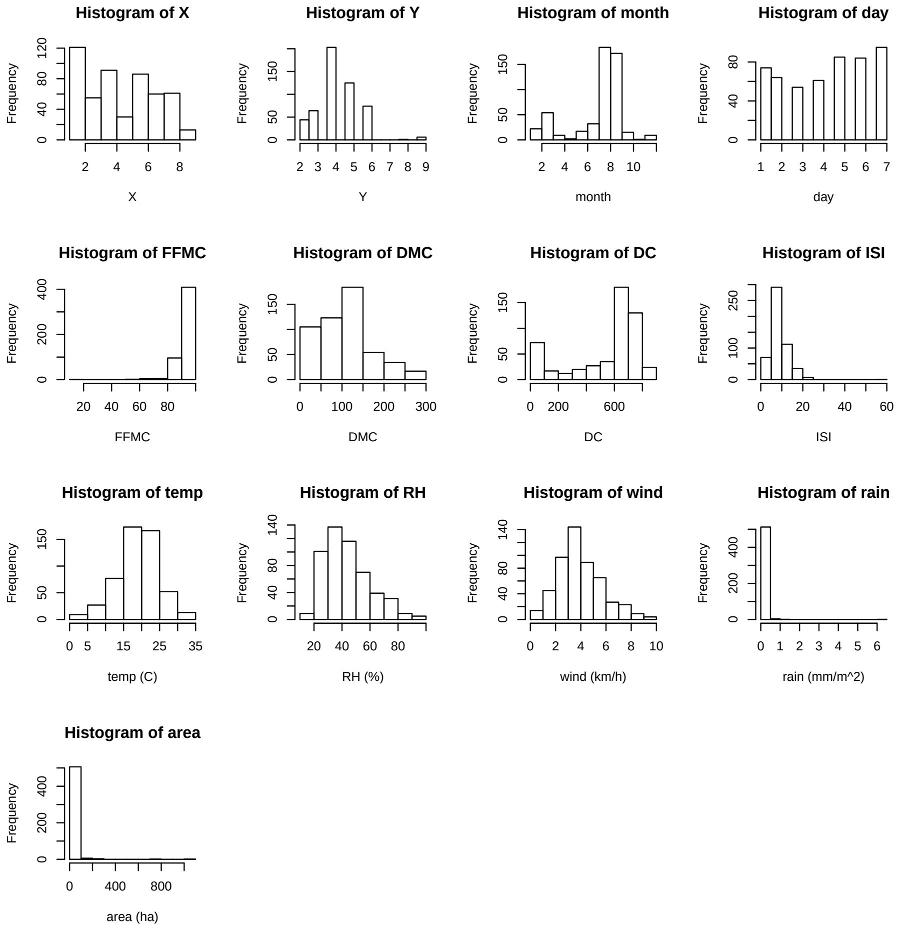


Figure 1: Histogram of variables

The histogram depicted in figure 1) show that area burned, FFMC, and rain have extremely skewed distributions. Most of the variables seem to have a unimodel distribution.

Table 2 is a frequency table for the number of forest fires in a given month. About 83% of forest fires happen in the summer or fall months. This fact suggests the month variable be coded as a two level categorical variable. One level being the summer/fall months and the other being the winter/spring months.

| | jan | feb | mar | apr | may | jun | jul | aug | sep | oct | nov | dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Freq | 2 | 20 | 54 | 9 | 27 | 17 | 32 | 184 | 172 | 15 | 1 | 9 |
| % | 0.4 | 3.9 | 10.4 | 1.7 | 0.4 | 3.3 | 6.2 | 35.6 | 33.3 | 2.9 | 0.2 | 1.7 |

Table 2: Frequency table for months

Table 3 shows that the days on which fires occurred are relatively evenly distributed. Though, there seems to be a slight bias towards the weekend with 51% of fires occurring between Friday and Sunday.

| | mon | tue | wed | thu | fri | sat | sun |
|---|---|---|---|---|---|---|---|
| Freq | 74 | 64 | 54 | 61 | 85 | 84 | 95 |
| % | 14.3 | 12.4 | 10.4 | 11.8 | 16.4 | 16.2 | 18.4 |

Table 3: Frequency table for days

Table 4 shows there is a slight bias towards gird locations where X is either 4 or 6. Not many fires seem to have occurred in locations where X was 9.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Freq | 48 | 73 | 55 | 91 | 30 | 86 | 60 | 61 | 13 |
| % | 9.3 | 14.1 | 10.6 | 17.6 | 5.8 | 16.6 | 11.6 | 11.8 | 2.5 |

Table 4: Frequency table for X

Table 5 reveals almost all recorded fires occurred in grid locations where $Y$ was between 2 and 6.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Freq | 0 | 44 | 64 | 203 | 125 | 74 | 0 | 1 | 6 |
| % | 0 | 8.5 | 12.4 | 39.3 | 24.2 | 14.3 | 0 | 0.2 | 1.2 |

Table 5: Frequency table for Y

Figure 2 depicts the pairs plot of the recorded variables along with their correlation.

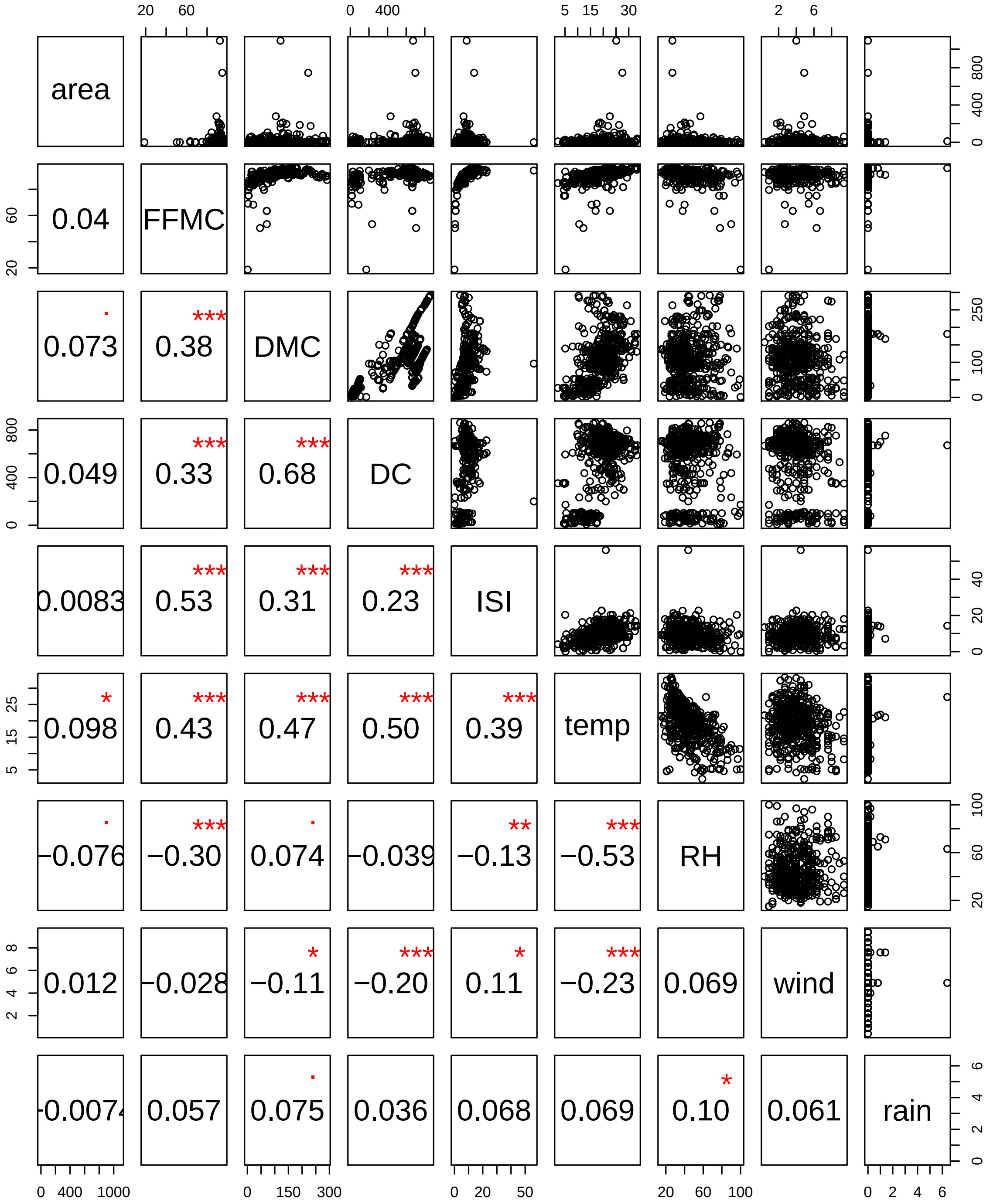


Figure 2: Pairs plot

From looking at figure 2, there does not seem to be any strong relationship between area and any of the possible regressors. Furthermore the four components of FWI are highly correlated among each other. This could lead to serious multicollinearity issues.

Figure 3 depicts area vs the categorical variables: X, Y, month, and day at each level. The box plots are skewed in the positive direction. The median indicates a majority of forest fires at any given level are small and tend not burn much land. Variability among the day of the week is slightly jittery, and seems to vary widely among the months.

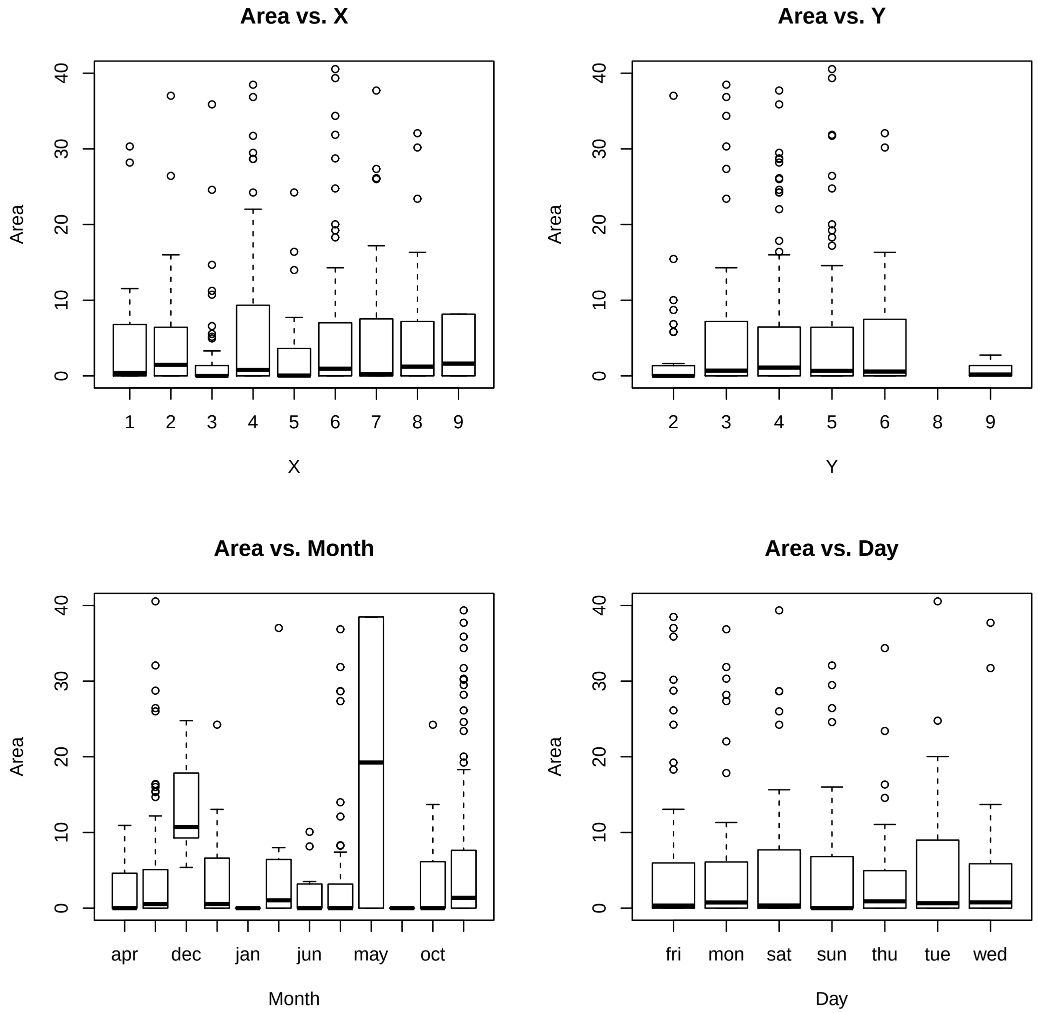


Figure 3: Boxplots

To build the initial model, we construct a fairly general model by including many regressors. Since rain was present 8 out of the 517 times, we treat rain as a two level categorical variable. We also break the days of the week into the two categories: the weekend and the work day. To prevent the model from becoming overly complicated, the X and Y coordinates are treated as ordinal categorical variables. This way they account for 2 terms as oppose to 16. With the previous comments in mind, the initial model is:

$$\begin{aligned}\widehat{\text{Area}} = \beta_0 + \beta_1\text{FFMC} + \beta_2\text{DMC} + \beta_3\text{DC} + \beta_4\text{ISI} + \beta_5\text{Month} + \beta_6\text{Day} + \beta_7\text{Temp} + \beta_8\text{RH} \\ + \beta_9\text{Rain} + \beta_{10}\text{Wind} + \beta_{11}\text{Rain*Temp}\end{aligned}$$

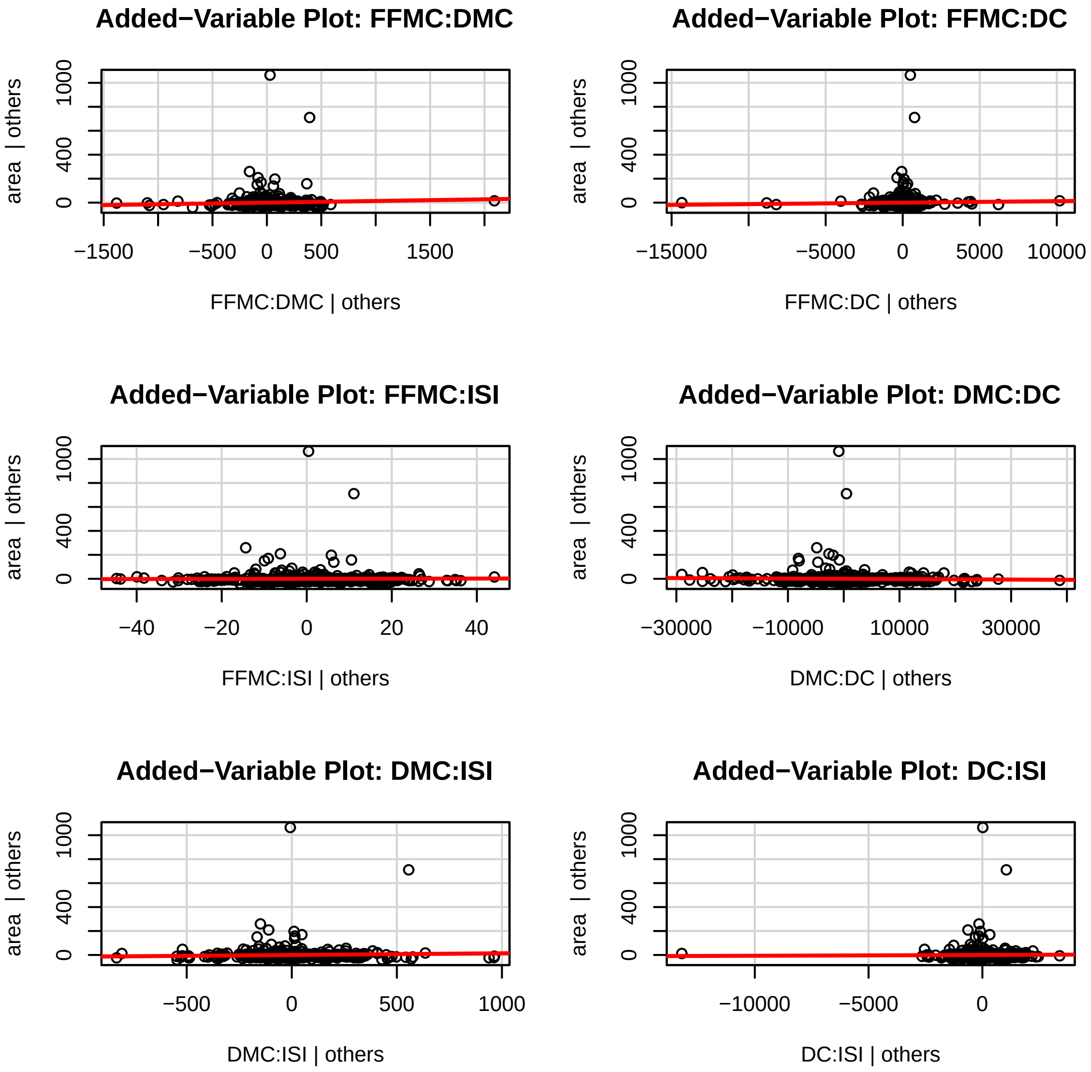
Where

$$\text{Month} = \begin{cases} 1 & \text{winter or spring months} \\ 0 & \text{summer or fall moths} \end{cases}$$

$$\text{Day} = \begin{cases} 1 & \text{Monday, Tuesday, Wednesday, Thursday} \\ 0 & \text{Friday, Saturday, Sunday} \end{cases}$$

$$\text{Rain} = \begin{cases} 1 & \text{if rain} > 0 \\ 0 & \text{if rain} = 0 \end{cases}$$

None of the added variable plots (figure 4) show any distinctive tends so we leave out interaction terms for the four FWI predictors.



Added−Variable Plot: FFMC:DMC

Added−Variable Plot: FFMC:DC

Added−Variable Plot: FFMC:ISI

Added−Variable Plot: DMC:DC

Added−Variable Plot: DMC:ISI

Added−Variable Plot: DC:ISI

Figure 4: Added variables plot

Figure 5 is the plot for residuals vs fitted values The plot clearly shows a pattern: a downward cutting line that has a slope around -1. There is also a very distinct Pac-Man mouth shape that indicates there may be a serious problem with the assumption of constant variance. Figure 6 shows the residual plot for FFMC. It has a very sharp peak when FFMC is between 90 and 95. The plots for the other FWI predictors share similar issues regarding constant variance.

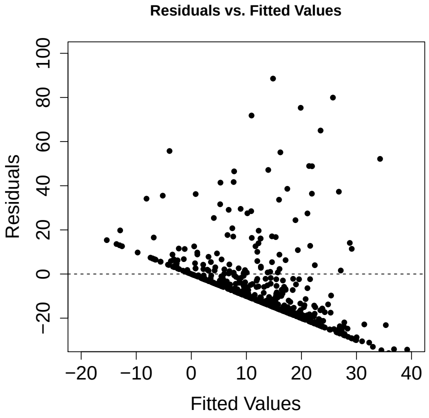


Figure 5: Residual plot

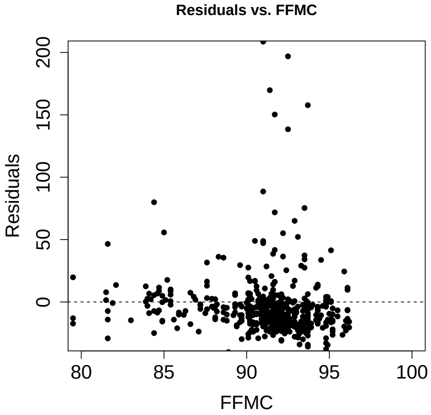


Figure 6: Residual Plot

The QQ-plot in figure 7 indicates the underlying distribution is positively skewed. The box-cox plot suggests we might be able to obtain a better fit by taking the predictor to the $-1/2$ power.

To identify possible outliers we studentize the residuals. Table 6 lists all the observations whose

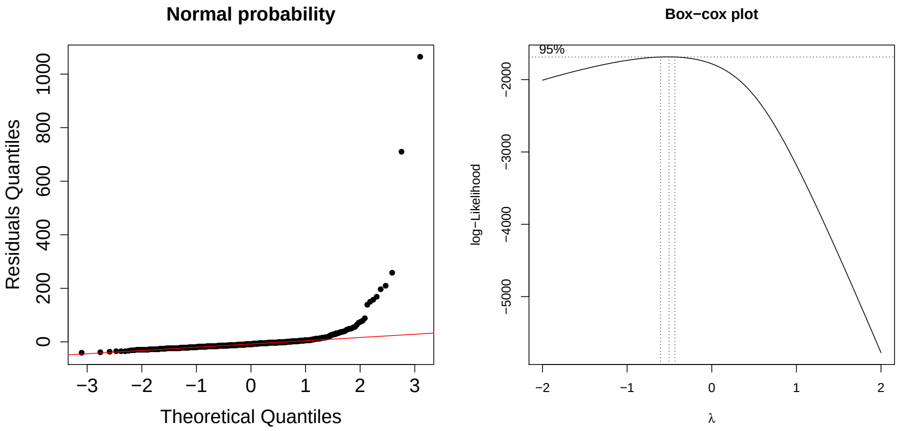


Figure 7: QQ-plot and Box-Cox plot

studentized residual is greater than 3.

| Obs # | Residual |
|---|---|
| 237 | 3.1 |
| 238 | 3.3 |
| 480 | 4.2 |
| 416 | 13.0 |
| 239 | 25.3 |

Table 6: Studentized Residuals

Observation 239 has a very large studentized residual. However removal of this observation along with the 4 other possible outliers does not change the model much. Furthermore, because the data does not indicate there was an error in data collection for these observation, these 5 potential outliers will be kept in the model.

# Diagnostics

The final model is built using the backwards method. The general idea is to keep removing the variable with the smallest $t$ value (in terms of absolute value) until all variables in the model have $p$ values less than a certain threshold. In this case there is one special case to consider. During the removal process, temperature was associated with having the smallest $t$ value. However the initial model contains a temp/rain interaction variable. Since the removal of both temperature and its interaction term was not significant on the partial $F$ test, they were both thrown out. The final model generated by the backwards method when we require $p < .1$ is:

$$\widehat{\text{Area}} = \beta_0 + \beta_1\text{Wind} + \beta_2\text{Rain} + \beta_3\text{Month} \tag{2}$$

Figure 9 shows the residual plots of model (2).

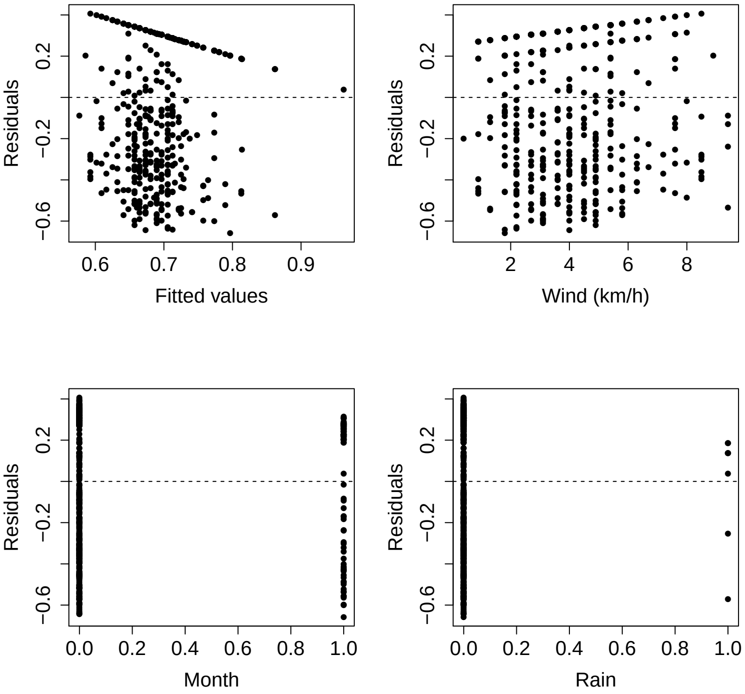


Figure 8: Transformed Residual Plots

There is still a distinctive line in the plot of residuals versus fitted values. Problems regarding the assumption of constant variance persist. However, the residual plots no longer have extreme points. Indeed the studentized residuals for model (2) are all less than 3 in magnitude.

The QQ-plot in figure 9 indicates that the transformation did not change the underlying error distribution in the desired manner. The pattern depicted suggests the distribution is more tail heavy relative to the normal. The small hump at 0 also suggests the distribution may be bimodel.

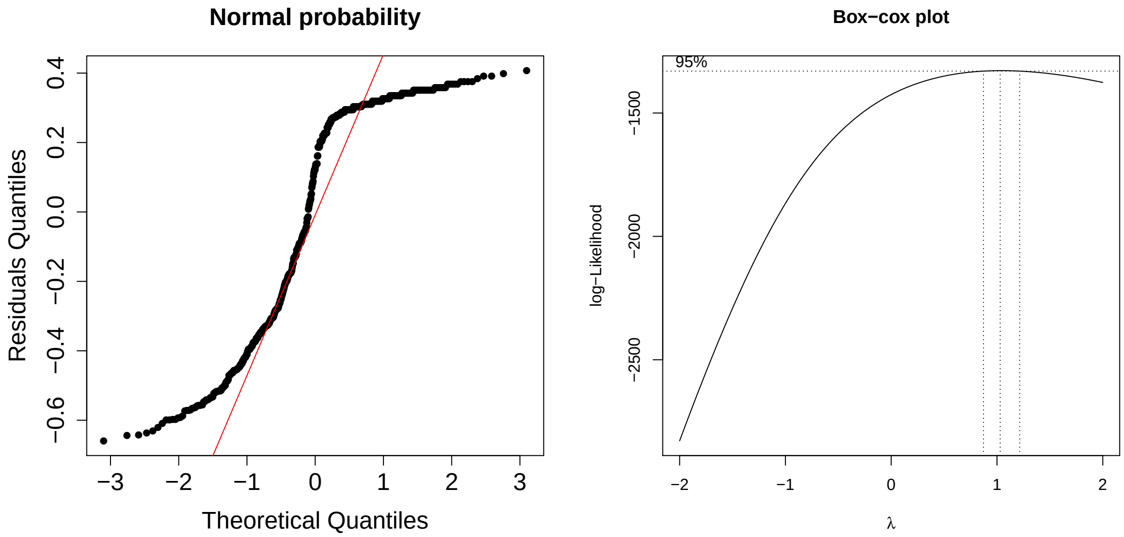


Figure 9: Transformed QQ-plot and Box-Cox plot

## Model Inference and Results

To determine how resources should be allocated spatially, we look at the scatter plot of $Y vs.X$ which is depicted in figure 10

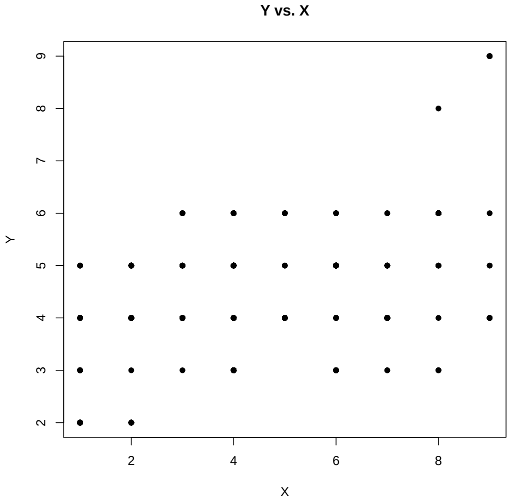


Figure 10: Scatter plot of Y vs. X

The plot demonstrates a very distinctive pattern: a vast majority of the 517 recorded forest fires occurred for $Y \leq 6$. Table 5 indicates that 1.4% of recorded forest fires happened in locations where $Y > 6$. Likewise looking at table 4, the grid locations with $X = 9$ totaled 2.5% of recorded forest fires. At least 9% of recorded forest fires occur in grid locations where $X < 9$. With this in mind, resources should be allocated so grid locations with $1 \leq X \leq 8$ and $Y \leq 6$ have the most support.

The model selection process did not find day of the week as a significant factor in determining the total area burned. However, there is evidence it influences whether or not a forest fire occurs. A chi squared test on the contingency table for days produces a $\chi^2$ value of 18.02 on 6 degrees of freedom. This corresponds to a $p$ value of about .006. Hence at an $\alpha$ level of .01 we conclude there is a statistically significant relationship between the day of the week and when forest fires occur.

The coefficients in the final model are as follows:

| Coefficient | Estimate | Std. Error | $p$ value |
|---|---|---|---|
| $\beta_0$ | .745 | .036 | $2\times10^{-10}$ |
| $\beta_1$ | -.0179 | .008 | .031 |
| $\beta_2$ | .205 | .119 | .086 |
| $\beta_3$ | .0840 | .039 | .033 |

Table 7: Estimates of various model parameters

The interpretation for each coefficient is:

$\beta_0$ : A forest fire would burn $(.745)^{-2} = 1.8$ $ha$ of land if there was no wind.

$\beta_1$ : Holding other variables constant, the response would decrease by $-.0179$ $ha^{-1/2}$ for every 1 km/h increase in wind speed.

$\beta_2$ : Compared to no rain, having rain increase the expected response by .205 $ha^{-1/2}$.

$\beta_3$ : Compared to summer and fall, forest fires in winter and spring increase the expected response by .0840 $ha^{-1/2}$.

The adjusted $R^2$ for this model is .014 which means this model will probably be quite bad at making predictions.

## Conclusion and Discussion

During the selection processes, none of the four FWI statistics or the interaction term between rain and temperature were significant in determining total area burned. There is no evidence to support that the amount of area burned is somehow influenced by the FWI statistics or the interaction of rain and temp.

A likely reason why more fires seem to start on the weekends is because this is probably the time most people go to the national park. The weekend would be an excellent time for, say, a barbecue.

This analysis suggests forest fires occurring during the summer and fall months happen more frequently and with more intensity. A possible explanation is that during the summer months long periods of dry weather can leave vegetation dried out. The dry vegetation in turn provides fuel for forest fires. During the fall months it is possible the turning of leaves provides fuel for fires even if the dry weather stops.

Table 8 shows that the $p$ values for the coefficients of wind, month, and rain are all less than .1. However, it would be unwise to conclude the results are statistically significant at the $\alpha = .1$ level. There is no reason to believe the model errors are normally distributed. Figure 11 strongly suggests the underlying distribution has heavier tails than the normal distribution. Since the $p$ value is based off the assumption that the underlying distribution is normal, it is likely the $p$ value seen in table 8 is smaller than the true $p$ value. At best, we can conclude this analysis points to wind, month, and rain being more likely to play a role in predicting the amount of area burned compared to the other possible predictors.

On a positive note, the final model at least makes sense. Since the response in model (2) is in $ha^{-1/2}$, a negative slope is associated with increasing $ha$. By similar logic, an increase in $\beta_0$ is associated with a decrease in $ha$. On other words model (2) says higher winds are associated with larger fires. And fires that occur in the winter/spring seasons or just after it has rained tend to be smaller.

The model may improve by only considering events where the amount of burned area is greater than .1 *ha*. It is possible forest fires need to reach a certain threshold in size before they can self sustain. Another idea is to look only at areas with with similar characteristics. In our data we were given only one value for each statistic to represent the entirety of the national park. In reality, because of geographic features such as mountains, the statistics can vary wildly based on location.

## Appendix

```
#Initialize

rm(list=ls(all=TRUE))

setwd("C:/Users/Kevin/Google Drive/36-401/TEST REPORT 2")

mydata <- read.csv(file="forestfires.csv",head=TRUE,sep=",")

attach(mydata)

library(MASS)

library(car)


#Format Data

Month = NULL

Month[length(month)] = NA

Month[month == "jan"] = 1

Month[month == "feb"] = 2

Month[month == "mar"] = 3

Month[month == "apr"] = 4

Month[month == "may"] = 5

Month[month == "jun"] = 6

Month[month == "jul"] = 7

Month[month == "aug"] = 8

Month[month == "sep"] = 9

Month[month == "oct"] = 10

Month[month == "nov"] = 11

Month[month == "dec"] = 12

month = Month


Day = NULL
```

```
Day[length(day)] = NA

Day[day == "mon"] = 1

Day[day == "tue"] = 2

Day[day == "wed"] = 3

Day[day == "thu"] = 4

Day[day == "fri"] = 5

Day[day == "sat"] = 6

Day[day == "sun"] = 7

tempvar = day

day = Day


#Generate Summary Statistics

vars = cbind(names(mydata))

M = matrix(nrow = length(names(mydata)),ncol = 5)

for(i in 1:length(names(mydata))){

M[i,1] = mean(get(vars[i]))

M[i,2] = median(get(vars[i]))

M[i,3] = max(get(vars[i])) - min(get(vars[i]))

M[i,4] = IQR(get(vars[i]))

M[i,5] = var(get(vars[i]))

}

round(M,2)


#Frequence Tables

table(month)

round(100*table(month)/length(month),1)

table(day)

round(100*table(day)/length(day),1)
```

```
table(X)
round(100*table(X)/length(X),1)
table(Y)
round(100*table(Y)/length(Y),1)

#Histograms
par(mfrow=c(4,4))
vars = cbind(names(mydata))

for(i in 1:8){
hist(get(vars[i]),breaks = 10,xlab = vars[i],main = paste("Histogram of", vars[i]
, sep = " "))
}
i =9
hist(get(vars[i]),breaks = 10,xlab = "temp (C)",main = paste("Histogram of", vars[i]
, sep = " "))
i = 10
hist(get(vars[i]),breaks = 10,xlab = "RH (%)",main = paste("Histogram of", vars[i]
, sep = " "))
i = 11
hist(get(vars[i]),breaks = 10,xlab = "wind (km/h)",main = paste("Histogram of", vars[i]
, sep = " "))
i = 12
hist(get(vars[i]),breaks = 10,xlab = "rain (mm/m^2)",main = paste("Histogram of", vars[i]
, sep = " "))
i = 13
hist(get(vars[i]),breaks = 10,xlab = "area (ha)",main = paste("Histogram of", vars[i]
, sep = " "))
```

```
#Initial Model
day = tempvar
Day = NULL
Day[length(day)] = NA
Day[day == "mon"] = 1
Day[day == "tue"] = 1
Day[day == "wed"] = 1
Day[day == "thu"] = 1
Day[day == "fri"] = 0
Day[day == "sat"] = 0
Day[day == "sun"] = 0
day = Day

summerFallMonth = ifelse(Month >= 6 & Month <= 12,0,1)
Rain = ifelse(rain > 0,1,0)

reg.line = lm(area~ X + Y + FFMC + DMC + DC + ISI + wind + Rain + summerFallMonth
+ day + temp + RH + Rain*temp)
summary(reg.line)
pdf(file = "initqqplot.pdf", width = 12, height = 6)
par(mfrow = c(1,2))
qqnorm(reg.line$res, pch = 16, main = "Normal probability", ylab = "Residuals Quantiles"
, cex.lab=1.5, cex.axis=1.5, cex.main=1.5, cex.sub=1.5)
qqline(reg.line$res, col = 2)
boxcox(area+1~ X + Y + FFMC + DMC + DC + ISI + wind + rain + summerFallMonth
+ day + temp + RH)
title(main = "Box-cox plot")
```

```
par(mfrow = c(1,1))
dev.off()


par(mfrow = c(4,4))
plot(reg.line$fit, reg.line$res, xlab = "Fitted values", ylab = "Residuals", pch = 16)
abline(h = 0, lty = 2)
plot(reg.line$res, ylab = "Residuals", pch = 16)
abline(h = 0, lty = 2)
plot(X, reg.line$res, xlab = "X",ylab = "Residuals", pch = 16)
abline(h = 0, lty = 2)
plot(Y, reg.line$res, xlab = "Y",ylab = "Residuals", pch = 16)
abline(h = 0, lty = 2)
plot(FFMC, reg.line$res, xlab = "FFMC",ylab = "Residuals", pch = 16)
abline(h = 0, lty = 2)
plot(DMC, reg.line$res, xlab = "DMC",ylab = "Residuals", pch = 16)
abline(h = 0, lty = 2)
plot(DC, reg.line$res, xlab = "DC",ylab = "Residuals", pch = 16)
abline(h = 0, lty = 2)
plot(ISI, reg.line$res, xlab = "ISI",ylab = "Residuals", pch = 16)
abline(h = 0, lty = 2)
plot(wind, reg.line$res, xlab = "Wind (km/h)",ylab = "Residuals", pch = 16)
abline(h = 0, lty = 2)
plot(temp, reg.line$res, xlab = "Temp (C)",ylab = "Residuals", pch = 16)
abline(h = 0, lty = 2)
plot(RH, reg.line$res, xlab = "RH (%)",ylab = "Residuals", pch = 16)
abline(h = 0, lty = 2)
plot(day, reg.line$res, xlab = "Day",ylab = "Residuals", pch = 16)
abline(h = 0, lty = 2)
```

```
plot(summerFallMonth, reg.line$res, xlab = "Month",ylab = "Residuals", pch = 16)
abline(h = 0, lty = 2)
plot(Rain, reg.line$res, xlab = "Rain",ylab = "Residuals", pch = 16)
abline(h = 0, lty = 2)

r = rstudent(reg.line)
sort(r)

#A Closer Look
plot(FFMC, reg.line$res, xlab = "FFMC", xlim = range(80:100)
, ylim = range(-30:200),ylab = "Residuals", pch = 16
, cex.lab=1.5, cex.axis=1.5, cex.main=1.5, cex.sub=1.5)
title("Residuals vs. FFMC")
abline(h = 0, lty = 2)
plot(reg.line$fit,reg.line$res, xlab = "Fitted Values", xlim = range(-20:40),
 ylim = range(-30:100),ylab = "Residuals", pch = 16
, cex.lab=1.5, cex.axis=1.5, cex.main=1.5, cex.sub=1.5)
title("Residuals vs. Fitted Values")
abline(h = 0, lty = 2)

#Added Variables Plots
par(mfrow = c(3,2))
reg.line = lm(area ~ X + Y + FFMC + DMC + DC + ISI + wind + rain + summerFallMonth
 + day + temp + RH + FFMC*DMC)
avPlot(reg.line, FFMC:DMC)
reg.line = lm(area ~ X + Y + FFMC + DMC + DC + ISI + wind + rain + summerFallMonth
 + day + temp + RH + FFMC*DC)
avPlot(reg.line, FFMC:DC)
```

```
reg.line = lm(area ˜ X + Y + FFMC + DMC + DC + ISI + wind + rain + summerFallMonth
 + day + temp + RH + FFMC*ISI)
avPlot(reg.line, FFMC:ISI)
reg.line = lm(area ˜ X + Y + FFMC + DMC + DC + ISI + wind + rain + summerFallMonth
 + day + temp + RH + DMC*DC)
avPlot(reg.line, DMC:DC)
reg.line = lm(area ˜ X + Y + FFMC + DMC + DC + ISI + wind + rain + summerFallMonth
 + day + temp + RH + DMC*ISI)
avPlot(reg.line, DMC:ISI)
reg.line = lm(area ˜ X + Y + FFMC + DMC + DC + ISI + wind + rain + summerFallMonth
 + day + temp + RH + DC*ISI)
avPlot(reg.line, DC:ISI)

reg.line = lm(area ˜ X + Y + FFMC + DMC + DC + ISI + wind + rain + summerFallMonth
 + day + temp + RH + rain*temp)
avPlot(reg.line, rain:temp , cex.lab=1.5, cex.axis=1.5, cex.main=1.5, cex.sub=1.5)

#Transformed Model
reg.line = lm(1/sqrt(area+1)˜ wind + Rain + summerFallMonth)
summary(reg.line)
pdf(file = "transqqplot.pdf", width = 12, height = 6)
par(mfrow = c(1,2))
qqnorm(reg.line$res, pch = 16, main = "Normal probability", ylab = "Residuals Quantiles"
, cex.lab=1.5, cex.axis=1.5, cex.main=1.5, cex.sub=1.5)
qqline(reg.line$res, col = 2)
boxcox(1/sqrt(area+1)˜ wind + Rain + summerFallMonth)
title(main = "Box-cox plot")
par(mfrow = c(1,1))
```

```
dev.off()

pdf(file = "transResid.pdf", width = 8, height = 8)

par(mfrow = c(2,2))

plot(reg.line$fit, reg.line$res, xlab = "Fitted values", ylab = "Residuals", pch = 16

 , cex.lab=1.5, cex.axis=1.5, cex.main=1.5, cex.sub=1.5)

abline(h = 0, lty = 2)

plot(wind, reg.line$res, xlab = "Wind (km/h)",ylab = "Residuals", pch = 16

, cex.lab=1.5, cex.axis=1.5, cex.main=1.5, cex.sub=1.5)

abline(h = 0, lty = 2)

plot(summerFallMonth, reg.line$res, xlab = "Month",ylab = "Residuals", pch = 16

 , cex.lab=1.5, cex.axis=1.5, cex.main=1.5, cex.sub=1.5)

abline(h = 0, lty = 2)

plot(Rain, reg.line$res, xlab = "Rain",ylab = "Residuals", pch = 16

 , cex.lab=1.5, cex.axis=1.5, cex.main=1.5, cex.sub=1.5)

abline(h = 0, lty = 2)

dev.off()


#Backwards Method

reg.line = lm(1/sqrt(area+1)~ X + Y + FFMC + DMC + DC + ISI + wind + Rain

+ summerFallMonth + day + temp + RH + Rain*temp)

summary(reg.line)

reg.line = lm(1/sqrt(area+1)~ X + Y + FFMC + DMC + DC + ISI + wind + Rain

+ summerFallMonth + day + RH)

summary(reg.line)

reg.line = lm(1/sqrt(area+1)~ X + Y + FFMC + DMC + ISI + wind + Rain

+ summerFallMonth + day + RH)

summary(reg.line)

reg.line = lm(1/sqrt(area+1)~ X + Y + FFMC + DMC + ISI + wind + Rain
+ summerFallMonth + RH)
summary(reg.line)
reg.line = lm(1/sqrt(area+1)˜ X + FFMC + DMC + ISI + wind + Rain
+ summerFallMonth + RH)
summary(reg.line)
reg.line = lm(1/sqrt(area+1)˜ X + FFMC + DMC + ISI + wind + Rain
+ summerFallMonth)
summary(reg.line)
reg.line = lm(1/sqrt(area+1)˜ X + FFMC + ISI + wind + Rain
+ summerFallMonth)
summary(reg.line)
reg.line = lm(1/sqrt(area+1)˜ X + ISI + wind + Rain + summerFallMonth)
summary(reg.line)
reg.line = lm(1/sqrt(area+1)˜ X + wind + Rain + summerFallMonth)
summary(reg.line)
reg.line = lm(1/sqrt(area+1)˜ wind + Rain + summerFallMonth)
summary(reg.line)
```