RPS data analysis for Niceville WWTF

<http://www.frecologic.com>

This document reports the analyses done for the MSWRF WQBEL project. This was last assembled:

2016-08-23 16:24:18

First, run the following code. Also load pander and set options for outputting for proper tables:

make function for geomean and geosd:

# Background

FDEP plans to place the creek recieving spray field effluent from Niceville on the study list due to RPS failures. Frydenborg Ecologic was contracted to better characterize the stream, as all previous sampling points were taken frome the same area.

Frydenborg Ecologic sampled five sites along the stream for RPS and canopy cover. These results were combined with previous sampling done by Frydenborg Ecologic, and with sampling conducted by FDEP.

# Methods

## Field

FDEP RPS protocol was followed.

## Statistical

The response variable of interest is the percent of RPS scores greater than 4. Most of these analyses are interested in examining RPS response at each transect of the 11 taken during a typical FDEP RPS, though each overall site result is also considered. For transect level analyses, the number of 4, 5, and 6 scores are counted (max of 9), then the percentage is calculated by dividing the count by 9 and multiplying by 100.

Canopy cover was taken with a densiometer, with a maximum of 96 points. Percent canopy cover was calculated by dividing the obverseved number of canopy points out of the maximum of 96, then multiplying by 100.

### Data import

A CSV file containing all RPS observations, along with metadata, was imported into R. Dataframe was manipulated

### Data cleaning

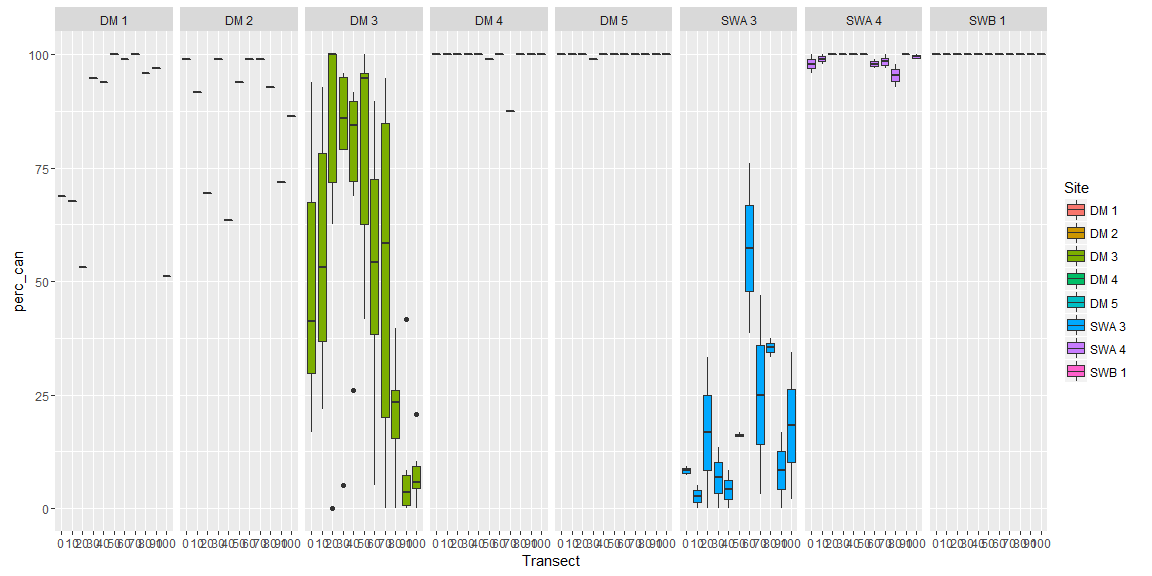
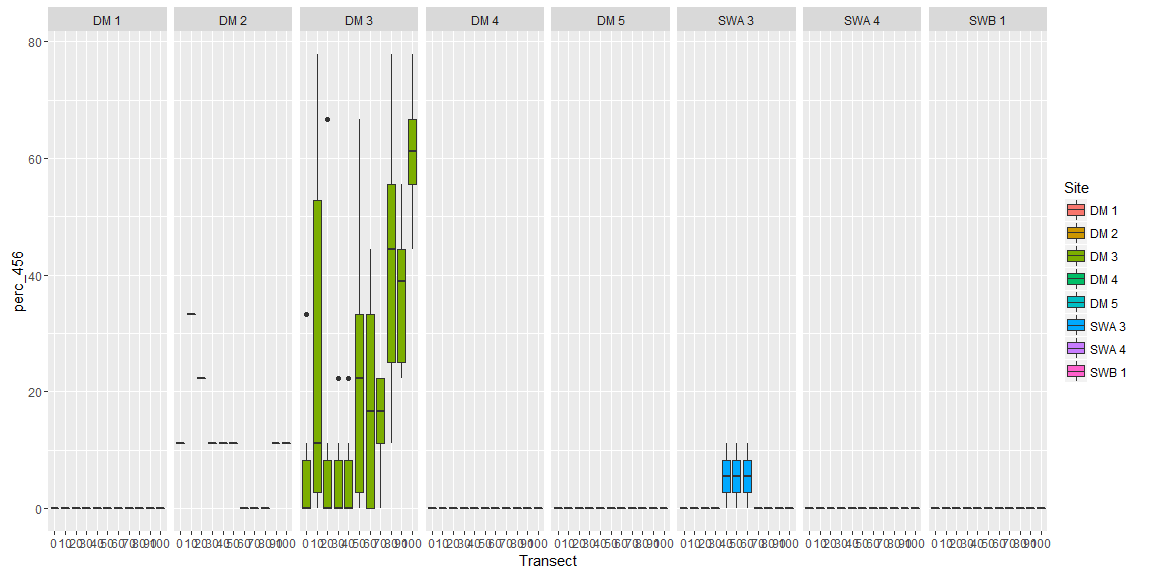
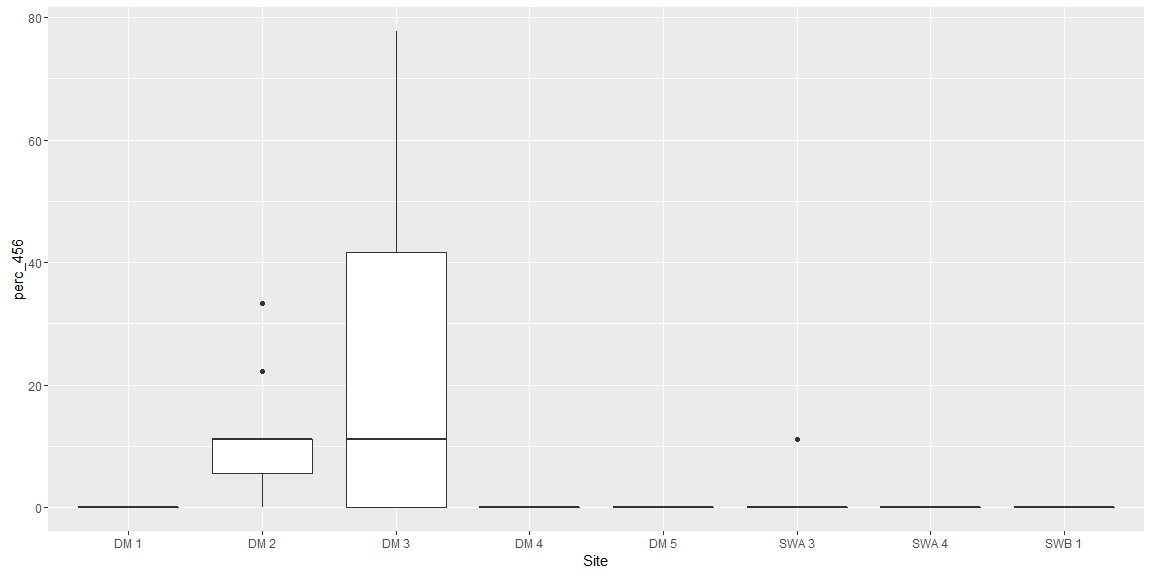
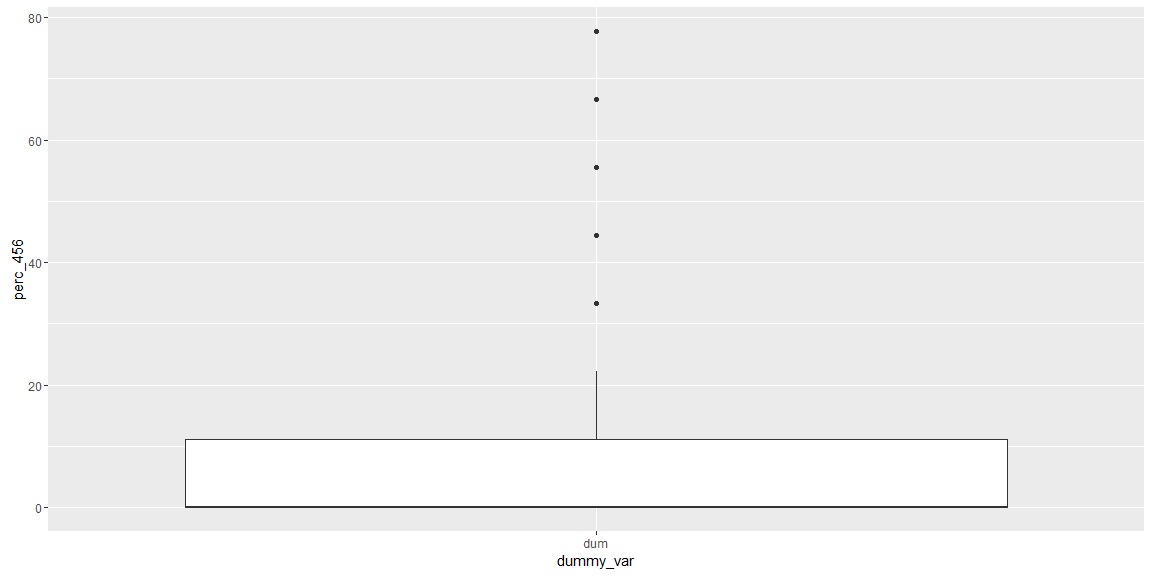
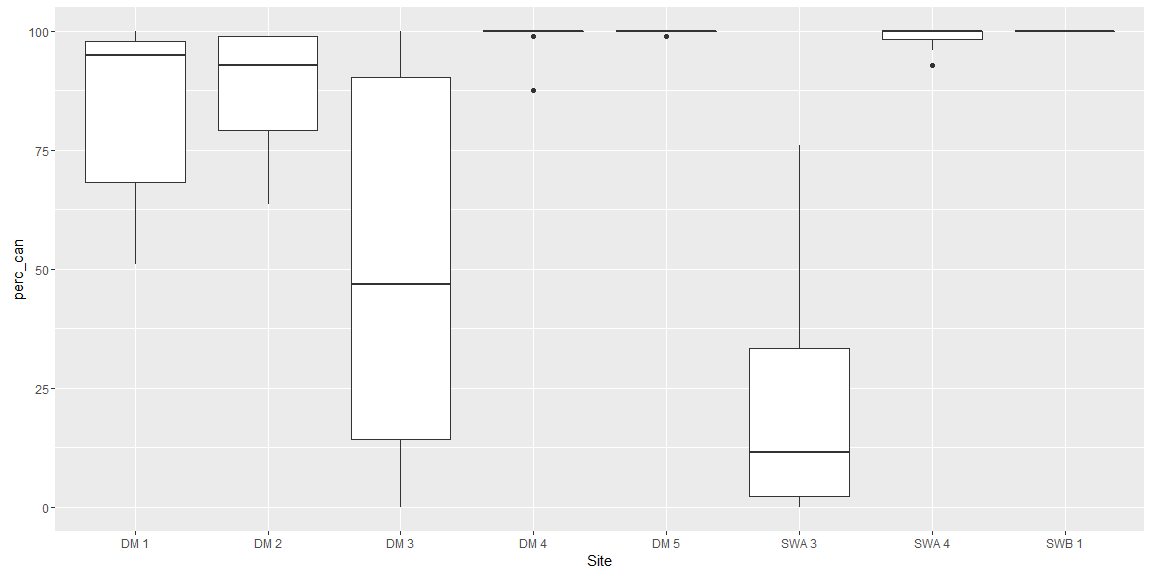
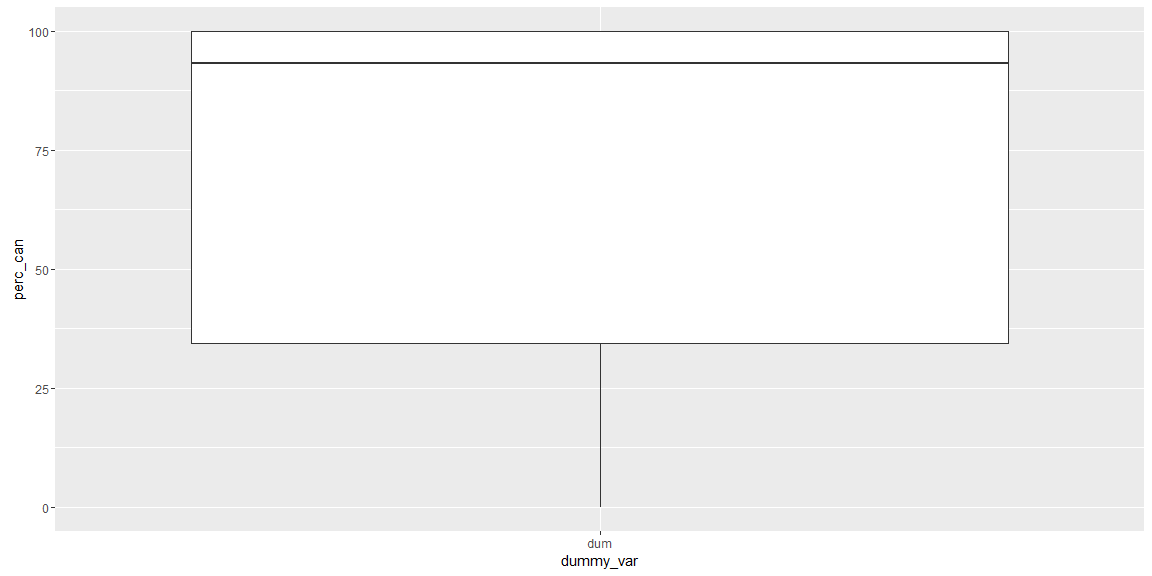
# Evaluating hypotheses for RPS response

## Excess nitrogen leads to RPS

FDEP wants to list Deer Moss Creek for failing the RPS, blaming the excess TN coming from the spray field. If this is the case, one should see a relationship between failure of the RPS and amount of nitrogen in the water. A linear regression using the average RPS percent 4-6 per site along with the annual geometric mean of TN was performed. This resulted in 4 data points: the upstream control of Deer Moss Creek, and the three streams that recieve effluent. Only Deer Moss Creek has any sort of RPS response. There is no relationship apparent between nutrient level in the stream and RPS observed.

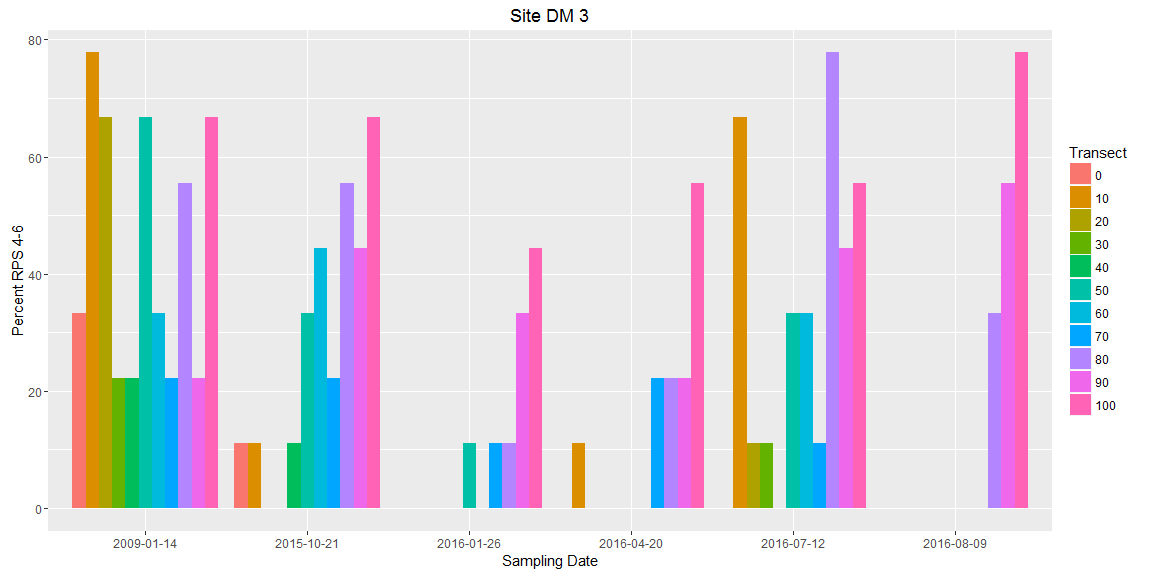
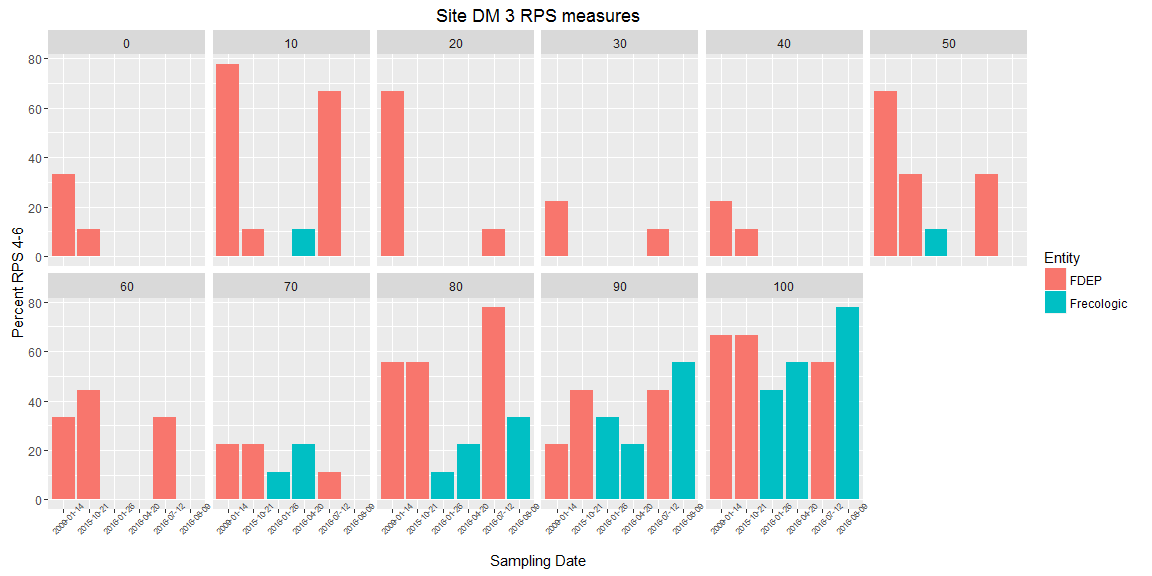
## lack of Canopy cover leads to RPS

## Check assumptions for linear regression



### Data exploration of DM 3

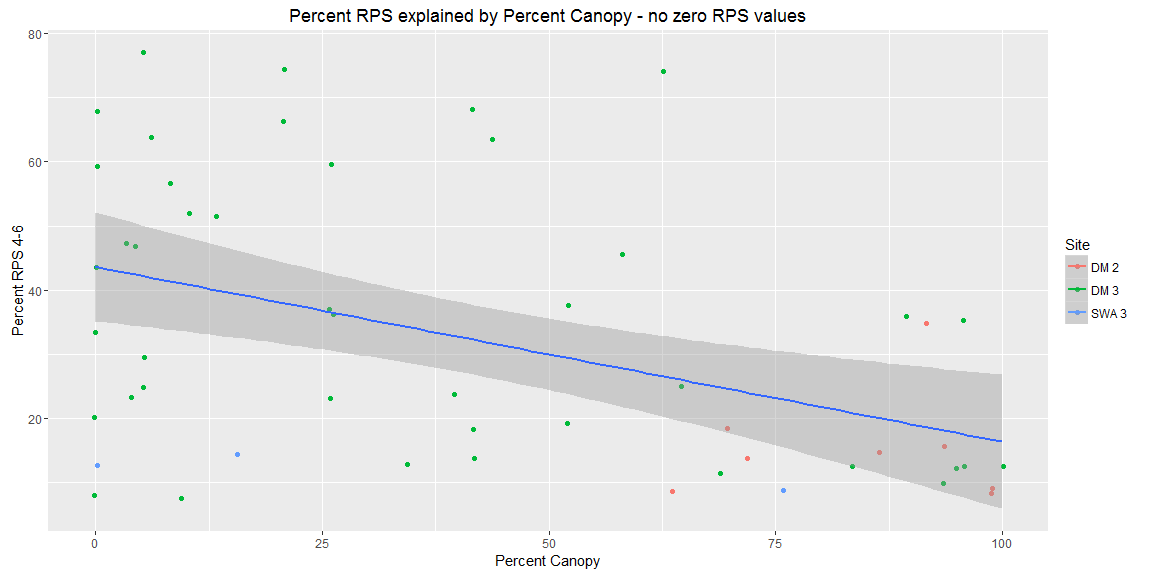
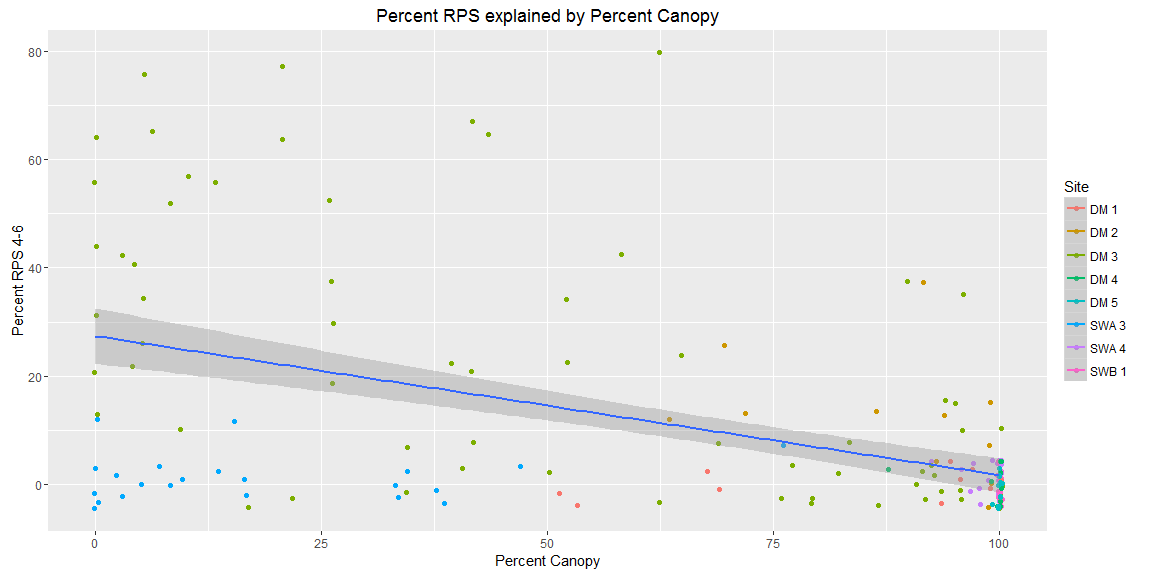
Exploratory plots of site DM 3, the only site sampled by both Frydenborg Ecologic and FDEP.



# Results

## relationship between canopy cover and rps at each transect

The percent of 4, 5, and 6 algea was determined at each transect. This is then matched to the canopy reading determined at the transect using a densiometer. A linear relationship is examined.



R squared and other stats for percent 4, 5, 6 explained by percent canopy coverage

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| **perc\_can** | -0.26 | 0.033 | -7.8 | 4.8e-13 |
| **(Intercept)** | 27 | 2.6 | 11 | 1.3e-20 |

Fitting linear model: perc\_456 ~ perc\_can

|  |  |  |  |
| --- | --- | --- | --- |
| Observations | Residual Std. Error |  | Adjusted |
| 176 | 16 | 0.26 | 0.26 |

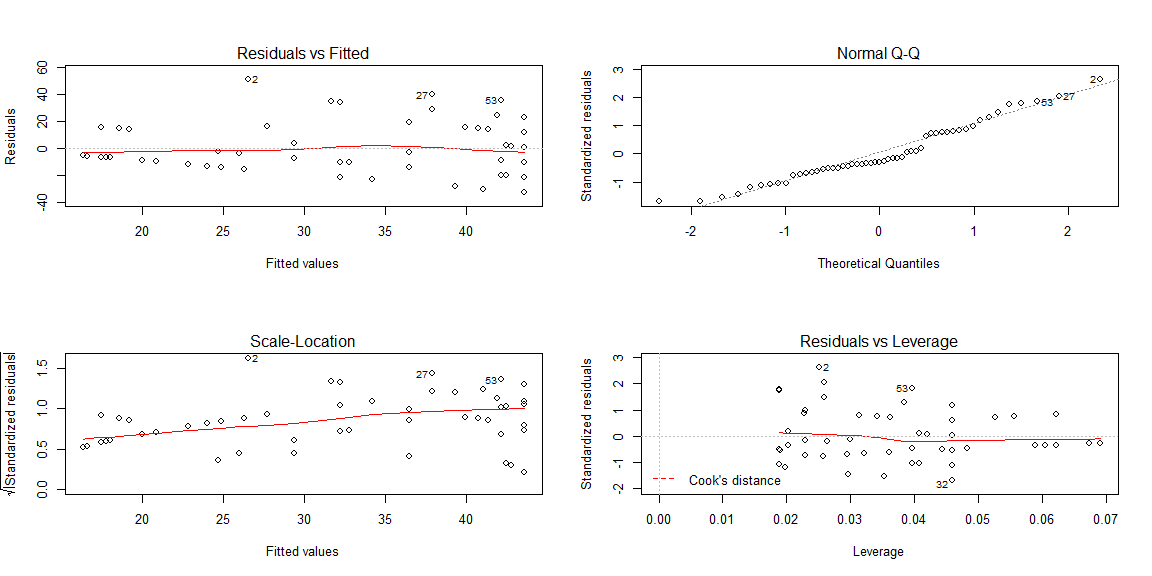
## 

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| **perc\_can** | -0.27 | 0.077 | -3.6 | 0.00082 |

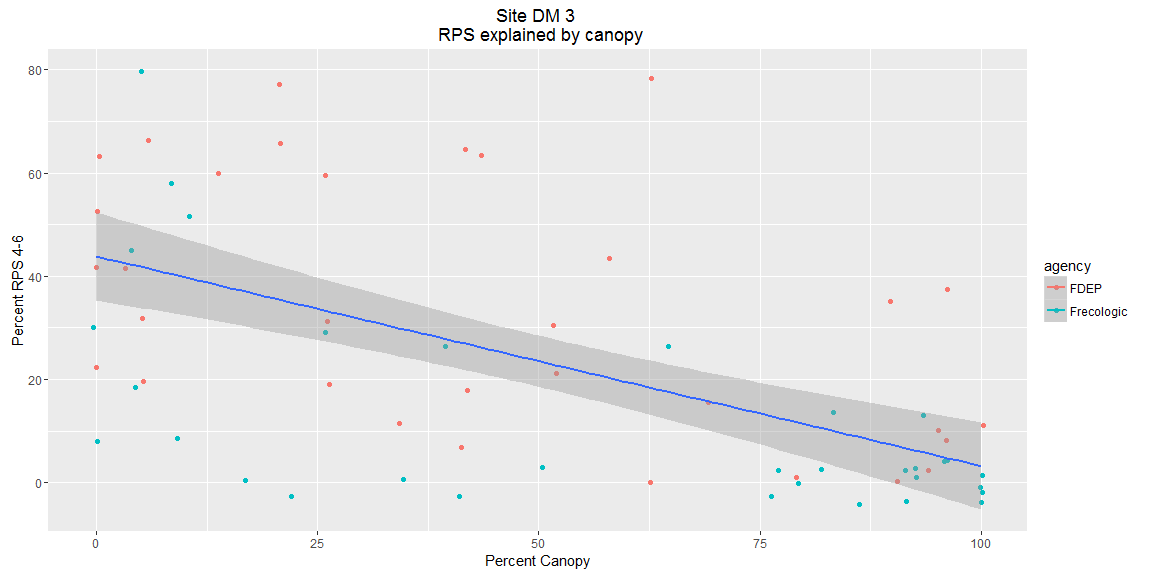
## **(Intercept)** 44 4.2 10 4.5e-14

Fitting linear model: perc\_456 ~ perc\_can

|  |  |  |  |
| --- | --- | --- | --- |
| Observations | Residual Std. Error |  | Adjusted |
| 53 | 20 | 0.2 | 0.18 |



Examining site DM 3 data only.

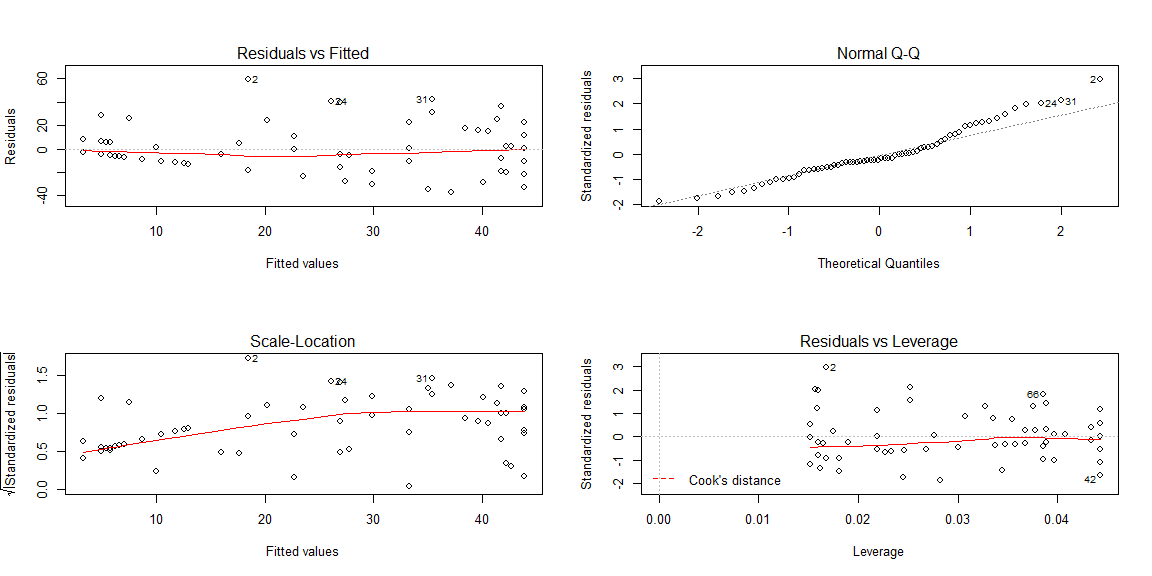


Examine the R squared and intercept for DM 3 only

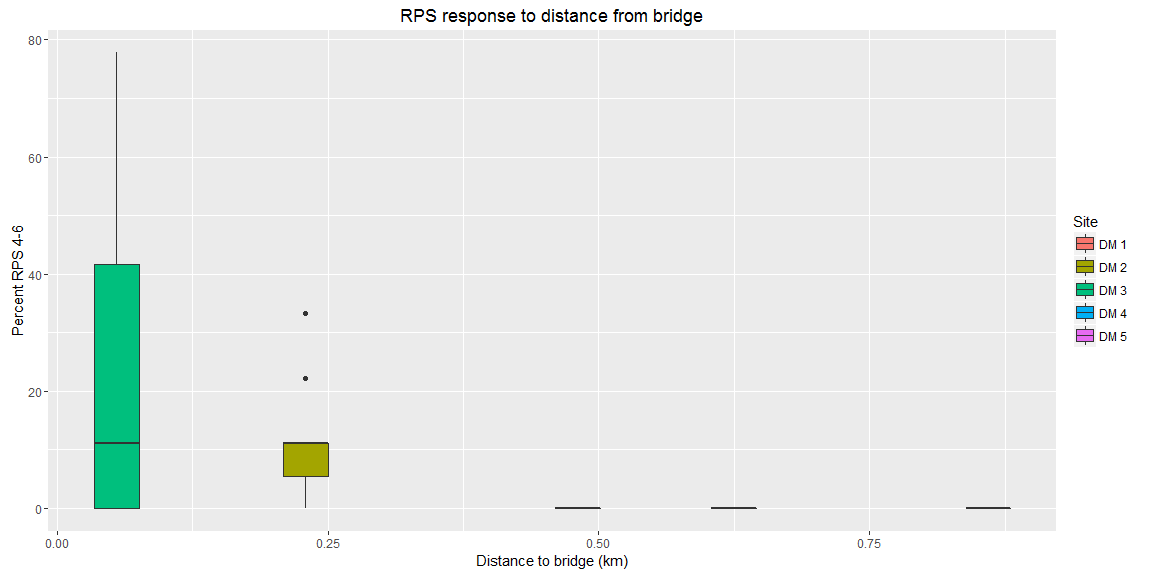
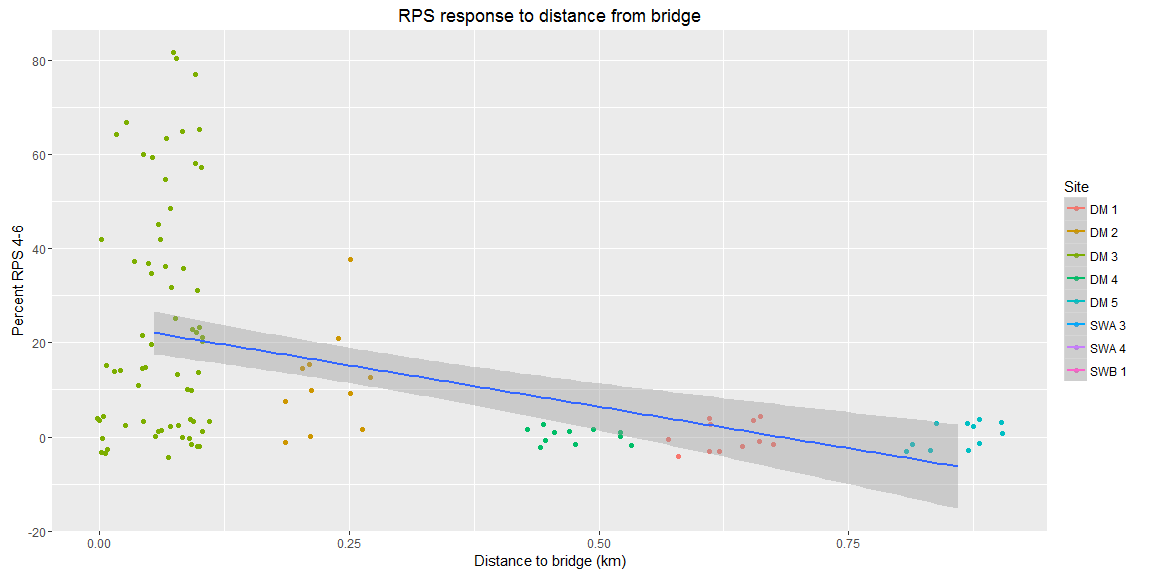
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| **perc\_can** | -0.41 | 0.069 | -5.9 | 1.3e-07 |
| **(Intercept)** | 44 | 4.3 | 10 | 3.2e-15 |

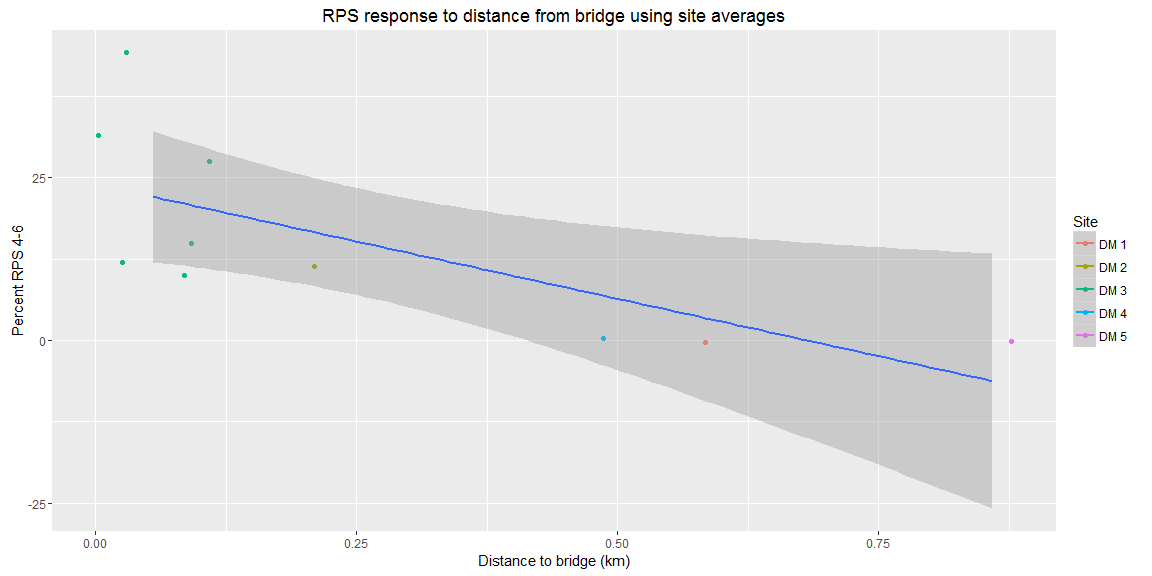
Fitting linear model: perc\_456 ~ perc\_can

|  |  |  |  |
| --- | --- | --- | --- |
| Observations | Residual Std. Error |  | Adjusted |
| 66 | 20 | 0.35 | 0.34 |



## Relationship between distance to highway and RPS/canopy



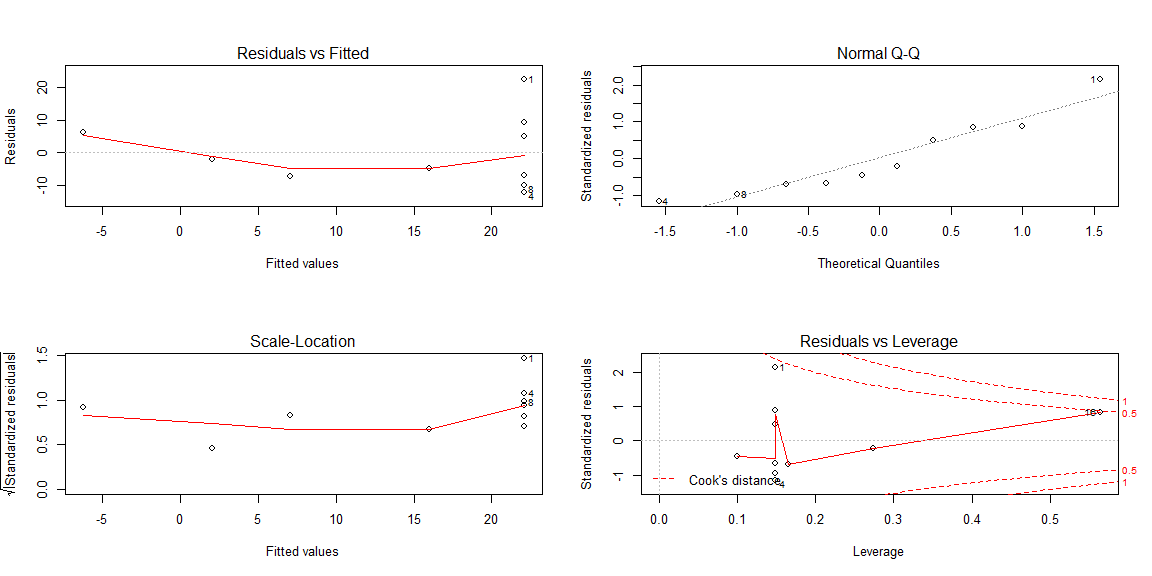


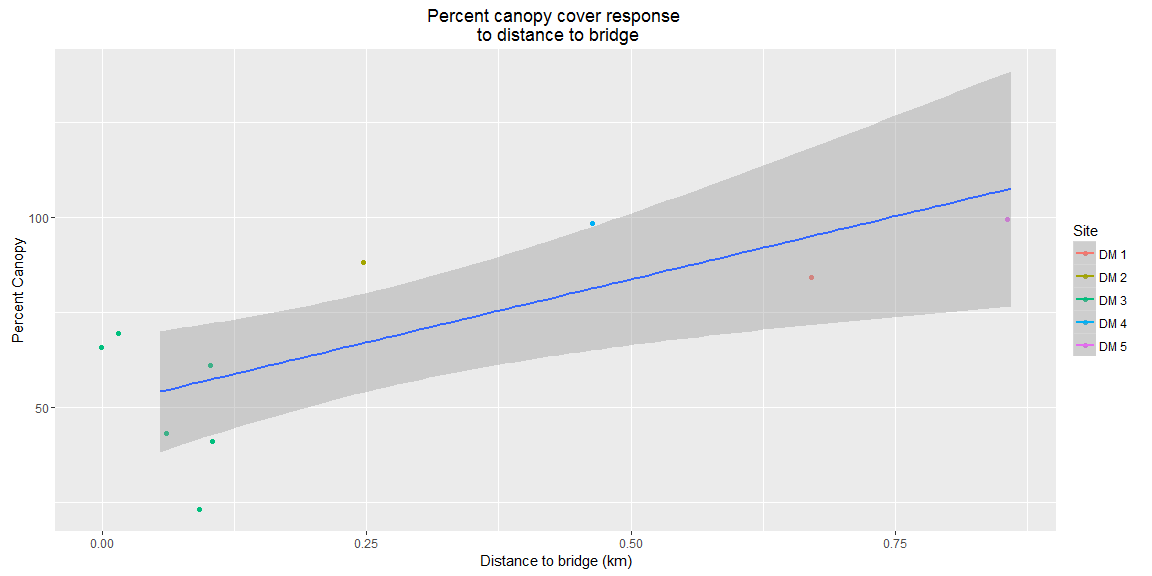
Linear model for percent 456 explained by distance to bridge (using site average, not transect):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| **distance\_from\_bridge\_km** | -35 | 13 | -2.8 | 0.024 |
| **(Intercept)** | 24 | 4.8 | 5 | 0.001 |

Fitting linear model: perc\_456 ~ distance\_from\_bridge\_km

|  |  |  |  |
| --- | --- | --- | --- |
| Observations | Residual Std. Error |  | Adjusted |
| 10 | 11 | 0.49 | 0.43 |



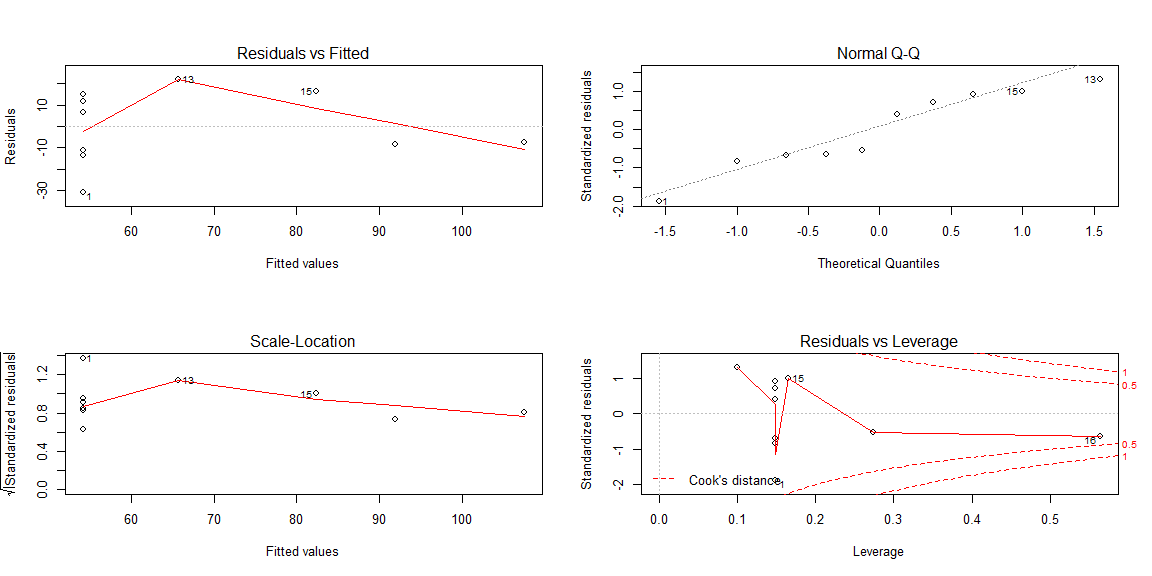


Linear model for percent canopy explained by distance to bridge

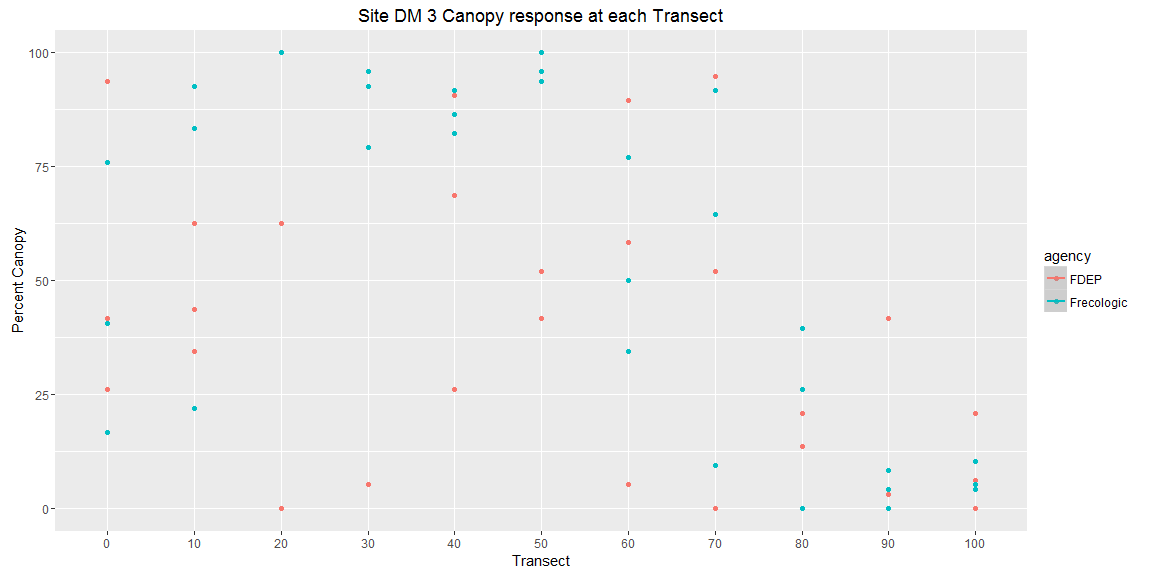
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| **distance\_from\_bridge\_km** | 66 | 20 | 3.3 | 0.011 |
| **(Intercept)** | 51 | 7.6 | 6.7 | 0.00016 |

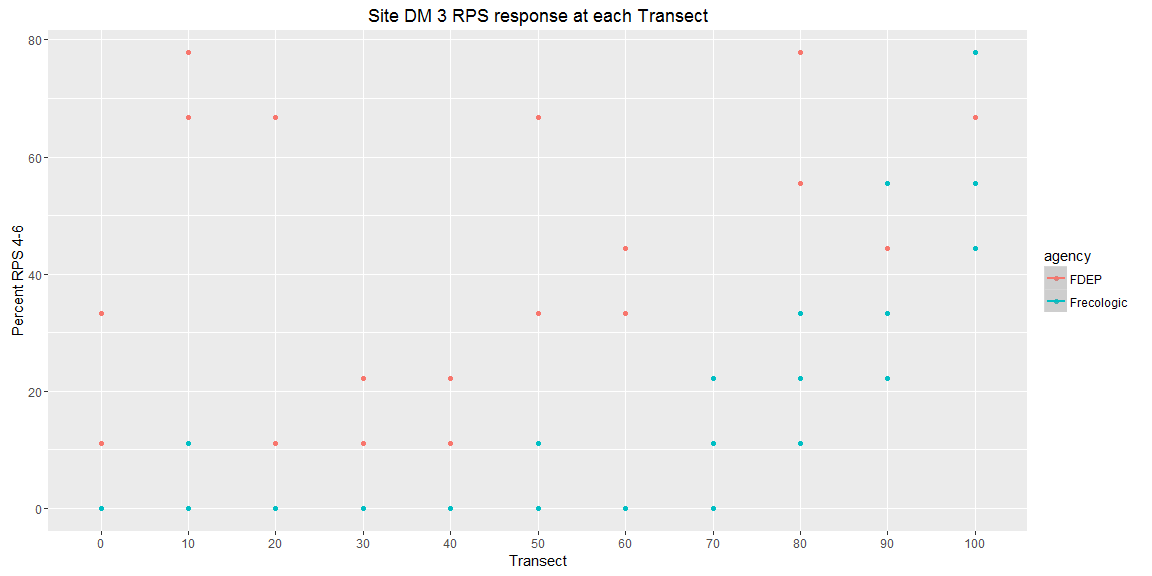
Fitting linear model: perc\_can ~ distance\_from\_bridge\_km

|  |  |  |  |
| --- | --- | --- | --- |
| Observations | Residual Std. Error |  | Adjusted |
| 10 | 18 | 0.58 | 0.52 |



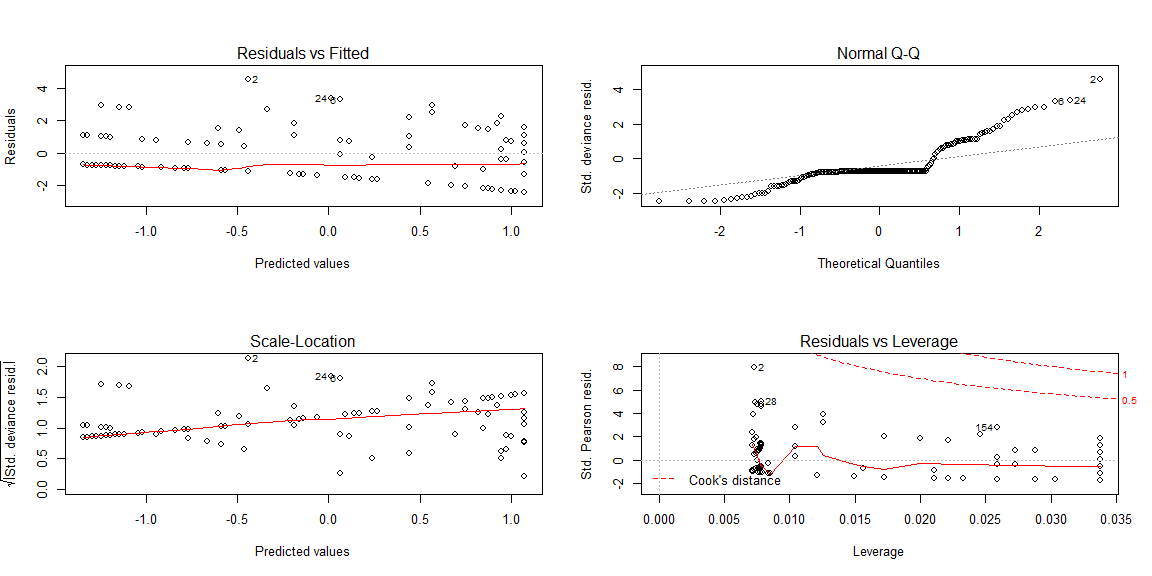
Site DM 3 only. Look at percent canopy cover at each transect. Map RPS perc at each transect too.





# generalized lienar model approach

##   
## Call:  
## glm(formula = sum\_rps ~ Canopy, family = poisson, data = rps\_df\_test2)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.4190 -0.7880 -0.7214 -0.0388 4.5470   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 1.073579 0.107474 9.989 <2e-16 \*\*\*  
## Canopy -0.025205 0.002331 -10.811 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for poisson family taken to be 1)  
##   
## Null deviance: 434.61 on 175 degrees of freedom  
## Residual deviance: 297.51 on 174 degrees of freedom  
## AIC: 447.15  
##   
## Number of Fisher Scoring iterations: 6



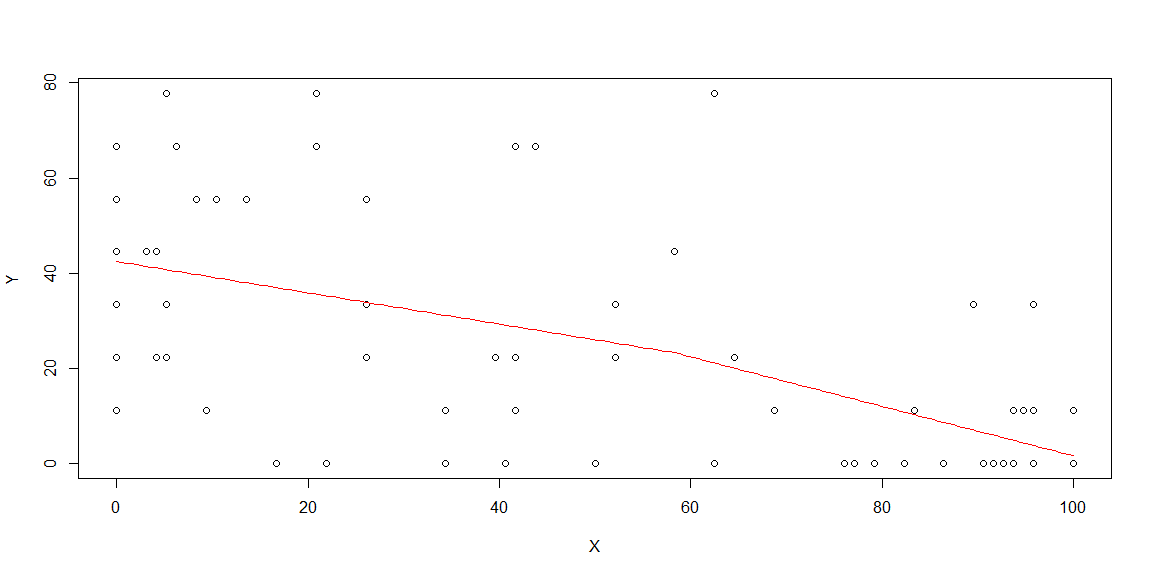
### Change point analysis

Change point analysis for RPS data, assuming change point based by visual inspection.

Estimating the change point using SiZeR package.

## [1] "Threshold alpha: 58.3333682080848"  
## [1] ""  
## [1] "Model coefficients: Beta[0], Beta[1], Beta[2]"  
## (Intercept) x w   
## 42.4249884 -0.3288375 -0.1916187   
## Change.Point Initial.Slope Slope.Change Second.Slope  
## 2.5% 6.242878 -1.289134 -4.055058 -1.0406801  
## 97.5% 86.460444 3.588162 1.266636 0.4238199

## [1] "Threshold alpha: 58.3333682080848"  
## [1] ""  
## [1] "Model coefficients: Beta[0], Beta[1], Beta[2]"  
## (Intercept) x w   
## 42.4249884 -0.3288375 -0.1916187   
## Change.Point Initial.Slope Slope.Change Second.Slope  
## 2.5% 5.952882 -1.214656 -3.883543 -0.9869641  
## 97.5% 85.855504 3.339866 1.140272 0.3289907



### rps vs tn

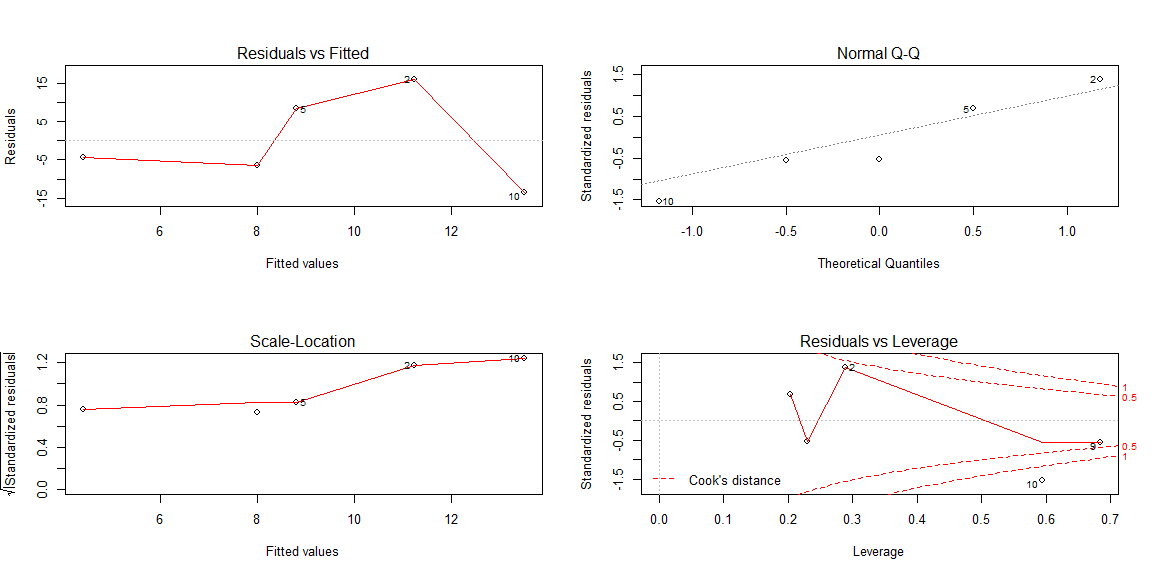
## 

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| **TN** | -2.4 | 4.7 | -0.5 | 0.65 |

## **(Intercept)** 15 13 1.1 0.33

Fitting linear model: perc\_456 ~ TN

|  |  |  |  |
| --- | --- | --- | --- |
| Observations | Residual Std. Error |  | Adjusted |
| 5 | 14 | 0.077 | -0.23 |



### multiple linear regression

Not doing this for now