

Final Project - EXST 7142

Dina Dinh

2023-11-22

Contents

Introduction	2
Data Description	2
Packages Needed for the Analysis	4
Importing and Cleaning the Data	4
Splitting the Dataset into 20 Subsets Based on GridId	4
Analyzing the Data	4
Summary Statistics of the Dataset	4
Visualizing the Correlation	5
Visualizing Outliers	6
Distribution of the Response Variable	7
Scatter Plots of the Numeric Variables vs. Response	8
Comparing the Different Grids	24
Summary Statistics of each Grid	24
Comparing Grid 21 to Grid 52	24
Relative Elevation vs. Yield	24
Slope vs. Yield	25
Terrain Ruggedness vs. Yield	26
Topographic Position vs. Yield	27
Elevation vs. Yield	28
Application 7 Nitrogen Rate vs. Yield	29
Application 10 Nitrogen Rate vs. Yield	30
pH Mean vs. Yield	31
Clay Mean vs. Yield	32

Silt Mean vs. Yield	33
Sand Mean vs. Yield	34
Ksat Mean vs. Yield	35
OM Mean vs. Yield	36
Conclusion for Comparing Grid 21 and Grid 52	37
Prediction Model	37
Random Forest (RF)	37
Removing Outliers in Response Variable, Yield	42
Optimized RF on the Reduced Outlier Dataset	44
Results of Optimized RF on Reduced Outlier Data	45
Optimized RF on Columns 25-45	47
Results of Un-Optimized RF	47
Results of Optimized RF	48
Conclusion	48
Appendix	49

Introduction

In the fast-evolving landscape of agriculture, a technological transformation is underway, propelled by the integration of Machine Learning (ML) and Artificial Intelligence (AI). Termed Digital Agriculture, this innovative approach leverages data analytics, automation, and intelligent decision-making to revolutionize traditional farming practices. By assimilating vast datasets from sensors, satellites, and IoT devices, ML and AI empower farmers with unprecedented insights, enabling precise management of resources, predictive analytics for crop health, and the implementation of smart, efficient farming practices. The objective of this analysis is to examine the correlation between the explanatory variables and the response and assessing whether precise predictions can be made using these explanatory factors. This will give a better understanding of how these factors affect crop yield.

Data Description

In this section, we will describe the variables used in the soybeans dataset provided by Ag Analytics. The variables, “fid” and “field_1,” have been excluded as they serve as non-informative identifiers in this analysis. The variables “application_4_N_rate” and “application_5_N_rate” were also removed due to no values were collected in this dataset for those variables.

1. GridId: Identification numbers given by the program QGIS of the 20 different grids created on the given field (see Figure 1).
2. x: data point’s longitude coordinate using the WGS-84 coordinate reference system (CRS) with precision of 1.11×10^{-6} m (± 4 cm).

3. y: data point's latitude coordinate using the WGS-84 coordinate reference system (CRS) with precision of 1.11×10^{-6} m (± 4 cm).
4. VRYieldVOI: the response variable, which stores volumetric yield. The units are user-defined and are usually either kg/ha, lbs/ac, or bushels/ac.
5. Row & Column: row and column of the data point's location within Ag Analytic's 10ft x 10ft data point sample grid for a given field (see Figure 2). Row and col identifiers are 0-indexed (i.e. column numbering starts with 0 instead of 1). The Values are calculated from raw data using proprietary algorithms. Data are of the integer data type and should be treated as being categorical in most cases.
6. Relative_Elevation1: decimal value representing the standardized elevation value (z-score) of a given data point relative to the mean elevation of the field. Relative elevation can affect water and nutrient status of a given area due to how water and nutrients flow to and from a given area due to elevation differences. Values are positive or negative numeric with 9 decimals of precision.
7. Slope1: holds the maximum slope value present in the 10ft x 10ft cell represented by a given data point. Slope can affect water and nutrient status of a given point due to how water and nutrients flow over that area due to the degree of slope.
8. TRI1: terrain ruggedness index value. The amount of hilliness and slope amount present within a given cell. Terrain ruggedness can affect water and nutrients flow.
9. TPI1: Topographic position index which is the difference of target cell to the average cells around it.
10. Elevation1: absolute elevation value of a given point. Generally, meters above sea level (ASL).
11. Application_<#>_<N_Rate>: these fields hold the application rates of nitrogen (N) for each record. The number (#) is equivalent to the crop's growth stage. Nitrogen application rate, in conjunction with the percentage of N applied is very strongly correlated to plant growth and yield in most grain crops. Units are user-defined and are generally gallons/ac, lbs/ac, or kg/ha.
12. ph_mean_30_60: mean soil pH value present between 30cm and 60 cm below the surface for a given 10ft x 10ft cell. pH is the -log of hydrogen ion activity present in a sample. Values below 7 are acidic and values above 7 are alkaline. Soil pH is generally between 4 and 10, except in extreme cases. Soil pH affects nutrient availability to plants. Optimum ranges are between 5.5 and 6.5 for most agronomic crops. A magnet is used for variable measurements that are underground.
13. Clay_mean_30_60, silt_mean_30_60, and sand_mean_30_60: denote the percentages of clay, silt, and sand, respectively, present in the soil between 30cm and 60cm for a given 10ft x 10ft cell. The relative percentages of each affect the soil's texture, nutrient holding capacity, and other chemical and physical soil properties.
14. ksat_mean_30_60: mean saturated hydraulic conductivity of soil between 30cm and 60cm for a given 10ft x 10ft cell record. This value indicates how easily water can percolate through soil once the soil is fully saturated. Higher values indicate greater flow rates, meaning that the soil allows water to flow through it more freely than areas of soil having lower ksat values. Soil's hydraulic conductivity can influence water and nutrient availability to plants by influencing how long water and nutrients are present in a given area and how quickly mobile nutrients can leach away. Soils having higher percentages of clay generally have lower ksat values where soils having higher percentages of sand generally have higher ksat values.
15. om_mean_30_60: mean organic matter (OM) percentage of the soil for a given 10m x 10m cell record between 30cm and 60cm. OM generally improves soil texture and soil nutrient and water holding capacity, and nutrient availability to plants.

Packages Needed for the Analysis

```
library(easypackages)
libraries("tidyverse", "boot", "randomForest", "psych", "AUC", "MASS", "car",
          "viridis", "caret", "ggplot2", "corrplot", "gridExtra", "mlbench", "neuralnet")
```

Importing and Cleaning the Data

Initially, I removed the variables that aren't needed for the analysis. Subsequently, I made the appropriate variables categorical. Given the minimal presence of missing values in the extensive dataset, I excluded the 20 observations with missing values. Assuming that "x" and "y" variables were employed to determine rows and columns, and considering that the levels of the respective pairs were consistent, both "x" and "y" were omitted from the analysis.

```
df1 = read.csv("C:/Users/ddinh4/Documents/EXST 7142/EXST 7142 Final Project/Data/YY_varying_withGrid_id.csv")
df1 <- df1[, -which(names(df1) %in% c("fid", "field_1", "X", "x", "y", "Application_4_N_rate", "Application_10_N_rate"))]
df1 <- df1 %>%
  drop_na() %>%
  mutate(GridId = as.factor(GridId)) %>%
  mutate(row = as.numeric(row)) %>%
  mutate(col = as.numeric(col))
```

Splitting the Dataset into 20 Subsets Based on GridId

I decided to break down the data by GridId because I believe a crop's location in the field could play a role in its yield. The idea is that different parts of the field might have varying effects on crop growth. By splitting the data this way, I hope to uncover any patterns or influences tied to specific locations. This approach allows for a more detailed analysis compared to looking at the entire dataset in one go.

The grid layout can be seen in Figure 1.

```
gridid_subsets <- split(df1, df1$GridId)
```

Analyzing the Data

Summary Statistics of the Dataset

Taking a look at the data summary, it's apparent that certain variables exhibit significant skewness, such as VRYieldVol, Slope1, and TRI1. Additionally, there are variables with notably wide ranges, as observed in the case of Application_10_N_rate. These findings suggest that specific variables could have distinct distributions or extreme values that could influence the overall analysis.

Table 1: Summary Table

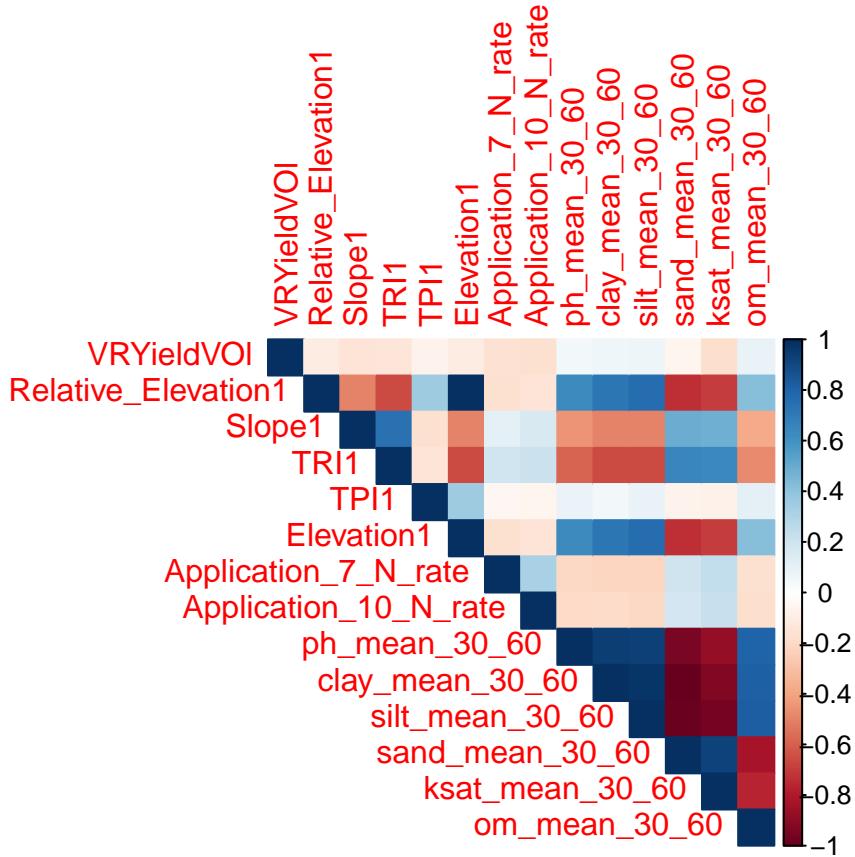
	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se	
GridId*		1	3315	10.84	5.53	11.00	10.86	7.41	1.00	20.00	19.00	-	-1.16	0.10

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
VRYieldVOL	2	3315	78.86	24.32	81.67	81.47	18.00	0.00	127.21	127.21	-	2.65	0.42
row	3	3315	72.22	8.36	72.00	72.26	10.38	58.00	86.00	28.00	-	-1.20	0.15
col	4	3315	62.30	33.07	62.00	62.35	43.00	2.00	119.00	117.00	-	-1.19	0.57
Relative_Elevation15	5	3315	0.01	0.98	0.48	0.16	0.62	-	1.47	5.27	-	1.13	0.02
Slope1	6	3315	0.01	0.01	0.01	0.01	0.00	0.00	0.05	0.05	2.68	9.82	0.00
TRI1	7	3315	0.35	0.27	0.26	0.29	0.13	0.09	1.62	1.54	2.08	4.30	0.00
TPI1	8	3315	0.00	0.83	0.07	0.02	0.77	-	3.65	8.01	-	0.86	0.01
Elevation1	9	3315	54.90	1.36	55.55	55.11	0.86	49.60	56.93	7.33	-	1.13	0.02
Application_7_N_10ate	10	3315	0.24	0.07	0.24	0.24	0.00	0.15	2.26	2.10	20.36	511.74	0.00
Application_10_N1lrate	11	3315	81.03	13.08	79.87	79.84	2.10	38.91	332.41	293.50	12.27	181.53	0.23
ph_mean_30_60	12	3315	7.60	0.05	7.61	7.61	0.04	7.39	7.68	0.29	-	5.30	0.00
clay_mean_30_60	13	3315	43.72	1.62	44.22	43.98	0.68	37.45	46.30	8.86	-	1.86	0.03
silt_mean_30_60	14	3315	25.16	0.91	25.46	25.33	0.34	21.34	26.21	4.87	-	6.52	0.02
sand_mean_30_60	15	3315	27.73	2.31	27.14	27.35	1.17	24.33	36.52	12.19	2.04	4.69	0.04
ksat_mean_30_60	16	3315	0.50	0.44	0.35	0.40	0.09	0.25	2.68	2.43	3.57	13.03	0.01
om_mean_30_60	17	3315	1.21	0.03	1.21	1.21	0.02	1.10	1.26	0.17	-	3.25	0.00
											1.30		

Visualizing the Correlation

The examination reveals numerous variables with significant correlations, both positive and negative. Notably, clay and sand means exhibit a strong negative correlation, while clay and silt display a pronounced positive correlation.

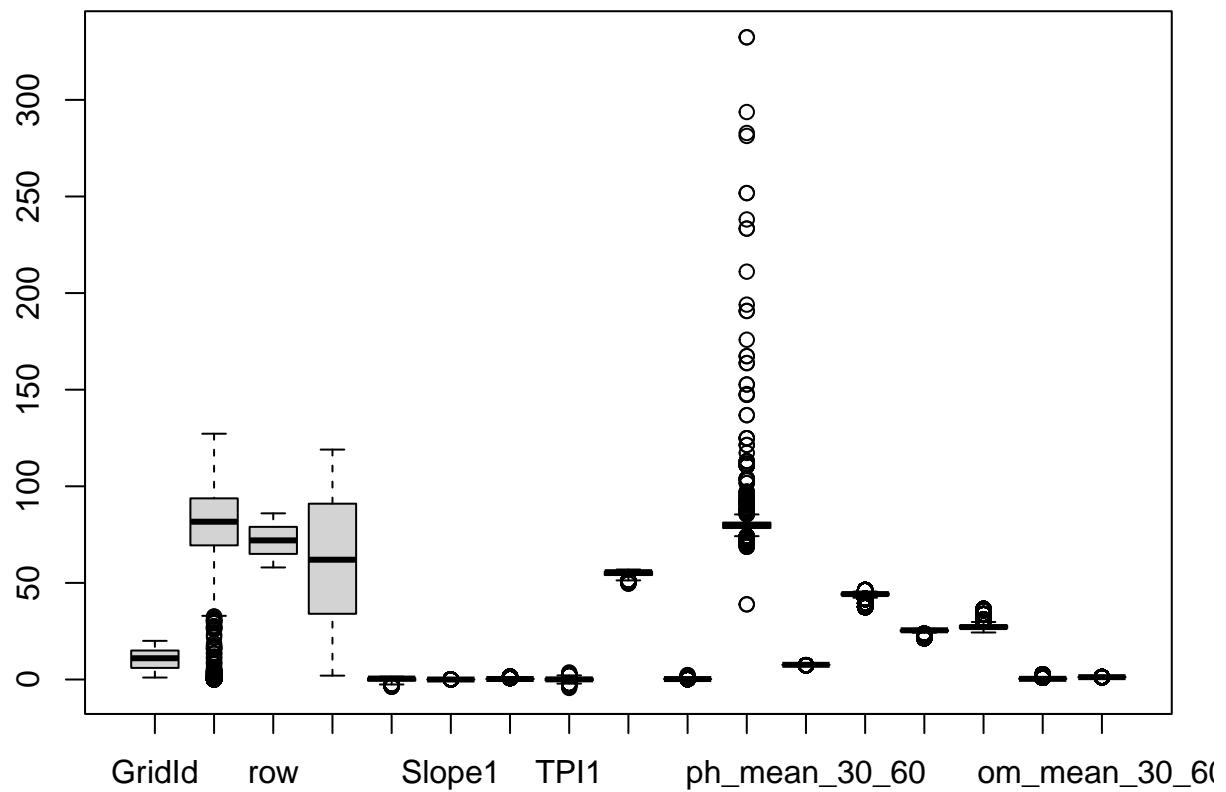
```
corr.matrix = cor(df1[,-c(1,3,4)])
corrplot(corr.matrix, method = "color", type = "upper")
```



Visualizing Outliers

The boxplots for each variable reveal that VRYieldVol and Application_10_N_Rate contains numerous outliers. This corresponds to the high ranges seen in the summary of the data.

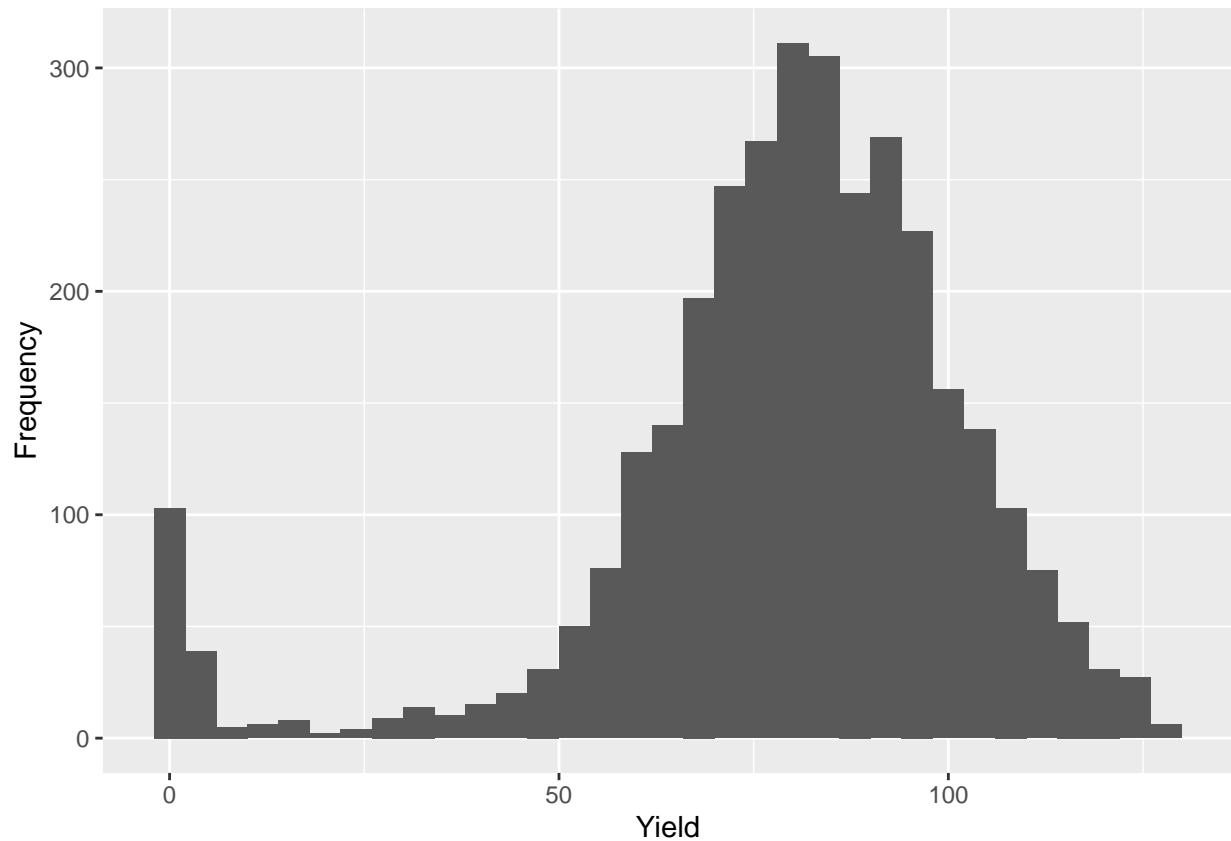
```
par(mfrow = c(1, 1), mar = c(3, 3, 1, 1) + 0.1)
boxplot(df1)
```



Distribution of the Response Variable

The measurements resemble a normal distribution; however, the presence of numerous zeroes and low values introduces a skewness that departs from the expected pattern.

```
ggplot(df1, aes(x = VRYieldV01)) +
  geom_histogram(binwidth = 4, aes(y = ..count..)) +
  labs(x = "Yield", y = "Frequency")
```



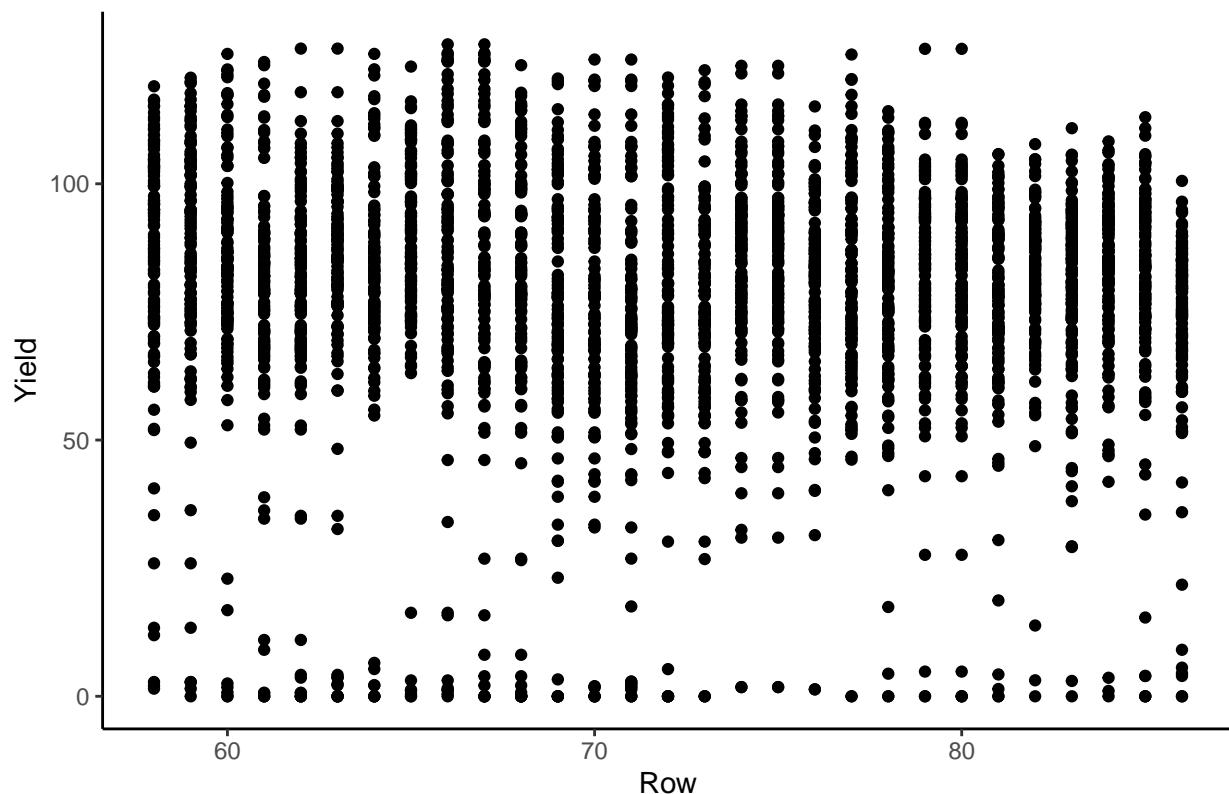
Scatter Plots of the Numeric Variables vs. Response

Row vs. Yield

In the plot, there's an observable undulating pattern.

```
ggplot(df1, aes(x = row, y = VRYieldV01)) +  
  geom_point() +  
  labs(title = "Scatter Plot of Row and Yield",  
       x = "Row",  
       y = "Yield") +  
  theme_classic()
```

Scatter Plot of Row and Yield

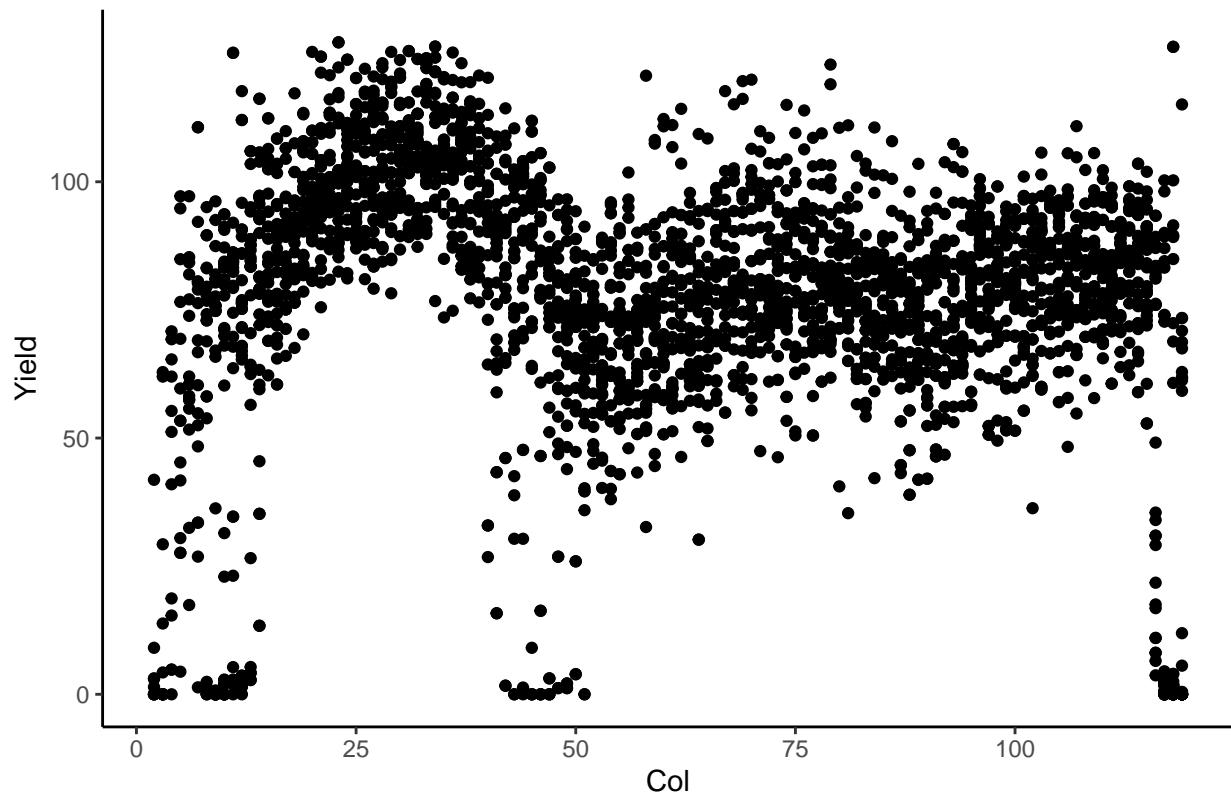


Column vs. Yield

A distinct undulating pattern is noticeable, particularly between the column and yield. It appears that plants between column 25 and 45 consistently had the highest yield.

```
ggplot(df1, aes(x = col, y = VRYieldV01)) +  
  geom_point() +  
  labs(title = "Scatter Plot of Column and Yield",  
       x = "Col",  
       y = "Yield") +  
  theme_classic()
```

Scatter Plot of Column and Yield

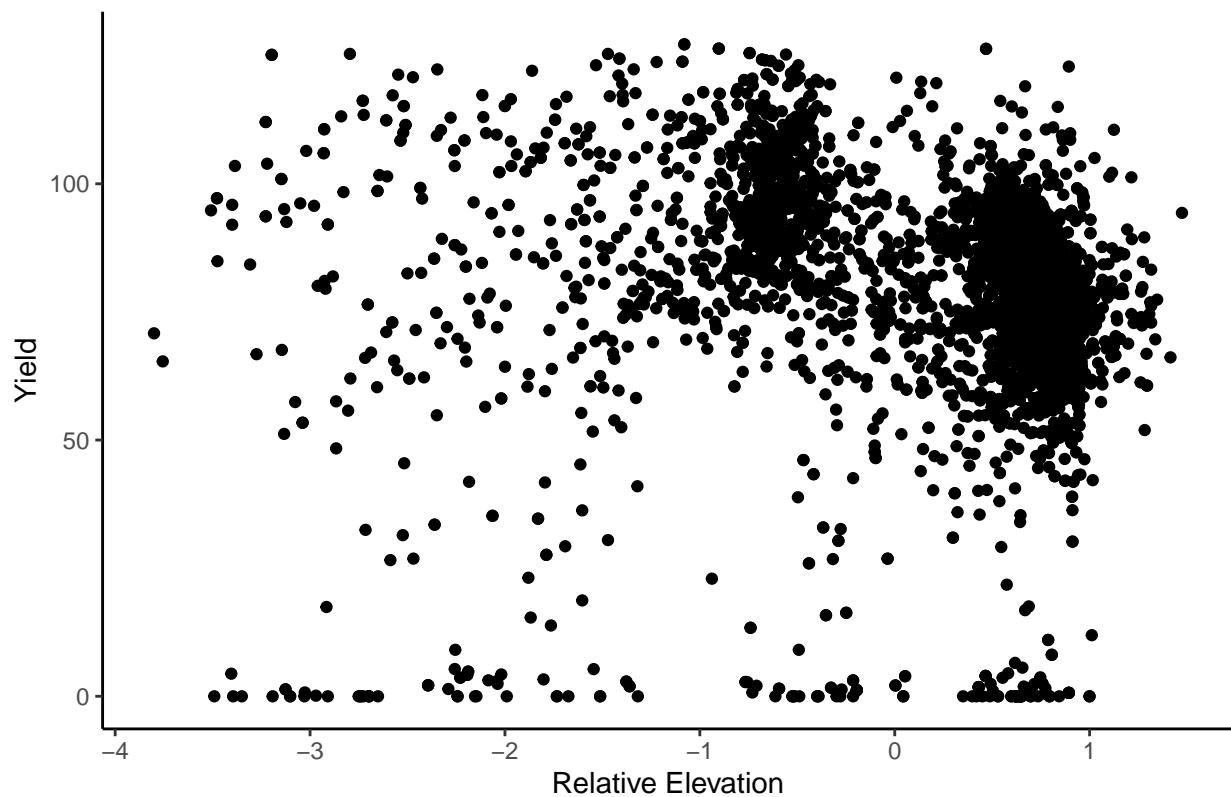


Relative Elevation vs. Yield

According to the scatter plot, there's a notable trend where areas with lower elevation correspond to higher yields. This finding is somewhat unexpected, as lower-lying areas are typically prone to water pooling, which could impede crop growth. One possible explanation for this counter intuitive observation is the high water requirement for optimal soybean growth, suggesting that the soil in lower elevations may be conducive to providing the necessary moisture for robust crop development.

```
ggplot(df1, aes(x = Relative_Elevation1, y = VRYieldV01)) +  
  geom_point() +  
  labs(title = "Scatter Plot of Relative Elevation and Yield",  
       x = "Relative Elevation",  
       y = "Yield") +  
  theme_classic()
```

Scatter Plot of Relative Elevation and Yield

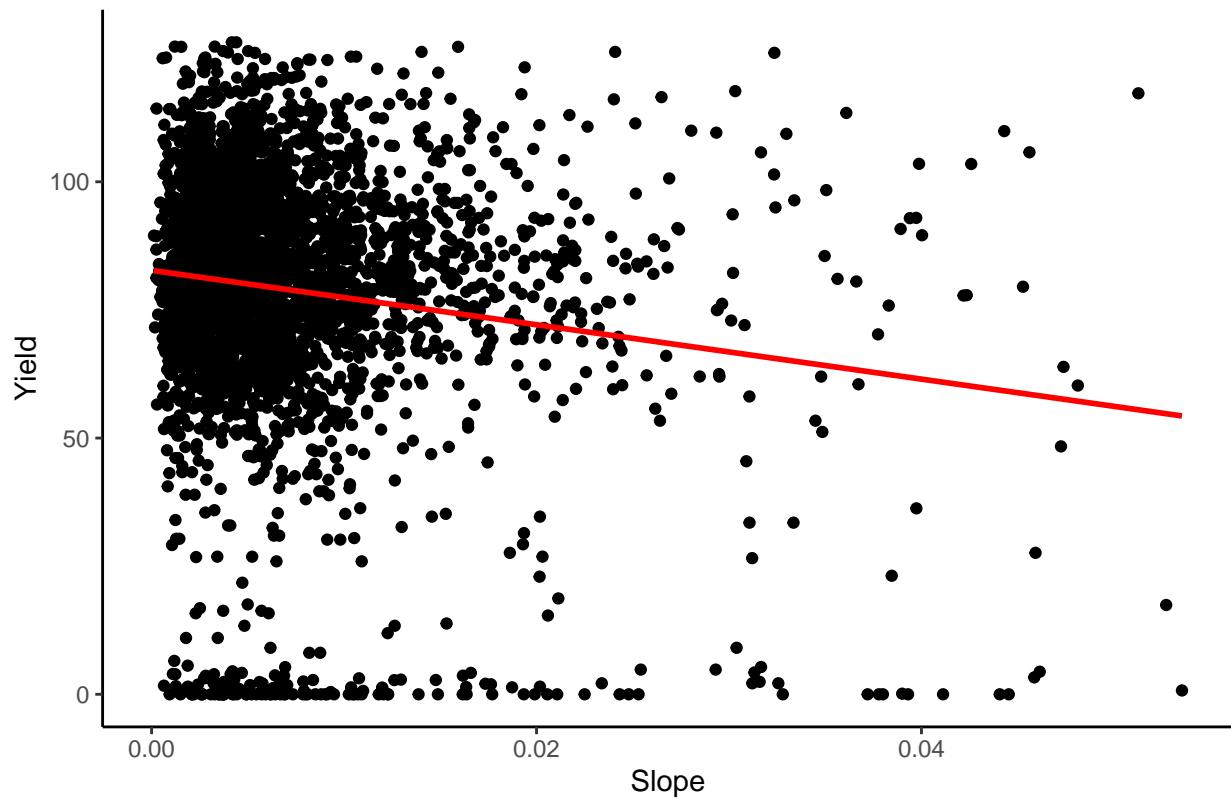


Slope vs. Yield

Determining the relationship is challenging due to a substantial imbalance in the data, with a significant concentration on one side. This skewed distribution makes it difficult to draw clear conclusions.

```
ggplot(df1, aes(x = Slope1, y = VRYieldV01)) +  
  geom_point() +  
  geom_smooth(method = "lm", se = FALSE, color = "red") +  
  labs(title = "Scatter Plot of Slope and Yield",  
       x = "Slope",  
       y = "Yield") +  
  theme_classic()
```

Scatter Plot of Slope and Yield

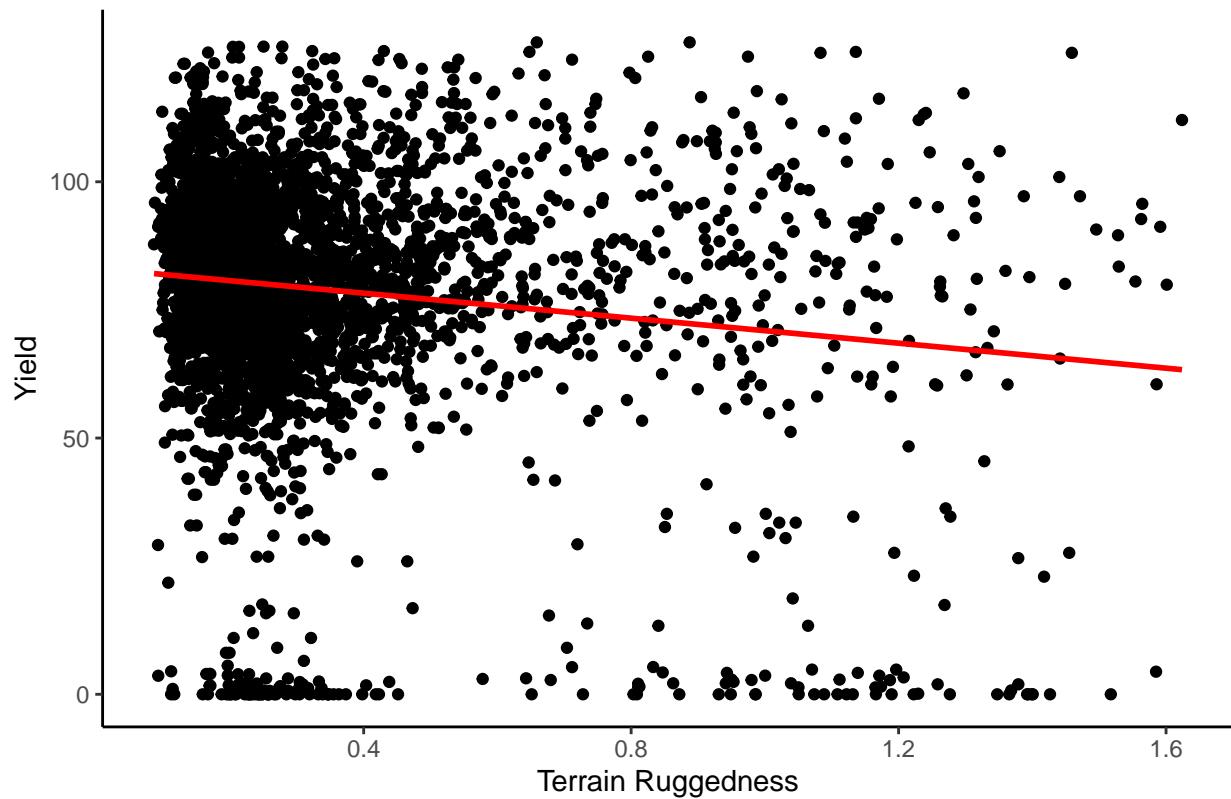


Terrain Ruggedness vs Yield

Again, due to the heavy imbalance, the relationship is difficult to determine.

```
ggplot(df1, aes(x = TRI1, y = VRYieldV01)) +  
  geom_point() +  
  geom_smooth(method = "lm", se = FALSE, color = "red") +  
  labs(title = "Scatter Plot of Terrain Ruggedness and Yield",  
       x = "Terrain Ruggedness",  
       y = "Yield") +  
  theme_classic()
```

Scatter Plot of Terrain Ruggedness and Yield

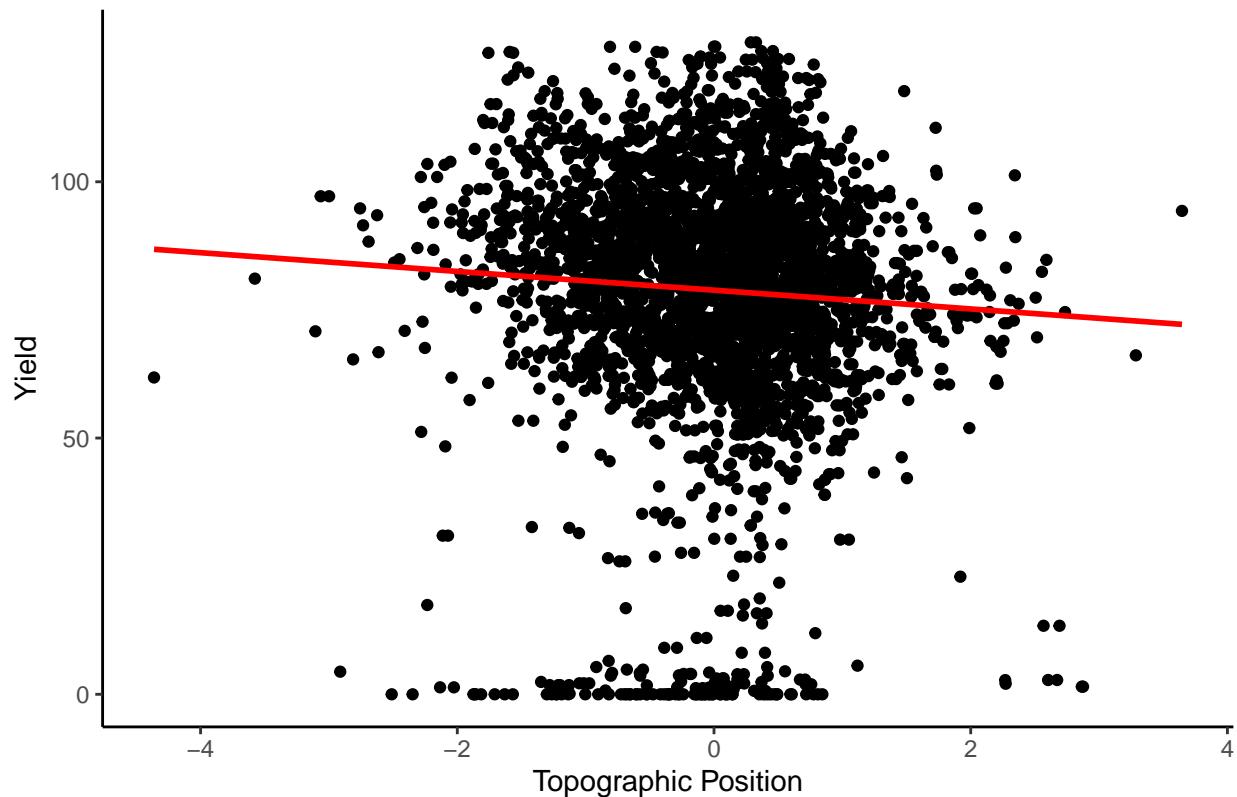


Topographic Position vs. Yield

The majority of data points are concentrated around zero.

```
ggplot(df1, aes(x = TPI1, y = VRYieldV01)) +  
  geom_point() +  
  geom_smooth(method = "lm", se = FALSE, color = "red") +  
  labs(title = "Scatter Plot of Topographic Position and Yield",  
       x = "Topographic Position",  
       y = "Yield") +  
  theme_classic()
```

Scatter Plot of Topographic Position and Yield

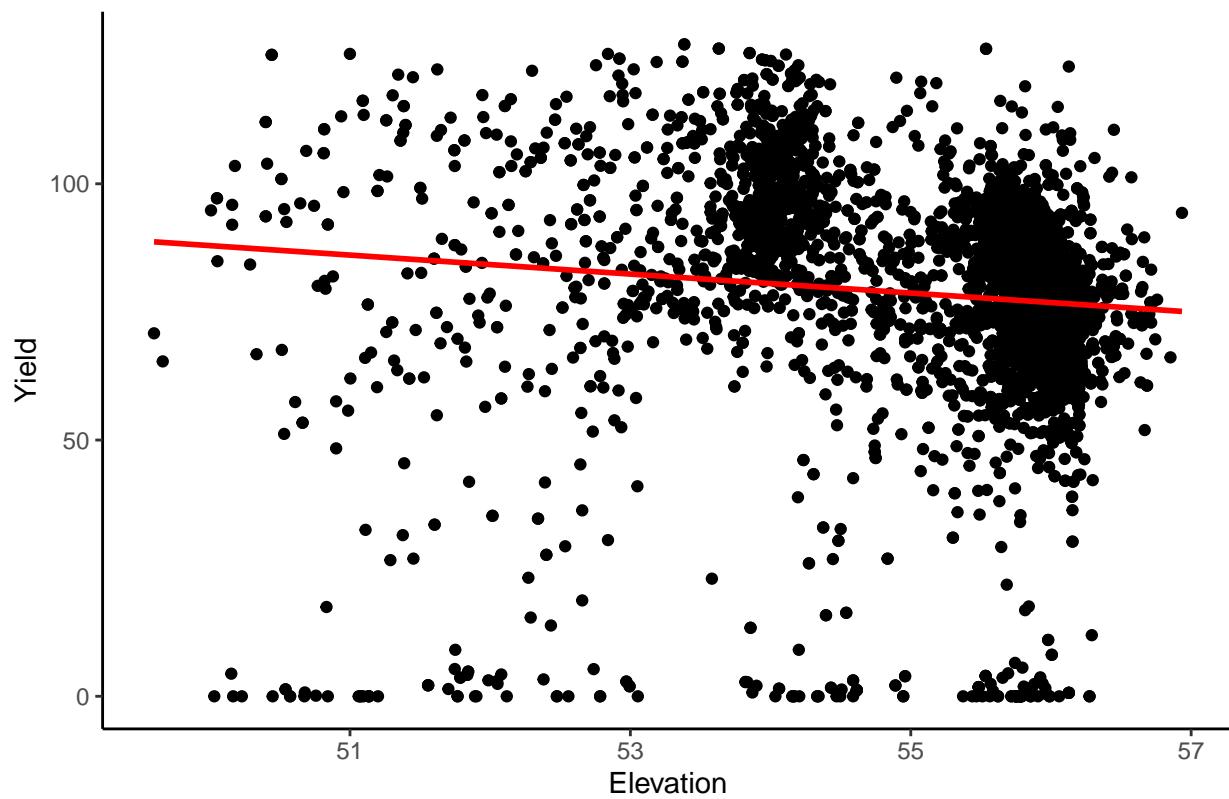


Elevation vs. Yield

The data suggests that the elevation range between 53 meters and 55 meters above sea level correlates with the highest yield. The scatter plot looks almost exactly like relative elevation vs. yield. This is caused by the strong positive correlation between relative elevation and elevation.

```
ggplot(df1, aes(x = Elevation1, y = VRYieldV01)) +  
  geom_point() +  
  geom_smooth(method = "lm", se = FALSE, color = "red") +  
  labs(title = "Scatter Plot of Elevation and Yield",  
       x = "Elevation",  
       y = "Yield") +  
  theme_classic()
```

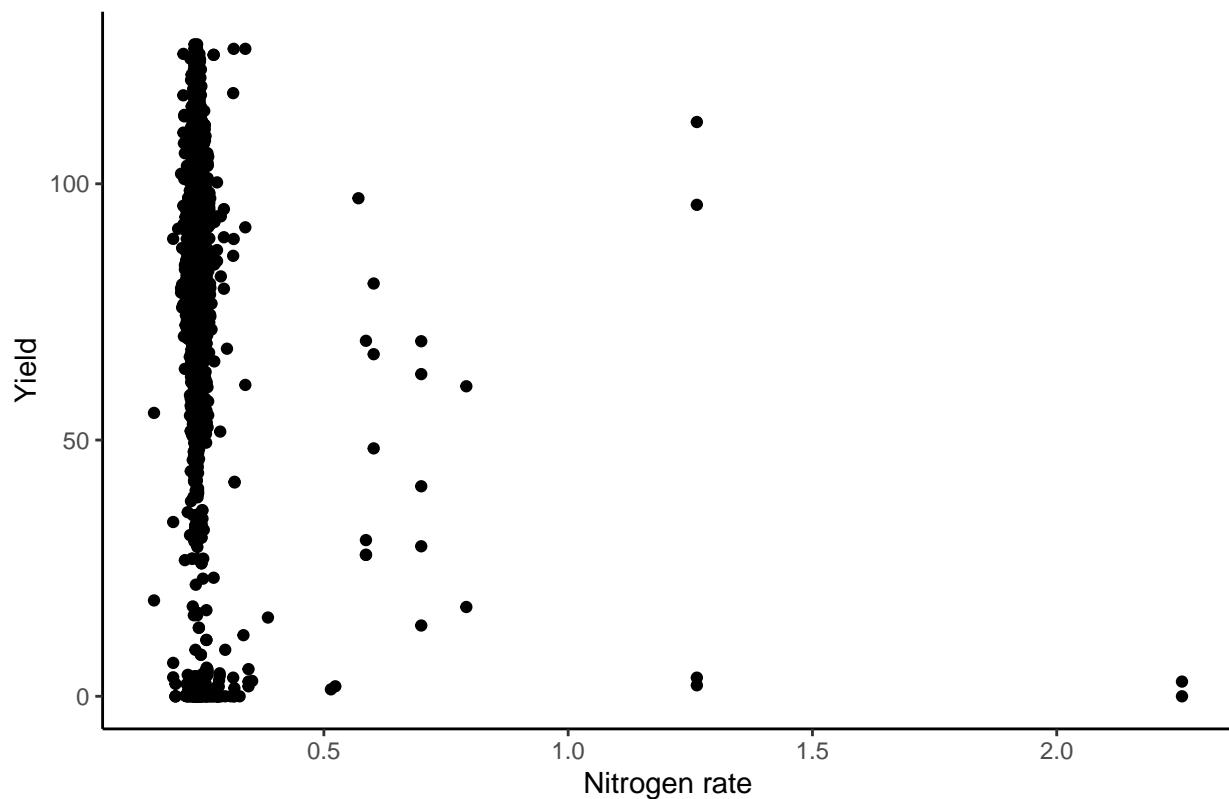
Scatter Plot of Elevation and Yield



Application 7 Nitrogen Rate vs. Yield

No clear relationship can be seen.

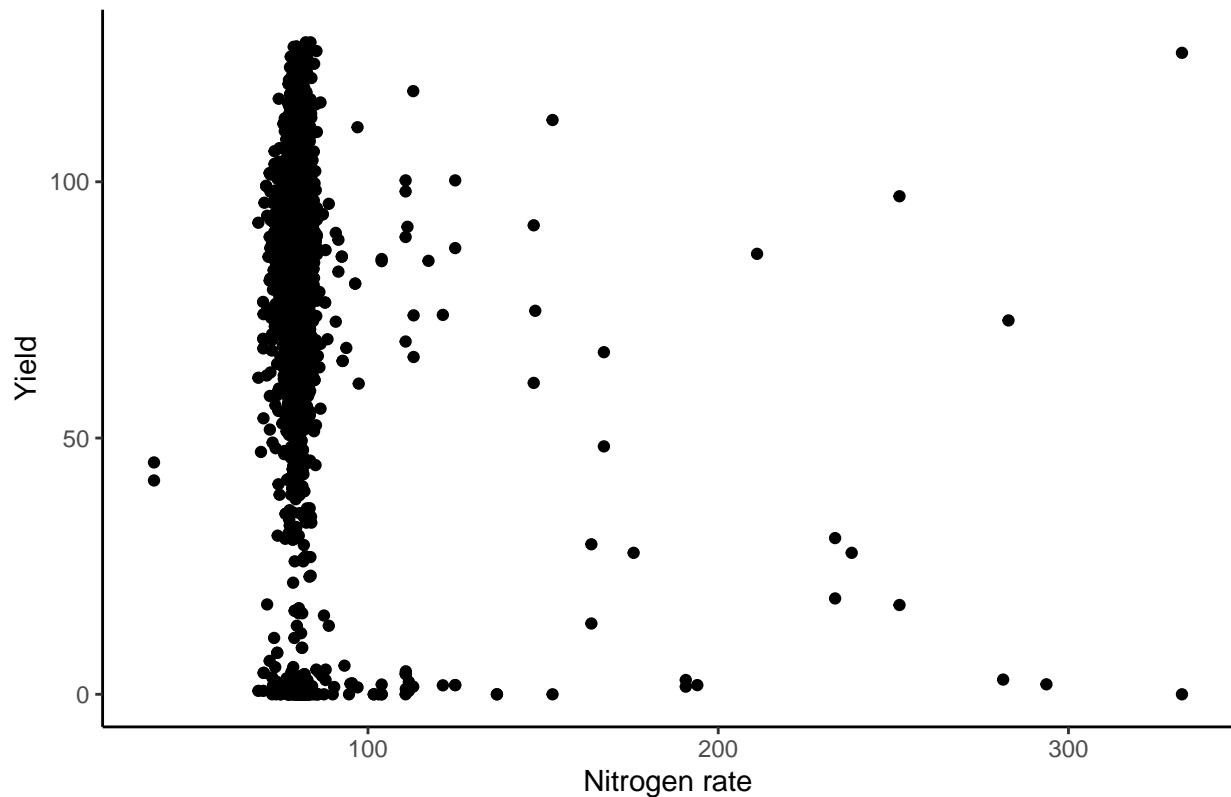
Scatter Plot of Application 7 and Yield



Application 10 Nitrogen Rate vs. Yield

Again, no clear relationship can be seen.

Scatter Plot of Application 10 and Yield

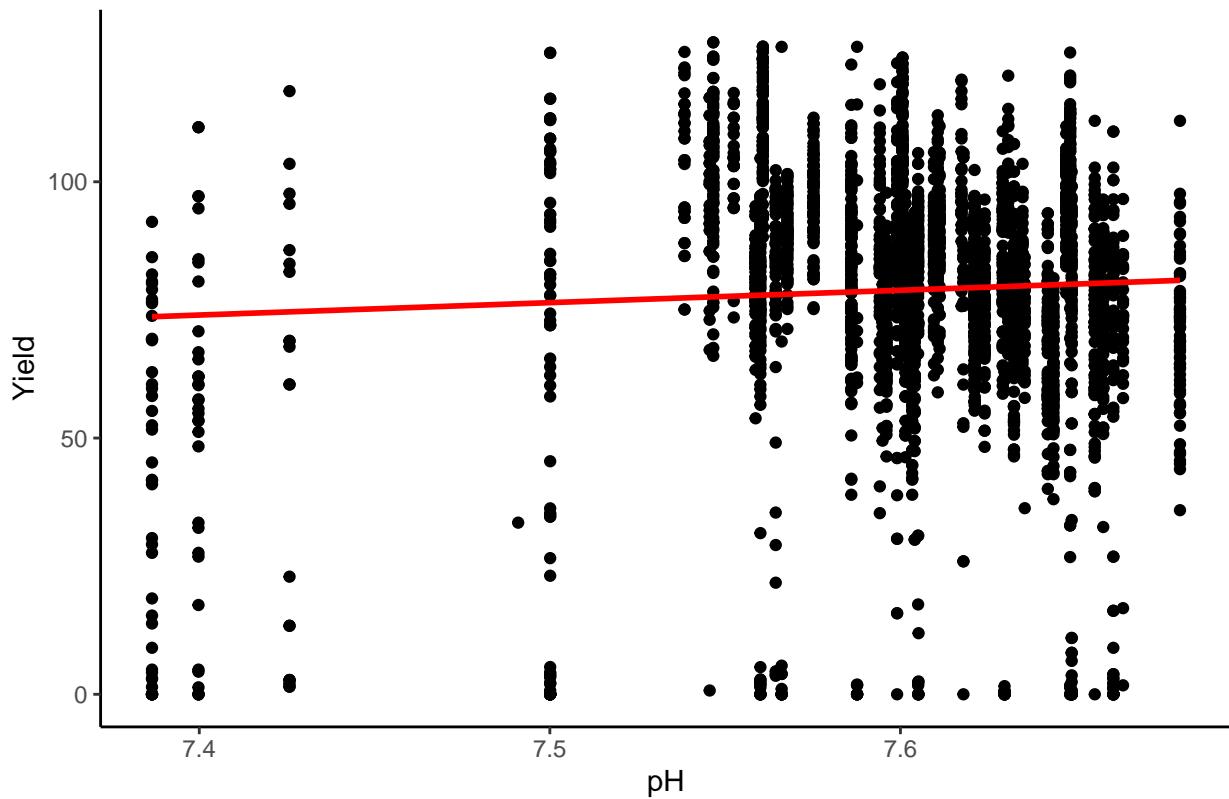


pH Mean vs. Yield

The pH level around 7.55 is associated with the highest yield, and notably, there are no instances of zero yields in this pH except one.

```
ggplot(df1, aes(x = ph_mean_30_60, y = VRYieldV01)) +  
  geom_point() +  
  geom_smooth(method = "lm", se = FALSE, color = "red") +  
  labs(title = "Scatter Plot of pH Mean and Yield",  
       x = "pH",  
       y = "Yield") +  
  theme_classic()
```

Scatter Plot of pH Mean and Yield

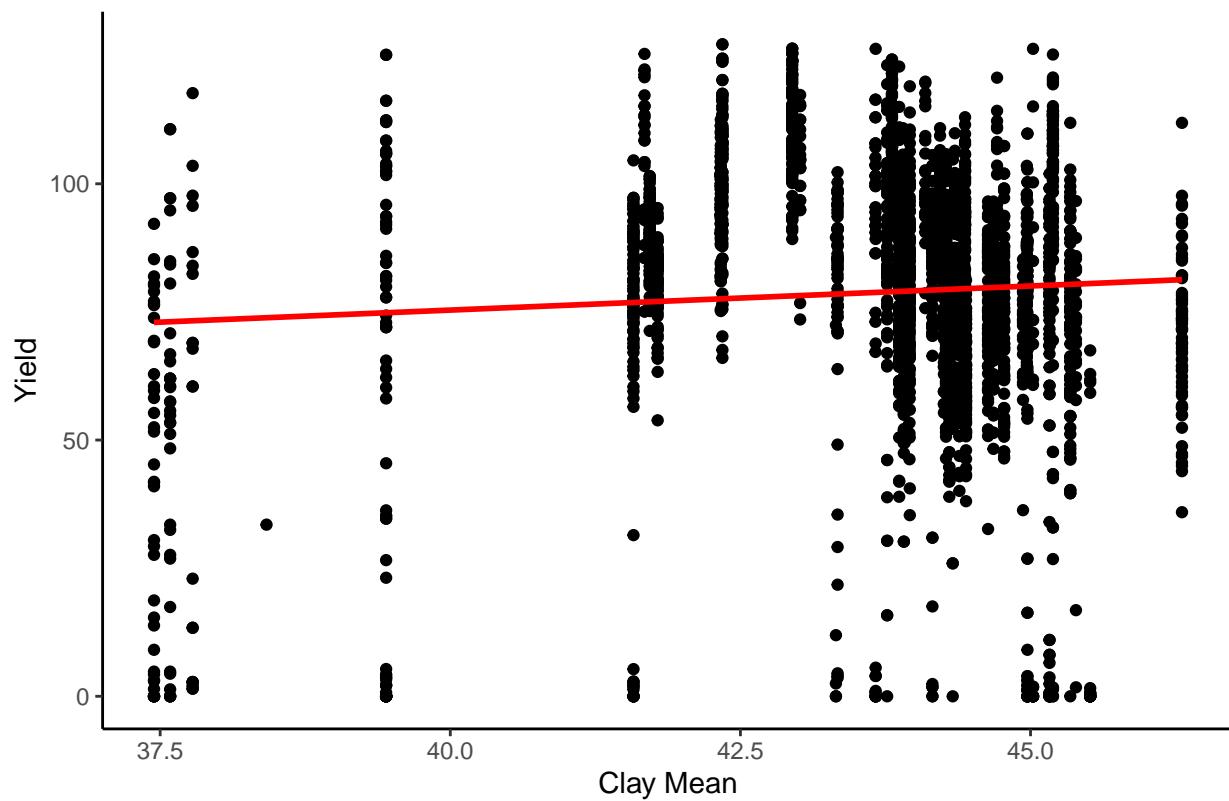


Clay Mean vs. Yield

Clay means around 42.5 produces the highest yield with no instances of zero yields.

```
ggplot(df1, aes(x = clay_mean_30_60, y = VRYieldV01)) +  
  geom_point() +  
  geom_smooth(method = "lm", se = FALSE, color = "red") +  
  labs(title = "Scatter Plot of Clay Mean and Yield",  
       x = "Clay Mean",  
       y = "Yield") +  
  theme_classic()
```

Scatter Plot of Clay Mean and Yield

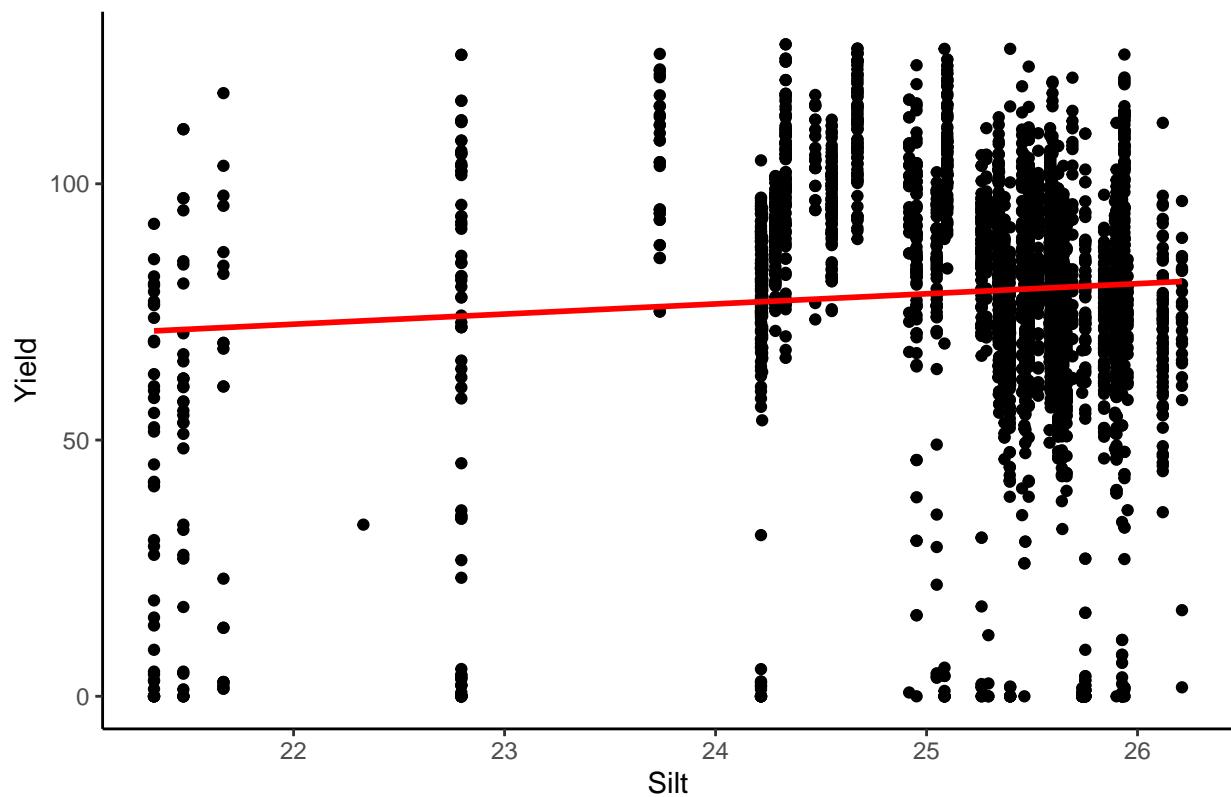


Silt Means vs. Yield

Consistently higher yields are observed in areas where the silt content falls within the range of 23.7 to 24.5.

```
ggplot(df1, aes(x = silt_mean_30_60, y = VRYieldV01)) +  
  geom_point() +  
  geom_smooth(method = "lm", se = FALSE, color = "red") +  
  labs(title = "Scatter Plot of Silt Mean and Yield",  
       x = "Silt",  
       y = "Yield") +  
  theme_classic()
```

Scatter Plot of Silt Mean and Yield

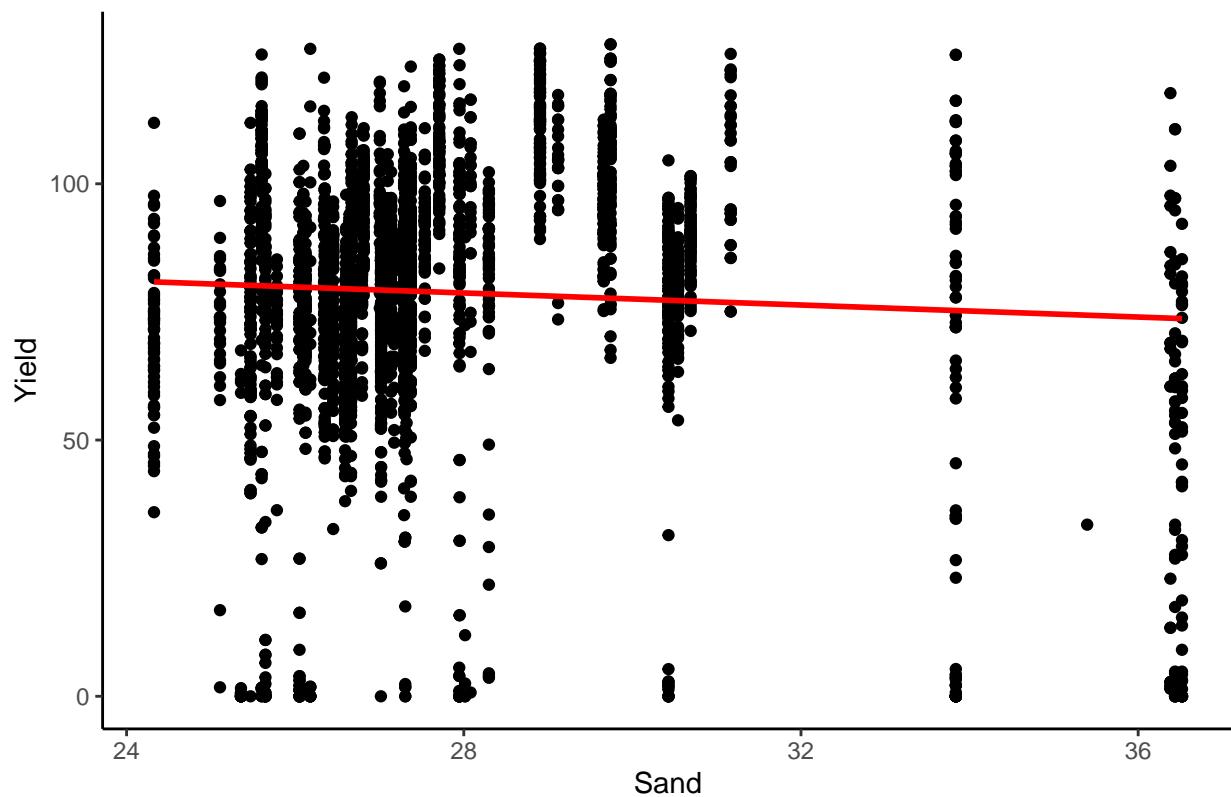


Sand Mean vs. Yield

The data reveals that the mean sand content, particularly around 29-30, is associated with the highest crop yields.

```
ggplot(df1, aes(x = sand_mean_30_60, y = VRYieldV01)) +  
  geom_point() +  
  geom_smooth(method = "lm", se = FALSE, color = "red") +  
  labs(title = "Scatter Plot of Sand Mean and Yield",  
       x = "Sand",  
       y = "Yield") +  
  theme_classic()
```

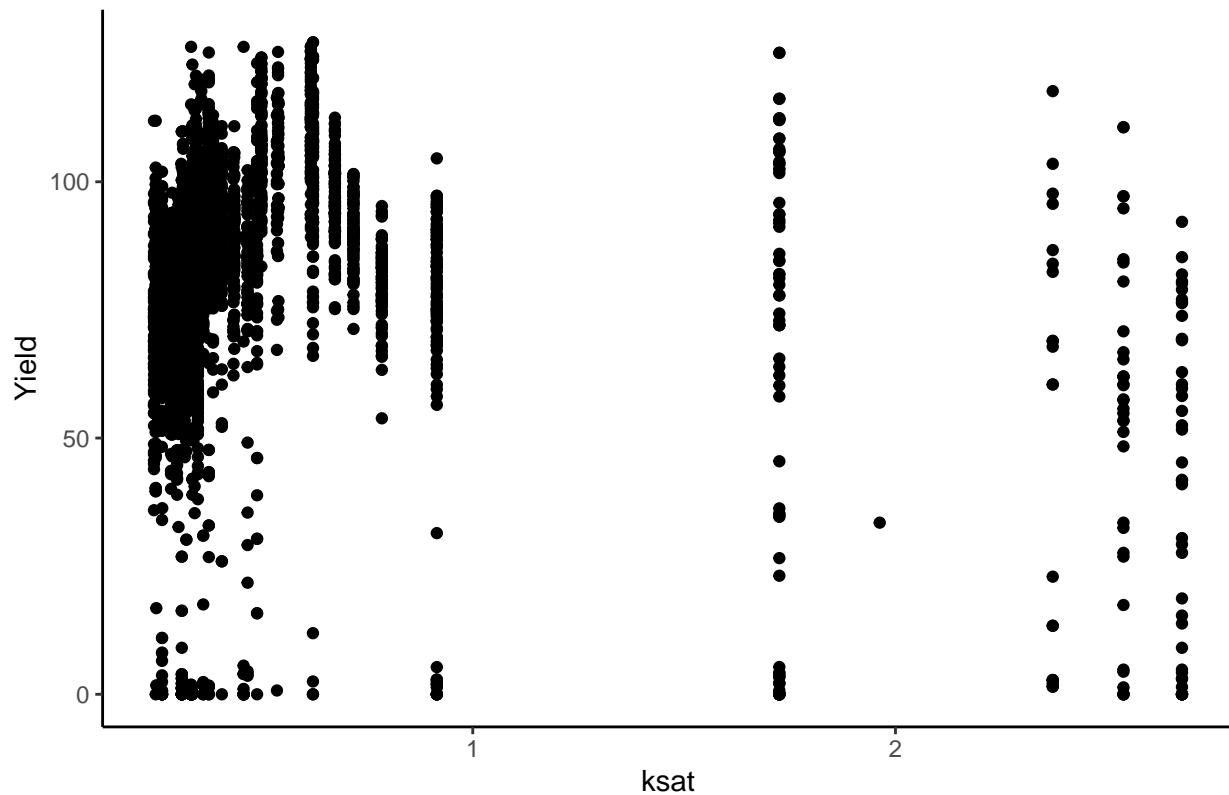
Scatter Plot of Sand Mean and Yield



Hydraulic Conductivity vs. Yield

It is difficult to determine the relationship from the scatter plot

Scatter Plot of Hydraulic Conductivity and Yield

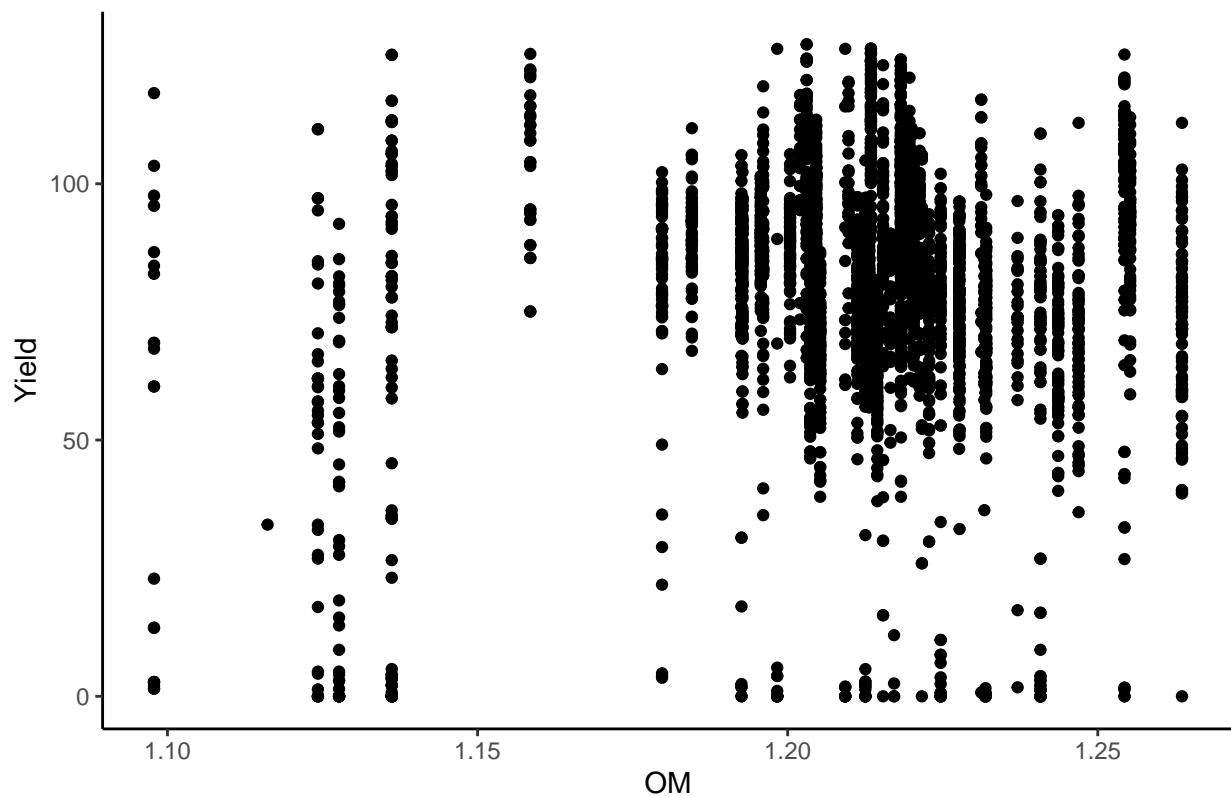


Organic Matter (OM) vs. Yield

Note that OM around 1.16 produced high yields with no zero value yields.

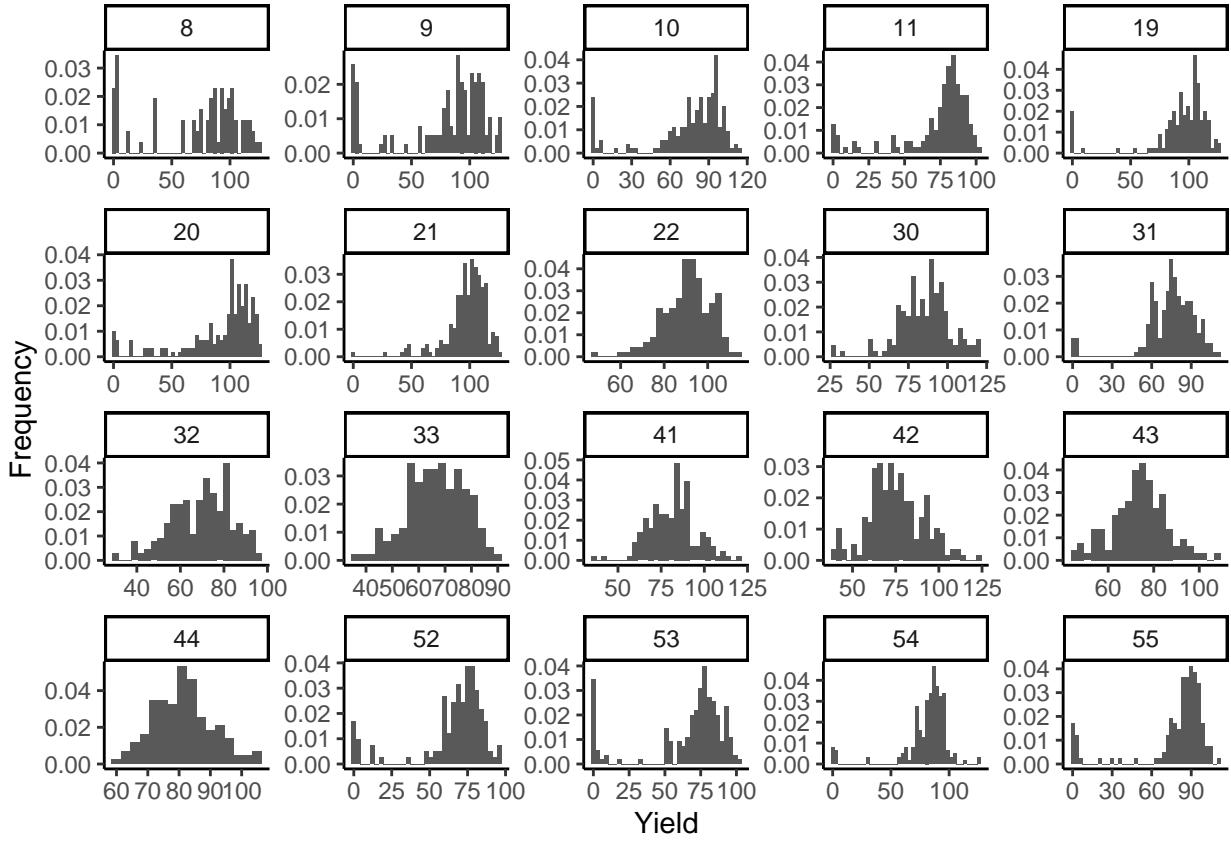
```
ggplot(df1, aes(x = om_mean_30_60, y = VRYieldV01)) +  
  geom_point() +  
  labs(title = "Scatter Plot of OM and Yield",  
       x = "OM",  
       y = "Yield") +  
  theme_classic()
```

Scatter Plot of OM and Yield



Yield Based on GridId

```
ggplot(df1, aes(x = VRYieldV01)) +  
  geom_histogram(binwidth = 3, aes(y = ..density..)) +  
  facet_wrap(~GridId, scales = "free") +  
  labs(x = "Yield", y = "Frequency") +  
  theme_classic()
```



We can see that the area of the grid influences yield. The grids along the left and right of the field tend to have more zero yields. Grids 8-11 are the left border and Grids 52-55 are the right border. Grids 8, 19, 30, 41, and 52 create the top border. Grids 11, 22, 33, 44, and 55 create the bottom border. Some of the distributions are more skewed than the others.

Comparing the Different Grids

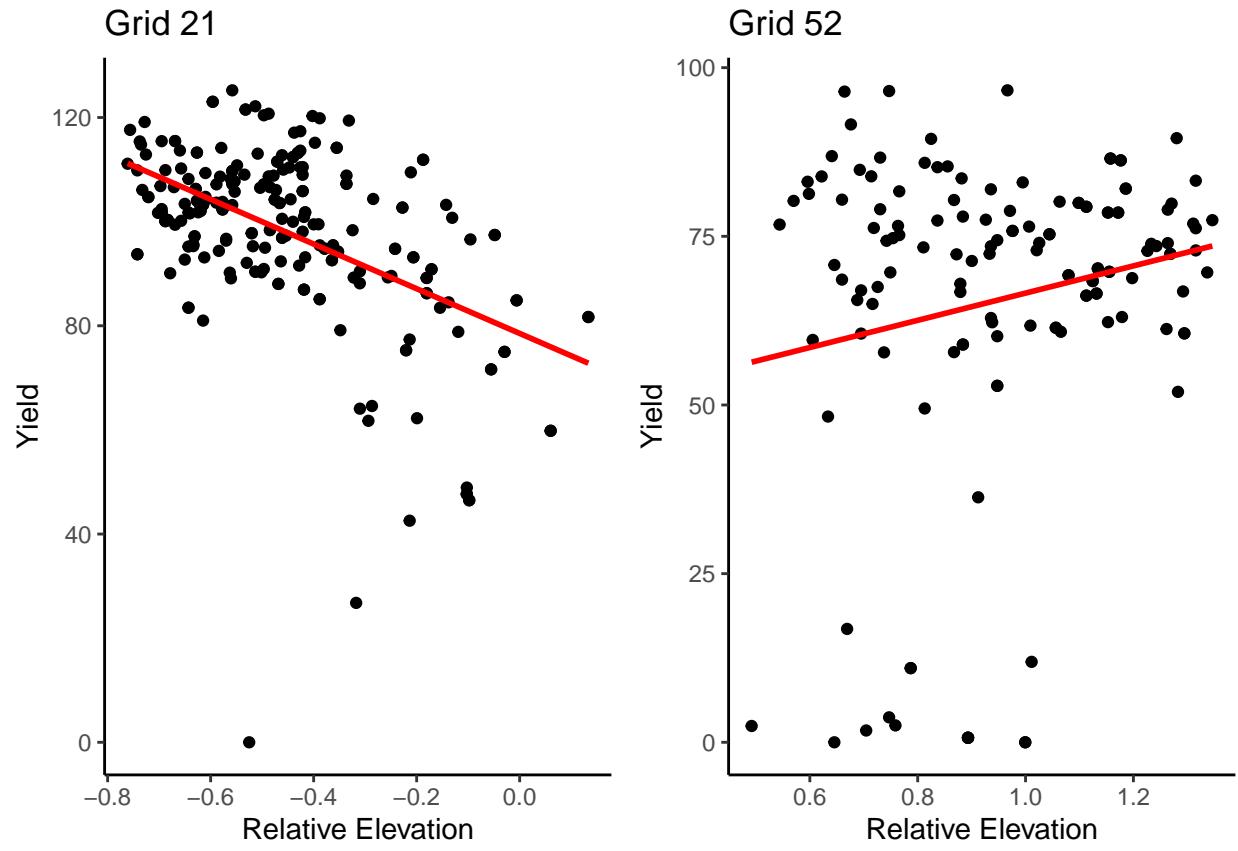
Summary Statistics of each Grid

All of the grids in the 2nd column (Grid 19-22) of the field had the largest yield with each mean being over 90. Grid 21 having the highest yield. The lowest average yield of a grid came from Grid 52 (see Table 2 - Table 21).

Comparing Grid 21 to Grid 52

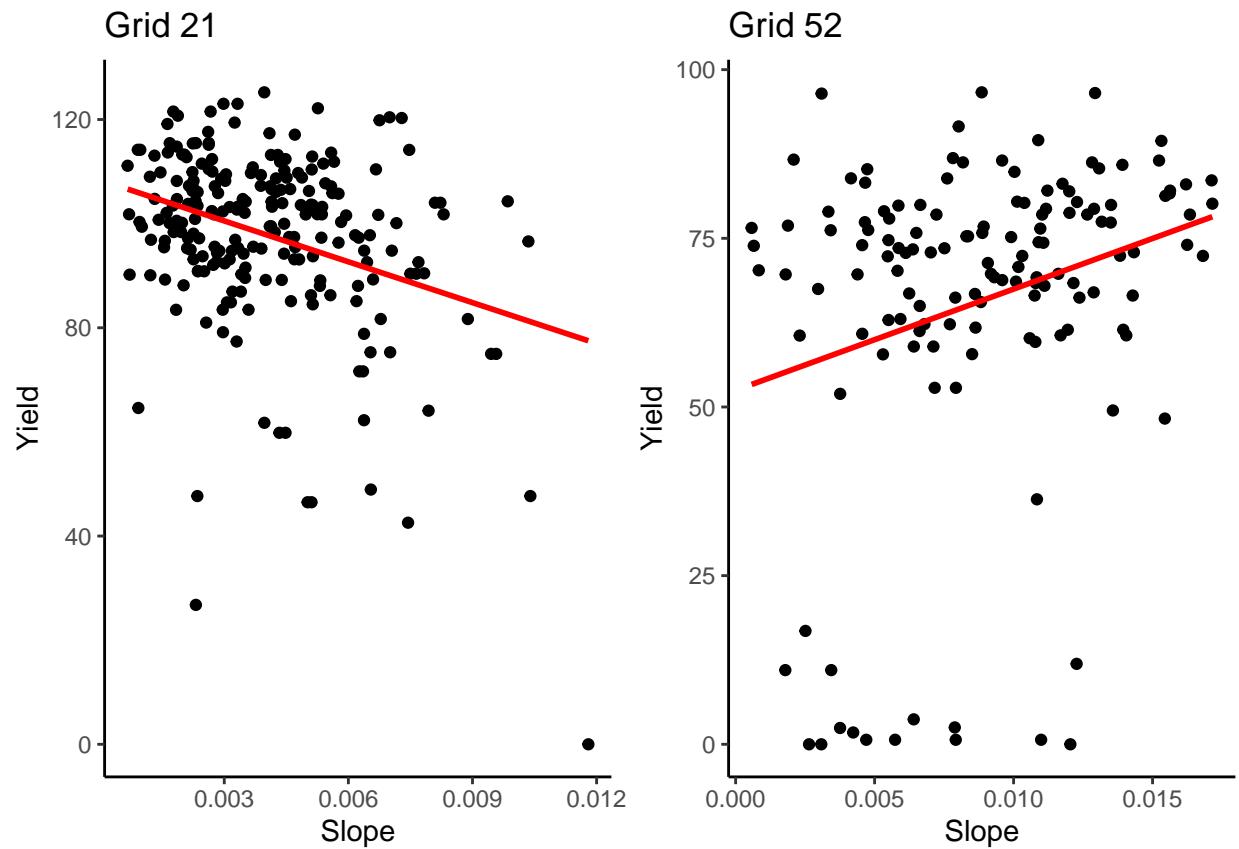
Relative Elevation vs. Yield

We can see a clearer relationship between the response and relative elevation in Grid 21. It seems like areas of lower relative elevation produce higher yield. This might be due to soybeans needing more water for proper growth.



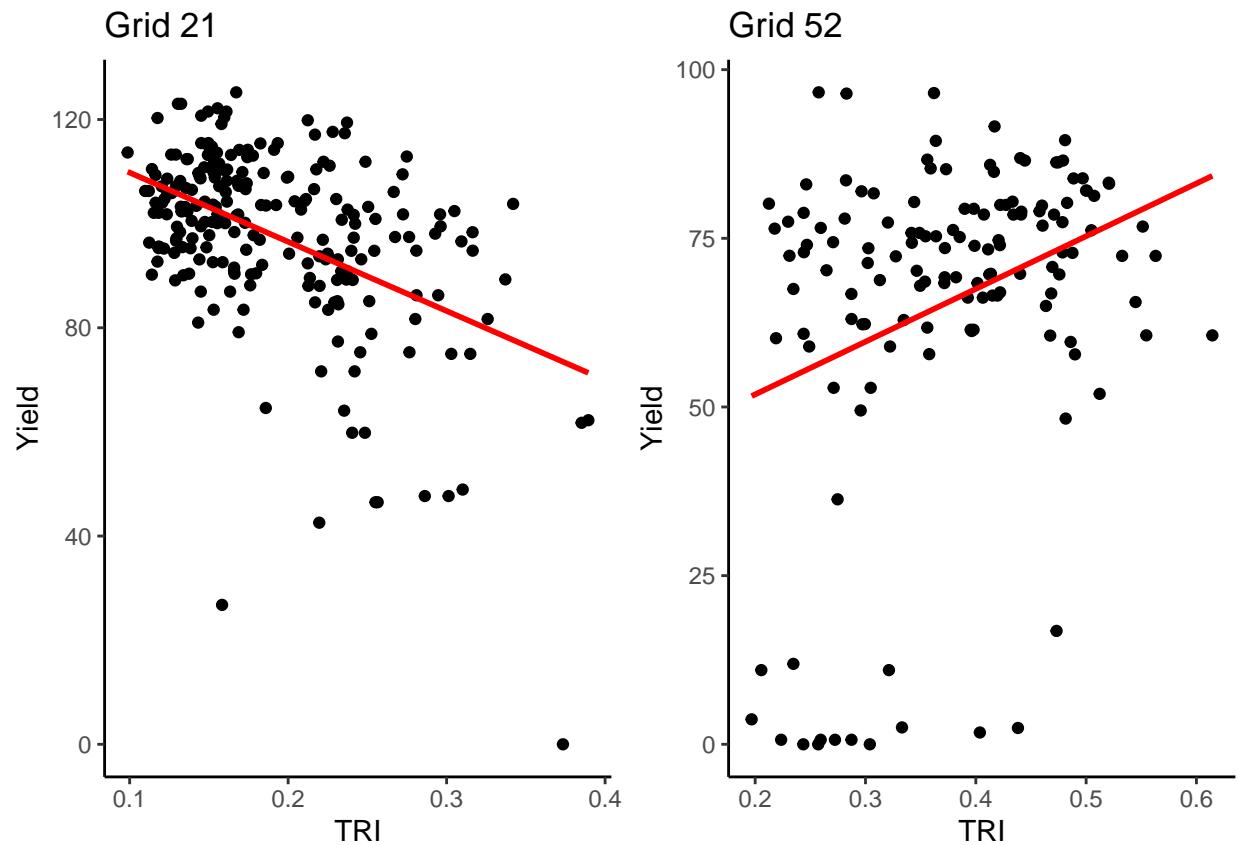
Slope vs. Yield

In Grid 52, there is a notable concentration of data points with zero yield. Considering that Grid 21 exhibits the highest yields, it raises the possibility of a correlation between minimal slope (little to no slope) and enhanced production.



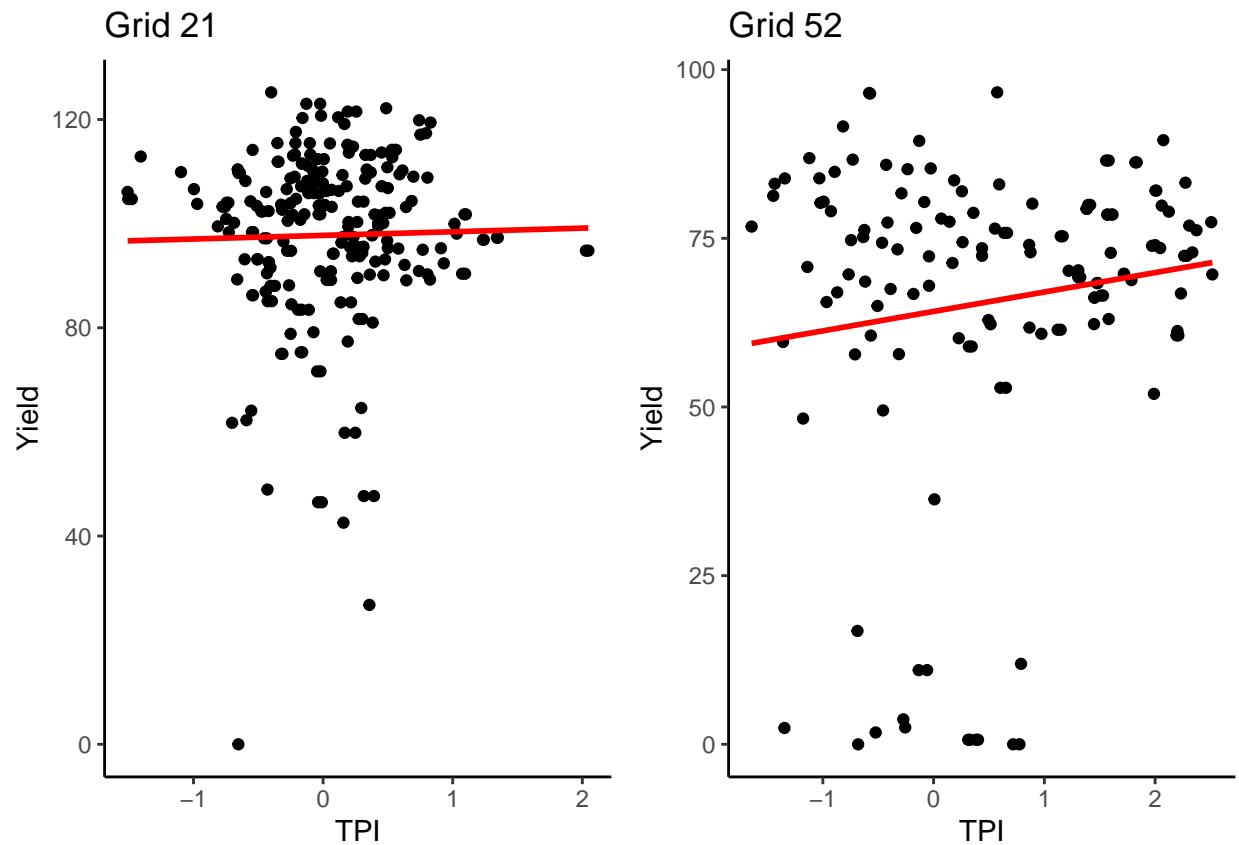
Terrain Ruggedness vs. Yield

The results are similar to the previous plots.



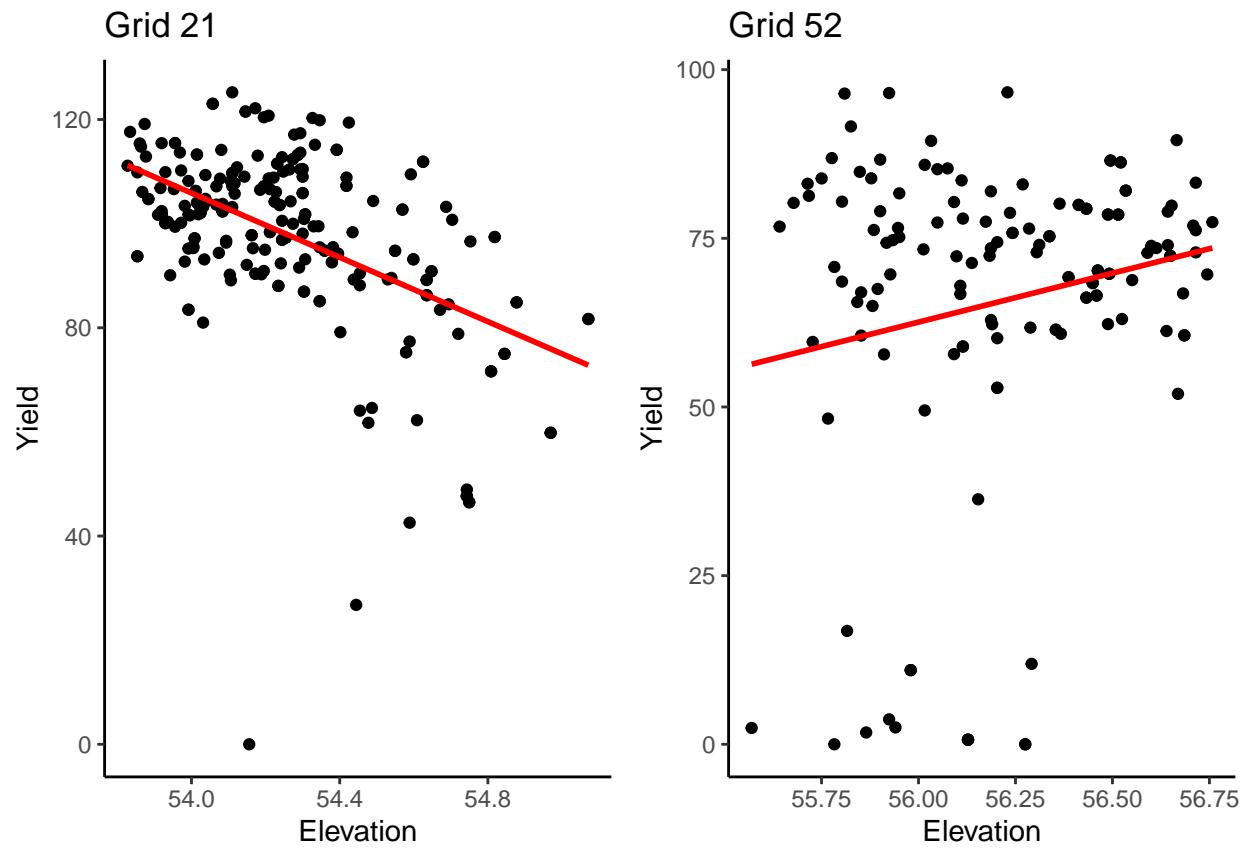
Topographic Position vs. Yield

Overall the relationship between the response and TPI is hard to distinguish.



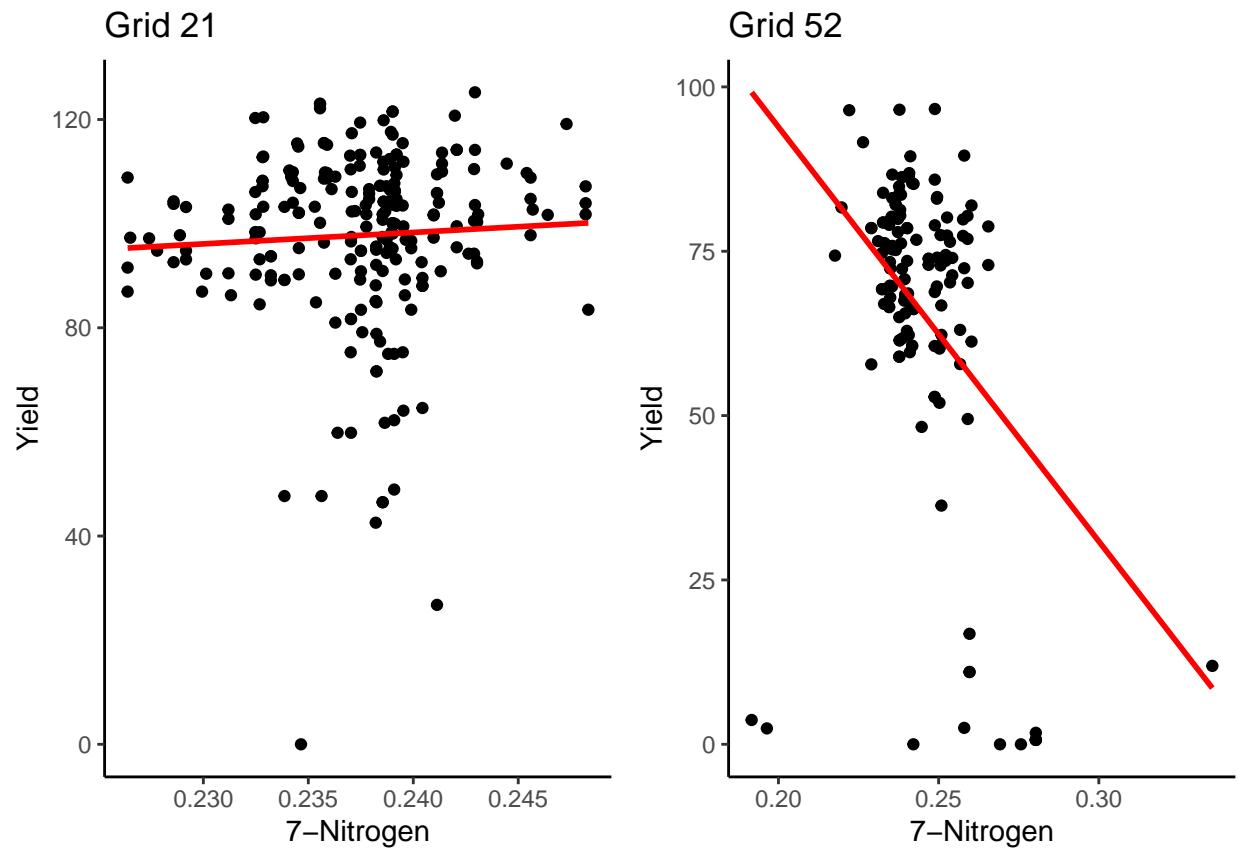
Elevation vs. Yield

The results are similar to the relative elevation plots.



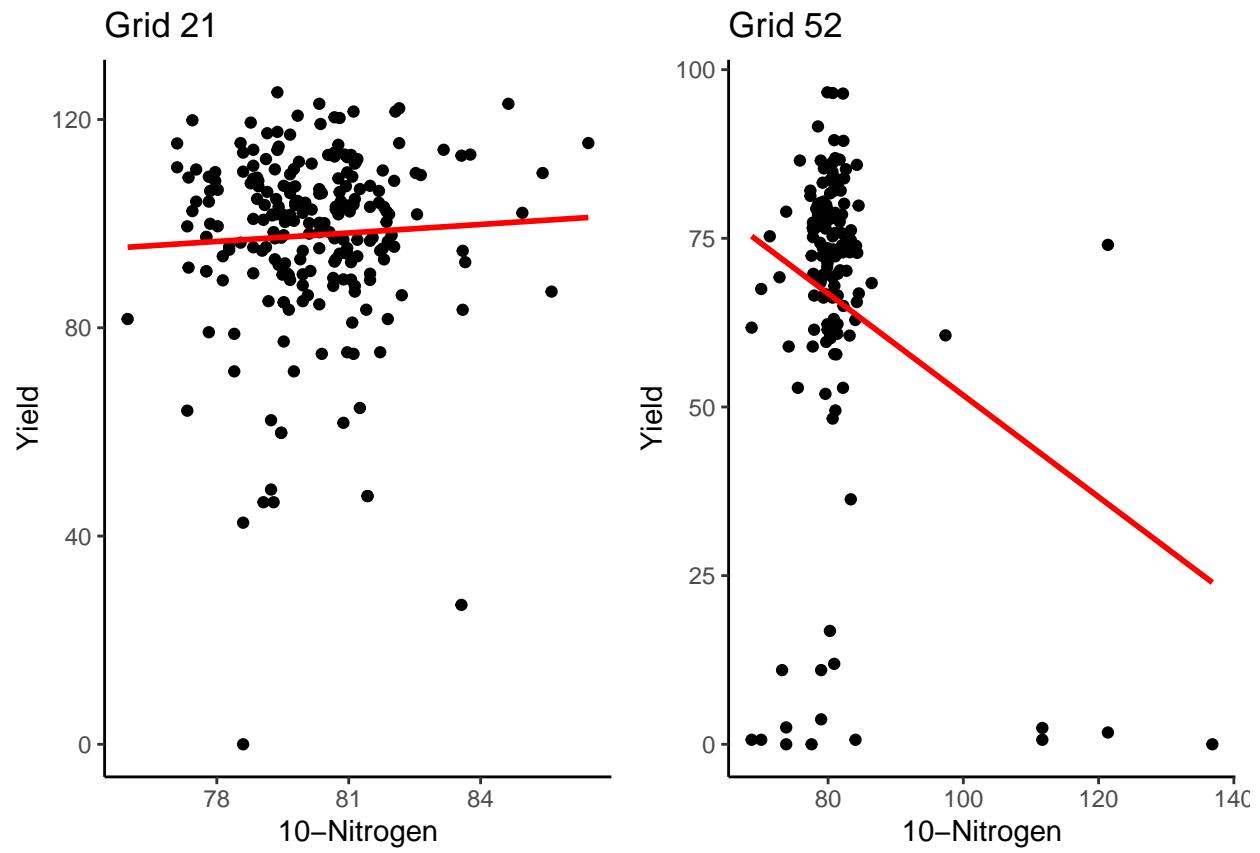
Application 7 Nitrogen Rate vs. Yield

No clear relationship can be seen from either grids.



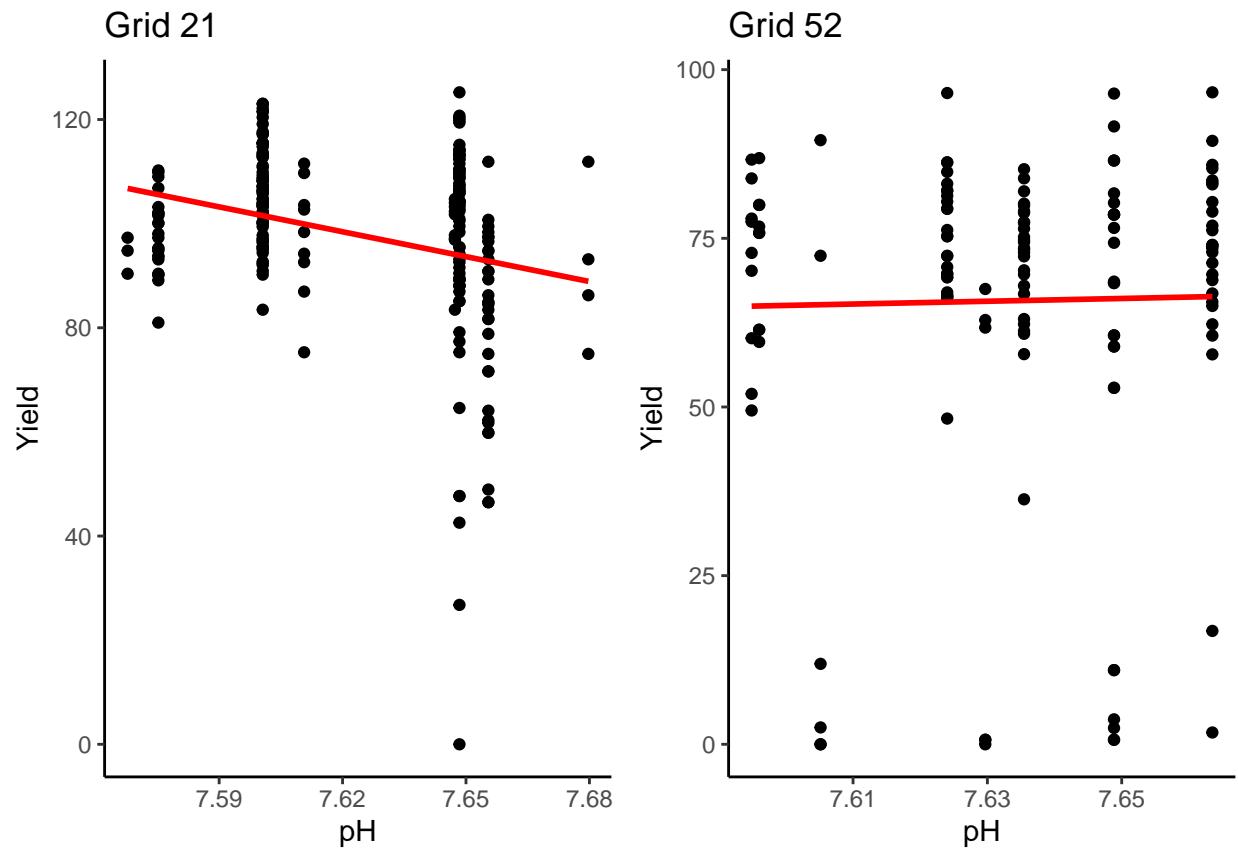
Application 10 Nitrogen Rate vs. Yield

No clear relationship can be seen from either grids.



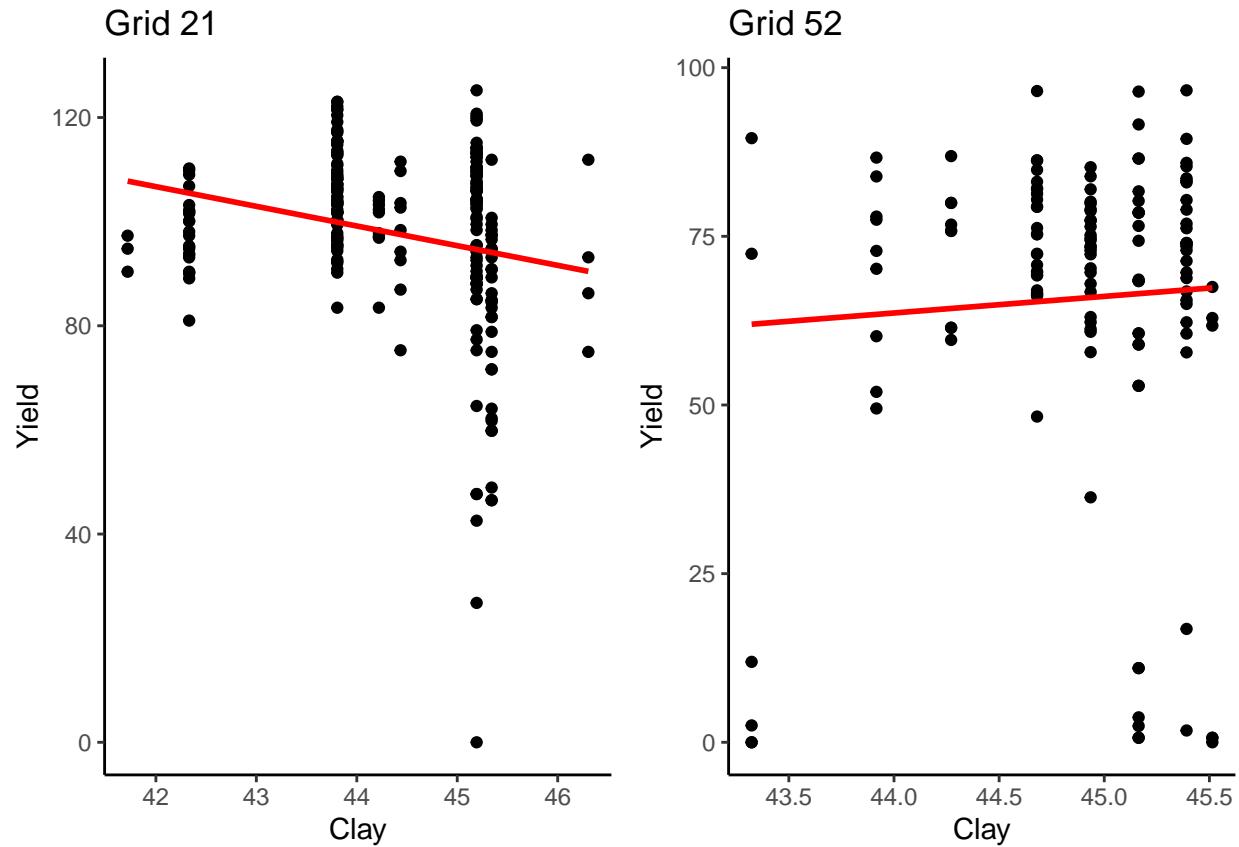
pH Mean vs. Yield

In Grid 21, pH's under 7.6 consistently give high yields.



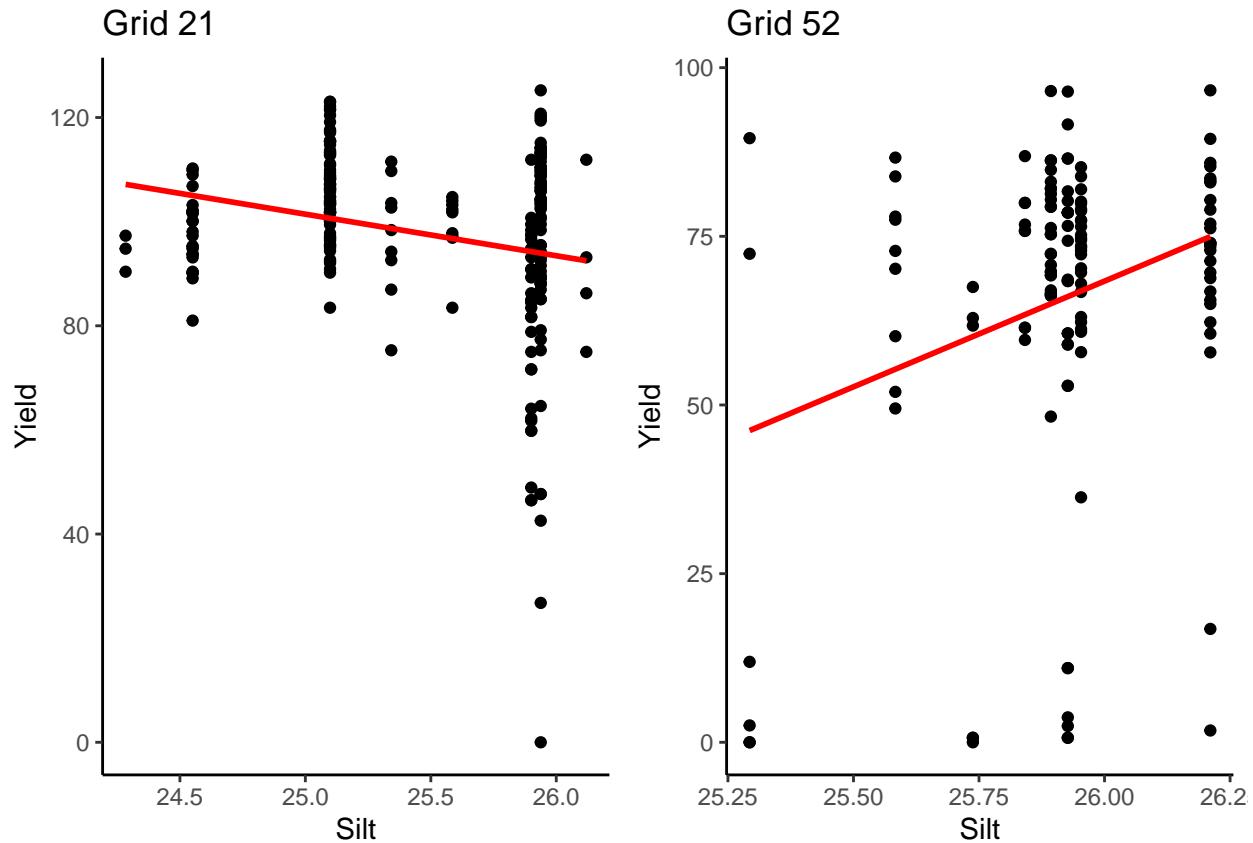
Clay Mean vs. Yield

In Grid 21, clay means below 45 consistently give high yields.



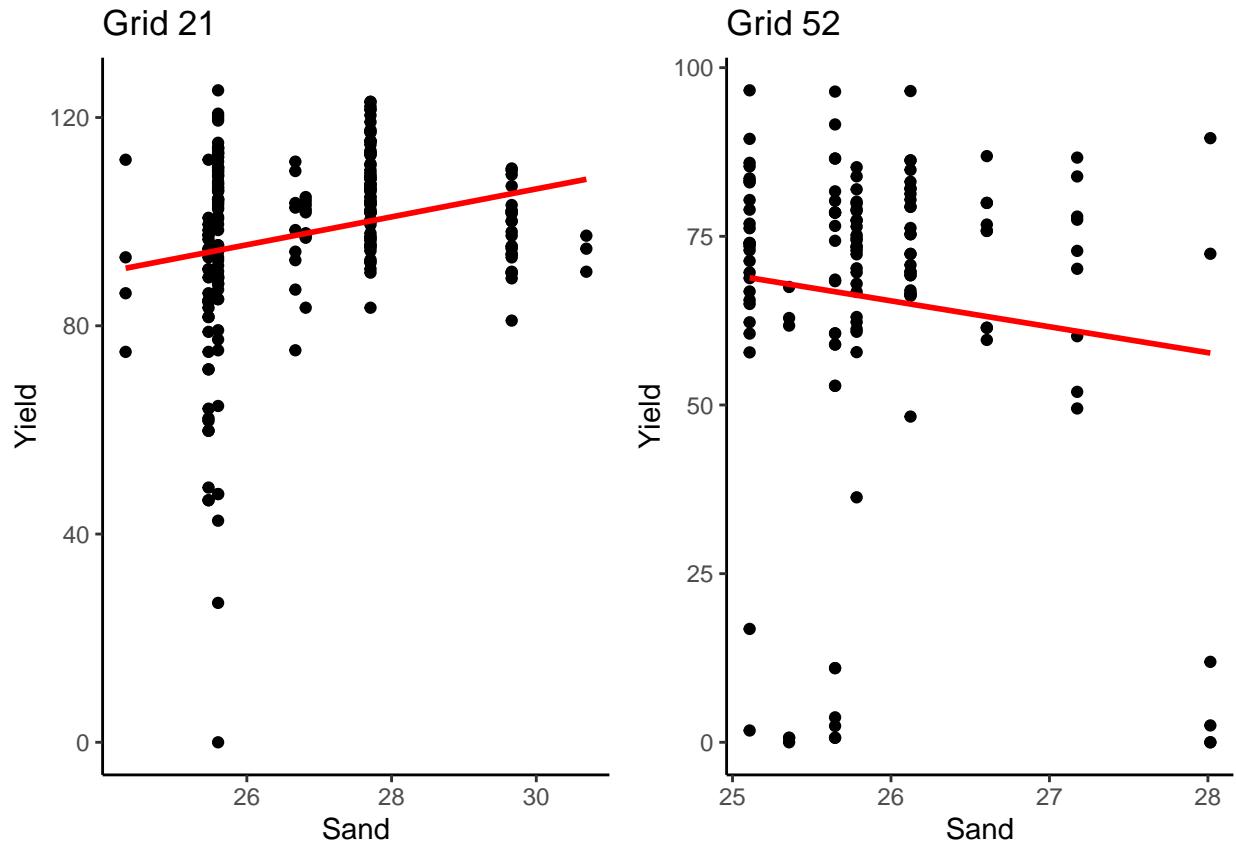
Silt Mean vs. Yield

In Grid 21, there's a consistent trend indicating that lower levels of silt are associated with higher yields. This finding aligns with expectations, as clay and silt are positively correlated. The observed pattern suggests that reduced silt content contributes to increased crop productivity in this specific grid.



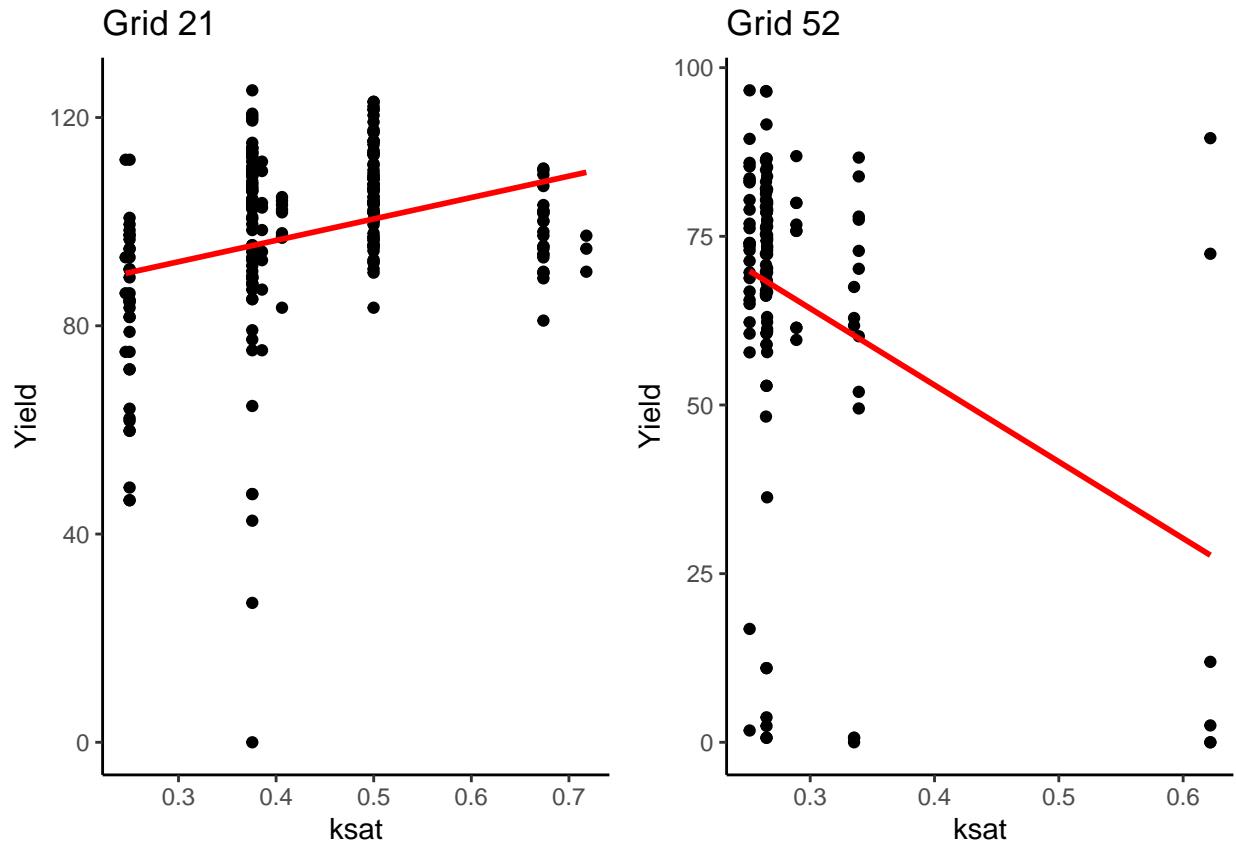
Sand Mean vs. Yield

In Grid 21, a consistent pattern emerges where higher levels of sand are correlated with consistently high yields. This relationship aligns with expectations, as higher sand content is associated with increased Ksat, representing hydraulic conductivity or the soil's ability to facilitate water flow. The observed connection implies that the enhanced hydraulic conductivity linked to higher sand content contributes to improved water movement in the soil, potentially influencing the higher crop yields in this grid.



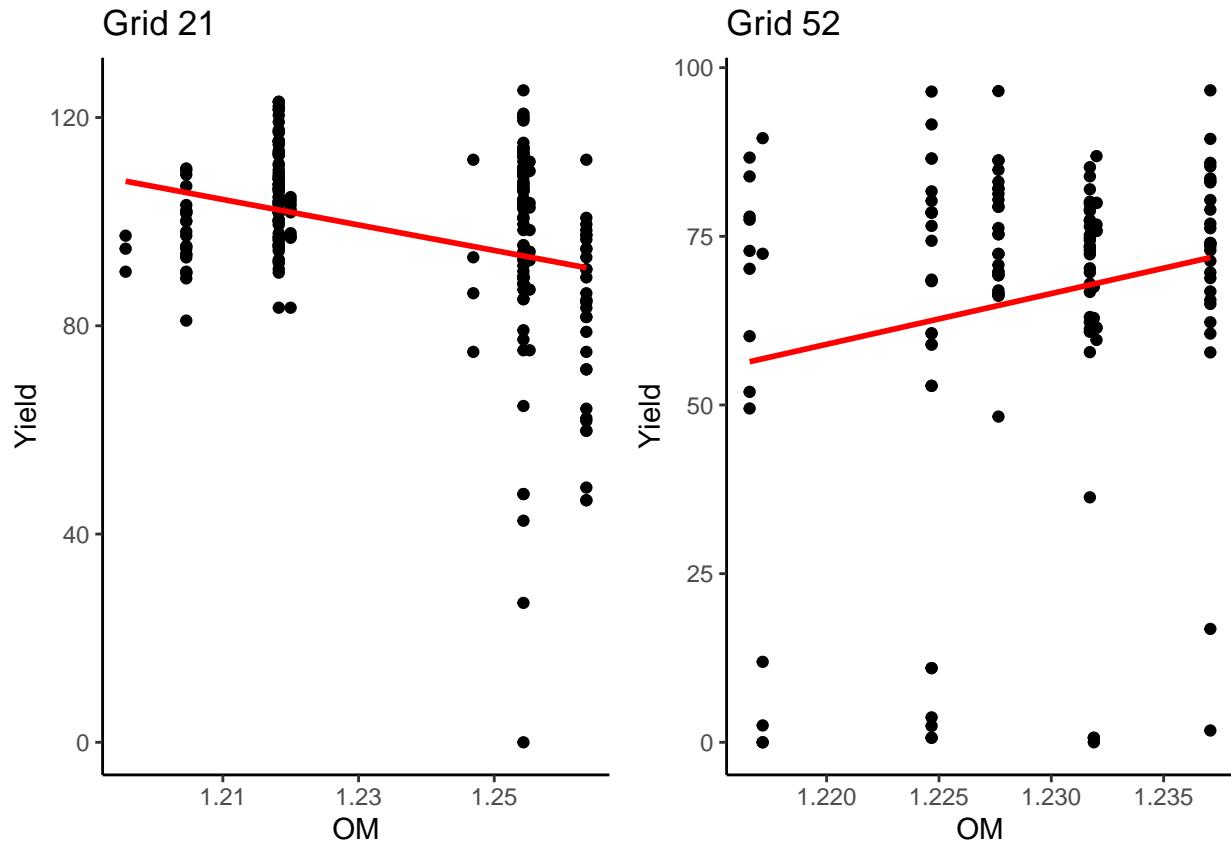
Ksat Mean vs. Yield

In Grid 21, there's a consistent trend indicating that higher Ksat (saturated hydraulic conductivity) is associated with consistently high yields. Conversely, in Grid 52, numerous high-yield instances are observed with lower Ksat values. This aligns with the earlier findings related to sand mean results. Given that Grid 52 has fewer data points with increased sand mean, it's plausible that the grid would exhibit many data points with low Ksat values, contributing to the observed pattern of high yields associated with lower Ksat in that grid.



OM Mean vs. Yield

In Grid 21, a consistent pattern emerges where lower organic matter (OM) levels are associated with consistently high yields. Notably, there is considerable yield variation when OM exceeds 1.23. This finding suggests that reduced organic matter content may contribute to enhanced crop yields in Grid 21, while higher OM levels beyond a certain threshold introduce variability in yield outcomes.



Conclusion for Comparing Grid 21 and Grid 52

When comparing Grid 21 to Grid 52, noticeable differences become apparent, likely attributable to the distinct characteristics of each grid's location. In Grid 21, factors such as low elevation, low slope, low TRI, varied nitrogen 7 and nitrogen 10 applications, pH under 7.6, clay means under 45, and sand means greater than 26, contributing to high Ksat means, collectively seem to be associated with increased yields. These variations highlight the multifaceted nature of factors influencing crop productivity in different locations.

Prediction Model

Random Forest (RF)

Given the dataset's substantial size and the presence of numerous explanatory variables, utilizing Random Forest (RF) as a predictive model appears to be a fitting choice. The robustness of RF in handling complex datasets with multiple predictors makes it well-suited for capturing intricate relationships and achieving accurate predictions.

Optimizing the RF

Before optimizing the RF, let's see how well it performs with only the default settings.

```

set.seed(123)
rf_model1 <- randomForest(VRYieldV01~, data = train, xtest = test_x,
                           ytest=test_y, keep.forest = TRUE)
rf_model1

##
## Call:
##   randomForest(formula = VRYieldV01 ~ ., data = train, xtest = test_x,      ytest = test_y, keep.forest
##                 Type of random forest: regression
##                 Number of trees: 500
##   No. of variables tried at each split: 5
##
##           Mean of squared residuals: 225.582
##           % Var explained: 61.21
##           Test set MSE: 194.28
##           % Var explained: 69.13

```

We can see that the test MSE is quite high without any optimization.

Let's look at the MSE when using all the trees to predict the train samples.

```

pred.train1 = predict(rf_model1, train)
mean((pred.train1-train_y)^2)

```

```

## [1] 45.96056

```

The MSE is very different and better than the test MSE. The MSE on predicting the train set will be used to evaluate the model.

Now let's try to optimize the RF by tuning the hyperparameters and evaluating the models using the MSE on the model predicting the train set.

Finding Optimal Number of Trees Start with finding the optimal number of trees for the RF.

```

trees <- seq(from = 200, to = 1000, by = 100)
z <- length(trees)
rf.mse1 <- vector(mode = "numeric", length = z)
N = nrow(df1)
for (i in 1:z) {
  set.seed(123)
  index <- sample(1:N, size = N*.8, replace = F)
  train_x <- df1[index, 2:17]
  test_x <- df1[-index, 2:17]
  train_y <- df1[index, 1]
  test_y <- df1[-index, 1]
  train <- df1[index,]
  test <- df1[-index,]
  rf_model1 <- randomForest(VRYieldV01~, data = train, xtest = test_x,
                             ytest=test_y, ntree=trees[i], keep.forest = TRUE)
  pred.train1 = predict(rf_model1, train)
  rf.mse1 [i] <- mean((pred.train1-train_y)^2)
}

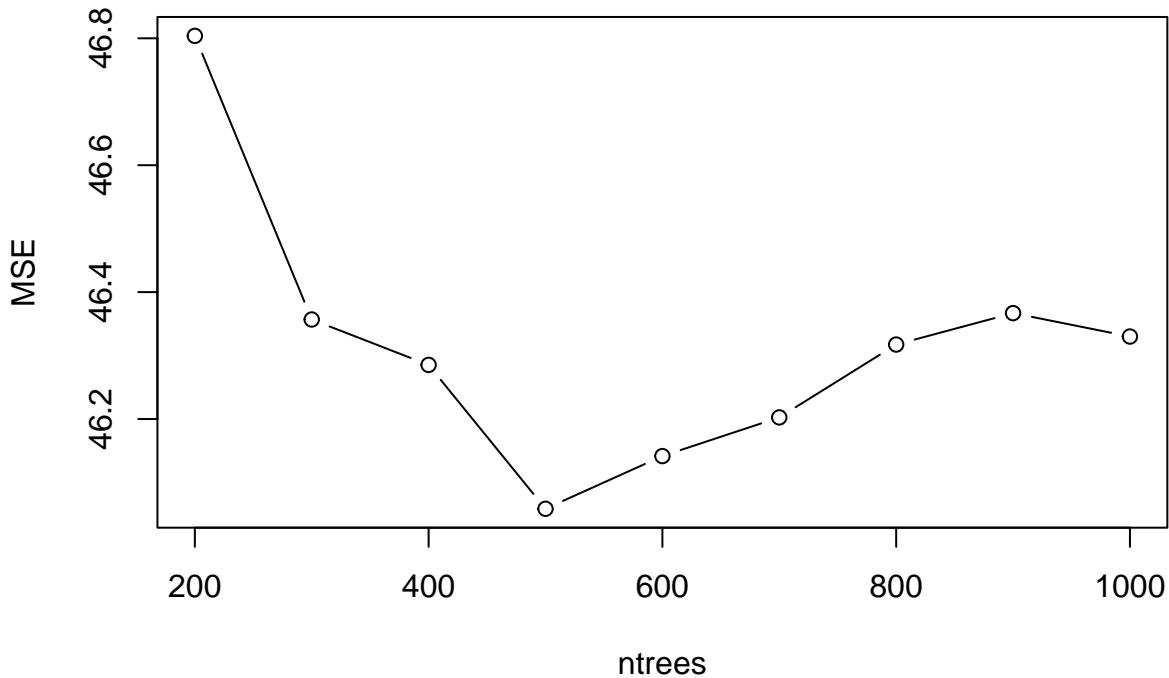
```

```

}

optimal.ntrees <- trees[which.min(rf.mse1)]
plot(trees, rf.mse1, type = "b", xlab = "ntrees", ylab = "MSE")

```



Finding Optimal m We will now find the optimal m using the optimal number of trees based on MSE. As we know, too large m will have high correlation and low bias.

```

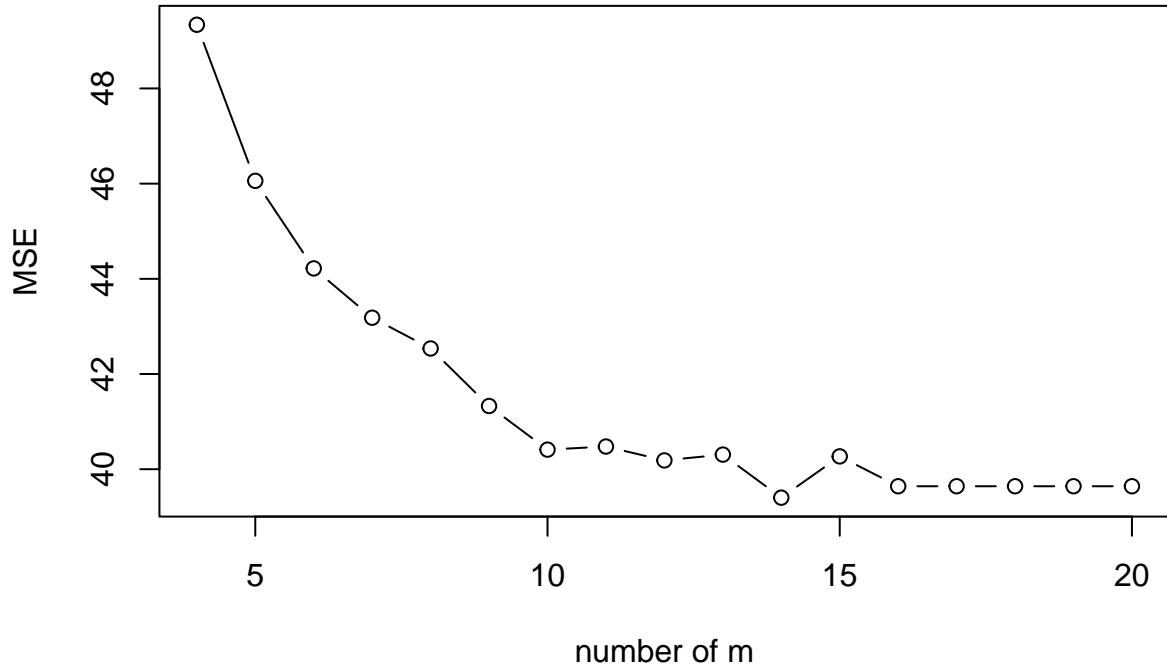
m <- seq(from = 4, to = 20, by = 1)
h <- length(m)
rf.mse1 <- vector(mode = "numeric", length = h)
for (i in 1:h) {
  set.seed(123)
  index <- sample(1:N, size = N*.8, replace = F)
  train_x <- df1[index, 2:17]
  test_x <- df1[-index, 2:17]
  train_y <- df1[index, 1]
  test_y <- df1[-index, 1]
  train <- df1[index,]
  test <- df1[-index,]
  rf_model1 <- randomForest(VRYieldV01~., data = train, xtest = test_x,
    ytest=test_y, mtry = m[i],
    ntree=optimal.ntrees, keep.forest = TRUE)
  pred.train1 = predict(rf_model1, train)

```

```

rf.mse1 [i] <- mean((pred.train1-train_y)^2)
}
optimal.m <- m[which.min(rf.mse1)]
plot(m, rf.mse1, type = "b", xlab = "number of m", ylab = "MSE")

```



Finding Optimal Tree Depth We will now find the optimal tree depth using the optimal number of trees and optimal m based on MSE. The smaller the node size, the deeper the tree.

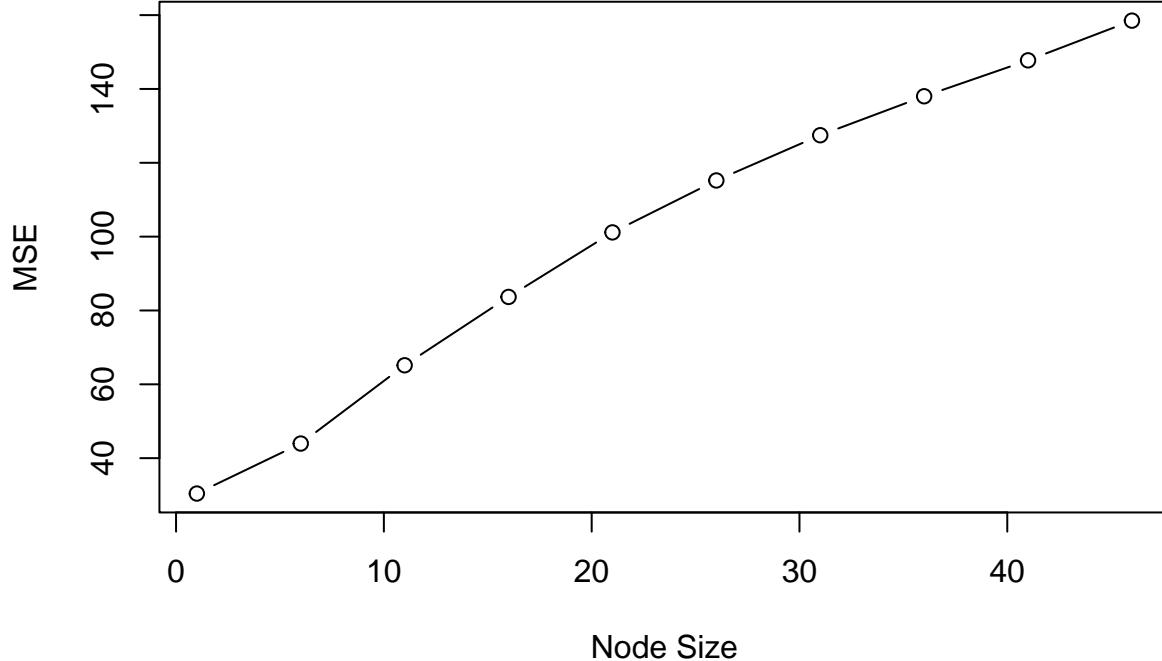
```

depth <- seq(from = 1, to = 50, by = 5)
d <- length(depth)
rf.mse1 <- vector(mode = "numeric", length = d)
for (i in 1:d) {
  set.seed(123)
  index <- sample(1:N, size = N*.8, replace = F)
  train_x <- df1[index, 2:17]
  test_x <- df1[-index, 2:17]
  train_y <- df1[index, 1]
  test_y <- df1[-index, 1]
  train <- df1[index,]
  test <- df1[-index,]
  rf_model1 <- randomForest(VRYieldV01~, data = train, xtest = test_x,
                               ytest=test_y, mtry = optimal.m, nodesize = depth[i],
                               ntree=optimal.ntrees, keep.forest = TRUE)
  pred.train1 = predict(rf_model1, train)
}
```

```

rf.mse1 [i] <- mean((pred.train1-train_y)^2)
}
optimal.depth <- depth[which.min(rf.mse1)]
plot(depth, rf.mse1, type = "b", xlab = "Node Size", ylab = "MSE")

```



Optimized RF

```

n=10
rf.mse1 <- vector(mode = "numeric", length = n)
for (i in 1:n) {
  set.seed(i)
  index <- sample(1:N, size = N*.8, replace = F)
  train_x <- df1[index, 2:17]
  test_x <- df1[-index, 2:17]
  train_y <- df1[index, 1]
  test_y <- df1[-index, 1]
  train <- df1[index,]
  test <- df1[-index,]
  rf_model1 <- randomForest(VRYieldV01~, data = train, xtest = test_x,
                               ytest=test_y, mtry = optimal.m, nodesize = optimal.depth,
                               ntree=optimal.ntrees, keep.forest = TRUE)
  pred.train1 = predict(rf_model1, train)

```

```

rf.mse1 [i] <- mean((pred.train1-train_y)^2)
}

```

Results of Optimized RF

The Mean Squared Error (MSE) across all iterations indicates that the optimized Random Forest (RF) model yields lower values compared to the unoptimized RF. This outcome suggests that the tuning or optimization of the model has led to improved predictive performance, resulting in reduced errors in predicting the target variable.

RF_MSE
30.23014
30.12148
27.94697
29.55104
29.07963
29.00461
28.30417
29.92180
29.33213
29.82861

Removing Outliers in Response Variable, Yield

The result of the Shapiro-Wilk test suggests that, before removing outliers, the distribution of the response variable significantly deviates from a normal distribution, given that the p-value is below the threshold of 0.05.

```

## 
## Shapiro-Wilk normality test
## 
## data: df1$VRYieldV01
## W = 0.89336, p-value < 2.2e-16

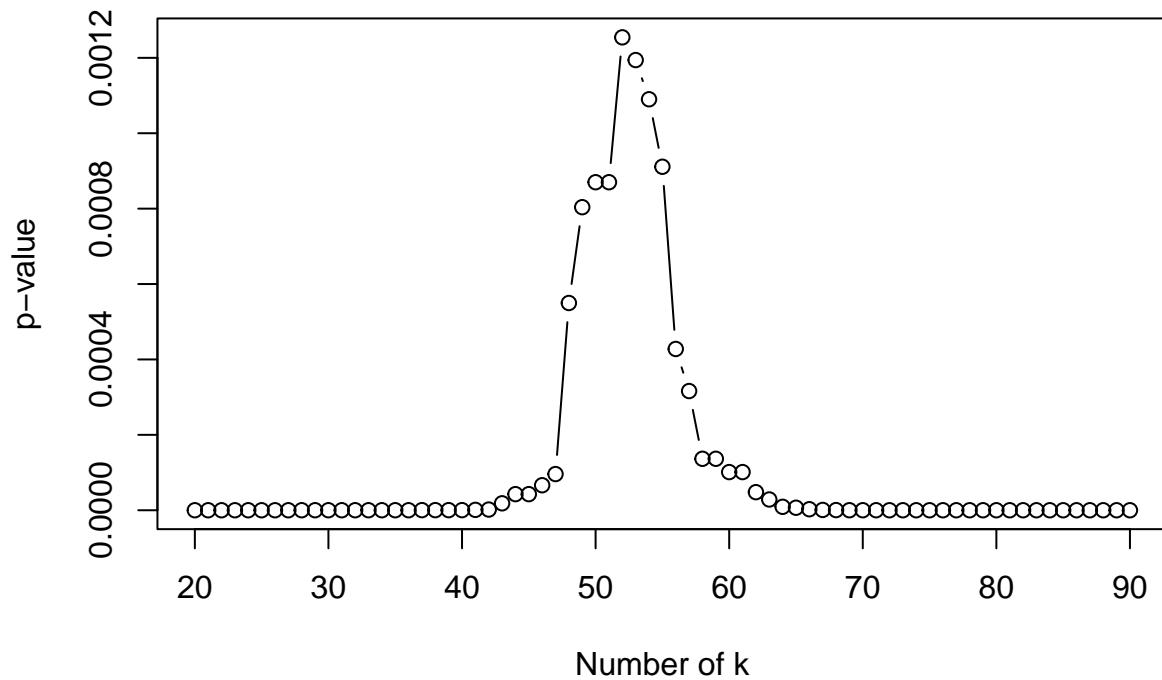
```

The leftward skewness in the distribution prompts the consideration of reducing the number of outliers. Incrementally introducing a lower value threshold appears to align the response variable more closely with a normal distribution.

```

a = 0.005*max(df1$VRYieldV01)
k <- seq(from = 20, to = 90, by = 1)
h <- length(k)
p.value <- vector(mode = "numeric", length = h)
for (i in 1:h) {
  set.seed(1)
  subset_condition <- df1$VRYieldV01 >= k[i] * a
  df2 <- df1[subset_condition, ]
  p.value[i] <- shapiro.test(df2$VRYieldV01)$p.value
}
optimal.k <- k[which.max(p.value)]
plot(k, p.value, type = "b", xlab = "Number of k", ylab = "p-value")

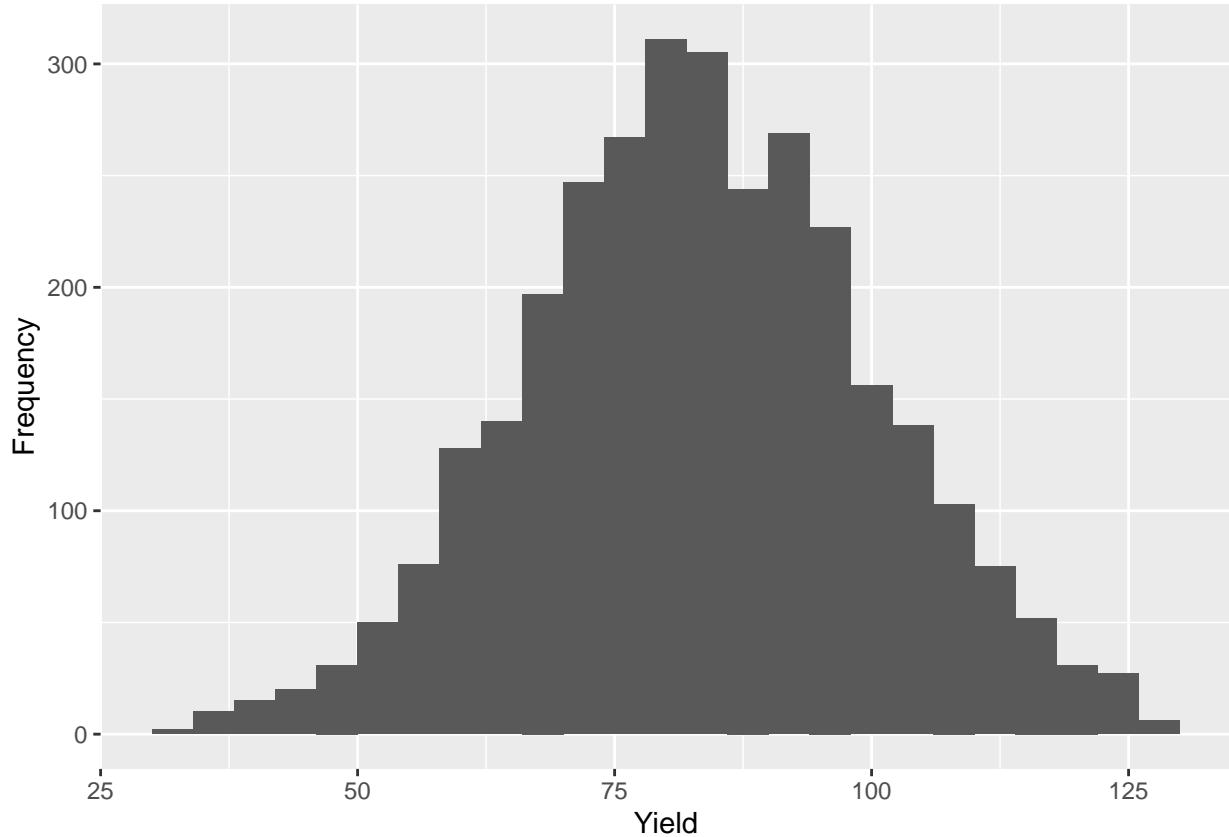
```



For the normality test to not be rejected, $k = 190$. This leaves only 32 observation which is not good, and the normality test does not guarantee the data is normally distributed. Therefore, using the first that there is a significant change in the p-value will help reduce many low value outliers. There's a significant change in the p-value between 40 and 60 as seen in the plot.

Let's see the normality test and distribution of the response variable now after the outliers have been removed.

```
##  
## Shapiro-Wilk normality test  
##  
## data: df3$VRYieldV01  
## W = 0.99819, p-value = 0.001254
```



The response variable now exhibits increased symmetry and less skewness.

Optimized RF on the Reduced Outlier Dataset

```

n=10
R = nrow(df3)
rf.mse2 <- vector(mode = "numeric", length = n)
for (i in 1:n) {
  set.seed(i)
  index2 <- sample(1:R, size = R*.8, replace = F)
  train_x2 <- df3[index2, 2:17]
  test_x2 <- df3[-index2, 2:17]
  train_y2 <- df3[index2, 1]
  test_y2 <- df3[-index2, 1]
  train2 <- df3[index2,]
  test2 <- df3[-index2,]
  rf_model2 <- randomForest(VRYieldV0~., data = train2, xtest = test_x2,
                               ytest=test_y2, mtry = optimal.m, nodesize = optimal.depth,
                               ntree=optimal.ntrees, keep.forest = TRUE)
  pred.train2 = predict(rf_model2, train2)

  rf.mse2 [i] <- mean((pred.train2-train_y2)^2)
}

```

Results of Optimized RF on Reduced Outlier Data

Here is the MSE of all the iterations using the optimized RF on the reduced outlier data.

```
##  
## Call:  
##   randomForest(formula = VRYieldV01 ~ ., data = train2, xtest = test_x2,      ytest = test_y2, mtry =  
##               Type of random forest: regression  
##               Number of trees: 500  
## No. of variables tried at each split: 14  
##  
##               Mean of squared residuals: 107.5626  
##               % Var explained: 62.48  
##               Test set MSE: 91.67  
##               % Var explained: 67.12
```

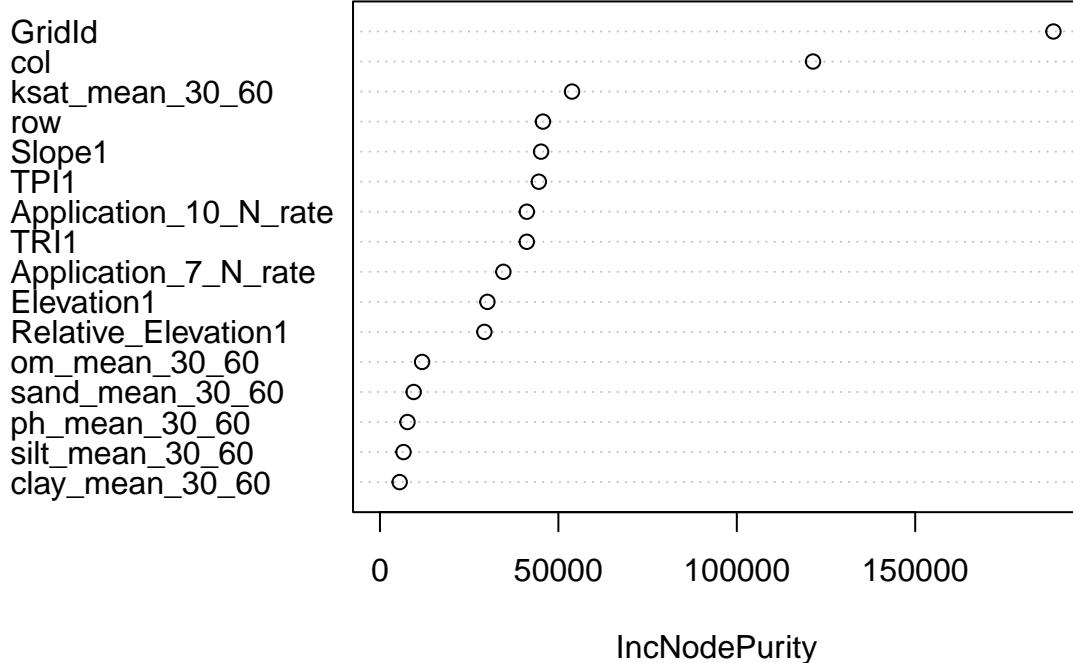
RF_MSE
14.13086
14.22481
14.70413
14.42047
14.10298
13.66012
14.28290
14.21335
14.49230
14.51861

The performance of the Random Forest (RF) prediction model showed a substantial improvement in Mean Squared Error (MSE) when certain lower bound outliers were removed. This suggests that the removal of these outliers has a positive impact on the model's accuracy and its ability to make predictions, resulting in lower prediction errors. The refinement of the dataset through outlier removal appears to contribute to the overall enhancement of the RF model's predictive performance.

Variable Importance

```
varImpPlot(rf_model2, sort = T, main = "Relative Variable Importance")
```

Relative Variable Importance



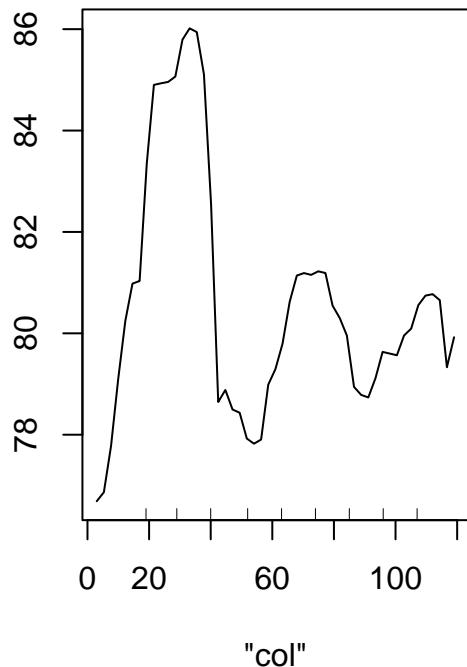
The analysis indicates that GridId holds the highest importance in the prediction model, followed by col (column). This suggests that the geographical location, as represented by GridId, plays a crucial role in predicting crop yield.

Partial Dependency Plot

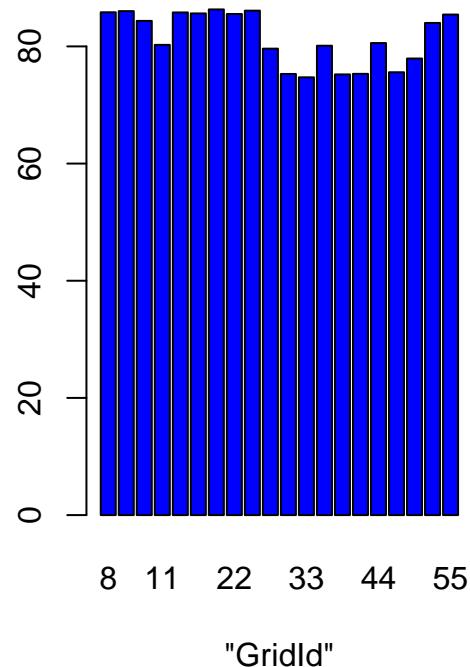
The partial dependency plot reveals a discernible pattern characterized by both increasing and decreasing trends in relation to the crop's location. Specifically, Grids 30-33 exhibit the lowest values according to the column variable.

```
par(mfrow=c(1,2))
partialPlot(rf_model2, train2, 'col')
partialPlot(rf_model2, train2, 'GridId')
```

Partial Dependence on "col"



Partial Dependence on "GridId"



Optimized RF on Columns 25-45

Since the highest yielding grids have a clearer relationship between the explanatory and response variable, let's try predicting on this subset!

Results of Un-Optimized RF

```
rf_model3 <- randomForest(VRYieldV01~., data = train3, xtest = test_x3,
                           ytest=test_y3, keep.forest = TRUE)
rf_model3

##
## Call:
##  randomForest(formula = VRYieldV01 ~ ., data = train3, xtest = test_x3,      ytest = test_y3, keep.f.
```

```

pred.train3 = predict(rf_model3, train3)

mean((pred.train3-train_y3)^2)

```

```
## [1] 51.05158
```

The MSE for predicting using the train and the test MSE are both worse than using the original dataset.

Results of Optimized RF

The MSE showed a slight improvement from the unoptimized RF model. However, when compared to the MSE obtained using the reduced outlier dataset, using this subset performed significantly worse.

```

n=10
rf.mse3 <- vector(mode = "numeric", length = d)
for (i in 1:n) {
  set.seed(i)
  index3 <- sample(1:G, size = G*.8, replace = F)
  train_x3 <- col_df[index3, 2:16]
  test_x3 <- col_df[-index3, 2:16]
  train_y3 <- col_df[index3, 1]
  test_y3 <- col_df[-index3, 1]
  train3 <- col_df[index3,]
  test3 <- col_df[-index3,]
  rf_model3 <- randomForest(VRYieldV01~, data = train3, xtest = test_x3,
                             ytest=test_y3, mtry = optimal.m2, nodesize = optimal.depth2,
                             ntree=optimal.ntrees2, keep.forest = TRUE)
  pred.train3 = predict(rf_model3, train3)

  rf.mse3 [i] <- mean((pred.train3-train_y3)^2)
}

```

RF_MSE
37.28003
35.56850
35.27519
36.70898
35.91438
38.79728
37.35393
34.67273
36.35823
34.59119

Conclusion

The unexpected findings from the prediction model using the subset that generates the highest yields, particularly with a worse Mean Squared Error (MSE) compared to other subsets, indicate that predicting crop yield accurately involves more complexities than the identified relationships in Grid 21. Despite clear

connections between explanatory and response variables in this grid, achieving precise yield predictions remains challenging, especially when considering unaccounted factors like weather effects.

The observed importance of location in predicting high yields emphasizes the intricate nature of crop growth, with environmental and spatial variables playing pivotal roles. The unpredictability of certain outcomes may also highlight the impact of unmeasured factors, such as weather conditions, in outdoor crop cultivation.

As digital agriculture continues to advance, ongoing potential exists for technological innovations to enhance various aspects of crop growth and collection. However, the intricacies and subtleties involved in predicting crop yield stress the need for continuous research and technological developments in the field to address the multitude of influencing factors.

Quick Note: The analysis was conducted with “col” and “row” treated as factors initially. However, due to the large number of levels in the “col” variable (118 levels), I merged some levels for RF to run. Interestingly, the results remained consistent when using “col” and “row” variables as numeric. Opting for numeric representation provided a slight speed advantage in code running, and since no information was lost from the merging process, this approach was retained for efficiency.

Appendix

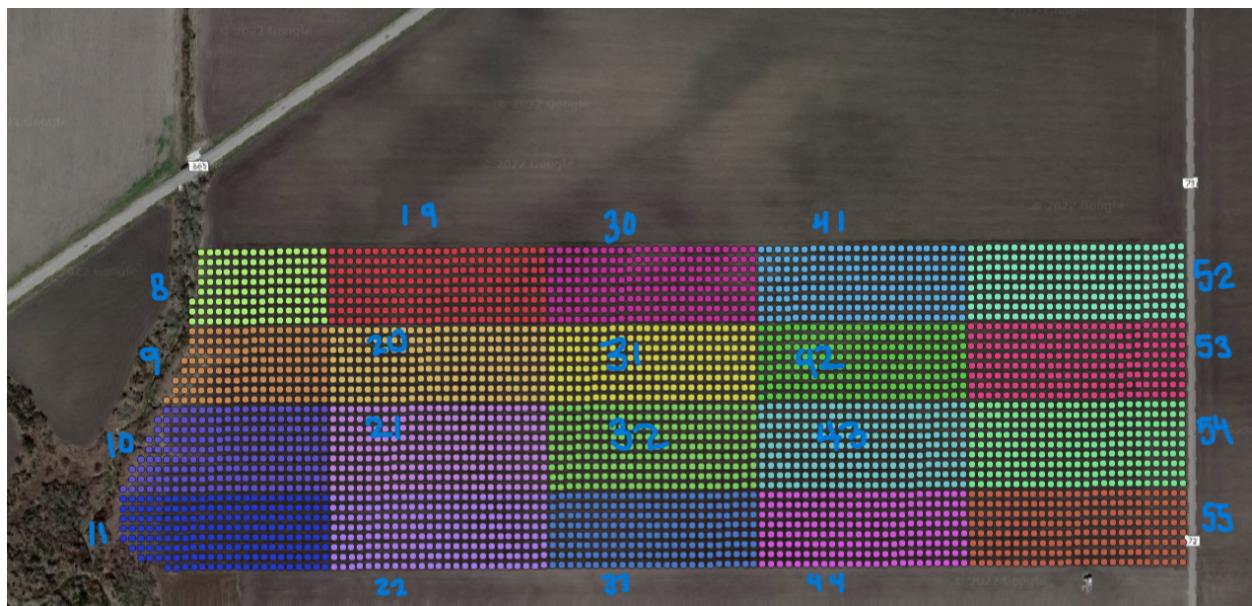


Figure 1: The 20 different grids created using QGIS for the variable “GridId”

Table 5: Grid 08

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
GridId*	1	87	1.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	NaN	NaN	0.00
VRYieldVOL	2	87	71.70	39.66	85.52	74.30	25.10	0.00	125.34	125.34	-	-0.86	4.25
											0.76		
row	3	87	60.55	1.72	61.00	60.56	1.48	58.00	63.00	5.00	-	-1.30	0.18
											0.05		
col	4	87	16.24	4.22	16.00	16.25	5.93	9.00	23.00	14.00	-	-1.24	0.45
											0.01		

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
Relative_Elevation15	87	-	0.72	-2.04	-1.90	0.79	-	-0.59	2.55	0.22	-1.07	0.08	
			1.89				3.15						
Slope1	6	87	0.02	0.01	0.02	0.02	0.01	0.00	0.05	0.05	1.14	1.01	0.00
TRI1	7	87	1.00	0.24	1.00	1.01	0.23	0.43	1.56	1.13	-	-0.28	0.03
											0.20		
TPI1	8	87	-	1.47	-0.39	-0.15	1.57	-	2.87	5.16	0.50	-1.01	0.16
			0.06				2.28						
Elevation1	9	87	52.26	1.00	52.05	52.25	1.10	50.51	54.06	3.55	0.22	-1.07	0.11
Application_7_N_10ate	87	0.25	0.11	0.24	0.24	0.02	0.21	1.26	1.05	8.55	74.54	0.01	
Application_10_N_11rate	87	86.10	22.55	80.94	81.45	2.74	70.23	211.09	140.86	4.29	18.62	2.42	
ph_mean_30_60	12	87	7.50	0.05	7.50	7.51	0.06	7.43	7.55	0.12	-	-1.04	0.01
											0.74		
clay_mean_30_60	13	87	40.28	1.80	39.45	40.33	2.47	37.78	42.34	4.56	-	-1.58	0.19
											0.20		
silt_mean_30_60	14	87	23.12	1.00	22.80	23.15	1.39	21.67	24.33	2.66	-	-1.35	0.11
											0.27		
sand_mean_30_60	15	87	32.82	2.52	33.84	32.76	3.77	29.74	36.38	6.64	0.19	-1.45	0.27
ksat_mean_30_60	16	87	1.33	0.76	1.73	1.30	0.96	0.54	2.37	1.83	0.20	-1.69	0.08
om_mean_30_60	17	87	1.15	0.04	1.14	1.15	0.03	1.10	1.20	0.11	0.14	-1.15	0.00

Table 6: Grid 09

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
GridId*	1	129	2.00	0.00	2.00	2.00	0.00	2.00	2.00	0.00	NaN	NaN	0.00
VRYieldVOL	2	129	78.63	37.96	90.97	82.79	23.09	0.00	127.21	127.21	-	-0.10	3.34
											1.09		
row	3	129	67.64	2.31	68.00	67.68	2.97	64.00	71.00	7.00	-	-1.26	0.20
											0.08		
col	4	129	15.40	4.72	15.00	15.44	5.93	6.00	23.00	17.00	-	-1.17	0.42
											0.04		
Relative_Elevation15	129	-	0.83	-1.97	-1.92	1.00	-	-0.66	2.83	-	-	-1.20	0.07
			1.93				3.49				0.09		
Slope1	6	129	0.02	0.01	0.02	0.02	0.01	0.00	0.05	0.05	0.46	-0.66	0.00
TRI1	7	129	1.02	0.28	1.01	1.02	0.28	0.37	1.62	1.26	0.01	-0.53	0.02
TPI1	8	129	-	0.98	-0.34	-0.37	1.21	-	1.22	3.47	-	-1.18	0.09
			0.39				2.25				0.10		
Elevation1	9	129	52.20	1.15	52.15	52.22	1.40	50.03	53.96	3.93	-	-1.20	0.10
											0.09		
Application_7_N_10ate	129	0.28	0.24	0.24	0.24	0.01	0.21	2.26	2.04	6.24	42.25	0.02	
Application_10_N_11rate	129	88.01	41.35	79.61	79.66	3.81	68.70	332.41	263.71	4.99	24.46	3.64	
ph_mean_30_60	12	129	7.53	0.03	7.55	7.53	0.04	7.40	7.58	0.18	-	2.15	0.00
											1.07		
clay_mean_30_60	13	129	40.95	1.46	41.58	41.03	1.14	37.59	42.34	4.76	-	-1.45	0.13
											0.37		
silt_mean_30_60	14	129	23.65	0.84	24.22	23.69	0.50	21.48	24.55	3.07	-	-1.16	0.07
											0.50		
sand_mean_30_60	15	129	31.65	2.10	30.43	31.54	1.14	29.66	36.44	6.78	0.41	-1.49	0.18
ksat_mean_30_60	16	129	1.16	0.55	0.91	1.13	0.43	0.62	2.54	1.92	0.42	-1.28	0.05
om_mean_30_60	17	129	1.18	0.04	1.20	1.18	0.01	1.12	1.21	0.10	-	-1.87	0.00
											0.29		

VRYIELDVOL Classified by Std Deviation

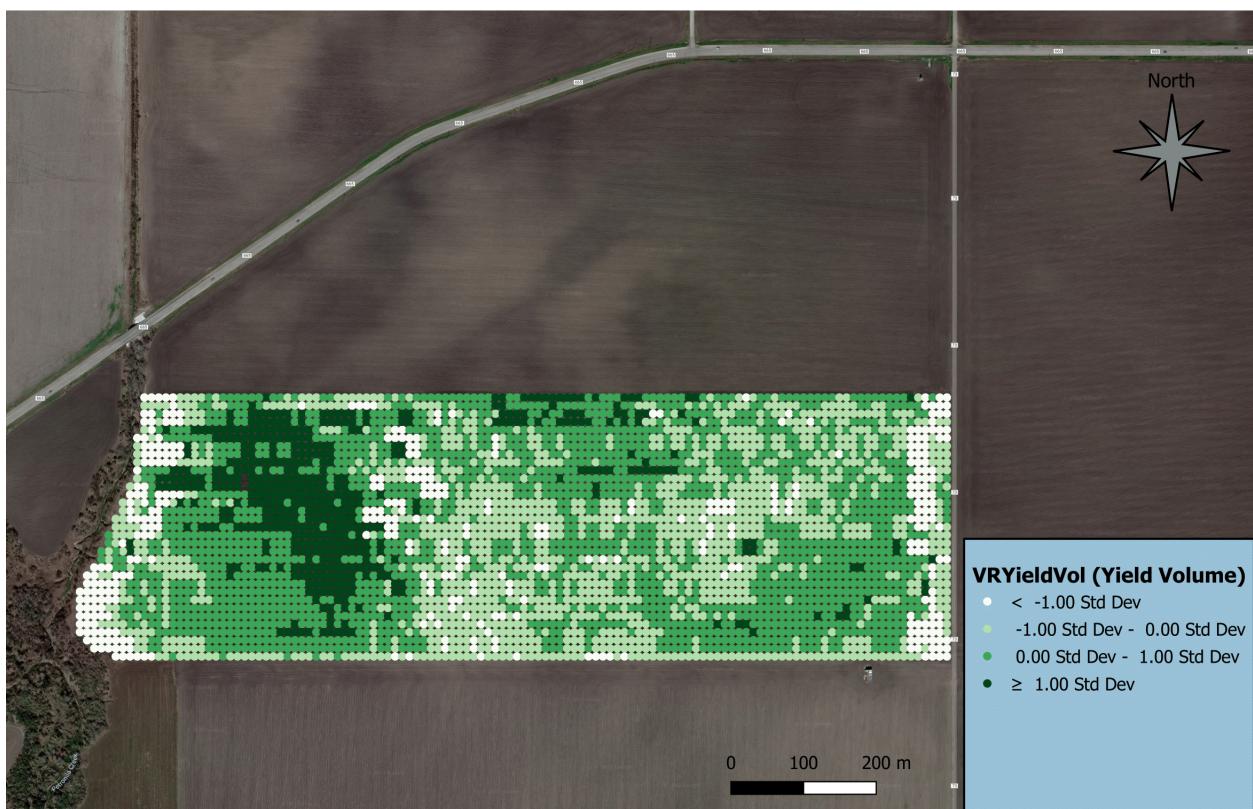


Figure 2: Each dot represents a 10ft x 10ft cell in the given field

Table 7: Grid 10

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
GridId*	1	183	3.00	0.00	3.00	3.00	0.00	3.00	3.00	0.00	NaN	NaN	0.00
VRYieldVOL	2	183	75.31	28.85	83.84	80.67	18.30	0.00	112.51	112.51	-	1.47	2.13
row	3	183	76.17	2.58	76.00	76.22	2.97	72.00	80.00	8.00	-	-1.24	0.19
col	4	183	13.28	5.96	13.00	13.33	7.41	2.00	23.00	21.00	-	-1.16	0.44
Relative_Elevation15	5	183	-	0.95	-1.35	-1.48	0.98	-	-0.44	3.36	-	-0.96	0.07
				1.57				3.80			0.61		
Slope1	6	183	0.01	0.01	0.01	0.01	0.01	0.00	0.05	0.05	1.14	1.06	0.00
TRI1	7	183	0.73	0.39	0.78	0.72	0.48	0.14	1.59	1.44	0.10	-1.05	0.03
TPI1	8	183	-	1.08	0.14	-0.04	0.97	-	1.14	4.24	-	-0.03	0.08
				0.18				3.11			0.96		
Elevation1	9	183	52.71	1.31	53.01	52.83	1.36	49.60	54.27	4.67	-	-0.96	0.10
										0.61			
Application_7_N_1@te	10	183	0.27	0.17	0.24	0.24	0.01	0.21	2.26	2.05	8.97	95.93	0.01
Application_10_N1@ratd	11	183	85.39	28.76	79.84	79.72	1.93	72.54	282.87	210.34	5.17	27.10	2.13
ph_mean_30_60	12	183	7.53	0.07	7.56	7.54	0.02	7.39	7.58	0.19	-	-0.09	0.01
										1.36			
clay_mean_30_60	13	183	40.97	1.80	41.58	41.22	1.12	37.45	42.33	4.88	-	-0.19	0.13
										1.28			
silt_mean_30_60	14	183	23.73	1.20	24.22	23.91	0.50	21.34	24.55	3.21	-	-0.12	0.09
										1.34			
sand_mean_30_60	15	183	31.49	2.61	30.43	31.11	1.14	29.66	36.52	6.86	1.32	-0.14	0.19
ksat_mean_30_60	16	183	1.18	0.73	0.91	1.07	0.36	0.67	2.68	2.00	1.32	-0.13	0.05
om_mean_30_60	17	183	1.19	0.03	1.20	1.20	0.01	1.12	1.21	0.09	-	-0.13	0.00
										1.34			

Table 8: Grid 11

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
GridId*	1	132	4.00	0.00	4.00	4.00	0.00	4.00	4.00	0.00	NaN	NaN	0.00
VRYieldVOL	2	132	72.61	25.12	80.77	77.66	11.99	0.00	101.47	101.47	-	2.15	2.19
										1.74			
row	3	132	83.50	1.71	83.50	83.50	2.22	81.00	86.00	5.00	0.00	-1.29	0.15
col	4	132	12.50	6.37	12.50	12.50	8.15	2.00	23.00	21.00	0.00	-1.23	0.55
Relative_Elevation15	5	132	-	0.46	-1.02	-1.08	0.45	-	-0.53	1.76	-	-0.18	0.04
								2.29			0.84		
Slope1	6	132	0.01	0.01	0.01	0.01	0.01	0.00	0.03	0.03	1.35	1.86	0.00
TRI1	7	132	0.46	0.24	0.46	0.45	0.31	0.12	1.10	0.98	0.46	-0.63	0.02
TPI1	8	132	0.41	0.28	0.43	0.44	0.25	-	0.86	1.54	-	1.22	0.02
								0.69			0.98		
Elevation1	9	132	53.31	0.64	53.46	53.38	0.62	51.70	54.15	2.45	-	-0.18	0.06
										0.84			
Application_7_N_1@te	10	132	0.27	0.10	0.24	0.24	0.01	0.15	0.70	0.55	3.58	12.14	0.01
Application_10_N1@ratd	11	132	81.59	22.53	78.14	78.39	3.55	38.91	233.37	194.46	5.26	31.12	1.96
ph_mean_30_60	12	132	7.51	0.08	7.56	7.52	0.01	7.39	7.57	0.18	-	-0.99	0.01
										1.00			

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
clay_mean_30_60	13	132	40.58	1.93	41.72	40.82	0.10	37.45	41.79	4.34	-	-0.99	0.17
											1.01		
silt_mean_30_60	14	132	23.45	1.30	24.22	23.61	0.09	21.34	24.28	2.95	-	-0.99	0.11
											1.01		
sand_mean_30_60	15	132	32.22	2.64	30.69	31.90	0.22	30.54	36.52	5.98	1.01	-0.99	0.23
ksat_mean_30_60	16	132	1.28	0.86	0.78	1.18	0.10	0.72	2.68	1.96	1.01	-0.99	0.07
om_mean_30_60	17	132	1.18	0.03	1.20	1.18	0.01	1.13	1.20	0.08	-	-1.01	0.00
											0.97		

Table 9: Grid 19

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
GridId*	1	150	5.00	0.00	5.00	5.00	0.00	5.00	5.00	0.00	NaN	NaN	0.00
VRYieldVOL	2	150	92.47	28.06	99.94	98.06	13.93	0.00	126.38	126.38	-	4.83	2.29
											2.24		
row	3	150	60.50	1.71	60.50	60.50	2.22	58.00	63.00	5.00	0.00	-1.29	0.14
col	4	150	36.00	7.24	36.00	36.00	8.90	24.00	48.00	24.00	0.00	-1.23	0.59
Relative_Elevation	15	150	-	0.53	-0.95	-1.03	0.47	-	0.07	2.59	-	0.06	0.04
											0.72		
Slope1	6	150	0.01	0.00	0.01	0.01	0.00	0.00	0.02	0.02	1.06	1.62	0.00
TRI1	7	150	0.37	0.14	0.33	0.36	0.14	0.19	0.75	0.56	0.71	-0.46	0.01
TPI1	8	150	-	0.40	-0.42	-0.44	0.32	-	0.71	2.50	-	1.72	0.03
											0.73		
Elevation1	9	150	53.40	0.74	53.56	53.46	0.66	51.38	54.99	3.61	-	0.06	0.06
											0.72		
Application_7_N_10	16	150	0.24	0.01	0.24	0.24	0.01	0.23	0.26	0.03	0.87	0.27	0.00
Application_10_N11	17	150	80.34	1.59	80.55	80.38	1.48	76.46	83.97	7.50	-	-0.30	0.13
											0.23		
ph_mean_30_60	12	150	7.57	0.04	7.56	7.57	0.02	7.54	7.66	0.12	1.17	0.30	0.00
clay_mean_30_60	13	150	43.40	0.79	43.66	43.40	0.96	41.67	44.97	3.30	-	0.08	0.06
											0.02		
silt_mean_30_60	14	150	24.80	0.47	24.92	24.80	0.36	23.74	25.75	2.02	0.05	0.40	0.04
sand_mean_30_60	15	150	28.43	1.17	28.08	28.43	1.22	26.05	31.17	5.12	0.07	0.49	0.10
ksat_mean_30_60	16	150	0.52	0.08	0.54	0.53	0.07	0.31	0.62	0.31	-	0.63	0.01
											1.02		
om_mean_30_60	17	150	1.21	0.02	1.22	1.22	0.02	1.16	1.24	0.08	-	2.55	0.00
											1.30		

Table 10: Grid 20

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
GridId*	1	200	6.00	0.00	6.00	6.00	0.00	6.00	6.00	0.00	NaN	NaN	0.00
VRYieldVOL	2	200	91.59	31.91	101.83	97.50	19.79	0.00	125.50	125.50	-	1.56	2.26
											1.52		
row	3	200	67.50	2.30	67.50	67.50	2.97	64.00	71.00	7.00	0.00	-1.26	0.16
col	4	200	36.00	7.23	36.00	36.00	8.90	24.00	48.00	24.00	0.00	-1.22	0.51
Relative_Elevation	15	200	-	0.30	-0.57	-0.57	0.25	-	0.15	1.78	-	1.31	0.02
											0.66		
Slope1	6	200	0.01	0.00	0.00	0.00	0.00	0.00	0.02	0.02	1.85	3.94	0.00

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
TRI1	7	200	0.28	0.15	0.23	0.25	0.09	0.12	0.88	0.76	1.84	3.34	0.01
TPI1	8	200	0.21	0.31	0.26	0.22	0.36	-	0.81	1.42	-	-0.81	0.02
Elevation1	9	200	54.07	0.41	54.10	54.09	0.35	52.61	55.09	2.47	-	1.31	0.03
Application_7_N_1@te	200	0.24	0.00	0.24	0.24	0.00	0.23	0.25	0.02	0.42	0.73	0.00	
Application_10_N1@rate	200	79.78	1.99	79.81	79.73	2.20	74.56	85.30	10.74	0.16	-0.14	0.14	
ph_mean_30_60	12	200	7.60	0.04	7.60	7.60	0.06	7.55	7.66	0.11	0.43	-1.11	0.00
clay_mean_30_60	13	200	43.71	0.94	43.76	43.69	1.21	42.33	45.34	3.01	0.33	-1.04	0.07
silt_mean_30_60	14	200	25.04	0.51	24.95	25.02	0.42	24.33	25.94	1.61	0.63	-0.87	0.04
sand_mean_30_60	15	200	27.86	1.34	27.95	27.92	1.41	25.47	29.74	4.27	-	-0.94	0.09
ksat_mean_30_60	16	200	0.50	0.12	0.49	0.51	0.19	0.25	0.67	0.42	-	-0.86	0.01
om_mean_30_60	17	200	1.22	0.02	1.22	1.22	0.00	1.20	1.26	0.06	1.18	0.11	0.00

Table 11: Grid 21

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
GridId*	1	225	7.00	0.00	7.00	7.00	0.00	7.00	7.00	0.00	NaN	NaN	0.00
VRYieldVOL	2	225	97.77	17.19	101.58	100.02	12.33	0.00	125.22	125.22	-	6.02	1.15
											1.92		
row	3	225	76.00	2.59	76.00	76.00	2.97	72.00	80.00	8.00	0.00	-1.25	0.17
col	4	225	36.00	7.23	36.00	36.00	8.90	24.00	48.00	24.00	0.00	-1.22	0.48
Relative_Elevation	15	225	-	0.20	-0.47	-0.46	0.21	-	0.13	0.89	0.69	-0.18	0.01
								0.76					
Slope1	6	225	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.83	0.55	0.00
TRI1	7	225	0.19	0.06	0.17	0.18	0.06	0.10	0.39	0.29	0.86	0.18	0.00
TPI1	8	225	0.03	0.55	-0.01	0.02	0.51	-	2.05	3.55	0.31	1.14	0.04
							1.51						
Elevation1	9	225	54.26	0.28	54.23	54.24	0.29	53.83	55.07	1.24	0.69	-0.18	0.02
Application_7_N_1@te	225	0.24	0.00	0.24	0.24	0.00	0.23	0.25	0.02	-	-	0.52	0.00
								0.23					
Application_10_N1@rate	225	80.19	1.64	80.10	80.12	1.46	75.97	86.45	10.48	0.62	0.62	1.23	0.11
ph_mean_30_60	12	225	7.62	0.03	7.65	7.63	0.01	7.57	7.68	0.11	-	-1.39	0.00
								0.26					
clay_mean_30_60	13	225	44.37	1.03	44.44	44.49	1.12	41.72	46.30	4.59	-	-0.29	0.07
silt_mean_30_60	14	225	25.46	0.51	25.59	25.51	0.52	24.28	26.12	1.84	-	-1.08	0.03
								0.49					
sand_mean_30_60	15	225	26.83	1.47	26.67	26.66	1.58	24.33	30.69	6.37	0.69	-0.46	0.10
ksat_mean_30_60	16	225	0.43	0.12	0.39	0.43	0.17	0.25	0.72	0.47	0.44	-0.29	0.01
om_mean_30_60	17	225	1.24	0.02	1.25	1.24	0.01	1.20	1.26	0.07	-	-1.65	0.00
								0.16					

Table 12: Grid 22

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
GridId*	1	150	8.00	0.00	8.00	8.00	0.00	8.00	8.00	0.00	NaN	NaN	0.00

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
VRYieldVOL	2	150	90.52	10.94	91.31	91.19	9.54	46.86	112.99	66.13	-	1.08	0.89
row	3	150	83.50	1.71	83.50	83.50	2.22	81.00	86.00	5.00	0.00	-1.29	0.14
col	4	150	36.00	7.24	36.00	36.00	8.90	24.00	48.00	24.00	0.00	-1.23	0.59
Relative_Elevation15	150	-	0.22	-0.54	-0.51	0.14	-	0.39	1.18	1.44	1.85	0.02	
Slope1	6	150	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	1.06	1.35	0.00
TRI1	7	150	0.21	0.09	0.18	0.20	0.07	0.10	0.50	0.40	1.00	0.11	0.01
TPI1	8	150	-	0.63	-0.55	-0.45	0.56	-	1.63	3.50	0.41	0.27	0.05
Elevation1	9	150	54.22	0.31	54.13	54.18	0.20	53.78	55.42	1.64	1.44	1.85	0.03
Application_7_N_1@te	150	0.24	0.01	0.24	0.24	0.01	0.22	0.26	0.04	0.55	0.86	0.00	
Application_10_N_1@te	150	80.02	2.68	79.83	79.99	2.57	73.55	85.61	12.07	0.09	-0.49	0.22	
ph_mean_30_60	12	150	7.63	0.03	7.65	7.63	0.05	7.57	7.68	0.11	-	-0.69	0.00
clay_mean_30_60	13	150	44.33	1.21	44.44	44.41	0.32	41.72	46.30	4.59	-	0.76	0.10
silt_mean_30_60	14	150	25.43	0.50	25.59	25.48	0.36	24.28	26.12	1.84	-	0.97	0.04
sand_mean_30_60	15	150	26.83	1.68	26.67	26.66	0.21	24.33	30.69	6.37	0.93	1.15	0.14
ksat_mean_30_60	16	150	0.41	0.13	0.39	0.39	0.03	0.25	0.72	0.47	1.37	1.64	0.01
om_mean_30_60	17	150	1.23	0.02	1.25	1.24	0.01	1.20	1.26	0.06	-	-1.17	0.00
											0.44		

Table 13: Grid 30

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
GridId*	1	144	9.00	0.00	9.00	9.00	0.00	9.00	9.00	0.00	NaN	NaN	0.00
VRYieldVOL	2	144	85.47	16.80	87.31	85.99	14.20	25.94	120.70	94.76	-	1.52	1.40
row	3	144	60.50	1.71	60.50	60.50	2.22	58.00	63.00	5.00	0.00	-1.29	0.14
col	4	144	60.50	6.95	60.50	60.50	8.90	49.00	72.00	23.00	0.00	-1.23	0.58
Relative_Elevation15	144	0.16	0.35	0.21	0.17	0.37	-	0.82	1.42	-	-	-0.85	0.03
Slope1	6	144	0.01	0.00	0.01	0.01	0.01	0.00	0.02	0.02	0.36	-0.42	0.00
TRI1	7	144	0.49	0.15	0.48	0.48	0.14	0.26	1.04	0.79	0.91	0.65	0.01
TPI1	8	144	-	0.71	-0.49	-0.37	0.89	-	0.85	2.94	0.04	-1.20	0.06
Elevation1	9	144	55.10	0.48	55.18	55.12	0.51	54.06	56.03	1.97	-	-0.85	0.04
Application_7_N_1@te	144	0.24	0.01	0.24	0.24	0.01	0.23	0.26	0.03	0.85	-0.31	0.00	
Application_10_N_1@te	144	80.46	1.60	80.39	80.40	1.52	77.01	84.69	7.68	0.29	-0.01	0.13	
ph_mean_30_60	12	144	7.63	0.02	7.63	7.63	0.02	7.59	7.66	0.07	0.08	-0.54	0.00
clay_mean_30_60	13	144	44.47	0.30	44.35	44.46	0.42	43.87	44.97	1.11	-	-0.98	0.02
silt_mean_30_60	14	144	25.61	0.09	25.60	25.61	0.10	25.45	25.75	0.30	-	-1.04	0.01
sand_mean_30_60	15	144	26.65	0.35	26.79	26.66	0.51	26.05	27.37	1.32	-	-1.09	0.03
ksat_mean_30_60	16	144	0.34	0.03	0.34	0.33	0.04	0.30	0.41	0.10	0.89	-0.01	0.00

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
om_mean_30_60	17	144	1.22	0.01	1.22	1.22	0.01	1.20	1.24	0.04	0.18	0.59	0.00

Table 14: Grid 31

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
GridId*	1	192	10.00	0.00	10.00	10.00	0.00	10.00	10.00	0.00	NaN	NaN	0.00
VRYieldVOL	2	192	74.54	20.05	76.48	76.48	15.87	0.00	109.86	109.86	-	4.75	1.45
row	3	192	67.50	2.30	67.50	67.50	2.97	64.00	71.00	7.00	0.00	-1.26	0.17
col	4	192	60.50	6.94	60.50	60.50	8.90	49.00	72.00	23.00	0.00	-1.22	0.50
Relative_Elevation15	192	0.47	0.28	0.53	0.48	0.29	-	0.91	1.10	-	-	-0.95	0.02
Slope1	6	192	0.01	0.01	0.00	0.01	0.00	0.00	0.03	0.03	1.53	1.90	0.00
TRI1	7	192	0.32	0.14	0.26	0.30	0.12	0.12	0.70	0.58	0.76	-0.60	0.01
TPI1	8	192	0.28	0.46	0.28	0.30	0.36	-	1.37	2.57	-	0.64	0.03
Elevation1	9	192	55.53	0.39	55.63	55.55	0.40	54.62	56.15	1.53	-	-0.95	0.03
Application_7_N_10ate	192	0.24	0.00	0.24	0.24	0.00	0.23	0.24	0.01	-	-	-0.10	0.00
Application_10_N1fratd	92	79.60	1.94	79.61	79.64	1.88	71.93	84.75	12.82	-	0.06	1.76	0.14
ph_mean_30_60	12	192	7.64	0.02	7.64	7.64	0.02	7.59	7.66	0.07	-	-0.14	0.00
clay_mean_30_60	13	192	44.52	0.37	44.39	44.52	0.37	43.87	45.34	1.47	0.20	-0.16	0.03
silt_mean_30_60	14	192	25.62	0.11	25.64	25.61	0.16	25.37	25.90	0.53	0.52	-0.06	0.01
sand_mean_30_60	15	192	26.57	0.45	26.66	26.58	0.31	25.47	27.37	1.90	-	0.26	0.03
ksat_mean_30_60	16	192	0.31	0.02	0.31	0.31	0.01	0.25	0.34	0.09	-	4.18	0.00
om_mean_30_60	17	192	1.23	0.01	1.23	1.23	0.01	1.21	1.26	0.05	1.28	1.25	0.00

Table 15: Grid 32

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
GridId*	1	216	11.00	0.00	11.00	11.00	0.00	11.00	11.00	0.00	NaN	NaN	0.00
VRYieldVOL	2	216	69.81	13.74	71.67	70.29	14.63	30.18	97.38	67.21	-	-0.32	0.93
row	3	216	76.00	2.59	76.00	76.00	2.97	72.00	80.00	8.00	0.00	-1.25	0.18
col	4	216	60.50	6.94	60.50	60.50	8.90	49.00	72.00	23.00	0.00	-1.22	0.47
Relative_Elevation15	216	0.66	0.25	0.71	0.69	0.25	-	1.01	1.09	-	-	-0.19	0.02
Slope1	6	216	0.01	0.00	0.01	0.01	0.00	0.00	0.01	0.01	0.26	-0.67	0.00
TRI1	7	216	0.25	0.07	0.25	0.25	0.08	0.11	0.43	0.32	0.27	-0.57	0.00
TPI1	8	216	0.37	0.58	0.39	0.41	0.55	-	1.88	3.49	-	0.61	0.04
Elevation1	9	216	55.81	0.34	55.88	55.84	0.35	54.78	56.29	1.51	-	-0.19	0.02
om_mean_30_60	17	216	1.23	0.01	1.23	1.23	0.01	1.21	1.26	0.05	1.28	1.25	0.00

	vars	n	mean	sd	median	trimmedmad	min	max	range	skew	kurtosis	se
Application_7_N_10	te	216	0.24	0.00	0.24	0.24	0.00	0.23	0.25	0.02	-	0.40
											0.12	
Application_10_N_11	rate	216	80.14	1.70	80.03	80.09	1.80	76.12	84.37	8.26	0.25	-0.51
ph_mean_30_60	12	216	7.63	0.02	7.64	7.63	0.02	7.60	7.68	0.08	-	-1.34
											0.06	
clay_mean_30_60	13	216	44.44	0.58	44.39	44.36	0.71	43.91	46.30	2.40	1.20	0.83
silt_mean_30_60	14	216	25.64	0.18	25.66	25.62	0.29	25.37	26.12	0.75	0.67	-0.27
sand_mean_30_60	15	216	26.62	0.74	26.66	26.71	0.94	24.33	27.32	3.00	-	0.56
											1.11	
ksat_mean_30_60	16	216	0.29	0.03	0.29	0.30	0.05	0.25	0.35	0.10	-	-1.01
											0.20	
om_mean_30_60	17	216	1.24	0.02	1.24	1.24	0.03	1.21	1.26	0.05	0.25	-1.19
											0.00	

Table 16: Grid 33

	vars	n	mean	sd	median	trimmedmad	min	max	range	skew	kurtosis	se
GridId*	1	144	12.00	0.00	12.00	12.00	0.00	12.00	12.00	0.00	NaN	NaN
VRYieldVOL	2	144	66.53	10.95	66.89	66.95	11.17	35.92	89.86	53.94	-	-0.31
											0.30	
row	3	144	83.50	1.71	83.50	83.50	2.22	81.00	86.00	5.00	0.00	-1.29
col	4	144	60.50	6.95	60.50	60.50	8.90	49.00	72.00	23.00	0.00	-1.23
Relative_Elevation	15	144	0.72	0.21	0.78	0.74	0.14	0.05	1.06	1.01	-	0.60
											1.10	
Slope1	6	144	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.64	0.00
TRI1	7	144	0.22	0.08	0.19	0.22	0.10	0.10	0.48	0.37	0.50	-0.84
TPI1	8	144	0.38	0.57	0.36	0.39	0.51	-	2.15	3.60	-	1.08
											0.14	
Elevation1	9	144	55.88	0.29	55.97	55.92	0.19	54.96	56.36	1.40	-	0.60
											1.10	
Application_7_N_10	te	144	0.24	0.01	0.24	0.24	0.00	0.22	0.27	0.05	1.59	4.32
Application_10_N_11	rate	144	79.53	4.13	79.70	79.58	2.65	69.45	113.03	43.59	3.33	28.23
ph_mean_30_60	12	144	7.64	0.02	7.64	7.64	0.03	7.62	7.68	0.06	0.73	-0.82
clay_mean_30_60	13	144	44.83	0.76	44.44	44.69	0.02	44.43	46.30	1.88	1.42	0.02
silt_mean_30_60	14	144	25.72	0.21	25.64	25.69	0.08	25.59	26.12	0.53	1.38	-0.04
sand_mean_30_60	15	144	26.14	0.94	26.59	26.30	0.08	24.33	26.65	2.32	-	0.02
											1.42	
ksat_mean_30_60	16	144	0.30	0.04	0.29	0.31	0.07	0.25	0.35	0.10	-	-1.36
											0.08	
om_mean_30_60	17	144	1.22	0.01	1.21	1.22	0.00	1.21	1.25	0.03	1.41	0.01
											0.00	

Table 17: Grid 41

	vars	n	mean	sd	median	trimmedmad	min	max	range	skew	kurtosis	se
GridId*	1	144	13.00	0.00	13.00	13.00	0.00	13.00	13.00	0.00	NaN	NaN
VRYieldVOL	2	144	81.55	13.54	82.79	81.34	12.37	35.34	119.02	83.68	-	0.64
											0.08	
row	3	144	60.50	1.71	60.50	60.50	2.22	58.00	63.00	5.00	0.00	-1.29
col	4	144	84.50	6.95	84.50	84.50	8.90	73.00	96.00	23.00	0.00	-1.23
											0.58	

	vars	n	mean	sd	median	trimmedmad	min	max	range	skew	kurtosis	se
Relative_Elevation15	144	0.68	0.21	0.66	0.69	0.19	0.22	1.11	0.90	-	-0.59	0.02
Slope1	6	144	0.01	0.00	0.01	0.01	0.00	0.02	0.01	0.04	-0.64	0.00
TRI1	7	144	0.37	0.09	0.36	0.37	0.10	0.22	0.64	0.42	0.46	-0.18
TPI1	8	144	-	0.89	-0.28	-0.34	1.08	-	1.20	3.30	-	-1.11
Elevation1	9	144	55.84	0.29	55.81	55.85	0.27	55.19	56.43	1.25	-	-0.59
Application_7_N_10ate	144	0.24	0.01	0.24	0.24	0.00	0.23	0.26	0.03	1.00	1.13	0.00
Application_10_N1rate	144	80.56	1.38	80.42	80.51	1.20	77.50	85.47	7.97	0.43	0.64	0.11
ph_mean_30_60	12	144	7.59	0.00	7.59	7.59	0.00	7.59	7.60	0.01	-	-1.72
clay_mean_30_60	13	144	43.96	0.13	43.94	43.93	0.07	43.87	44.27	0.40	1.80	1.98
silt_mean_30_60	14	144	25.53	0.12	25.48	25.50	0.05	25.45	25.84	0.39	1.85	1.92
sand_mean_30_60	15	144	27.22	0.24	27.29	27.28	0.12	26.60	27.37	0.77	-	2.33
ksat_mean_30_60	16	144	0.33	0.02	0.34	0.34	0.00	0.29	0.34	0.05	-	2.85
om_mean_30_60	17	144	1.21	0.01	1.22	1.21	0.01	1.20	1.23	0.04	-	-1.34
										0.07		

Table 18: Grid 42

	vars	n	mean	sd	median	trimmedmad	min	max	range	skew	kurtosis	se
GridId*	1	192	14.00	0.00	14.00	14.00	0.00	14.00	14.00	0.00	NaN	NaN
VRYieldVOL	2	192	75.18	15.95	73.56	74.91	14.33	38.94	122.87	83.93	0.20	-0.04
row	3	192	67.50	2.30	67.50	67.50	2.97	64.00	71.00	7.00	0.00	-1.26
col	4	192	84.50	6.94	84.50	84.50	8.90	73.00	96.00	23.00	0.00	-1.22
Relative_Elevation15	192	0.83	0.16	0.83	0.83	0.11	0.33	1.47	1.15	0.56	2.53	0.01
Slope1	6	192	0.01	0.00	0.00	0.01	0.00	0.00	0.03	0.03	1.60	2.97
TRI1	7	192	0.27	0.13	0.25	0.25	0.11	0.10	0.94	0.84	1.76	4.55
TPI1	8	192	0.42	0.81	0.40	0.40	0.58	-	3.65	5.45	0.49	1.94
Elevation1	9	192	56.04	0.22	56.04	56.04	0.16	55.34	56.93	1.59	0.56	2.53
Application_7_N_10ate	192	0.24	0.00	0.24	0.24	0.00	0.23	0.24	0.01	0.03	-0.39	0.00
Application_10_N1rate	192	79.46	2.98	79.19	79.15	1.99	74.41	96.35	21.94	2.84	12.77	0.21
ph_mean_30_60	12	192	7.59	0.01	7.59	7.59	0.00	7.59	7.63	0.05	1.99	3.85
clay_mean_30_60	13	192	44.05	0.26	43.87	44.01	0.00	43.87	44.77	0.90	1.31	0.94
silt_mean_30_60	14	192	25.54	0.15	25.48	25.52	0.00	25.37	25.84	0.47	1.15	-0.12
sand_mean_30_60	15	192	27.12	0.36	27.37	27.18	0.00	26.35	27.37	1.02	-	-0.66
ksat_mean_30_60	16	192	0.32	0.02	0.34	0.32	0.00	0.29	0.34	0.05	-	-1.26
om_mean_30_60	17	192	1.22	0.01	1.22	1.22	0.00	1.20	1.23	0.03	0.15	-0.45
										1.01		
										0.67		

Table 19: Grid 43

	vars	n	mean	sd	median	trimmedmad	min	max	range	skew	kurtosis	se
GridId*	1	216	15.00	0.00	15.00	15.00	0.00	15.00	15.00	0.00	NaN	NaN

	vars	n	mean	sd	median	trimmedmad	min	max	range	skew	kurtosis	se
VRYieldVOL	2	216	73.74	12.13	74.90	73.83	10.95	44.72	107.36	62.65	-	0.13
										0.05		0.83
row	3	216	76.00	2.59	76.00	76.00	2.97	72.00	80.00	8.00	0.00	-1.25
col	4	216	84.50	6.94	84.50	84.50	8.90	73.00	96.00	23.00	0.00	-1.22
Relative_Elevation15	216	0.75	0.15	0.74	0.75	0.17	0.44	1.09	0.65	0.12	-0.82	0.01
Slope1	6	216	0.01	0.00	0.01	0.01	0.00	0.00	0.01	0.01	0.41	-0.74
TRI1	7	216	0.23	0.08	0.21	0.22	0.08	0.10	0.47	0.37	0.66	-0.17
TPI1	8	216	-	0.98	-0.16	-0.11	1.16	-	2.02	4.71	0.06	-0.71
							0.10		2.69			0.07
Elevation1	9	216	55.93	0.21	55.92	55.93	0.24	55.50	56.41	0.91	0.12	-0.82
Application_7_N_10ate	216	0.24	0.00	0.24	0.24	0.00	0.23	0.25	0.02	-	0.02	0.00
									0.18			
Application_10_N_11rate	216	80.20	1.70	80.07	80.13	1.71	76.55	85.44	8.89	0.38	-0.02	0.12
ph_mean_30_60_12	216	7.61	0.01	7.60	7.61	0.00	7.60	7.63	0.03	0.87	-1.15	0.00
clay_mean_30_60_13	216	44.34	0.32	44.30	44.32	0.51	43.96	44.97	1.01	0.39	-1.05	0.02
silt_mean_30_60_14	216	25.46	0.11	25.39	25.45	0.04	25.34	25.65	0.31	0.82	-1.25	0.01
sand_mean_30_60_15	216	26.91	0.40	27.02	26.94	0.45	26.10	27.32	1.23	-	-1.02	0.03
									0.65			
ksat_mean_30_60_16	216	0.31	0.02	0.31	0.31	0.01	0.29	0.43	0.14	4.40	21.51	0.00
om_mean_30_60_17	216	1.21	0.00	1.21	1.21	0.00	1.20	1.22	0.02	1.05	0.81	0.00

Table 20: Grid 44

	vars	n	mean	sd	median	trimmedmad	min	max	range	skew	kurtosis	se
GridId*	1	144	16.00	0.00	16.00	16.00	0.00	16.00	16.00	0.00	NaN	NaN
VRYieldVOL	2	144	81.12	9.39	80.34	80.87	9.34	61.40	105.78	44.38	0.27	-0.28
row	3	144	83.50	1.71	83.50	83.50	2.22	81.00	86.00	5.00	0.00	-1.29
col	4	144	84.50	6.95	84.50	84.50	8.90	73.00	96.00	23.00	0.00	-1.23
Relative_Elevation15	144	0.69	0.13	0.68	0.69	0.13	0.32	1.00	0.68	0.17	-0.14	0.01
Slope1	6	144	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.95	1.59	0.00
TRI1	7	144	0.15	0.04	0.15	0.15	0.04	0.09	0.36	0.27	1.37	3.26
TPI1	8	144	-	0.76	-0.24	-0.24	0.76	-	1.86	4.49	0.13	0.41
				0.23			2.62					0.06
Elevation1	9	144	55.84	0.18	55.83	55.84	0.18	55.33	56.28	0.94	0.17	-0.14
Application_7_N_10ate	144	0.24	0.01	0.24	0.24	0.01	0.22	0.27	0.05	1.63	4.21	0.00
Application_10_N_11rate	144	79.03	2.89	79.57	79.22	2.52	72.16	84.11	11.95	-	-0.34	0.24
									0.64			
ph_mean_30_60_12	144	7.62	0.01	7.62	7.62	0.02	7.61	7.63	0.03	-	-1.48	0.00
									0.21			
clay_mean_30_60_13	144	44.61	0.31	44.43	44.61	0.23	44.27	44.97	0.70	0.22	-1.82	0.03
silt_mean_30_60_14	144	25.55	0.13	25.59	25.57	0.09	25.34	25.65	0.31	-	-0.81	0.01
									1.00			
sand_mean_30_60_15	144	26.53	0.41	26.65	26.51	0.67	26.10	27.10	1.00	0.19	-1.49	0.03
ksat_mean_30_60_16	144	0.33	0.06	0.31	0.33	0.03	0.29	0.43	0.14	1.07	-0.75	0.00
om_mean_30_60_17	144	1.21	0.01	1.21	1.21	0.01	1.20	1.22	0.02	-	-1.09	0.00
									0.66			

Table 21: Grid 52

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
GridId*	1	138	17.00	0.00	17.00	17.00	0.00	17.00	17.00	0.00	NaN	NaN	0.00
VRYieldVOL	2	138	65.76	23.77	72.87	70.31	11.03	0.00	96.64	96.64	-	2.26	2.02
											1.77		
row	3	138	60.50	1.71	60.50	60.50	2.22	58.00	63.00	5.00	0.00	-1.29	0.15
col	4	138	108.00	6.66	108.00	108.00	8.90	97.00	119.00	22.00	0.00	-1.23	0.57
Relative_Elevation15	138	0.96	0.22	0.95	0.96	0.28	0.49	1.35	0.85	-	-	-1.11	0.02
											0.04		
Slope1	6	138	0.01	0.00	0.01	0.01	0.00	0.00	0.02	0.02	0.02	-0.86	0.00
TRI1	7	138	0.38	0.09	0.38	0.38	0.12	0.20	0.61	0.42	0.01	-0.93	0.01
TPI1	8	138	0.56	1.12	0.51	0.56	1.44	-	2.52	4.16	-	-1.19	0.09
							1.64				0.01		
Elevation1	9	138	56.22	0.30	56.20	56.22	0.39	55.57	56.76	1.19	-	-1.11	0.03
								55.57			0.04		
Application_7_N_10ate	138	0.24	0.02	0.24	0.24	0.01	0.19	0.34	0.14	1.36	8.21	0.00	
Application_10_N1ratd	138	81.38	8.52	80.37	80.32	2.18	68.70	136.82	68.12	4.05	19.55	0.73	
ph_mean_30_60_12	138	7.63	0.02	7.64	7.64	0.02	7.59	7.66	0.07	-	-0.78	0.00	
								7.59			0.38		
clay_mean_30_60_13	138	44.86	0.53	44.94	44.94	0.38	43.32	45.51	2.19	-	1.08	0.05	
								43.32			1.21		
silt_mean_30_60_14	138	25.92	0.21	25.93	25.94	0.05	25.29	26.21	0.92	-	1.70	0.02	
								25.29			0.92		
sand_mean_30_60_15	138	25.91	0.70	25.78	25.82	0.50	25.11	28.02	2.91	1.27	1.52	0.06	
ksat_mean_30_60_16	138	0.29	0.08	0.26	0.27	0.00	0.25	0.62	0.37	3.75	13.63	0.01	
om_mean_30_60_17	138	1.23	0.01	1.23	1.23	0.01	1.22	1.24	0.02	-	-0.37	0.00	
								1.22			0.49		

Table 22: Grid 53

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
GridId*	1	184	18.00	0.00	18.00	18.00	0.00	18.00	18.00	0.00	NaN	NaN	0.00
VRYieldVOL	2	184	66.42	28.68	76.08	70.98	13.50	0.00	101.95	101.95	-	0.90	2.11
											1.45		
row	3	184	67.50	2.30	67.50	67.50	2.97	64.00	71.00	7.00	0.00	-1.26	0.17
col	4	184	108.00	6.65	108.00	108.00	8.90	97.00	119.00	22.00	0.00	-1.22	0.49
Relative_Elevation15	184	0.79	0.14	0.79	0.79	0.15	0.44	1.19	0.75	0.28	0.00	0.01	
Slope1	6	184	0.01	0.00	0.01	0.01	0.00	0.00	0.02	0.02	0.78	0.49	0.00
TRI1	7	184	0.26	0.08	0.24	0.25	0.05	0.11	0.57	0.46	1.21	1.16	0.01
TPI1	8	184	0.17	0.58	0.18	0.18	0.58	-	1.65	3.22	-	-0.04	0.04
								1.65			0.07		
Elevation1	9	184	55.99	0.19	55.98	55.98	0.20	55.50	56.54	1.04	0.28	0.00	0.01
Application_7_N_10ate	184	0.24	0.02	0.24	0.24	0.00	0.19	0.32	0.13	1.40	6.61	0.00	
Application_10_N1ratd	184	79.71	6.02	79.18	79.05	2.14	71.21	136.82	65.61	5.75	45.84	0.44	
ph_mean_30_60_12	184	7.63	0.02	7.62	7.63	0.03	7.59	7.65	0.06	-	-0.83	0.00	
								7.59			0.26		
clay_mean_30_60_13	184	44.75	0.42	44.68	44.75	0.63	44.15	45.51	1.36	0.07	-1.20	0.03	
silt_mean_30_60_14	184	25.75	0.24	25.89	25.79	0.05	25.26	25.93	0.67	-	-0.37	0.02	
								25.26			1.18		
sand_mean_30_60_15	184	26.20	0.60	26.12	26.16	0.71	25.36	27.31	1.95	0.58	-0.84	0.04	
ksat_mean_30_60_16	184	0.29	0.04	0.26	0.29	0.00	0.26	0.36	0.10	0.84	-1.06	0.00	

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
om_mean_30_60 17		184	1.22	0.01	1.22	1.22	0.00	1.19	1.23	0.04	-	-0.42	0.00

Table 23: Grid 54

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
GridId*	1	207	19.00	0.00	19.00	19.00	0.00	19.00	19.00	0.00	NaN	NaN	0.00
VRYieldVOL	2	207	80.93	20.85	85.69	84.34	10.52	0.00	126.33	126.33	-	7.20	1.45
											2.43		
row	3	207	76.00	2.59	76.00	76.00	2.97	72.00	80.00	8.00	0.00	-1.25	0.18
col	4	207	108.00	6.65	108.00	108.00	8.90	97.00	119.00	22.00	0.00	-1.22	0.46
Relative_Elevation15	5	207	0.54	0.14	0.55	0.54	0.13	0.01	0.85	0.84	-	0.40	0.01
											0.20		
Slope1	6	207	0.01	0.00	0.01	0.01	0.00	0.00	0.02	0.02	0.78	0.93	0.00
TRI1	7	207	0.23	0.07	0.21	0.22	0.06	0.10	0.62	0.52	1.23	2.89	0.01
TPI1	8	207	-	0.84	-	-0.55	0.72	-	1.18	5.54	-	1.87	0.06
											0.65		
Elevation1	9	207	55.64	0.20	55.65	55.64	0.17	54.90	56.07	1.17	-	0.40	0.01
											0.20		
Application_7_N_10ate	10	207	0.24	0.02	0.24	0.24	0.01	0.22	0.34	0.12	3.51	15.41	0.00
Application_10_N1lrat	11	207	82.16	12.77	79.58	79.58	1.93	72.10	194.03	121.92	5.40	34.76	0.89
ph_mean_30_60 12	12	207	7.61	0.02	7.61	7.61	0.01	7.56	7.63	0.07	-	1.33	0.00
											1.10		
clay_mean_30_60 13	13	207	44.28	0.35	44.25	44.27	0.15	43.34	45.02	1.68	0.10	1.25	0.02
silt_mean_30_60 14	14	207	25.33	0.13	25.34	25.32	0.12	25.05	25.62	0.58	0.75	1.38	0.01
sand_mean_30_60 15	15	207	27.11	0.47	27.14	27.12	0.25	26.18	28.30	2.12	-	0.75	0.03
											0.24		
ksat_mean_30_60 16	16	207	0.36	0.04	0.35	0.35	0.02	0.31	0.47	0.15	1.68	2.72	0.00
om_mean_30_60 17	17	207	1.19	0.01	1.19	1.19	0.00	1.18	1.21	0.03	0.56	0.70	0.00

Table 24: Grid 55

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
GridId*	1	138	20.00	0.00	20.00	20.00	0.00	20.00	20.00	0.00	NaN	NaN	0.00
VRYieldVOL	2	138	78.11	27.61	87.14	84.18	11.18	0.00	110.84	110.84	-	2.79	2.35
											1.99		
row	3	138	83.50	1.71	83.50	83.50	2.22	81.00	86.00	5.00	0.00	-1.29	0.15
col	4	138	108.00	6.66	108.00	108.00	8.90	97.00	119.00	22.00	0.00	-1.23	0.57
Relative_Elevation15	5	138	0.56	0.08	0.56	0.56	0.10	0.40	0.74	0.34	0.12	-0.83	0.01
Slope1	6	138	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.47	-0.71	0.00
TRI1	7	138	0.16	0.05	0.16	0.16	0.05	0.09	0.31	0.22	0.61	0.31	0.00
TPI1	8	138	0.03	0.79	0.10	0.04	0.86	-	1.73	3.38	-	-0.78	0.07
											1.65		
Elevation1	9	138	55.66	0.11	55.66	55.66	0.13	55.44	55.91	0.48	0.12	-0.83	0.01
Application_7_N_10ate	10	138	0.24	0.01	0.24	0.24	0.01	0.20	0.32	0.11	2.61	10.82	0.00
Application_10_N1lrat	11	138	81.45	8.78	79.75	79.65	2.89	70.43	110.81	40.38	2.59	6.15	0.75
ph_mean_30_60 12	12	138	7.59	0.02	7.61	7.59	0.00	7.56	7.61	0.05	-	-2.00	0.00
											0.09		
clay_mean_30_60 13	13	138	43.71	0.33	43.90	43.69	0.55	43.34	44.27	0.94	0.17	-1.28	0.03

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
silt_mean_30_60	14	138	25.18	0.12	25.28	25.18	0.08	25.05	25.34	0.29	-	-1.90	0.01
											0.04		
sand_mean_30_60	15	138	27.81	0.44	27.54	27.84	0.65	27.10	28.30	1.20	-	-1.41	0.04
											0.13		
ksat_mean_30_60	16	138	0.45	0.02	0.44	0.45	0.00	0.43	0.47	0.03	0.15	-1.92	0.00
om_mean_30_60	17	138	1.19	0.01	1.18	1.18	0.01	1.18	1.20	0.02	1.09	-0.42	0.00