BIRMINGHAM CITY
University

Ecological Analysis of Forest Cover Types, Soil Variations, and Environmental Influences

Author: Rohan Kalu

Date: 10/08/2024

Student ID: 23189640

Course: Data Visualization (CMP 5352)

Page count: 26

Abstract

This report explores the interplay between forest cover types, soil characteristics, and environmental variables in the Rocky Mountains of northern Colorado. By examining key species such as Lodgepole Pine, Spruce/Fir, and Krummholz, we assess the effects of elevation, distance to water sources, and fire risk on their distribution. Our results indicate that elevation has a profound influence on both soil types and tree species density, with the most favorable growth conditions found between 2500 and 3000 meters, attributed to greater sunlight exposure. Additionally, the proximity to water bodies significantly affects the distribution of certain tree species, while soil composition is vital for sustaining a variety of vegetation. We also investigate areas susceptible to fire hazards, underscoring the adaptive features of specific species in environments prone to fire. Ultimately, this report deepens our comprehension of ecological dynamics in this area, offering essential insights for forest management and conservation strategies

Contents

List of Figures

1 Introduction

Data visualization is an essential tool in the field of data analysis, allowing researchers to translate complex datasets into visually interpretable formats. By representing data graphically, visualization helps reveal patterns, trends, and insights that might not be immediately apparent through raw data alone. It enhances understanding and facilitates decision-making by enabling users to grasp significant information quickly and effectively.

1.1 Aim

The aim of the report is to explore and analyze the dataset related to forest cover types in the Rocky Mountains of northern Colorado. It seeks to identify patterns and relationships between environmental variables such as elevation, slope, aspect, distances to hydrology and roadways, and various soil and wilderness area types. By examining these relationships, the report aims to improve the understanding of how these factors influence the distribution of different tree species across various wilderness areas.

1.2 Achievements

The most successful achievement of the report is to identify the distribution of the cover types across the northern mountain of Colorado and the studying the environmental components such as elevation, slope, distance from hydrology and many more that influences the type of tree growing there.

2 Motivation and Objective

I decided to choose the data 'forest' because it provides a comprehensive view of the ecological characteristics of forest ecosystems, including critical attributes such as elevation, slope, soil types, and the distribution of various tree species. This dataset is particularly valuable for analyzing the complex interactions between environmental factors and biodiversity. Understanding these relationships is essential for effective forest management, conservation efforts, and studying the impacts of climate change on tree populations.

2.1 Introduction to Dataset

The dataset is from a study of cover type of Mountains of Colorado, US. The dataset has 581012 rows and 55 columns. The columns are:
• **Elevation:** Elevation in meters above sea level.
• **Aspect:** Aspect in degrees azimuth, indicating the compass direction that a slope faces.
• **Slope:** Slope in degrees, representing the steepness of the terrain.
• **Horizontal_Distance_To_Hydrology:** Horizontal distance to the nearest surface water features.
• **Vertical_Distance_To_Hydrology:** Vertical distance to the nearest surface water features.
• **Horizontal_Distance_To_Roadways:** Horizontal distance to the nearest roadway.
• **Hillshade_9am:** Hillshade index at 9 a.m. during the summer solstice, ranging from 0 to 255.
• **Hillshade_Noon:** Hillshade index at noon during the summer solstice, ranging from 0 to 255.
• **Hillshade_3pm:** Hillshade index at 3 p.m. during the summer solstice, ranging from 0 to 255.
• **Horizontal_Distance_To_Fire_Points:** Horizontal distance to the nearest wildfire ignition points.
• **Wilderness_Area:** Four binary columns indicating the presence (1) or absence (0) of different wilderness areas.
• **Soil_Type:** Forty binary columns indicating the presence (1) or absence (0) of different soil types.
• **Cover_Type:** Forest cover type designation, represented by integers from 1 to 7, indicating different forest types.

2.2 Data cleaning:

Since the data was too large with 55 columns I had to clean it thoroughly.
Missing values:
Although the dataset was too large, there were no missing values.
Feature Engineering:

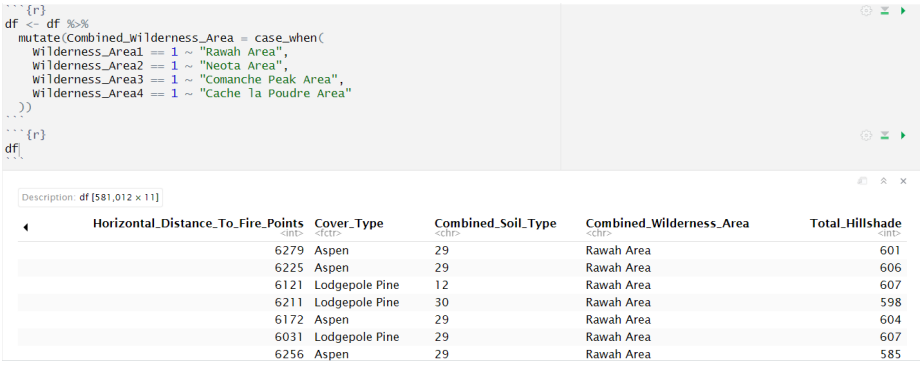
• Check for any null values

```{r}
any(is.na(df))
```

[H]
```
[1] FALSE
```

Figure 1: Checking for Missing values

Since the columns soil type were only binary values, I merged them into a single one in a new column which was named 'Combined_Soil_Type'. Other column such as Wilderness_Area met the same fate. But for Hillshade, a sum was imputed into a new column which dramatically reduced the number of columns.



```{r}
df <- df %>%
  mutate(Combined_Wilderness_Area = case_when(
    Wilderness_Area1 == 1 ~ "Rawah Area",
    Wilderness_Area2 == 1 ~ "Neota Area",
    Wilderness_Area3 == 1 ~ "Comanche Peak Area",
    Wilderness_Area4 == 1 ~ "Cache la Poudre Area"
  ))
```

```{r}
df
```

Description: df [581,012 × 11]

Horizontal_Distance_To_Fire_Points <int>	Cover_Type <fctr>	Combined_Soil_Type <chr>	Combined_Wilderness_Area <chr>	Total_Hillshade <int>
6279	Aspen	29	Rawah Area	601
6225	Aspen	29	Rawah Area	606
6121	Lodgepole Pine	12	Rawah Area	607
6211	Lodgepole Pine	30	Rawah Area	598
6172	Aspen	29	Rawah Area	604
6031	Lodgepole Pine	29	Rawah Area	607
6256	Aspen	29	Rawah Area	585

Figure 2: A glimpse of feature engineering

As for the number of rows, since it was almost impossible to make a proper point visualization, a sample of 3000 data were randomly chosen which showed similar trend to that of the original dataset.

2.3 Questions about the dataset:

Q1. What are the largely present species of plant present in the northern mountains of Colorado, US?
Ans: **Lodgepole Pine** was the species of plant that was in abundance compared to the rest of the plants.

Q2. What are the soil types present and does it affect the type of trees that grow in that area?
***Ans*:** There are 40 types of soils and they do affect the type of trees that grow in the area.

Q3. Does elevation affect the distribution of tree species?
Ans: Yes.

Q4. Does elevation level determine the Wilderness Area?
Ans: Yes.

Q5. Does elevation affect the soil types?
Ans: Yes elevation affects the type of soils found.

Q6. Does the sources of water also determine the types of soil found?
Ans: It is hard to determine but some types of soils are found only near water resources.

Q7. Does water the availability of water resource also determine the cover type?
Ans: Yes, the availability of water determines the type of cover as some plants are found near sources of water.

Q8. Is there any relationship between elevation and sunlight received?
Ans: Yes as some places receive more sunlight throughout the day compared to others.

Q9. Are some areas more prone to fire than others?
Ans: Yes some areas are more prone to fire than others.

Q10. Does roadway influence the cover type in any type of way? ***Ans:*** No

3 Visualization

This section will try to answers all the non trivial questions above through the method of visualization using various graphs and plots to study trends and patterns in the data. By presenting information visually, they enhance our understanding of the dataset and allow for more effective interpretation and decision-making. The plots are made using R language in the RStudio IDE.

3.1 Distribution of Cover types

What are the species of trees present across the vicinity of northern mountains of Colorado US? Do the cardinal direction also influence the cover type?

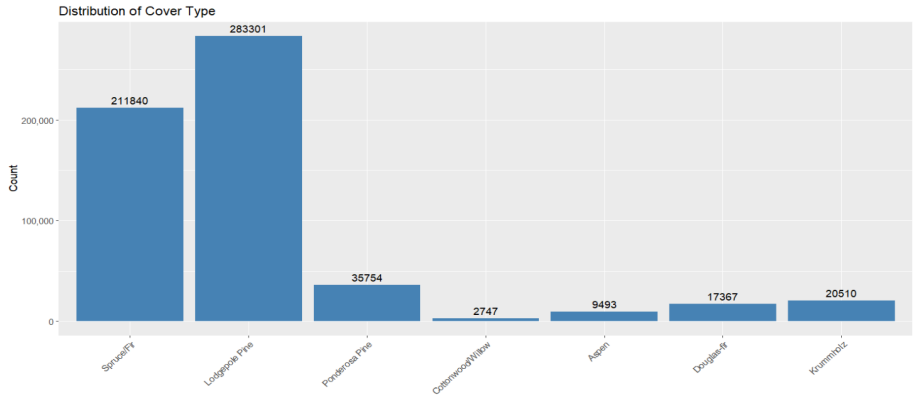


Figure 3: Distribution of Cover Type

Across the northern mountains of Colorado, the most prevalent tree is Lodgepole Pine at 283301 which accounts to almost half of the plant species present in the area. Second most abundant species is the Spruce/Fir tree at 211840. The other species of trees present are Ponderosa Pine, Cottonwood, Aspen, Douglas at 35754, 2747, 9493, 17367 respectively which are nothing compared to the numbers of aforementioned trees.

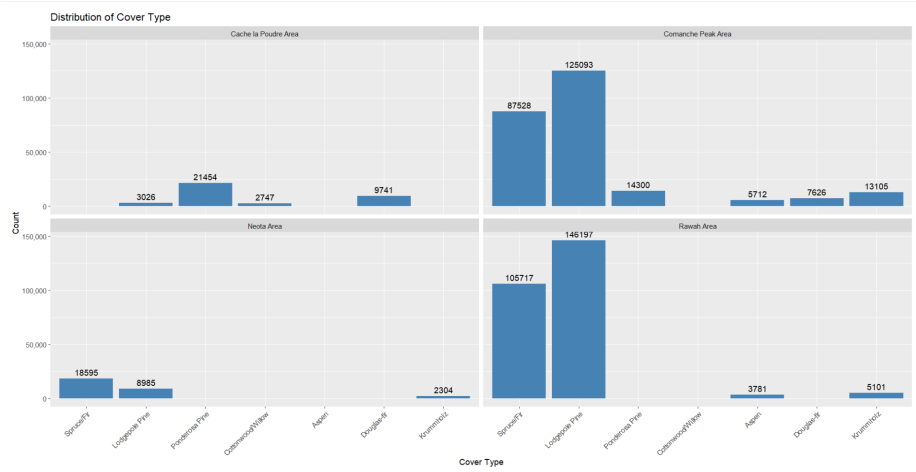


Figure 4: Distribution of Cover Type by Wilderness Area

We can see that Rawah Area hosts the most number of plants but Comanache Peak Area hosts the most species of plants. Comanache Peak has every type of trees except for the Cottonwood tree. On the other hand, Cache la Poudre and Neota Area hosts the least number of plants.

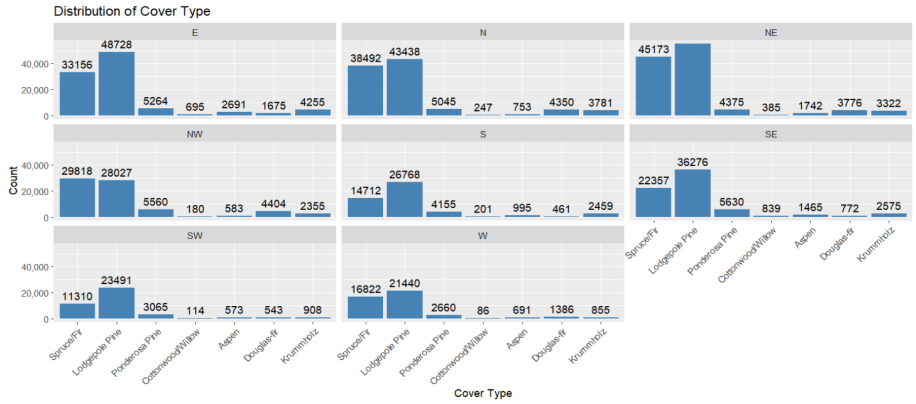


Figure 5: Distribution of Cover type in cardinal directions

The cover types seem to be almost equally spread across every cardinal directions.

3.2 Analysis of Distribution of Soil Type

What are the soil types present in the area of northern mountains of Colorado and does it have any influence on the types of trees they harbour?

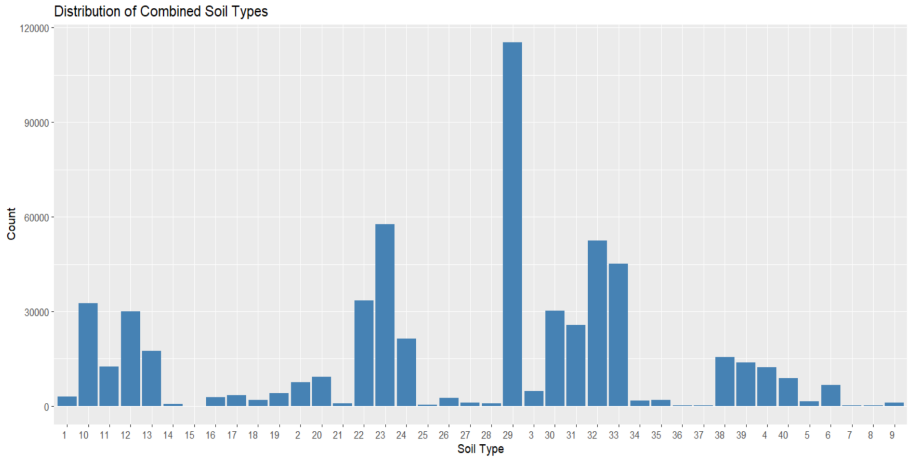


Figure 6: Distribution of Soil

We can see that the most type of prevalent soil is of type 29 which is almost 1200000. Soil type 23, 32 and 33 are also present in large quantity compared to others but is just half compared to that of soil type 29. On the other hand, most of the soil types are almost non-existent like soil type 13,14,37 and many more which indicates their presence in somewhat small amount compared to others.

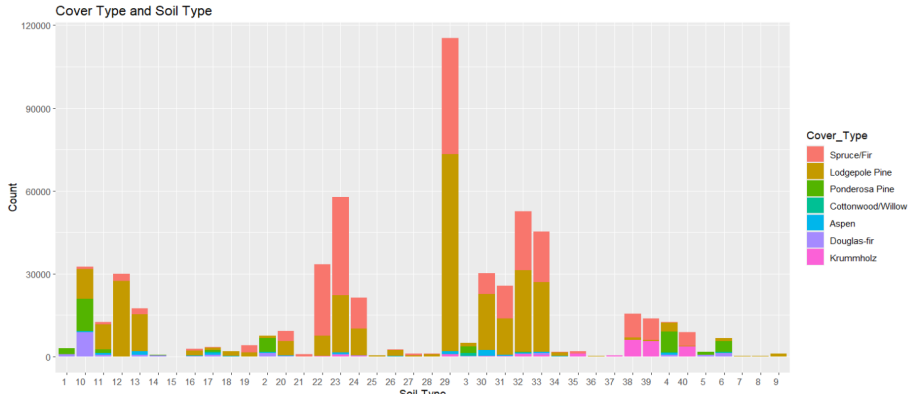


Figure 7: Soils and Cover Type

From the bar plot we can see that soil 29 is the most suitable for **spruce/fir**

tree and **Lodgepole Pine** tree which are also the most present type of trees on other soils compared to other plants indicating their capacity to grow on any type of situation. We can also observe that Spruce and **Lodgepole Pine** are the trees that grows the most in almost every type of soil which is why they are most prevalent in the area as shown by figure 1.

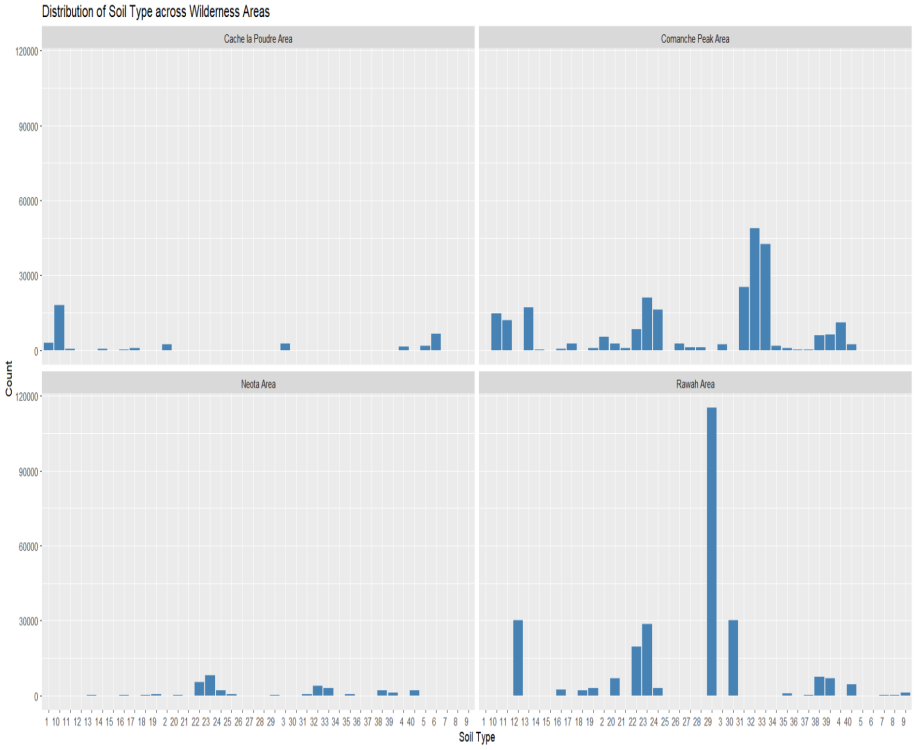


Figure 8: Distribution of Soil Type across Wilderness Area

The plot above shows that **Cache Poudre Area** and **Neota Area** have very less amount of soil which if we take a look at figure 4 explains why they have low vegetation. Again, by looking at figure 4 we can also see that **Comanache Area** has variety of soils compared to **Rawah Area** which explains why one hosts more species of trees than the other.
Overall, the findings highlight how important soil type is for shaping forest ecosystems in the northern mountains of Colorado. It affects which tree species grow in the area and contributes to the overall diversity of plants

3.3 Elevation and Cover Type

How does elevation influence the distribution of tree species?

First let's study the distribution of elevation throughout the mountain.

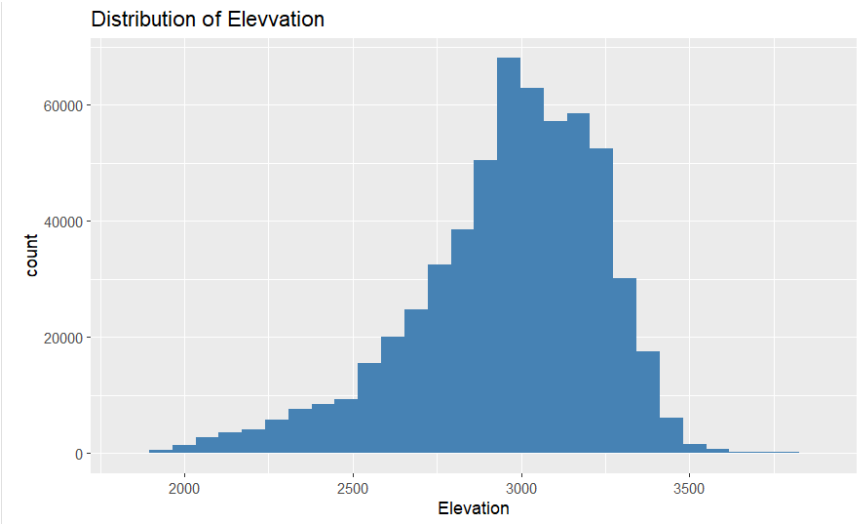


Figure 9: Distribution of Elevation

The area ranges from about 1600 meters to 3700 meters approximately above sea level.

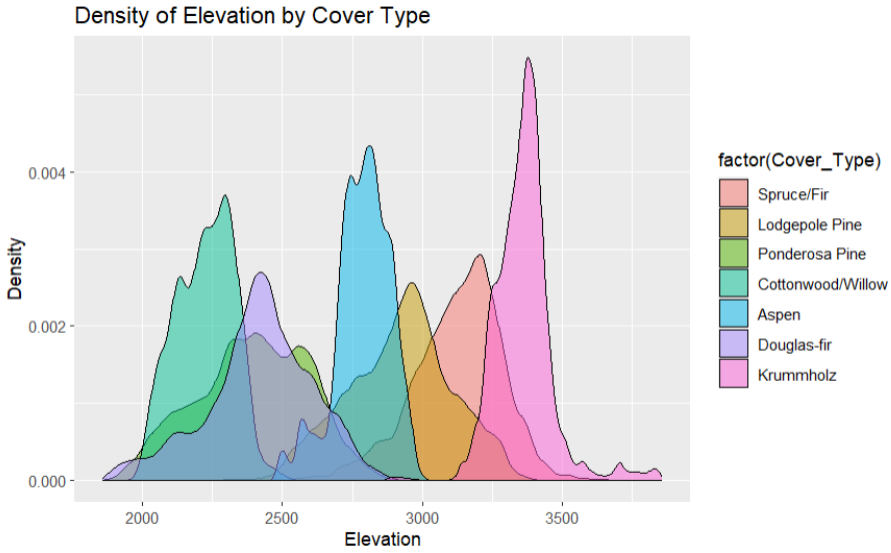


Figure 10: Density of Elevation by Cover Type

The density plot shows distinct elevation preference for different tree species. **Cottonwood** predominantly grows at lower elevations, ranging from around 1500 to 2500 meters. In contrast, **Douglas Fir** and **Ponderosa Pine** also thrive at lower elevations around 1500 meters but can also grow at elevation of 3000 meters.

At higher elevations, both **Spruce/Fir** and **Lodgepole Pine** are present. Spruce/Fir trees typically grow at elevations between 2500 and 3300 meters, reaching their highest density around 3000 meters. Similarly, Lodgepole Pine grows at elevations between 2500 and 3600 meters, with its greatest density around 3250 meters.

Finally, the Krummholz tree, though less abundant, is concentrated at elevations above 3100 meters, reflecting its preference for higher-altitude environments

3.4 Elevation and Wilderness

Does elevation affect the distribution of Wilderness Areas across the mountain?

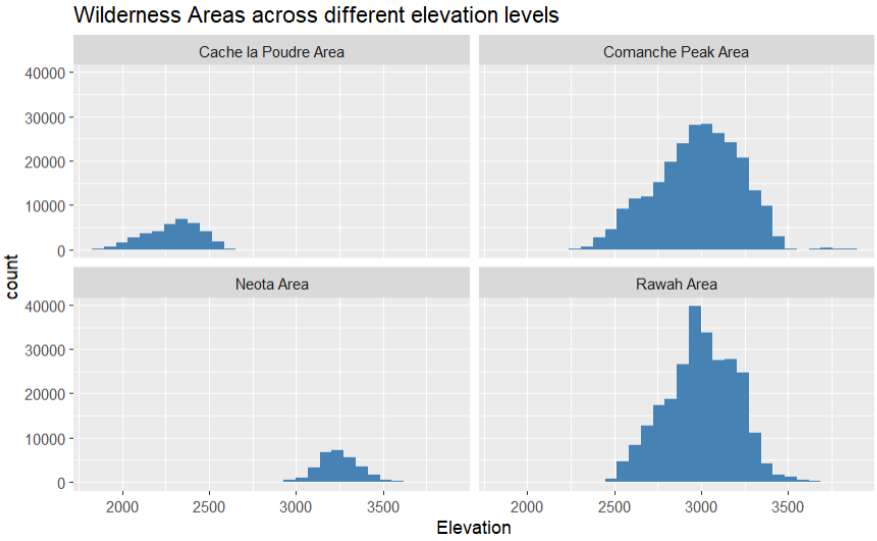


Figure 11: Distribution of Wilderness Area across different Elevation levels

The histogram shows that elevation areas do affect the distribution of Wilderness Areas across different altitudes.
Cache la Poudre Area is found at below 2500 meters. If we look at figure 10 and figure 4 this is the area where **Pendorsa Pines** are found in abundance.

The **Comanacche Area** and **Rawah Area** are found at an elevation of 2200 to 3700 meters whereas **Neota Area** is found at an elevation of above 2800 meters only.
Since **Krummholz** grows at higher elevation, is is not found in **Cache la Poudre Area** and vice versa is true for **Pendorsa Pine**.

3.5 Further analysis of soil

4.5.1. Does soil type also change depending on elevation?

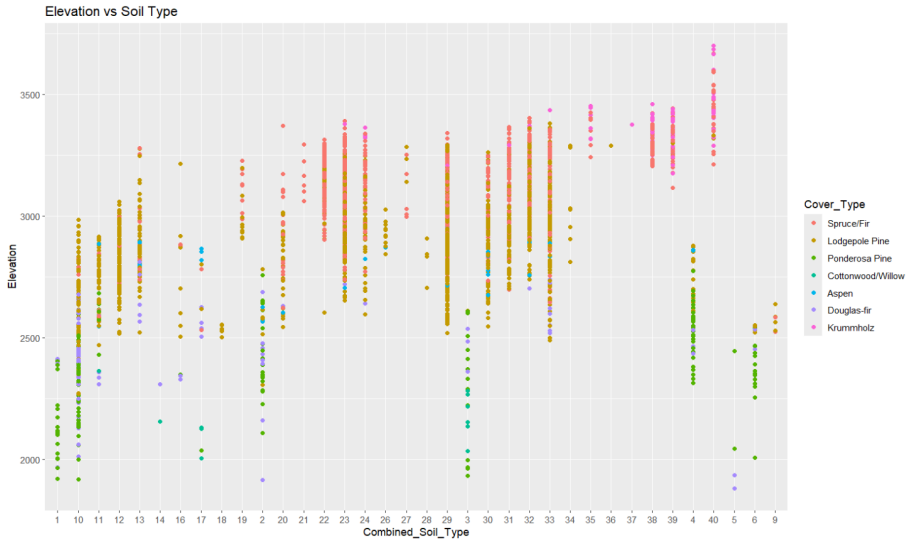


Figure 12: Elevation vs Soil type

The relationship between soil type and elevation reveals that different soil types are distributed across varying elevation levels. As elevation increases, the type of soil also changes progressively. For instance, soil type 1 is predominantly found at lower elevations, while soil type 40 is located at the highest elevations. This suggests a natural gradient where specific soil types are more suited to particular altitude ranges, possibly due to variations in factors like climate, vegetation, and drainage conditions at different elevations. The plot illustrates this trend, highlighting how elevation influences soil distribution.

4.5.2. Does the distance from water sources also influence the type of soil?

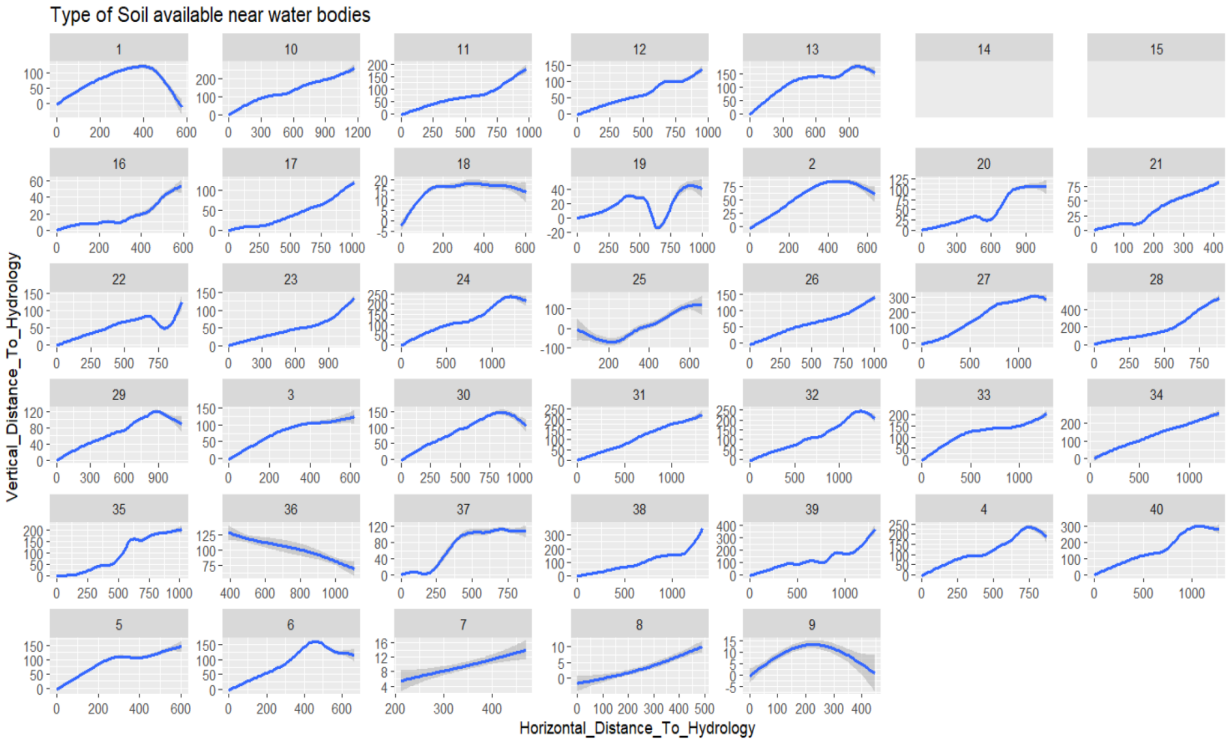


Figure 13: Type of soil near water bodies

The plots reveal a challenging relationship between soil types and their proximity to bodies of water, as many soil types appear to be present across a wide range of horizontal distances. For instance, soil types 1, 16, 18, 2, 21, 8, 9, 5, 6, and 7 are commonly found within 600 meters of a body of water, while other soil types extend as far as 1000 meters away.

Regarding vertical distance from hydrology, most soil types are located within 100 to 150 meters above the water level. However, soil types 8 and 25 are found at lower elevations, beneath the water line. Notably, soil types 4, 24, 27, 28, 31, 32, 38, 39, and 40 can be identified at elevations exceeding 200 meters above any body of water, whereas soil types 7, 8, 9, and 25 are situated as close as 40 meters from the water's edge. Although some soil type have negative vertical distance from water resources but still have positive horizontal distance, it is because of different levels of elevation.

3.6 Water resources and Cover Type

Does the distance from water sources also affect the distribution of Cover Type?

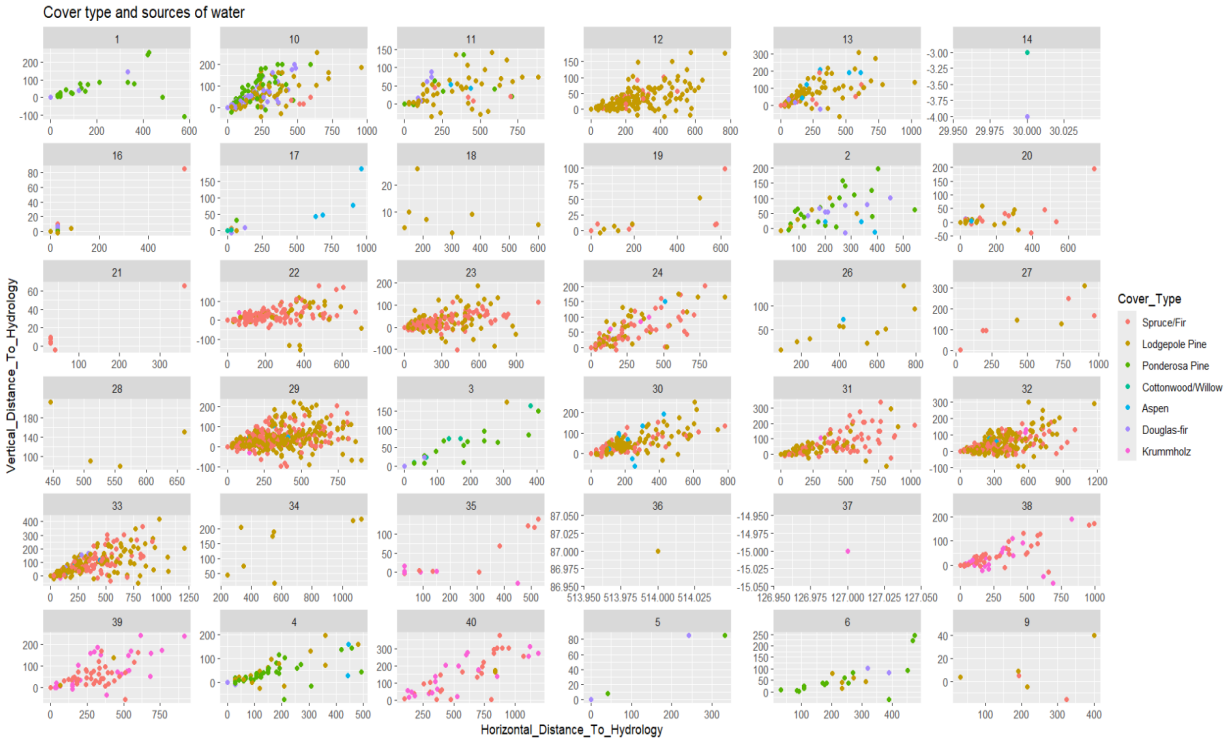


Figure 14: Cover type and sources of water

The plot above illustrates that **Krummholz** trees are predominantly located closest to water sources and are primarily associated with soil type 38. In contrast, **Aspen** trees are typically found at a distance of about 200 meters from water bodies. Both **Ponderosa Pine** and **Douglas Fir** are also commonly found near these water sources. Unlike the other species, **Cottonwood**, **Spruce**, and **Lodgepole Pine** are distributed across various distances from water, appearing both near and far from these sources. However, it is important to note that **Cottonwood** trees are exclusively found at lower elevations. In conclusion **Krummholz** , **Penderosa Pine** and **Douglas Fir** are found near sources of water while **Cottonwood**, **Lodgepole Pine** and **Spruce** are found across various distances from water resources.

3.7 Hillshade Distribution

Does elevation influence the amount of sunlight received by a specific area, and how does it impact the distribution of different cover types?

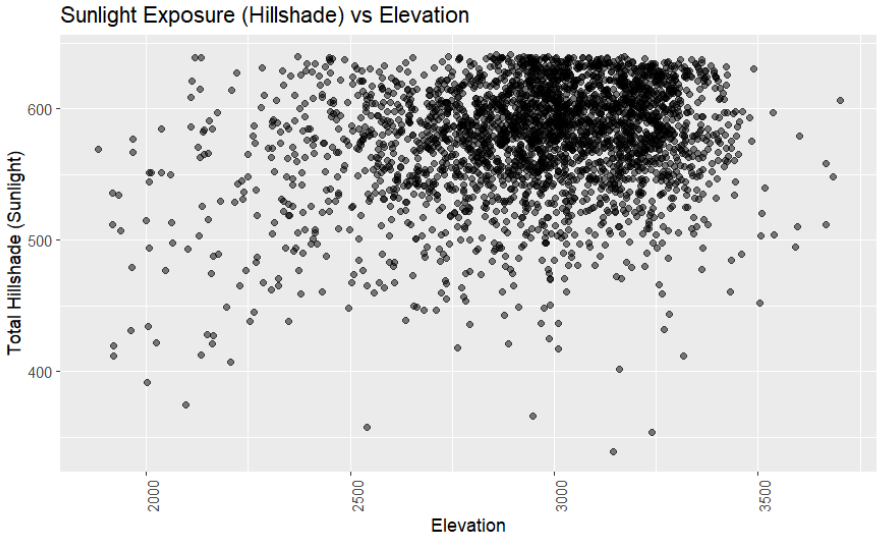


Figure 15: Effect of Elevation on Hillshade

From the graph above, it appears that elevation has a large influence over the total amount of sunlight received in an area. Although the sunlight received by every elevation point is the same (ranging from approximately 300 to 650 units), However, we observe a higher density of sunlight at elevations between 2500 and 3000 meters, indicating that these elevations may receive slightly more sunlight overall. If we look at figure 10, then we can see that **Spruce** and **Lodgepole Pine** grows at that elevation level which are the 2 of the most prevalent trees on the vicinity of the the mountain.
Thus, we can conclude that elevation affects the amount of sunlight received by a certain area which in turn affects the distribution of cover types.

3.8 Fire Hazard Areas

What areas are more prone to fire hazard?

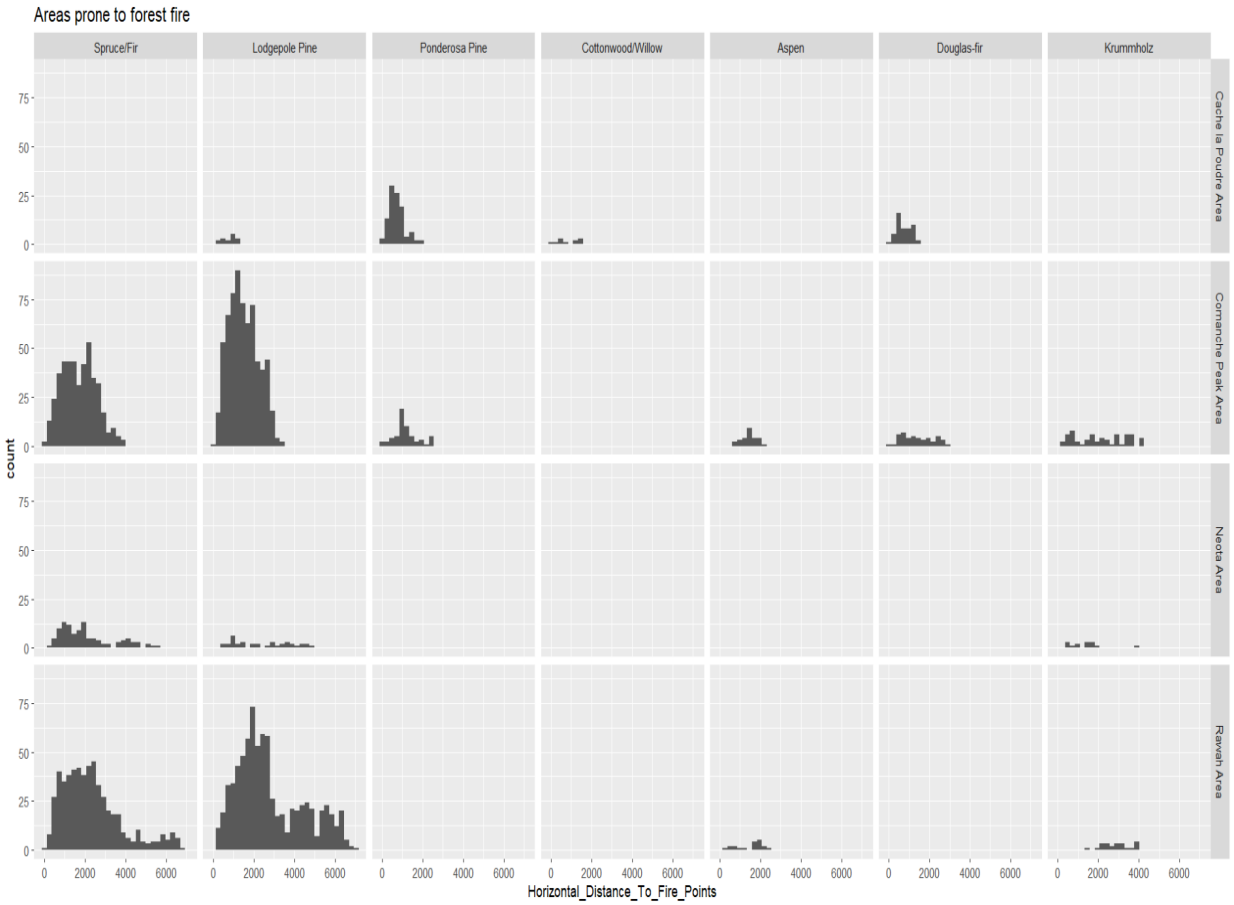


Figure 16: Areas prone to fire hazard

The figure reveals that **Comanache** and **Rawah** area are more prone to fire hazards compared to others. It is because those areas hosts large number of plants/trees compared to other areas. The other interesting fact to observe is that Pine trees and Fir trees are planted more closer to fire points compared to others. According to research papers, **Lodgepole Pine** has serotinous cones that release seeds in response to fire heat, allowing for regeneration post-fire (Zimmerman & Omi, 1991). Similarly, **Ponderosa Pine**'s thick bark insulates it from low-intensity fires, making it well-suited for fire-prone environments (Stephenson et al., 1991).

3.9 Influence of Roadways on Cover type

In these rocky mountains, where are the roads built and do they influence the distribution of cover types around these areas?

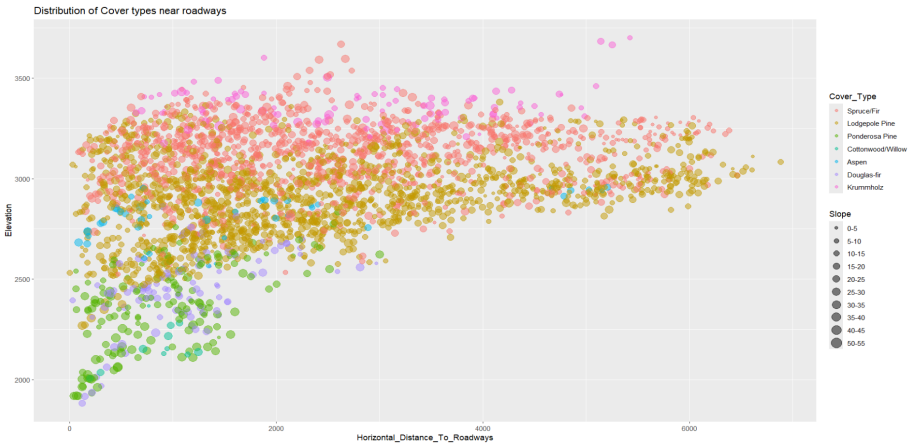


Figure 17: Distribution of Cover Types near roadways

From the plot above, we can confirm that the cover type does not change accordingly no matter how near or far they are from roadways as every type of trees are found near and far from the roadways. However, roads are mostly built at an elevation of 2500-3000 meters where **Lodgepole Pine** are largely present and if we observe the bubbles we can also derive that the roadways are built at an slope of 0 to 5 degrees but mostly built on slopes with lower degrees.

4 Observations

1. Cover Type distribution:
The **Spruce** and **Lodgepole Pine** are the most prevalent trees while **Cottonwood** is the least prevalent in the areas. Other trees like **Ponderosa Pine**, **Aspen**, **Douglas fir** and **Krummholz** are also present but are not notable compared to the number of **Lodgepole Pine** and **Spruce**.
Across different wilderness areas, **Comanache Peak Area** hosts the most number of trees but **Rawah Area** hosts the most species trees. Cover type is equally distributed on every Cardinal Directions.

2. Soil Type:
There are 40 types of soil and only few of them are very suitable for plantation. Soil type is poorly distributed among the different wilderness areas as some areas like **Cache la Poudre** and **Neota** have little to no distribution of soil while the other 2 have a massive amount of soil.

3. Elevation:
Elevation plays the most major role in determining the type of trees that grow there. Eg: **Pendorsa Pine** grows only at lower altitude while **Krummholz** trees grow at an higher.
Cache la Poudre Area is found at an elevation of below 2500 meters whereas other areas are found above 2500 meter range. Elevation also affects the type of soils found in the vicinity as lower altitudes seems to have soil type of low integer value (i.e 1,2,3...) and vice versa.
Also the area at an elevation of 2500-3300 meters receive a large amount of sunlight throughout the day.

4. Distance from Hydrology: It is quite hard to determine the relationship between soil type and distance from hydrology may it be vertical or horizontal because some soil types are found only near water sources while others are found across various distance from the water sources.
As for cover type, **Krummholz** , **Ponderosa Pine** and **Douglas Fir**are found nearer to water sources while other species such as **Spruce**, **Cottonwood** and **Lodgepole Pine** are found distributed across various distance from sources of water.

5. Sunlight: Sunlight levels remains constant across various elevations with every area receiving 300 to 650 units of sunlight. However, the density varies as elevation below 1500 is very low and is highest at 2500 meters above sea level.

6. Distance from fire points: Trees like fir and pine trees are planted near fire points because pine trees requires heat to release seeds allowing for a post fire regeneration while fir trees have thick bark which protects them from fire.

7. Distance from road ways: Roadways are found mostly on an ele-

vation level of 2500-3000 meters and does not seem to have any effect over the distribution of cover type at all.

5 Conclusion

In this report, we examined the relationships between different forest cover types, environmental factors, and soil types in the Rocky Mountains of northern Colorado.

First, we found that tree species like **Lodgepole Pine** and Spruce/Fir are strongly influenced by elevation and soil type. These factors affect both the types of trees present and their density in various areas. **Lodgepole Pine** was the most common species, thriving in a variety of soil types and elevations, while other species showed specific preferences based on their ecological needs.

Additionally, certain soil types, especially Soil Type 29, support a wide range of tree species, highlighting the role of soil composition in vegetation patterns. We also found that trees like **Krummholz** and Aspen tend to grow closer to water sources, while other species have more flexible distribution patterns.

The analysis of how elevation affects hillshade showed that sunlight availability is crucial for tree growth. Our findings indicate that areas with optimal sunlight, particularly between 2500 and 3000 meters in elevation, are favorable for dominant species like Spruce and **Lodgepole Pine**.

Overall, this report improves our understanding of the ecological interactions in the northern Colorado mountains and offers valuable insights for forest management and conservation efforts.

6 Link to the Shiny App

My Shiny App

References

[1] Zimmerman, G. T., & Omi, P. N. (1991). Ecology of lodgepole pine in fire-prone areas. In *Fire in the Rocky Mountains* . Tall Timbers Research Station.

[2] Stephenson, N. L., M. R. H. N. Omi, & K. W. Omi. (1991). Fire as a management tool in the Ponderosa Pine ecosystem. In *Fire in the Rocky Mountains* . Tall Timbers Research Station.