UNIVERSITY OF CALIFORNIA

Los Angeles

Alpine Wildfire Risk Analysis

with Monte Carlo Simulations in Mount Rainier National Park

A thesis submitted in partial satisfaction

of the requirements for the degree Master of Science

in Statistics

by

Nikhil Pavan Auradkar

2024

ABSTRACT OF THE THESIS

Alpine Wildfire Risk Analysis

via Monte Carlo Simulations in Mount Rainier National Park

by

Nikhil Pavan Auradkar

Master of Science in Statistics

University of California, Los Angeles, 2024

Professor Mark S. Handcock, Chair

This thesis explores an approach to analyzing wildfire risk in forests, with a focus on Mount Rainier National Park. Monte Carlo simulations utilize terrain and elevation data, point process methods, and existing wildfire spread knowledge to generate risk heatmaps at a fine-grained scale. The research is driven by the following questions: How can we accurately and efficiently simulate wildfire start points and spread using terrain and elevation data and various statistical methods? What can we learn about risk factors, drawing from the simulation results in conjunction with existing wildfire knowledge?

The study aims to capture the inherent uncertainties and complexities associated with wildfire behavior. Grounded in the relationship between topographical features, forest density, and climatic conditions, the simulation model effectively integrates the key factors contributing to wildfire ignition and spread. The probabilistic nature of the Monte Carlo method allows for the

exploration of a wide range of wildfire scenarios, providing a nuanced understanding of risk and potential impacts. Results from the simulation identified areas of increased wildfire susceptibility and gave rise to Bayesian conclusions about risk in variables such as seasonality, moisture levels, and terrain.

The thesis of Nikhil Pavan Auradkar is approved.

Nicolas Christou

Frederic R. Paik Schoenberg

Yingnian Wu

Mark S. Handcock, Committee Chair

University of California, Los Angeles

2024

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

Introduction

This project is a continuation of a much more basic simulation assignment started in 2023. Improvements to the basic model will be discussed, along with the change in methodology, results, and conclusions.

1.1 Background and Context

Washington State, particularly its alpine regions, has a significant history of wildfires that have had environmental, economic, and human impacts. These regions, characterized by dense forests and rugged terrain, are particularly susceptible to fires due to dry summers and abundant fuel sources. Over recent years, these wildfires have grown in both frequency and intensity, likely driven both by climate change and increased human activities.

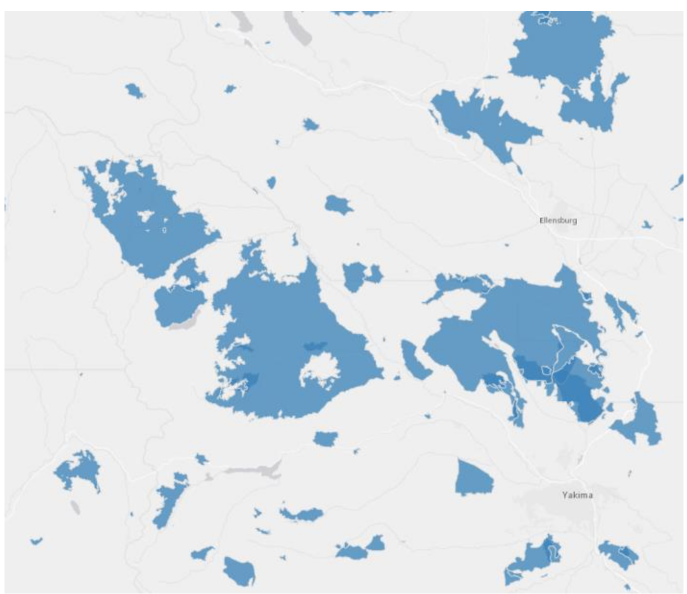

Figure 1.1: Large wildfires since 1973 in the Mount Rainier area

Wildfires have led to the loss of human lives, property damage running into billions of dollars, destruction of wildlife habitats, and the release of vast amounts of carbon dioxide into the atmosphere, contributing further to global warming. The after-effects of wildfires also pose significant challenges. Post-fire landscapes are often susceptible to soil erosion, landslides, and flash flooding, further exacerbating the environmental devastation. Additionally, the smoke from wildfires affects air quality not just in the immediate vicinity but also across large distances, impacting public health.

Given these severe impacts, there is a need for effective risk analysis to manage and mitigate the threat of wildfires. Predicting where and when a wildfire might occur, understanding its potential spread and intensity, and identifying the areas and communities at risk are all crucial for early warning, strategic resource allocation, planning effective responses, and ultimately reducing the damage caused by these catastrophic events. Risk modeling techniques, such as the Monte Carlo simulations performed in this project, can play a critical role in this context by providing probabilistic and scenario-based risk assessments, helping decision-makers understand and prepare for a range of possible wildfire events.

1.2 Problem Statement and Objectives

Despite the growing recognition of the wildfire risks, identifying areas that are most susceptible to these disasters is still a significant challenge. This challenge is rooted in the complex and dynamic nature of wildfires. A multitude of factors influence the spread of wildfires, including weather conditions (temperature, humidity, wind speed), terrain information (vegetation type, tree density, moisture content, proximity to bodies of water/ice), and the topography of the land (slope, aspect, elevation). These factors are further complicated by the complex variables involved in starting a wildfire – involving both weather conditions (primarily lightning) and human factors (campgrounds, roads). For the sake of simplification and interpretability, this model will assume ignition via lightning strikes, which have been shown to cause a majority of large alpine fires.

The interplay of these factors often leads to nonlinear and unpredictable fire behavior. Traditional methods for predicting wildfire risk often fall short in capturing these complexities and uncertainties, and thus, may not provide the level of detail and accuracy needed for effective planning and response. Therefore, there is a need to develop a more sophisticated and comprehensive approach to wildfire risk analysis in Washington's alpine regions that leverages our understanding of fire behavior and incorporates the detailed topographical and terrain data available.

The primary objective of this project is to develop and implement a Monte Carlo simulation-based model for assessing wildfire risk in Washington's alpine regions. This probabilistic model will seek to integrate the various factors influencing wildfire ignition and spread, including but not limited to topographical features, terrain features, and climatic conditions. By incorporating the inherent uncertainties in these factors and simulating a broad range of wildfire scenarios, the model aims to provide a comprehensive and nuanced understanding of wildfire risk.

Specifically, the model will aim to identify areas most susceptible to wildfires, the most likely ignition points for wildfires, and the potential extent and intensity in these areas under different scenarios. These findings, synthesized, should be able provide valuable inputs for disaster risk management strategies, including preventive measures, emergency response planning, and resource allocation. Ultimately, this research aims to contribute towards a more resilient and prepared fire response system.

CHAPTER 2

Existing Literature and Modeling

2.1 Literature

2.1.1 The Rothermal Model

The Rothermel Surface Fire Spread Model is a foundational model used to predict the spread of wildfires. The model calculates the rate of fire spread based on factors such as fuel type, fuel moisture, wind, and slope. The comprehensive formula looks like:

$$R = \frac{I_R \xi (1 + \phi_w + \phi_s)}{\rho_b \varepsilon Q_{ig}} \tag{2.1}$$

R is the rate of spread (flaming front speed)

I_R is the reaction intensity (energy release)

ξ is the propagating flux ratio (proportion of heat transfer)

Φ_W and Φ_P are dimensionless multipliers based on wind/slope

P_b is bulk density (fuel metric)

ε is the effective heating number

Q_{ig} is the heat of preignition

The model is designed to estimate surface fire behavior (fires that burn within 6 feet of the ground). The Rothermel model, with modifications, has been widely used for fire and fuels management since its development in 1972. It includes equations for the rate of spread, flame length, and fire intensity. The model's main inputs are fuel particle properties (such as moisture content), environmental variables (wind speed and slope), and the fuel bed.

Not all variables here can be effectively transferred to the simulation – however specific parts of the model were helpful in construction the spread model. Specifically, factors influencing the bulk density and heat of preignition were incorporated – the moisture of extinction M_X and the moisture content M_F. The Rothermal model calculates fire spread as a balance between heat source and heat sink, with inputs like fuel moisture content playing a key role in the outcome. The model factors in live and dead fuel separately, with dead fuel moisture of extinction serving as a critical threshold where fire can no longer propagate through dead fuel.

The simulation focuses on how live fuel moisture of extinction is calculated. According to the Rothermel Model, live fuel moisture is dynamically influenced by the proportion of live-to-dead fuels, meaning as the live fuel moisture content drops, the fire’s ability to spread increases. This behavior was the basis for the spread behavior, which simulates how changing moisture levels affect fire behavior over time.

The model's reliance on slope and wind speed also plays a significant role in how moisture impacts fire spread. For instance, the propagating flux ratio adjusts the heat release to account for wind and slope effects, which was incorporated into my simulation to mimic how fires spread faster uphill and in dry, windy conditions.

2.1.2 Lightning Strike Probability Risks

The University of Georgia explores the probabilities and risk assessments associated with lightning strikes on trees and other tall objects. It details formulas used to estimate the probability of a lightning strike, including factors like historic lightning strike density, tree height, and surrounding topography. The guide explains how to use lightning strike density maps and gives practical examples, such as the frequency of lightning strikes in the southeastern U.S. The manual also addresses factors that influence lightning attraction, such as the height of objects, the electrical field created by tall objects, and different types of lightning strokes.

The lightning strike probability model was used to determine ignition points based on environmental factors. The model described a relatively simply probabilistic estimate of where lightning will strike, focusing on the influence of tree height and the surrounding area:

$$\textbf{Number of Years Between Lightning Strikes} = \frac{1}{(\text{N mile}^2 \text{ per year}) \text{ X } \{[(3.142) \text{ X } ((\text{HT feet}) \text{ X } (3))^2] \;/\; (5{,}280)^2\}} \tag{2.2}$$

N mile^2 per year = Number of ground strikes per square mile per year.
HT feet = Tree height in feet above its surroundings.

This formula was used to simulate ignition in areas with trees. The guide also provided insight into lightning attraction distances, which helped refine how I positioned potential ignition sources in my simulation. For instance, taller trees have an enhanced electric field, increasing the likelihood of a lightning strike, a factor that was included to simulate higher ignition risk in specific terrain. By understanding how different current levels (up to 100kA for positive polarity strikes) affect the likelihood of lightning starting a fire, ignition behavior could be better predicted.

2.1.3 Other Sources

Other sources were used to better understand wildfire behavior and general knowledge. One emphasizes the role of topography in influencing wildfire spread. It describes how slope steepness, aspect, and elevation affect fire behavior. Fires tend to spread more quickly uphill due to pre-heating of fuels by the flames. Conversely, fires on downhill slopes or in complex terrain, such as valleys and ridges, can spread more unpredictably due to wind turbulence. This document also referenced historical Australian bushfires, explaining how weather conditions and topography contributed to some of the country's most destructive wildfires, like the 2009 Black Saturday fires.

2.2 Modeling

Multiple studies have been done with similar intents. Exploring methodologies helps with determining pros and cons to this given approach.

2.2.1 Assessing Fire Risk Using Monte Carlo Simulations of Fire

This paper focuses on modeling fire spread using Monte Carlo simulations combined with FARSITE, a well-known fire behavior simulator. The study area covers 300 km^2 of Mt. Carmel, Israel, where a series of 500 simulations were run to model the spatial distribution of fires. The key inputs for this model include topographical data, vegetation/fuel maps, and climatic data. The simulations randomly selected ignition locations, fire length, and climatic data for each run. The ignition risk was modeled using proximity to roads as a proxy for human-caused fires. This study emphasizes the complexity of fire behavior, as it is influenced by a combination of vegetation, topography, and human activity. The Monte Carlo process generated fire-risk maps with high spatial resolution, allowing fine-scale strategic fire prevention planning. This approach confirmed that simulated fire frequencies closely correspond to historical fire events, suggesting that these simulations can serve as a reliable predictor for fire risk.

The Monte Carlo approach here is particularly relevant as it underscores the importance of combining various spatial factors (ignition location, fire length, topography) and temporal factors (calendar date, climactic data) to assess fire spread risk comprehensively. Also, the paper’s validation with historical data reinforces the utility of such simulations in predicting risk. The fine-grained simulations performed were a good starting point to look at and improve upon.

2.2.2 Spatiotemporal Wildfire Modeling Through Point Processes with Moderate and Extreme Marks

This paper presents a sophisticated statistical model to analyze wildfire ignition and spread based on spatiotemporal point processes. The primary modeling technique is the log-Gaussian Cox process (LGCP), which incorporates both moderate and extreme wildfire occurrences. By combining extreme-value theory with point process modeling, this approach allows for the joint analysis of wildfire occurrence and burnt area sizes, using Bayesian hierarchical models. The study was conducted on wildfire data from the French Mediterranean Basin from 1995 to 2018. Key factors influencing fire ignition in this model include weather (Fire Weather Index), forest cover, and historical fire data.

By jointly modeling both moderate and extreme fires, the method improves the accuracy of predicting large-scale wildfires, especially in regions susceptible to significant fire risks. Nonlinear effects of covariates are captured to improve predictions, making the model adaptable to different seasonal conditions and locations. Focusing on the distribution of fire size is an important feature to look at during analysis and interpretation. This paper is relevant to the planning of a simulation as it provides a detailed framework for modeling fire ignition and size distributions based on spatiotemporal factors. Its emphasis on Bayesian modeling and uncertainty quantification aligns with the goal of simulating ignition behavior in complex terrains and understanding how varying conditions (such as weather and forest cover) influence wildfire spread.

2.2.3 FireCast: Leveraging Deep Learning to Predict Wildfire

The FireCast paper explores the use of deep learning techniques to predict wildfire spread. The neural network (NN) model developed for this study relies on input data such as satellite imagery, elevation, weather conditions, and historical fire events. The system produces predicted fire perimeters and risk maps, which are then evaluated using metrics like accuracy, recall, and F-scores. The authors advocate for improvements in granularity by incorporating Monte Carlo simulations to refine predictions. Additionally, the paper briefly discusses the integration of a Wildfire Fire Decision Support System (WFDSS), a geospatial tool for fire management that also employs predictive modeling.

This paper uses GIS data and satellite inputs for accurate fire prediction, which are very important to the planning of the simulation. While the simulation may focus more on deterministic factors like elevation and topography, the combination of deep learning and Monte Carlo simulations presented here offers a pathway to enhance predictions, especially in integrating spatial data to create more granular and precise fire risk maps. It also established an important fact in that current official wildfire risk modeling via Farsite is very slow (in the order of days to months) and very situational.

CHAPTER 3

Data

The data utilized in this research comprises detailed topographical and terrain information of the alpine regions in Washington State. Data was taken primarily from the United States Geological Survey (USGS), which has extensive data in a wide variety of formats. One dataset provided by the USGS is 3D Elevation Program (3DEP) - the topographic data provided includes elements such as elevation, recognizable features, and landmarks, providing a detailed overview of the surface characteristics of the landscape at a very fine granularity (~2x2 meters in some areas, including Mount Rainier National Park). Another such dataset is offered by the USGS's National Land Cover Database (NLCD), which provides high-resolution terrain data collected using satellite techniques. This data categorizes the entire continental US into different terrains at a granularity of around 30x30 meters. It is especially useful for detailed terrain analysis, including the study of surface and subsurface features. In addition to these, the TNM datasets also include comprehensive hydrography and bathymetry data, which map out the water bodies across the country. This includes information about the depth and contours of these bodies, offering essential insights for any analysis related to water flow, flood risk, or aquatic ecosystems.

3.1 USGS 3DEP Dataset

The USGS 3DEP is a major initiative to collect high-resolution, three-dimensional elevation data for the United States. It provides topographic maps in GeoPDF and GeoTIFF formats, designed for various uses, including fire modeling. The 7.5-minute quadrangle maps cover 1.15 miles per minute, making each map span approximately 8.6 by 8.6 miles with a granularity of less than 2 meters. The maps used contain elevation data, which critically can be used to derive other important variables such as slope and aspect. Together they are important for modeling how fires spread over different landscapes. The historical topographic maps also offer insight into past landscapes, while more modern maps are available for contemporary analyses.

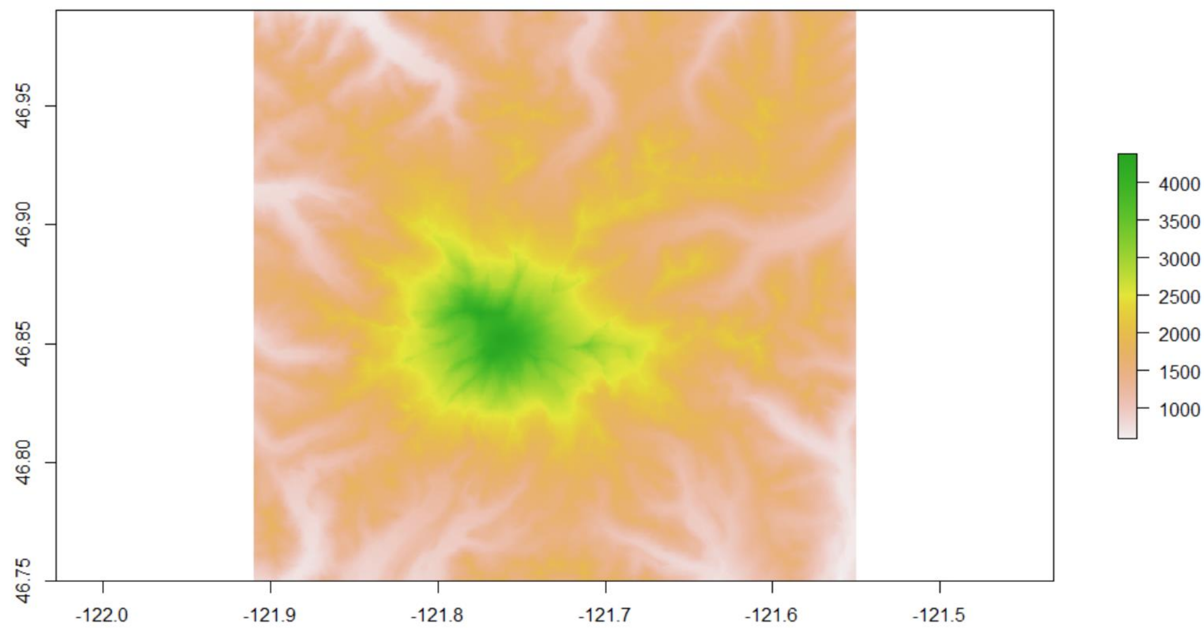

Figure 3.1: USGS 3DEP data in Mount Rainier National Park

In addition, USGS also provides LIDAR (Light Detection and Ranging) data can include trees and river features, providing valuable inputs for fire spread simulations. LIDAR is particularly useful for identifying fine-scale variations in terrain, helping to understand the obstacles or pathways fires may encounter. However, NLCD terrain data was eventually chosen over LIDAR due to its direct relevant to the Rothermal spread model.

3.2 USGS NLCD Dataset

The NLCD 2021 dataset provides comprehensive land cover data for the contiguous United States (CONUS), including vegetation types, water bodies, and developed areas. It is periodically updated to reflect changes in land use, natural vegetation, and development patterns. For wildfire modeling, the vegetation type is critical since different terrain have vastly different fire behaviors. Grasslands, forests, and shrublands have distinct fuel properties, such as fuel moisture and moisture of extinction, as mentioned above in the Rothermal model. This dataset, provided at a 30 meter granularity, helps to determine the terrain's fuel load and structure, which can significantly influence wildfire ignition, spread rates, and overall fire behavior. The dataset also provides information on terrain and land cover dynamics, such as changes in vegetation over time.

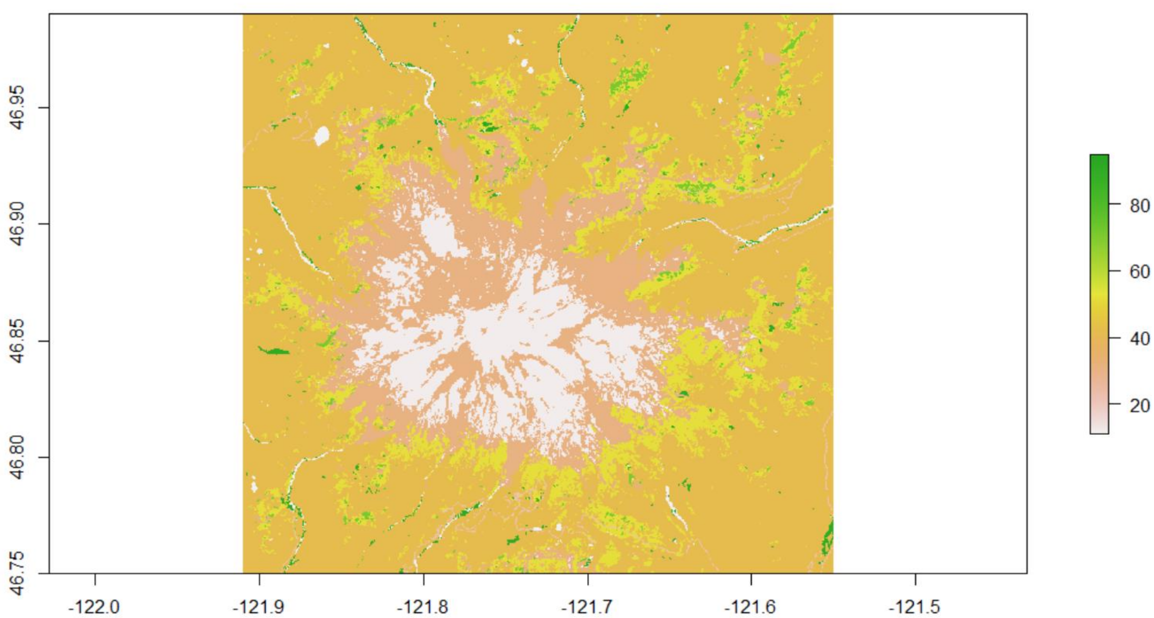

Figure 3.2: USGS NLCD data in Mount Rainier National Park

Figure 3.3: USGS NLCD legend (note different vegetation types)

3.3 Washington Large Fires Dataset (1973–2023)

The Washington Large Fires Dataset offers detailed historical data on large wildfires in the state, including the date, size, acreage, boundary/perimeter, and cause of each fire. This dataset is essential for understanding historical fire behavior, helping validate fire models by comparing simulated fire spread and size to real-world events. The dataset also allowed for analysis of trends in fire frequency, size, and causes, offering insights into how factors like climate change or human activities are impacting wildfire risks. By including boundary and perimeter data, this dataset is particularly useful for visualizing how fires expanded over time, and for refining the simulation of fire behavior based on real-world observations.

CHAPTER 4

Simulation

4.1 Initial Model

The initial simulation logic was relatively simple and did not capture much of the Rothermal model behavior, data sourcing, or model logic listed above.

4.1.1 Methodology

One of the first thoughts was to find a determine what temperature and wind speed were more likely to cause large, dangerous burns. Thus, they could be the parameters that were varied the most in testing to determine their 'optimal' values. However, they were not measured as temporal variables existing explicitly in the simulation. Temperature was instead tied directly to spread probability, i.e. the likelihood that a neighboring cell catches fire in each iteration of the simulation. Wind speed was tied to spread probability, with higher speeds increasing the probability and direction increasing/decreasing the probability.

Several parameters, directly influencing the occurrence and spread of wildfires, were identified for the simulation. One such parameter was the fire start point, which was determined by elevation data and other factors such as the likelihood of lightning strikes at higher elevations or human involvement in camping/hiking areas. From research, it was determined that the two most likely causes of a wildfire were:

1. Human intervention, either from a campground or with machinery
2. Lightning strikes

However, data from the internet widely varies on the proportion of wildfires that are caused by humans versus inclement weather. To try and accurately simulate both impetuses, the model was tuned to favorably select start points that were either close to local minima or local maxima. In doing so, both human intervention and lightning strikes were treated as equal contributors to wildfires.

Importance sampling was used to focus on more likely fire start points and spread patterns, increasing the efficiency of the simulation. In the model, not all start points were equally likely; this required a non-uniform sampling approach, achieved through importance sampling. Mathematically, this can be described as:

$$E[f(X)] \approx \frac{1}{n} \sum_{i=1}^{n} \frac{f(X_i)}{p(X_i)} \tag{4.1}$$

In this context, X represents the start points, and f(X) represent the wildfire risk at each start point. This technique helped create a more representative dataset and consequently a more accurate model.

For terrain data NLCD data was not used – a simplified satellite imagery dataset. The spread of the algorithm from cell to cell was determined by a single variable tied to the RGB value of each of the cells in the satellite image.

Figure 4.1: Satellite imagery of a portion of Mount Rainier National Park

The spread of fire was also modeled as an uncertain parameter, influenced by the terrain, wind direction and speed, and climatic conditions like dry or wet periods. Here, the 'green value' of the satellite data was utilized heavily. As heavily forested areas burn more easily than less forested areas, the model dictated that a burn would be more likely if the 'green value' in the RGB coding for the satellite data was very high. From research, it was also determined that elevation difference and wind speed/direction had major impacts on the spread of wildfires. Fires generally traveled both uphill and with the wind; the model was thus tuned to enhance spread probabilities when either of these conditions were met.

Estimating critical values of the equivalent parameters for temperature and wind speed can be viewed in a Bayesian setting. Let's denote T as the temperature, W as the wind speed, and B as a binary event representing whether a large, dangerous burn occurred. By running numerous simulations varying T and W, the likelihood P(B|T, W) could be estimated. The prior beliefs about how T and W affect B could then be updated using Bayes' theorem, which can be expressed as:

$$P(T, W|B) = \frac{P(B|T, W) \cdot P(T, W)}{P(B)} \quad (4.2)$$

Here, P(T, W|B) is the posterior distribution of T and W given B. P(T, W) is the prior distribution of T and W. The resulting posterior distribution represents the updated beliefs about the distributions of T and W that lead to large, dangerous burns, given the data observed. By examining this distribution, it was possible to make informed statements about the equivalent parameters that were more likely to result in large, dangerous burns.

Other parameters identified were burnout probability – the likelihood that a given cell on fire would enter a 'burnt' state, and stop probability – the likelihood that the entire wildfire would stop burning (in lieu on other parameters for rain). The simulation model was defined to capture the interplay of each the identified parameters. The process was repeated thousands of times, each time with different combinations of the uncertain parameters, leading to different potential fire paths. The result was a theoretical but comprehensive set of wildfire scenarios, capturing a wide range of possible events.

4.1.2 Results

Bayesian strategies were employed to tune the model parameters, particularly the equivalent parameters for temperature and wind speed. Dimensional tiered sampling was applied until the parameters produced a large burn (over 1000 square meters affected) at least once out of every hundred runs. The spread probability was found to need to be at least 0.15 and the modulation factor (based on optimal wind and elevation) was found to need to be at least 1.5.

Once these parameters were fixed, sampling began to find the areas in the specified region with the most risk. N = 1024 simulations were run with parameter values spread_prob = 0.15, burnout_prob = 0.15, stop_prob = 0.005, and modulation factor = 2.

Below are starting points and heatmap-plotted results from this run of simulations:

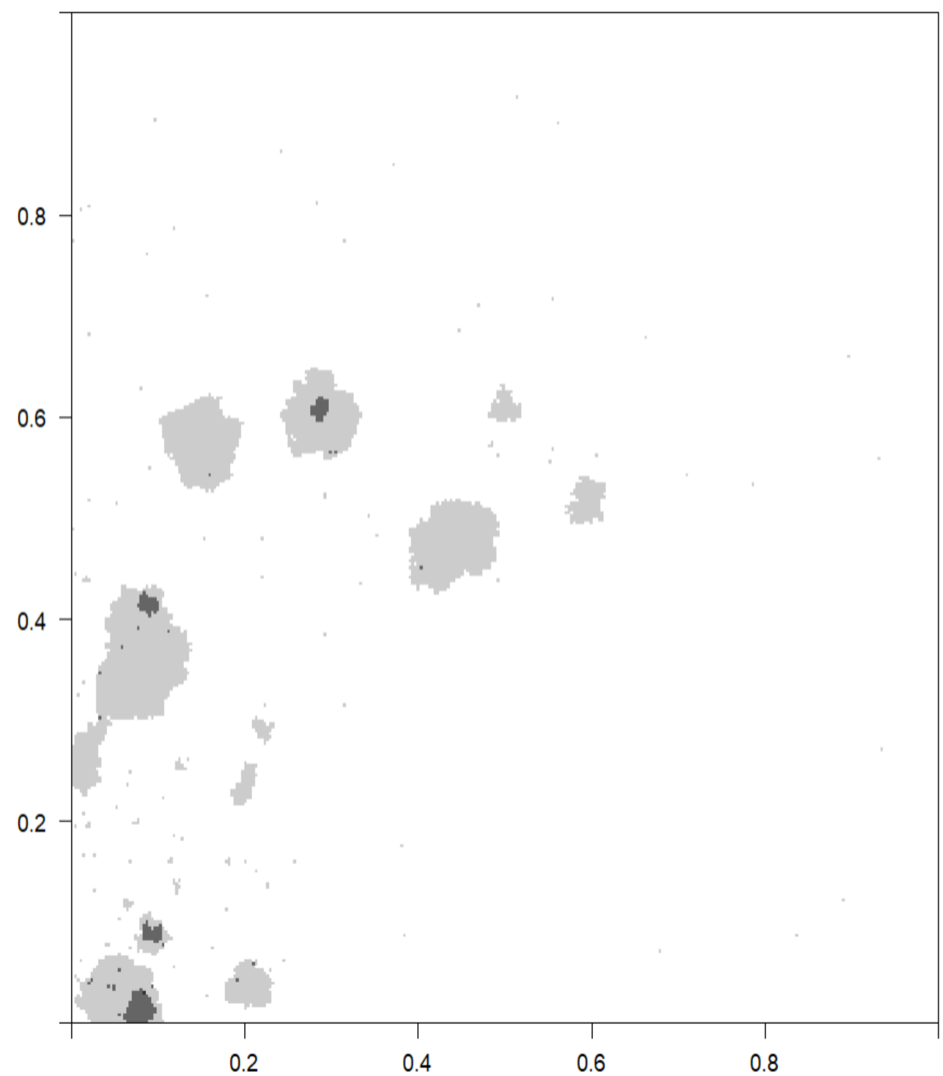

Figure 4.2: Cumulative damage from n = 1204 simulated wildfires

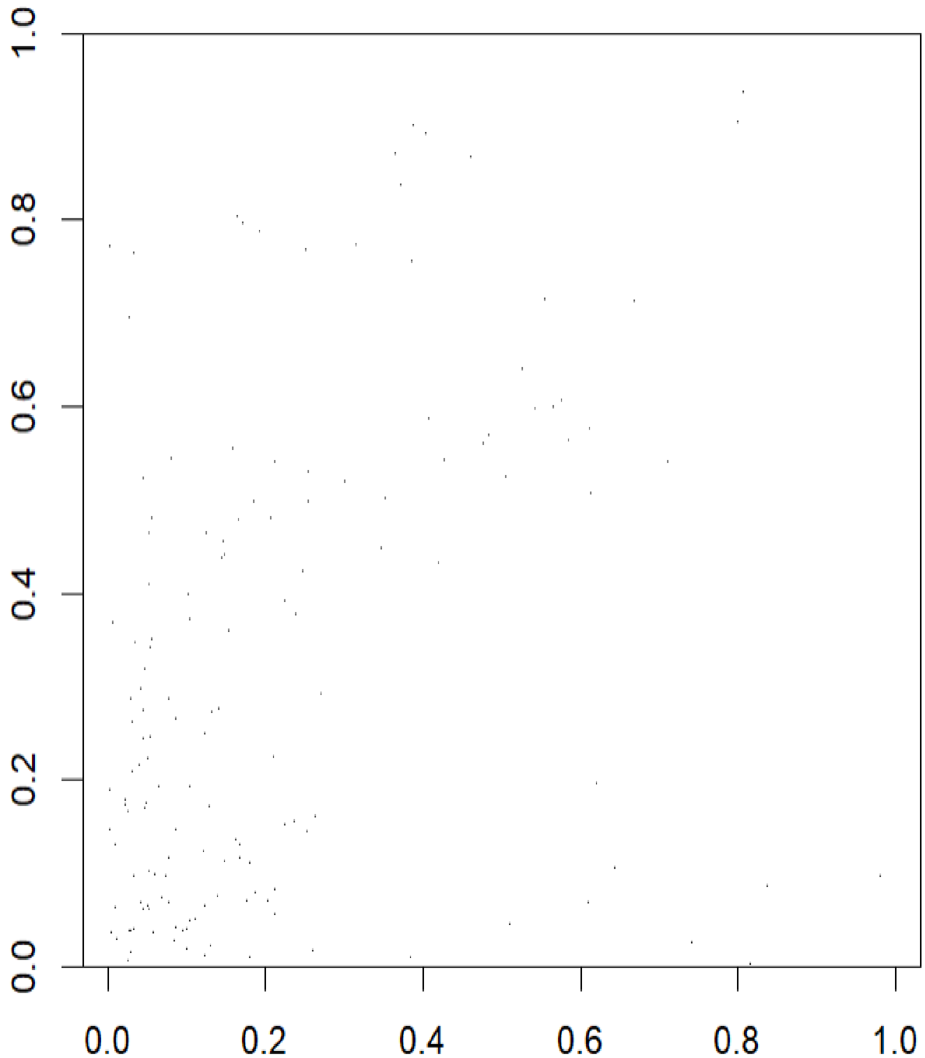

Figure 4.3: Starting points from n = 1204 simulated wildfires

Areas of high risk density were found to coincide with one of two area profiles:

1. Local maxima, with a higher-than-average RGB green value, with a negative gradient on at least two sides. These can be classified as narrow and long hills.
2. Flat areas, with very high RGB green values, with a local minima close by (presumably where the fire started). These can be classified as lush plateaus.

Overall, out of the 1204 simulations run, 24 were categorized as ‘large burns’. Burns heavily favored greener portions of the specified area. These results are in line with the model’s tuning.

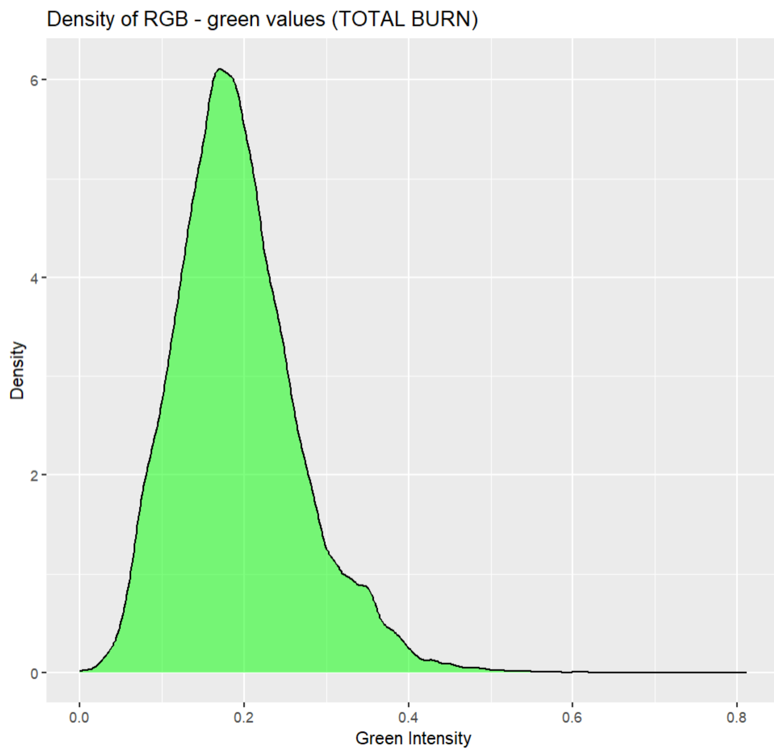

Figure 4.4: Distribution of green RGB values in burnt cells

4.1.3 Limitations

To start, the start/ignition algorithm needed improvement. By attempting to account for both human intervention and lightning strike ignitions, the algorithm had inherently become cloudy. The model considered local minima and maxima as the most likely starting points for wildfires; minima due to human factors associated with roads and campgrounds, which are typically at lower altitudes, and maxima due to the higher propensity of lightning strikes at higher altitudes. The model could benefit from independently addressing these factors, perhaps through a tiered approach that considers multiple variables in concert. Eventually, it was decided to just focus on lightning strike ignitions.

Improvements in terrain analysis were identified to potentially enhance the model's performance. As it is tied to satellite imagery, it does not properly recognize water bodies, large elevation differences, or previously burnt forest areas, which can all act as natural barriers to fire spread. A more detailed terrain analysis (through NLCD) would incorporate these factors, providing a more realistic representation of how fires spread and are stopped in the real world.

The spread probability definition could also be improved. The initial model tied just one variable to spread probability modifications. The probability of fire spread based on terrain and elevation was significantly lacking in this regard. Incorporating these elements would yield more realistic simulations of fire behavior and spread patterns. Furthermore, the stoppage probability used in the model is a simplification in place of variables such as rain and firefighting efforts, implemented to make the simulation more efficient. With more computing power, a more accurate representation of these factors could be used.

Lack of temporal variables was also an issue – no real risk factors could be identified easily, without looking into how temperature and wind tied themselves to other simulation variables. Even the definition of this behavior was somewhat nebulous.

Finally, one of the limitations of this study was the number of simulations performed. Due to computational constraints, the total number of simulations across the final run barely exceeded 1.2K. With more computational power or efficient algorithms, future studies would run more simulations, improving the robustness and reliability of the results. Additionally, more data points would provide a more comprehensive understanding of the multivariate nature of wildfire behavior and risk.

4.2 Current Model

The wildfire simulation model presented here integrates the several key data sources and algorithms mentioned above to predict wildfire behavior under varying environmental conditions. Data collection and preparation involve terrain, elevation, moisture levels, and fire perimeters, while the simulation models fire ignition, spread, and burnout through probabilistic algorithms. This section goes into detail about the improvements made from the initial model.

4.2.1 Methodology

The simulation is still structured around a Monte Carlo (MC) method, running multiple iterations to produce a range of potential fire outcomes. By simulating numerous fires with varying initial conditions—such as different ignition points, weather, and moisture levels—the model can capture the inherent randomness of fire behavior.

Elevation data is sourced from the USGS 3DEP product. This data is used to derive critical terrain characteristics such as slope and aspect. Slope, which measures the steepness of the terrain, and aspect, which identifies the dominant cardinal direction a slope faces, are vital in predicting fire spread rates.

The maxima score, another terrain feature, is calculated using a diagonal-inclusive Laplacian kernel as opposed to a categorization based on neighboring cells like before. This score helps identify prominent features in the landscape, such as the height of trees relative to their surroundings. The maxima score is particularly relevant for modeling ignition probabilities since lightning tends to strike taller objects, making trees with high maxima scores more likely ignition points.

The terrain data is derived from the National Land Cover Database (NLCD), which categorizes different land cover types. Different vegetation categories—such as forests, grasslands, and urban areas—exhibit different fire behaviors. Forested areas, with dense tree coverage and leaf litter, provide substantial fuel, leading to higher fire intensities and longer burn durations. In contrast, grasslands may ignite more quickly but burn out sooner due to lower fuel loads. This categorization allows the model to assign base moisture levels, moisture of extinction, and base spread probabilities to each terrain type. Vegetation type directly impacts the likelihood of ignition and how quickly fires spread once they start.

Two temporal random variables – season and rainfall – are created at simulation time to reflect environmental variability and allow for easier risk analysis. The season influences baseline conditions such as temperature and fuel moisture, while rainfall is dynamically updated during the simulation, affecting moisture levels. Both temporal variables interact with the spatial parameters to determine the likelihood of ignition and fire spread.

The spread model has been completely overhauled based on the Rothermal model - moisture plays a crucial role in the simulation, as it directly influences the probability of a fire starting and spreading. The model calculates moisture as a function of baseline terrain moisture, season, and rainfall. Each cell in the simulation grid is assigned a moisture level, and this level is dynamically updated based on weather changes during the simulation. Additionally, each terrain type has a pre-determined moisture of extinction, representing the point at which the fuel is too wet to sustain a fire. These values directly impact the fire's catch/spread probability, which dictates whether a cell ignites when exposed to flames.

NLCD_Category	Terrain_Type	Moisture_base	Moisture_of_Ext	Spread_Prob_base
11	Open Water	100	1	0
12	Perennial Snow/Ice	90	1	0
21	Developed, Open Space	15	20	50
22	Developed, Low Intensity	12	20	40
23	Developed, Medium Intensity	10	20	40
24	Developed, High Intensity	10	20	40
31	Barren Land	5	10	65
41	Deciduous Forest	20	35	80
42	Evergreen Forest	25	35	85
43	Mixed Forest	22	35	82
52	Shrub/Scrub	10	20	75
71	Herbaceous	10	25	70
81	Pasture/Hay	8	25	90
82	Cultivated Crops	20	30	50
90	Woody Wetlands	50	40	25
95	Emergent Herbaceous Wetlands	55	40	25

Figure 4.5: Base moisture, moisture of extinction, and base spread probability for all terrains

Values from this table are mostly derived from the Rothermal documentation. Others are taken from other sources mentioned in the literature review above. Rain and seasonality have impacts on moisture to start each iteration of the simulation (e.g. summer results in dryer climate, 3” rainfall results in moister climate). Below is an example of the dynamic nature of the moisture algorithm:

Given:

Terrain Type: Evergreen forest (NLCD 42)

Season: Summer

Recent Rainfall: 1 inch

Step 1: Calculate Moisture Level

$$ML_{base} = 25\%$$

$$SAF = -0.1 \quad \text{(assuming summer dries out the terrain)}$$

$$RA = 1 \times 3 = 3\%$$

$$ML = 25 + (25 \times -0.25) + 3 = 25 - 6.25 + 3 = 21.75\%$$

Step 2: Calculate Moisture Damping Coefficient

$$Mx = 35\%$$

$$\eta = \frac{21.75}{35} = 0.621$$

Step 3: Calculate Reaction Intensity Reduction

$$I = I_0 \times (1 - 0.621) = I_0 \times 0.379$$

Figure 4.6: Moisture algorithm (spread probability modifier is calculated as 0.379)

Historical fire perimeter data provides information on fire size distribution, shape, and the types of vegetation burned in past fire events. This historical data is used to tune the simulation model and ensure it accurately reflects real-world fire behavior. By analyzing the size and shape of past fires, the model refines its spread algorithms to predict similar behavior under analogous conditions. Specifically, convex hulls were utilized here – making sure large historic and large simulated fires had similar ratio between their areas and the areas of their convex hulls.

4.2.2 Runtime

The ignition, spread, and burnout of fires are driven by both spatial and temporal variables. Each simulation run starts by selecting ignition points probabilistically, influenced by the maxima score, which reflects the likelihood of lightning strikes. Cells with higher maxima scores, typically taller trees or elevated terrain, are more likely to be selected as ignition points.

The spread algorithm operates on a cellular automaton framework, where each cell can be in one of three states: not burning (0), burning (1), or burnt (0.9). The first iteration of each simulation ignites the starting cell, after which the fire spreads to neighboring cells in subsequent iterations. Fire spread is influenced by several factors:

1. Stop Probability: This parameter determines whether the fire will continue spreading or if the simulation should stop at a given point. If the fire reaches areas with insufficient fuel or moisture levels, the simulation may stop earlier than anticipated.
2. Burnout Probability: For cells that are burning, the burnout probability determines whether they transition to the burnt state. This probability is influenced by the cell's moisture level and the burnout prob parameter. Moist cells are less likely to burn completely, and the fire may smolder without fully transitioning to the burnt state.

3. Spread Probability: For non-burning cells adjacent to burning cells, the spread probability determines whether the fire will spread to those cells. This probability is influenced by terrain type, slope, and moisture levels. Cells on the diagonals are not considered neighboring cells for spread purposes; only cells in the four cardinal directions are evaluated. The model checks each adjacent cell's terrain type and moisture level to determine whether the fire will propagate into that cell.

The result of these combined processes is a realistic simulation of fire spread that reflects how wildfires move through different terrains and under different environmental conditions.

The model provides several outputs that capture the results of each simulation run. The primary outputs include:

1. Raster Maps showing the cumulative spread of all wildfires over multiple runs.
2. Length Maps that quantify the amount of damage caused by the fire, reflecting the fire's final perimeter after it stops spreading.
3. Time Maps indicating the number of iterations it took for the fire to spread.
4. Ignition Maps highlighting the ignition points for all simulation runs.
5. Temporal variables (season, rainfall) vs. Wildfire spread.
6. Terrain vs. Wildfire spread.

4.2.3 Results

The simulation was conducted over two geographic areas – one a smaller subset of Mount Rainier National Park, and two the entire national park. As the simulation was also made more efficient by optimizing data types and loop behaviors, tens of thousands of runs were able to be completed in each simulation in the smaller subset. Note the patterns around rolling hillsides and lush areas:

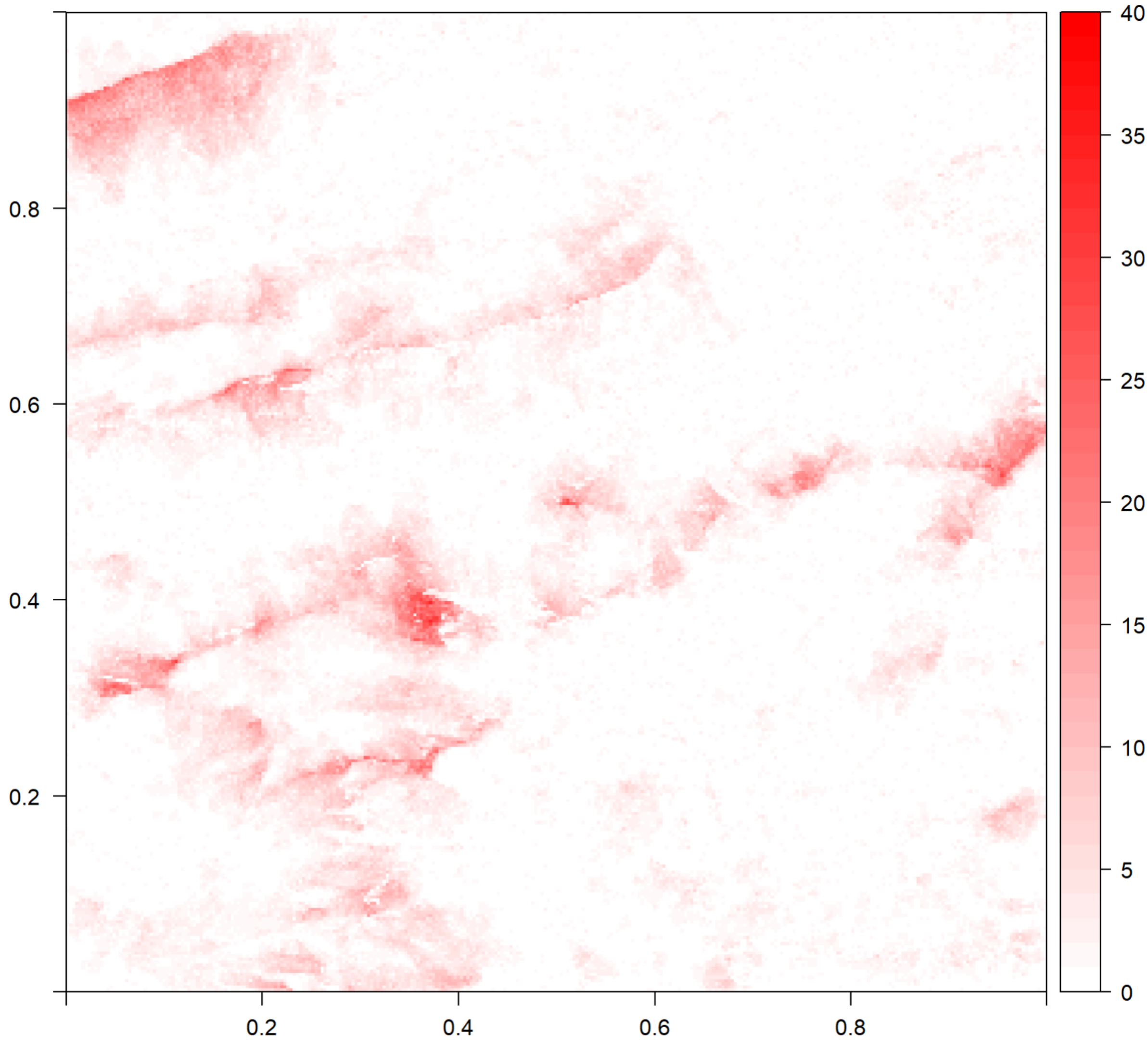

Figure 4.7: Cumulative damage from n = 56K simulated wildfires

Even in the much larger national park bounds, many thousands of runs were able to be completed in each simulation. Note the pattern around the perimeter of Mount Rainier:

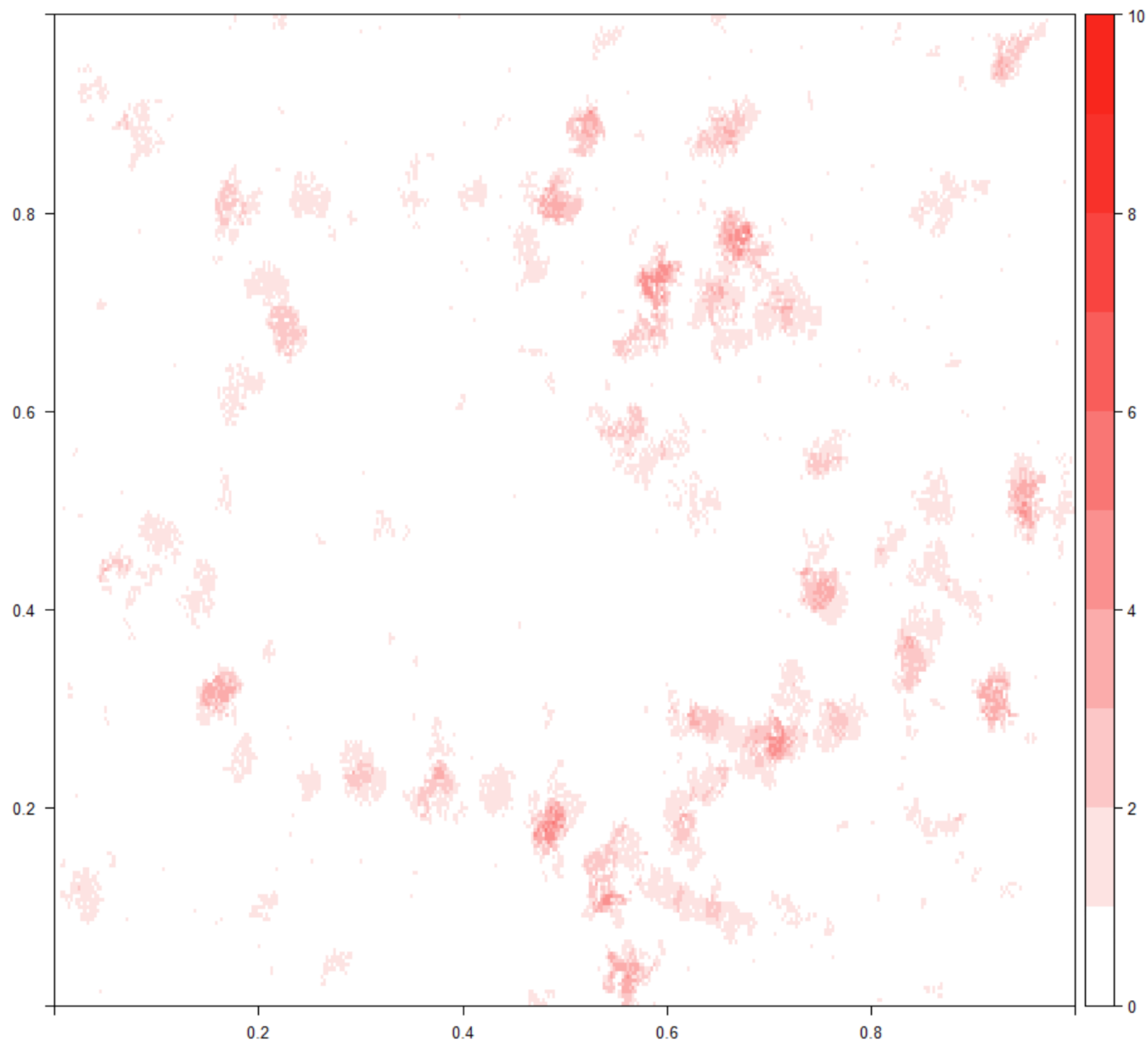

Figure 4.8: Cumulative damage from n = 7K simulated wildfires

Wildfire damage resembled a distribution similar to that of historic fire areas. A cutoff was imposed at 100K cells due to computational restrictions.

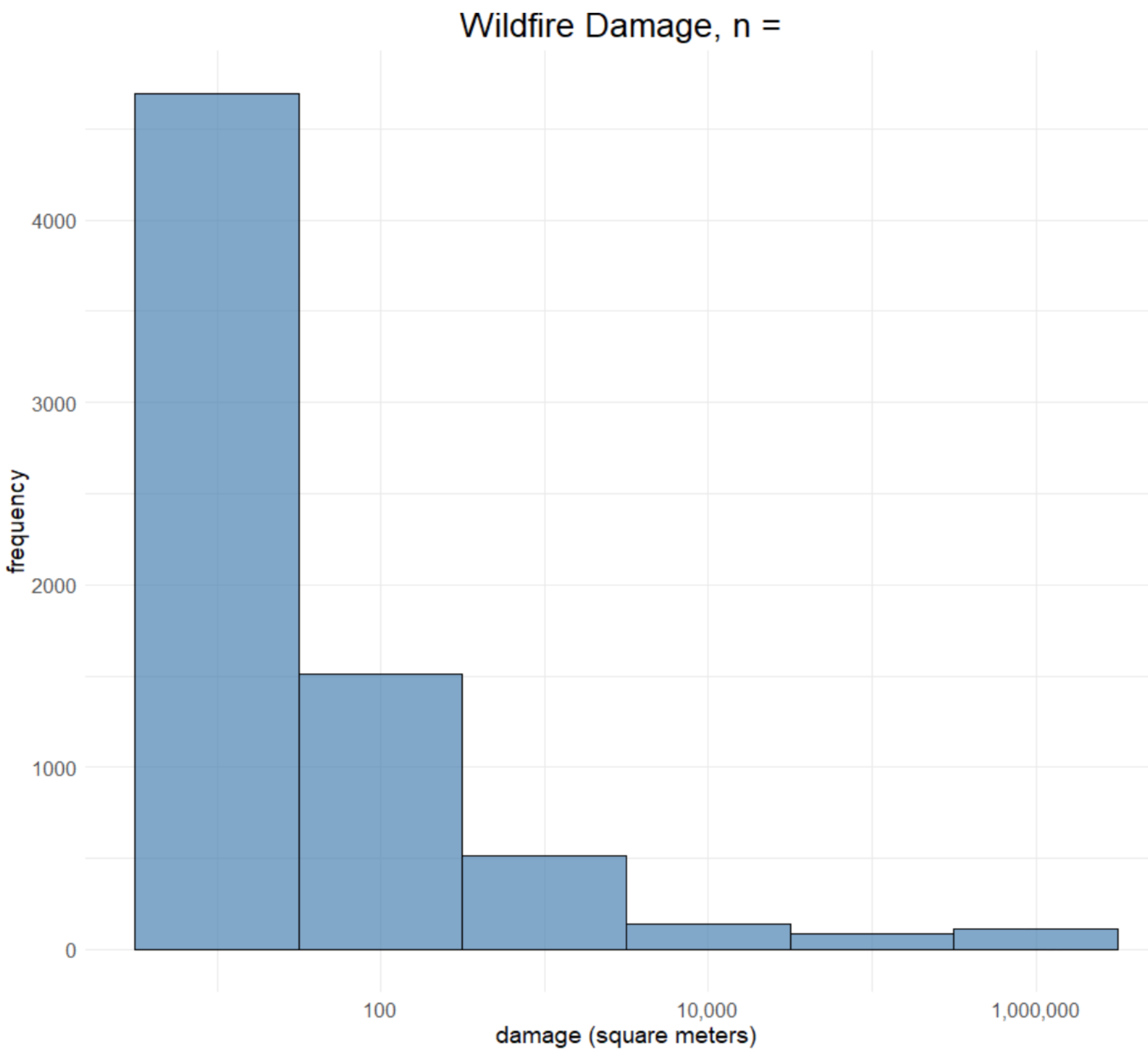

Figure 4.9: Distribution of damage

Summer dominated the average fire damage average.

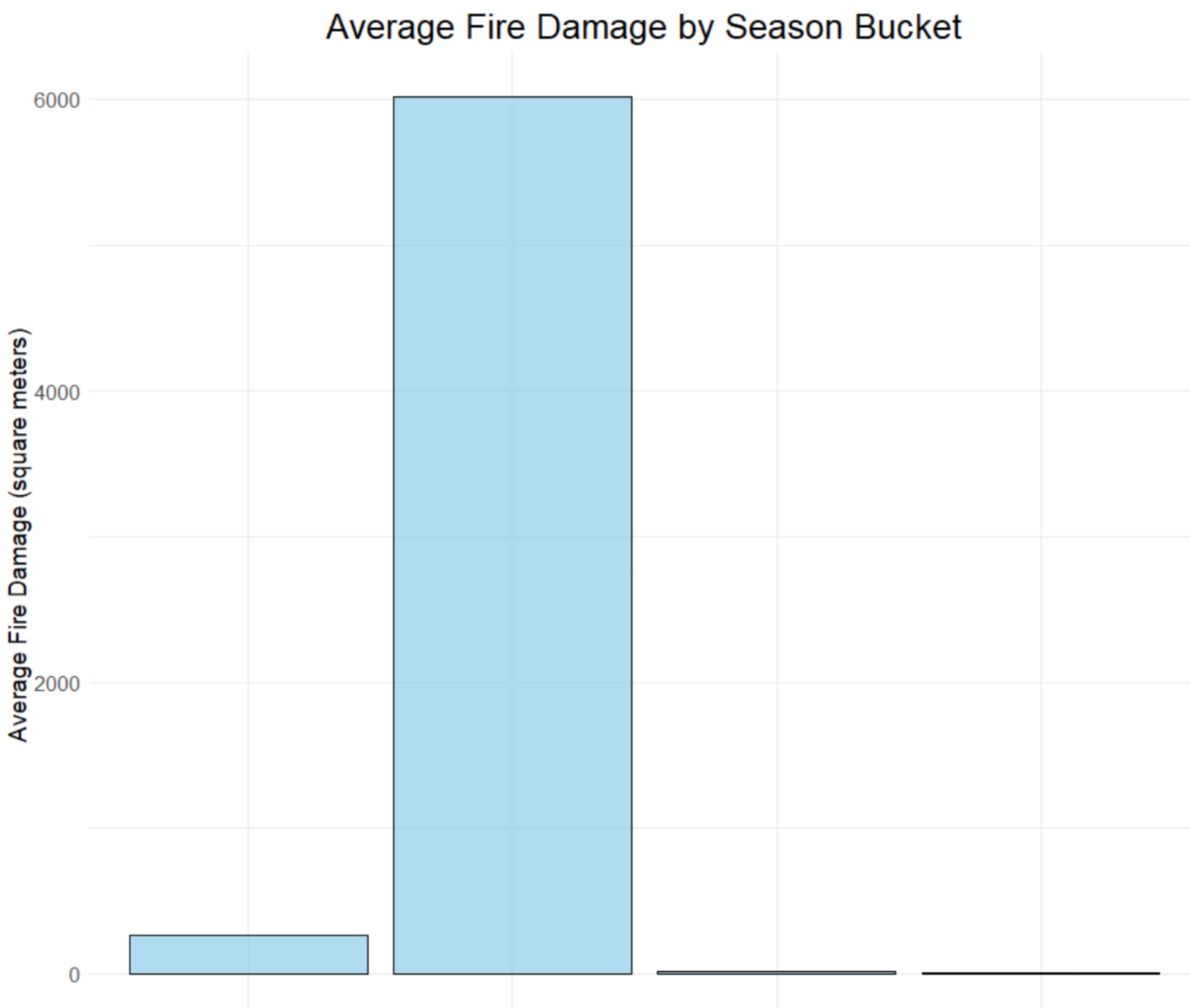

Figure 4.10: Seasonality (spring/**summer**/fall/winter) vs. Damage

Zero rainfall dominated the average fire damage average.

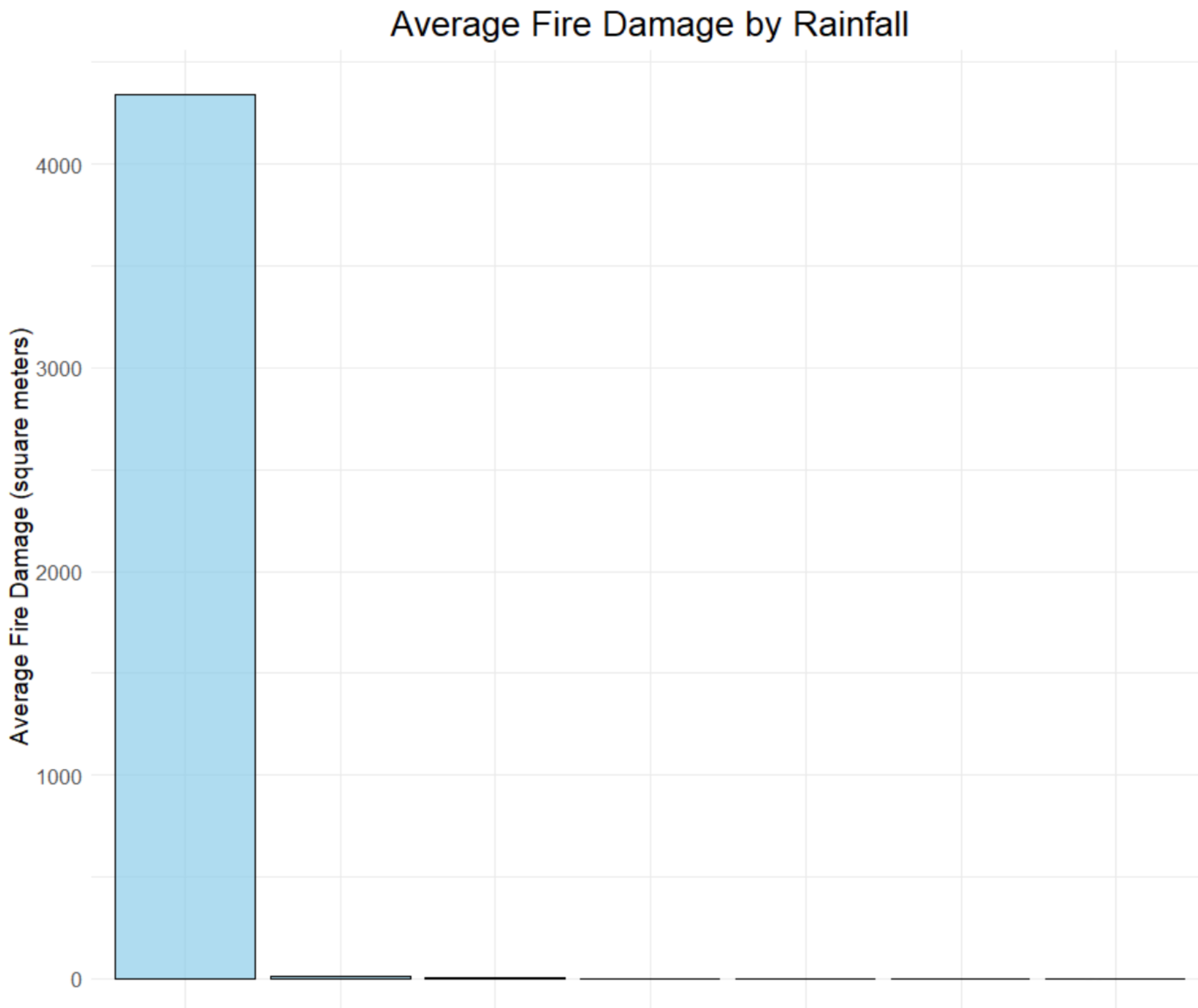

Figure 4.11: Rainfall vs. Damage

APPENDIX

```
### Nikhil Auradkar
### Simulating Wildfires in Mt. Rainier National Park
### 2024

#################################################################################

## Theoretical packages to use for plots

install.packages(c("terra", "raster", "rasterVis", "jsonlite"))
install.packages(c("magick", "jpeg", "imager", "spatstat", "ggplot2"))

library(terra)
library(raster) # working with GeoTIFF and IMG
library(rasterVis)
library(jsonlite)

library(magick)
library(jpeg) # JPEG terrain
library(imager)
library(spatstat)
library(ggplot2)

#################################################################################

## ELEVATION and TERRAIN

# Establishing borders of analysis area
# TESTING(-121.875, -121.825, 46.775, 46.825)
# SIMULATION(-121.910, -121.550, 46.750, 46.990)
lat_min <- 46.750
lat_max <- 46.990
lon_min <- -121.910
lon_max <- -121.550

lat <- 46.870
lon <- -121.730

# 3DEP NED image from USGS
ned_elevation_1 <- "C:/Users/aurni/OneDrive/Documents/UCLA/Research - Dr. Handco
ned_elevation_2 <- "C:/Users/aurni/OneDrive/Documents/UCLA/Research - Dr. Handco

# National Land Cover Data (NLCD) terrain image
nlcd_terrain <- "C:/Users/aurni/OneDrive/Documents/UCLA/Research - Dr. Handcock/

# Reading 3DEP NED images of West Rainier, 1/9 arcsecond granularity (3 meters)
myImage1 <- raster(ned_elevation_1)
myImage2 <- raster(ned_elevation_2)
myImage <- merge(myImage1, myImage2)
ext_subset <- extent(lon_min, lon_max, lat_min, lat_max)
westRainier_elev <- crop(myImage, ext_subset)
dim_elev <- dim(westRainier_elev)
plot(westRainier_elev)
```

```r
rm(myImage1, myImage2, myImage)

# Reading NLCD terrain image of North America
myTerrain <- raster(nlcd_terrain)
ext_subset_terr <- extent(-1975000, -1915000, 2890000, 2950000) # slightly larger e
westRainier_terr <- crop(myTerrain, ext_subset_terr)
westRainier_terr_latlon <- projectRaster(westRainier_terr, crs = "+proj=longlat +da
westRainier_terr_latlon <- crop(westRainier_terr_latlon, ext_subset)
dim_terr <- dim(westRainier_terr_latlon)
westRainier_terr_latlon_resampled <- resample(westRainier_terr_latlon, westRainier_
plot(westRainier_terr_latlon_resampled)

rm(myTerrain, ext_subset, ext_subset_terr, westRainier_terr, westRainier_terr_latlo

rm(ned_elevation_1, ned_elevation_2, nlcd_terrain)

##############################################################################

## TERRAIN CONSTANTS and SAF

# Table w/ NLCD categories, base moisture and moisture of extinction levels, and ba

terrain_info <- data.frame(
  NLCD_Category = c(11, 12, 21, 22, 23, 24, 31, 41, 42, 43, 52, 71, 81, 82, 90, 95)
  Terrain_Type = c("Open Water", "Perennial Snow/Ice", "Developed, Open Space", "De
                   "Developed, Medium Intensity", "Developed, High Intensity", "Bar
                   "Evergreen Forest", "Mixed Forest", "Shrub/Scrub", "Herbaceous",
                   "Woody Wetlands", "Emergent Herbaceous Wetlands"),
  Moisture_base = c(100, 90, 15, 12, 10, 10, 5, 20, 25, 22, 10, 10, 8, 20, 50, 55),
  Moisture_of_Ext = c(1, 1, 20, 20, 20, 20, 10, 35, 35, 35, 20, 25, 25, 30, 40, 40)
  Spread_Prob_base = c(0, 0, 50, 40, 40, 40, 65, 80, 85, 82, 75, 70, 90, 50, 25, 25
  Moisture = numeric(16)

)

# SAF - seasonal adjustment factors (moisture changes based on season)
# Reference Table

seasonal_adjustment_factors <- list(
  Spring = -0.05,
  Summer = -0.2,
  Fall = 0.05,
  Winter = 0.15
)

##############################################################################
```

```
## MAXIMA SCORING and START PROBABILITIES

# Convolution with uniform 5x5 Laplacian kernel to generate maxima score
conv_kernel <- matrix(-1, nrow=5, ncol=5)
conv_kernel[3, 3] <- 24

maxima_score <- focal(westRainier_elev, w=conv_kernel, pad=TRUE, na.rm=FALSE)
maxima_score <- reclassify(maxima_score, cbind(NA, 0))
maxima_score <- maxima_score / 24

# Starting point probability calculations
abs_diff_vec <- as.vector(maxima_score)
abs_diff_vec[abs_diff_vec < 0] <- 0
start_probs <- abs_diff_vec / sum(abs_diff_vec[1:length(abs_diff_vec)])

rm(conv_kernel, abs_diff_vec)

##############################################################################

## SLOPE/ASPECT SCORING

# Calculate slope and aspect
slope <- terrain(westRainier_elev, opt='slope', unit='degrees')
aspect <- terrain(westRainier_elev, opt='aspect', unit='degrees')

##############################################################################

## COMBINING DATA

# Removing NAs and NaNs
slope[is.nan(slope)] <- 0
aspect[is.nan(aspect)] <- 0

westRainier_terr_latlon_resampled[is.na(westRainier_terr_latlon_resampled)] <- 0

# Combining data (elevation, terrain) and derived features (maxima score, slope, aspe
westRainier <- stack(westRainier_elev, westRainier_terr_latlon_resampled,
                     maxima_score, slope, aspect)

rm(westRainier_elev, westRainier_terr_latlon_resampled, maxima_score, slope, aspect)


##############################################################################
##############################################################################
##############################################################################
```

```
## SIMULATION PREP

# Programmable parameters
iter <- 10000 # number of times the loop is run
spread_prob_modifier <- 1.5
burnout_prob <- 0.15
stop_prob <- 0.0001

rows <- dim_elev[1]
cols <- dim_elev[2]

master_wildfire_spread <- array(0L, dim = c(rows, cols)) # cumulative spread area
master_wildfire_start <- array(0L, dim = c(rows, cols)) # cumulative start points
master_wildfire_lengths <- c()
master_wildfire_times <- c()

# Generating RVs - season and rainfall
season <- runif(iter)
rainfall <- rpois(iter, 1)

# Selecting starting points using sample
start_index <- sample.int(length(start_probs), size = iter, prob = start_probs, replac

# Optimizing by converting start_index to row and column once and storing it
start_indices <- lapply(start_index, function(idx) {
  list(row = (idx - 1) %% rows + 1, col = (idx - 1) %/% rows + 1)
})

# Pre-calculate moisture levels
season_moisture <- ifelse(season < 0.25, 0.9, ifelse(season < 0.50, 0.75, ifelse(seaso


##################################################################################
```

```
## SIMULATION

# Precompute neighbor offsets to avoid recalculating every time [1st Change]
neighbor_offsets <- matrix(c(-1, 0, 1, 0, 0, -1, 0, 1), ncol = 2, byrow = TRUE)

for (a in 1:iter) {
  print(paste("ITERATION", a))
  wildfire_spread <- array(0L, dim = c(rows, cols))

  # Getting starting point
  start_cell <- c((start_index[a]-1) %% rows + 1,
                  (start_index[a]-1) %/% rows + 1)

  master_wildfire_start[start_cell[1], start_cell[2]] <- master_wildfire_start[start_ce
  wildfire_spread[start_cell[1], start_cell[2]] <- 1

  # Calculate moisture values for the current iteration
  terrain_info$Moisture <- pmin(100, terrain_info$Moisture_base * season_moisture[a] +

  ## Determining spread
  total_damage <- 1  # Set to 1 to include the initial fire [2nd Change]
  total_time <- 0
  print(paste("Total damage", total_damage))

  # Create a list to track active cells on fire [3rd Change]
  active_fire <- list(start_cell)

  while (length(active_fire) > 0 && (runif(1) > stop_prob) && total_damage < 100000) {
    total_time <- total_time + 1
    new_fire <- list()  # Track newly spread cells in this iteration

    # Spread algorithm - iterating through all cells on fire
    for (i in seq_along(active_fire)) {
      current_cell <- active_fire[[i]]
      x <- current_cell[1]
      y <- current_cell[2]

      # Spontaneous fire burnout
      if (runif(1) < burnout_prob) {
        wildfire_spread[x, y] <- 0.9
        next
      }

      terrain <- westRainier[x, y][2]
      spread_probability = spread_prob_modifier *
        (terrain_info$Spread_Prob_base[terrain_info$NLCD_Category == terrain]) *
        (terrain_info$Moisture_of_Ext[terrain_info$NLCD_Category == terrain] - terrain_
        (terrain_info$Moisture_of_Ext[terrain_info$NLCD_Category == terrain]) / 100
      spread_probability = pmin(spread_probability, 1)

      # Get neighbors using precomputed offsets [1st Change]
```

```
    # Get neighbors using precomputed offsets [1st Change]
    neighbors <- neighbor_offsets + matrix(c(x, y), nrow=4, ncol=2, byrow=TRUE)

    # Check boundaries and valid neighbors
    valid_neighbors <- neighbors[neighbors[,1] >= 1 & neighbors[,1] <= rows & neighb

    for (j in 1:nrow(valid_neighbors)) {
      neighbor_x <- valid_neighbors[j, 1]
      neighbor_y <- valid_neighbors[j, 2]

      # Checking if neighbor is already on fire or burnt out
      if (round(wildfire_spread[neighbor_x, neighbor_y]) == 1) {
        next
      }

      # Updating spread probability based on slope and aspect
      current_slope <- westRainier[x, y][4]  # Slope of current cell
      current_aspect <- westRainier[x, y][5]  # Aspect of current cell
      neighbor_slope <- westRainier[neighbor_x, neighbor_y][4]  # Slope of neighbor
      neighbor_aspect <- westRainier[neighbor_x, neighbor_y][5]  # Aspect of neighbor

      slope_diff <- neighbor_slope - current_slope
      if (slope_diff > 0) {
        spread_probability <- pmin(1, spread_probability * (1 + slope_diff / 90))
      } else {
        spread_probability <- pmax(0, spread_probability * (1 + slope_diff / 180))
      }

      aspect_diff <- abs(current_aspect - neighbor_aspect)
      if (aspect_diff > 180) {
        aspect_diff <- 360 - aspect_diff
      }
      aspect_modifier <- 1 + (aspect_diff / 180) * -0.2
      spread_probability <- spread_probability * aspect_modifier

      ## Check if fire spreads to the neighboring cell
      if (is.null(spread_probability) || is.na(spread_probability) || length(spread_
        spread_probability <- 0
      }
      if (runif(1) < spread_probability) {
        wildfire_spread[neighbor_x, neighbor_y] <- 1
        new_fire <- append(new_fire, list(c(neighbor_x, neighbor_y)))  # Add to new 
      }
    }

    wildfire_spread[x, y] <- 0.9  # Mark the cell as burnt out [5th Change]
  }
```

```
    # Update the total damage directly without recalculating [2nd Change]
    total_damage <- total_damage + length(new_fire)
    print(paste("     TOTAL DAMAGE", total_damage))

    # Update active_fire for the next iteration
    active_fire <- new_fire
  }

  master_wildfire_lengths <- c(master_wildfire_lengths, total_damage)
  master_wildfire_times <- c(master_wildfire_times, total_time)
  master_wildfire_spread <- master_wildfire_spread + round(wildfire_spread)
}
```

BIBILIOGRAPHY

Andrews, Patricia L. *The Rothermel Surface Fire Spread Model and Associated ...*, United States Department of Agriculture, Mar. 2018, www.fs.usda.gov/rm/pubs_series/rmrs/gtr/rmrs_gtr371.pdf.

Carmel, Yohay, et al. "(PDF) Assessing Fire Risk Using Monte Carlo Simulations of Fire Spread." *ResearchGate*, Israel Institute of Technology, 2009, www.researchgate.net/publication/235990366_Assessing_fire_risk_using_Monte_Carlo_simulations_of_fire_spread.

Coder, Kim D. *Lightning Strike Probability Risks*, University of Georgia, 2022, bugwoodcloud.org/resource/files/25278.pdf.

CRC, Tropical Savannas. "Topography." *Fire Fundamentals: Topography*, Charles Darwin University of Australia, 2024, learnline.cdu.edu.au/units/env207/fundamentals/topography.html.

"Current Wildfire Incident Information: WA - DNR." *Washington State Department of Natural Resources*, 2024, www.dnr.wa.gov/Wildfires.

Koh, Jonathan, et al. "SPATIOTEMPORAL WILDFIRE MODELING THROUGH POINT PROCESSES WITH MODERATE AND EXTREME MARKS." *Arxiv*, Institute of Mathematics, EPFL, 2021, arxiv.org/pdf/2105.08004.pdf.

Radke, David, et al. "FireCast: Leveraging Deep Learning to Predict Wildfire ... - IJCAI." *IJCAI*, David R. Cheriton School of Computer Science, University of Waterloo, 2019, www.ijcai.org/Proceedings/2019/0636.pdf.

USGS - U.S. Geological Survey. *Datasets (3DEP, NLCD)*, 2024, apps.nationalmap.gov/datasets/.

USGS - U.S. Geological Survey. *EarthExplorer*, 2024, earthexplorer.usgs.gov/.

"Wildfire Causes and Evaluations (U.S. National Park Service)." *National Parks Service*, U.S. Department of the Interior, 2017, www.nps.gov/articles/wildfire-causes-and-evaluation.htm.

"Wildfire Investigations." *Wildfire Investigations | Indian Affairs*, 2002, www.bia.gov/service/wildfire-prevention/wildfire-investigations.

"Wildland Fire Information." *National Parks Service*, U.S. Department of the Interior, 2023, www.nps.gov/mora/learn/news/fire.htm.