a. Our objective is to simulate the spread of forest fires with different wind conditions, providing insights into fire behavior. The main task is to identify high-risk areas within the forest that are most likely to be affected by or adjacent to a burnt-out grid cell. We will select around 40 km squared area (150x150 grid) from the real-world data and proceed with our simulation to find out which parts of the selected area need extra protection from forest fire.

From our simulation, we will be able to measure the number and location of cells burnt in each simulation, identification of grid cells with the highest risk of being burnt or adjacent to burnt cells, the impact of wind strength and density, and simulation variability which is the analysis of fire spread under random wind conditions to understand the variability in fire behavior.

b. The forest is modeled as a two-dimensional grid, with each cell representing a specific area and each cell having its density which is a value between 0 and 1 (0 representing a cell with no trees and 1 representing a cell full of trees). Each cell can be in one of three states: unaffected, burning, or burnt. Each simulation run uses randomized wind direction and strength to mimic the variability of real-world conditions. Fire spreads from burning cells to adjacent cells based on tree density and wind conditions using the Moore neighborhood approach.

Tree Density and Fire Spread Probability:

Each cell in the grid has an associated tree density value, which represents the amount of fuel available for the fire. This value is used as the base probability of the cell catching fire. A higher tree density means a higher probability of the fire spreading to that cell. The simulation incorporates wind direction and strength to realistically model how wind influences the spread of fire. The wind can increase the likelihood of the fire spreading in its direction.

In each simulation step, the algorithm checks each burning cell (cells in state 1) and evaluates the probability of fire spread to its adjacent cells (using the Moore neighborhood, which includes diagonals). For each adjacent cell, the algorithm calculates the spread probability considering both the tree density of that cell and the wind effect. A random value is then generated. If this value is less than the calculated probability, the adjacent cell is set to a burning state in the next simulation step.

c. Tree density in each cell is assumed constant throughout a simulation run. Keeping tree density constant simplifies the model, making it easier to focus on other variables like wind and the spread of the fire itself.

The assumption that tree density in each cell remains constant throughout a simulation run means that the density of trees in any given cell does not change as the simulation progresses. In other words, regardless of whether the cell burns or not, the value representing the density of trees in that cell remains the same from the start to the end of the simulation.

Wind direction and strength are uniform across the grid but vary between different simulation runs. The fire starts at a random place but will only start at a cell with non-zero density. However, if the density is 0.0001 which is non-zero, the fire will still start but it will not spread that far. This could influence our results but we could run it as many times as possible to reduce its effect. The model does not account for external factors like topography, human intervention, or changing weather conditions other than wind. This would be invalid in a real-world scenario since there are many more factors that contribute to the scenario.

These assumptions simplify the complex nature of forest fires, making the model suitable for educational and basic predictive analysis. However, it may not fully capture the intricacies of real-world forest fires.

d. The parameters:

- Grid Size (150x150): Defines the size of the simulated forest area.
- Tree Density Grid: Determines the likelihood of fire spreading in each cell.
- Wind Direction and Strength: Randomized per run, these parameters influence the direction and rate of fire spread.
- Number of Simulation Steps: Dictates how long the simulation observes the fire's progression.

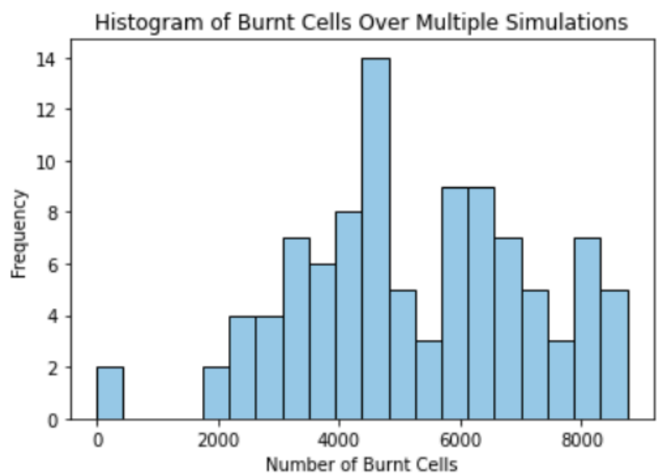
e. Identified through multiple simulations, we will identify the high-risk areas on the grid. These areas have a higher frequency of burning or being adjacent to burnt cells. This output helps in strategic planning for fire prevention measures by pinpointing where resources and efforts should be concentrated. The most direct measure of the fire's impact is the number of burnt cells. It represents the total area affected by the fire within the simulation grid.

After finding out the high-risk areas, we will implement different fire-protection strategies. Beyond just the average number of burnt cells, the frequency distribution of burnt cells would provide more insight into the consistency and reliability of each strategy.

Lastly, we will find the 95% confidence intervals for each strategy which provide a statistical measure of the reliability of the simulation results. Narrower intervals would suggest that we can be more confident in the average outcomes predicted by the simulation.

Interpretation

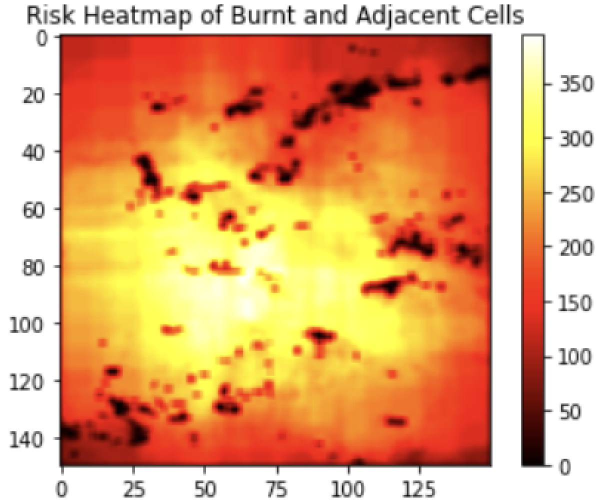
a.



The histogram illustrates the distribution of burnt cells across multiple simulations of our model. As we can see, there are several peaks that show a multi-modal distribution. The tallest peak occurs between 4000 and 5000 burnt cells, indicating that this is the most frequently occurring range of outcomes in our simulation. There are also noticeable peaks in other ranges, suggesting that the fire spread can vary quite a bit.

The spread of the distribution, with significant numbers of simulations leading to both lower and higher numbers of burnt cells, indicates variability in fire spread. This could be due to the stochastic nature of the simulation.

We just looked at the histogram that shows the frequency of the number of burn cells after 100 simulations. Now, let’s look at the high-risk area after 100 simulations.



The heatmap shows the frequency of cells being burnt or adjacent to burnt cells across multiple simulations. The color gradient, ranging from dark to bright, indicates low to high frequency.

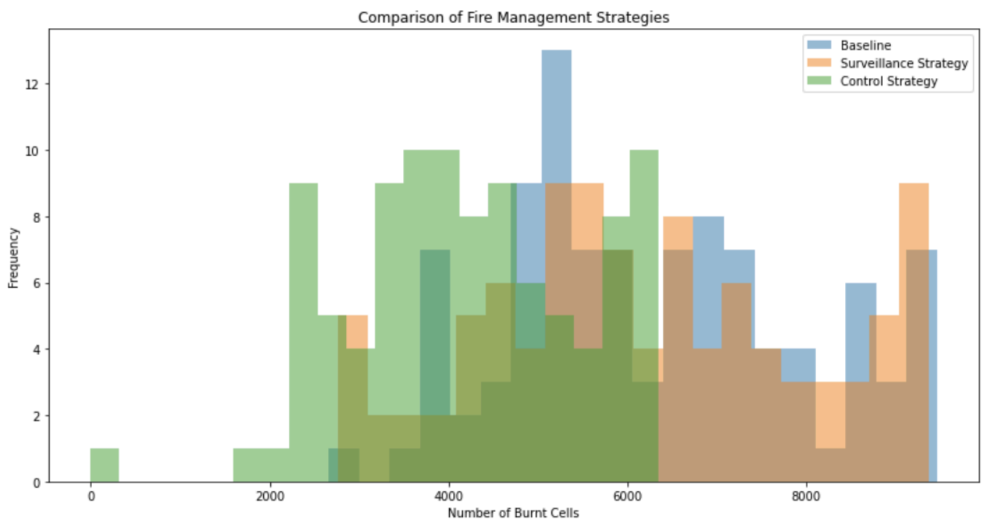
Brighter areas represent parts of the forest that are most susceptible to fire or being near fire. These could be areas with higher tree density, or they might be more exposed to prevailing winds in the simulation. From the graph, the area in the highest risk would be an area that is bordered by an x-axis: of 25 to 75 and a y-axis: of 60 to 100 since it shows the brightest color.

b. For this part of the assignment, I chose the following two strategies:

1. Enhanced surveillance in high-risk areas. This approach focuses on monitoring and rapid response in areas identified as high-risk based on simulation results. With this strategy, we would have quicker detection and response to fires in high-risk areas, potentially reducing the overall impact.

2. Preemptive control measures on high-risk areas. This strategy creates firebreaks in high-risk areas before the "fire season". It reduces the available fuel in high-risk areas. In other words, we cut down some trees to decrease the density, potentially stopping or slowing down the spread of fires.

I implemented the strategies and ran them 100 times to see the comparison between the two strategies and the scenario without any intervention.



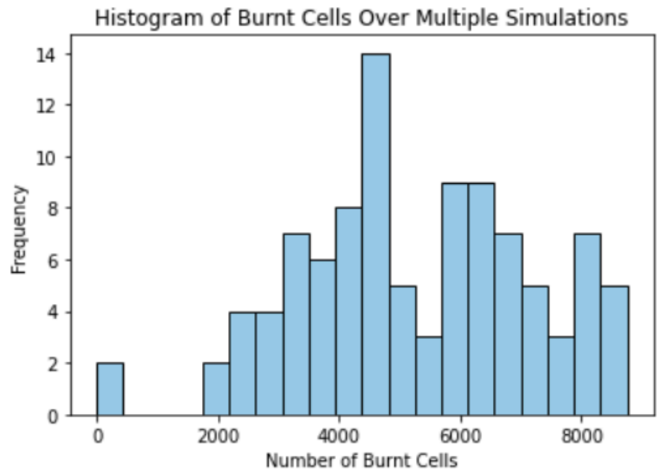
Baseline Average Burnt Cells: 6311.65

Surveillance Strategy Average Burnt Cells: 6236.1

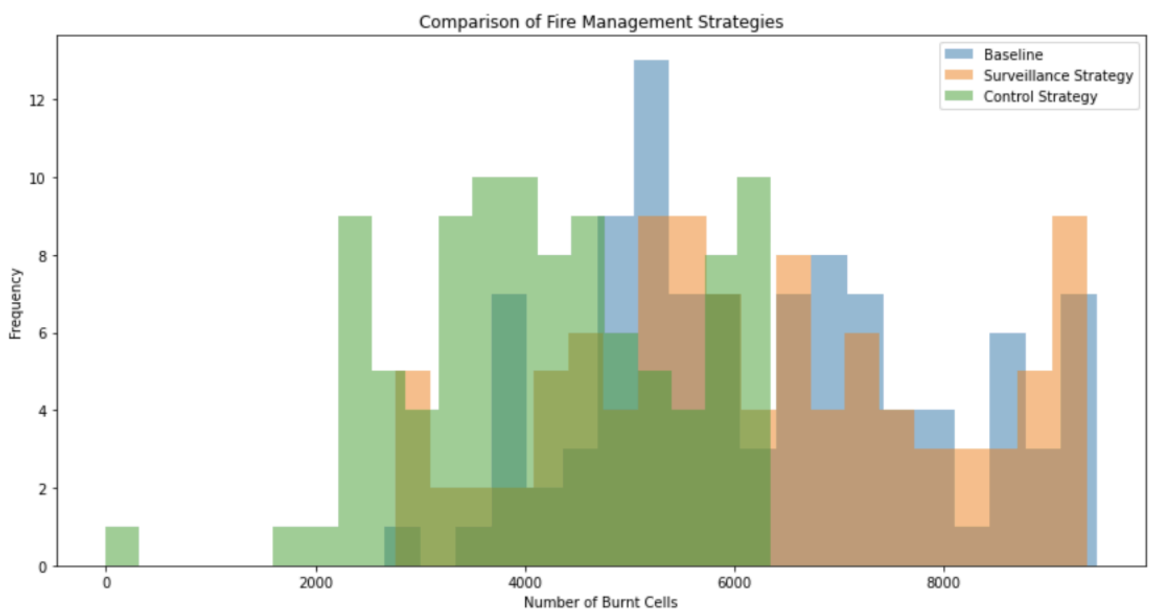
Control Strategy Average Burnt Cells: 4189.04

As we can see, the Surveillance strategy does not differ that much from the case where there is no strategy. On the other hand, the Control strategy seems to be working much better.

c.



The histograms tell us that without any protection plan, the average number of burnt cells is around 5000-6000. From the histogram “Comparison of Fire Management Strategies” below, we can see that the Control Strategy gives a different result from the other two results.



Baseline: This represents the scenario without any intervention, providing a control group to understand the natural spread of fire.

Surveillance Strategy: This would involve increased monitoring and rapid response measures to detect and combat fires early.

Control Strategy: This means measures such as controlled burns, cutting trees off, or creating firebreaks to prevent the spread of fire.

Note that the control was only put around the highest-risk areas. If we use the control strategy in more areas, the number of burnt cells would decrease significantly.

d. The data clearly indicates that the control measures, which include creating firebreaks and conducting controlled burns, are significantly more effective in reducing the number of burnt cells. This strategy has resulted in lower average numbers of burnt cells, showing effectiveness despite varying wind conditions. On the other hand, while the enhanced surveillance strategy might have shown some benefits in terms of early fire detection and response, its impact on reducing the spread of fires is small.

Given these findings, our recommendation is to prioritize the implementation of preemptive control measures as the primary strategy for forest fire management. Investing resources in creating strategic firebreaks and controlled burns in high-risk areas will likely result in a significant reduction of fire spread potential.

Analysis

a. Baseline Strategy: Mean = 6471.63, 95% CI = (6089.53, 6853.73)

Surveillance Strategy: Mean = 6697.94, 95% CI = (6356.91, 7038.97)

Control Strategy: Mean = 3865.15, 95% CI = (3698.02, 4032.28)

b. The confidence interval for baseline is [6089.53,6853.73], which is relatively wide, indicating a moderate level of uncertainty in the estimate of the mean. This variability might be due to the inherent randomness in the fire spread without any intervention.

The CI for surveillance strategy is [6356.91,7038.97], which is also quite wide, and interestingly, the mean is higher than the baseline strategy. This may suggest that the surveillance strategy, as implemented, is not effectively reducing the fire spread or that there is significant variability in its effectiveness.

Control Strategy: The CI is [3698.02,4032.28], which is narrower than the other two strategies. This narrower interval indicates a higher level of precision in the estimate and suggests that the control strategy is consistently more effective at reducing the number of burnt cells.

The width of the confidence interval is inversely proportional to the square root of the number of runs. Therefore, to make the interval narrower by a factor of a, we would need to increase the number of runs by a factor of a^2. For example, if we wanted to make the CI for the Control Strategy 10% narrower, we would need to increase the number of runs by $(1/0.9)^2$ ≈1.52, or 52% more runs. In other words, if we're running 100 simulations, we would need to run approximately 152 simulations to achieve this new width.