CS166 Final Project

Minerva University

CS166: Modeling and Analysis of Complex Systems

Dec 15, 2023

Simulation Description

The simulation models the spread of fire on a 2D grid representing a forest area. The goal of the simulation is to understand how various factors such as wind direction, and strength impact the fire spread as well as to find out which forest cells are most prone to catching fire regardless of the fire spread start. To achieve this goal, I will have to be able to measure the spread rate, which will help me evaluate how a certain parameter like wind strength impacts the rate at which the fire spreads through the system. I will also have to measure the overall directionality of the fire (or its “harm” on the system). This will help me to understand how much of an impact the wind directionality has on fire spread. To understand which forest areas are most likely to get burnt or be close to getting burnt regardless of the fire start location, I will have to measure the burn probability per cell and the average time it takes the given cell to burn. It will also be helpful to simply analyze the fire effects visually.

In regards to the rules of the simulation, I am modeling the forest fire on a 2D grid with every cell representing a forest area with the tree density obtained from the NASA dataset of global tree coverage. Using this resource adds realism to the simulation and considers real-life conditions such as non-uniform tree coverage. In my simulation, the fire can start from any point on the grid. Initially, I am using coordinates (70, 50) to create all the visualizations prior to starting to experiment with various starting positions.

Most important rules are incorporated into the spread mechanism within the simulation. More specifically, the fire spread probability is calculated considering the wind direction and strength as well as the tree density. This effectively helps to model the forest fire scenario because in nature areas with more vegetation are likely to burn more intensely due to more available fuel for the fire. By incorporating tree density as a factor in each cell, the simulation realistically models how some areas are more susceptible to fire than others. Similarly to the tree density factor, wind is of critical importance to fire spread in real life. In particular, wind carries the heat to new areas, thus accelerating the fire spread. The simulation incorporates both wind density and direction which captures how wind can influence the fire spread in real-life scenarios. Specifically, my simulation considers how much the wind direction is aligned with the location of the given cell relative to the fire cell. If the directions are aligned, then the fire is more likely to spread. If the wind is blowing in the opposite direction ("wind_factor" = 0), then the fire spread is not influenced by the wind, however, it remains under the impact of tree density and the base fire spread rate. As mentioned above, wind strength is considered in the calculation of the wind factor, which models real-life scenarios well because, in nature, a stronger wind is able to carry more heat and fire to adjacent territories than a lighter wind. The calculation also incorporates the base fire spread rate which aims to capture how the fire spreads in neutral conditions, which can be impacted by combustibility of vegetation and environmental conditions. The base spread rate aims to capture these factors and provide a "control" condition to enable comparisons given various parameters. Overall, the spread mechanism implemented in the simulation provides the basis on which the cell state is updated and effectively incorporates various real life aspects such as wind and tree density.

Mathematically, I've created the following spread probability formula:

$$P_s = b_r * d * (1 + w)$$

In the equation above, P_s is the spread probability, b_r is the base spread rate, d is the tree density factor, and w is the wind factor. The wind factor is added to 1 to account for situations where the wind factor is 0. In such cases, the fire spread will only depend on the inherent base spread rate and the tree density because while the wind wouldn't contribute to fire spread, there is still a chance it spread in the given direction.

The wind factor is calculated in the following way:

$$w = (w_d w_s).\bar{v}_{norm}$$

In the wind factor equation, w is the wind factor we are computing, w_d is the wind direction, which is multiplied by w_s, wind strength. $\bar{v}_{norm}$ is the normalized neighbor direction from the fire cell perspective, it is calculated in the following way:

$$\bar{v}_{norm} = \frac{\bar{v}}{||\bar{v}||},$$

where $\bar{v}$ is the neighbor direction in relation to the fire cell, from the fire cell view. $||\bar{v}||$ is the magnitude of the neighbor direction vector. While there are four cardinal neighbor vectors that have length of 1 and, therefore, do not need to be normalized, there are 4 diagonal directions which have the length of $\sqrt{2}$ and, hence, need to be normalized so that their magnitude does not impact the alignment calculation between the wind and the neighbor directions.

After the spread probability is calculated using the described formula, it is compared to a random number from 0 to 1. If the number falls below the spread probability, then the fire spreads to the nearby area. However, if the probability does not fall in that interval, the fire will not spread to the given neighboring cell. As mentioned above, such an update mechanism enables the simulation to capture key aspects of forest fires such as wind and tree density.

The neighborhood the simulation considers for fire spread is a Moore neighborhood, where we consider all 8 cells next to the current cell. This means that for every cell, the simulation calculates the probability to catch fire from all of the 8 neighboring cells given they are on fire. This approach models the real life situation more effectively than a Von Neumann neighborhood that does not consider diagonal directions. This is because in real forests, trees are surrounded by trees from every side, and are not specifically planted in a Von Neumann manner.

The simulation considers three states for an area, - alive trees with a given density, fire, or char. These three states allow the model to incorporate the realistic state transitions from a healthy tree to a burning tree to a dead tree.

However, it should be noted that one of the limitations of these states is that once the cell is on fire, the simulation considers the entire cell to be on fire, which represents a forest patch. This means the simulation cannot monitor the fire spread on a micro level of a single cell containing multiple trees. Hence, the fact that the simulation views a single cell as a smallest unit can be seen as one of its assumptions.

Another assumption the simulation makes is that the wind strength and direction are uniform throughout the grid. This assumption might not hold when we are considering a patch at the edge of a forest, where the trees growing at an intersection of a field and the forest are most susceptible to wind. However, the trees inside this patch can be more protected from the wind and, hence, the wind density there might be significantly less. The assumption is valid when the area we are considering is small and has a relatively uniform landscape, which would ensure that wind travel patterns are relatively stable over the given patch.

The simulation also assumes a simplified fire spread mechanism (described above), where the fire spread probability is mostly determined by the wind and tree density. In real-world scenarios, fire spread is influenced by a variety of other factors including topography, fire suppression efforts, and specific types of vegetation. However, this is a valid simplification if the simulation's goal is to study a general fire behavior.

Moreover, the simulation assumes only three states,- unburnt, burning, char. In real life, a patch of the forest might not burnt all at once and might have some smoldering, which is important to consider. However, the assumption can be seen as valid for a high level overview of general fire dynamics and behavior.

The simulation also assumes periodic boundary conditions, where the left/top end of the grid is connected to the right/down edge. This is not true in real life. However, this assumption allows one to understand how the fire would behave on a larger grid given similar landscape conditions. Moreover, periodic boundary conditions enable one to effectively assess how strong the fire actually is because it would not stop at the edge of the grid, allowing for a better contextual understanding of fire spread.

The simulation parameters are the base spread rate, which affects how quickly the fire spreads in neutral conditions without interference of wind or tree density. Another parameter is wind direction which controls which direction the wind is blowing. As described above, if the wind is blowing North, then the fire is more likely to spread North due to the acceleration provided by the wind. Wind strength parameter determines how intense the wind is. In the simulation, the wind ranges from 1 (mild) to 5 (strong). These values are arbitrary. Tree density is determined by the NASA coverage data. The more densely trees are growing on a given patch, the more likely the fire is to spread to that patch.

In the analysis of the simulation, I am mostly using the number of burnt (char) and burning cells. I then use them to obtain the fraction of burning/burnt cells. I use the fraction or the raw number of burnt cells as a measure of the fire impact on the grid.

Python Implementation and experiments/tests

To understand whether the python implementation is correct, it is helpful to look at the landscape of the area the simulation is modeling.

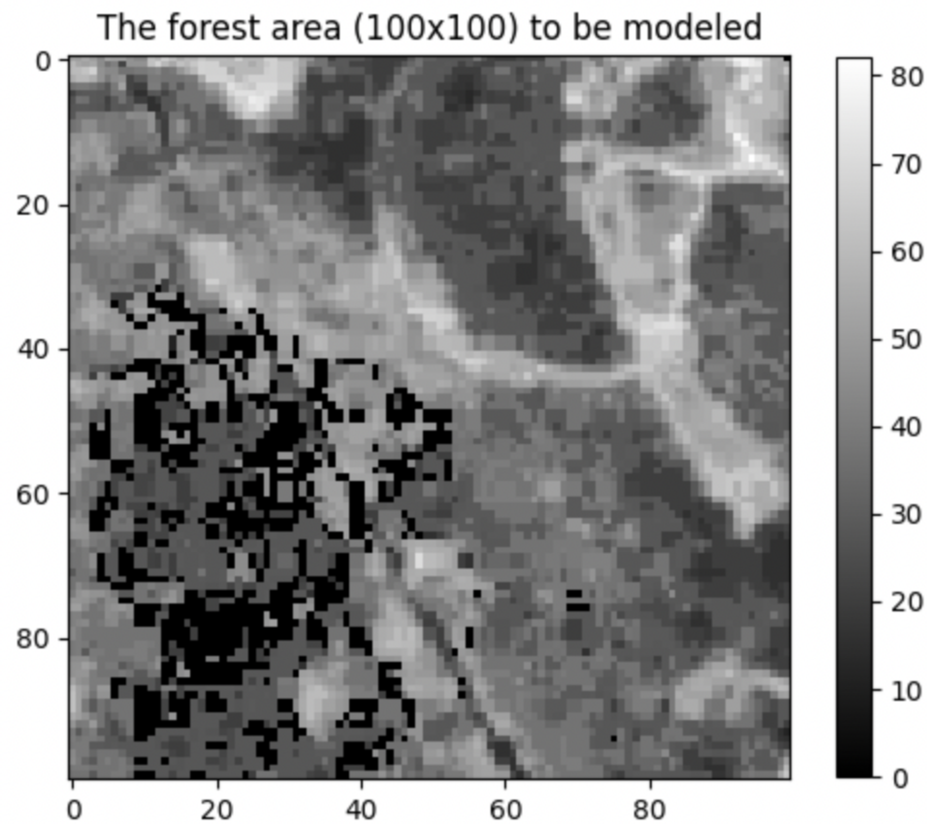

Figure 1. Showcasing the real forest area to be modeled in the simulation. The chosen area is a 100x100 square with the following coordinates: 00N, 020E. The colorbar corresponds to the tree density on a particular piece of the area, 0 (black) means there is no vegetation and a value of 100 (white) means that the area has 100% vegetation.

Looking at the image above, we can see that the area is not uniform. That is, there are areas where there is no vegetation (on the left), and areas with 80% vegetation (on the top right). Therefore, when evaluating the simulation implementation, one can watch the fire propagation through these areas to understand if the simulation is capturing the landscape correctly. Moreover, this area also represents a great fit for modeling due to its diversity. This means that we are likely to arrive at very different results depending on the fire starting point. For instance, if the fire starts near the down left part with little to no vegetation, it is unlikely to move far. However, if the fire starts in the top right corner, the fire is very likely to move around the square due to increased ability to propagate through the forest.

The first way to understand whether the simulation works as expected is to monitor the fire progress as it propagates through the forest via a dynamic animation. For the dynamic animation and the tests that follow I selected the fire start location to be x=70, y =50. I chose these coordinates since there is some vegetation and the point is close enough to the areas with the highest percentage of trees on the entire square, yet it is also close to the areas with no trees. This will allow us to see how the fire behaves around both types of landscape.

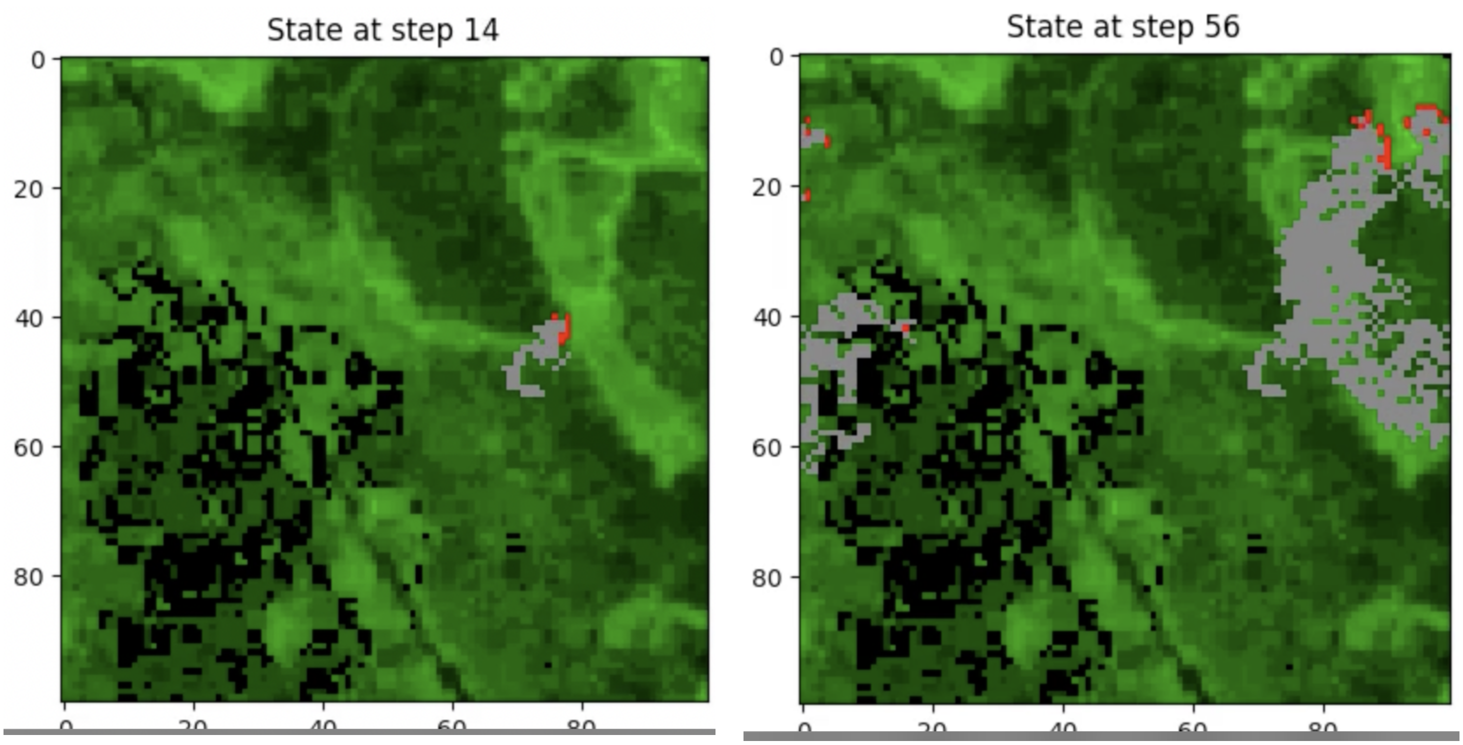

Figure 2. Showing the screenshots from a dynamic forest fire animation. The left screenshot shows the forest fire state at step 14, and the right screenshot shows the state at time step 56. The black areas are areas without any vegetation, and light green areas are the patches with most vegetation. Fire is displayed in red, and char in gray. The fire starting location is (70, 50), the boundary conditions are periodic.

Analyzing figure 2, we can see that the fire behaves as expected. More specifically, the left screenshot shows how fire started out at a lower point on the y-axis and traveled "up" (North), which corresponds to the set wind direction. In the screenshot to the right, we can see how the fire propagated easily through the more densely populated areas, and how it died out in the areas with no vegetation, leaving it black.

Another experiment that would help understand whether the simulation works as supposed to is to change the wind strength and the base spread rate. With a stronger wind, we would expect to see a greater proportion of burnt cells at the end of the simulation. Similarly, with a stronger base spread rate, we should expect to see a greater percentage of the entire area burnt by the end of the simulation. To check whether these parameters were implemented correctly, I ran the simulation 50 times for 100 steps with 5 different wind strength parameters ranging from 1 (weak) to 5 (strong), and plotted the averages over these 50 runs:

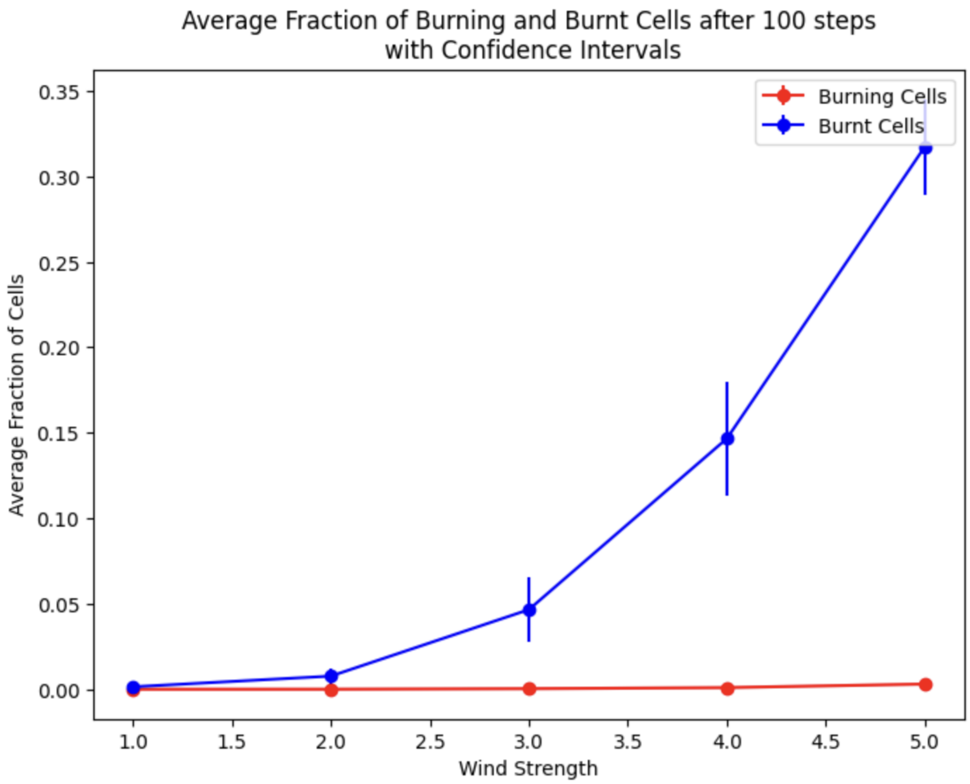

Figure 3. Showing the average fraction of cells burnt or burning at the final step. The simulation was run for 100 steps, 50 times. The points were obtained by averaging the 50 resulting values that also informed the 95% confidence intervals. The simulation was run for 5 wind strength values (0,1, 2, 3, 4, 5).

Analyzing the graph above, we can observe a general upward trend in the fraction of burnt cells as the wind strength increases. This is an expected result, since, as mentioned above, wind strength is integrated into the spread probability formula which aims to reflect the natural conditions. We can also see that the bruning cell numbers remain constant.This is because the forest usually stops burning by the end of 100 steps, which we can observe by running the dynamic animation.

Uncertainty Analysis

Due to the randomness in the simulation, we cannot be 100% sure in the results. To quantify the certainty level in the observed results, I created confidence intervals for simulation results for every wind strength. The confidence intervals on figure 3 are 95% confidence intervals, meaning that if we were to run the simulation 100 times, we would expect the true value to fall into the interval in 95 times. The confidence intervals for burnt cells widen as wind strength increases, indicating increased variability in the simulation outcomes at higher wind strengths. This could be because in case the wind factor is relatively small, the spread probability depends more heavily on the base rate and the density factor, which is keeping the average fraction of burnt cells low and leads to less variability due to low chance of occurrence. We see that the confidence intervals are greater with a higher wind strength, which reflects how a small change in natural conditions can lead to various outcomes in the environment. Wide confidence intervals also indicate high uncertainty in the simulation outcomes. To reduce the width of a confidence interval by half, one would need to increase the number of simulations by a factor of four, meaning that I would have to run each simulation 200 times to halve the intervals.

Overall, the result implies that managing wind exposure could be critical in controlling forest fires. Strategies such as creating windbreaks or conducting controlled burns in strategic locations to prevent the spread of fire could be effective.

Another way to test whether the simulation is implemented correctly and produces reasonable or expected results can be via running the simulation for different base spread rates. Similarly to the wind strength experiment, I ran the base spread rate experiment for 5 different base spread rate values from 0 to 1. Specifically, I ran the simulation for 100 time steps for 50 trials for every value. Then, I averaged the results over the number of runs and created 95% confidence intervals to display the change of the proportion of burnt/burning cells at a final simulation state with an increase in base spread rate:

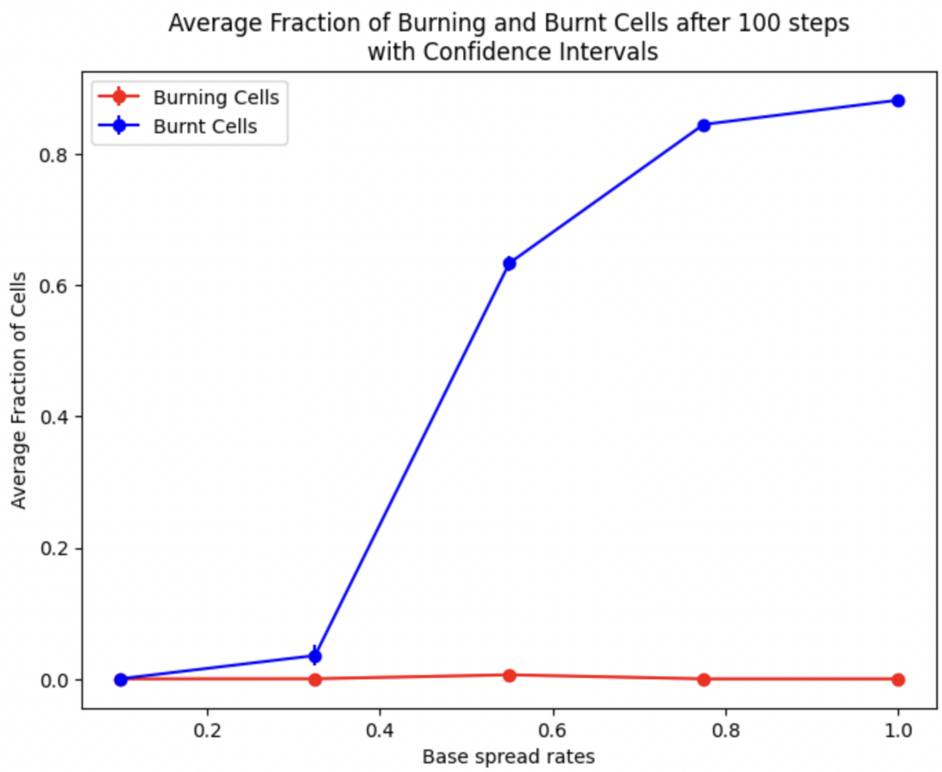

Figure 4. Showing the average fraction of cells burnt or burning at the final step. The simulation was run for 100 steps, 50 times. The points were obtained by averaging the 50 resulting values that also informed the 95% confidence intervals. The simulation was run for 5 base spread values from 0 to 1.

Observing the graph, we can notice that the "burnt cells" proportion follows the general upward trend like the previous graph did. Moreover, similarly to the previous graph, the burning cells proportion remains unchanged due to the same reason as above. In particular, the cells stop burning by step 100. However, compared to the previous graph, which showed a gradual increase with wind strength, the increase here is less gradual. The difference in the rate of increase of burnt cells suggests that the base spread rate is a more sensitive parameter in the model, with changes leading to more dramatic effects on the outcome. Once the base spread rate passes a certain threshold (< 0.4), the average fraction of burnt cells jumps dramatically, suggesting a tipping point or threshold effect where the fire becomes much more aggressive in its spread.

Uncertainty Analysis

The confidence intervals are very tight for burning cells, indicating a high level of consistency across simulations which is reasonable given that we have observed that cells stop burning by the 100th step. However, for burnt cells, the confidence intervals are only noticeable for base rate ~ 0.3, which indicates greater variability than for other rates. This can be because the simulation is highly susceptible to base rate changes and a higher base rate is very likely to lead to greater fires, making the outcomes predictable. The ~0.3 point could be an inflection point where the system's behavior starts to change, possibly becoming more chaotic or sensitive to initial conditions or random variations within the simulation.

The most susceptible areas

To understand the effect of fire starting in various locations, I ran the simulation with three different randomly set fire starting locations. I generated random coordinates and used them to start the fire at that point. I then ran the simulation for 100 steps for 100 trials/runs and created histograms to see the impact of various fire locations. It should be noted that "burnt cells" includes both burning and burnt cells at the end of each trial after 100 steps have passed.

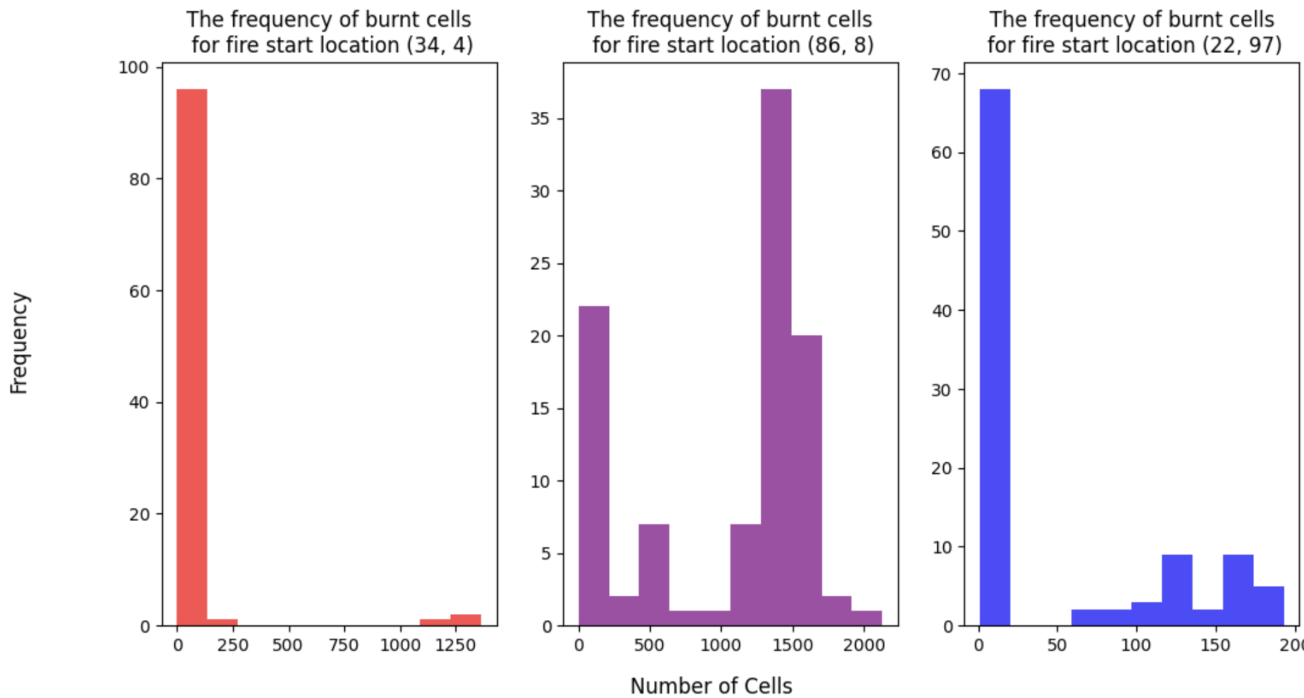

Figure 5. Showing the frequency of the total number of burnt/burning cells with various fire start locations ((34, 4), (86, 8), (22, 97)). The results were obtained by running the simulation for 100 steps for 100 trials for every fire starting location.

Looking at the plots above and specifically at the x-axis that indicates the number of cells close to being burnt or the number of burnt cells, we can see how the second fire starting location causes the most damage on average. In particular, when the fire starts at the coordinates [86, 8], the number of burnt cells can reach 2000, whereas for the other two locations, the maximum number of burnt cells is 1250 or 200. Moreover, both fire start location 1 and fire start location 2 have the most frequent number of burnt cells in the first bin. For location 1, it is 0-250 burnt cells, and for location 3, it is 0-50 burnt cells. Thus, we can infer that the fire is more likely to spread and have more extensive damage if it starts at (86, 8) coordinates. The fire is less likely to spread and cause a large number of burnt cells if it starts from (34, 4) coordinates, and the fire is the least dangerous if it starts at location (22, 97). To understand how these statistics can inform wildfire management strategies, it is helpful to visualize where the fire starting points are located on the map:

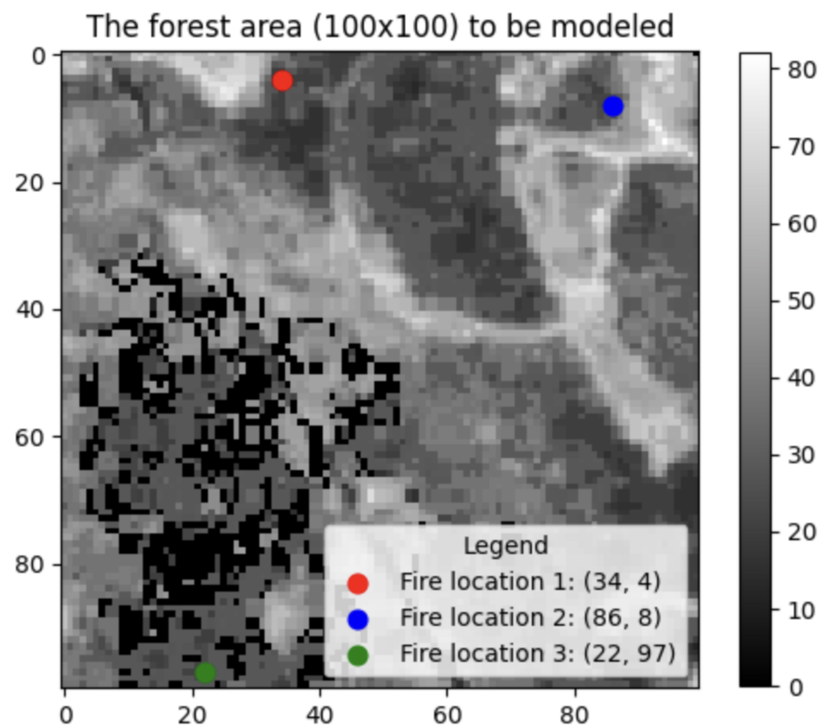

Figure 6. Showing the three randomly chosen fire start locations ((34, 4), (86, 8), (22, 97)) on the real tree coverage data visualization obtained from the NASA coverage data. The brighter areas indicate greater tree density, the darker areas represent lower tree density, the black patches are areas of no vegetation.

Looking at figure 6, we can see that fire start location 2 is situated in the area most densely populated with trees, while fire start location 1 is located in a lower density area and fire start location 3 is quite close to the “no vegetation” area. Synthesizing the information obtained from figure 5 and 6, we can infer that when the fire starts at a more densely populated location, it can lead to 10x times the number of burnt cells than when it starts near the no vegetation area. Thus, it might be a good preventative measure to keep the tree density at ~50 to avoid large forest fires.

We can now create a heatmap to observe which forest patches are most influenced by the fire in each scenario. This will allow us to infer the regions most prone to being burnt. Specifically, if the same region appears in multiple scenarios, we can infer it is a high risk forest fire area.

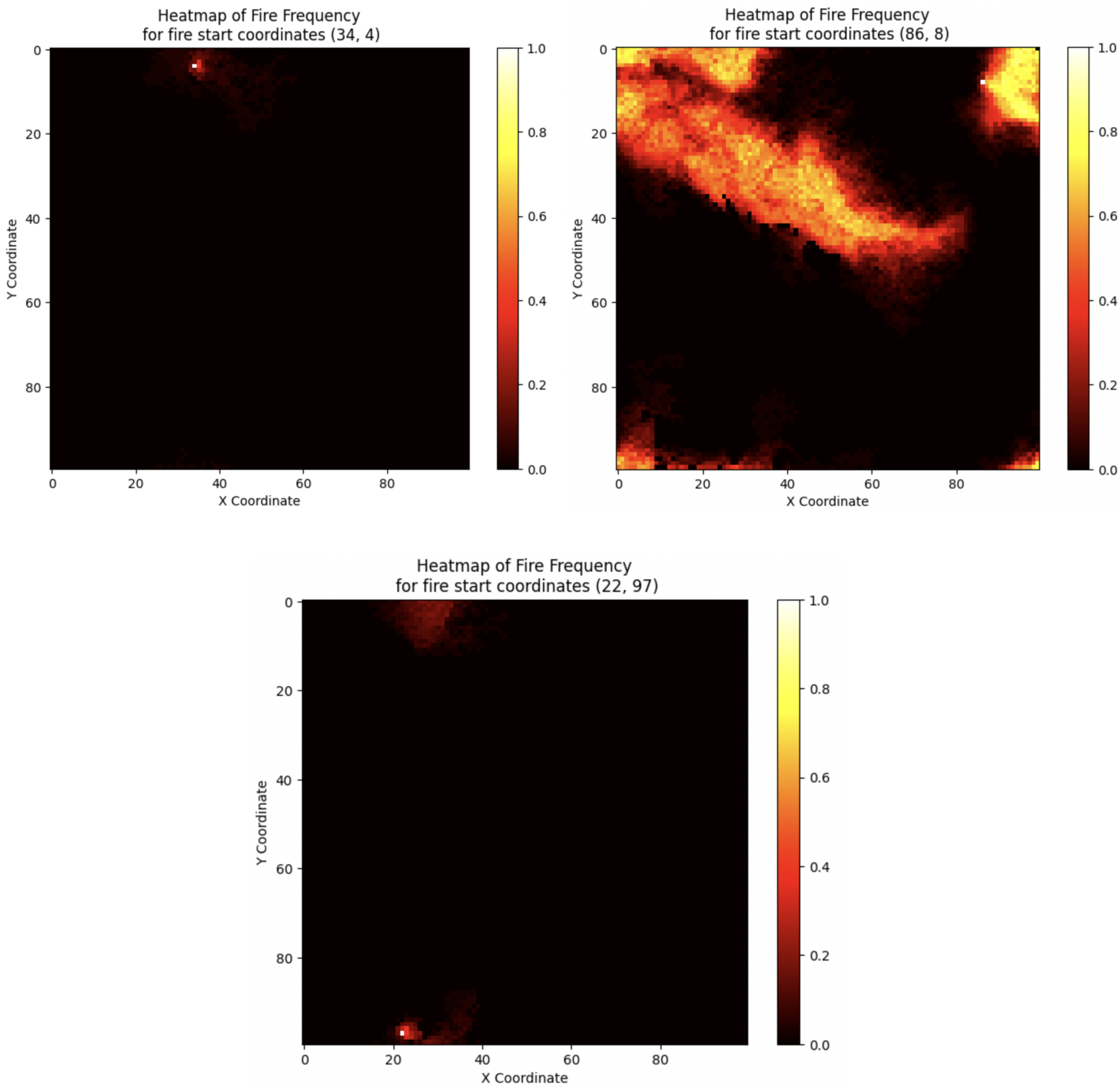

Figure 7. Showing three heatmaps of fire frequency corresponding to three different fire starting locations, - (34, 4), (86, 8), and (22, 97). The lighter areas indicate that the given patch is more likely to burn, and darker colored areas indicate the patches with lower burning probability. The simulation was run for 100 steps for 100 trials.

Analyzing figure 7, we can see that there is no significant overlap of the most frequently burnt/burning areas. The only overlapping area apparent in figure 7 is the triangular shape at ~ (30, 5) coordinates. This area seems to be burnt in the second and third scenarios and a little touched by the fire in scenario 1. However, there is no specific evident area that always catches fire. It should be noted that there might be areas that are more likely to catch fire that I could have missed since I have only looked at three fire starting scenarios. Looking at the fire starting locations in figure 6, additional fire starting locations can be the lower right and the center. I am also going to try the central right location because it appears to be highly vegetated. Therefore, the additional fire starting locations are: (80, 80), (50, 50), and (45, 85):

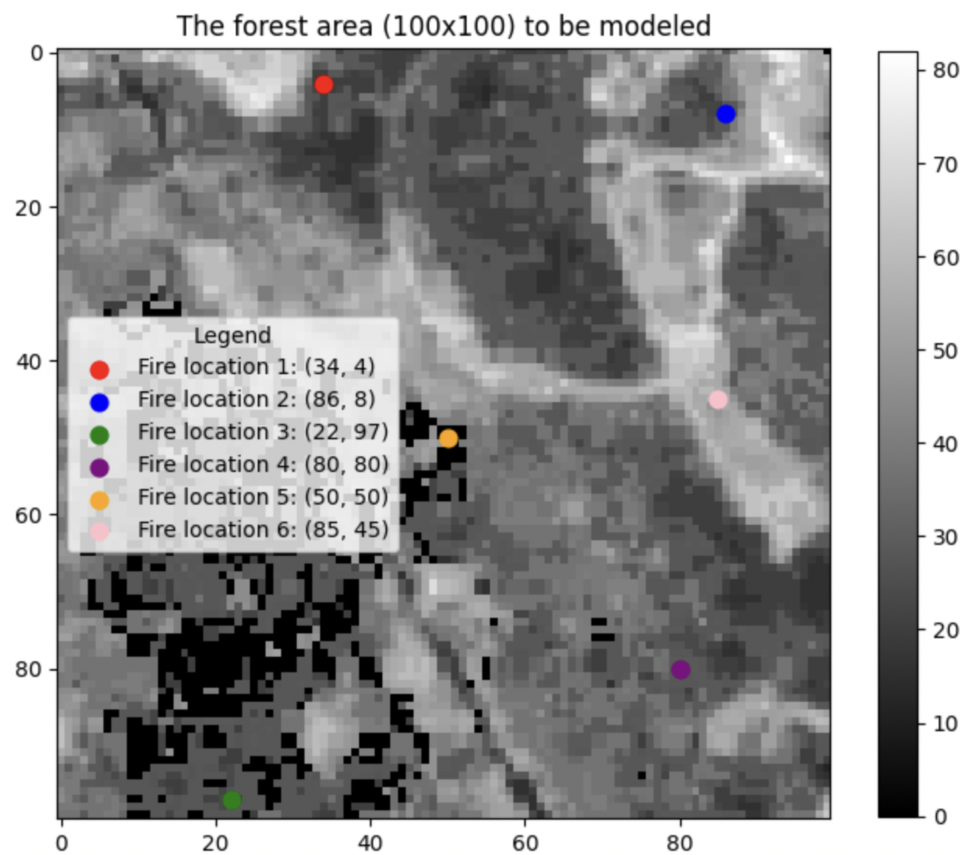

Figure 8. Showing the additional fire starting locations. The image displayed the real world tree coverage data obtained from the NASA database. The lighter areas indicate more vegetation and the darker ones less.

I ran the simulation with the same parameters as above for the three new locations to obtain the average frequency of the cells being burnt/close to burning for a fuller analysis. The results are presented as heat maps below. Since there are three heatmaps, it was best to display them on the next page together.

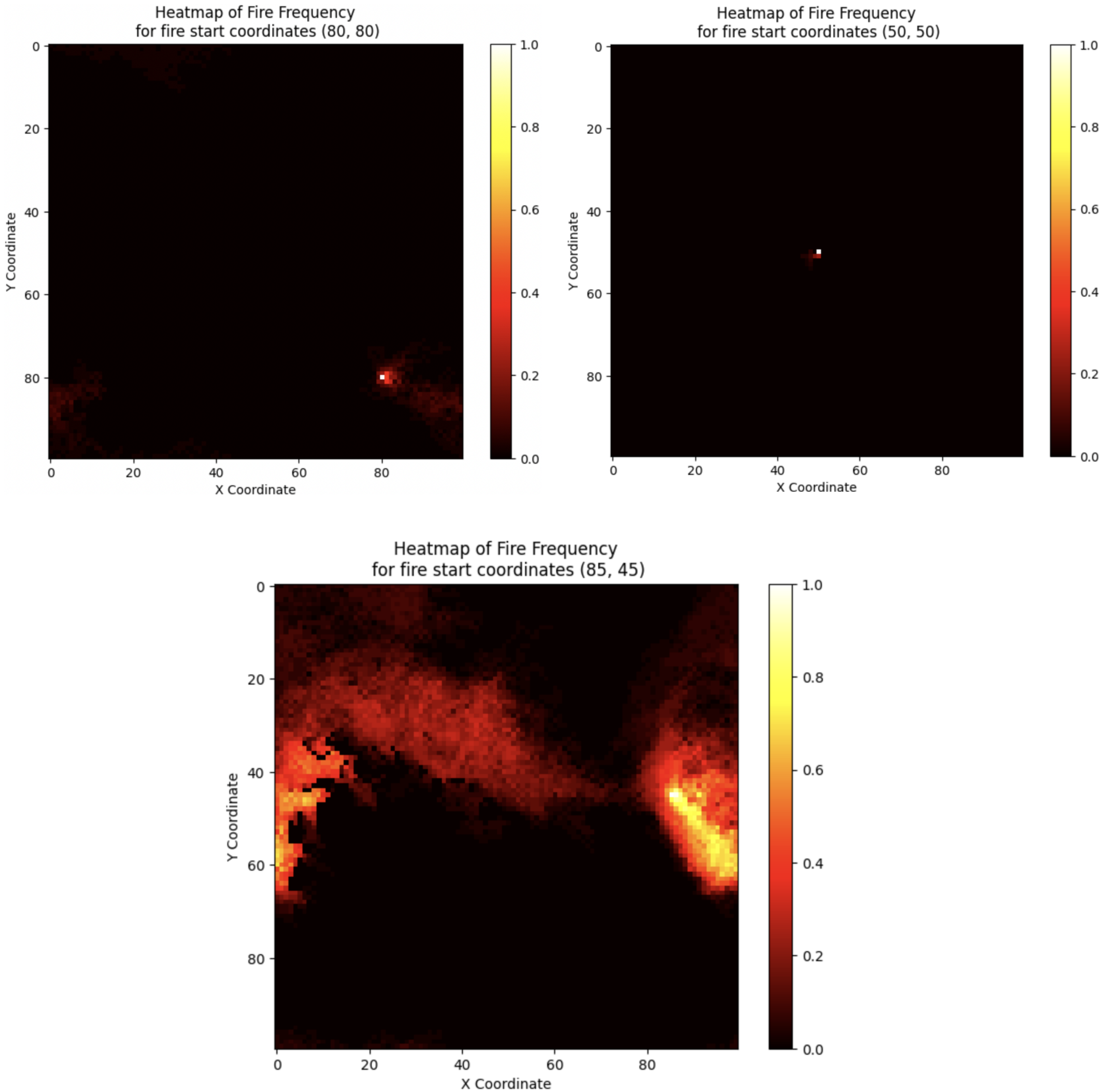

Figure 9. Showing three heatmaps of fire frequency corresponding to three different fire starting locations, - (80, 80), (50, 50), and (85, 45). The lighter areas indicate that the given patch is more likely to burn, and darker colored areas indicate the patches with lower burning probability. The simulation was run for 100 steps for 100 trials.

Looking at the additional three heatmaps, I do not observe much overlap in the affected areas with the start locations (80, 80) and (50, 50). However, there is significant overlay between the burnt areas between location 2 (86, 8) and location 6 (85, 45). This is not surprising given that the two locations are relatively close to one another and are situated within a greater patch of increased tree density. The overlap between the affected areas between these two scenarios is a strip starting from (0, 0) to coordinates (70, 40). It should be noted that these coordinates are approximations. However, the fire spread in that direction makes sense due to the boundary conditions. Specifically, both location 2 and 6 are at the far right of the modeled area, which has increased level of vegetation and connects to the area of moderate vegetation to the left. Fires starting in both locations follow the above described route.

Considering all six heatmaps together, we see that there are no areas that are consistently at a risk of being burnt down unless the fire starts in the center/upper right of the area (locations 2 and 6). This means that environmental specialists and policymakers should pay particular attention to those areas or areas with similar tree coverage and conditions. This might also mean that a useful preventative measure can be to attempt to “separate” the more densely populated areas from one another. However, it should be noted that there is a multitude of fire starting locations one can simulate for a fuller analysis. Moreover, these results were obtained under a specific combination of parameters (wind strength = 3, base spread rate = ⅓, wind direction = North). Having observed which areas need particular attention in case of fire, we can move on to implementing some of the preventative measures mentioned in the analysis above.

Measures to implement

Looking back at figure 3, we see that the fraction of burnt cells increases with increasing wind strength. Up until this point, the wind strength parameter was set to 3, which corresponds to "moderate" strength. However, figure 3 indicates that we can lower the fraction of burnt cells by some proportion by decreasing the wind strength parameter value. In real life, such a measure can correspond to installing wind breakers. I used coordinates (70, 50) for the analyses prior to the fire start location experiments where I intentionally varied the fire start locations. To assess the effectiveness of the proposed measure, I will implement it with the (70, 50) fire start coordinates for consistency. These coordinates are also located in the center-upper right region of the grid which was found to be the most likely to catch fire and burn in the previous sections. This means that this area needs particular attention since it is high risk due to increased vegetation. Therefore, I am testing the wind breaker intervention given (70, 50) as the fire start position.

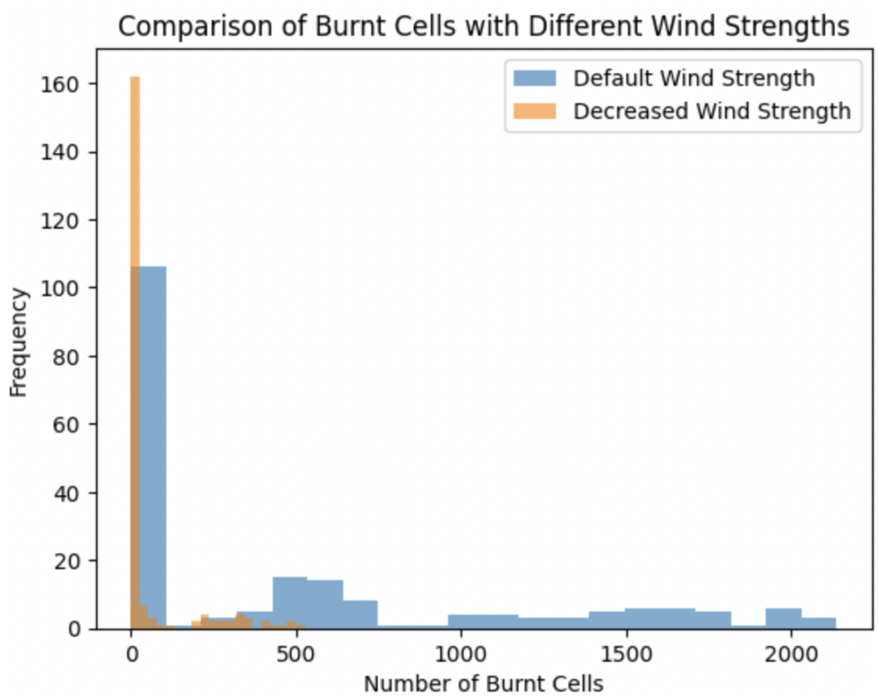

Figure 10. Showing the comparison of burnt cells at the final state of the simulation after 100 time steps. The histogram was obtained by running the simulation 200 times. The blue histogram is the default wind strength value of 3 that was used up until this point. The orange histogram corresponds to the results of the simulation with decreased wind strength of 2.

Figure 10 shows that the simulation with the default wind parameter of 3 has much more variability in the final number of burnt cells compared to the outcome of the simulation with wind strength of 2. More specifically, the outcomes of the simulation with a greater wind strength parameter value spread up to 2000 burnt cells whereas lower wind density results in a maximum of ~500 burnt cells. This suggests that when the wind strength is decreased, the fire spreads less aggressively, resulting in a lower number of simulations with a high count of burnt cells. We can also observe first bin dominance in both scenarios, which indicates that there is a smaller chance of the fire spreading extensively overall. However, the difference in the spread of the results conveys that decreased wind strength can be an effective solution to reduce and avoid extensive damage that can occur with greater wind intensity. In real life, these measures can be windbreakers or blowing wind in the opposite direction of the natural wind to counteract its effect.

Another strategy we can try to implement is changing the direction of the wind. Up until this point in the report, the default wind direction was set to “North”, which is also where most of the vegetation is in relation to the fire starting point. Hence, wind direction might be playing a crucial role in exacerbating the fire impact on the forest. Looking, at the figure 8, we can see that in order to get away from the most heavily tree populated area, we need to move to the center-lower left of the grid, which is southwest from the (70, 50) fire starting point. Therefore, I am running the simulation with the default parameter of wind direction (“North”), and with the new “Southwest” direction.

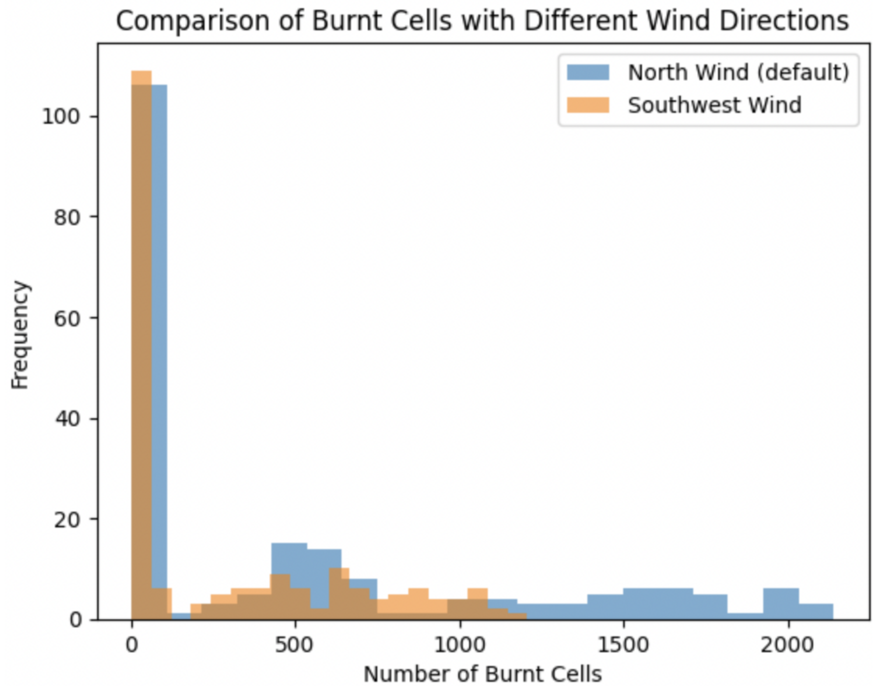

Figure 11. Showing the burnt cell frequency comparison for the simulations with the default wind direction parameter "North" (blue) and "Southwest" direction (orange). The simulation was run for 100 time steps for 200 runs.

We can see that the general trend is the same as on figure 10. Specifically, the intervention helps reduce the variability of the fire damage, however, not as significantly as it did in the wind strength reduction case. In this case, the maximum number of burnt cells in the intervention case is ~1200, while in the default case it is ~2000. This is definitely an improvement, however, not as significant as the one observed in figure 10. We can also see that both histograms display first bin dominance and are skewed to the right. This means that mostly the simulations result in few burnt cells with fewer occurrences of more extensive damage. While figure 11 shows that wind redirection can help reduce the maximum number of burnt cells and minimize the fire impact, it would do so less effectively than the wind strength reduction measure.

The combination of the aforementioned measures can help improve the outcome even more.

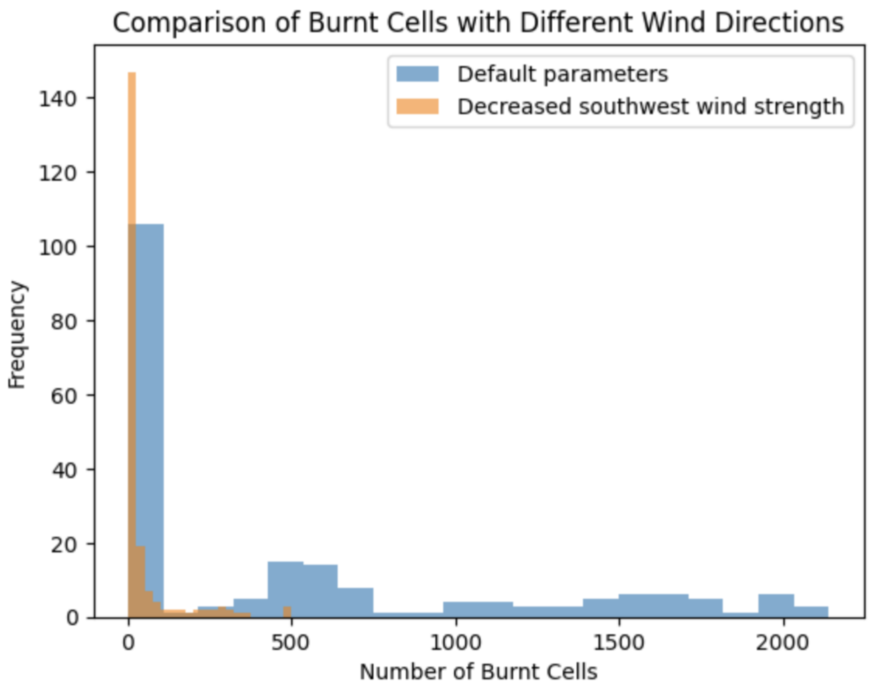

Figure 12. Showing the frequency of burnt cells comparison between the default simulation parameters (blue) and decreased wind strength and "Southwest" wind direction (orange). The simulation was run for 100 time steps 200 times.

Figure 12 shows that the combined measures (Southwest wind direction and reduced wind strength) resulted in similar results to the ones displayed on figure 10. The reduction in the variability of the number of burnt cells and the increase in the frequency of observing fewer burnt cells as compared to the default simulation parameters follow those shown on figure 10. This suggests that the wind strength reduction alone might be sufficient to achieve the improvement on figure 12. Therefore, it might be unnecessary to spend the resource attempting to redirect the wind given that the improvement is not as significant as can be achieved via reducing the wind strength. Thus, I would suggest the policy-makers invest in wind breakers in particularly densely vegetated regions that are surrounded by areas with moderate to high tree coverage.

Summary

In general, I would recommend monitoring the areas with tree density coverage > 50 because, as shown in the report, these areas are easily inflammable and could propagate fire effectively. Fires starting at locations with little to no vegetation such as the lower left corner of the grid this paper was modeling are very unlikely to propagate fire to other areas and result in significant damage. Therefore, it is essential to prevent fire starting in areas surrounded by other moderately/highly vegetated areas. However, if the fire has already started in such an area, I would recommend putting all the resources and efforts into breaking the wind and stopping it from contributing to the fire. This can allow the firefighters sufficient time to stop the fire since it would not be accelerated by the wind. This report also highlighted that putting efforts into redirection of the wind will not yield much effect. Thus, the two main recommendations this report puts forward is to maintain a lower tree density coverage such that the damage does not propagate to other areas, and decreasing the wind impact if the fire has started.

AI Statement

I used AI for writing my doc strings for the code.