Empirical Wildfire Spread & Intervention Analysis in Western Germany

CS166 - Modeling and Analysis of Complex Systems
Minerva University
Tambasco, TTh@3PM UTC – Fall 2025

December 19, 2025

Contents

1 Executive Summary **2**

2 Introduction **3**

3 Model Description and Python Implementation **3**
3.1 Modeling Goals . . . 3
3.2 Rules of the Simulation . . . 4
3.2.1 Initial Model: Density and Wind-Dependent Spread . . . 4
3.2.2 Model Extension 1: Fuel Usage . . . 5
3.2.3 Model Extension 2: Rain Events . . . 5
3.3 Modeling Assumptions . . . 5
3.4 Simulation Outputs and Quantities of Interest . . . 6
3.5 Test cases . . . 6
3.6 Sensitivity Analysis . . . 7

4 Fire Risk Assessment **9**
4.1 Stable State Analysis . . . 9
4.2 Complete Risk Map . . . 9

5 Proposed Interventions **10**
5.1 Intervention 1: Forest Thinning . . . 10
5.1.1 Deriving the critical thinning factor through renormalization . . . 11
5.2 Intervention 2: Forest Network Intervention . . . 13

6 Acknowledgments **15**

1 Executive Summary

This report presents a computational risk analysis of wildfire vulnerability in the "Bergisches Land" region near Cologne, my home town whose ecological health I deeply care about, utilizing a probabilistic cellular automaton model created using local satellite data. The primary objective of this report is to provide authorities with data-based insights to mitigate the increasing threat of wildfires in this spruce-dominated landscape which has earned the "high wildfire risk" label alongside cities in Spain, Italy, and Greece.

Our analysis identifies that the region's high connectivity and dense forest clusters create a critical state where, without intervention, a single ignition event has a $> 75\%$ probability of consuming the majority of the forest. This report evaluated three intervention strategies to counter this risk:

1. **Uniform Thinning:** Reducing tree density in the entire modeled area. While effective, this requires removing over 35% of total biomass to achieve safety, which is ecologically and logistically straining.

2. **Targeted Thinning:** If we instead focus thinning efforts on the densest parts of the forest, this improves logistical efficiency as less areas need to be thinned but still demands significant widespread deforestation ($> 30\%$ of targeted biomass) to stop fire spread.

3. **Network-Based Firebreaks:** Through an algorithm that models the most likely paths of the fire spread, we can identify specific areas in the forest that, if cut down, can significantly reduce the probability of a continued and large-scale wildfire.

Recommendation. Based on the simulation results stated above, **the Network-Based Firebreak strategy is strongly advised**. By removing only **50 strategic cells** (less than 0.2% of the forest), the average expected burnt area can be reduced by approximately 50%. Among the three approaches, this intervention creates safety while *minimizing ecological disruption and implementation costs*. For more information, continue to the following sections.

2 Introduction

After the horrifying casualties and destruction caused by the flooding in the West-German Ahr Valley in 2021, local authorities have had to admit their low preparedness for the occurrence of natural disasters. Given their predicted increase in frequency amid the effects of climate change, emergency response protocols have become a large matter of public attention, forcing many authorities to revise their climatic risk assessments. This also includes wildfires, which have been studied in detail in Mediterranean areas like Greece and Spain, but not so much in Central European countries like Germany. This is somewhat surprising given that the European Forest Fire Information System (EFFIS) is assigning the highest level of wildfire risk to many western and central forests in Germany, mostly due to an overly homogeneous forest composition dominated by spruce trees - a legacy of requiring high-yield timber after the Second World War - that are particularly vulnerable to heat events and pests. One of these areas is the "Bergisches Land" around Cologne, a very green area for which wildfire spread has not yet been studied much. This analysis will provide insights into the vulnerability of a 8.3km by 8.3km grid around the towns of Dattenfeld, Hönrath, and Schladern.

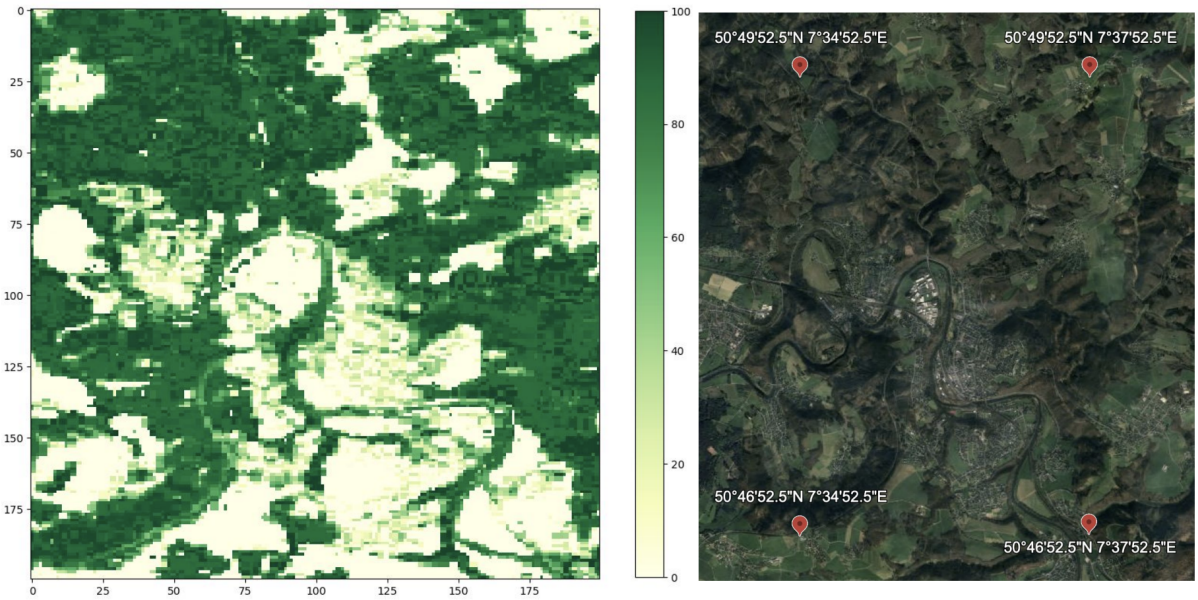

Figure 1: Satellite data plotted alongside an image on Google Earth to confirm the location and 'look' of these two areas. The chosen grid has coordinates $7.58125 - 7.63125$E and $50.781111 - 50.83111$N.

3 Model Description and Python Implementation

3.1 Modeling Goals

Given that this analysis aims to help model the potential spread of wildfire in this chosen region, we would aim to learn about several different metrics:

1. Firstly, we want to learn which areas are at the highest risk of catching fires. As a local authority, this is highly relevant for targeting prevention measures as well as emergency responses to these areas.

2. Secondly, this model should provide information on the maximum fire intensity, which is the maximum number of cells that are burning at the same time for a given set of parameters. This is key to identify the 'scope' of an appropriate emergency response

3. We might also want to see how 'fast' the fire spreads. It is plausible to assume that fire spreads faster in high-density areas

4. Lastly, given the topography of the forest density on this map, we want to identify specific areas (or cells in the graph representation) that forest rangers can cut down in order to prevent the spread

of fires across forest clusters. What are some strategies we can apply here, and how effective are they?

3.2 Rules of the Simulation

3.2.1 Initial Model: Density and Wind-Dependent Spread

The core simulation is a probabilistic Cellular Automaton (CA) operating on a 2D grid representing forest density. The state of the system evolves in discrete time steps according to the following rules:

1. Grid Representation. The selected area is represented as a grid G of size $H \times W$. Each cell (x, y) possesses a static property representing tree density, $D_{x,y} \in [0, 1]$, derived from the normalized satellite data. Note that while each cell carries the density attribute, they are still in one of three categorical cells. These are tracked using two boolean masks:

- $M_{burning}$: Cells currently on fire. Cells of this state have the ability to pass on the fire to neighboring cells.
- M_{burnt}: Cells that have been consumed and can no longer burn, catch fire, or pass on fire.

If a cell has neither of the two masks, its forest is intact, meaning that a third categorical state, $M_{healthy}$ is implied by the absence of the other two. At the beginning of a simulation, a fire is initiated at a coordinate (x_0, y_0) where $D(x_0, y_0) > 0$ by setting $M_{burning}(x_0, y_0) = \text{True}$. All other cells are initialized as unburnt.

2. Update logic at each time step. For every time step t, the simulation iterates through all currently burning cells. For each burning cell at location (x_B, y_B), the fire attempts to spread to its 8 neighbors. A Moore neighborhood was chosen as a more realistic assumption as fires are not restricted to spread solely to their left and right neighbors. Especially if the wind flows in a diagonal direction, the diagonal neighbors become important. For a neighbor at (x', y'), the probability of ignition, P_{spread}, should depend both on the tree density of that neighbor, as well as the wind direction. This is because trees are not uniformly distributed within each cell, even though the density value gives an average value. For instance, consider the three pictures below, each of which represents a *single* cell with the black squares representing trees. Note that for any neighboring cell, it would be much more likely for fire so spread when the density of this cell is high and more trees at the 'edge' that could catch and transmit the fire are to be expected.

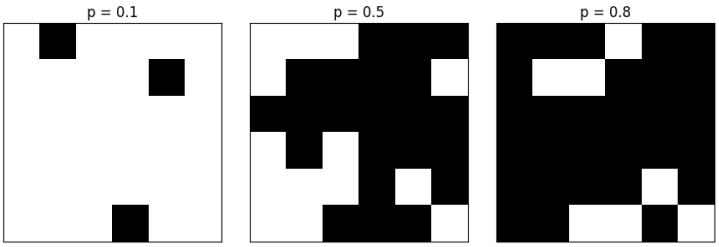

Figure 2: Sample Cells with different densities. The p-values represent the probability of any cell to be a tree during the initialization process.

Additionally, the direction and strength of the wind also affects the probability of fire spread, however, it does so differently depending on each of the neighbors. For instance, if the second grid in Figure 2 was burning, and the wind was blowing from the left to the right, the probability of fire spread to cell 3 should be higher than if the wind was blowing in the opposite direction.

3. Wind Dynamics The reason that wind matters is because it carries embers to neighboring trees in its direction (SOURCE). To simulate this physical effect, we calculate the alignment between the wind direction and the fire's "propagation vector". Assuming that wind takes a uniform direction across the entire grid, which is not unreasonable given that it is only about 8×8 km large, let $\vec{W}$ be the global wind vector and $\vec{v_{ij}}$ be the normalized vector pointing from the burning cell i to the neighbor j. Each

cell has a maximum of 8 of these vectors $\vec{v}$, given that at most all eight surrounding cells have a density larger than zero and are able to catch cell i's fire.

If $\vec{W}$ and $\vec{v}$ are aligned (and the angle between them $\theta = 0°$), the dot product is 1.0, maximizing the spread probability. If they are opposing ($\theta = 180°$), which models crosswind, the dot product is -1.0, significantly reducing the probability. If perpendicular, the term is 0, and the modifier is neutral. Note that for this simulation, a wind direction of 0° corresponds to moving South (↓). An exponential link function was used to model this behavior to get strictly positive values. To ensure the wind effect remains bounded even at potentially very high wind speeds, a hyperbolic tangent link function was used. The modifier ψ_{wind} is defined as:

$$\psi_{wind} = 1 + \tanh(S_{wind} \cdot (\vec{W} \cdot \vec{v})) \tag{1}$$

where $S_{wind} \geq 0$ represents the wind intensity. Unlike an exponential link function, which can grow indefinitely, the tanh function saturates which means the total range of ψ_{wind} is strictly bounded between $1 + (-1) = 0$ for strong opposing winds and $1 + 1 = 2$ for strong aligned winds, i.e., $\psi_{wind} \in [0, 2]$. This formula ensures that stronger winds ($S_{wind} > 0$) create a directional bias by nearly doubling the spread probability in the wind's direction while reducing it to near zero against the wind, causing the fire to spread in narrow cone shapes aligned with the wind direction. This allows us to complete the probability of fire spread from a burning to a healthy cell. Note that we also multiply by a baseline probability of P_{base} which can model additional differences in forest nature, e.g. the type of wood: spruce trees after a long hot summer are *highly* inflammable, meaning they would have a higher P_{base} than a humid forest where transmission is less likely. Therefore:

$$P_{spread} = P_{base} \cdot D_{x,y} \cdot \psi_{wind}. \tag{2}$$

3.2.2 Model Extension 1: Fuel Usage

The extended model introduces variable burn duration (fuel load) and environmental stochasticity (rain) to increase realism. In the initial model, a cell burns for exactly one time step. This is a strong simplification as wildfires can 'linger' for sustained periods of time. In the extension, we introduce a dynamic fuel variable $F_{x,y}$, initialized to the tree density:

$$F_{x,y}(t = 0) = D_{x,y}$$

In each time step, if a cell is burning, its fuel is depleted by a constant burn rate β:

$$F_{x,y}(t + 1) = F_{x,y}(t) - \beta \tag{3}$$

A cell remains in the $M_{burning}$ state as long as $F_{x,y} > 0$. This is a more realistic model because existing fires can repeatedly 'attempt' to pass the fire on to their neighbors. This implies that dense areas burn longer, increasing the likelihood of spreading the fire to their neighbors, while sparse areas burn out quickly and naturally dampen the spread of the fire.

3.2.3 Model Extension 2: Rain Events

Alongside wind, rain is another natural event that can affect the spread of the wildfire. While only long, sustained periods of rain can extinguish wildfires, short periods can help reduce the speed at which fires spread as they dampen materials and hence reduce the probability of spread (source). At the beginning of each step, a global rain event occurs with probability P_{rain}, which is plausible as the modeled area is relatively small (8×8km). If it rains, two system parameters are temporarily modified to simulate the suppression of fire. Firstly, the ignition probability is multiplied by a dampening factor $\lambda < 1$, so that $P'_{spread} = P_{spread} \cdot \lambda$. Secondly, the burn rate is doubled to simulate wet fuel extinguishing faster with $\beta' = 2 \cdot \beta$.

3.3 Modeling Assumptions

This simulation relies on several simplifying assumptions in order to make the wildfire dynamics computationally tractable while keeping the main mechanisms of fire spread that were identified above.

First, wind is assumed to be uniform across the entire grid and constant throughout each simulation run. This assumption is reasonable given the relatively small spatial scope of the modeled region (approximately 8×8 km), where larger scale weather events reasonably dominate local fluctuations. However,

this assumption may not hold when there is a lot of variation in elevation or urban structure, where the wind might flow very heterogeneously throughout the grid.

Second, each grid cell is characterized by a single tree density value derived from satellite data. This assumes that (1) trees are evenly distributed within each cell and that (2) tree density alone determines how flammable an area is. While this is a reasonable approximation for regional risk assessment, especially on an 8x8k m scale, it does not capture differences in tree species, moisture content, or other fuels like bushes or leaves, which may significantly affect the spread of real wildfires.

Third, fire spread is modeled as a probabilistic process that depends only on local interactions on a Moore neighborhood. This assumes that the spread is only limited to the immediate neighborhood, while the "long-range transport" of embers has been shown to affect areas further away, especially under extreme winds. A further extension of this model would involve modeling the probability of ignition for cells that are in the second or third or nth row of the burning cell's neighborhood. The probability of ignition would be significantly lower, but positive. Windier conditions would then bias the transmission into a specific direction and increase the probability of a larger distance ignition.

Fourth, the rain events are modeled as unlikely and global stochastic shocks that uniformly slow down the fire spread while accelerating fuel depletion. Similarly to the uniformity assumption of wind, this assumption is valid when rain falls evenly over the entire region, but it does not capture localized precipitation or prolonged rainfall patterns.

Lastly, it should be stated that this grid also models a closed area, while wildfires can both be caused by an area outside of the interval and passed on to them. While this zoomed-in approach might more accurately model local conditions, it does not accurately represent the real forest structure. This problem is exacerbated by the potential outdatedness of the satellite data which was collected in 2000.

3.4 Simulation Outputs and Quantities of Interest

The simulation produces several outputs that summarize the outcome of each stochastic wildfire simulation run. These quantities are used as inputs to a Monte Carlo analysis, allowing distributions over outcomes to be estimated.

The primary quantity of interest is the total burnt fraction of the forest, defined as the number of grid cells that are burnt divided by the number of cells with $D_{x,y} > 0$, by the end of a simulation. This metric directly measures the severity of a wildfire and serves as a proxy for the damage it causes. We will aim to minimize this quantity.

A second important quantity is the peak fire intensity, measured as the maximum number of cells burning simultaneously during a simulation. This value is directly relevant for emergency response planning, as it reflects the maximum strain placed on firefighting resources.

The fire duration, defined as the number of time steps until no burning cells remain, provides insight into how long fires persist under different environmental conditions and intervention strategies.

In addition to these scalar metrics, the simulation records spatial burn frequencies across Monte Carlo runs. By aggregating results over many simulations with randomized ignition locations, it is possible to estimate the probability that each grid cell is burned. This produces a spatial risk map that identifies particularly vulnerable regions and informs the placement of strategic firebreaks.

Together, these outputs allow the model to quantify both the expected severity and the variability of wildfire outcomes, enabling robust comparison of intervention strategies under uncertainty.

3.5 Test cases

Several test cases were implemented to show that the simulation runs as expected. For instance, the first three plots show the effects of the wind direction parameter, which, if given a large positive value, biases the spread of the fire into its direction. Additionally, the fourth plot shows that fire cannot be spread to cells with a tree density of 0 (which is created at $x = 30$). We also observe that rain has the intended

dampening effect which allows fire to spread, but more slowly. Lastly, the sixth plot confirms that the simulation correctly models the model's preference for high density cells.

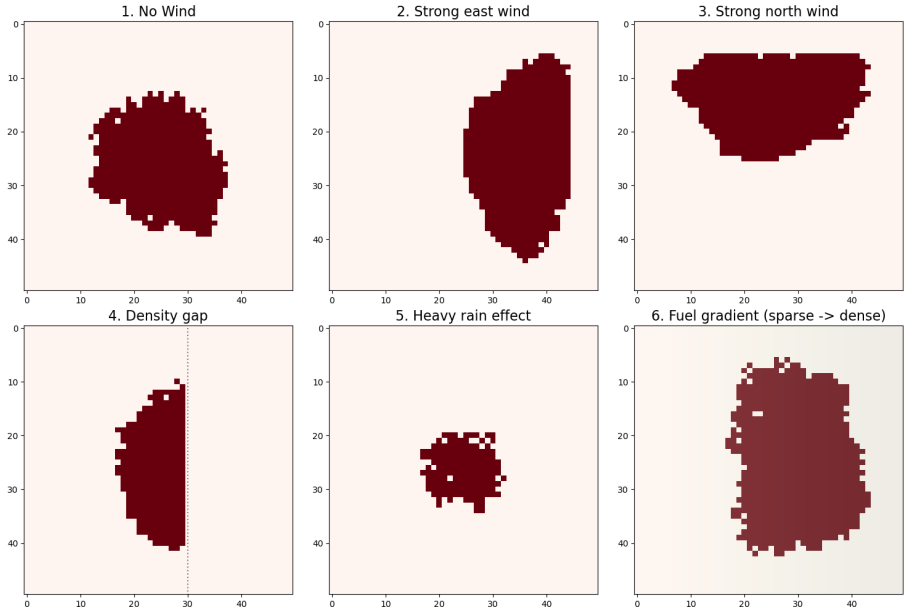

Figure 3: Visualizing test cases.

3.6 Sensitivity Analysis

To further validate the model dynamics, we performed a parameter sweep (sensitivity analysis) to observe how changes in environmental variables affect the total burnt area fraction. These are by no means an exhaustive analysis as the parameters likely interact with each other. The goal of these plots is to analyze whether changes in the parameters lead to plausible changes in the chosen metrics.

- **Rain Chance:** An increase in rain probability leads to a non-linear decay in burnt area, confirming the dampening effect of precipitation events. Due to the increase in the burning rate that is accompanied by a rain event, the fire duration is also decreased. As it reduces the probability of fire transmission from cell to cell, the peak intensity (maximum number of cells burning at the same time) also decreases with a higher chance of rain.

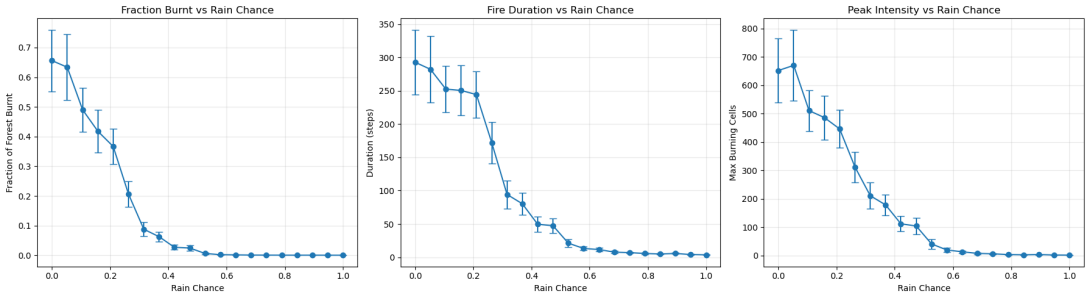

Figure 4: Parameter sweep with 95% confidence intervals of the rain chance parameter for the burnt fraction, fire duration, and the peak intensity.

- **Burn Rate:** A higher burn rate (β) inversely correlates with burnt area. When fuel is consumed too quickly (when $\beta > 0.6$), the fire extinguishes itself before it can spread effectively to neighbors. This is coupled with a short fire duration at a high burn rate. As expected, low burn rates cause fires to persist as they can stay in a single cell for multiple simulation steps. This wildfire scenario is particularly hard to put out as flames spread and persist at the same time.

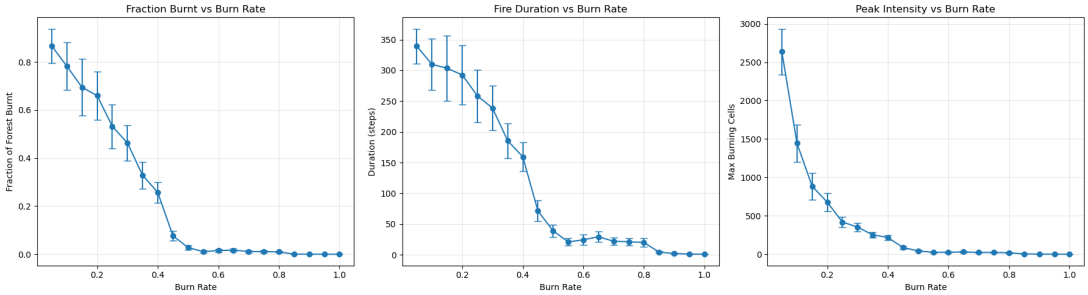

Figure 5: Parameter sweep with 95% confidence intervals of the burn rate parameter for the burnt fraction, fire duration, and the peak intensity.

- **Base Probability:** The relationship between the base ignition probability and burnt area follows a logistic-like curve, indicative of a phase transition where the system quickly moves from sub-critical (fire dies out) to super-critical (fire consumes the grid) behaviors. At high probabilities, the fire not only burns a large fraction, it also does so increasingly fast as $P_{base} > 0.2$. Lastly, the maximum intensity increases almost linearly with P_{base}. This is also plausible as, with a constant value of the burn rate, a higher ignition probability will cause the majority of the neighbors to catch fire in each step.

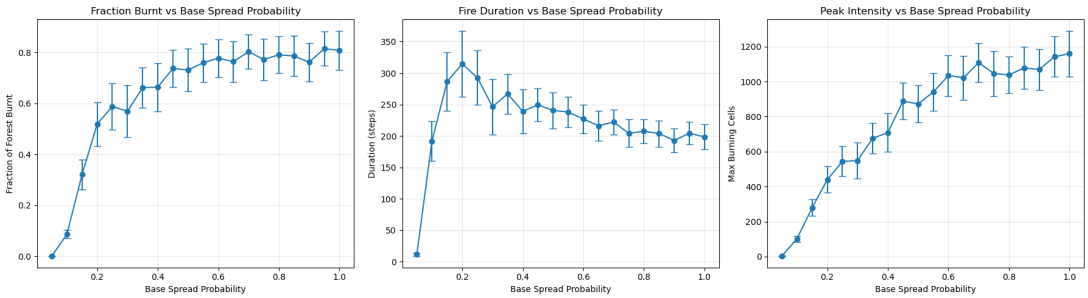

Figure 6: Parameter sweep of the base ignition probability parameter for the burnt fraction, fire duration, and the peak intensity.

These trends provide additional evidence that the simulation rules work coherently to produce plausible wildfire approximations.

4 Fire Risk Assessment

The most important part of assessing the wildfire risk is identifying the most vulnerable cells with the highest probability of catching fire r_p. Initializing the first spark at a random tree-populated cell, we observe multiple simulations and record the average risk values on an interval $r_p \in [0, 1]$ in a heatmap. However, to make sure that we are not using unstable, temporary simulation states for this comparison, convergence of the mean has to be ensured.

4.1 Stable State Analysis

For that, the simulation was run dynamically with an upper-bound maximum number of trials of 500. At each simulation step, the probability of each cell to catch fire,

$$r_p(C_{ij}, n) = \frac{1}{n} \sum_{n} I(C_{ij}),$$

where $I(C_{ij})$ is an indicator variable that is 1 when the cell c_{ij} caught fire, and 0 when it did not. At each subsequent time step, the mean risk value of each cell is updated and compared to that of the previous step. The simulation is continued until the maximum absolute change in mean of any cell is smaller than $\epsilon = 0.01$ for ten consecutive trials. This way, whenever a simulation with a given parameter set is run, its output will represent the converged state. Note that an even stricter strategy would be to calculate the relative difference between steps. However, due to computational limitations on this 200x200 grid, where single simulations take more than 500 steps to converge under a $\epsilon = 0.01$ threshold for relative differences, the absolute difference was chosen. Figure 4 shows an example simulation that converged after 105 simulation runs. The plot on the right shows how the absolute change decreases without much fluctuations, which implies that the simulation runs do not produce vastly fluctuating results. This is likely to the high connectivity whereby fire spreads through the very similar parts of the grid in each run.

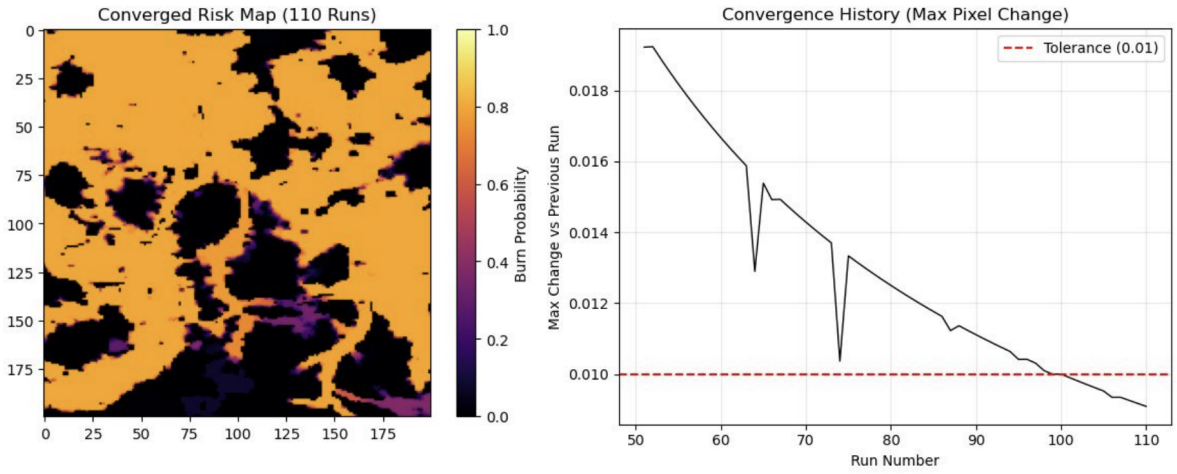

Figure 7: Sample simulation plot with the default parameters showing a relatively linear convergence trend. This shows how convergence is defined and established for the heat map.

4.2 Complete Risk Map

The plot was randomly initialized 10 times for each of the four wind directions each to also account for varying results due to different wind directions. The 40 resulting simulation results were then averaged to produce the complete risk heatmap below. It is noticeable that two colors dominate this plot: black and orange. This is also evident on the histogram over the burn probabilities which shows two peaks at the very edge of the empirical range: a lot of the cells have a burn probability of 0, which corresponds to the black cells in the heatmap where the tree density is 0 and fires cannot spread to. The other peak is at the other end of the interval at around 0.75 and corresponds to the orange areas on the map. These

are the areas that almost always burn down in the simulations, which likely implies that if the fire sparks anywhere in the orange area, its entirety will burn down over time with the given parameter set.

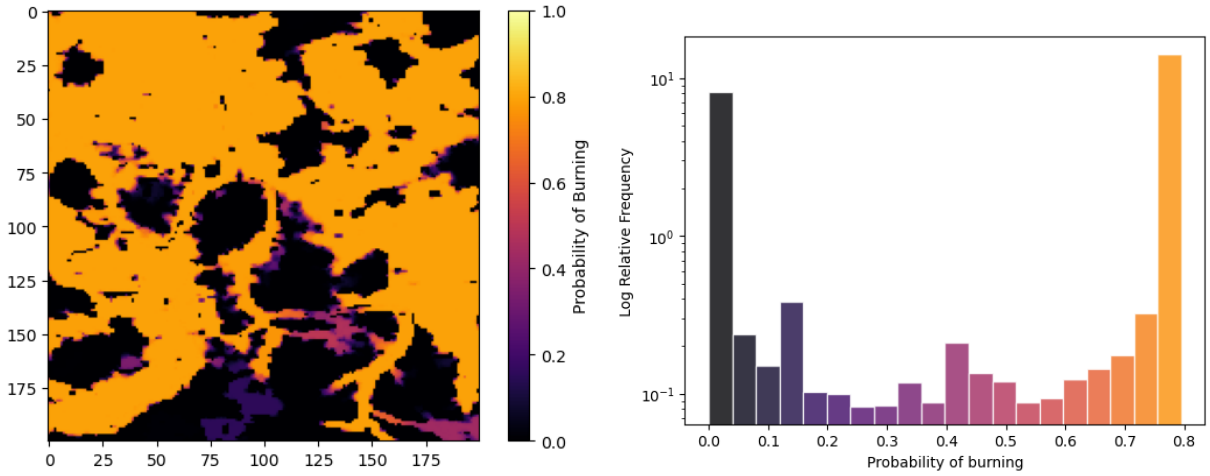

Figure 8: Left: Burn Risk Heatmap where orange signifies the highest levels of risk. **Right:** Bimodal distribution over the probability of burning.

5 Proposed Interventions

The fire risk assessment revealed a concerning result: after accounting for different wind directions, the majority (75% of all populated cells) has a 75% or higher risk of burning down if left uncontrolled. This is due to the relatively high tree density across the entire grid, as well as the presence of connectors between larger clusters of trees. These can be seen in 'thin' funnel-shaped areas on the heat-risk map: if fire breaks in either one of the large clusters, it needs to pass through funnels to ignite the other cluster, too.

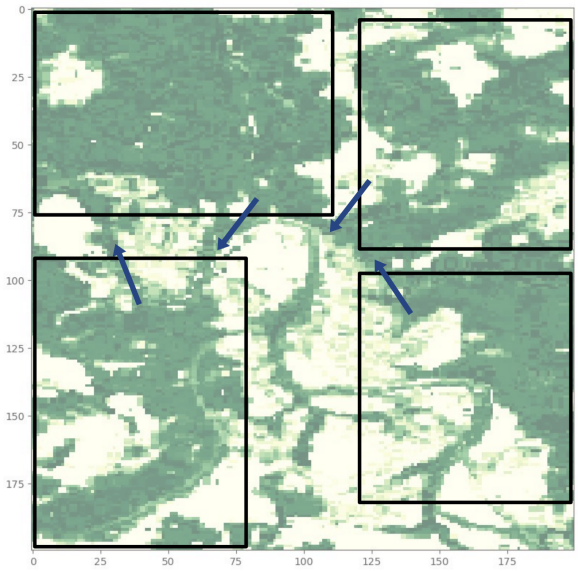

Figure 9: Expected location funnels as well as rough estimation of cluster locations.

5.1 Intervention 1: Forest Thinning

Another commonly applied strategy in wildfire prevention is "forest thinning", where the tree density is deliberately reduced to slow down the speed of the spread and allow more time for emergency responses.

As the tree density is an input parameter to the spread probability equation in (2), we can extend this by adding a thinning factor ν that models the percentage of the original forest that is preserved, while $1 - \nu$ shows the percentage of trees that is cut down:

$$P_{spread} = P_{base} \cdot D_{x,y} \cdot \psi_{wind} \cdot \nu. \tag{4}$$

5.1.1 Deriving the critical thinning factor through renormalization

While the simulation in the next section identifies an effective thinning factor empirically, we can attempt to find a result by deriving the theoretical critical threshold using renormalization group analysis which allows us to predict how much the forest density must be reduced (ν) to stop the spread of fire. This strategy is used to create an *approximation* by assuming the absence of pre-defined spatial patterns, wind, and rain events.

Finding the Absolute Density Threshold (p_c). First, we calculate the absolute density required to stop a fire. The simulation uses a Moore neighborhood (with 8 neighbors), meaning fire can spread in all directions. For the fire to be stopped completely, it must be surrounded by a connected barrier of empty space because. We simplify the scale of the grid to standard 2×2 blocks. Let q be the probability that a cell is empty. The probability $R(q)$ that a 2×2 block forms a vertical blocking wall (a path of empty cells from top to bottom) is the sum of probabilities for having 4, 3, or 2 (vertical pair) empty cells [1]:

$$q' = q^4 + 4q^3(1 - q) + 2q^2(1 - q)^2 \tag{5}$$

Solving for the stable state (aside of that where $q = 1$, which happens when all cells are empty) where the system is scale-invariant ($q' = q$) yields the critical value for the density of empty space q_c:

$$q^3 - 2q + 1 = 0 \implies q_c = \frac{\sqrt{5} - 1}{2} \approx \mathbf{0.61803}.$$

Since a cell is either empty (q) or occupied by a tree (p), we have $q + p = 1$, and therefore the critical tree density p_c is:

$$p_c = 1 - q_c \approx \mathbf{0.38197} \tag{6}$$

This implies us that on this simplified version of the model, fire on this grid cannot sustain itself if the tree density is below **38.2%**. However, it is important to note that the uniform thinning intervention we propose does not set the density on the grid to a fixed value. Instead, it multiplies the existing density by a thinning factor ν. Given the threshold for the density $p_c = 0.382$ above, the product of the initial density and the thinning factor must not exceed p_c:

$$D_{initial} \cdot \nu < p_c \tag{7}$$

Solving for the critical thinning factor ν_c:

$$\nu_c = \frac{p_c}{D_{initial}} = \frac{0.382}{D_{initial}} \tag{8}$$

From the tree density plot (Figure 1), we observe that the tree clustesr have a local density $D_{initial}$ ranging roughly between 0.7 and 0.9. If we assume an average density of $D_{initial} \approx 0.70$ within a cluster, the predicted critical factor is $\nu_c = 0.382/0.60 \approx \mathbf{0.545}$. If we assume a denser cluster of $D_{initial} \approx 0.9$, the predicted factor is $\nu_c = 0.382/0.65 \approx \mathbf{0.42}$. Let us proceed with the empirical analysis and cosequently compare the results.

Uniform Thinning. Firstly, we model the case that the tree density on every grid of the cell is reduced by a global factor of $1-\nu$. This is an unrealistic and extremely resource- and logistics-intensive assumption as it requires the cutting down and removal of trees across large areas (Figure 9, left). This will also have severe ecological implications as wildlife and ecosystems are disturbed. However, at extremely high wildfire risks, this might be preferable to a large-scale forest fire. The plot below shows how the average burnt area decreases almost linearly until it dampens and reaches areas extremely close to 0. Depending on the local authorities risk preferences and updated data on the wildfire risk, a critical threshold with regard to the initial tree covered area under no intervention can be defined. The value for the average

[1] This derivation is used from the reading and the work we did in class.

burnt area falls below a threshold of 10% (an arbitrary choice of a relatively low value) when each tree density is decreased by 35%. Additionally, reducing it by another 5 or 10 percent brings this probability down to values extremely close to 0, implying that the fire in this simulation virtually always dies out. As this horizontally compresses the histogram over the tree densities, it strongly moves all of the probability mass of the fire risk within the $[0, 0.2]$ interval (as compared to the initial distribution in Figure 8, right). Ultimately, the theoretical prediction ($\nu_c \approx 0.42 - 0.545$) aligns remarkably well with the empirically produced threshold of $\nu_c \approx 0.6$. The slight deviation is all but surprising given that wind conditions as well as rain events are not modeled in the theoretical exploration, and that the grid has strong local clusters where the scale matters.

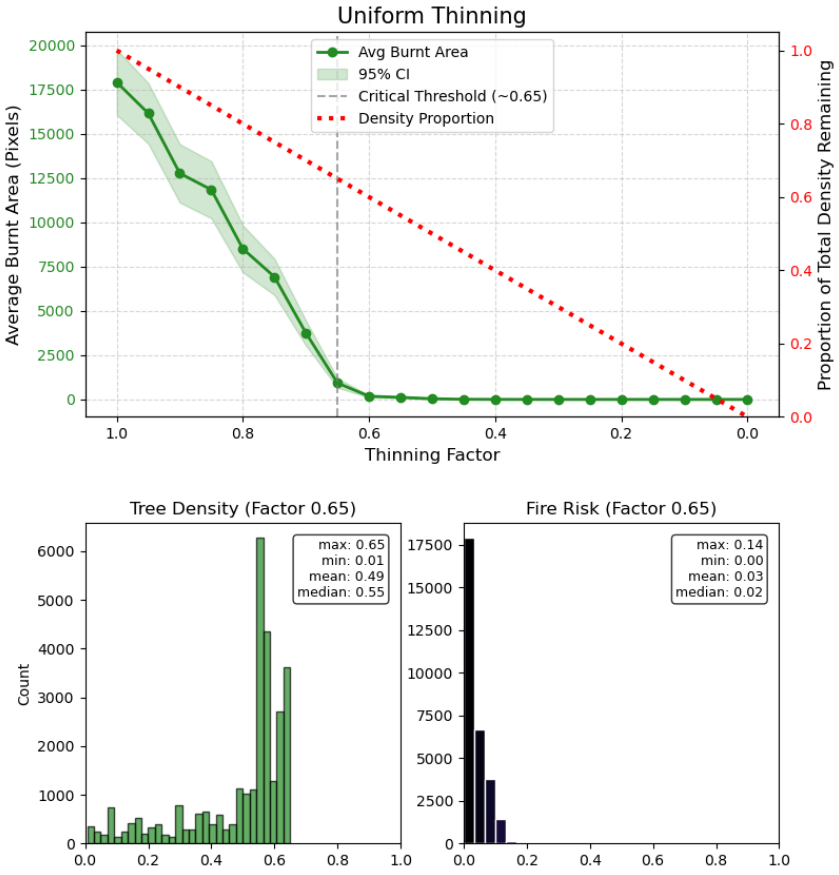

Figure 10: Empirical Effects of the uniform thinning strategy. **Top:** Plotting the thinning factor against the average burnt area with 95% CIs (green) and the remaining total tree density (red). Each simulation was run until convergence, which is defined as staying under the 1% relative change threshold in average values for 10 consecutive trials. **Bottom:** Histograms over the tree density values and fire risk at the critical threshold value of $\nu = 0.65$.

Targeted Thinning. Secondly, we model a more targeted intervention that applies the thinning factor to only the top 50% most tree-dense cells which are more likely to transmit fire (Figure 9, right). On the graph, we also observe a decreasing trend for the average burnt area that is, however, less steep than that of the graph for uniform thinning. This is expectable as the absolute number of trees that are cut down is smaller in the second scenario, where we only target the top 50% densest cells. The 10% threshold is reached for the first time when the tree density in the targeted areas is reduced by 60%, while a probability very close to 0 is only reached when more than 80% of the trees in the highest-density trees are cut down. Note that while a lot of high-density cells remain on the plot (left histogram), a lot of of the formerly most densely populated values have been shifted to the left by the factor of 0.4). Despite the remaining heterogeneity of the tree densities, the fire risk is decreased in a very similar way as for the uniform thinning strategy. Note that in both scenarios for the 10% threshold, a very similar proportion of the initial tree density is removed, leaving about 70% of the original tree 'mass' amid both approaches.

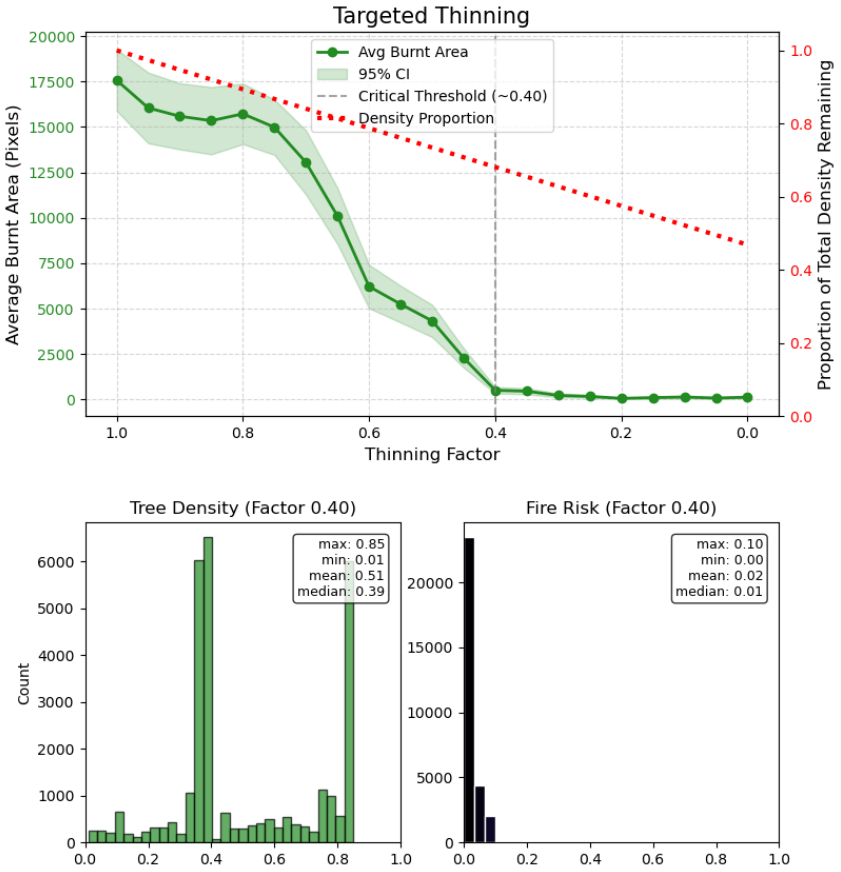

Figure 11: Empirical Effects of the targeted thinning strategy. **Top:** Plotting the thinning factor against the average burnt area with 95% CIs (green) and the remaining total tree density (red). Each simulation was run until convergence, which is defined as staying under the 1% relative change threshold in average values for 10 consecutive trials. **Bottom:** Histograms over the tree density values and fire risk at the critical threshold value of $\nu = 0.65$.

5.2 Intervention 2: Forest Network Intervention

As a response to the 'clustered' nature of the tree distribution across this grid, tools from network theory might be appropriate to model the places with the highest expected 'flow' of fire. If these super-transmitters of fire can be cut down, fires might be prevented from spreading from one cluster to the next. Using the `networkx` library, we can transform the simulation grid into a weighted directed graph $G = (V, E)$, where each cell with biomass constitutes a node v, and edges e_{ij} represent potential fire spread between neighbors. This means that edges exist only between trees that both have a tree density larger than 0, and the weight of the edge is calculated using the fire spread probability equation (2).

Defining shortest paths. Standard algorithms determine shortest paths by minimizing the *sum* of edge weights between two nodes. However, fire spread is probabilistic: the probability of a fire traveling a specific path is the *product* of the individual transmission probabilities along that chain of nodes. To include this in the model gap, we defined the weight of an edge from cell i to cell j using a negative log-likelihood transformation:

$$w_{ij} = -\ln(P_{spread}(i \to j)) \tag{9}$$

This calculation maps the problem of maximizing transmission probability ($\max \prod P_{spread}$) to the equivalent problem of minimizing the path resistance ($\min \sum -\ln P$) which is much easier to evaluate in additive log space than multiplicative linear space. Consequently, the "shortest path" in our graph corresponds mathematically to the "most likely path" of fire spread in the physical simulation.

Betweenness Centrality. With the topology defined, we calculated the *Betweenness Centrality* for every node. This metric quantifies the influence of a node as a "bridge" or "bottleneck" by counting how many shortest paths pass through it. For a node v, the centrality $C_B(v)$ is given by:

$$C_B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{10}$$

where σ_{st} is the total number of most likely fire paths from source s to target t, and $\sigma_{st}(v)$ is the subset of those paths that traverse the node v. High-centrality nodes represent critical "funnels" where fire flow is concentrated; removing them should disproportionately disrupt the connectivity of the network.

Iterative implementation. A critical additional feature of this algorithm is its **iterative nature**. Rather than identifying and removing all N targets based on the initial graph state, we proceed in small batches: we remove the top 10 nodes at a time and recalculate the centrality of the remaining network after each step. This is necessary for stability because the importance of a node depends on the entire network. The removal of a central node affects the flow of the network, potentially shifting the "critical paths" to a completely different region or making previously high-ranking nodes irrelevant.

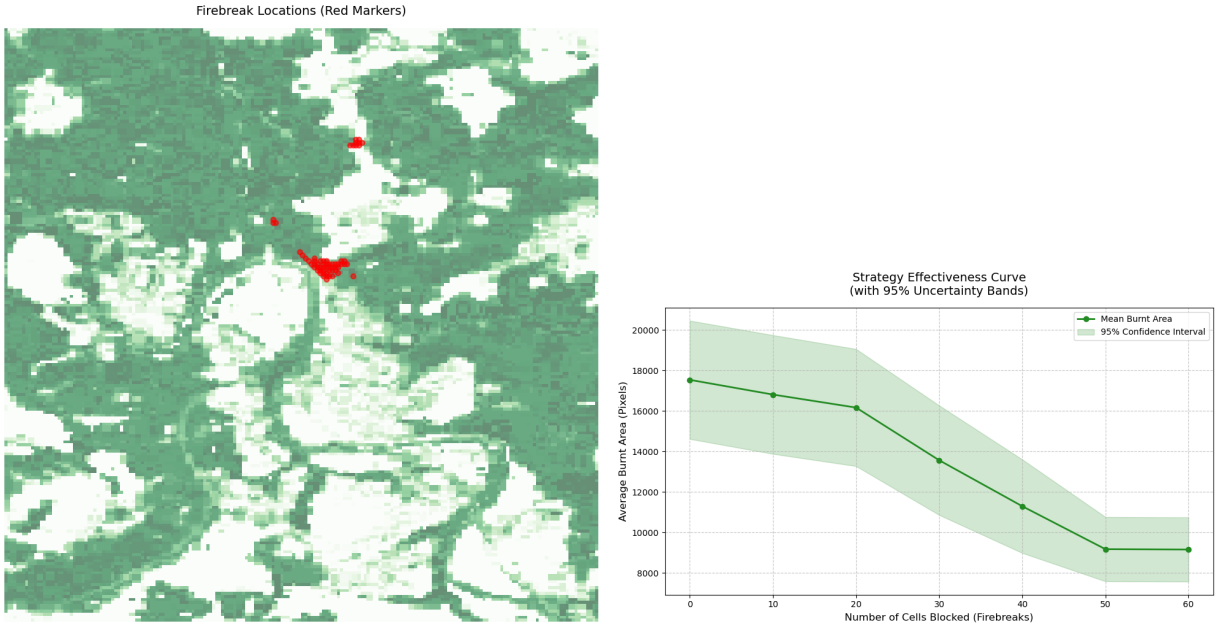

Figure 12: Left: Nodes with the highest betweenness centrality identified by the algorithm are marked in red. **Right:** Average burnt area with 95% confidence intervals after each iteration of the algorithm where $n = 10$ cells are set to zero density.

The determined locations of the highest centrality nodes aligns quite well with their expected locations in Figure 9 as they sit right on the funnel-shaped areas that connect larger clusters. Note also how a few iterations of this algorithm are required before the average burnt area drops faster. The algorithm terminates prematurely after 60 as its highly computationally intense and determined no significant change between blocking 50 or 60 nodes. Ideally, this could be continued for larger numbers to block more of the funnel shapes. However, this shows us that by only blocking 50 of the 29905 tree-populated cells, the average burnt area is effectively reduced by 50%.

In comparison to the first proposed intervention, this strategy is a lot more targeted and provides a significant reduction in the average burnt area by only cutting down 50 trees. On the other hand, a thinning factor of $\nu \approx 0.8$ and $\nu \approx 0.7$ would be needed to achieve a similar result through the uniform and targeted thinning technique respectively, which, in practice, imply either the thinning of *all* forest area by 20%, or the thinning of the densest half of the forest by 30%. Running the first strategy with more computational power to determine the likely further reductions by blocking the paths the fire can take is highly recommended. It should, however, also be noted that the 95% confidence intervals are significantly larger for the second than for the first strategy: only after 40 blocked cells do the confidence intervals of the

average burnt area not overlap anymore and represent a significant effect. While the narrower confidence intervals make us more certain about the resulting average burnt area, the practical implications of this strategy likely outweigh this benefit. As the standard errors are inversely proportional to the square root of the number of trials, halving the size of the confidence intervals would require increasing the number of runs by a factor of 4. While this will require very high computational power, it would increase the level of confidence into the effect of these strategies.

6 Acknowledgments

Thank you again to my friends and classmates Meri Avetisyan & Pedro Paiva for productive study sessions for this assignment.

I also want to state that some of the code was reused and adapted from the Bunny Fox Simulation assignment, which involved some similar computational approaches.

AI Statement: I used AI to help me with some of the numpy vectorization implementations.

References

[1] DLG. (2025, April 3). *Germany's Forests: The Current Situation*. Dlg.org. https://www.dlg.org/en/magazine/germanys-forests-the-current-situation

[2] EFFIS. (2022). *EFFIS - Wildfire Risk Assessment*. Copernicus.eu. https://forest-fire.emergency.copernicus.eu/about-effis/technical-background/wildfire-risk-assessment

[3] EFFIS. (2025). *EFFIS Wildfire Risk Viewer*. Forest-Fire.emergency.copernicus.eu. https://forest-fire.emergency.copernicus.eu/apps/fire.risk.viewer/

[4] Gray, R. (2025, January 9). *Five images that explain why the LA fires spread so fast*. BBC.com. https://www.bbc.com/future/article/20250109-five-images-that-explain-why-the-la-fires-spread-so-fast

[5] Moore, A. (2021, December 3). *Explainer: How Wildfires Start and Spread*. College of Natural Resources News; North Carolina State University. https://cnr.ncsu.edu/news/2021/12/explainer-how-wildfires-start-and-spread/

[6] Westover, R. (2021, August 19). *Thinning the Forest for the Trees*. US Forest Service. https://www.fs.usda.gov/about-agency/features/thinning-forest-trees

https://doi.org/10.1016/j.amc.2022.126986