CS166 Final Project: Wildfire Risk Simulation

Marina Levay

December 19, 2025

Contents

1 Introduction

Wildfires are complex spatiotemporal phenomena driven by local interactions with environmental factors and inherent randomness. At the scale of kilometers and hours, detailed physical modeling of combustion and fluid dynamics may not be computationally tractable. Thankfully, it is not strictly necessary for assessing wildfire risk. Instead, wildfire spread can be successfully modeled as a stochastic process on a spatial grid, where large-scale outcomes emerge from simple local rules governing fuel connectivity and probabilistic spread.

In this project, wildfire dynamics are modeled using a discrete-time cellular automaton informed by percolation theory [1, 6] and analyzed through Monte Carlo simulation [2, 7]. Tree density controls local connectivity in the landscape, giving rise to a percolation-driven phase transition between localized fires and system-spanning burn events. Because ignition location and fire spread are stochastic, repeated simulation runs are required to characterize the distribution of possible outcomes and quantify uncertainty.

By combining percolation-based connectivity intuition with Monte Carlo methods, this report focuses on estimating wildfire risk probabilistically rather than predicting a single deterministic outcome.

A central focus of this analysis is the evaluation of fuel management strategies, specifically mechanical thinning, as a policy tool for risk reduction. We aim to demonstrate whether targeted reductions in fuel density can effectively disrupt fire pathways.

Our key finding: strategic thinning is not merely an incremental improvement but triggers a phase transition, drastically reducing the probability of catastrophic, system-spanning fires. The results produced here can support the analysis of mitigation strategies and can provide statistically grounded guidance for decision-making of authorities, including immediate rescue firefighters and long-term policy makers, under uncertainty.

2 Theoretical Background

2.1 Connectivity and percolation intuition

The implemented model treats the landscape as a 2D grid with a fuel (tree-cover) density $d_i \in [0, 1]$ at each cell. Fire spread is only possible through connected regions of nonzero fuel, so large fires emerge when there exist long connected pathways of sufficiently high-density cells. This is the main way percolation theory informs the simulation: it provides intuition that small reductions in fuel density (e.g., in the thinning technique) can disconnect pathways and cause nonlinear reductions in fire size.

2.2 Stochastic spread and Monte Carlo estimation

The spread mechanism in code is stochastic: for each candidate neighbor ignition, a Bernoulli trial is performed by comparing a uniform random draw to a spread probability. Because ignition location and subsequent spread outcomes are random, we estimate quantities of interest by repeated simulation runs (Monte Carlo) [2, 7], reporting sample means and 95% confidence intervals using the normal approximation of $\bar{X} \pm 1.96\, s/\sqrt{N}$.

3 Model

Building on the connectivity and Monte Carlo ideas described above, we now specify the wildfire simulation model used in this project. The model translates these ideas into a discrete computational system suitable for Monte Carlo analysis on real-world tree-cover data.

3.1 Modeling framework

Wildfire spread is modeled as a stochastic process on a two-dimensional grid. Each Monte Carlo run starts from a single ignition location and then expands a set of burned cells by repeatedly attempting to ignite neighboring cells with a probability that depends on (i) the destination cell's tree density and (ii) the alignment of the spread direction with a prescribed wind vector.

In the implemented code, the simulation does not explicitly track a multi-state cellular automaton (burning vs. burned) or perform globally synchronous time-step updates. Instead, it maintains only a Boolean burned mask and a frontier (stack) of cells whose neighbors have not yet been processed. This produces an "absorbing" terminal state when the frontier is empty. This choice is computationally simple while still capturing the key percolation-style mechanism: connectivity through high-density regions controls whether the burn cluster remains localized or grows large.

Because ignition location and fire spread are stochastic, repeated simulation runs are required to characterize the distribution of possible outcomes. We estimate quantities of interest by reporting sample means and 95% confidence intervals using the normal approximation $\bar{X} \pm 1.96\, s/\sqrt{N}$.

3.2 State variables

Each grid cell i is characterized by two key variables:

- A fixed tree density $d_i \in [0, 1]$, obtained from empirical data.
- A burn indicator $b_i \in \{0, 1\}$ indicating whether the cell ultimately burns in a given run.

Tree density is treated as a static property of the landscape, while the burn indicator is a random outcome induced by the stochastic spread process.

3.3 Percolation-based spread dynamics

Fire spread is modeled using concepts from site percolation theory. Each cell is treated as a potentially occupiable site whose ability to transmit fire depends on its tree density.

For each cell i, a percolation (or occupation) probability is defined as

$$p_i = f(d_i),$$

where $f : [0, 1] \to [0, 1]$ is a monotonic mapping from tree density to the probability that fire can spread into and through that cell. In the simplest case, $f(d_i) = d_i$, but alternative functional forms can be used to reflect nonlinear fuel effects.

Because the landscape is spatially heterogeneous (with spaces of different tree density, as well as land areas and water bodies), there is no single global percolation probability. Instead, fire spread depends on whether local regions of high p_i form connected clusters. This framing explains why small changes in tree density or mitigation strategies can lead to large, nonlinear changes in fire outcomes.

3.4 Local update rule

For a currently burned cell i and a neighboring cell j, the probability of igniting j can be defined as

$$P(i \to j) = \text{clip}\Big(\beta \cdot d_j^{\alpha} \cdot (1 + \gamma \cdot \text{align}_{ij}),\ 0,\ 1\Big),$$

where β is a base spread rate, α is a density exponent, and γ is a wind strength. The alignment term is the cosine similarity between the neighbor direction $i \to j$ and a global wind vector, so spreads aligned with the wind are more likely and spreads against the wind are less likely. The results are clipped because there can exist strong winds in the opposite direction ($\text{align}_{ij} < 0$), and we must constrain the results to probabilities.

In the code, neighbor processing is performed via a frontier (stack) rather than a globally synchronous time step. However, the spread decision for each neighbor is still a Bernoulli trial driven by a uniform random draw [3], which matches the Monte Carlo interpretation above.

3.5 Neighborhood structure

Fire spread in the simulation is governed by a fixed local neighborhood structure that defines which nearby cells may ignite from a burning cell. In this analysis, we adopt a **Moore neighborhood** (8-neighbor connectivity).

The neighborhood definition directly affects spatial connectivity, cluster geometry, and the effective percolation threshold of the system. A Moore neighborhood allows diagonal connectivity, which increases overall connectivity and lowers the effective critical fuel density required for large-scale fire spread.

This better reflects real wildfire behavior at the chosen spatial resolution, where fire spread is not confined to cardinal directions due to irregular vegetation patterns, ember transport, and wind-driven effects.

3.6 Boundary conditions

The grid uses fixed (non-periodic) boundary conditions. Fire is not allowed to spread beyond the edges of the grid, reflecting the finite spatial extent of the real geographic region being modeled.

While periodic boundary conditions are often used in theoretical studies to eliminate edge effects, they are not appropriate for this application because they introduce artificial wraparound connectivity. Fixed boundaries preserve physical interpretability, though they may reduce estimated risk near the edges of the grid. This effect is acknowledged when interpreting per-cell risk maps.

3.7 Ignition distribution and terminal state

If an ignition location is not specified, the ignition cell is sampled with probability proportional to its density value d_i. Cells with $d_i = 0$ (e.g., water/no-data after masking) therefore never ignite. A run terminates when no additional neighbor ignitions occur, yielding a final burned mask.

3.8 Model parameters

All parameters of the model are defined explicitly and grouped as follows:

- Tree-density to percolation mapping function $f(d)$.
- Wind direction and wind strength parameters defining w_{ij}.
- Mitigation parameters, such as thinning intensity or firebreak placement, which modify local tree densities and can shift regions below the percolation threshold.

This organization ensures that all assumptions and control variables are clearly specified in a single location, supporting reproducibility and interpretability.

3.9 Assumptions and scope

The model assumes that fire spread is primarily governed by fuel connectivity and wind effects. Processes such as moisture variation, terrain slope, and heterogeneous fuel types are abstracted away. As a result, the model is best interpreted as a tool for comparative risk analysis and strategy evaluation rather than a physically exact fire prediction system.

4 Extensions (optional)

4.1 Domain masking with `datamask`

An immediate extension enabled by the Hansen GFC layers is to treat the landscape grid as a *masked domain* rather than a fully active lattice. The `datamask` raster distinguishes land (1), permanent water (2), and no-data (0). After cropping and downsampling to the simulation resolution, we aggregate `datamask` over each coarse cell (e.g., by majority or by threshold on the fraction of water pixels). Coarse cells classified as water or no-data are assigned density. 0. This has two important effects:

- It prevents permanent water from contributing fuel, so such cells never ignite.
- It alters connectivity by introducing holes and barriers in the lattice, changing the geometry of percolation pathways and therefore the distribution of burned cluster sizes.

In this masked-domain view, mitigation strategies (thinning, firebreaks) operate only on the remaining land cells, and summary statistics such as burned fraction can be computed either over all grid cells (treating masked cells as always-unburnable) or restricted to land cells only.

4.2 Time-varying fuel maps from `lossyear` and `gain`

A second extension is to run the same stochastic fire spread model on a *time series* of fuel maps derived from the baseline year 2000 canopy cover. Given the baseline density field $d^{(2000)}$ obtained from `treecover2000`, we construct annual densities $d^{(y)}$ by applying two simple transformations:

- **Forest loss:** if `lossyear` is nonzero and corresponds to a loss event in or before year y, we set that cell's density to 0 from year y onward.
- **Forest gain:** if `gain`=1, we allow a modest density increase after 2012 (capped at 1.0) to represent regrowth, acknowledging that `gain` is an indicator rather than a year-by-year measurement.

The `datamask` constraint is enforced for all years so that water/no-data cells remain at density 0 throughout the series. This time-varying setup supports longitudinal questions such as how simulated fire risk evolves under observed loss patterns, and it enables comparing mitigation policies across years while holding the spread mechanism fixed.

5 Empirical data and preprocessing

5.1 Hansen Global Forest Change layers

We use the Hansen/UMD/Google/USGS/NASA Global Forest Change dataset (GFC 2019 v1.7), which provides global rasters in $10^\circ \times 10^\circ$ tiles at approximately 30m resolution. The primary layer used for fuel density is `treecover2000`, which encodes canopy closure (0–100) for vegetation taller than 5m in the year 2000. We also incorporate:

- `datamask`: distinguishes no-data (0), land (1), and permanent water (2).
- `lossyear`: year of forest loss event (0 for no loss; 1–19 for 2001–2019).
- `gain`: forest gain indicator during 2000–2012 (0/1).

5.2 Region selection and downsampling

We focus on a region in the wilderness of northern Canada using the Hansen tile `60N_120W` (approximately 60°N, 120°W). From this selected tile, we crop a square window of the raster and then downsample it by mean-pooling to a simulation grid of size 200×200 (satisfying the assignment requirement of at least 100×100 cells). The resulting density values are normalized to $[0, 1]$ by dividing `treecover2000` by 100.

5.3 Water/no-data masking

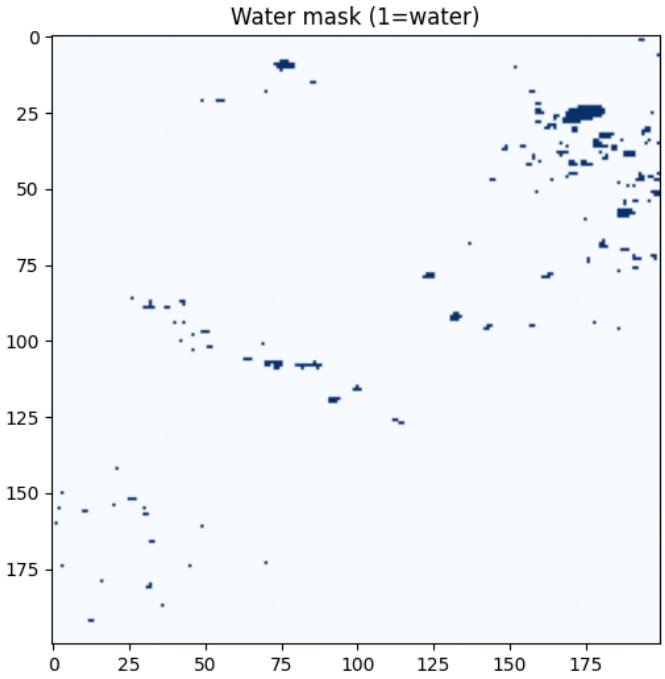

Figure 1: Water mask applied to the tree coverage region.

To prevent treating permanent water as forest fuel, we apply `datamask` after downsampling. Cells dominated by water are assigned density 0. Likewise, cells with no data are assigned density 0. This ensures such cells never ignite and never transmit fire in the simulation.

6 Empirical experiments

6.1 Time series construction (2000–2019)

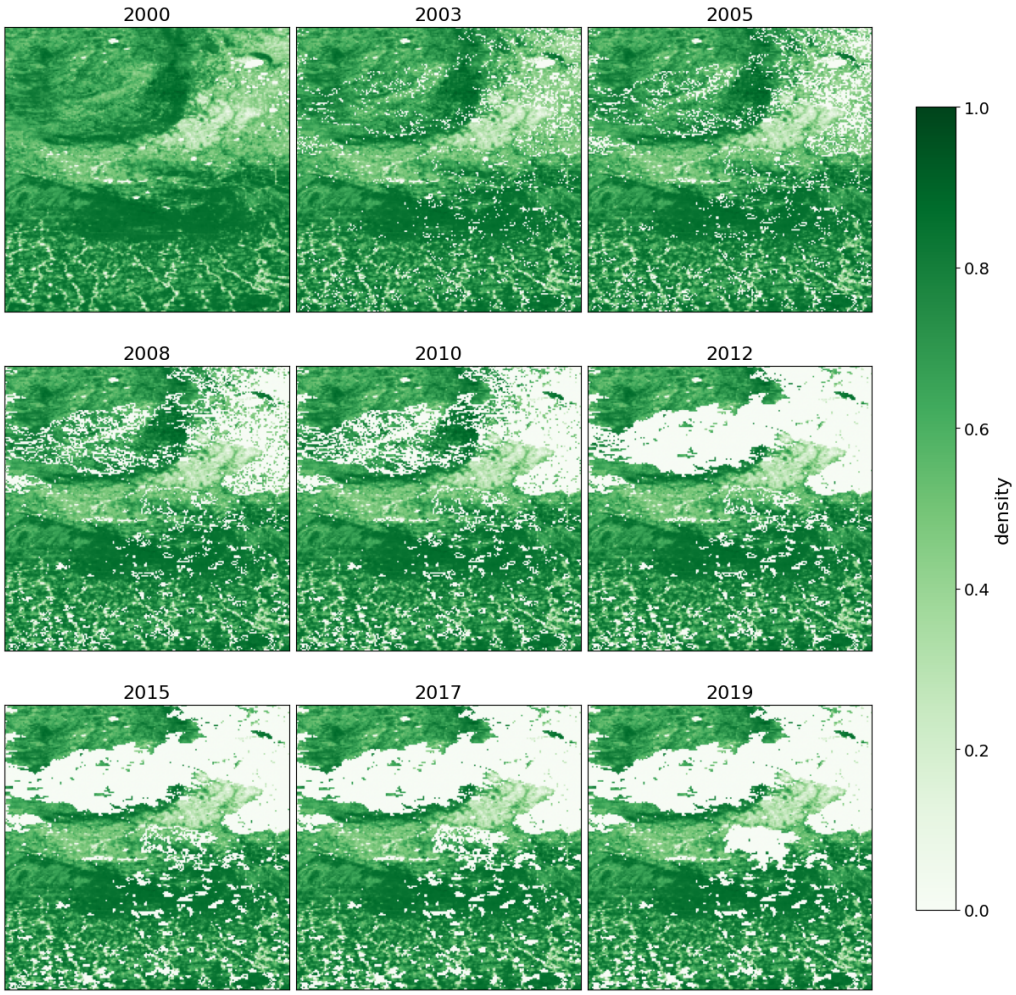

Figure 2: Forest tree density in the study region from 2000 to 2019, derived from the Global Forest Change dataset (Hansen et al., 2013). Darker green indicates higher vegetation density. Over time, large contiguous forest patches become increasingly fragmented and reduced in extent. This decline in tree cover likely reflects a combination of wildfire impacts and other potential disturbances (such as land-use change, disease). This results in reduced fuel connectivity that directly influences fire-spread dynamics in the percolation-based simulations.

To produce a consistent time series of fuel maps, we define a derived density field $d^{(y)}$ for each year $y \in [2000, 2019]$ from the 2000 baseline density $d^{(2000)}$ using simplified, explicit assumptions:

- **Loss:** if `lossyear` indicates a loss event in or before year y, that cell's density is set to 0 from year y onward (stand-replacing disturbance assumption).
- **Gain:** if `gain`=1 and $y \geq 2012$, density is increased by a small constant (capped at 1.0) to represent modest regrowth.
- **Masking:** `datamask` is enforced for all years, so water/no-data cells remain at density 0.

6.2 Experimental conditions

All simulations use the same fire spread parameters unless otherwise stated:

- Wind vector: $(1.0, 0.0)$
- Wind strength: 0.4
- Density exponent: 1.2
- Base spread: 0.7
- Neighborhood: Moore (8-neighborhood)
- Monte Carlo runs per condition: $N = 200$

We analyze the years $\{2000, 2003, 2005, 2008, 2010, 2012, 2015, 2017, 2019\}$ to illustrate how risk changes under forest loss/gain over time.

6.3 Compared strategies

We compare two strategies for each analyzed year:

- **Baseline:** no mitigation.
- **Thinning:** density multiplied by a factor of 0.6 (uniform thinning), representing fuel reduction.

These two strategies provide a controlled comparison of wildfire outcomes with and without fuel reduction, with uniform thinning at 60% density used as a simple, interpretable intervention to reduce fuel connectivity while preserving substantial forest cover.

6.4 Measured outputs and visualizations

For each year and strategy, we compute:

- **Burned fraction:** fraction of cells that burn.
- **Affected fraction:** fraction of cells that are burned or adjacent to a burned cell. This measures the broader "zone of danger," including areas not directly consumed but threatened by proximity to the fire front.
- **Risk map:** for each cell, the Monte Carlo estimate of the probability of being affected.

We visualize outcome distributions with histograms of burned fraction, and visualize spatial risk with heatmaps of the risk map. We also summarize the mean burned fraction across selected years.

6.5 Preliminary results and interpretation

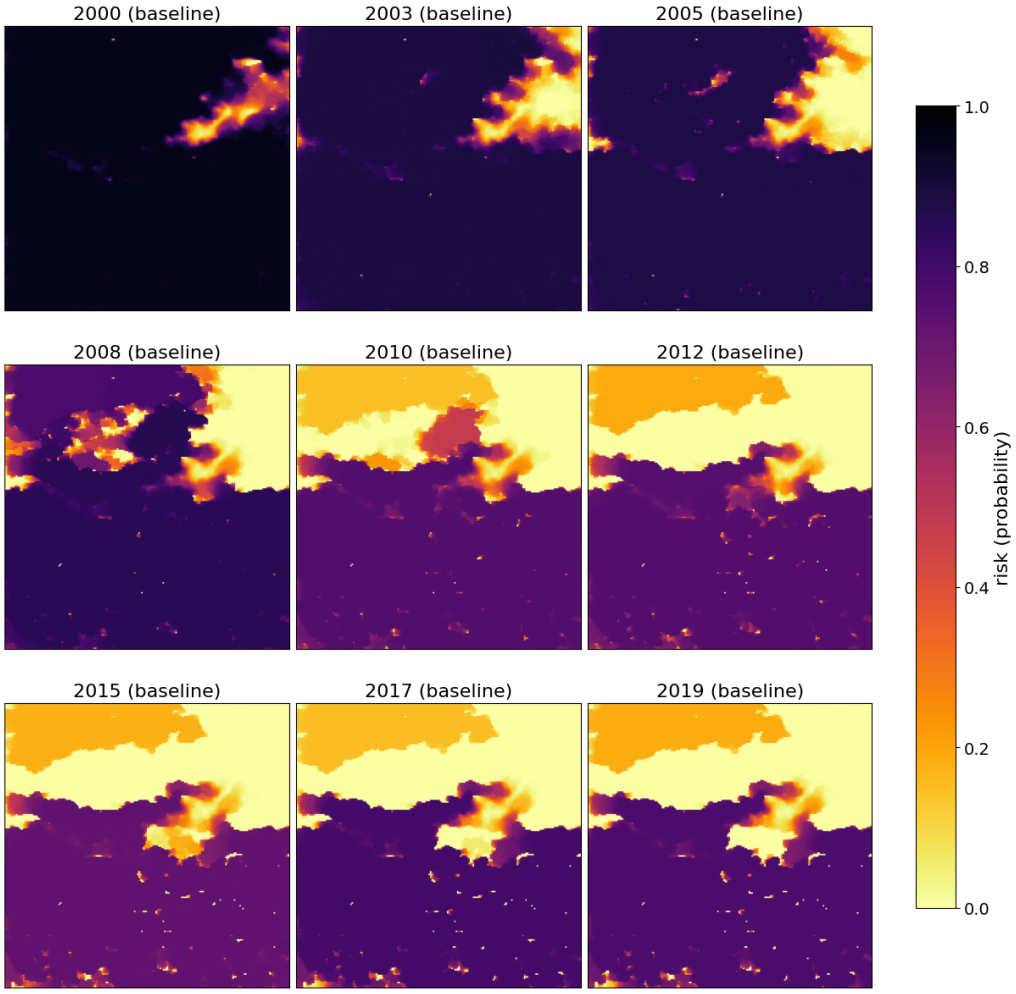

Figure 3: Spatial wildfire risk maps under baseline (no mitigation) conditions across multiple years. Color indicates the probability that each grid cell is burned or adjacent to a burned cell across Monte Carlo runs. Persistent high-risk corridors indicate stable fuel-connectivity patterns that dominate fire spread.

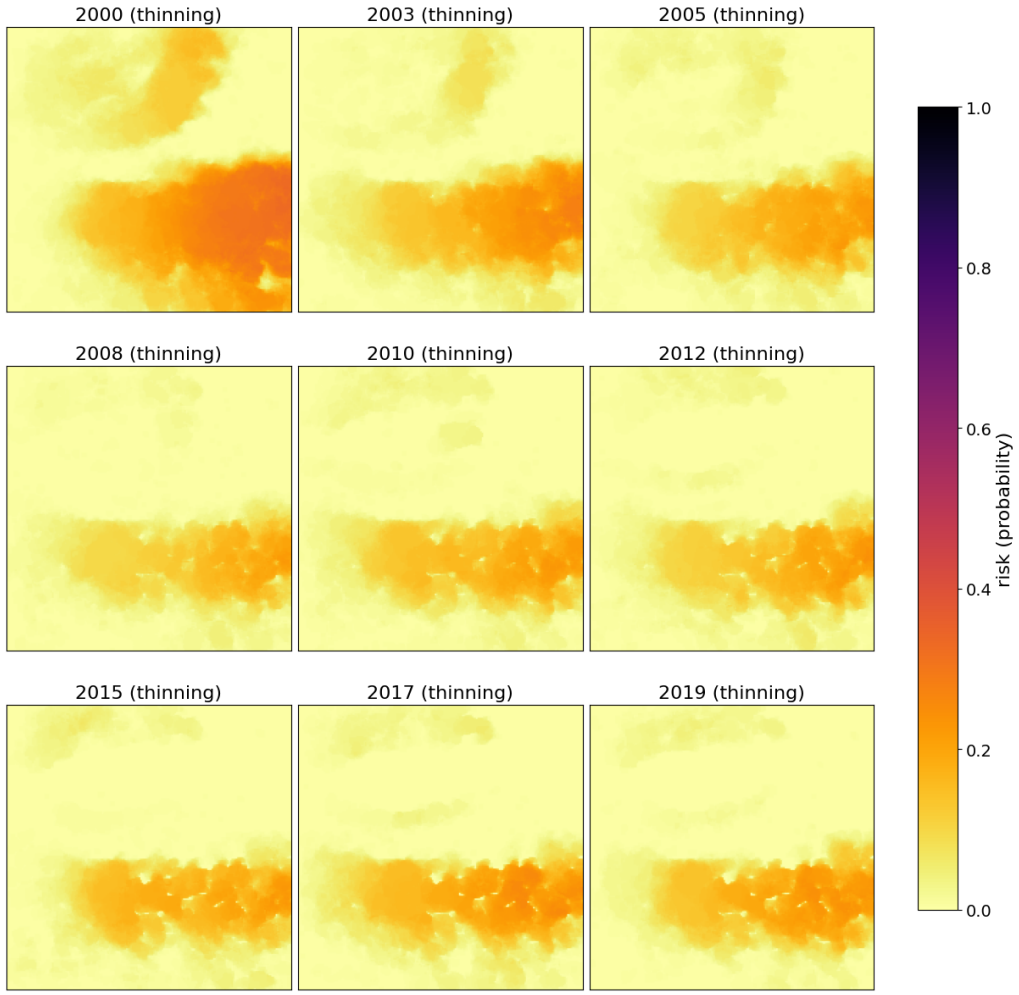

Figure 4: Spatial wildfire risk maps under uniform thinning (density factor 0.6) across multiple years. Overall risk is substantially reduced and more spatially diffuse, reflecting disrupted fuel connectivity and reduced likelihood of large-scale fire spread.

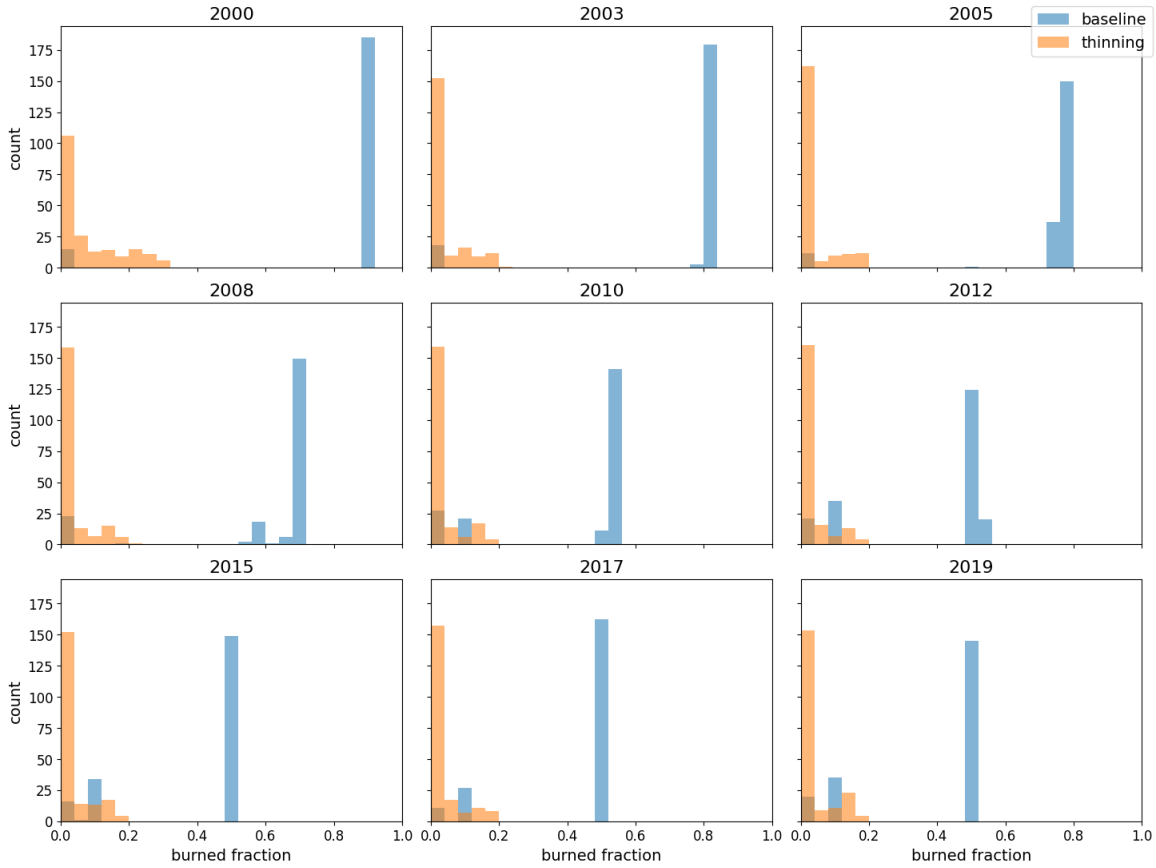

Figure 5: Distributions of burned fraction across Monte Carlo simulation runs for multiple years, comparing baseline conditions (blue) and thinning treatment (orange). Baseline distributions are mainly bimodal, while thinning shifts outcomes toward low burned fractions across all years.

Figure 5 shows that, across all analyzed years, baseline simulations exhibit a strongly bimodal distribution of burned fraction. Most Monte Carlo runs result either in negligible burns or in system-spanning fires that consume a large fraction of the domain. This behavior is characteristic of a supercritical percolation regime, in which connected fuel pathways allow fire to propagate across much of the landscape once ignition occurs within the dominant connected cluster.

The spatial manifestation of this behavior is visible in the baseline risk maps (Figure 3), which reveal persistent, spatially coherent high-risk corridors that remain stable across years. These corridors indicate that fuel connectivity, rather than stochastic ignition alone, governs fire spread. Water and no-data regions act as effective barriers, shaping fire pathways and confirming that domain masking materially alters connectivity.

Applying uniform thinning (density factor 0.6) produces a qualitative shift in fire behavior. Under thinning, burned-fraction distributions collapse toward low values, and system-spanning fires become rare across all years (Figure 5). Additionally, risk maps under thinning (Figure 4) show substantially reduced and more diffuse risk, with previously continuous high-risk corridors weakened or fragmented.

Taken together, these results indicate that thinning operates not merely by reducing average fire size, but by altering the connectivity structure of the landscape. The observed transition from predominantly system-spanning fires to mostly localized burns is robust across years and aligns with theoretical expectations from percolation-based wildfire models.

6.6 Temporal robustness of wildfire impacts

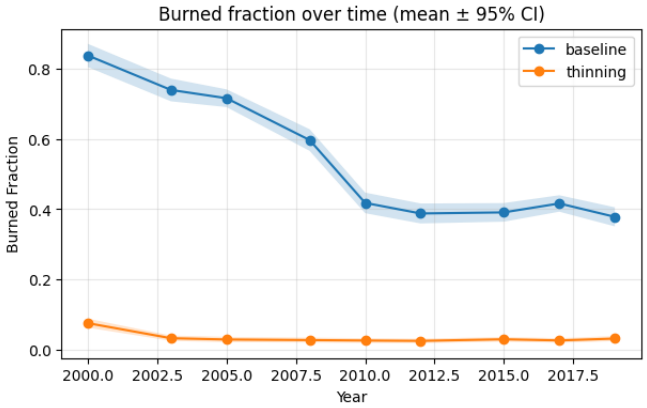

Figure 6: (a) Mean burned fraction

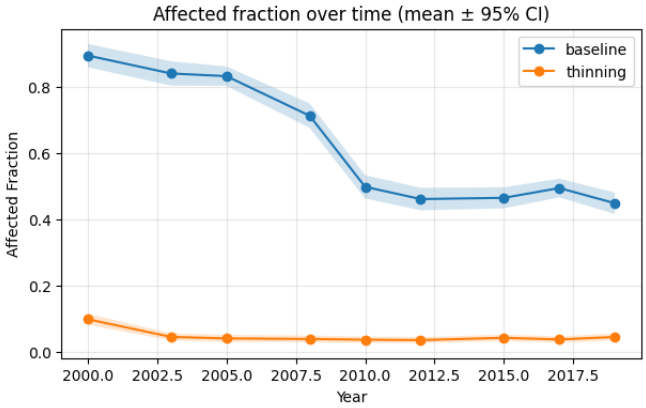

Figure 7: (b) Mean affected fraction

Figure 8: Temporal evolution of wildfire impact metrics under baseline conditions and uniform thinning (density factor 0.6). Lines show mean values across Monte Carlo runs, with shaded regions indicating 95% confidence intervals.

Figure 8 compares the temporal evolution of burned fraction and affected fraction under baseline and thinning scenarios. Under baseline conditions, both metrics remain high across all years, indicating a persistent large-scale fire impact despite gradual changes in forest density over time. In contrast, thinning consistently suppresses both burned and affected fractions to low levels, with narrow confidence intervals indicating that the reduction is statistically robust across Monte Carlo runs.

The close correspondence between burned and affected fractions reflects the strong spatial coupling created by fuel connectivity: when fires spread extensively, a large surrounding region is also impacted. Thinning disrupts this connectivity, leading to localized fires with limited secondary impact. Together, these results demonstrate that thinning reduces not only direct fire damage but also the broader spatial footprint of wildfire, and that this effect is stable across the multi-year time series.

6.7 Uncertainty quantification

We explicitly quantified the uncertainty of our Monte Carlo estimates. With $N = 200$ runs, the maximum 95% confidence interval (CI) width observed was 0.0661 (Baseline, Year 2000), with an average width of 0.0569 across all scenarios. This number of runs was selected due to practical computational constraints (the total running time is currently approximately 10 minutes).

This indicates that our intervals are relatively wide ($> 5\%$), reflecting the moderate uncertainty inherent in the stochastic simulation near a supercritical percolation threshold. While the current sample size is sufficient to distinguish the large effect sizes between baseline and thinning scenarios, more precise point estimates would require increasing the number of runs. Specifically, to achieve a target CI width of 0.02 (mean $\pm 1.0\%$) for the worst-case year, we estimate that approximately $N \approx 2181$ runs would be required, an increase of 1981 runs over the current setup. This scaling follows the standard error proportionality $1/\sqrt{N}$.

6.8 Significance of mitigation

To confirm that the observed reduction in fire impact is statistically significant, we performed a one-sided permutation test comparing baseline and thinning scenarios for each year, correcting for multiple comparisons using the Holm–Bonferroni method [4, 5]. The null hypothesis (H_0) for each test is that the thinning

strategy does not reduce the mean fire impact compared to the baseline (i.e., $\mu_{\text{thin}} \geq \mu_{\text{base}}$).

The results confirm that thinning produces a statistically significant reduction in both burned and affected fractions for every year analyzed (adjusted $p < 0.002$ for all years; null hypothesis rejected). The estimated effect sizes (differences in means) are substantial, ranging from $\Delta \approx 0.35$ to $\Delta \approx 0.79$, with the largest reductions observed in earlier years when baseline fuel connectivity was highest.

Year	Δ Burned	p_{adj} (Burned)	Reject?	Δ Affected	p_{adj} (Affected)	Reject?
2000	0.7627	0.0018	Yes	0.7951	0.0018	Yes
2003	0.7073	0.0018	Yes	0.7942	0.0018	Yes
2005	0.6874	0.0018	Yes	0.7907	0.0018	Yes
2008	0.5698	0.0018	Yes	0.6730	0.0018	Yes
2010	0.3921	0.0018	Yes	0.4608	0.0018	Yes
2012	0.3630	0.0018	Yes	0.4252	0.0018	Yes
2015	0.3616	0.0018	Yes	0.4220	0.0018	Yes
2017	0.3903	0.0018	Yes	0.4564	0.0018	Yes
2019	0.3468	0.0018	Yes	0.4032	0.0018	Yes

Table 1: Significance testing results for the efficacy of thinning, corrected for multiple comparisons across years.

6.9 Theoretical vc Empirical validation

To validate that the simulation captures the expected nonlinear relationship between density and connectivity, we performed a density scaling sweep (Figure 9). By scaling the density of the landscape uniformly by a factor $c \in [0, 1]$, we observed the characteristic phase transition predicted by percolation theory.

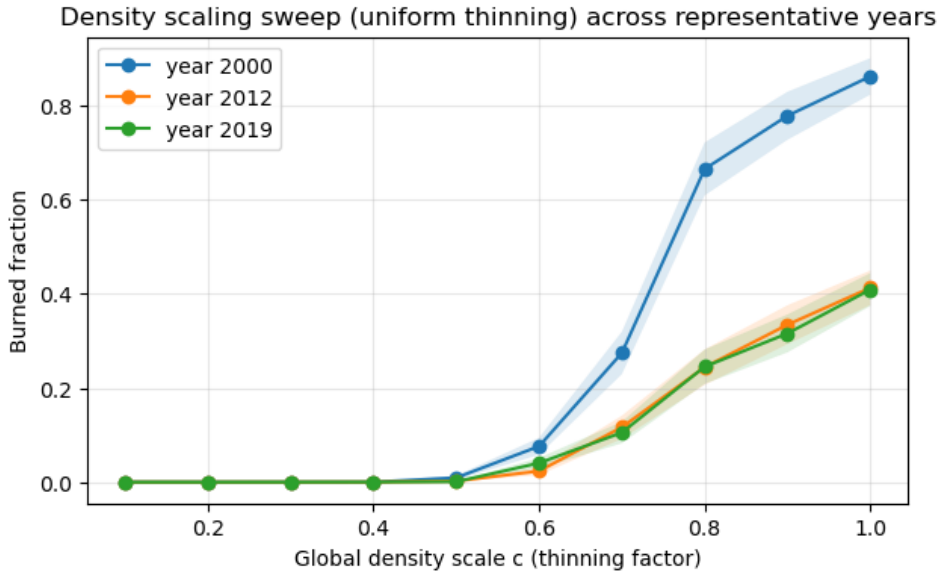

Figure 9: Empirical density scaling sweep for three representative years. The curves show the mean burned fraction as a function of a uniform density scaling factor c. The sharp nonlinearity, where outcomes transition rapidly from negligible to system-spanning fires, is a hallmark of a percolation phase transition.

The transition from low to high burned fraction occurs sharply between scaling factors of 0.6 and 0.8. This aligns with our main results, explaining why thinning to 0.6 (just below the transition) is so effective, while the baseline density (near or above the transition) enables system-spanning fires. The consistency of

this curve across the years 2000, 2012, and 2019 suggests that the fundamental connectivity structure of the landscape remains similar despite local changes in forest cover.

7 Conclusion

Executive Summary for Stakeholders This simulation study demonstrates that wildfire risk in the target region is driven primarily by fuel connectivity. Our analysis reveals that the landscape currently sits in a "supercritical" state, where continuous vegetation pathways allow small ignitions to grow into system-spanning mega-fires. We find that a uniform thinning strategy (reducing density to 60%) is highly effective. It does not simply reduce fire size linearly; rather, it triggers a phase transition that fragments the fuel network, effectively eliminating the possibility of catastrophic spread. This intervention is robust across 20 years of historical data and is statistically significant. For policymakers, this suggests that fuel management strategies focused on breaking connectivity are far more effective than those that merely suppress ignition.

Technical Summary This project successfully modeled wildfire spread on a real-world landscape using a percolation-based cellular automaton analyzed via extensive Monte Carlo simulations. By integrating empirical data from the Hansen Global Forest Change dataset, we constructed a spatially explicit simulation that captures the critical role of fuel connectivity in driving large-scale fire risks.

Our results demonstrate that the landscape exists in a supercritical percolation regime, where high fuel connectivity facilitates system-spanning fires. We showed that a simple uniform thinning strategy (reducing density to 60%) is highly effective, not just reducing fire size linearly, but triggering a phase transition that shatters the connected fuel network and suppresses large fires almost entirely. This effect was shown to be statistically significant and robust across a 20-year time series.

The uncertainty analysis highlights that while our findings are statistically robust, the inherent stochasticity of the system near the critical point leads to moderate variability in individual outcomes. Future work could improve precision by increasing simulation runs by an order of magnitude or by incorporating more complex factors such as variable wind fields and moisture dynamics. Nevertheless, the current analysis provides compelling evidence that connectivity-based mitigation is a powerful tool for wildfire risk reduction.

8 Appendix

8.1 Code Repository Setup

The codebase is available in a public GitHub repository (necessary to clone the code locally into an external notebook). It's organized into a modular Python package for the core simulation logic and a Jupyter Notebook for experiment visualization.

Simulation Package The core modeling logic is encapsulated in the `simulation/` directory, consisting of three files:

- `forest.py`: Defines the `Forest` class, which manages the landscape grid state. It handles the storage of fuel density values and implements mitigation strategies such as thinning by creating modified copies of the landscape.
- `monte_carlo_model.py`: Contains the `MonteCarlo` class, which executes the stochastic fire spread algorithm. It implements the optimized percolation logic, handling neighbor connectivity, wind-biased spread probabilities, and repeated simulation runs to estimate risk distributions. It also includes the `FireModelParams` data structure for managing physical parameters.
- `data_collector.py`: Provides the `DataCollector` class, responsible for aggregating results from multiple Monte Carlo runs. It handles the storage of burned and affected fractions and computes statistical metrics, including 95% confidence intervals.

Experimentation and Visualization The `wildfire-visualization.ipynb` notebook serves as the main entry point for the analysis. It integrates the simulation modules to:

1. Load and preprocess the Hansen Global Forest Change data (creating density and mask arrays).
2. Configure and execute the Monte Carlo simulations across different years and mitigation scenarios.
3. Generate the statistical plots (histograms, time series) and spatial risk maps presented in this report.

References

[1] CS166 Session 8. (2024). *Percolation and renormalization.* Minerva University. https://forum.minerva.edu/app/courses/3704/sections/12807/classes/96243

[2] CS166 Session 10. (2024). *Simulations with random numbers.* Minerva University. https://forum.minerva.edu/app/courses/3704/sections/12807/classes/96247

[3] CS166 Session 11. (2024). *Random numbers from simulations.* Minerva University. https://forum.minerva.edu/app/courses/3704/sections/12807/classes/96249

[4] Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65–70. https://www.jstor.org/stable/4615733

[5] Laerd Statistics. (n.d.). *The Holm–Bonferroni method.* https://statistics.laerd.com/statistical-guides/holm-bonferroni-method-statistical-guide.php

[6] Sayama, H. (2015). *Introduction to the modeling and analysis of complex systems.* Open SUNY Textbooks, Milne Library. https://milneopentextbooks.org/introduction-to-the-modeling-and-analysis-of-complex-systems/

[7] Shonkwiler, R. W., & Mendivil, F. (2009). *Explorations in Monte Carlo methods.* Springer. https://doi.org/10.1007/978-0-387-87858-2

8.2 AI Statement

Given the depth and length of the report, I used most of my time to derive a thorough, written analysis. Consequently, I used Gemini 3 Pro in my IDE for faster code support. Namely, the auto-complete feature for writing docstrings efficiently, formatting the output images on 3x3 grids, and adding helper methods in `monte_carlo_model.py` for speeding up the simulation by pre-computing weights and probabilities. This was essential to cut down simulation time in the notebook from roughly 1 hour to 10 minutes. I also used Grammarly AI throughout my report to avoid typos and grammatical mistakes for #professionalism.

The Holm-Bonferroni correction section was inspired by both FA50 and FA51 knowledge, as well as drawing intuition from CS130: Causal Inference in measuring the effectiveness of alternative hypotheses against the null.

The LaTeX-formatted table was adapted from prior Mathematics assignments.