
Reinforcement Learning for Wildfire Mitigation in Simulated Disaster Environments

Alexander Tapley* Marissa Dotter Michael Doyle Aidan Fennelly
Dhanuj Gandikota Savanna Smith Michael Threet Tim Welsh
The MITRE Corporation - *AI Security and Perception*

{atapley, mdotter, mdoyle, afennelly, dgandikota, sosmith, mthreet, twelsh}@mitre.org

Abstract

Climate change has resulted in a year over year increase in adverse weather and weather conditions which contribute to increasingly severe fire seasons. Without effective mitigation, these fires pose a threat to life, property, ecology, cultural heritage, and critical infrastructure. To better prepare for and react to the increasing threat of wildfires, more accurate fire modelers and mitigation responses are necessary. In this paper, we introduce SimFire, a versatile wildland fire projection simulator designed to generate realistic wildfire scenarios, and SimHarness, a modular agent-based machine learning wrapper capable of automatically generating land management strategies within SimFire to reduce the overall damage to the area. Together, this publicly available system allows researchers and practitioners the ability to emulate and assess the effectiveness of firefighter interventions and formulate strategic plans that prioritize value preservation and resource allocation optimization. The repositories are available for download at <https://github.com/mitrefireline>.

1 Introduction

The global effects of climate change such as drought and increased temperatures are exacerbating the frequency and intensity of wildfires [1] and in-turn increasing the effects of climate change through excessive carbon-dioxide emissions [2] and significant terrain change. Severe wildfires pose a significant threat to life, property, ecology, cultural heritage, and critical infrastructure - in 2022 alone, over 7.6 million acres of land was burned due to wildfires across the United States [3] costing over 3.5 billion dollars in suppression costs [4]. While wildfires are an essential natural occurrence for maintaining healthy ecological systems [5], uncontrolled fires, particularly those close to the Wildland Urban Interface (WUI), can present significant risks to public health, life and property, necessitating effective management or suppression measures.

In this paper, we introduce SimFire and SimHarness, a Python-based system to accurately model wildland fire spread and generate appropriate mitigation strategy responses via Reinforcement Learning (RL). SimFire utilizes the Rothermel fire spread formula [6] to simulate the movement of wildland fire through an environment generated from real-world operational or procedurally generated fictional data. Simulated agents can be added to the generated environment and place firefighter-induced mitigations to visualize how wildfires will react to specific mitigation strategies. SimHarness is a machine learning harness designed to train RL agents within SimFire to identify the optimal mitigation strategies for a given wildfire scenario. The combination of SimFire and SimHarness provides a customizable system designed to simulate the spread of wildland fire through a generated

*Corresponding author.

environment and suggest optimal mitigation strategies for the given scenario. The repositories are available for download at <https://github.com/mitrefireline>.

1.1 Related Works

Existing fire spread models [7, 8] and visualization tools [9, 10] have brought value to the decision making and planning process for fire chiefs, burn bosses, and land managers. SimFire and SimHarness aims to derive even more insights for planners and mitigators by leveraging agent-based machine learning that identifies the optimal strategies for addressing wildland fire scenarios. In recent years, there has been an increase of academics studying RL for disaster relief and response. In [11, 12], both provide open-source RL environments and models for training agents to mitigate the spread of wildfire to limit overall damage. Altamimi [13] similarly trains an RL agent to mitigate the spread of wildfire through an environment, but does not include open-source code. In all these cases, the environments do not support using real-world data during terrain generation or true fire-spread models such as Rothermel [6]. SimFire aims to fill this gap with realistic environments, research-backed fire spread capabilities, and a design that supports further improvement by the open-source community. Similarly, SimHarness' modular structure makes it compatible with any disaster modeler that utilizes the required Simulation API - not just wildland fire - making SimHarness more flexible than current frameworks and extensible for a variety of disaster scenarios.

2 Background and Preliminaries

2.1 Rothermel Fire Projection

The Rothermel surface fire spread model has been widely used in the field of fire and fuels management since 1972 [6]. To model the spread of fire across a given surface, the Rothermel equation takes fuel moisture and wind into account for weather conditions, and slope and elevation into account for terrain conditions. These environmental conditions, weather and terrain, pair with input fuel complexity values to determine the spread of a fire throughout an environment. While Rothermel is considered a valuable tool for estimating the rate of fire spread under certain conditions, its accuracy and applicability can vary depending on factors such as the specific environmental conditions, fuel types, and terrain. Researchers and practitioners often use Rothermel as part of a suite of tools with weather data to better understand and manage wildfires.

2.2 Reinforcement Learning

Reinforcement learning is an agent-based sub-field of machine learning that utilizes a user-designed reward function to train an intelligent agent how to interact within an environment and achieve a desired goal [14]. In RL, the environment is defined as a Markov Decision Process [15], MDP, $M = (S, A, P, \rho_0, R, \gamma, T)$, where S is the state space, A is the action space, $P : S \times A \times S \rightarrow [0, 1]$ is the state transition probability, $\rho_0 : S \times A \rightarrow [0, 1]$ is the initial state probability, $R : S \times A$ is the reward function, γ is the discount factor, and T is the maximum episode length. The policy $\pi_\theta : S \times A$ assigns a probability value to an action given a state.

Throughout training, the agent receives an observed state from the environment $s_t \in S$, representing the state space information currently available to the agent, and performs an action $a_t \in A$ according to its policy π_θ , or, at times, a random policy to encourage exploration. After the agent interacts with the environment via the given action, the environment returns both a next state $s'_t \in S$ and reward $r_t \in R$. The agent is trained to find a policy that optimizes the user-defined reward function R .

3 System Components

As wildfire frequency and severity increases, it is clear innovation is needed to generate more effective wildfire management and mitigation strategies. SimFire and SimHarness work as a single system to both accurately simulate the spread of a fire through a user-specified environment and suggest mitigation strategies for the given scenario to reduce the fire severity risk and lessen the financial, ecological, and public health impacts that agencies manage.

3.1 SimFire

SimFire is an open-source Python tool that simulates realistic wildfire spread over generated environments. These generated environments can be created using procedurally generated fictional data or using topographic and fuel data sources available through LANDFIRE [16] to model real-world environments. When using real-world data, users can specify a year to gather data from and an area’s GPS coordinates to create a realistic training environment.

SimFire introduces the `Simulation` class, which is a base class to support a simulated disaster environment. This parent class provides the API necessary for `SimHarness` to train agents within the environment. `FireSimulation`, a child class of `Simulation`, provides a simulated wildfire disaster environment based off the Rothermel equations provided by Andrews [17] as the basis for the fire spread model. Through a configuration file, users can adjust the simulated environment’s size, terrain, fuel sources, and wind dynamics - including complex wind dynamics based off of Navier-Stokes equations [18]. SimFire provides a variety of fuel configurations out-of-the-box, including the standard 13 Anderson Behavior Fuel Models [16, 19], and supports the addition of user-specified fuel types as well. Additionally, users can configure aspects of the fire itself, such as ignition location, rate of spread attenuation, and max fire duration for a single space. The library allows researchers and wildland fire managers control over the scenarios used in their mitigation experiments.

In addition to the fire spread model, SimFire supports the placement of different mitigations to control the spread of fire. Mitigations such as firelines, scratchlines, and wetlines can be placed at any pixel within the simulated environment, allowing users to experiment with different mitigation strategies to see how the fire reacts in certain scenarios. SimFire employs PyGame [20], a scalable and highly-optimized game simulation Python library to visualize the fire spread, agent movements, and agent interactions within the environment. The implemented algorithms and formulas along with the flexibility provided by SimFire allow researchers to define different test scenarios and locations for their mitigation experiments. Additional information about SimFire’s fire spread verification, data layers, and agent actions can be found in Appendix A along with example fire scenarios.

3.2 SimHarness

SimHarness is a Python repository designed to support the training of RLlib [21] RL algorithms within simulated disaster environments. SimHarness takes as input an instance of the `Simulation` class, such as SimFire’s `FireSimulation`, as the training environment. The `Simulation` object provides an API that allows SimHarness to move agents around the simulated environment and interact with it by placing mitigations. The `FireSimulation` agents represent firefighters moving through an environment as a wildfire spreads, placing mitigations such as firelines to limit the spread of the fire within the area.

SimHarness utilizes Hydra [22] as a hierarchical configuration management tool to allow users to configure the training parameters of SimHarness. The configuration files provided by SimHarness mirror the structure of the Algorithm Configs used by RLlib for model training, such as `training`, `evaluation`, and `resources`. Users can also configure the parameters used for initializing the `Simulation` and the agents within the environment. For example, users can configure the `agent_speed`, which determines the number of actions an agent can take before the simulation is run, `interactions`, which are the mitigation techniques an agent can apply to the landscape, and `attributes`, which determine which attributes are passed as an input dimension to the RL model during training and inference.

Another configurable aspect of the SimHarness environment is the reward function. Users can create a custom reward function for training that emphasizes user-specific goals. This allows for tuning of the agent policy to better suit the user’s goals. For example, some users may want policies that prioritize ending the fire as quickly as possible, while others may focus more on limiting the fire spread to specific areas. An example workflow of SimHarness can be found in Appendix B.1.

4 Preliminary Experiments

SimFire and SimHarness provide a novel system for generating mitigation strategies for scenarios including real-world data. For this reason, comparisons between SimHarness and other available meth-

ods are not one-to-one, but we hope the open-sourcing of SimFire and our preliminary experiments can expand current benchmarks to include the testing of strategies in real-world scenarios.

The following experiment applies the SimFire and SimHarness tools to an area of land in Coalinga, CA near the location of the Mineral Fire of 2020. The fuel and terrain data is the true observed data from 2019 pulled from LANDFIRE [16] to simulate the fuels and terrain prior to fire ignition. The area of land simulated covers a 128 unit \times 128 unit square area with a left-corner GPS lat-long location of (36.09493, -120.52193), where each unit in the grid represents 30 square meters.

In the scenario, the fire starts at a random location within the simulated area and the simulated "firefighter" controlled by a trained DQN [23] policy always begins at location [0, 64], halfway down the left-hand side of the simulated environment.

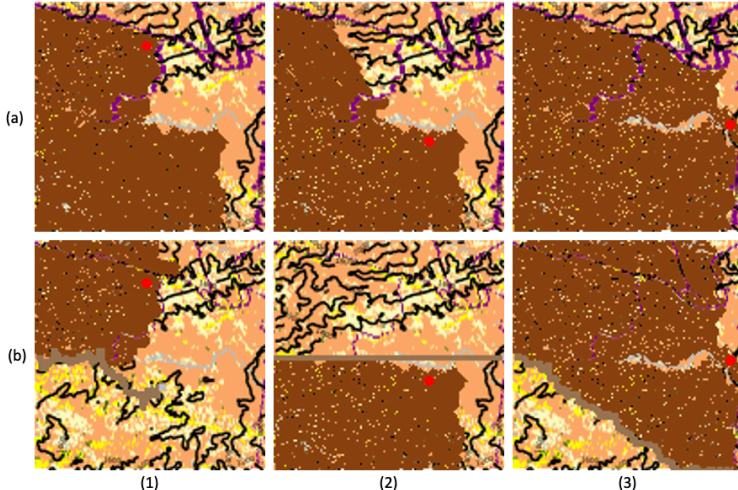


Figure 1: The difference in fire spread for 3 separate scenarios (1, 2, 3) when (a) no mitigations are placed and (b) mitigations are placed using SimHarness. Ignition locations are marked by a red circle.

As shown in Figure 1, the trained agent has generalized to generate successful mitigation strategies for random fire ignition scenarios. In all three cases, the agent’s mitigation strategy successfully saved a large section of land from being burned and limited the rate of spread of the fire. In scenario 2, the fire ignites close to latitude of the agent, resulting in the agent needing to move in a straight line to cut the fire off before it can spread to the top half of the environment. The training parameters and quantitative metrics from the training run can be found in Appendix B.

5 Discussion

The SimFire and SimHarness repositories together create an open-source Python-based system to accurately model wildland fire spread and generate appropriate mitigation strategy responses via RL. Researchers and wildfire managers can leverage SimFire and SimHarness to identify new land management strategies that limit the fire severity risk and lessen financial, ecological, and public health impacts caused by increasingly severe wildfires.

In the future, we aim to incorporate additional agent constraints, like agent distance to the fire and realistic movements [24], to the training process to produce more accurate strategies. We also aim to add agent types and capabilities to more accurately model the range of equipment and crews available to land managers and add data layers to the Simulation to more accurately model the landscape, including buildings and areas of cultural importance such that the environment more accurately models the real-world.

Acknowledgments and Disclosure of Funding

The authors acknowledge the help of Chris Kempis in the development of SimFire. This work was funded under MITRE’s 2022 and 2023 Independent Research and Development Program.

References

- [1] D.R. Reidmiller et al. “Fourth national climate assessment”. In: *Volume II: Impacts, Risks, and Adaptation in the United States, Report-in-Brief* (2019).
- [2] Omar Mouallem. *The Impossible Fight to Stop Canada’s Wildfires*. URL: <https://www.wired.com/story/canada-wildfires-future/>.
- [3] Katie Hoover and Laura A Hanson. *Wildfire Statistics*. 2023. URL: <https://sgp.fas.org/crs/misc/IF10244.pdf>.
- [4] National Interagency Fire Center. *Suppression Costs*. 2023. URL: <https://www.nifc.gov/fire-information/statistics/suppression-costs>.
- [5] Andrew Burchill. *Are wildfires bad?* May 2021. URL: <https://askabiologist.asu.edu/explore/wildfires>.
- [6] C. R. Rothermel. “A mathematical model for predicting fire spread in wildland fuels”. In: *U.S. Department of Agriculture, Intermountain Forest and Range Experiment Station* (1972), Res. Pap. INT-115.
- [7] C. Lautenberger. “Wildland fire modeling with an Eulerian level set method and automated calibration”. In: *Fire Safety Journal* (Nov. 2013), Volume 62, Part C, 289–298. URL: <https://doi.org/10.1016/j.firesaf.2013.08.014>.
- [8] R.R. Linn et al. *QUIC-fire: A fast-running simulation tool for prescribed fire planning*. 2020. URL: <https://www.fs.usda.gov/research/treesearch/59686>.
- [9] Ilkay Altinaş. *BurnPro3D*. URL: <http://wifire.ucsd.edu/burnpro3d>.
- [10] David Saah. *Pyrecast*. URL: <https://pyrecast.org>.
- [11] Travis Hammond. *Wildfire-Control-Python*. URL: <https://github.com/dashdeckers/Wildfire-Control-Python>.
- [12] Emanuel Becerra Soto. *gym-cellular-automata*. <https://github.com/elbecerrasoto/gym-cellular-automata>. 2021.
- [13] Altamimi et al. “Large-Scale Wildfire Mitigation Through Deep Reinforcement Learning”. In: *Frontiers in Forests and Global Change* 5 (2022).
- [14] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. Second. The MIT Press, 2018. URL: <http://incompleteideas.net/book/the-book-2nd.html>.
- [15] Richard Bellman. “A Markovian Decision Process”. In: *Journal of Mathematics and Mechanics* 6.5 (1957), pp. 679–684. ISSN: 00959057, 19435274. URL: <http://www.jstor.org/stable/24900506> (visited on 09/22/2023).
- [16] LANDFIRE. *13 Anderson Fire Behavior Fuel Models, Elevation, LANDFIRE 2.0.0*, 2021. DOI: <http://www.landfire/viewer>.
- [17] Patricia L. Andrews. “The Rothermel surface fire spread model and associated developments: A comprehensive explanation”. In: *Gen. Tech. Rep. RMRS-GTR-371. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station* (2018), p. 121.
- [18] Jos Stam. “Real-Time Fluid Dynamics for Games”. In: *Proceedings of the game developer conference* (2003), Vol. 18, p. 25.
- [19] Robert E. Burgan Joe H. Scott. *Standard Fire Behavior Fuel Models: A Comprehensive Set for Use with Rothermel’s Surface Fire Spread Model*. <https://www.nwcg.gov/sites/default/files/training/docs/s-290-usfs-standard-fire-behavior-fuel-models.pdf>.
- [20] Pete Shinnars et al. *Pygame*. URL: <https://www.pygame.org>.
- [21] Eric Liang et al. *RLLib: Abstractions for Distributed Reinforcement Learning*. 2018. arXiv: 1712.09381 [cs.AI]. URL: <https://docs.ray.io/en/latest/rllib/index.html>.
- [22] Omry Yadan. *Hydra - A framework for elegantly configuring complex applications*. Github. 2019. URL: <https://github.com/facebookresearch/hydra>.
- [23] Volodymyr Mnih et al. “Playing Atari with Deep Reinforcement Learning”. In: (2013). arXiv: 1312.5602 [cs.LG].
- [24] National Wildland Fire Coordinating Group. *Fire Line Production Tables*. https://www.fs.usda.gov/t-d/nwcg/files/NWCG_production_tables_2021.pdf.
- [25] M. Dotter et al. “BurnMD: A Fire Projection and Mitigation Modeling Dataset”. In: *International Conference of Learning Representations* (2023).

[26] Mark Towers et al. *Gymnasium*. Mar. 2023. DOI: 10.5281/zenodo.8127026. URL: <https://zenodo.org/record/8127025> (visited on 07/08/2023).

A SimFire

A.1 Available Data Layers and Actions

SimFire provides Fuel, Terrain, and Wind data layers that are ingested into SimHarness to create the following state spaces at each pixel:

- **Fuel:**
 - w_0*: Oven-dry Fuel Load (lb/ft^2)
 - sigma*: Surface-Area-to-Volume Ratio (ft^2/ft^3)
 - delta*: Fuel Bed Depth (ft)
 - M_x*: Dead Fuel Moisture of Extinction
- **Terrain:**
 - elevation*: Elevation (ft) in relation to sea level.
- **Wind:**
 - wind_speed*: Wind Speed (mph)
 - wind_direction*: Direction of the Wind (degree)

These layers are provided to SimHarness as dimensions of the input observation to the RL model. Users can specify which data layers are passed to the model and what order they are within the observation. SimFire also provides the minimum and maximum bounds for each data layer which helps SimHarness normalize observations dimensions, if desired by the user.

The *Fuel* data layer is set based on the type of fuel present in the given scenario, determined by the fire start location and the size of the simulation specified. SimFire supports the usage of the 13 Anderson Behavior Fuel Models [19].

SimFire also supports 3 types of firefighting mitigations:

- **Fireline:** Flammable material has been removed by scraping or digging down to mineral soil.
- **Wetline:** Water/suppressant sprayed in an area.
- **Scratchline:** Preliminary, quick-action fireline where flammable material is removed but not entirely and not completely down to mineral soil.

Each mitigation has different properties which effect the speed and movement of the fire when the fire is in contact with the mitigation.

A.2 Additional Fire Scenarios

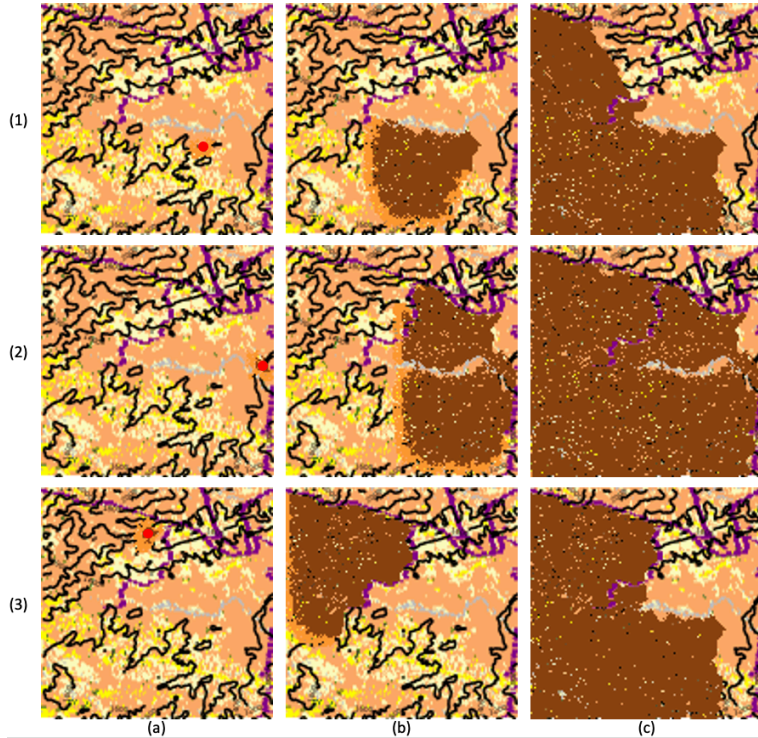


Figure 2: The progression of fire spread in 3 scenarios (1,2,3) using SimFire. Frames are captured at three separate points of the episode: a.) near ignition, b.) mid-episode, and c.) after the fire spread has stopped. Each scenario represents a different fire ignition location, marked by a red circle in frame (a).

A.3 Validation and Verification

Future work will provide a detailed comparison of the validation and verification process for the fire spread simulator, SimFire, and the underlying fire spread model, Rothermel to other fire spread models including ElmFire. This study is evaluated using the historical database, BurnMD [25], which includes 308 medium sized fires with near real-time mitigations and daily wildfire perimeters. For more details about the database, its contents, and the data sources, please see the referenced publication or visit the BurnMD website, <https://fireline.mitre.org/burnmd>.

B SimHarness

B.1 Workflow

SimHarness allows users to train RL agents within any simulated disaster environment, assuming the disaster environment implements the methods required by the Simulation API. In the case of SimFire, SimHarness can generate mitigation strategies for firefighters to limit the damage caused by wildfires. The general workflow for SimHarness is shown in Figure 3.

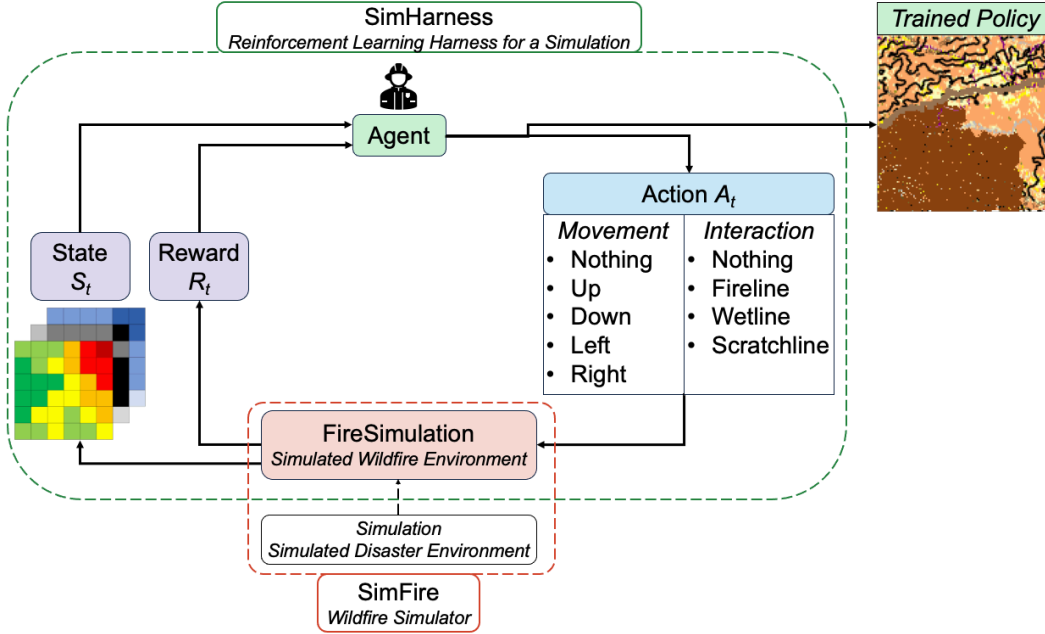


Figure 3: Conceptual workflow for training an RL model using SimHarness within the SimFire environment.

The SimHarness training loop functions similarly to a traditional RL training loop, except it expects the passed-in environment to be a child class of `Simulation` as opposed to a gym [26] environment. `Simulation` is currently a class within the SimFire package, but is expected to be moved to a separate, non-disaster-specific package in the future. The simulated environment outputs training signals such as observations and rewards to the SimHarness agent(s) which use the observations to predict optimal actions. The actions produced by the model provide both *movement* and *interaction* information. *Movements* are how the agent is traversing across the environment, such as [nothing, up, down, left, right]. *Interactions* are how the agent is changing the environment itself. In the case of SimFire, this can be [nothing, fireline, wetline, scratchline]. These actions are relayed back to the simulated environment, which then affects the overall disaster scenario simulated by the environment.

B.2 Training Parameters and Reward Function

The table below provides a detailed overview of the training parameters leveraged by the environment, agent, learning models, and harness for the experimental results that are presented in Section 3.2. The operational metrics of performance for this experiment are displayed in Section B.3.

A notable aspect of this experiment was the inclusion of a parallel benchmark simulation during each episode. The benchmark simulation simulates the spread of the current wildfire scenario without any agent placed mitigations. The input observation to the model at each step includes 3 arrays:

- Fire Map: Array representing the current fire spread at a given timestep including the placed mitigations and agent locations.
- Benchmark Fire Map: Array representing how the fire spread would look at the current timestep if no mitigations were placed.
- Final Benchmark Fire Map: Array representing how the fire spread would look at the final timestep if no mitigations were placed.

The inclusion of the above information within the training environment is used for both the input observation to the agent as well as the reward function generation, as the reward function compares the total area burned within the mitigated episode to the unmitigated counterpart.

Class	Experiment Parameter	Value
Env	Type	Wildfire
Env	Simulator	SimFire
Env	Type	Operational
Env	Left Corner (Lat,Lon)	(36.09493, -120.52193)
Env	Geographic Area	491520 Square Meters
Env	Grid Cells	128x128
Env	Fire Start Location	Random
Env	Simulation Data	[Terrain, Fuel, Wind]
Agent	Type	Single Agent
Agent	Start Location	(64,0)
Agent	Observation Space	[current fire map, benchmark fire map, final benchmark fire map]
Agent	Action Space Movements	[up, down, left, right]
Agent	Action Space Interactions	[Fireline]
Agent	Speed	4
Training	Algorithm	DQN
Training	Algorithm Config	[Dueling + Distributional + Noisy]
Training	Episodes	14000
Training	Timesteps	8273288
Training	Exploration	Epsilon Greedy [1.0, 0.01]
Training	Replay Buffer	Episodic Prioritized Replay
Training	[lr, gamma]	[0.0045, 0.99]
Harness	Type	Ray[Rllib]
Harness	CPU	16
Harness	GPU	2

The corresponding reward function for this experiment was based on the incremental proportion of Area saved, or area that was not burned, burning, or mitigated, at each timestep (t), when comparing the mitigated simulation (Sim) to the unmitigated benchmark simulation (Bench).

$$Damaged_{Sim_t} = Sim[Burned_t] + Sim[Burning_t] + Sim[Mitigated_t] \quad (1)$$

$$Damaged_{Bench_t} = Bench[Burned_t] + Bench[Burning_t] \quad (2)$$

$$NewDamaged_{Sim_t} = Damaged_{Sim_t} - Damaged_{Sim_{t-1}} \quad (3)$$

$$NewDamaged_{Bench_t} = Damaged_{Bench_t} - Damaged_{Bench_{t-1}} \quad (4)$$

$$TotalEndangered = Bench_{final_t}[Burned] \quad (5)$$

$$Reward_t = \frac{NewDamaged_{Bench_t} - NewDamaged_{Sim_t}}{TotalEndangered} \quad (6)$$

Equation 1 and 2 represent the total number of pixels that are "lost" in the simulation at a given timestep, including burned pixels, currently burning pixels, and pixels that have been mitigated by the agent, for the mitigated sim and unmitigated benchmark sim, respectfully. Equation 3 and 4 represent the new number of pixels that are "lost" in the simulation at the given timestep. Equation 5 is the total number of pixels that will burn if no mitigations are placed in the environment. The final reward in Equation 6 is the difference in new pixels burned between the unmitigated and mitigated simulations as a percentage of the total pixels burned in the benchmark simulation. A positive value represents more pixels being saved in the mitigated scenario than in the unmitigated scenario, with a higher value corresponding to more area saved. A value of 0 means the unmitigated and mitigated scenarios saved the same amount of pixels, and a negative value means the mitigated scenario saved less land than the unmitigated scenario. This ensures that the total sum of the rewards within an

episode directly corresponds to the total proportion of Area 'saved' for the entire episode (Equation 8).

A small positive reward (Equation 7 is applied to the final reward (Equation 6) when the agent places a mitigation on an unburned square with no prior mitigations. This addition encourages the agent to place more firelines overall, which helps with training as the agent will get better training examples of how the fire spread reacts to placed mitigations.

$$Reward_t = Reward_t + \frac{0.25}{Area_{sim}} \quad (7)$$

B.3 Training Metrics

The graphs below report the experiment results and the operational metrics of performance for the experiment detailed in Section B.2. The graphs illustrate the Episode Reward Mean, Mean Area Saved, Mean Timesteps Saved, and Mean Burn Rate Reduction over the training of 14,000 episodes.

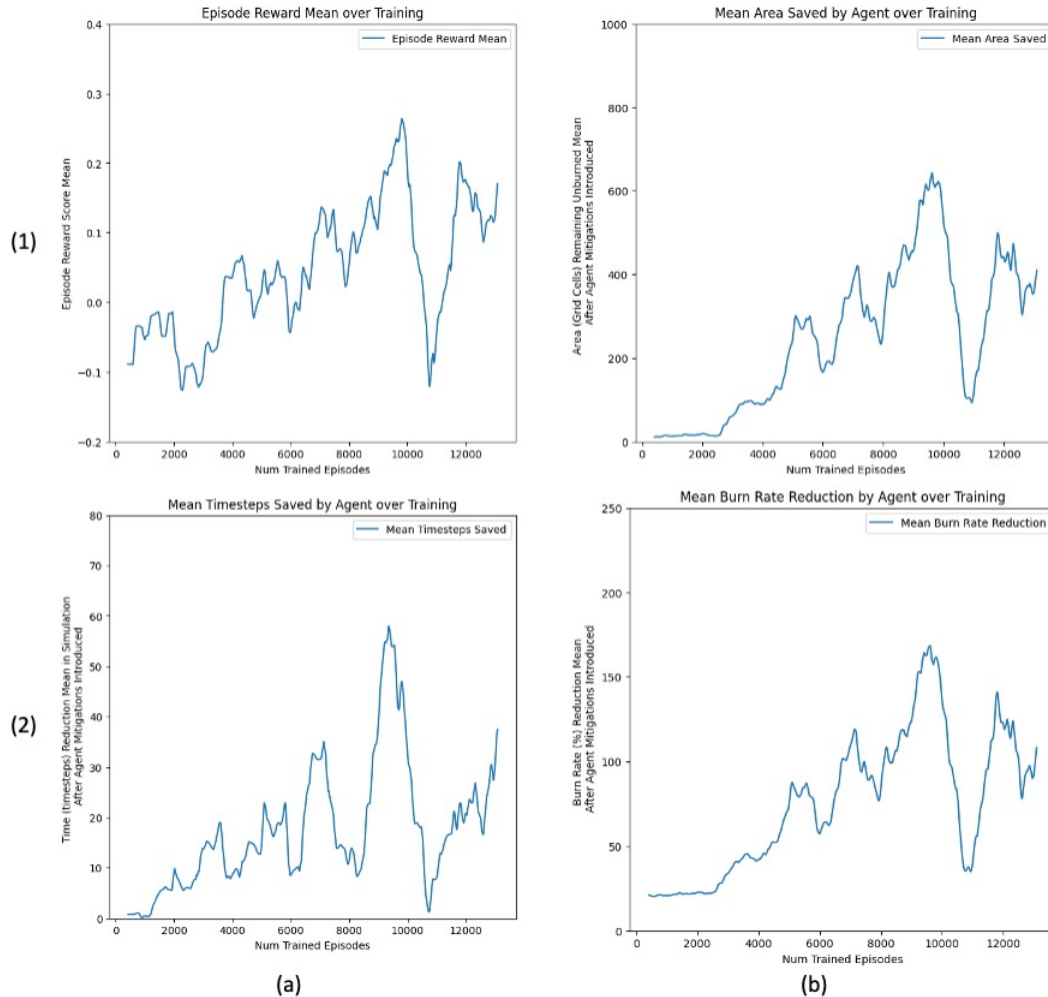


Figure 4: 1a: Mean Episode Reward: mean episode received over an episode when mitigations are placed. 1b: Mean Area Saved: mean number of pixels that are left unburned or unmitigated at the end of an episode after mitigations have been placed. 2a: Mean Timesteps Saved: mean number of timesteps that an episode is reduced by when mitigations are introduced compared to the unmitigated scenario. 2b: Mean Burn Rate Reduction: mean reduction in overall episode burn rate (%) when mitigations are placed compared to the unmitigated scenario.

The metrics per episode (eps) of Mean Area Saved (Equation 8), Mean Timesteps Saved (Equation 9), and Mean Burn Rate Reduction (Equations 10, 11) are based on operational metrics that are utilized to estimate wildfire severity and mitigation efficacy. These metrics also serve as heuristic measurements to monitor to validate that the agent is learning effective policies.

$$AreaSaved_{eps} = (Sim[Burned_{eps}] + Sim[Mitigated_{eps}]) - Bench[Burned_{eps}] \quad (8)$$

$$TimestepsSaved_{eps} = Sim[timesteps_{eps}] - Bench[timesteps_{eps}] \quad (9)$$

$$BurnRate_{eps} = \frac{(Burned_{eps} + Mitigated_{eps})}{timesteps_{eps}} * 100 \quad (10)$$

$$BurnRateReduction_{eps} = Sim[BurnRate_{eps}] - Bench[BurnRate_{eps}] \quad (11)$$