Being the Fire A CNN-Based Reinforcement Learning Method to Learn How Fires Behave Beyond the Limits of Physics-Based Empirical Models

William L. Ross
Department of Earth, Energy, and Environmental Sciences
Stanford University
Stanford, CA 94305
wlross@stanford.edu

Abstract

Wildland fires pose an increasing threat in light of anthropogenic climate change. Fire-spread models play an underpinning role in many areas of research across this domain, from emergency evacuation to insurance analysis. We study paths towards advancing such models through deep reinforcement learning. Aggregating 21 fire perimeters from the Western United States in 2017, we construct 11-layer raster images representing the state of the fire area. A convolution neural network based agent is trained offline on one million sub-images to create a generalizable baseline for predicting the best action - burn or not burn - given the then-current state on a particular fire edge. A series of online, TD(0) Monte Carlo Q-Learning based improvements are made with final evaluation conducted on a subset of holdout fire perimeters. We examine the performance of the learned agent/model against the FARSITE fire-spread model. We also make available a novel data set and propose more informative evaluation metrics for future progress.

1 Introduction

The performance of fire-spread models, which aim to predict the spatial spreading process of an active fire across a given area, is important to protecting our communities from wildfire. Most contemporary fire spread models can be traced back to a single 1972 paper – *A Mathematical Model for Predicting Fire Spread in Wildland Fuels* – authored by Richard Rothermel [1]. While Wells (2008) points out that the Rothermel Model's empirical, physically-informed approach is "still running like a champ", many experts recognize that the model is now being asked to do things it was never meant to do [2].

The last decade has seen marked progress in the fields of deep learning and reinforcement learning and has spurred a new era for machine learning and artificial intelligence [3,4]. In the field of deep learning, convolutional neural networks exhibit unique predictive ability in image recognition tasks, including those that use remote sensing [5,6]. Deep reinforcement learning, meanwhile, has demonstrated the ability to solve complex optimization problems dynamically and over time in the presence of uncertainty [7].

Combining these techniques, there is initial evidence to suggest that deep reinforcement learning can be used to learn wildfire dynamic models from historic observations and remote sensing data. We extend the work of Subramanian and Crowley – *Using Spatial RL to Build Forest Wildfire Dynamics Models From Satellite Images* – in hopes of unifying the latest remote sensing data, machine learning algorithms, and physical techniques to advance fire spread modeling [8].

Tackling Climate Change with Machine Learning: workshop at NeurIPS 2021.

2 Review of Literature

2.1 Fire-Spread Modeling and Prediction

The vast majority of today's fire-spread models represent small changes to individual characteristics within the framework provided by Rothermel. Models such as FARSITE and BehavePlus are widely adopted in commercial and government work today but typically focus on improving select parameters, with incremental progress in each new generation [9,10]. But the reality in the words of Rothermel pupil Brett Butler, is that "(these models describe) very well a fire burning in a field of wheat. As you get further away from that uniformity, the less accurate (they) become [11]."

Among the most meaningful areas of such progress has been the improvement of topographic wind speed modeling. Because most wildfires do not burn in a field of wheat, understanding how wind changes speed in complex topography is important to assessing speed and direction of fire spread. Wagenbrenner et al. (2016) make use of physical conservation of mass and momentum to downscale surface wind predictions or measurements in complex terrains [12]. While such solvers are intended to improve the Rothermel framework, they yield equally useful inputs for machine learned approaches.

2.2 Machine Learning and Remote Sensing in Fire-Spread Models

The science of remote sensing has advanced as the resolution, coverage, and frequency of such data improves [13]. Government funded projects such as Landsat 8 (2013) and Sentinel 1-A/B (2014-16) provide high resolution (20-30m) data at a consistent frequency [14,15]. Private companies such as Planet provide further coverage through projects like RapidEye (5m) and Planetscope (3m), both of which provide data from much of the planet on a daily frequency or better.

Such data has opened the door for the use of machine learning in various applications in widlfire. For example, Zhang et al. (2011) provide a hybrid model that makes use of satellite imagery and is now used in the Canadian Forest Fire Weather Index (FWI) [16]. The use of sequential models in the form of markov decision processes (MDP) offers another path forward particularly relevant to fire-spread models. In Subramanian and Crowley (2019), a number of methods including Q-Learning, monte carlo tree search, and deep reinforcement learning are identified as promising opportunities.

3 Problem Formulation/Methods

We evaluate the spread of wildfire in a grid-based 30m resolution environment on the USGS Contiguous Albers Equal Area Conic coordinate reference system as an MDP S,A,P,R. Our continuous state space, S, represents the then-current state of a given cell on the fire edge as represented by an 11x3x3 raster of that cell and all adjacent cells. 10 layers represent constants over the observation period and 1 layer represents the dynamic condition of where the fire has or has not spread at a given time step T. Our binary action space, A, is a simple burn, not burn choice for each unburned grid cell on the fire edge at each time step. Our transition probability P is represented as a convolutional neural network (CNN) and estimates the likelihood that a burn or not burn action will maximize our reward R - the negative binary cross-entropy loss of the CNN at each time step.

All code and data used for model/agent training and analysis are publicly available for reuse: `github.com/wlross/Being_The_Fire_Final`.

3.1 Data Acquisition and Processing

Critical to this approach is the state space as represented by historical data from each fire perimeter. A total of 42 fire perimeters representing a T=0 and T=Final perimeter for each of 21 fires (see Appendix A) from the 2017 wildfire season in the Western United States were collected via GeoMAC [17]. Fires were manually curated to ensure consistent measurement methodologies and a geographical, topographical, and fuel load distribution consistent with the full set of 7418 GeoMAC perimeters from the 2017 fire season.

For each fire boundary, a bounding-box representing the edges at T=Final was created and gridded into 30m cells. For each grid point, 10 data characteristics were gathered from several sources including: *Planet 5m Resolution RapidEye Program* - Red (1), Green (2), Blue (3), Red Edge (4), and

Near Infrared (5) Imagery, *US Geological Survey* - 30m Resolution Topography, *National Weather Service* - Average Wind Speed and Direction (7,8) and Maximum Wind Speed and Direction (9,10) [18,19,20]. All values were imputed to the final 30m resolution grid using mean or nearest neighbor approaches as appropriate.

3.2 Training and Evaluation

In order to train our agent, two distinct phases of model training were used. The initial offline training approach was introduced to increase the generalizability of the online model. The offline environment was also used for experimentation and hyper-parameter tuning as detailed in Appendix B. The model architecture used in both offline and online training is visualized as follows:

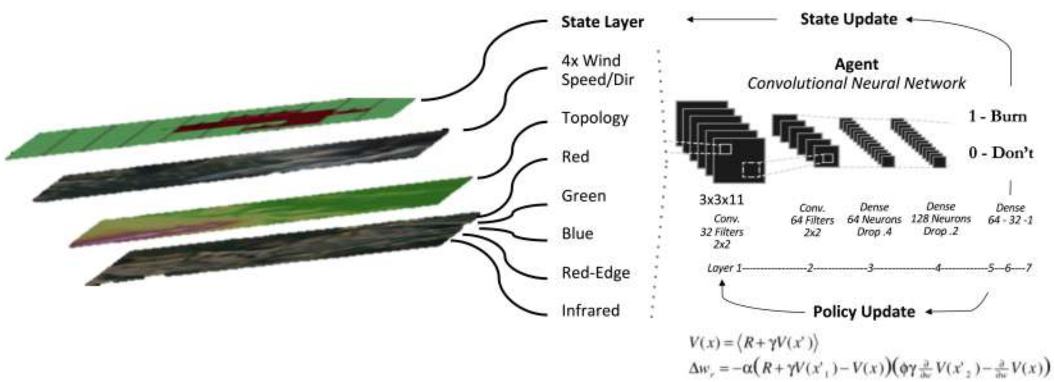

Figure 1: Model architecture for convolutional neural network and TD(0) Monte Carlo Q-Learning

For online training, weights from offline training were transferred and additional training was conducted using a TD(0) Monte Carlo Q-Learning algorithm. The reinforcement learning aspect of this approach was consistent with the work of Subramanian and Crowley with the primary difference being the CNN representation of the agent and the use of more, higher resolution data layers.

Results for final evaluation were generated using the trained agent on the four holdout fires. In parallel, the FlamMap 6 package was used to generate benchmark data via the FARSITE model using default parameters and landscape files available via the LANDFIRE program [21].

4 Analysis of Results

4.1 Quantitative Model Performance

Quantitative model performance was measured using the Weighted Average F-1 Score as the primary metric, recognizing that accuracy measures may overstate performance of "under-burn" or "over-burn" models depending on the denominator used. For this research, all grid squares within the bounding box that were not already ignited at T=0 were used for analysis in order to fairly weight both "under-burn" and "over-burn" behaviors. Results were not compared to Subramanian and Crowley as the accuracy metrics presented did not provide a reasonable means for direct comparison and reproducing this work was challenging.

Table 1: Reinforcement Learning (RL) and FARSITE (FS) Model Performance

	Reinforcement Learning			**FARSITE Benchmark**		
Fire Name	**Precision**	**Recall**	**F-1**	**Precision**	**Recall**	**F-1**
Buck	.82	.78	**.74**	.64	.45	**.44**
Highline	.77	.69	**.59**	.62	.43	**.39**
Pinal	.84	.84	**.81**	.84	.20	**.08**
Sulfur	.78	.72	**.64**	.79	.73	**.74**

Weighted average 0s and 1s in t=0 unburned sample area

The RL-Model outperformed the FARSITE model on 3/4 test fires, though both had low F-1 scores. In general, this was due to "under-burn" by the RL Model (low class 1 recall) and "over-burn" by the FARSITE model (low class 0 recall). Both methods performed similarly on class 0 precision but the RL model significantly outperformed the FARSITE model on class 1 precision, providing some evidence of a better "fit" by the RL model.

4.2 Qualitative Model Performance

Given the relatively low F-1 scores exhibited by both the RL and FARSITE model, a smoothing function was applied to the fire perimeter so that the final fire boundaries could be inspected qualitatively. This is consistent with expected use in the field - see Appendix C.

In both cases, models appeared to be performing in ways consistent with our understanding of physical fire spread - burn was driven by wind direction, slope, and vegetation and obstructed by roads, rivers, and lakes. A visual inspection of the fire spread patterns provides some indication of superior performance by the RL model. For instance, the fire road present in the Highline fire and the river present in the Pinal both seem to have influenced a closer fit to the ground truth data for the RL model when compared to FARSITE, which crossed these boundaries easily - see Figure 2.

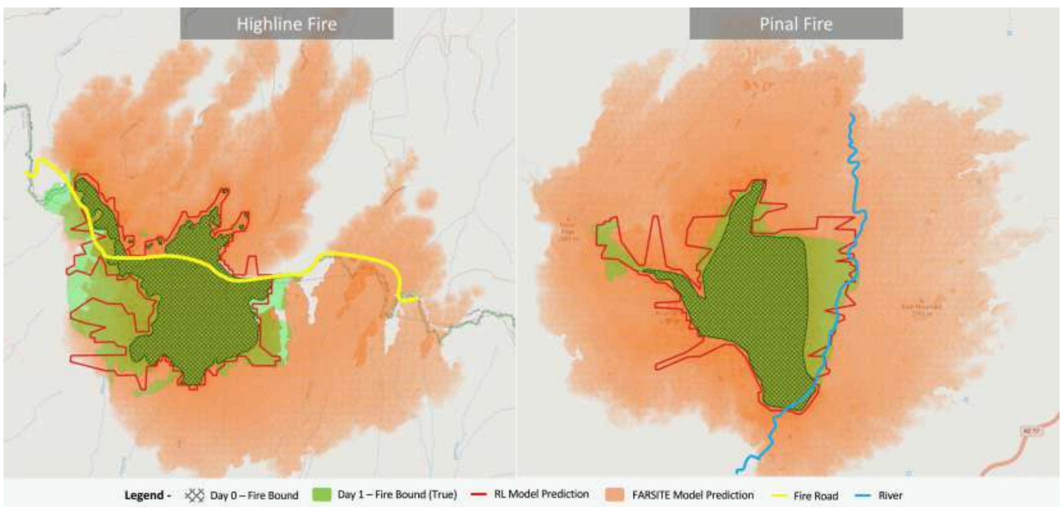

Figure 2: RL + FARSITE Models of Highline and Pinal Fires from Test Set

5 Discussion

The reality of fire spread models is that they are attempting to model highly stochastic physical processes. But given such models serve as a critical building block for climate adaptation to wildland fires, progress is important. When compared to existing methods like FARSITE, the results of this work support continued exploration of deep reinforcement learning approaches in this domain.

The CNN-based RL methods proposed in this paper have the advantage of tailwinds in both machine learning research and remote sensing data availability. One challenge to progress, however, is the availability of remote sensing data at high resolution and high frequency. Notably, these dependencies are often also present (and sometimes less obvious) when working with physical models.

Another source of challenge in this direction is the lack of standardization. Metrics like accuracy that have been previously reported lack sufficient context for determining model performance. The use of metrics like the weighted average F-1 score, which factors in both "under-burn" and "over-burn", alongside qualitative assessments provide an opportunity to establish new benchmarks.

It is clear that reinforcement learning methods for fire-spread modeling are not without their challenges. Still, this work demonstrates the potential for learned methods to, over time, add value to progress in fire-spread modeling.

References

[1] Rothermel, Richard C. A mathematical model for predicting fire spread in wildland fuels. Vol. 115. Intermountain Forest Range Experiment Station, Forest Service, US Department of Agriculture, 1972.

[2] Wells, Gail. "The Rothermel Fire-Spread Model: still running like a champ." (2008).

[3] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521.7553 (2015): 436-444.

[4] Watkins, Christopher JCH, and Peter Dayan. "Q-learning." Machine learning 8.3-4 (1992): 279-292.

[5] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems 25 (2012): 1097-1105.

[6] Zhang, Wei, Ping Tang, and Lijun Zhao. "Remote sensing image scene classification using CNN-CapsNet." Remote Sensing 11.5 (2019): 494.

[7] Mnih, Volodymyr, et al. "Playing atari with deep reinforcement learning." arXiv preprint arXiv:1312.5602 (2013).

[8] Subramanian, Sriram Ganapathi, and Mark Crowley. "Learning forest wildfire dynamics from satellite images using reinforcement learning." Conference on reinforcement learning and decision making. Ann Arbor MI, 2017.

[9] Finney, Mark A. FARSITE, Fire Area Simulator–model development and evaluation. No. 4. US Department of Agriculture, Forest Service, Rocky Mountain Research Station, 1998.

[10] Andrews, Patricia L. "BehavePlus fire modeling system: past, present, and future." In: Proceedings of 7th Symposium on Fire and Forest Meteorology; 23-25 October 2007, Bar Harbor, Maine. Boston, MA: American Meteorological Society. 13 p.. 2007.

[11] Gabbert, Bill. "Throwback Thursday: The origin of the model for predicting the spread of wildland fires." In: Wildfire Today, June 2018.

[12] Wagenbrenner, Natalie S., et al. "Downscaling surface wind predictions from numerical weather prediction models in complex terrain with WindNinja." Atmospheric Chemistry and Physics 16.8 (2016): 5229-5241.

[13] Sandau, Rainer, and Klaus Brieß. "Potential for advancements in remote sensing using small satellites." The International Archives of the (2008).

[14] Williams, Darrel L., Samuel Goward, and Terry Arvidson. "Landsat." Photogrammetric Engineering Remote Sensing 72.10 (2006): 1171-1178.

[15] Torres, Ramon, et al. "GMES Sentinel-1 mission." Remote Sensing of Environment 120 (2012): 9-24.

[16] Zhang, Jia-Hua, et al. "Detection, emission estimation and risk prediction of forest fires in China using satellite sensors and simulation models in the past three decades—An overview." International journal of environmental research and public health 8.8 (2011): 3156-3178.

[17] Walters, Sandra P., Norma J. Schneider, and John D. Guthrie. "Geospatial Multi-Agency Coordination (GeoMAC) Wildland Fire Perimeters, 2008." US Geological Survey Data Series 612.6 (2011).

[18] Marta, Santa. "Planet Imagery Product Specifications." Planet Labs: San Francisco, CA, USA (2018): 91.

[19] Gesch, Dean, et al. "The national elevation dataset." Photogrammetric engineering and remote sensing 68.1 (2002): 5-32.

[20] Glahn, Harry R., and David P. Ruth. "The new digital forecast database of the National Weather Service." Bulletin of the American Meteorological Society 84.2 (2003): 195-202.

[21] Rollins, Matthew G. "LANDFIRE: a nationally consistent vegetation, wildland fire, and fuel assessment." International Journal of Wildland Fire 18.3 (2009): 235-249.

Appendix A - Sample of 21 Fires from 2017 Western US Fire Season

Train Fires

Powerline - ID - Jul	Cove - CA - Jul	Oak - CA - Aug
Swiss Helms - AZ - Jun	Steele - CA - Jul	Indian Ridge - ID - Sep
Creek - CA - Dec	Preacher - NV - Jul	Cub Creek - MT - Sep
Saddle - AZ - Jun	Little Hogback - MT - Aug	Mammoth Cave - ID - Aug
Gutzler - CO - Jul	North Pelican - OR - Aug	Helena - CA - Oct
Sheep - AZ - Jul	Nena Springs - OR - Au	

Test Fires

Pinal - AZ - May
Highline - AZ - June
Sulfur - CA - Oct
Buck - CA - Sept

Appendix B - Details around model hyper parameters for training

The neural network's input was a 3x3 cell array with 11 bands. The first convolutional layer creates 32 filters of the 3x3x11 with a kernel size of 2x2 and the second convolutional layer creates an additional 64 2x2 filters. The third layer is fully connected with 64 neurons followed by a dropout layer of .4. The fourth layer consists of 128 fully connected neurons followed by a dropout layer of .2. The fifth, sixth, and seventh/output layers are fully connected with 64, 32, and 1 neurons respectively. The ReLU activation function is used for all layers with the exception of the binary output, which uses a sigmoid function. Binary cross-entropy loss is used with the adam optimizer in both the offline and online setting with epochs, batch size, learning rates (α), and class weights specified below.

For offline training, a random sample of one million 3x3x11 images was assembled across all fires. The eleventh band of data representing the then-current state of the fire was substitute with random noise. The model was trained over 300 epochs with a batch size of 40, and a learning rate of α=1e-5. Class weights of 1 (no burn) and 4 (burn) to account for the uneven distribution of the randomly generated dataset and to maximize recall of the burned area. This approach was thought to be advantaged when moved to the online environment.

The agent/model made burn or no burn decisions for each cell and the fire edge over a number of iterations equal to 1.7 times the maximum wind speed. This fixed parameter ν was determined via an independent linear regression of the number of cells burned in a fixed period as a function of the maximum wind speed of the fire, regardless of directional change. For each online session, predictions were initiated as random (ϵ=1) and allowed to become increasingly ϵ-greedy with an exponential decay function where λ=.75 for each iterative model/agent update.

Model/agent updates were performed online after every 10,000 predictions/decisions. The model was trained at each iteration over 80 epochs with a batch size of 400 and a learning rate of α=1e-3. Class weights of 1 (no burn) and 2.3 (burn) were used as these values were inversely proportional to their respective frequencies in a random sample of the online training data.

Appendix C - Example of gridded vs smooth RL output

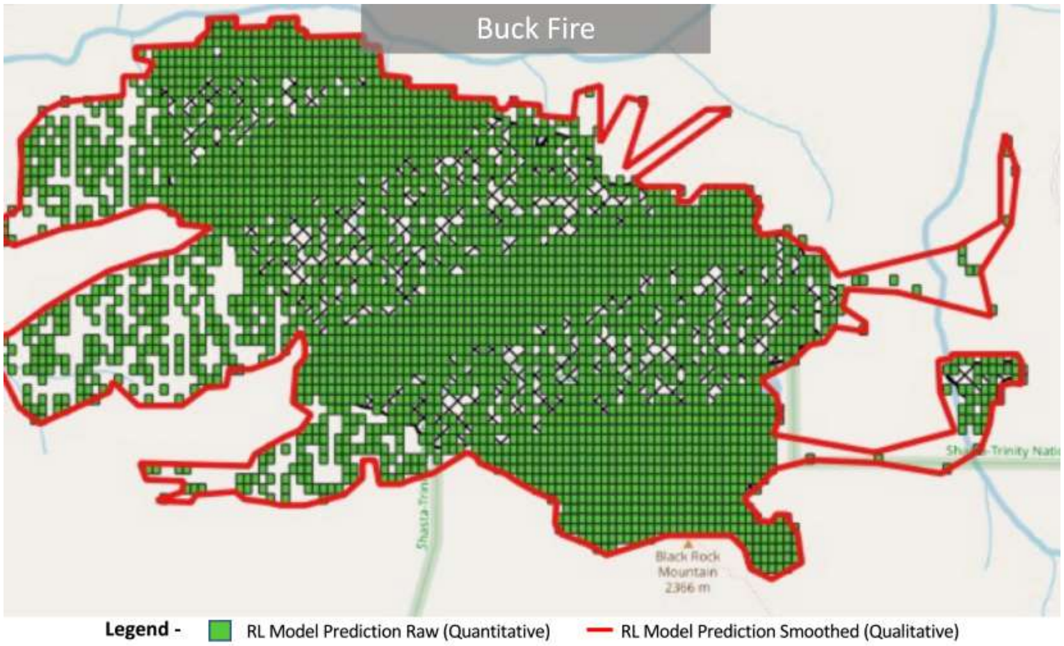

Figure 3: Raw vs Smooothed RL Model Prediction for Buck Fire

Appendix D - Example of RL vs FARSITE on RapidEye imagery

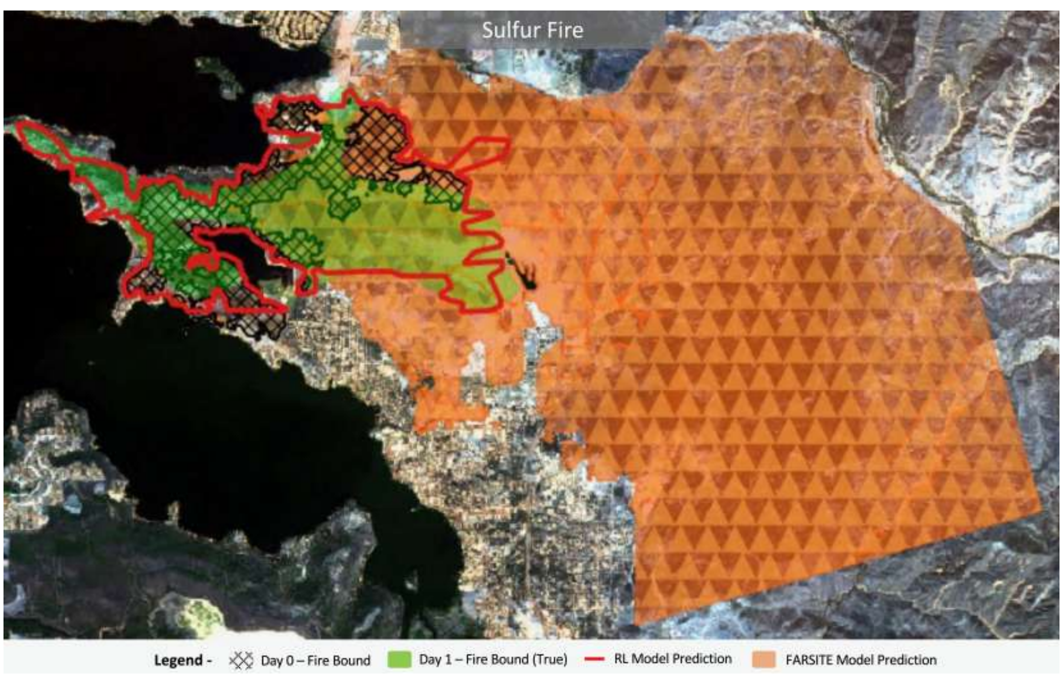

Figure 4: Imagery of site of Sulfur Fire via RapidEye program with RL and FARSITE predictions