Colorado Fire Acreage Predictor Final Report

Alec Bell
Computer Science - Master's Student
University of Colorado
Boulder, Colorado
alexander.bell-1@colorado.edu

Sijae Schiefer
Computer Science - Master's Student
University of Colorado
sijae.schiefer@colorado.edu

Abstract

Fires are a major hazard to wilderness areas throughout the western half of the United States. The Colorado wildfires of 2020 were the largest to date. Being able to effectively estimate total burn area will enable decision makers to more effectively prepare in advance. Using live fuel moisture content scoring from Stanford University and historic fire data from the Monitoring Trends in Burn Severity group, we formed a grid of feature vectors to train multiple machine learning models. Results are discussed at the end of this report.

1 Introduction

Throughout recent years, forest fires have continued to grow in size, frequency, and severity in the western half of the United States as well as in other geographic areas with dry climates like Australia. Several factors contribute to this. Variation in fuel conditions, an increase in human activity, and rising temperatures are three examples [1]. In response to the growing threat of wildfires, researchers are developing tools to analyze different aspects of the phenomenon. The University of Waterloo and Colorado College teamed up to develop a product called FireCast [2]. Historically, physics-based models and mathematical growth predictors have been used to study wildfire growth. These models are computationally expensive. FireCast leverages data from geographical information systems and utilizes machine learning techniques to predict which areas surrounding a wildfire are likely to experience wildfire spread. This development motivated our predictor strategy.

The goal of our project was to predict the total number of acres burned for a geographical area should the location experience a wildfire event. Our work used historical data to accomplish this goal. We were unable to find historical, open-source, satellite imagery data of Colorado within the time-frame of this project. Subsequently, we decided upon a different approach to formulating the input data. In our search for satellite imagery, we found a dataset of historical fire records from the Monitoring Trends in Burn Severity research group and a dataset from Stanford University containing a live fuel moisture content measurement for various locations throughout the United States. To represent the geographic information, we used latitude, longitude, and elevation to capture any spatial information that theoretically is relevant when combined with moisture content measurements or total burn acreage.

We chose to develop a convolutional neural network (CNN) along with three other models: a linear regression model, a random forest regression model, and a support vector regression model. The CNN was chosen because our initial understanding of the live fuel moisture content data was that

Preprint. Under review.

there existed spatially relevant information, especially in combination with geographic location and elevation. We hypothesized that this model would do best. Our experiments and results are discussed below.

2 Data Overview and Discussion

2.1 Feature Motivation and Sources

Fuel Moisture Content is a measure of the amount of water in fuel (i.e. vegetation) that is available to a fire. It is frequently used for determining wildfire risk, because it conveys a measurable idea of the dryness of an area. We collected this information at a number of different geographical locations in Colorado from a database released by Stanford researchers [3]. Because their data often spanned multiple dates for the same locations, they referred to it as Live Fuel Moisture Content (LFMC). They managed to generate this data using deep learning on SAR data, giving us insight into 45 different fuel types in Colorado and their respective LFMC values at many locations. These fuel types included (but not limited to) tree species such as the Common Juniper, Lodgepole Pine, and Douglas Fir. We suspected this data would enable our model to get a general idea of dry areas that are at risk of a fire.

Elevation is also known to have a significant influence on wildfire frequency and spread. Research from University of Colorado shows that high-elevation subalpine forests tend to have less frequent fires of greater spread, whereas low-elevation forests tend to have more frequent fires of lesser spread [4]. We collected elevation data using Google's Elevation API, and we suspected that this data would help our model to identify areas where a wildfire could realistically spread to as well as identify the areas of higher risk for spread.

2.2 Label Source

The features listed above (LFMC and elevation data) were used by our model to predict wildfire burn acreage. We retrieved this data from Monitoring Trends in Burn Severity (MTBS), a US Forest Service sponsored research project that aggregates information on almost every wildfire in the western US since 1984 [5].

2.3 Preprocessing Methods

As our data was collected from a variety of sources, we had to employ a number of varying techniques to organize it into a format that we could use to train our models. To begin, we simply parsed the MTBS wildfire data dump for Colorado, decompressing a package for each wildfire and extracting information from its XML definition files. This gave us the burn acreage, boundaries, and names of each of the 310 documented wildfires.

Second, we constructed a 400x700 grid-map of Colorado, where each grid coordinate corresponded to a specific geo-coordinate (latitude-longitude pair). We then ran a number of operations on the LFMC data using pandas to find the average LFMC value for each fuel type at each geo-coordinate in their data. From there, for each point in the grid-map, we assigned its LFMC values to be equal to those of the nearest geo-coordinate in their data based on geodesic distance. We also ran a batch script executing HTTP requests to the Google Elevation API to collect the elevation data at every point in the grid-map using asynchronous calls due to the sheer size of the data we needed to collect (280,000 elevation values).

Lastly, we used this 400x700 grid-map to construct 50x50 sub-grid-maps for each wildfire using the boundaries from the MTBS data.These sub-grid-maps became the input examples to our models, and had a corresponding burn acreage label from the MTBS data.

2.4 Note on the Quality of the LFMC Data

Unfortunately, there were some unforeseen flaws with the LFMC dataset we used that we were not aware of initially due to the format. For example, different locations had varying amounts of dates for which LFMC measurements were recorded, so we had to average them to ensure that we could have our data aligned in our input matrices for our models. Similarly some locations had many recordings in close proximity whereas others did not have any recordings for potentially hundreds of kilometers,

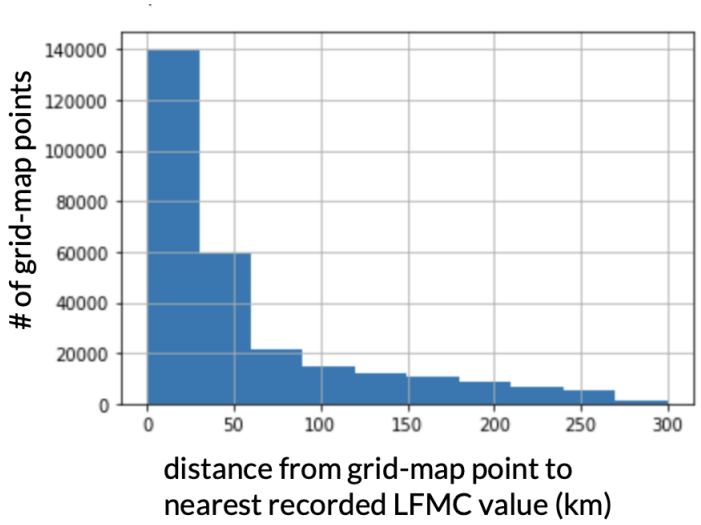

Figure 1: The distance (km) from grid-map points to the nearest LFMC value.

meaning a substantial portion of the points on our grid-map have LFMC values for a location that is very far away.

As seen in the figure above, the majority of points in the grid-map were assigned LFMC values from a location less than 50km away. However, as mentioned, there is still a large number of them that were greater than 50km away, which is likely too far as a dryness estimate for a region. This method resulted in a lot of data duplication across grid-map points in close proximity which was also not ideal. We will mention some methods in the *Next Steps* section for how we could circumvent these issues as we continue the research.

3 Models

3.1 CNN

We initially developed and trained a CNN model. The input to the CNN were 50x50 grids of feature vectors. The feature vectors contained the latitude, longitude, elevation, the LFMC scores for the 45 fuel types, and finally the distance to the nearest fuel type. (In retrospect, the last feature was likely not necessary and potentially only added complexity.) We saved the first 62 fires for testing and validation. This set included fires varying in size. The remaining 248 fires were used as our training set. The labels were the total number of acres burned.

The CNN architecture was composed of 2 convolutional layers followed by 4 fully connected layers. The first convolutional layer had 32 filers and a kernel size of 3x3; the second layer had 64 filters and a kernel size of 3x3. The fully connected layers had 64, 32, 8, and 1 neuron(s), respectively. We used rectified linear unit activation functions for all layers because negative predictions did not make sense for our use case.

Overall, the CNN performed poorly. Once we established our training and test sets, we hypothesized that this would be the case due to lack of data and non-consistent LFMC measurements. We ran experiments with normalized and non-normalized data and found the mean squared error (MSE) to be quite different. Our first phase of experiments used 25 epochs and the second phase used 50 epochs. We ran larger experiments with 200 epochs, but saw negligible difference to the 50 epoch experiments. The results are shown in the Table 1 below. We also considered different architectures (more convolutional layers, less fully connected layers, etc..) and found the test MSE for these models to be significantly higher.

Table 1: CNN Experimental Results

Normal/Non-Normal Data	Num Epochs	MSE (millions)
Non-Normalized	25	$\sim$164.5
Non-Normalized	50	$\sim$105.7
Normalized	25	$\sim$99
Normalized	50	$\sim$98.5

We observed two different loss value phenomena. For the non-normalized data, the loss was completely unstable for the training and validation sets, but on the normalized data, the loss saturated to an extremely high value and did not fluctuate for the remainder of the training. Given the lower test set MSE, we suspect the loss saturation on the normalized data indicates that the model may have performed better on larger fires than smaller ones. The larger MSE for the non-normalized data may indicate that this model performed better on smaller fires and greatly underperformed on larger fires. Of course, it is also possible that the models simply performed poorly on the whole test dataset.

3.2 Other Models

We trained three other models to compare these against the performance of the CNN using the scikit learn library. These models included a linear regression, a random forest regression, and a support vector regression. We split the data into training and test sets in the same manner as we did for the CNN, the first 62 fires in our processed data being assigned to the test set and the remaining 248 fires being assigned to the training set.

Unsurprisingly, the linear regression model yielded harsh results. We observed a mean squared error of 1.17e+28 on the test set. We suspect that this was not a good model to represent the problem we are trying to solve because the data is non-linear.

The random forest regression and support vector regression models yielded surprisingly comparable results to the CNN. First, with a maximum allowed depth of 20 and a total of 30 estimators within the random forest regression, we observed a mean squared error of 1.97e+8 on the test set. Second, with the RBF kernel, a slack parameter of 1.0, and a penalty parameter of 0.1, we observed a mean squared error of 8.3e+7 on the test set. We suspect this lower error (relative to the linear regression model) was achieved due to the non-linearity of the data.

4 Conclusion

In conclusion, the support vector regression model produced the lowest MSE. This is contrary to our expectations as we presumed that the convolutional layers of the neural network would help to identify "spread-prone" areas more accurately than other models. We believe that the poor performance yielded by the CNN is due to the significant lack of data and the inconsistencies found in the LFMC. Had the LFMC data contained more consistent measurements both across a time span and geographically, we believe the CNN could have performed better.

Fortunately, we were able to observe a massive decrease in MSE on the test set between the linear regression model and the 3 nonlinear models (CNN, SVR, and RFR). Because of this, we are able to conclude that LFMC and elevation likely do have some sort of relationship with the burn acreage in a nonlinear pattern, but more research is needed to learn more about this relationship.

5 Next Steps

In the future, we would like to explore using other datasets that are more complete. We originally investigated datasets that include a "greenness index," and at the time did not find one with enough historical data. However, we could potentially combine several of these datasets. Another option would be to use the ESA Sentinel-2 1.6 micrometer band dataset. This band is useful for visualizing combustion. It is comprised of greyscale imagery data. The downside to using this is that it is unclear how much history was captured. We would ideally need a few decades of history. We would also be

required to label this data ourselves, and that would take a considerable amount of time and could be prone to errors.

Using our current dataset, we could include data from all western U.S. states. We could not do this for the project due to computational limitations (for example, running the LFMC expansion job took 4 hours for Colorado alone). We could also consider outputting a perimeter map instead of simply total acres. However, this would almost certainly require more data. We would also like to consider an object detection approach to predicting fire likelihood. The input would be satellite imagery before a wildfire event (like the Sentinel-2 data, for example, assuming we had the correct timeframe), and the labels would be "fire" or "not fire" for each pixel. Looking at the "fire" class, you could use the confidence score as the measure of how likely a geographic area was to experience fire.

References

[1] M.L. Chas-Amil, J. Touza, J.P. Prestemon, C.J. McClean. (2012). *Natural and social factors influencing forest fire occurrence at a local spatial scale*. Modelling Fire Behavior and Risk. Global Fire Monitoring Center: Freiburg, Germany, 181-186.

[2] David Radke, Anna Hessler, Dan Ellsworth. (2019). *FireCast: Leveraging Deep Learning to Predict Wildfire Spread*. International Joint Conference on Artificial Intelligence.

[3] Krishna Rao, A. Park Williams, Jacqueline Fortin Flefil, Alexandra G. Konings. (2020) *SAR-enhanced mapping of live fuel moisture content*. Remote Sensing of Environment, Volume 245.

[4] Jessica Lowrey, PhD. (2007) *Drought, Wildfire, and Climate Connections: Research in Forests of the Western U.S. Intermountain West Climate Summary*. University of Colorado.

[5] Finco, Mark; Quayle, Brad; Zhang, Yuan; Lecker, Jennifer; Megown, Kevin A.; Brewer, C. Kenneth. 2012. *Monitoring Trends and Burn Severity (MTBS): Monitoring wildfire activity for the past quarter century using landsat data*. In: Morin, Randall S.; Liknes, Greg C., comps. Moving from status to trends: Forest Inventory and Analysis (FIA) symposium 2012; 2012 December 4-6; Baltimore, MD. Gen. Tech. Rep. NRS-P-105. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. [CD-ROM]: 222-228.