Wildfire Detection

Final Project Proposal

Undergraduate Laboratory at Berkeley

Data Science ULAB

Mentors

Hunter McCoy
Nandita Radhakrishnan

Group Members

Aashritha Srirambhatla
Neel Datta
Peter F. Grinde-Hollevik
Pratyush Das
Shani Lyubomirsky
Sifath Mannan
Vaibhav Mohata

Certification

I certify that this project proposal is an original collaborative effort that I have carried out in conjunction with only those listed here and with the assistance of the subgroup mentors. I affirm that I have properly cited all references, including journals, textbooks, and other resources.

Peter F. Grinde-Hollevik
Pratyush Das
Shani Lyubomirsky
Neel Datta
Sifath Mannan
Vaibhav Mohata
Aashritha Srirambhatla
Hunter McCoy
Nandita Radhakrishnan

Background

Wildfires are a growing natural hazard in most regions of the United States, posing a threat to life and property, particularly where native ecosystems meet developed areas. (US Geological Survey, 2006)

The societal consequences are clear: Just in the United States alone, over 13 million acres of land burned, $3.2 trillion were spent in wildfire suppression costs, and over 14.700 structures were lost in the 2020 wildfire season. (National Large Incident Year-to-Date Report, 2020) Besides from that, wildfires carry both short- and long-term negative impacts, as illustrated in the following table.

Short Term Negative Impacts	Long Term Negative Impacts
• Ravaged Local communities • Loss of Sensitive Infrastructure • Loss of Timber • Loss of Forage • Loss of Wildlife habitats • Watershed Pollution	• Loss of Cultural and Economic resources • Reduced health from air pollution • Loss of access to recreational areas

Fig 1: Short Term & Long Term Negative Impacts of Wildfires
(US Geological Survey, 2006)
(Walsh, C, 2020)

Besides from leaving its mark on Earth, wildfires can be observed from several satellites collecting data about surface temperature, moisture level, and forest density including others. These large amounts of data are openly available and can be enriched with historical data and data collected at Earth's surface to create solid datasets suitable for prediction analysis.

Previous research on wildfire prediction has utilized a variety of methodologies to build and test models such as neural networks and support vector machines (Sayad et al., 2019) or decision tree classifiers (Cortez et al., 2007). These studies focussed primarily on parameters such as Normalized Difference Vegetation Index (NDVI), Land Surface Temperature (LST), and Thermal Anomalies collected from the Terra Satellite's MODIS-instrument. Other approaches have included weather-prediction and fire-behavior models that combine fuel conditions and terrain (Coen et al., 2020).

Thus, there appears to be very limited research that effectively incorporates instantaneous weather conditions such as air temperature, wind and soil moisture as potential parameters in wildfire prediction models.

Research Question:

Beyond the current metrics that determine the present likelihood of a wildfire, is there a way to incorporate for climate change in these functions such that we can identify the instantaneous and future likelihoods of fires in a certain location? Can we achieve a higher prediction accuracy using a temporal data-based approach in comparison to current measurements of wildfire prediction?

The current metrics for measuring the probability of wildfires in an area use a fuel flammability index that takes into account the soil moisture and NDVI (normalized difference vegetation index). There are datasets available that would allow us to compare soil moisture and fuel flammability with temperature and atmospheric conditions, such as *Long-Term Soil Moisture Data sets* from Dr. Paul Dirmeyer at UCAR. We can use atmospheric conditions as the link between climate change and fuel flammability index in order to create our projections for the risk of wildfire in a certain location.

Our research question is composed of 2 distinct components:

1. We predict that there exist latent factors that have contributed to the increase of the intensity and reach of wildfires that we can identify based on the pattern and spread of wildfires throughout the last 50 years.
2. We will create an architecture that successfully predicts the likelihood of wildfires in a given region at a true positive rate greater than 95 percent.

Some of the factors that we need to consider are source and quality of the data, past historical data, and wildfire regions that lack satellite imagery / records. Using historical satellite imagery and recorded statistics from previous wildfires, can we successfully predict future wildfires?

We realize that wildfires are random events but with the recent influx of wildfires in California, by implementing climate change into the models of prediction, we will be able to better mitigate the likelihood of wildfires, through periodic sprinkling and positioning of fire stations. As of Oct 4, 2020, over 4 million acres of land had been burned in wildfires, more than double the previous single year record. By performing this research and generating a prediction engine, we hope to flatten the California wildfire curve. While the current research is good at generating models for current or immediately-close fires, there is still a gap in the use of data in considering the effects of climate change on the intensity of wildfires. We look to close that gap with our research and through the development of our prediction engine.

We hope to achieve a better architecture than the team of USGS by having true positive accuracy > 95% for both fire/no fire predictions. These results would allow for government agencies and other firefighting groups to identify and focus on areas that are at a higher risk for wildfires.

Methods

To build our predictive model, we will use a combination of multivariate regression and Convolutional Neural Network (CNN). Our data will be gathered from several wildfire datasets, including satellite imagery from MODIS and other wildfire datasets (see resources section). Our data will be sourced from California, and will be augmented to balance the ratio of fire/no fire. The data will then be cleaned and regularized after sourcing, and we will use a multivariate regression model and principal component analysis to identify the variables which most affect the risk of a region being lit on fire.

We will then use our cleaned data to map wildfires/statistics to regions, and then divide a region into smaller blocks and treat each block as a singular entity. We will then use those blocked variables to train a CNN using the modified data to predict the block in a region where a fire could start or spread to based on historical data.

All required libraries are publicly available, but we will need to create a free account to access the data on MODIS. If we have additional time, we will extend the project to all of North America and test the efficacy of other model types such as XGBOOST or Bagged decision trees.

Resources

- https://fas.org/sgp/crs/misc/IF10244.pdf
- MODIS
- Wildfires Data Pathfinder | Earthdata
- Combined wildfire datasets for the United States and certain territories, 1878-2019
- Remote Sensing Dataset on Wildfires
- Landsat Satellite Imagery on Burned Land
- Can we use ML to predict Wildfires?
- Ipython Notebook on Wildfire Prediction
- Hazardous Fuels Treatments
- Indexes (Live Fuel Moisture, Dead Fuel Moisture, etc.)
    - Normalized Burn Ratio (NBR)
    - Normalized Difference Vegetation Index (NDVI)
- US Department of Forestry Satellite Data
- API Tool for California Weather/Fire Data
- NASA Forest Fire Prediction
- Google Tree Mapping

Budget and Equipment

- If necessary, we will need to purchase supercomputing time or AWS credits to generate data and train the model.
- If possible, we will get access to Soda Hall to use their computers for modeling and for other research purposes

Proposed Collaborations

1) National Oceanic and Atmospheric Administration

   The NOAA is an agency in the United States Department of Commerce that studies the conditions of oceans, waterways, and the atmosphere.

2) Pacific Southwest Research Station

   The Pacific Southwest Research Station represents the USDA Forest Service and Development in the states of California, Hawaii, and the U.S affiliated Pacific Islands. It conducts a very broad research on natural resources to keep a sustainable forest ecosystem.

3) Perimeter

   Perimeter is a start-up company focused on creating an app that detects wildfires.

4) HUMNET Lab

   HUMNET is an environmental engineering lab that does computational research on quantifying events in the natural environment.

6) Landscape Research Group

   The Landscape Group under the College of Environmental Design at UC Berkeley uses GIS and spatial modeling in order to understand and prevent dangerous landscape changes.

7) The Center for Fire Research and Outreach at Berkeley

   A collaborative effort between faculty, researchers and stakeholders, the Center for Fire Research and Outreach promotes wildfire research and aims to apply a multidisciplinary approach to wildfire management.

8) Kelly Research and Outreach Lab

   The Kelly Research and Outreach Lab works with map technologies to observe how California landscapes are changing in order to aid wildlife conservation groups.

9) UC Berkeley Forest Fire Experts

- Professor Scott Stephens (Email : sstephens@berkeley.edu)
- Professor Bill Stewart (Email : billstewart@berkeley.edu)
- Professor Brandon Collins (Email: bcollins@berkeley.edu)
- Justin Brashares, Professor Environmental Science, Policy and Management at UC Berkeley (Website)
- Duncan Callaway, Professor Energy and Resources at UC Berkeley (Website)

References

a. U.S. Geological Survey, U. S. D. of the I. (2006). Wildfire Hazards—A National Threat https://pubs.usgs.gov/fs/2006/3015/2006-3015.pdf

b. National Interagency Fire Center. (2020). National Large Incident Year-to-Date Report.

https://gacc.nifc.gov/sacc/predictive/intelligence/NationalLargeIncidentYTDReport.pdf

c. Walsh, C. (2020, January 31). *The long-term effects of wildfires*. Harvard Gazette. https://news.harvard.edu/gazette/story/2020/01/the-long-term-effects-of-wildfires/

d. Y.O. Sayad, H. Mousannif, H. Al Moatassime, Predictive modeling of wildfires:A new dataset and machine learning approach, Fire Safety Journal (2019), doi: https://doi.org/10.1016/j.firesaf.2019.01.006

e. Coen, J. Douglas, C. (2020) Computational Modeling of Large Wildfires: A Roadmap https://opensky.ucar.edu/islandora/object/articles%3A18989/datastream/PDF/download/citation.pdf

f. P. Cortez and A. Morais, "A Data Mining Approach to Predict wildfires using Meteorological Data," Proc. 13th Port. Conf. Artif. Intell., pp. 512–523, 2007https://dergipark.org.tr/tr/pub/dumf/issue/56929/661925

g. O. Rios, E. Pastor, M.G. Valero, E. Planas (January 2016) "Short-term fire front spread prediction using inverse modelling and airborne infrared images" https://www.researchgate.net/publication/309531321_Short-term_fire_front_spread_prediction_using_inverse_modelling_and_airborne_infrared_images