

Shootout 2017 Problem Statement

Understanding Wildfires

1 Background

In the fall of 2016, a number of wildfires occurred in the southeastern United States. Spanning over 5 states and several weeks, more than 33 wildfires broke out and caused extensive damage. As a result, over 80,000 acres were burned, dozens of people were injured, and at least 14 fatalities were reported [1]. Conditions and the physical landscape of these areas were very favorable for wildfires to occur. Many of the counties affected were experiencing significant and extreme droughts. Causes for the fires were varied, but at least some were suspected to have been started by arson.

Wildfires are a frequent problem for many parts of the world. They may even be increasing in some areas of the United States. They start unexpectedly and may cause extensive damage to homes, businesses, wildlife habitat and air/water quality, require a great deal of money and resources to manage, and affect the quality of life of those living in the areas they occur. Understanding the conditions, causes, and impacts of wildfire activity may help communities and individuals better prepare for the likelihood of their occurrence in the future, saving both money and, more importantly, lives.

2 Problem Overview

The formation of a wildfire is based off two factors; conditions and causes. *Conditions* are the factors that contribute to wildfire development are based on several physical and environmental factors, and they will change throughout the year by geography. *Causes* are the events that initially ignite the fire. These can be sporadic or deliberate and hard to predict.

Wildfires are strongly associated with weather. Changes in weather and precipitation can either increase or decrease the chance of wildfires breaking out. Warmer temperatures and less precipitation are clear indicators for wildfire potential. Additionally, research has shown that earlier-than-normal spring seasons in the US may be linked to an increase in wildfire activity [2]. Because some regions in the country tend to be drier and warmer than others, differences in geography and time of year in relation to wildfire activity should be considered.

The physical geography of an area is highly indicative of whether a wildfire is likely to even occur. The type of land cover is like fuel for a fire. Some land cover types are more likely to burn, while others may not be able to burn at all. There are several ways to classify land cover. In a broad categorization by the Multi-Resolution Land Characteristics Consortium's (MRLC) "National Land Cover Database", land cover types are put into 8 classes; Water, Developed, Barren, Forest, Shrubland, Herbaceous, Planted/Cultivated, and Wetlands. Each of these categories can be further broken down into more specific classes.

The link between increased and longer periods of drought with wildfire activity has been studied and shown to be correlated [3]. As a drought period goes on and precipitation is rare, the intensity of the drought may increase. The National Integrated Drought Information System classifies droughts into 5 categories; D0 (abnormally dry), D1 (moderate drought), D2 (severe drought), D3 (extreme drought), and D4 (exceptional drought). While the intensity of a drought may decrease with significant precipitation events, the period of drought in an area can last several weeks, months, or even years. When an area is experiencing a drought period, there should be extra concern for the potential of a wildfire event.

When it comes to wildfire causes, they can occur rather spontaneously. Some wildfires start due to natural and unavoidable circumstances. Other wildfires are started due to some sort of human activity, either accidental or deliberate. Because they are hard to predict, it is difficult to identify when and where a fire will break out. However, fires in some areas may be more likely to start due to specific factors. For example, lightning is a known cause of some wildfires. Lightning frequency, however, varies by season and region. It is important to understand the causes of wildfire events to better prepare for their occurrences.

Not all wildfires are alike in terms of damage. Depending on where and when the fire occurs, wildfire impacts vary greatly. Some wildfires can be contained quickly, while others can go on for days or even weeks. This may be due to several factors including the weather and drought conditions, cause of the fire, and the preparedness of the area.

After wildfires occur, information about the events are captured through narratives put together by newspaper articles, the National Weather Service, and various government agencies that help fight fires. This information can give the background of the conditions before the wildfire occurred, the potential causes of the fire, and information about the damage and intensity. This text contains information that may not be fully captured in the wildfire events database; text mining methods may be used to extract useful data.

3 Primary Tasks

Using SAS software and analytics methods, your team is being tasked with examining wildfire events and the underlying factors that contribute to their formation. Your research of wildfire events and their impacts may help communities and individuals plan better according to appropriate risks.

3.1 Wildfire Conditions

As previously mentioned, conditions for wildfires change based on a number of factors. At any particular time and location, given a cause, the chances of a wildfire occurring will change. Using any of data provided, develop a model for predicting propensity of wildfire at any given time based on the current conditions at that same time. What type of model techniques did you find effective? What are the conditions most favorable for a wildfire?

3.2 Wildfire Causes

A fire cannot start without cause (spontaneous combustion, spark, etc.). Sometimes it can start from a natural event. Other times it can be linked directly to human activity. The cause of a wildfire is often determined after the event has occurred. Using the data provided and the “Wildfire Narratives”, extract additional information that may help classify wildfires by cause. For example, how often are you able to determine the initial cause of a wildfire? Produce a model that will predict the probability of potential causes of a wildfire by time and geography.

3.3 Wildfire Impacts

The “impact” of a wildfire can be defined in many ways, including area burned, people displaced, cost to contain, cost in damage, and lives lost. Impacts vary greatly based on the cause, the physical geography, and the weather conditions before and after ignition. In the data provided, the impact of the wildfire events is not always captured adequately in the variables. Using the “Wildfire Narratives”, extract additional information to help quantify how impactful the wildfires are in terms of *acres burned*. Build a model that predicts acres burned from the provided variables.

4 Scenario Task

Using your previously created models and building off your initial work, you are given two additional tasks aimed at putting your models to use.

4.1 Score the scenario year of conditions

Now that you have developed models that examine wildfire conditions, causes, and potential impacts, you'll want to put them use in a real way. The "risk" of a wildfire can be developed using the product of wildfire propensity, sum of cause probabilities, and potential impact (in acreage). Long-range forecasts from the National Weather Service are provided containing a single year of variables. Using your previously developed models, create a new "risk model" that incorporates the outputs from your previous three models. In what county and week did your risk model find the highest potential of acres burned?

5 Data

To perform your analysis, your team is provided with several datasets, including:

- Wildfire Episodes/Events (4,956 records)
 - *Wildfire events in the United States, by county/zone, by date from 2001 to 2015.*
- Wildfire Narratives (4,956 records)
 - *Supplemental text data providing additional details about specific wildfires not included in the "Wildfire Episodes/Events" dataset, by episode/event from 2001 to 2015.*
- Drought Severity (2,520,477 records)
 - *Drought severity in 5 categories, by county, by week from 2001 to 2015.*
- Weather Station Data (13,122,253 records)
 - *Temperature and Precipitation variables, by weather station, daily from 2001 to 2015.*
- Storm Events (881,599 records)
 - *Storms and other significant weather events, by county/zone, 2001 to 2015.*
- Land Cover (3,109 records)
 - *20 land cover type percentages, by county.*
- Population
 - *Total population, by county, 2000-2009 (3,143 records).*
 - *Total population, by county, 2010-2015 (3,220 records).*

Additional information on every dataset can be found in the "Data Dictionary."

6 References

[1] Schelzig, Erik. "TOURISTS STREAM BACK TO GATLINBURG, BUT REBUILDING LOOMS." *News from The Associated Press*. N.p., 9 Dec. 2016. Web.

[2] Westerling, Anthony L., et al. "Warming and earlier spring increase western US forest wildfire activity." *science* 313.5789 (2006): 940-943.

[3] Australia Parliament, House of Representatives Select Committee on the Recent Australian Bushfires. "A Nation Charred: Inquiry Into the Recent Australian Bushfires". Parliament of the Commonwealth of Australia, Canberra ACT 2601, Australia. (2003)

