A Data Science Approach to Increasing the Efficiency and Effectiveness of Suppressing Wildland Fires

Janaar Harbour, Olivia Samples, Kun Zhao

Focus of Project

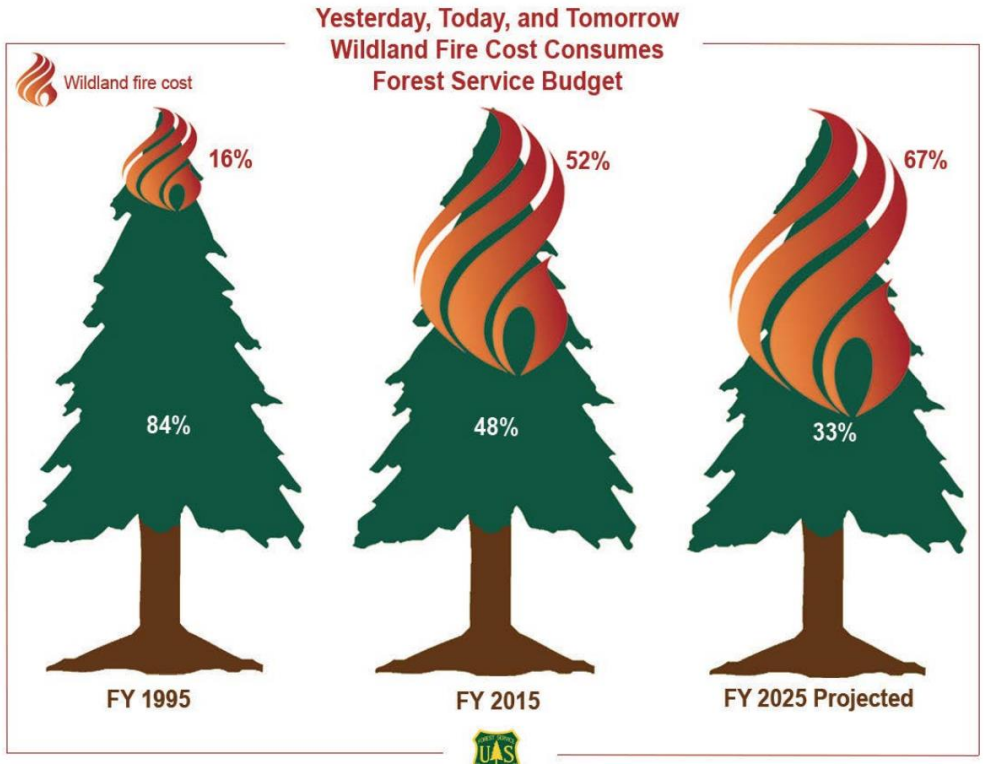
We propose a real-time model that will predict the optimal number of personnel needed to effectively fight a wildfire without using unnecessary resources, leaving units to be nimble as other fires and preventative needs arise. Additionally, we assess how changes in weather, region, and prescribed burning affect resource deployment decisions and fire suppression success.

Research Question

- **What is the optimal number personnel to be deployed in order to efficiently and effectively suppress wildland fires?**
  - Do **prescribed fires** decrease the number of deployed personnel in future wildland fires?
  - How does the **region** of the US in which the fire is located affect the number of needed personnel?
  - Do the wet and dry seasons (**weather**) influence the need for personnel differently?

Significance of Question

- Rapid increase in fire management cost that is harmful to the overall forest service budget
  - The US budget for wildfires **tripled** from 1995 to 2015 and is expected to exceed $1.8B in cost by 2025. Our project is designed to increase efficiency and help lower the cost of this rapidly growing budget.
  - Timothy Ingalsbee claims that fire management is accountable for **at least half** of fire suppression costs.



418% increase in fire management cost

- Lee, Y., Fried, J. S., Albers, H. J., & Haight, R. G. (2013). Deploying initial attack resources for wildfire suppression: spatial coordination, budget constraints, and capacity constraints. Canadian Journal of Forest Research, 43(1), 56-65.
- Cortez, P., & Morais, A. D. J. R. (2007). A data mining approach to predict forest fires using meteorological data.
- Martell, D. L. (2015). A review of recent forest and wildland fire management decision support systems research. Current Forestry Reports, 1(2), 128-137.

Successful Papers

Optional Techniques

Real Time Model

- Martin-fernÁndez, S., Martínez-Falero, E., & Pérez-González, J. M. (2002). Optimization of the resources management in fighting wildfires. Environmental Management, 30(3), 352-364.

XGBoost Model, Shap and Rank Order Label Encoding (ROLE)

- John Carr, Matt Lewis (2019) Predicting Number of Personnel to Deploy for Wildfire Containment

Review of John Carr's Paper

- Carr and Lewis (2019) created a practicum project focused on a unique data set received from the US Department of Interior. Their analysis began with acquiring and cleaning the data, and then applied tools such as Rank Order Label Encoding and the model of XGBoost and Shap on data for optimal predictions and explanations. Their model was able to explain 61% of the data variation. In a discussion with John Carr, one of the areas that would extend this project was to add differentiating data to the resource table. Their data set is the basis for our project.

Describe Data

- IRWIN Observer – US Department of the Interior's complete records for fires
- Jan 1st, 2015 to Dec. 31st, 2019
- The incident dataset included 135 features that describe 125,367 fire incidents.
- The resource dataset included 16 features that describe resources (i.e. personnel, equipment) used in 39,837 incidents.

Describe Approach to Analyzing Data

- Obtain permission from the US Department of Interior to access data
- Acquire data
- Create a key for all features
- Examine data to determine the most significant features
- Analyze data utilizing tools such as ANOVA
- Create the model
- Determine the different effects that prescribed fires, region and weather have on our model

Challenges

- Time constraint
- Lack of thorough content knowledge
- Reliable data or not?
- Implementation of solution may challenge status quo
- Cost savings in fire management may not directly benefit the forest service