Environmental Statistics Final Project Proposal - Group 3

Matthew David Wallace, Raymond Owino, Alex Salce

November 15th, 2024

Topic and Datasets

Our primary research questions will be centered around wildfire data in (AZ only? US? Other states? More than one state?), and will be utilizing the following datasets.

- NIFC | Wildland Fire Incident Locations
  - The primary dataset for our research questions will utilize the Wildland Fire Incident Locations dataset. This dataset includes 97 variables including point location data for wildfire incidence in the US from 2014 to the present as recorded from the IRWIN (Integrated Reporting of Wildland Fire Information) system, which aggregates fire data from many data sources. The dataset contains many features, but we will primarily utilize the point data for fire origin, as well as total acreage the fire went on to burn. Some of the assigned tasking will include investigation of other possibly useful features in the dataset.
- NIFC | WFIGS Interagency Fire Perimeters
  - This dataset is the best available fire perimeter data for individual wildfires in the US. Our research may necessitate a deeper dive into some of the individual wildfires and the areas they covered spatially.
- LiDAR data
  - The LiDAR data should give us features of the terrain that can be used as predictor variables for our model fits.
- rFIA data
  - We may incorporate abundance data for standing trees (using `TPA & BAA` data) as predictor variables for our model fits.

Research Questions

Using the above data (not limited to the above), we plan to approach the following three research questions.

(1) Research question 1?
(2) Research question 2?
(3) Research question 3?

Statistical Methods

The statistical methods we plan to use to address answer these questions are the following.

- Spatial linear model
  - Use additional datasets for prediction variables
- Point process model
  - Use additional datasets for prediction variables

Team Responsibilities

Media	Raymond	Matthew	Alex
Report	test	test	test
Code	test	test	test
Video	test	test	test

Anticipating Challenges

Challenges that we anticipate:

- Combining data, cleaning data, and filtering data
  - We intend to use data from multiple sources, which will present challenges for combining
- Deciphering available data
  - These datasets have many variables included, some of which will require researching to better understand whether they can be useful in answering any of our research questions or providing useful predictor data for our models.
- Dataset size
  - Can we reasonably answer all of our research questions, given that we are working with fairly large datasets? We may encounter computational challenges getting our models to work, given that we will be working with so much data. If we do run into challenges, we may need to refine our research questions.
- Available data
  - Will we have enough data available to utilize to fit the models we hope to support our research questions with the data that we currently have? We may need to acquire more data, or refine our research questions.