CPSC368 Project Group 8
March 22, 2024

Investigating the Connection Between Annual Precipitation and Subsequent Wildfire Severity

Introduction

This project investigates the connection between annual precipitation and river flow on wildfire severity in the BC interior. As the meteorological processes that cause precipitation or drought are too complex to use as direct inputs for the scope of this project, we use historical annual precipitation and river flow data to make high-level predictions of the severity of wildfires the subsequent year. We use historical weather data captured from a federal weather station located at the Kelowna International Airport. We use river flow data from Vernon Creek. This river flow data was collected very close to the Kelowna International Airport - the location of the precipitation data and river flow data is essentially the same. As we could not find access to specific geographic wildfire data, to measure the severity of a wildfire season we use the total hectares burned in BC for that season.

We chose to focus our study in the Kelowna region because we know the Okanagan region is regularly impacted by wildfires. However, the methods used in this analysis could be reused with data from a broader range of locations. In recent years, climate change has impacted traditional seasonal conditions. As such, examining the relationship between weather patterns and the corresponding severity of wildfire is important for governmental planning as well as for environmental protection.

Visualizations and Analysis

1. Wildfire total hectares and total precipitation by year
    a. We plotted both annual total wildfire size as well as total precipitation by year. Visually, we notice that both wildfire size and total precipitation are increasing with time. We also notice that peaks in wildfire size correspond to lows in precipitation and vice versa. This makes sense with our predictions. Lower rain and snowfall leads to dryer land that is more susceptible to fires.

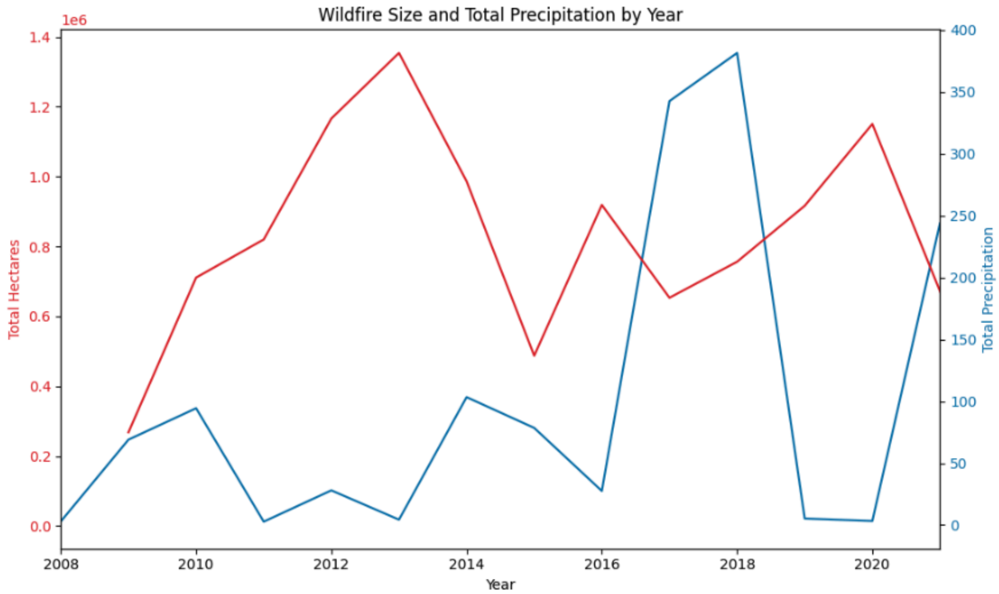

Figure 1. Python generated visualization of total hectares of wildfires in BC and total precipitation at the selected location (Kelowna Airport) by year.

2. Wildfire total hectares and river flow by year
   a. We plotted both annual total wildfire size as well as river flow by year. There are no glaring takeaways from the visualization. River flow appears to be quite variable, with some years experiencing a flow (in m^2/s) close to 1 and other years above 12. Past the year 2016, there does appear to be an inverse relationship between flow and wildfire size, similar to the relationship for total precipitation. This makes sense as river flow is a function of precipitation.

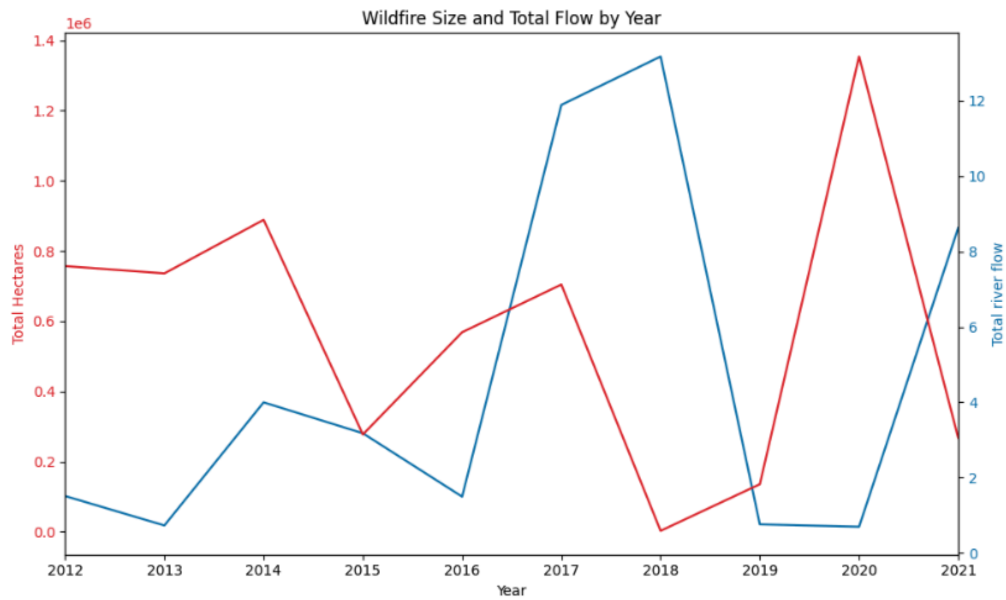

Figure 2. Python generated visualization of total hectares of wildfires in BC and average river flow at the selected location (Middle Vernon Creek) by year.

3. Wildfire total hectares versus total precipitation
    a. We plotted total precipitation versus wildfire size. As predicted, there is a negative relationship between the two, with a slope of -2171.57. This means the lower levels of precipitation correlate to greater wildfire size. However, this correlation is visual. The R-squared value is only 0.15; values below 0.4 generally suggest low correlation.

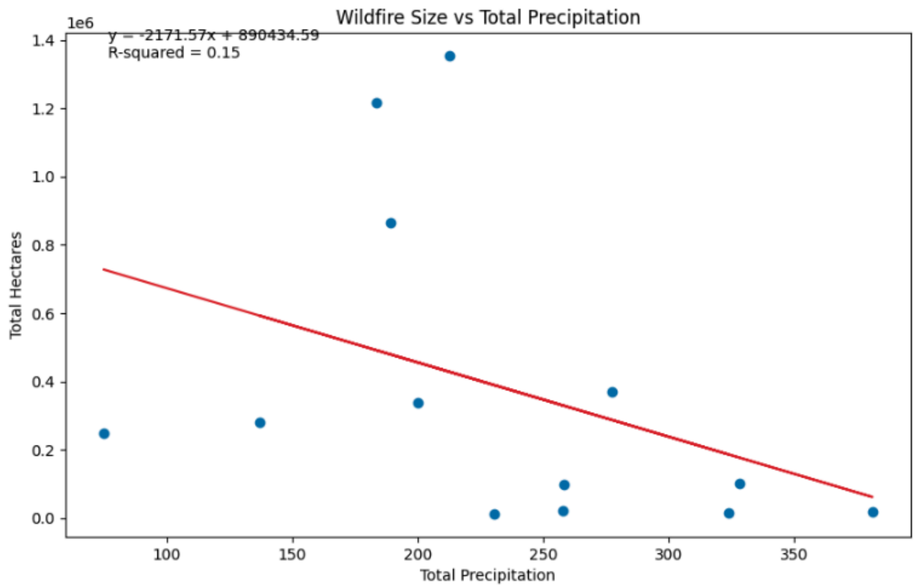

Figure 3. Python generated visualization of total hectares of wildfires in BC versus total precipitation at the selected location (Kelowna Airport) by year.

4. Wildfire size prediction model
    a. Our initial intended goal of this project was to create a simple model capable of predicting wildfires from precipitation data. We used Linear Regression (sklearn machine learning model) to find coefficients for precipitation to create a model. We determined model coefficients of: [-1883.1470197] and an intercept of: 803793.5676173045. Our model has a mean squared error of 63392155520.28 and a coefficient of determination (R-squared) of 0.38. This mean squared error is quite high, indicating the model may not be fitting the data well. This may also be because we skipped hyperparameter tuning steps and other machine learning steps because that is not required in the scope of this project. The R-squared value is relatively low but suggests that ~38% of the variability in wildfire size is explained by the precipitation variable.
    b. We tested the model with a sample annual precipitation of 250 mm, which predicted a wildfire size of 333,006.81269139197 hectares. This prediction is high but realistic compared to other annual sums. A high prediction makes sense as an annual sum of 250 mm of precipitation is relatively low so this corresponds to a dryer season.