Elaine Zhang

Common Analysis – Visualizations and Reflection

Figure Descriptions

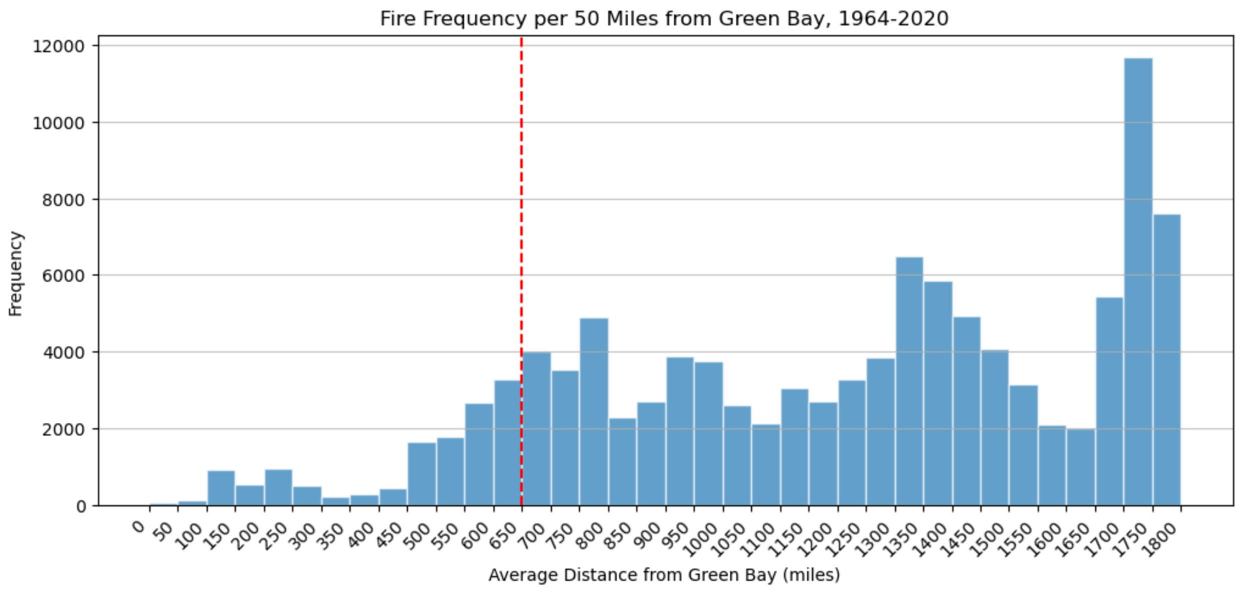

Fig. 1: Fire Frequency per 50 Miles from Green Bay, WI, 1964-2020

This visualization shows the total frequency of fires at different distances from Green Bay, Wisconsin in 50-mile buckets, from 0 to 1800 miles. The figure can be interpreted as follows: from 1964 to 2020, there were a little under 2000 fires between 500-550 miles of Green Bay, WI. The visualization includes wildfire data from the last 60 years and displays a red dashed line marking the cut-off at 650 miles for inclusion in the smoke estimate analysis. Fires right of the dotted line are not included in the smoke estimate calculations and model fit for smoke estimate predictions. Note that although we intend to include the last 60 years in this visualization, the wildfire data only includes fires as recent as 2020. As such, the data is not a complete representation of the last 60 years. The x-axis uses average distance as the basis for each bucket. Average distance is calculated as a proxy for the center of the fire using the average distance of the largest ring of each wildfire to Green Bay, WI. Ring data is provided in the wildfire GeoJSON data. Wildfires were filtered to include only those with an average distance less than 1800 miles to Green Bay, and then grouped into 50-mile buckets to create the count for each 50-mile grouping. There is no distinction for if the fire was east or west of the city, and no distinction for which year the fire occurred.

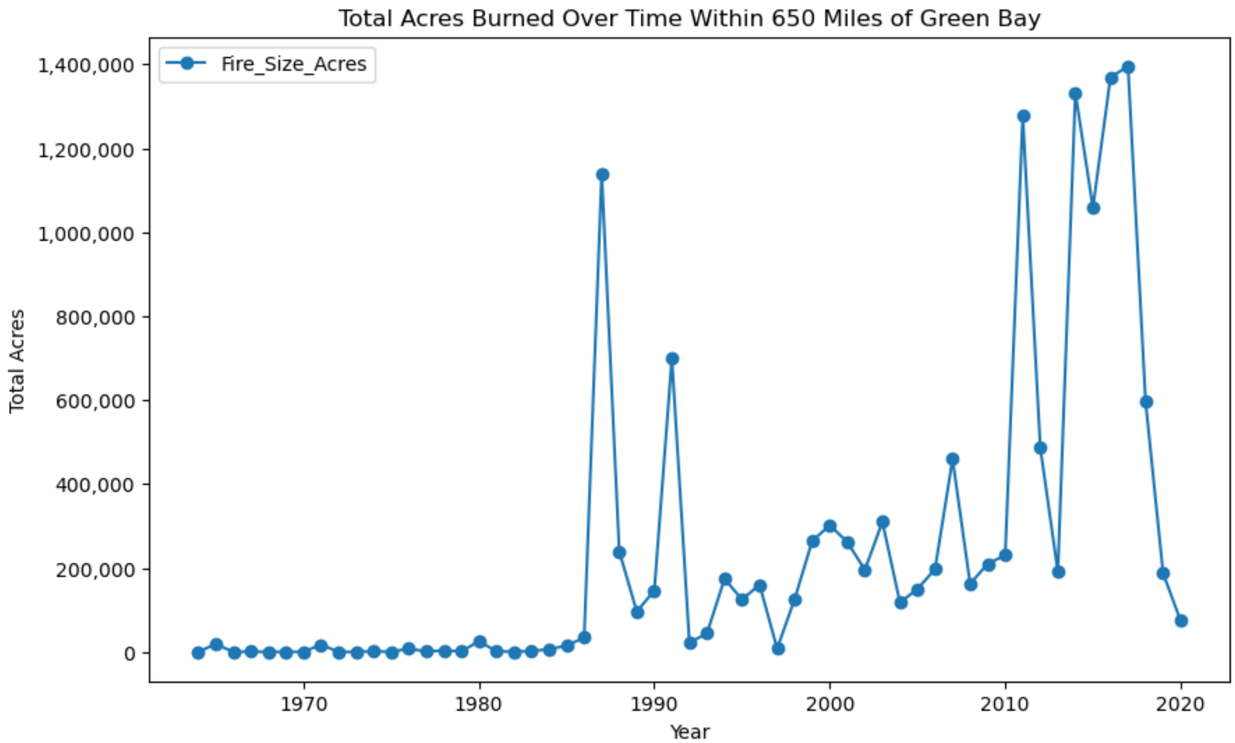

Fig. 2: Total Acres Burned from 1964-2020 Within 650 Miles of Green Bay, WI

This visualization displays the total number of acres burned each year in the last 60 years where the center of the fire is within 650 miles of Green Bay, Wisconsin. We plot years on the x-axis, and total acres on the y-axis. Note that the visualization cuts off at 2020 due to the lack of wildfire data past that year. The figure can be interpreted as follows: In 1987, almost 1.2 million total acres were burned in fires within 650 miles of Green Bay. For each year in the past 60 years, we sum each fire's size in acres for fires where the center is within 650 miles of Green Bay. We use average distance as a proxy of the distance between the center of the fire and Green Bay. It is possible for a fire's center to be within 650 miles but have a portion of the fire outside of the 650-mile boundary. In such cases, the fire's total burned acreage is included in this visualization regardless of how much of the fire was within 650 miles of Green Bay. It can be noted that fire distance is not represented in this graph, meaning that a large fire 650 miles of Green Bay and a large fire 2 miles of Green Bay are represented with the same weight in this graph.

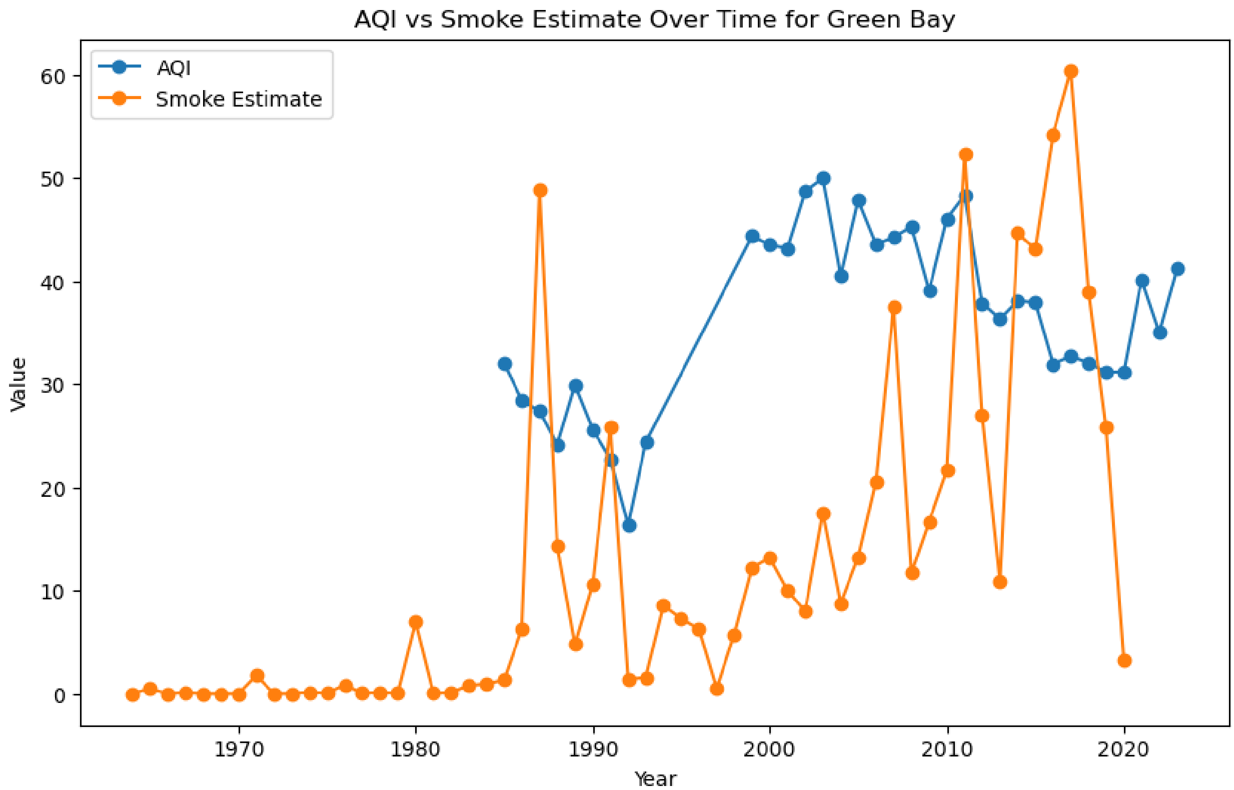

Fig. 3: Yearly AQI Estimates vs Smoke Estimates from 1964-2024 for Green Bay, WI

This visualization compares the AQI estimates acquired from the AQS API to the smoke estimate we created in the common analysis. We plot years on the x-axis, and AQI/Smoke Estimate values on the y-axis. To facilitate comparison between the two metrics, the smoke estimate calculation includes a scaling factor to keep the values on a similar scale as the AQI values. The figure can be interpreted as follows: For 1987, the smoke estimate for the fire season is around 50, and the AQI estimate is a little under 30. We can also use this figure to compare the AQI and smoke estimate trends. Note that the AQI data is unavailable before 1985 and is missing for 1994-1998. The smoke estimates stop at 2020 due to the lack of wildfire data past that year. The AQI estimate is calculated using daily AQI data by taking the mean of the AQI value of the seven pollutants over that year: CO, SO2, NO2, O2, PM10, PM2.5, and Acceptable PM2.5. The smoke estimates are calculated using a function of fire size and distance, where a larger fire closer to Green Bay will have a larger estimate and a smaller fire further away will have a smaller estimate. The yearly estimate is calculated as a sum of the smoke estimates divided by the number of days in the fire season, where fire season is defined as May 1st through October 31st. A more detailed explanation of the calculation can be found in the common analysis notebook.

Reflection and Attributions

Through answering the research question posed in the common analysis portion of this project, I learned more about geographic coordinate systems, the complexity of smoke, and how AQI is calculated. Through calculating each fire's distance to Green Bay, WI, I learned that there are multiple geographic coordinate systems to handle complexities with the Earth's shape. To accurately calculate these curved distances, points need to be in the same geographic coordinate system. I also learned about the many different components of smoke while coming up with a simplified way to represent smoke. Though the smoke estimate I created only takes fire size and distance into account, smoke in the real world consists of a mixture of particulate matter and gases like carbon dioxide, carbon monoxide, and many other chemicals. Smoke coverage is affected by factors such as wind direction, fire intensity, ecosystems, and many others. As such, I learned that there isn't an easy way to simply capture the smoke from a fire accurately with the wildfire dataset I used in this project. I also learned that there is a plethora of ways to create a simplified smoke estimate metric through discussions with my peers, as I will discuss further in the next section. While formulating my smoke estimate, I did research on how the air quality index is calculated to gather ideas and understand why AQI works the way it does. I learned that the AQI value includes a combination of seven pollutants, with each pollutant's number scaled to reflect pollutant-specific impacts on air quality and methods of metric collection. Each of the pollutants are measured differently on different scales before they are converted to AQI values. For example, CO values are collected for 8-hour concentrations, while PM2.5 values are collected for 24-hour concentrations. Each pollutant has its own breakpoints for each AQI category (e.g. Good, Moderate, Unhealthy, etc.) and then scaled to a standard AQI scale using the pollutant-specific breakpoints.

In the process of this analysis, most of my discussion with my peers focused on understanding how others were representing smoke estimates. Before talking to others, I thought to represent the smoke estimate by adding functions of fire distance and fire size, where a larger size would add to the estimate and a larger distance would detract from the estimate. Through conversations with others, I learned about the inverse-square law, where the fire distance is in the denominator of the estimate, rather than an additional term in the equation. In my conversations with a friend, we concluded that using the average of smoke estimates in a year could not fully represent fire frequency. In a case where a year had multiple similarly sized fires, the average of these estimates would yield a moderate number and not reflect the potentially compounding nature of smoke if these fires occurred in the same period. As such, we decided that it would make more sense to sum the smoke estimates and divide them by the number of days in the period, thus retaining the frequency aspect of the fires while normalizing the smoke estimate. To adhere to the requirements for the smoke estimate, the number of days should be defined as the number of days in the fire season rather than the whole year. The possibility of collaboration and discussion with others allowed me to gather more diverse ideas to tackle this problem than I would've had on my own. By collaborating and talking through ideas and thought processes with others working on a similar analysis, I had the opportunity to bounce ideas off others, influence others, and take suggestions from peers.

In this common analysis, I reused example code developed by Dr. David W. McDonald for handling GeoJSON files, calculating fire distances, and utilizing the AQS API. I also referenced Python documentation for the various packages used, such as the statsmodels documentation for use in the model fitting section of this analysis.