Apratim Tripathi   **6th November, 2024**

trips@uw.edu

DATA 512
Part 2 – Extension Plan

Motivation

Background:

Recent years have witnessed a marked increase in the frequency and severity of wildfires, particularly in regions close to and upwind of Jackson, MS. The smoke from these wildfires carries a mix of gases and fine particles that can significantly degrade air quality. This degradation impacts public health, especially in urban areas like Jackson, where population density can exacerbate the effects of poor air quality. The direct link between environmental factors and public health outcomes necessitates a thorough analysis to understand and mitigate these impacts.

Problem Statement:

The primary concern driving this analysis is the observation that periods of significant wildfire activity correlate with spikes in hospital admissions and mortality rates related to respiratory and cardiovascular conditions in Jackson, MS. Despite anecdotal evidence and preliminary data suggesting a strong relationship between increased smoke exposure and adverse health outcomes, there has yet to be a comprehensive study that quantifies these effects within the local context of Jackson. This gap in knowledge hinders effective public health planning and response, making it difficult for healthcare providers and policymakers to prepare for and mitigate the impacts of wildfire smoke.

Significance:

Understanding the impact of wildfire smoke on health outcomes is crucial for several reasons:

- **Public Health:** Provides insights into the conditions that most strongly correlate with poor air quality days, helping to inform public advisories and healthcare readiness.
- **Policy Making:** Data-driven analysis can lead to better-targeted health policies and resource allocation, ensuring that preventive measures and healthcare services are directed where they are most needed.
- **Community Preparedness:** Educating the public about the risks associated with wildfire smoke can lead to better community responses and individual preparedness during wildfire seasons.

- **Environmental and Urban Planning:** Insights from this study could influence urban planning decisions, such as the placement of green spaces or air filtration infrastructure, to mitigate the effects of smoke.

Research Questions:

1. How does exposure to wildfire smoke affect the incidence of respiratory diseases, cardiovascular problems, and overall mortality in Jackson, MS?
2. Can we predict the public health impact of future wildfire smoke events based on historical data?

Impact Focus

Health Impacts as Primary Concern:

This project centers on understanding the health consequences of wildfire smoke in Jackson, MS. Wildfire smoke, comprised of a complex mixture of particulate matter and harmful gases such as carbon monoxide and volatile organic compounds, poses significant risks to human health. The tiny particles can penetrate deep into the lungs and even enter the bloodstream, leading to a range of health problems from minor irritations to severe respiratory conditions and cardiovascular events.

Specific Health Outcomes Under Investigation:

The focus will be on quantifying the correlation between exposure to wildfire smoke and:

- **Respiratory Diseases:** Conditions such as asthma, chronic obstructive pulmonary disease (COPD), and other acute respiratory infections have been shown to exacerbate due to poor air quality caused by smoke.
- **Cardiovascular Problems:** Exposure to particulate matter can trigger heart attacks, stroke, and other cardiovascular events, especially in older adults and those with pre-existing conditions.
- **Mortality Rates:** An analysis of changes in mortality rates during and following significant smoke events will provide a broad picture of the public health crisis associated with wildfire smoke exposure.

Relevance to Public Health Policy:

Understanding these correlations is crucial for several reasons:

- **Emergency Response Planning:** Health services can better prepare for increased demand during wildfire events, ensuring adequate staffing, resources, and patient care protocols are in place.

- **Public Health Advisories:** With detailed information on the health risks associated with wildfire smoke, public health officials can issue more timely and relevant health advisories to help residents protect themselves from smoke exposure.

- **Long-term Health Strategy:** Insights from this study can inform ongoing public health strategies that aim to mitigate the long-term health impacts of recurrent smoke exposure, including community health monitoring programs and chronic disease management strategies.

Data and Methods

Data Sources

1. Wildfire Smoke Impact Estimates:

- **Source:** Derived from Part 1 of the course project.

- **Description:** This dataset contains annual estimates of wildfire smoke impact for Jackson, MS, quantified based on factors such as smoke density, area covered, and duration. These estimates provide a measure of the potential exposure of the population to wildfire smoke each year.

2. MSTAHRS Health Data:

- **Source:** Mississippi Statistically Automated Health Resource System (MSTAHRS).

- **Description:** MSTAHRS is a web-based query system that provides statistical data related to health outcomes in Mississippi. For this project, data on annual death rates by cause in Hinds County will be used. This includes data categorized by year, age group, sex, race, ethnicity, and place of residence.

- **Accessibility:** Data can be accessed through MSTAHRS [1], allowing for customized queries to obtain relevant data for specific diseases.

- **Data Handling Notes:** The dataset suppresses values less than 5 to maintain privacy, and rates based on fewer than 20 events are marked as unstable.

Focused Health Data Files:

- **birth_defects.csv**: Tracks annual deaths attributed to birth defects.

- **COPD.csv**: Annual death rates from Chronic Obstructive Pulmonary Disease.

- **hypertension.csv**: Deaths attributed to hypertension each year.
- **lung_cancer.csv**: Annual mortality rates from lung cancer.
- **heart_diseases.csv**: Annual mortality rates from heart diseases.
- **total.csv**: Total annual deaths for all causes combined.

Each file follows a simple format with two columns, "Year" and "Number," where "Number" represents the total deaths for that category per year.

Methodology

Data Preparation:

- **Process:** Merge the smoke impact data with health outcomes data from MSTAHRS by the "Year" column to align the datasets chronologically.
- **Objective:** Create a comprehensive dataset that includes both environmental exposure and health outcome data for each year, facilitating direct comparisons and correlation analyses.

Statistical Analysis:

- **Technique:** Utilize Pearson correlation coefficients to determine the strength and significance of relationships between annual smoke exposure and the rates of health issues such as COPD, lung cancer, and other conditions influenced by poor air quality.
- **Purpose:** Identify whether years with higher smoke exposure correlate with increased health problems, providing a statistical basis to infer causation.

Unknowns and Dependencies

Data Completeness:

- **Assumption of Accuracy:** The analysis assumes MSTAHRS mortality data is accurate. However, reporting inaccuracies could skew the health impact results of wildfire smoke.
- **Data Suppression:** Privacy protections result in suppressed data for small numbers, potentially limiting analysis in years or demographic categories with few events.

External Factors:

- **Healthcare Access:** Variations in healthcare access or quality during the study period could influence health outcomes independently of smoke exposure.

- **Public Health Policies:** New or changing public health policies could affect health outcomes. These might include changes in air quality standards or emergency responses to wildfire smoke.

Timing of Disease Onset vs. Mortality Reporting:

- **Lack of Onset Data:** The mortality data lacks information on when diseases begin, only recording the year of death. This absence makes it challenging to directly link health outcomes with smoke exposure in the same year.

- **Estimation Challenges:** Diseases like COPD, lung cancer, and heart disease develop over many years, complicating direct correlations with specific environmental events like wildfire smoke. Without onset data, it's difficult to determine if the increased mortality in a given year directly results from smoke events of that year or earlier.

Timeline to Completion

- **November 10:** Data collection and preprocessing complete.

- **November 17:** Exploratory data analysis complete.

- **November 24:** Model development and initial modifications complete.

- **November 25:** Visualization of results and preparation of the presentation complete.

- **December 1:** Validation and refinement of the model complete.

- **December 2:** Final report finalized and submitted.

References

[1] Mississippi State Department of Health. (2022). *MSTAHRS: Mississippi Statistically Automated Health Resource System*. Retrieved from https://mstahrs.msdh.ms.gov/forms/morttable.html