Final Project Report

1. Introduction

Nebraska, often called the Cornhusker State, has agriculture deeply woven into its economic fabric. Production agriculture contributes more than $25 billion to Nebraska's economy each year, with farming and ranching operations spread across 92% of the state's total land area [1]. The state's agricultural prominence is particularly evident in its role as the nation's third-largest corn producer and fourth-largest soybean producer, contributing 12.3% and 7.86% of the national production respectively [2]. With 1 in 4 jobs tied to agriculture and 44,400 farms and ranches operating across the state, analyzing and understanding potential threats of wildfire smoke to agricultural productivity is crucial [3].

The urgency of this analysis is underscored by projections indicating a 50% increase in wildfire incidence globally by 2100 [4]. While immediate effects of smoke on crops may appear minimal, experts warn that the cumulative impact could become more significant over time. In Fig 1.1, we can already note the sharp rise in the number of wildfire incidents occurring in Omaha, NE over time, as indicated by the blue line. The red line in the graph is generated by Meta's Prophet time series model, and shows immense cause for concern. For Nebraska, where agriculture contributes more than 6.5% of the state's gross product [2], this analysis addresses a critical intersection between environmental changes and economic stability. Understanding and addressing the risks posed by these environmental changes is crucial for maintaining economic stability of the state and food security of the country.

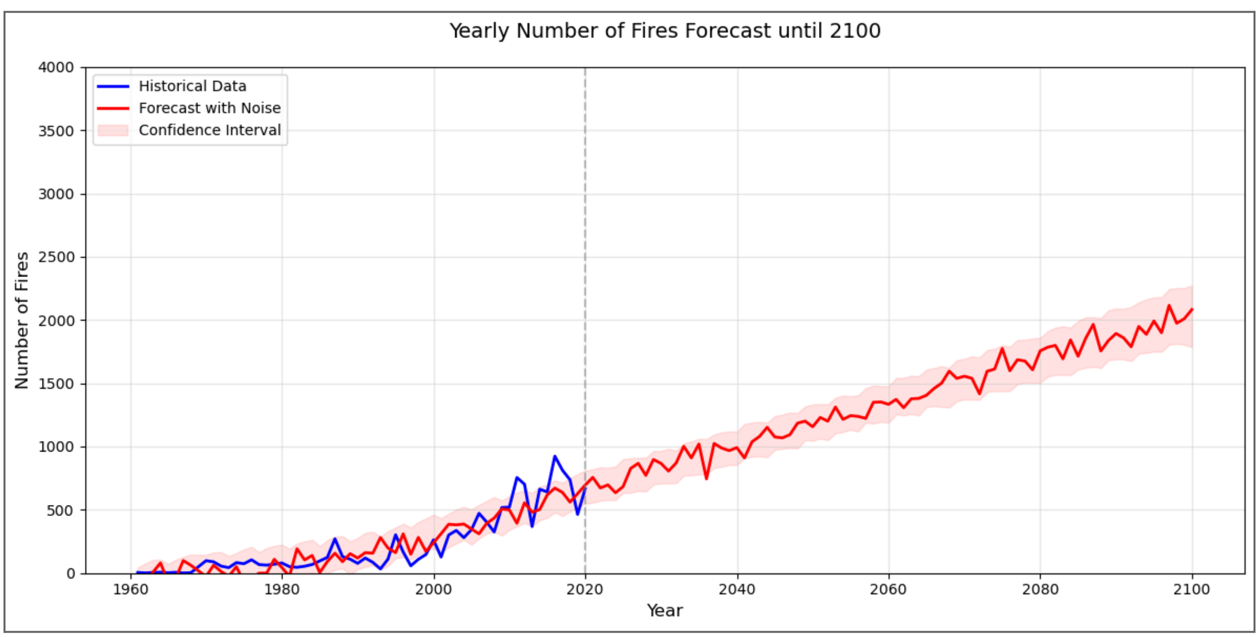

Fig 1.1. Forecasted # of wildfire incidents in Omaha, NE using data from 1960-2020

2. Background

2.1. Omaha’s Wildfire Risk

Nebraska, as a state, ranks #19 for fire risk in the country [5]. Omaha's climate presents a moderate increasing wildfire risk due to several key factors. Majorly, The Big Papillion-Mosquito watershed, which contains Omaha, has experienced drought conditions in 51% of weeks since 2000, with 5% of those weeks reaching Extreme or Exceptional drought levels [6]. These conditions are expected to continue, and even worsen, as the 2024 drought conditions demonstrate. Even now, nearly 75% of Nebraska is in drought conditions and undergoing record-breaking temperatures consistently 10-20°F above seasonal averages [7].

This persistent drought pattern creates conditions conducive to fire spread. In fact, the severity of this drought-fire relationship was demonstrated in 2022, which became Nebraska's second-worst fire season on record with over 250,000 acres burned and historically low soil moisture levels [8]. Rare rainfall events were accompanied by lightning strikes that further exacerbated fire risks. Strong sustained winds throughout the region further compound the problem. These conditions collectively create an environment where wildfires can spread rapidly when ignited, particularly in the urban-wildland interface areas around the city.

2.2. Major Crops of Omaha, NE

Corn and soybeans are Omaha's two major crops, with massive production numbers that place the state among the nation's top agricultural producers:

- Corn Production [2] [12] [13] [14]
  - Nebraska ranks as the nation's 3rd largest corn producer.
  - Estimated yield in 2023 was 1.73 billion bushels.
  - 2024 forecast shows record production of 1.89 billion bushels.
  - Approximately 9.70 million acres harvested for grain.

- Soybean Production [2] [12] [13] [14]
  - Nebraska ranks 4th nationally in soybean production.
  - Estimated yield in 2023 was 267 million bushels.
  - 2024 production forecast is at 310 million bushels.
  - Approximately 5.25 million acres harvested.

The combination of corn and soybean production creates what's known as the "Golden Triangle," where the integration of crop production with livestock and ethanol provides significant economic value at every step of the production chain [14]. Therefore, it can be inferred that a large chunk of the state’s economy, and by association the city of Omaha’s economy, is tied to agricultural production and success.

2.3. *Impact of Wildfire Smoke on Crops*

Agricultural productivity is dependent on myriad factors like soil conditions, rainfall, geographical region, and more. In light of increasing pollution and particulate matter in the environment around the world, several studies have been conducted in the past to examine the effect of smoke and particulate pollution on agricultural productivity. Recent research reveals complex and sometimes counterintuitive effects of wildfire smoke on crop yields. In fact, to quote one article that summarized the effects of the 2023 Canadian wildfires on crop yields - "results are just about as clear as the sky over the Midwest has been lately" [9].

Key and interesting findings from some of these studies are listed as follows:

- Wildfire smoke effects based on incoming sunlight [10]
    - Smoke reflects a portion of incoming sunlight, reducing the amount of light available to plants for growth. Since plants depend on sunlight to carry out photosynthesis, this reduction in light is detrimental to crop productivity.
    - However, at low levels, wildfire smoke can refract sunlight significantly, thereby increasing the diffuse fraction of photosynthetically active radiation (PAR), which can benefit plants by increasing their light use efficiency. Plants with taller, higher leaf area index, and multilayer canopies are likely to benefit more than shorter plants from this.
    - A study conducted in 2018 in the California Central Valley found that total PAR was only reduced by 3.6% on average due to wildfire smoke, while the diffuse fraction increased by over 34%. The predicted effect on corn was a 2.5% increase in photosynthesis, as the positive effect of more diffuse PAR exceeded the negative effect of reduced total PAR.

- Wildfire smoke effects based on smoke pollution [10] [11]
    - Wildfires emit various air pollutants to form ozone when reacting with sunlight.
    - Ozone can cause harm to both corn and soybeans by entering the plant through the stomata, which burns plant tissue during respiration.
    - Other components of wildfire smoke that are known to harm plants include nitrogen dioxide (NO_2) and sulfur dioxide (SO_2).
    - A 2021 study estimated that the presence of ozone and these pollutants in the atmosphere has reduced yields of corn and soybean in the U.S. by roughly 5% over the past 20 years.

There is also very little evidence that wildfire smoke can indirectly impact the quality of crops, specifically the health of the stalk and root of the crop [9]. Clearly, it is very hard to quantify the complex interaction between smoke and agricultural metrics like yield and crop condition. Several factors like amount of atmospheric smoke, time in the crop cycle, etc., play into whether the impact of smoke will be negative or positive.

3. Methodology

3.1. Construction of Smoke Estimate

The smoke estimate calculation involved the following steps:

- Data Acquisition: Fire data was sourced from USGS. The dataset contained historical fire event information including burnt area and geospatial information like coordinate of fire event.

- Data Processing:
  - Converting geographic coordinates to EPSG:3857 (Web Mercator) projection.
  - Filtering fires within a 650-mile radius of Omaha.
  - Converting fire areas from acres to square meters (1 acre = 4046.86 m^2).
  - Calculating Euclidean distance between fire coordinates and Omaha's coordinates.

- Smoke Estimate Formulation: Two different approaches were implemented.
  - Basic Smoke Estimate = Burnt Area / (Distance from Omaha)^2
  - Enhanced Smoke Estimate = Basic Smoke Estimate * Fire Duration

Eventually, the enhanced smoke estimate was discarded due to data inconsistencies in collection of fire start and end dates. The basic smoke estimate formula was chosen as the best estimator because it captured fundamental information like:

- Larger fires produce more smoke, justifying direct proportionality with burnt area.
- The inverse square relationship with distance from center matches how particulate matter could potentially spread from the source.
- Scaled well, as larger fires nearby had maximum impact and vice versa.
- Resulted in a dimensionless quantity which can be used for comparison, like an index.

However, the estimate made several assumptions like:

- Assuming radial spread of smoke from its source.
- Assuming uniform dispersion of smoke.
- Ignoring topography of the region and assuming plain, flat land.
- Assuming static wind patterns and atmospheric conditions.

3.2. Smoke Forecast

To forecast the smoke, two different time-series modelling approaches were tried, namely:

- Exponential Smoothing
- Meta's Prophet Model

Exponential smoothing is a forecasting technique used to predict future values in time series data by assigning exponentially decreasing weights to past observations. This means more recent data points have a greater influence on the forecast compared to older data. It is particularly useful for data with trends and seasonality.

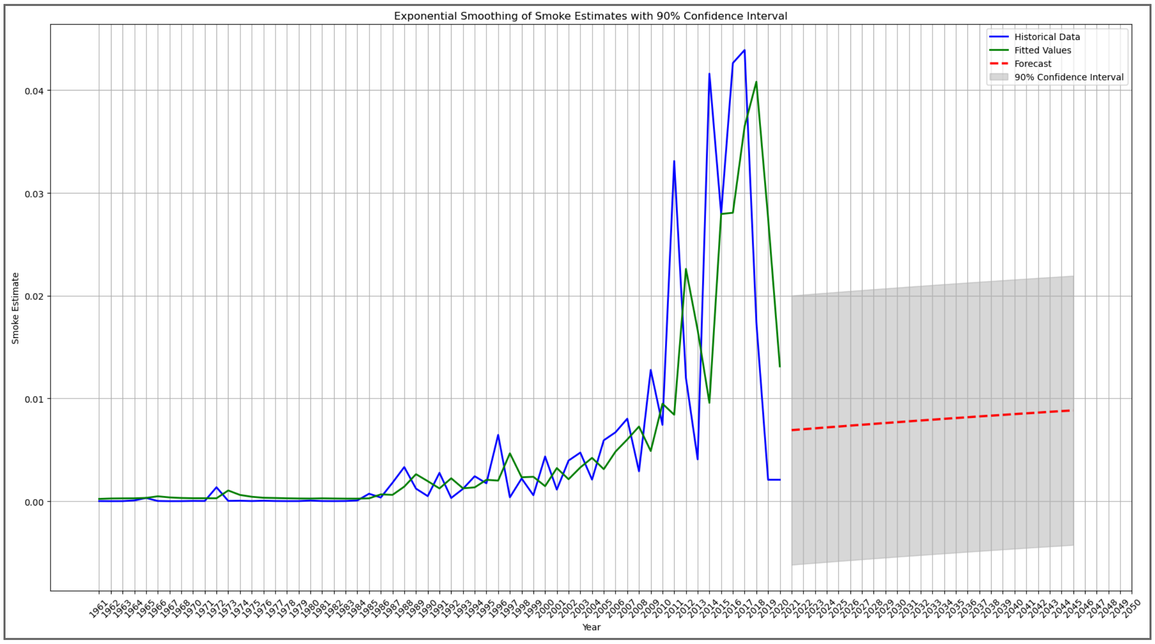

Fig 3.2.1. Forecasted smoke estimates for Omaha, NE using data from 1960-2020 using Exponential smoothing

Meta's Prophet is an open-source time-series forecasting model that uses an additive decomposition approach that breaks down time series into three main components: trend, seasonality, and holiday effects. It excels with data that has strong seasonal patterns and can handle common challenges like missing values, outliers, and trend shifts.

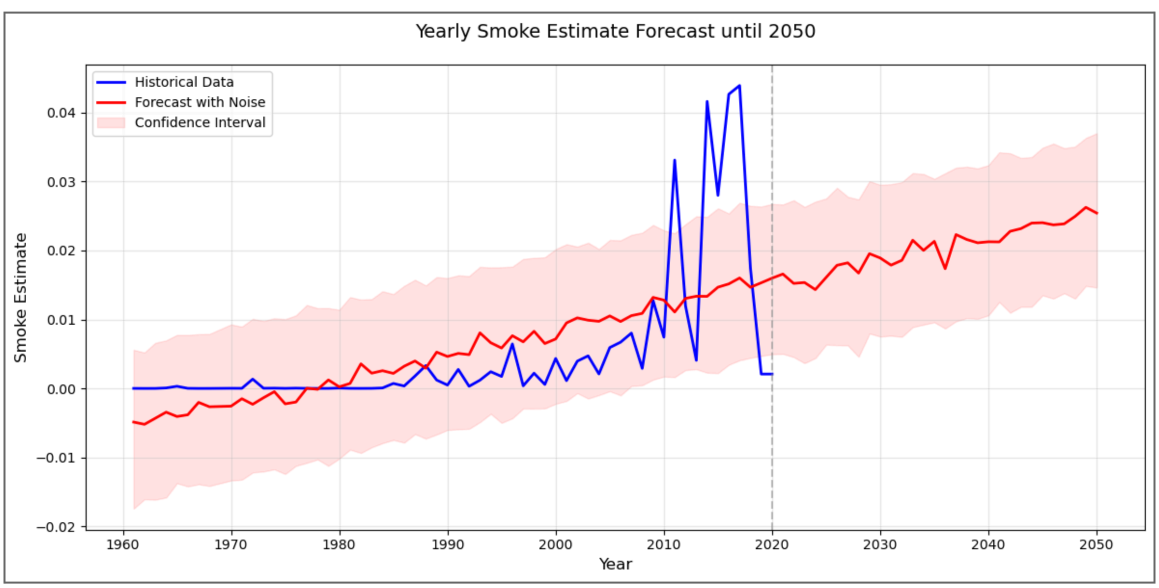

Fig 3.2.2. Forecasted smoke estimates for Omaha, NE using data from 1960-2020 using Meta's Prophet model

Due to Prophet's ability to decompose a time series into its components, it is able to isolate the effects of external regressors and factors. This enables stakeholders to better understand which factors drive the forecast, making the model more interpretable and human-centered. Adding confidence bands around the forecast further reinforces transparency around the certainty of the predictions. The model also offers comprehensive features like component plots and changepoint visualizations, which are useful for analyzing the overarching trends and seasonal patterns separately from influential external factors. This interpretability advantage makes Prophet particularly valuable for public policy and agricultural planning, where clear communication of environmental risks is crucial.

3.3. Relationship between Smoke Estimate and Crop Yields

- Data Acquisition: Crop Yield data was sourced from the USDA National Agricultural Statistics Service (NASS) Quick Stats database.

- Data Processing:
    - Soybean and corn yield data were loaded from CSV files.
    - Data was pivoted to extract yearly yield measurements in bushels/acre.
    - Historical data ranging from 1961-2021 was used.
    - Smoke estimate and agricultural data was merged to form a dataframe with the following columns:
        - Soybean yields (bushels/acre)
        - Corn yields (bushels/acre)
        - Annual smoke estimates
        - Number of fires
        - Total burned acres

- Exploratory Data Analysis:
    - Trend analysis of corn and soybean yield 1960 onwards.
    - Calculated the Pearson correlation coefficient between yields of both crops and the smoke estimate.
    - A Vector Autoregressive (VAR) model was initially developed to model the multiple inter-dependent time series to understand their dynamic relationships and potential predictive power.
    - VAR models allow us to test causality hypotheses, such as whether smoke estimates influence crop yields using Granger-Causality.

The forecasted smoke estimates from 3.2 were then used as external regressors to a Prophet forecasting model to predict crop yields for both corn and soybean.

3.4. Relationship between Smoke Estimate and Crop Conditions

- Data Acquisition: Crop Condition data was sourced from the USDA National Agricultural Statistics Service (NASS) Quick Stats database.

- Data Processing:
  - Soybean and corn condition data were loaded from CSV files.
  - Data was pivoted to extract yearly conditions measured in percentage of crop that was “excellent”, “good”, “fair”, “poor” and “very poor”.
  - Historical data ranging from 1961-2021 was used.

- Exploratory Data Analysis:
  - Trend analysis of corn and soybean conditions 1960 onwards.

An ordinal logistic regression model was utilized to explore the relation between crop conditions and smoke estimate.

4. Findings

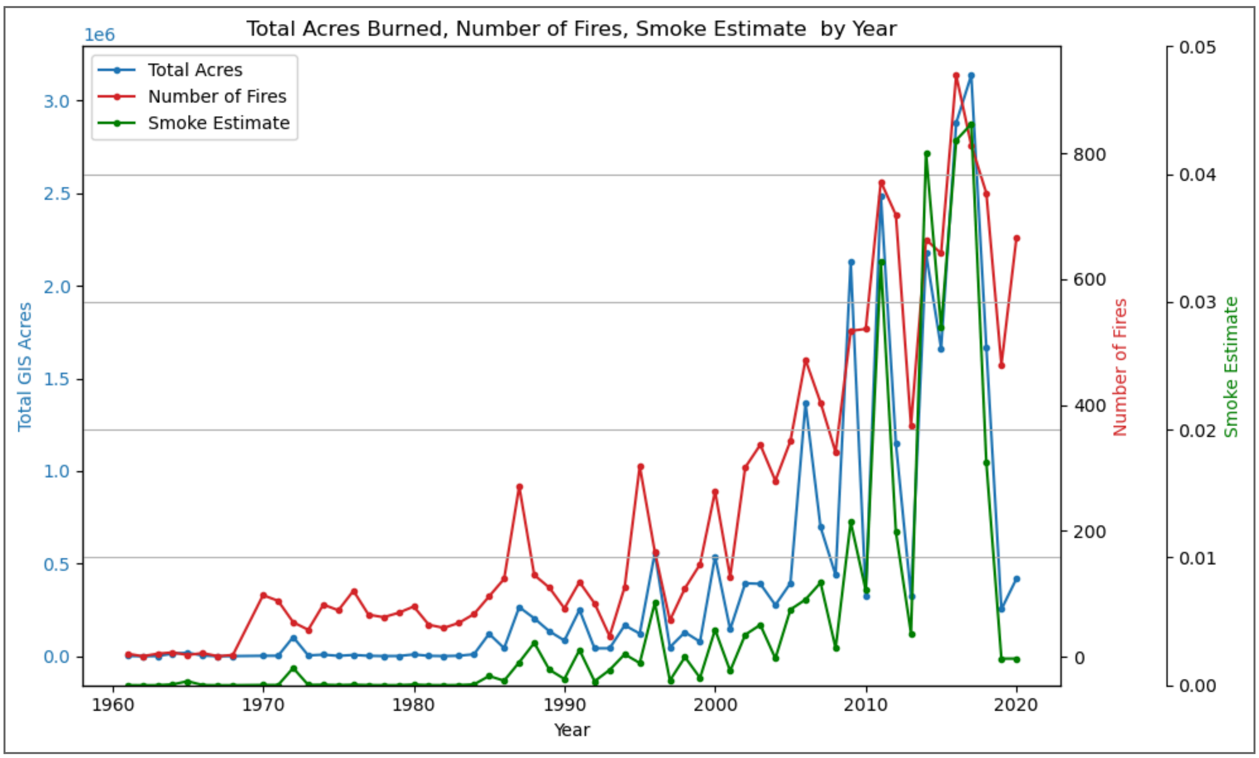

Fig 4.1 Fire data and smoke estimate for Omaha, NE from 1960 onwards.

Fig 4.1. sets the stage for the purpose of our entire analysis. It depicts the sharp increase in the number of fires around the Omaha region, and the amount of area burnt, indicated by the red and blue lines. Fig 4.2. goes a step further and contextualizes the geographical positions of the fires, showing that a sizable portion of fires occur further away from Omaha’s city center.

This essentially implies that while wildfires are becoming more commonplace, they generally occur further away from the city center. This is likely due to the fact that the city is more developed with lesser vegetation that could be susceptible to wildfire risk.

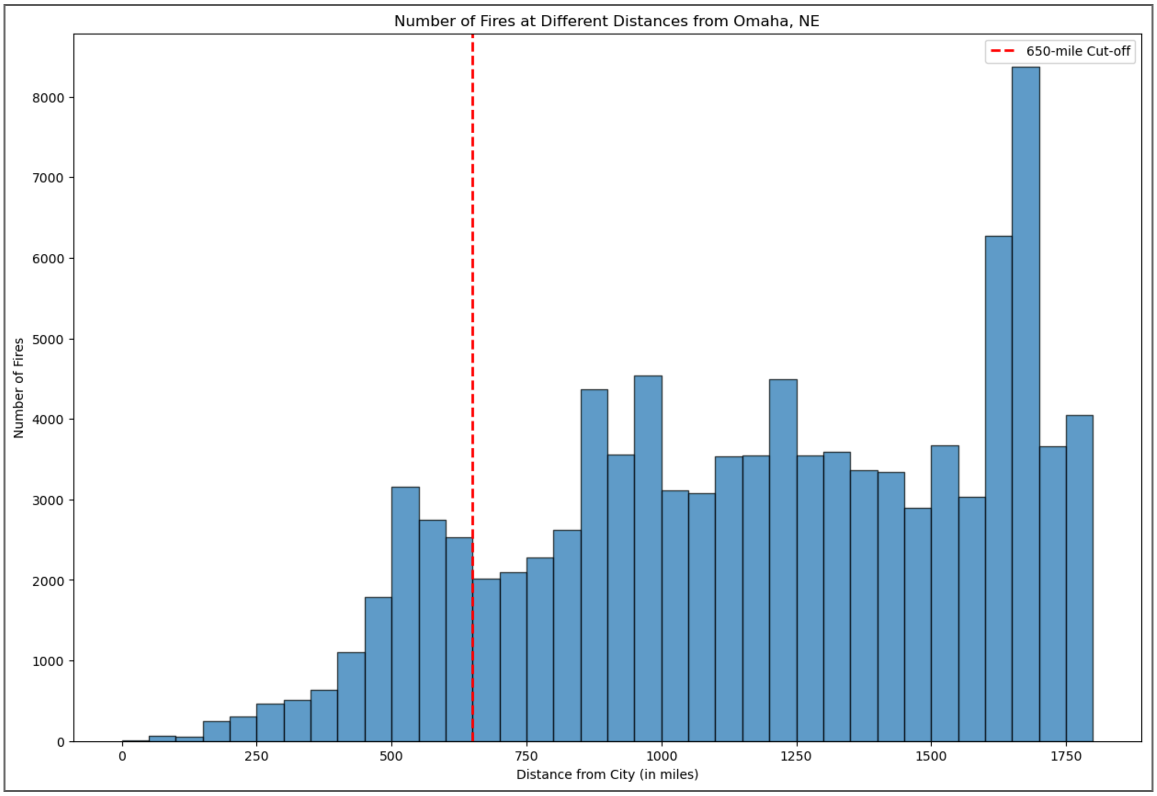

Fig 4.2. Histogram of fire occurrences at different distances from Omaha, NE (city centre)

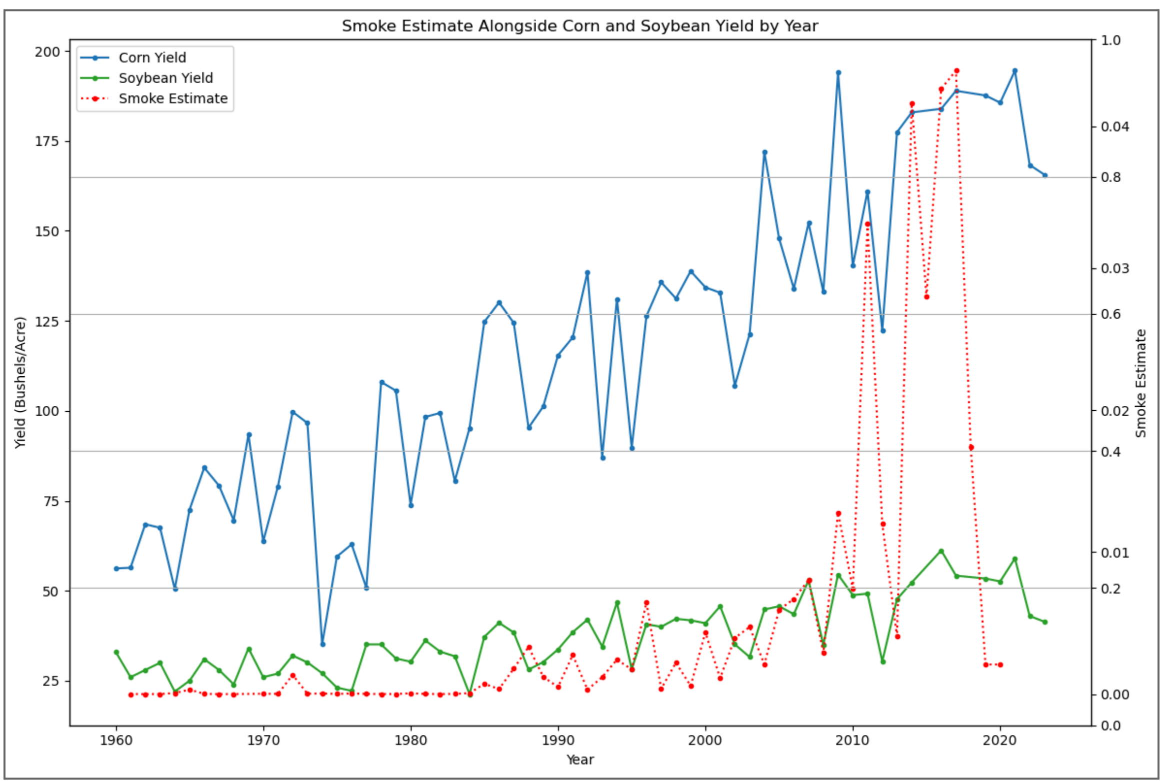

Fig 4.3. Crop yield trends for Corn and Soybean from 1960 onwards, alongside smoke estimate

Fig 4.3. showcases a strong upwards trend exhibited by corn yields over the years. Soybean yields, depicted by the green line, also show a steady growth trajectory, albeit not as dramatic as corn. The up and down nature of the trend lines are likely indicative of the seasonality associated with the crop cycle. Lastly, the dotted red line is present to showcase the smoke estimate, and how it suddenly spikes post 2010.

Since all variables under consideration show a rise over the years, a correlation analysis was run to quantify the relationship between them, as depicted by the matrix in Fig 4.4.. A moderate positive correlation was observed between yields of both crops and the smoke estimate, which warranted further analysis.

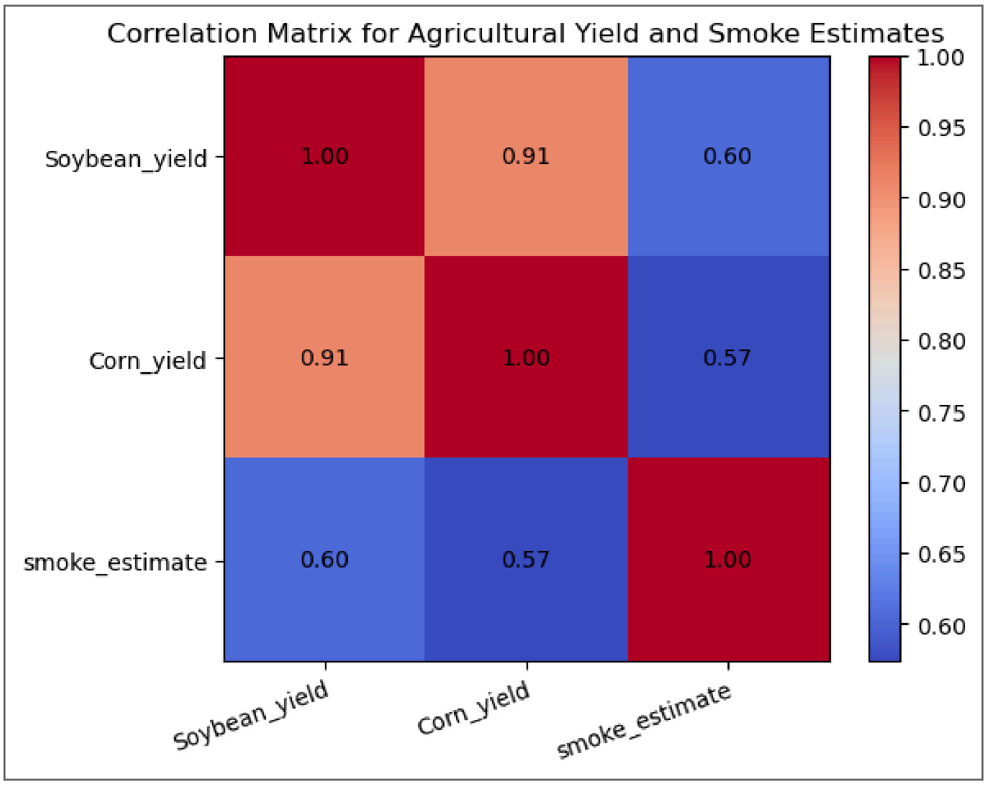

Fig 4.4. Correlation matrix between corn and soybean yields and smoke estimate

A simplistic VAR model was run to understand the relationships between all 3 time series. The results from the model were as follows:

- ADF results implied non-stationarity of the time series at $p > 0.05$.
- Granger causality implied past smoke estimate does not predict future crop yield at 5% significance level.
- There was no strong statistical evidence that smoke estimates significantly influence agricultural yields in Omaha.

Furthermore, when the smoke estimate variable was added as an external regressor to predict corn and soybean yields, it was observed that it contributed positively to the predictions, as depicted in Fig 4.5. and 4.6..

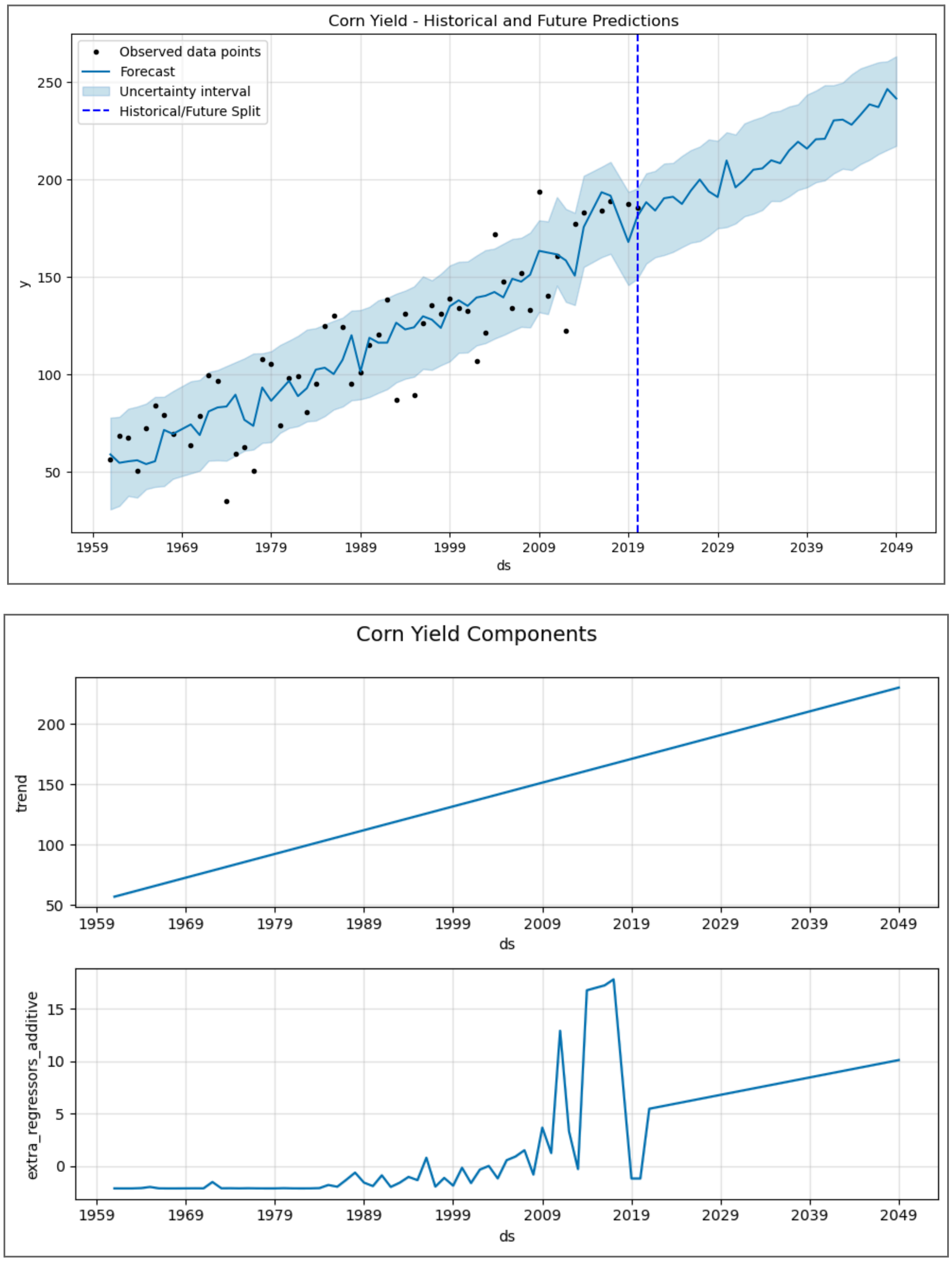

Fig 4.5. (top) Corn yield forecasted using smoke_estimate as external regressor; (bottom) Decomposition of corn yield forecast into trend and effect of smoke_estimate

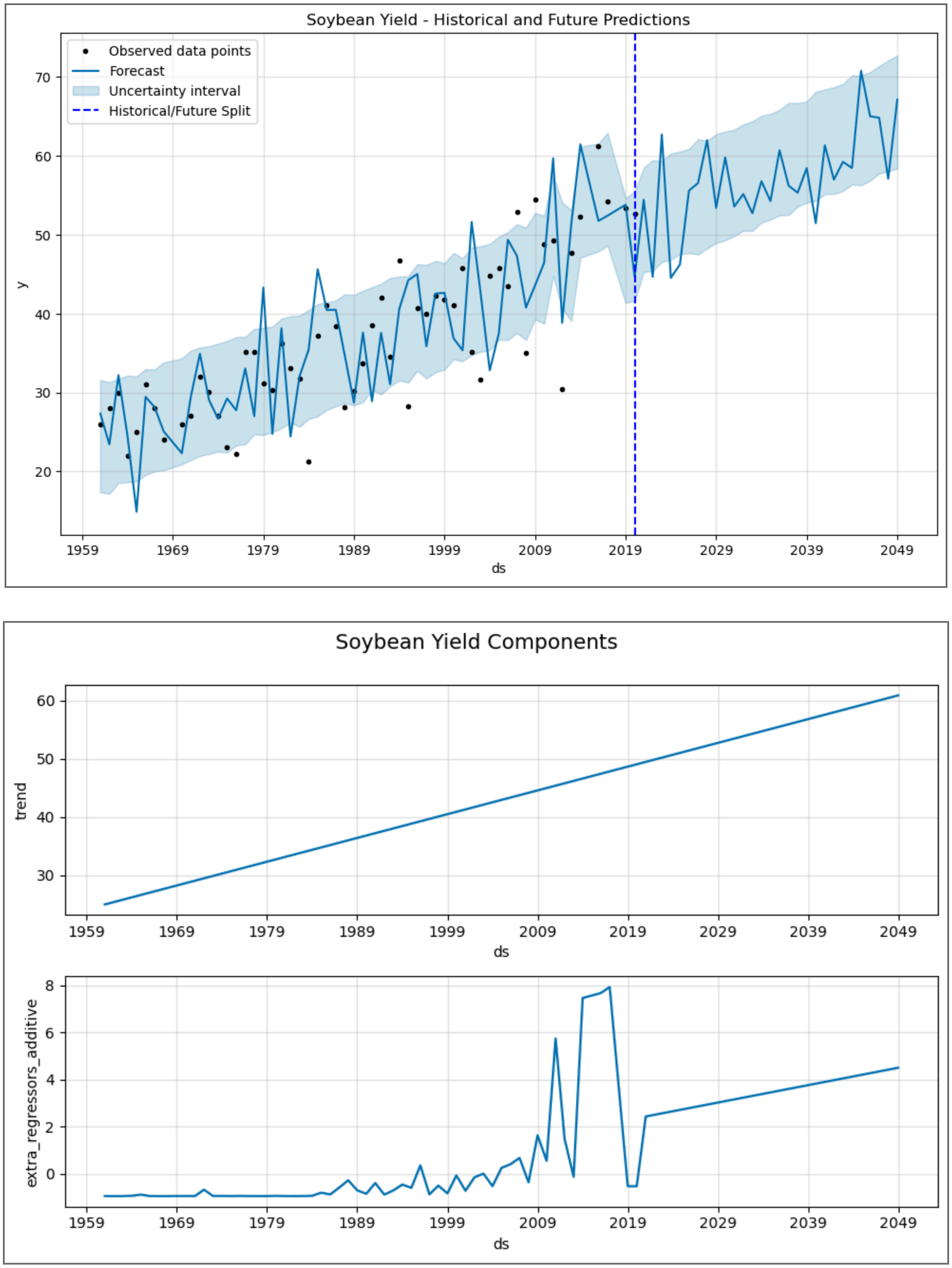

Fig 4.6. (top) Soybean yield forecasted using smoke_estimate as external regressor; (bottom) Decomposition of soybean yield forecast into trend and effect of smoke_estimate

The secondary analysis consisted of examining the impact of smoke estimate on the quality of crops harvested over the years in Omaha. For the purpose of data exploration, the conditions of both corn and soybean crops were plotted as stacked bar charts from 1960 onwards in Fig 4.7.

and 4.8.. In both cases, on average, the majority of the crops were in “good” condition. At initial glance, the quality of corn crops seems to be more consistent over the years than soybean crops. On further inspection however, both crops have performed similarly and reasonably well in the last decade.

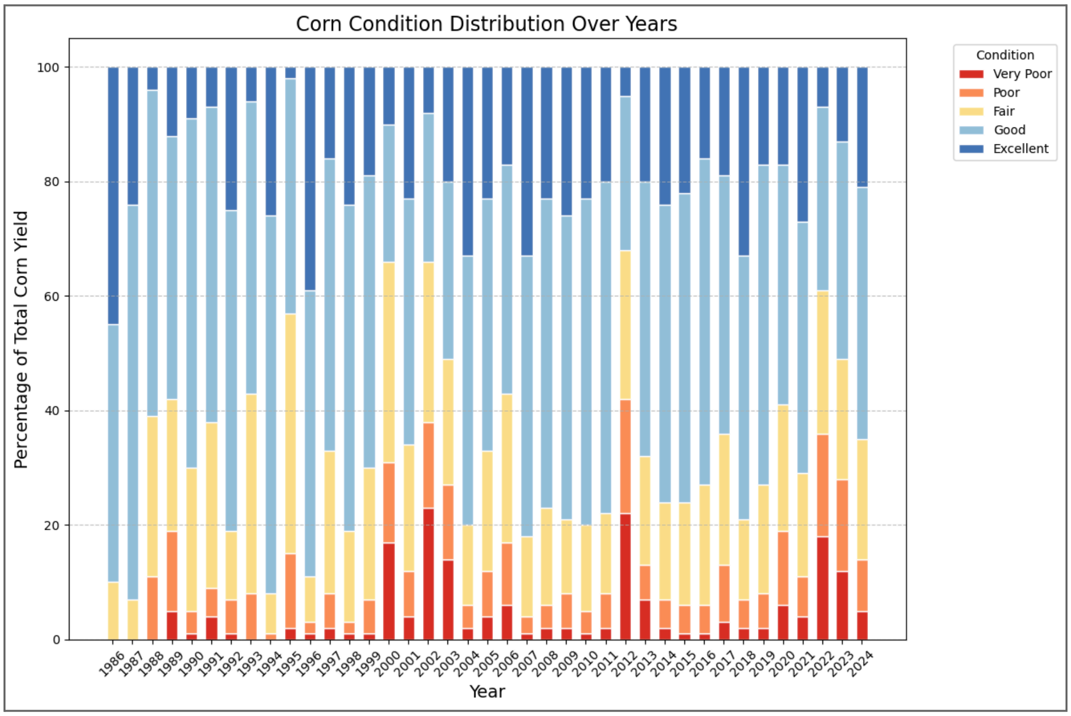

Fig 4.7. Stacked bar charts showcasing the % condition of corn crop from 1986 onwards

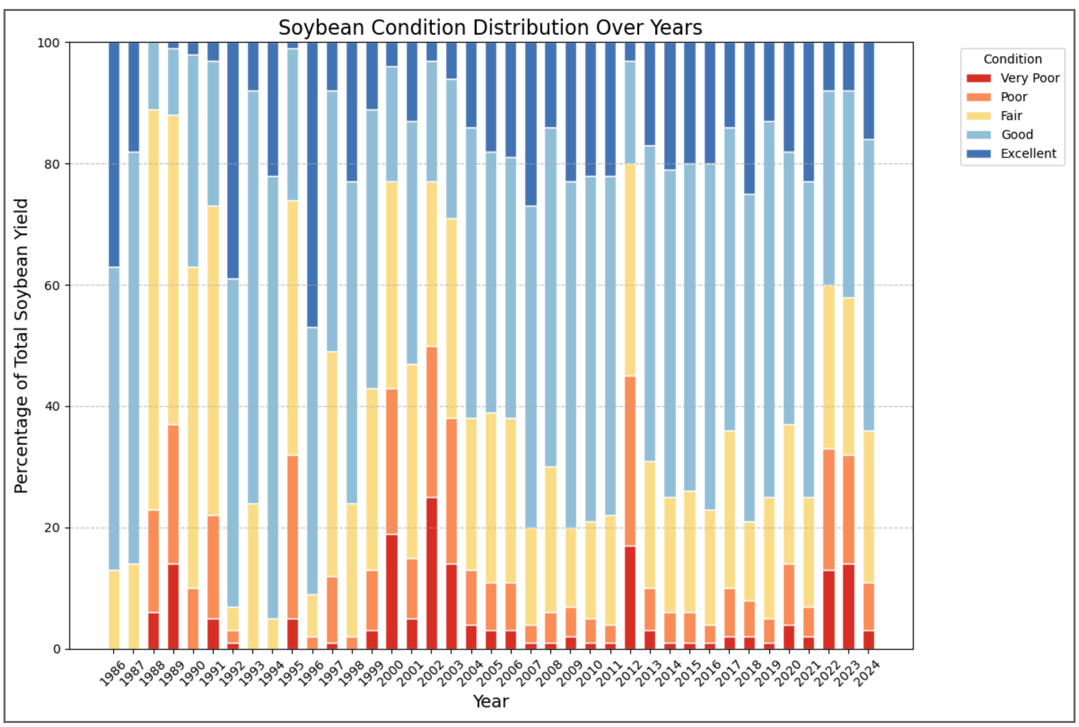

Fig 4.8. Stacked bar charts showcasing the % condition of soybean crop from 1986 onwards

The data was prepared for the ordinal logistic regression model by calculating a weighted condition score where a weight of 1 corresponded to “Very Poor” condition and 5 corresponded to “Excellent” condition. The smoke_estimate was used as a regressor, along with its lagged effects at 2 different levels. It was observed that corn showed stronger lagged effects with a high positive coefficient for the 1 year lag regressor and a small positive coefficient for the 2 year lag variable. Soybean, on the other hand, showed more immediate effects with a positive coefficient for the smoke_estimate and thereby a small negative coefficient for the 1 year lag variable. However, none of these effects were statistically significant at the $p<0.05$ significance level.

There were significant limitations to this approach, so a Granger Causality test was run to check for the effect of smoke on any category of condition. The results for the test were reported for lagged values 1 to 4. When done for corn, high p-values were observed across almost all conditions, indicating that smoke_estimate does not Granger-cause changes in the corn condition categories. This means the past values of smoke_estimate do not provide predictive information for the corn condition percentages. At lag 2, for poor condition of corn, there was some weak evidence of Granger causality (p-values close to 0.05). However, the overall results suggest no strong causal relationship between smoke_estimate and corn condition percentages. For soybeans, the p-values for all conditions were observed to be well above the threshold for all lags. Therefore, the Granger causality tests show no evidence that smoke_estimate Granger-causes soybean condition percentages at any lag.

5. Discussion

The findings of this project hold significant implications for Omaha, Nebraska, and its stakeholders, including the city council, mayor, and residents. The moderate positive correlation between wildfire smoke estimates and crop yields highlights a complex relationship, where smoke can influence agricultural productivity. While the direct effects of smoke are limited compared to broader technological advancements in irrigation technology and large-scale farming, understanding and mitigating potential risks is vital for ensuring the long-term sustainability of Omaha's agricultural economy.

Agriculture supports thousands of livelihoods and contributes significantly to food security. As wildfires become more frequent and severe, even the small, seemingly negligible impacts of smoke can compound over time. For example, while wildfire smoke’s refractive properties can marginally benefit photosynthesis under specific conditions, pollutants such as ozone, NO_2, and SO_2 present significant risks by damaging plant tissues. The observed correlations warrant further investigation into how long-term exposure to increased smoke levels might affect crop health, quality, and overall yields. Without proactive measures, Omaha’s agricultural resilience could be tested in the face of escalating environmental challenges.

5.1. Recommended Actions

To address these findings, the city council and mayor should collaborate with agricultural stakeholders like universities and farmers to create a smoke-resilience action plan within the next 18 to 24 months. Key recommendations include:

- Monitoring and Data Collection: Establish a regional air quality monitoring network to capture real-time data on smoke and pollutants. This would help quantify localized risks and inform timely interventions.

- Policy Initiatives: Introduce subsidies or incentives for farmers to adopt resilient agricultural practices, such as diversifying crops or implementing advanced irrigation systems to counteract drought and smoke-induced stress.

- Collaboration with Researchers: Work with universities and research institutions to explore the nuanced impacts of smoke on crop yields and identify adaptive strategies.

Given the increasing frequency and intensity of wildfires, a concrete plan should be implemented within 5 years. Immediate steps, such as air quality monitoring, can begin within 2 years, while longer-term initiatives, like policy and community programs, can unfold progressively.

5.2. Human-Centered Decisions

Human-centered data science principles played a pivotal role in shaping this project. The goal was to build a clear analysis that empowers lay-individuals to understand the effects of wildfire smoke on the crops that are tied to their livelihoods.

- The methodology, including data sources, processing steps, and limitations, are clearly documented. This transparency allows for peer review and replication of the study.

- The decision to use a basic smoke estimate rather than an enhanced version was guided by the principle of simplicity and accessibility, ensuring the findings could be effectively communicated to non-expert stakeholders.

- By utilizing interpretable models like Meta's Prophet, the analysis prioritized transparency and stakeholder understanding. Features such as component plots and confidence intervals offered clear insights into trends and the impacts of external factors, enabling informed decision-making.

- By framing recommendations for different stakeholders (city council, residents, farmers), the research acknowledges the diverse perspectives and needs within the community.

6. Limitations

This study, while comprehensive in scope, has several limitations that could significantly impact the results. Identifying and acknowledging these limitations is crucial to understanding the reliability and generalizability of the findings, and embody the spirit of transparency.

- Data Limitations
    - Incomplete or Inconsistent Data: The fire duration data used for the enhanced smoke estimate had inconsistencies, leading to its exclusion. This omission might have affected the overall robustness of the smoke estimate.
    - Missing Variables: The analysis did not account for critical environmental factors such as wind direction, precipitation, and topography, which significantly influence smoke dispersion.
    - Geographical Limitations: The crop yield data was acquired at the county level for Douglas county, while the crop condition data was acquired at the state level for Nebraska. This could lead to incorrect estimation of smoke impact on crops.

- Assumptions in Smoke Estimate Formulation
    - Simplistic Dispersion Model: The basic smoke estimate assumed radial and uniform smoke dispersion, ignoring topographical and atmospheric variations. This might have oversimplified the complex dynamics of smoke spread and its interaction with crops.
    - Static Conditions: The model presumed static wind patterns and atmospheric conditions, which may not hold true in real-world scenarios.

- Statistical and Analytical Constraints:
    - Non-Stationarity of Data: The time series data exhibited non-stationarity, potentially violating assumptions of the statistical models used, such as the Vector Autoregressive (VAR) and Granger Causality tests.
    - Lagged Effects Complexity: The ordinal logistic regression model identified lagged effects of smoke on crop conditions but did not sufficiently capture the variability and complexity of these relationships due to limited predictive power.
    - Simplistic Models: The models used were not optimized for capturing non-linear relationships and might have missed some complex dynamics.
    - Temporal Trends Dominated by Technology: The overall upward trend in crop yields was largely due to technological improvements, overshadowing the relatively smaller impact of smoke. This could result in underestimating the potential risks of cumulative smoke exposure over time

By recognizing these limitations, this study lays the foundation for future research to refine methodologies, incorporate additional variables, and expand the geographic and temporal scope. Addressing these issues can enhance the precision and applicability of findings, thereby contributing to more informed decision-making.

7. Conclusion

This study sought to explore the impact of wildfire smoke on agricultural productivity in Omaha, NE, with specific focus on two major crops—corn and soybeans. The research questions aimed to uncover whether wildfire smoke has a quantifiable impact on crop yields and quality, and to understand the broader implications of these findings for stakeholders, including farmers, policymakers, and city officials.

The findings revealed a moderate positive correlation between wildfire smoke estimates and crop yields, though no causal relationship was established at statistically significant levels. This suggests that while smoke might play a role, its impact is overshadowed by technological advancements and other environmental factors. Additionally, the condition of crops showed varying lagged effects with respect to smoke, but these too lacked statistical significance.

This analysis underscores the complexity of interactions between wildfire smoke and agricultural outcomes. While some effects, such as refracted sunlight increasing photosynthetic efficiency, may appear beneficial under specific conditions, the potential long-term risks of pollutants like ozone and nitrogen dioxide must not be underestimated. These findings stress the need for proactive measures to mitigate risks and build agricultural resilience.

By applying human-centered data science principles, this study prioritized transparency, interpretability, and stakeholder relevance. For example, Meta's Prophet model was chosen for its ability to decompose time-series data into clear components, aiding in transparent communication of results. The decision-making process also emphasized simplicity, opting for a basic smoke estimate to ensure accessibility for non-expert stakeholders.

This study contributes to the understanding of human-centered data science by demonstrating how data-driven insights can guide actionable recommendations for diverse stakeholders. It highlights the importance of aligning technical analysis with human and community needs, empowering informed decision-making at the intersection of agriculture and environmental change.

The findings invite further exploration into the nuanced relationships between wildfire smoke and crop productivity, as well as broader climate impacts on agriculture. Future research should aim to address the identified limitations, refine methodologies, and expand the scope of analysis to provide even more robust guidance for policymakers, agricultural communities, and urban planners.

8. References

1. University of Nebraska. (n.d.). *Dr. Martha Mamo*. Retrieved from https://nebraska.edu/nuforne/martha-mamo

2. University of Nebraska-Lincoln. (n.d.). *Nebraska: An agriculture leader in the region and country*. Retrieved from https://cap.unl.edu/news/nebraska-agriculture-leader-region-and-country/
3. Raising Nebraska. (n.d.). *Fast facts about Nebraska agriculture*. Retrieved from https://raisingnebraska.net/fast-facts-about-nebraska-agriculture/
4. Grow Further. (n.d.). *We must focus on fire risk in food security: Experts*. Retrieved from https://www.growfurther.org/we-must-focus-on-fire-risk-in-food-security-experts/
5. ClimateCheck. (n.d.). *Nebraska climate risk report*. Retrieved from https://climatecheck.com/nebraska
6. ClimateCheck. (n.d.). *Omaha, Nebraska climate risk report*. Retrieved from https://climatecheck.com/nebraska/omaha
7. University of Nebraska-Lincoln CropWatch. (2024). *Weekly weather update and outlook: Majority of the state now in drought*. Retrieved from https://cropwatch.unl.edu/2024/weekly-weather-update-and-outlook-majority-state-now-drought
8. Nebraska Forest Service. (n.d.). *Active fire season in Nebraska*. Retrieved from https://nfs.unl.edu/active-season
9. Bayer Crop Science. (n.d.). *Wildfire smoke on crops*. Retrieved from https://www.cropscience.bayer.us/articles/bayer/wildfire-smoke-on-crops
10. Pioneer. (n.d.). *Potential effects of wildfire smoke on crops*. Retrieved from https://www.pioneer.com/us/agronomy/wildfires-crop-yields.html#PotentialEffectsOfSmokeOnCrops_3
11. Purdue University. (2023, August). *Effects of wildfires on commodity crops: Good or bad? Much remains up in the air*. Retrieved from https://ag.purdue.edu/news/2023/08/effects-of-wildfires-on-commodity-crops-can-be-good-or-bad-but-much-remains-up-in-the-air.html
12. University of Nebraska-Lincoln CropWatch. (2024). *Nebraska 2023 annual crop production summary*. Retrieved from https://cropwatch.unl.edu/2024/nebraska-2023-annual-crop-production-summary
13. University of Nebraska-Lincoln CropWatch. (2024). *Nebraska crop production report, September 2024*. Retrieved from https://cropwatch.unl.edu/2024/nebraska-crop-production-report-september-2024
14. Nebraska Department of Agriculture. (n.d.). *Nebraska agriculture facts brochure*. Retrieved from https://nda.nebraska.gov/publications/ne_ag_facts_brochure.pdf

9. Data Sources

1. Fire Data (link) : Acquired from U.S. Geological Survey (USGS). This dataset contains historical fire event information, including burnt areas and geospatial information.
2. Corn Yield Data (link), Soybean Yield Data (link), Corn Condition Data (link), Soybean Condition Data (link) : Acquired from USDA QuickStats Database (link)