DATA 512 Course Project: Extension Plan

Impact of Wildfire Smoke on Agriculture in Rapid City, SD

Aditi Shrivastava

https://github.com/ad-iti/rapid-city-wildfire-smoke/tree/main

Introduction

Agriculture acts as the backbone of the state economy of South Dakota, unsurprisingly standing as the #1 industry in South Dakota and the #2 industry in the area surrounding Rapid City, and contributing over $25 billion to the economy annually [1]. With an increase from $1,781 million in 1997 to halfway over $7000 million in 2022 as reported by the BEA, the growth trajectory of agriculture in South Dakota is undoubtedly notable, clearly acting as a driver for the state's economy [2]. The agricultural sector's resilience solidifies its role as a key source of growth, acting not only as a huge economic driver, but also embodying a way of life for the farmers and ranchers working the land. And within Rapid City itself, agriculture remains the common thread of "linking the citizens, businesses and communities" all across the city [1]. Clearly, Rapid City's vast agricultural lands contribute significantly to the nation's economy, food, culture, and livelihood, making it crucial to examine the potential threats posed by deteriorating air quality.

As the vital agriculture industry continues to grow, contributing 129,753 jobs and $11.6 billion in household income to South Dakotans in 2021, it faces complex challenges in Rapid City [7]. Historically, the city has been known to struggle with a "history of poor air quality conditions" originating from particulate matters such as PM10 and PM2.5, and exacerbated by high winds and local topography [3]. Within the heart of this problem lies the exponential—nearly 900%— increase in wildfire acreage burned within 1,250 miles of Rapid City over the past 50 years, which acts as a damaging air pollutant with implications for both human health and plant growth [4]. And with an increase in wildfires comes a stark increase in wildfire smoke, which– aside from its impacts on businesses, construction, development– is highly detrimental to crop yields and threatens to disrupt the core of Rapid City's economy. This wildfire smoke leads to "reduced sunlight intensity, increased sunlight diffusion, and increased ozone (O_3) levels", all of which researchers predict to be extremely harmful to the state of agriculture in western South Dakota [5]. As per a case study by Pioneer Hi Bred Intl., "ample sunlight is critical for maximizing plant photosynthesis and crop yield, and lower than normal solar radiation during grain fill can be detrimental" [6]. Given this, we can infer that the smoke's influence on sunlight can act as a "hazy cloud cover", potentially diminishing the light essential for photosynthesis, and posing a direct threat to crop productivity [5]. Moreover, a surge in wildfires also poses a physical risk to crops, as flames and smoke can directly damage vegetation.

Related Work & Hypothesis

In the context of Rapid City's agricultural landscape, with nearly 600 working farms, over 1 million acres of farmland, and over 220,00 acres of cropland in the surrounding county alone, the stakes are high [1]. Preserving air quality is not only an environmental responsibility but a critical economic safeguard for the thriving agricultural sector. The potential consequences of wildfire smoke and its release of air pollutants undoubtedly compromise the industry's long-term sustainability and its economic vitality, urging a comprehensive understanding and mitigation of the connections between air quality, wildfire smoke, and crop productivity.

To this, I turn to the agricultural total factor productivity (TFP) as a pivotal metric. This statistic, coined and maintained by the USDA, acts as a comprehensive measure into the well-being and efficiency of local agriculture by quantifying how various inputs, such as labor, capital, and land, translate into agricultural output. Specifically, "to compute the agricultural TFP, the research team in the US Department of Agriculture collected the data of the quantities of 200 agricultural products (including 162 crops, 30 animal products, and 8 aquaculture products) and 6 categories of production inputs (labor, land, farm machinery, livestock, inorganic fertilizers, and animal feed) [9]. TFP as an analytical tool works to signify the overall efficiency and progress within agriculture, and is crucial for capturing changes in the quality of the agricultural sector over time. A positive trend in TFP correlates with advancements in technology or management practices, contributing to increased productivity, while the opposite may evidence challenges within the sector [8].

To contextualize this metric within the scope of my research paper, I propose an in-depth exploration of the potential future impacts of the increase in wildfire smoke on the state of agriculture and farming in Rapid City, South Dakota. More specifically, I plan to examine TFP trends over time in Rapid City's agricultural landscape, looking to the research study *Air pollution as a substantial threat to the improvement of agricultural total factor productivity* by Daxin Dong and Jiaxin Wang as my primary guiding model [9]. Building off this study, I will look towards the AQI index, which encompasses particulate matter and ozone pollution statistics such as PM2.5

and O_3, as well as wildfire smoke emission metadata, to create a predictive model for the agricultural TFP of Rapid City until the year 2050. This personalized approach will allow for a more targeted investigation of Rapid City, exploring correlations between fluctuations in agricultural TFP and instances of poor air quality caused by wildfire smoke, ultimately revealing how wildfires may influence the agriculture sector in Rapid City. I hypothesize that wildfire smoke in Rapid City will have a significantly negative relationship with agricultural TFP (model variables are explicitly detailed in the *Methods* section).

My overall goals of this exploration are to quantify and predict how the future of the agriculture industry in Rapid City, South Dakota will hold up in the face of environmental challenges posed by wildfire smoke. I hope to use the findings from this study to best inform and warn the city policymakers and residents of Rapid City of whether, and to what extent, air pollutants released by wildfire smoke affect the city's agriculture. More specifically, this report details the extent to which wildfire smoke might affect the future Agricultural TFP, and what exactly the discrepancy, if any, might be between the expected AgTFP and the actual TFP in the year 2050.

Methods & Data

As previously mentioned, I will be essentially reproducing the research performed by Daxin Dong and Jiaxin Wang on a smaller scale, localized to Rapid City, South Dakota. In this study, the authors establish agricultural TFP (*AgriculturalTFP*) to be the primary dependent variable, with *AQI*, *Fire Distance*, and *Fire Acreage* as the explanatory variables in a random forest regression model. Dong and Wang pinpoint the PM2.5 and O_3 pollutants to be especially pertinent to this model in light of the extensive potential damage to agriculture that has been reported in the past by Das et al., Sillmann et al., and others [9][12][13]. Based on this, I identify the aforementioned three variables to be most pertinent to explaining to *AgriculturalTFP*, source all necessary data, and design the following model.

Variable	**Description**
Y = *AgriculturalTFP*	Agricultural Total Factor Productivity (TFP) Index
X_1 = *AQI*	Air Quality Index
X_2 = *WildfireAcreage*	Average size of wildfires (acres)
X_3 = *WildfireDistance*	Average wildfire distance from Rapid City, SD (miles)

The wildfire data is primarily sourced from the *Combined wildland fire datasets for the United States and certain territories, 1800s-Present (combined wildland fire polygons)* dataset, which was collected and aggregated by the US Geological Survey [10]. The dataset is relatively well documented, fire polygons are available in ArcGIS and GeoJSON formats. For this exploration, I specifically rely upon the large .JSON formatted file, which can be found in the combined .ZIP file on the website. The data is filtered to include only wildfires taking place during or after 1963, and within 1250 miles of Rapid City, South Dakota. From this, X_2 = *WildfireAcreage* and X_3 = *WildfireDistance* are obtained.

Next, I request data from the US Environmental Protection Agency (EPA) Air Quality Service (AQS) API, which is a historical API that provides data on the Air Quality Index (AQI) statistic [11]. The EPA reports that they only started broad based monitoring with standardized quality assurance procedures in the 1980's. Some additional information on the Air Quality System can be found in the EPA FAQ on the system. I pull this data to extract the AQI indices produced by the EPA on a year to year basis.

Finally, the agricultural FTP data is acquired to serve as training data for the predictive model from the USDA [8].

I then construct the models with the specific following hypotheses in mind:

1. An increase in AQI should lead to a lower AgTFP based on the fact that it is calculated using the following pollutant levels, and is thereby harmful to agriculture and crops, as detailed in the Related Work section.
2. A fire that is larger in size, as well as closer in distance to Rapid City, should lead to a lower AgTFP.

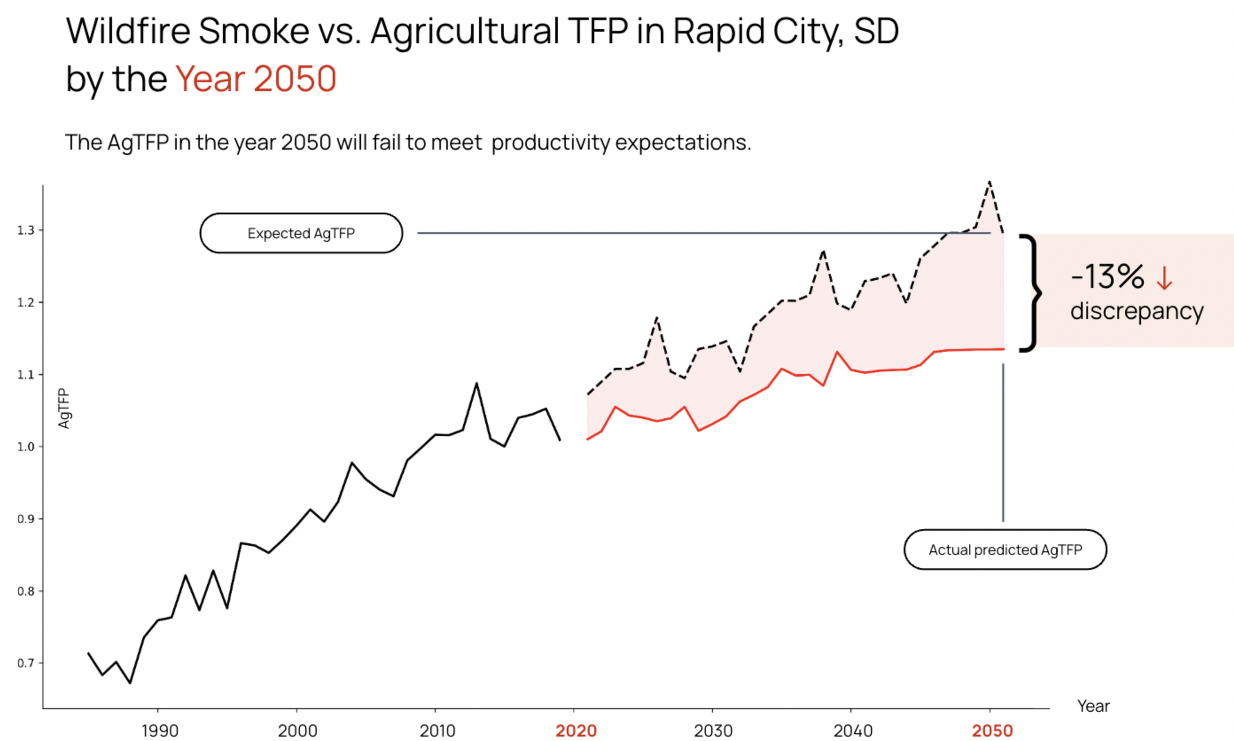
With this, I begin by forecasting the AgTFP into the year 2050 using an Autoregressive Integrated Moving Average (ARIMA) model. This model involves analyzing observations and relationships observed between time and AgTFP before the year 2020, essentially assuming the persistence of the same relationship and seasonality. In other words, the output of this first forecast model would describe what the AgTFP is expected to be in the year 2050, assuming that it will continue to hold the same trends that it has since 1985. To do this, the time series data undergoes decomposition into trend, seasonality, and residuals using the seasonal_decompose function. An ARIMA model is fitted using the Seasonal AutoRegressive Integrated Moving Average with eXogenous regressors (SARIMAX) class, with order and seasonal_order parameters determined through experimentation. This produces the final anticipated trajectory of AgTFP in the year 2050.

In tandem with the forecast model, I also wish to predict what the AgTFP might actually be in the year 2050 given the observed changes in fire acreage, fire distance, and air quality index (AQI). To do this, I first extrapolate the independent variables into the future, in order to serve as training data for the future model. Employing standard techniques, I extrapolate each of X_1 = *AQI,* X_3 = *WildfireAcreage,* and X_4 = *WildfireDistance* roughly 30 years into the future. For fire distances, a random uniform sample of existing data is employed, given by the absence of a discernible relationship between fire distance and time, as revealed during exploratory data analysis. For fire acreages and the Air Quality Index (AQI), I chose to use linear extrapolation given the discernable direct relationship between each of these variables and time. The generated future data is then utilized to predict AgTFP.

The goal with both of these models is to stay as true as possible to the various phenomena associated with wildfires while prioritizing the potential impacts of the study on individuals and communities in Rapid City, South Dakota. To mitigate biases and promote equity, both aforementioned models use comprehensive datasets that encompass the diverse communities within Rapid City. It is important to take an approach that adapts to underlying patterns without perpetuating biases, and I ensure this by not only thoroughly scrutinizing all data before using it, but also researching the nationally-recognized organizations that the data originates from. Since these models detect and account for disparities in the impacts of wildfires and air quality on different communities, they also therefore serve to explicitly encompass the potential differential effects on vulnerable populations.

Results & Implications

Below is the titular figure produced, with the forecast model in black and the regression model in red.



The findings of this exploration reveal a concerning disparity between the expected Agricultural Total Factor Productivity (AgTFP) in the year 2050 and the predicted AgTFP (which considers the increase in wildfire smoke and particulate matter). The expected AgTFP, estimated at approximately 1.3, stands in stark contrast to the predicted AgTFP of around 1.13. In other words, the AgTFP in the year 2050 will fail to meet productivity expectations. As such, this discrepancy confirms the hypothesized detrimental impact of escalating wildfire smoke on the agricultural sector in Rapid City, South Dakota. This more specifically implies that an increase in wildfire acreage, a decrease in wildfire distances, and an increase in the air quality index (AQI) all lead to a decrease in the Agricultural TFP.

The importance of these findings lies in their direct implications for the city's agriculture sector. As discussed earlier, a reduction in AgTFP could indicate potential reductions in crop yields, economic struggles for local farmers, and disruptions in the food supply chain that feeds Rapid City, SD. Furthermore, the broader community may experience negative repercussions, including increased unemployment and a strain on resources. These outcomes not only threaten the economic vitality of the city of Rapid City, but also pose risks to the well-being of its residents and the sustainability of the local agricultural way of life. Moreover, considering the current status of agriculture as the #1 industry in South Dakota and the #2 industry in the area surrounding Rapid City, the identified decline in AgTFP could have far-reaching consequences for the economic resilience of the region [1]. The growth trajectory of agriculture, as evidenced by the increase from $1,781 million in 1997 to over $7,000 million in 2022, reflects its significant role as a key driver of the state's economy, and this projected reduction in AgTFP could therefore severely impact all lives and business in the area by taking away billions of dollars and thousands of jobs [2].

In light of these findings, I believe that the city council members should consider discussing these findings, and taking immediate action to attempt to lower this impact via both short-term and long-term options. This could include investments in advanced air quality management systems, early warning systems for wildfires, and sustainable agricultural practices. Additionally, collaboration with environmental agencies, community stakeholders, and researchers could certainly help to confirm or further investigate the findings in this paper.

Unknowns & Limitations

In terms of data scientists hoping to reproduce the results, I find it necessary to highlight the following. The forecast model (ARIMA) and predictive model (RandomForestRegressor) produced in this study are all highly uncertain and likely not exactly representative of the next 30 years. While I took measures to ensure best data science practices in terms of modeling and data collection, it must be noted that models of this nature are always prone to some degree of uncertainty, and their results should therefore be taken lightly. Also, data extrapolation always holds some level of risk. The three variables used in the predictive ensemble model were all extrapolated nearly 30 years into the future to then predict the associated AgTFP, which could never be 100% representative of what the true values of these variables will be. Next, there is a varying level of granularity in the data sources used in this exploration, with the AgTFP metrics existing only at the state (South Dakota) level, rather than the city level. This of course implies that the results are likely more general to the larger area surrounding Rapid City. Finally, it's important to consider the difference in the type of smoke produced by different types of wildfires (such as prescribed fires), which was not handled in this exploration.

Conclusion

In Rapid City, South Dakota, the agricultural sector's resilience solidifies its role as a key source of growth, acting not only as a huge economic driver, but also embodying a way of life for the farmers and ranchers working the land. What threatens to disrupt this economy is the stark (over 900%) increase in wildfires and wildfire smoke since 1960 [4]. The Agricultural Total Factor Productivity (TFP) represents a critical metric in assessing the overall efficiency of agricultural production, encompassing various inputs such as labor, capital, and technology. In the context of this study, I focus on understanding how fluctuations in TFP, induced by exposure to wildfire smoke, can impact the agricultural landscape of Rapid City. Wildfire smoke poses potential threats to crop health, soil quality, and the overall productivity of the agricultural sector. Thus, a lowered TFP could lead to reduced yields, economic shortages for local farmers, and issues with the food supply chain that feeds this city. Moreover, the community may experience repercussions such as increased unemployment and a strain on resources. Understanding and quantifying these effects is pivotal for creating informed strategies to enhance the resilience of Rapid City's agricultural sector in the face of environmental challenges such as wildfire smoke.

In this report, I comprehensively analyze how wildfire smoke, an increasingly pervasive environmental factor, influences the efficiency and output of the agricultural sector in the city of Rapid City, South Dakota. Specifically, I forecast the Agricultural TFP in Rapid City to the year 2050 based on a number of wildfire smoke related factors, and find that the agriculture sector will fail to meet productivity expectations in the year 2050 if wildfire smoke continues to increase at the same pace.

Human-centered data science principles play a primary role in informing my decision-making throughout this project– the goal is always to prioritize the well-being and privacy of individuals who are being affected by a study such as this. I believe that transparency and open science practices facilitated the reproducibility of the results produced in this paper, and hopefully allow for scrutiny by the broader scientific community. In summary, this exploration offers readers a unique perspective on human-centered data science by finding a balance between ethical considerations and the practical implications of uncovering new insights from data. I hope to serve as a model for how data science can be used not only for analytical purposes, but also for meaningful contributions to the social and economic well-being of communities such as Rapid City, South Dakota.

Works Cited

[1] "Agriculture in Rapid City South Dakota." Town Square Publications, 6 Feb. 2019, https://www.townsquarepublications.com/agriculture-in-rapid-city-sd/ .

[2] "GDP by County, Metro, and Other Areas." GDP by County, Metro, and Other Areas | U.S. Bureau of Economic Analysis (BEA), U.S. Department of Commerce, https://www.bea.gov/data/gdp/gdp-county-metro-and-other-areas.

[3] "Air Quality General Information." Rapid City Government Website, City of Rapid City, https://www.rcgov.org/departments/community-planning-development/air-quality/air-quality-general-information-362.html#:~:text=Local%20topography%2C%20weather%20conditions%2C%20and,periods%20of%20strong%20temperature%20inversions.

[4] Shrivastava, Aditi. DATA 512 Course Project Part 1 Findings. https://github.com/ad-iti/data512-project-part1/tree/main

[5] Jeschke, Mark. "Does Smoke from Wildfires Affect Crop Yields?" No-Till Farmer, 29 June 2023, https://www.no-tillfarmer.com/articles/12574-does-smoke-from-wildfires-affect-crop-yields.

[6] Pioneer. "Is Smoke from Wildfires Affecting Crop Yields?" *Is Smoke from Wildfires Affecting Crop Yields? | Pioneer® Seeds*, Pioneer Hi Bred Intl., 29 Aug. 2023, https://www.pioneer.com/us/agronomy/wildfires-crop-yields.html.

[7] "2021 Economic Contribution Study of South Dakota Agriculture, Ethanol and Forestry ." South Dakota Department of Agriculture & Natural Resources, Decision Innovation Solutions, July 2021, https://danr.sd.gov/AboutDANR/docs/2021AgEcStudy.pdf.

Data Source #1

[8] "International Agricultural Productivity" USDA ERS, U.S. Department of Agriculture, https://www.ers.usda.gov/data-products/international-agricultural-productivity/.

[9] Daxin Dong, Jiaxin Wang. *Air pollution as a substantial threat to the improvement of agricultural total factor productivity: Global evidence*. Environment International, Volume 173, March 2023, https://doi.org/10.1016/j.envint.2023.107842.

Data Source #2

[10] "Combined Wildland Fire Datasets for the United States and Certain Territories, 1800s-Present (Combined Wildland Fire Polygons)." United States Geological Science, ScienceBase, 8 Dec. 2021, https://www.sciencebase.gov/catalog/item/61aa537dd34eb622f699df81.

Data Source #3

[11] "Air Quality System (AQS) API." United States Environmental Protection Agency, https://aqs.epa.gov/aqsweb/documents/data_api.html.

[12] Sillmann, Jana, et al. "Combined Impacts of Climate and Air Pollution on Human Health and Agricultural Productivity." Environmental Research Letters, IOPScience, 3 Sept. 2021, https://www.iopscience.iop.org/article/10.1088/1748-9326/ac1df8.

[13] Das, Sujit, et al. "Particulate Matter Pollution and Global Agricultural Productivity." SpringerLink, Springer International Publishing, 1 Jan. 1970, https://link.springer.com/chapter/10.1007/978-3-030-63249-6_4.

[14] "Federal Reserve Economic Data: Rapid City." Federal Reserve of Economic Data, St. Louis Fed, https://fred.stlouisfed.org/searchresults/?st=rapid+city. Accessed 16 Nov. 2023.

[15] "Weather Averages Rapid City, South Dakota." U.S. Climate Data, https://www.usclimatedata.com/climate/rapid-city/south-dakota/united-states/ussd0283. Accessed 16 Nov. 2023.