HIGH SCHOOL EDITION
Journal of Student Research

Volume 13 Issue 4 (2024)

Predicting International Commodity Prices and Land Usage Towards Reducing Agricultural Emissions

Sahir Gupta[1] and Landon Butler[#]

[1]Mission San Jose High School, USA
[#]Advisor

ABSTRACT

Our long-term goal for this research project was to analyze food insecurity and determine if we could predict the emissions and pricing of different food commodities. Towards this goal, we needed to understand the changes in food production, land use, efficiency, and pricing of these commodities in a sample of different countries. Given the rampant food insecurity heavily influenced by climate change, aiding in reducing the impacts of climate change could be a crucial step towards alleviating food insecurity. Our approach to answering this question was to first gather and analyze data and then use an univariate ARIMA model for time series to further develop our hypothesis and figure out if it was possible to predict future trends. Despite not including external factors like world conflicts, inflations, weather, and trade, the univariate ARIMA is able to predict the prices, production, and efficiency well on some commodities for some countries. Our major conclusions were that our research could be used to make cursory predictions for future values of the food prices and agricultural emissions.

Introduction

Global warming is increasing at a rapid pace and our agricultural practices are a major contributor to these emissions. While people will always need these food commodities and associated emissions will always be present in some form, through gaining a better understanding how various countries' food commodities affect emissions, we can strive to make agricultural practices more environmentally friendly. Agriculture plays a significant role in contributing to climate change through practices such as deforestation, methane emissions from livestock, and the use of synthetic fertilizers, accounting for 10% of the total greenhouse gas emissions [1]. To make agricultural processes more environmentally friendly, sustainable farming practices can be implemented, including more efficient land and resource usage.

Another major problem that many people face is food insecurity with some countries like Sub-Saharan Africa having "people who are food insecure at 51 percent" [2]. According to Nature.com, "food price inflation, which is estimated to outpace growth rates of real per capita Gross Domestic Product (GDP)" [2]. Through limiting the inflation of food prices, we may be able to combat food insecurity. Combining sustainable agricultural practices with efforts to enhance food security can contribute to a more resilient and environmentally-friendly food system.

Background

To achieve the goals of sustainable agriculture and address food insecurity, it is crucial to understand trends in production, land use, efficiency, and economic factors for key commodities like maize, wheat, and rice. Our investigation focuses on representative countries China, India, Ukraine, Peru, and Colombia. These were chosen to provide insights into diverse geographical, economic, and agricultural contexts. Analyzing the trajectories of production and land use can help identify areas where sustainable practices can be implemented. This approach allows for strategies that account for the specific challenges and opportunities, contributing to a more globally sustainable and secure food system.

In the remaining sections of this report, we outline our research regarding combating food insecurity and agricultural emissions. We discuss our data collection procedures and data cleaning effort. Following this, we dive into an exploration of our time series prediction model. We then discuss the successes and shortcomings observed in these models concerning their predictive capabilities across various aspects such as production, land use, efficiency, and prices. By thoroughly examining both the achievements and limitations, we aim to provide a balanced perspective on the model's performance and how it can be used in the long term to help with food insecurity and agricultural emissions.

Datasets

In our project, we sourced and cleaned two different datasets. The first dataset we used contained prices for different commodities in different countries [3]. In the raw data, there were mismatched names of commodities, so we made sure the names were consistent. The data was also rife with inconsistency in the times associated with the prices of different commodities. We used null infilling practices to smooth out the data.

The second dataset we used was a food production and emissions data set [4]. This data had the agricultural statistics of food emissions and production in different countries. For each commodity, the data had three different categories: TONNE_HA which is 1000 kilograms per 10,000 square meters (density/efficiency), THND_HA which is thousand 10,000 square meters (land usage), and THND_TONNE which is thousand 1000 kilograms (weight of crop production).

The food commodities we chose to use were wheat, rice, and maize since these are major foods that most people consume and there was a lot of data on those commodities. From the vast collection, we chose five different countries – China, Ukraine, India, Peru, and Colombia – since they were present in both datasets. We also reviewed the data for each country until we found countries that had overlapping data between the two datasets. These countries also had a plethora of data on them compared to other countries.

We cleaned up any remaining data by making sure all the dates matched up in chronological order and then decided to graph both the prices dataset as well as the food production and emissions data set to get a better visual of our data. These are shown below in Figure 1, 2, and 3.

HIGH SCHOOL EDITION
Journal of Student Research Volume 13 Issue 4 (2024)

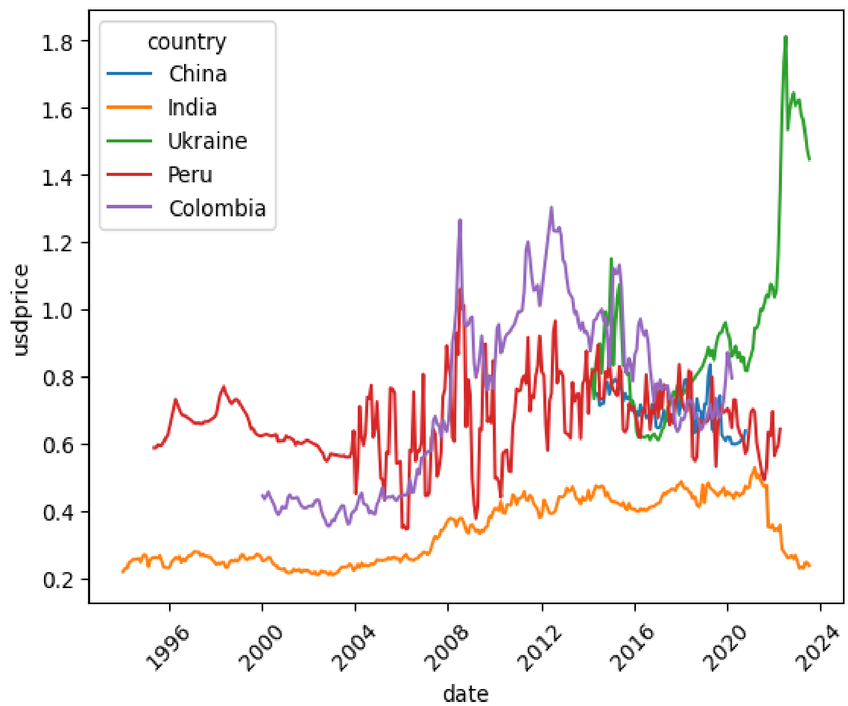

Figure 1. Graph of prices for rice from 1994 - 2024.

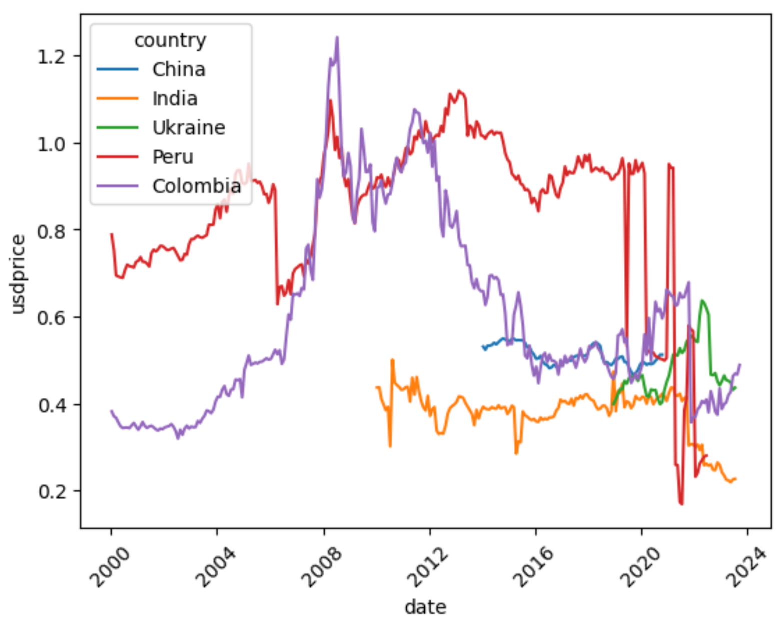

Figure 2. Graph of prices for wheat from 2000 - 2024.

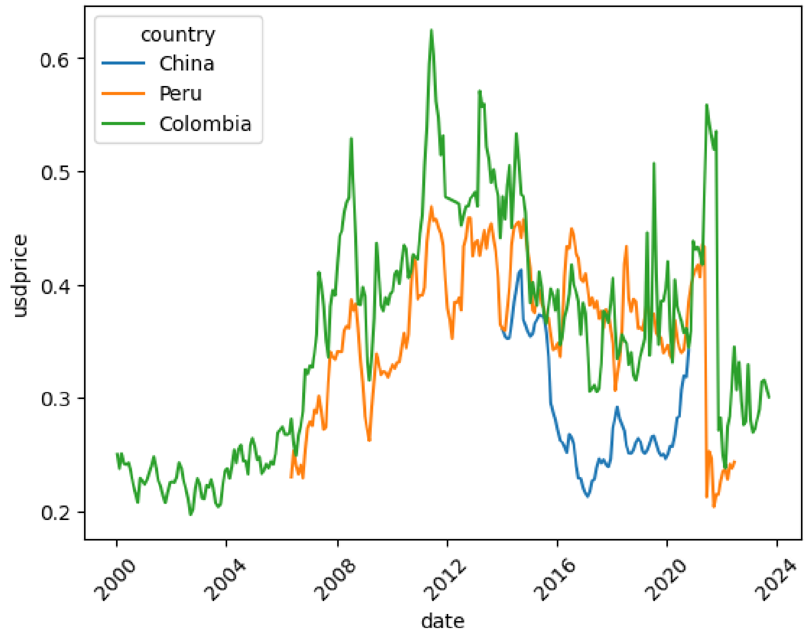

Figure 3. Graph of prices for maize from 2000 - 2024.

At first, these graphs appear very volatile and hard to predict. Upon closer inspection, patterns emerge. Most of the lines on the graphs follow the same trend as the other graphs and other countries. One major thing we noticed was a big spike in 2021 and 2022. This could possibly be due to inflation and supply chain issues during the Covid-19 pandemic.

We also explored the production density of these three commodities, shown in Figure 4, 5, and 6.

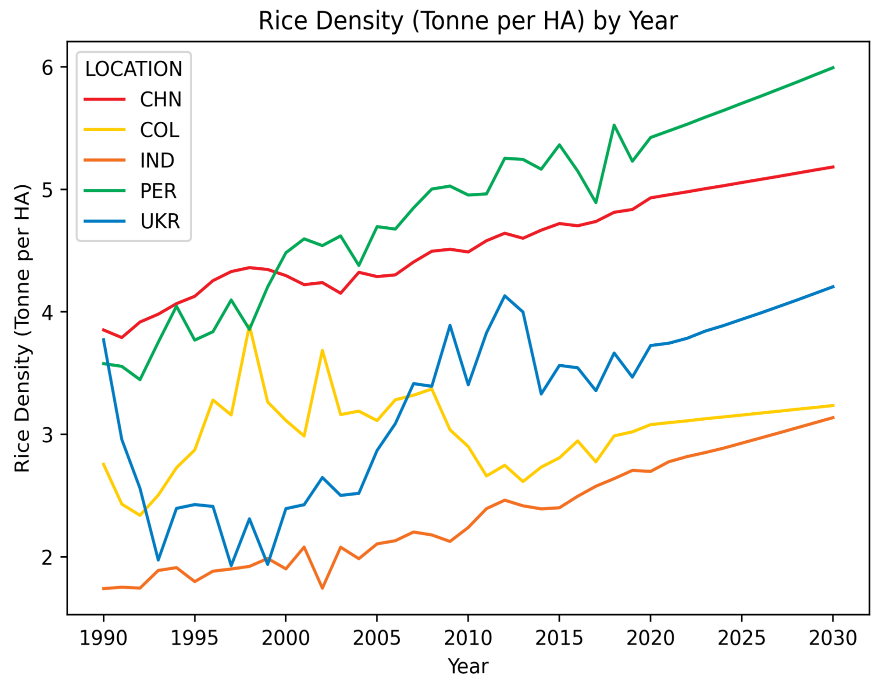

Figure 4. Graph of density for rice from 1990 - 2030.

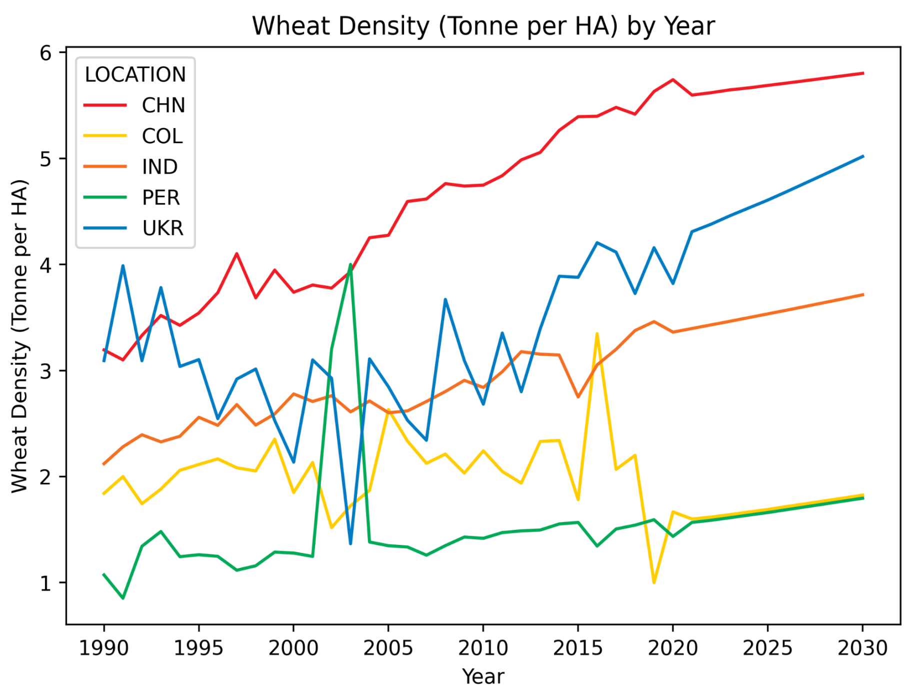

Figure 5. Graph of density for wheat from 1990 - 2030.

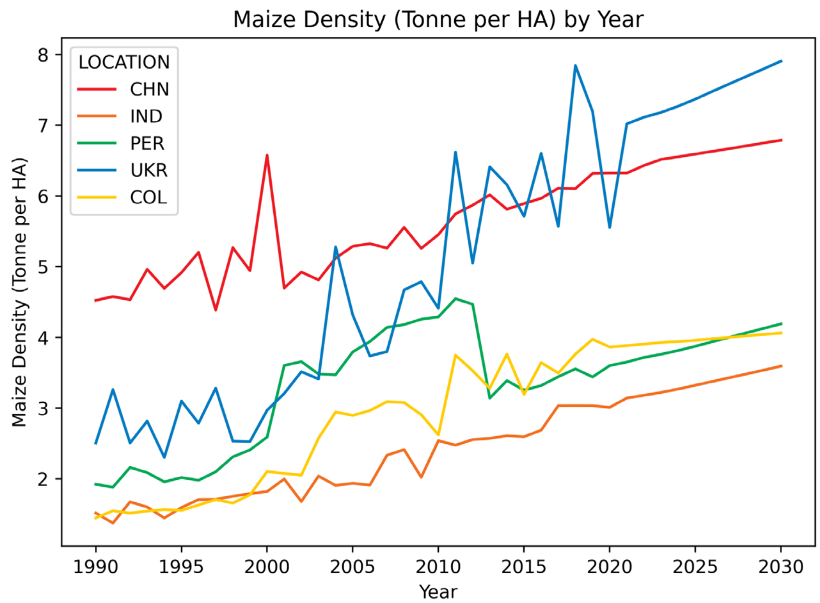

Figure 6. Graph of density for rice from 1990 - 2030.

These graphs show that over time, the efficiency of agricultural farming has been improving. We can also see that in the countries we have chosen, China is usually a frontrunner in terms of land-use efficiency while India and Columbia are usually at the bottom. There are also a lot of spikes in the graphs that seem random, but are likely due to external factors such as weather.

Methodology / Models

To predict the prices of the food commodities and agricultural emission, we decided to use the ARIMA model which stands for Autoregressive Integrated Moving Average. ARIMA is a time series forecasting model used for analyzing and predicting time-dependent data. This model essentially works by taking the points of the previous known data and using the trend to predict how the graph will look point by point.

We started with the food prices data set. We split the column of data that contained the actual prices into train and test data. The training data would be the data that the model would see and look for a trend. Then it would output data for what it thinks the graphs would look like. We would then compare the outputted values to the test data which is the actual data to see if the model was able to accurately predict the results.

We did this same thing for the land use and food production dataset. We split it into train and test data and then compared it to the real values.

Results and Discussion

After running the ARIMA model on the prices dataset and comparing to the realized values from the held-out set, some of the graphs closely matched the realized data while some were inaccurate. We illustrate three of the accurate results in Figures 7, 8, and 9 below.

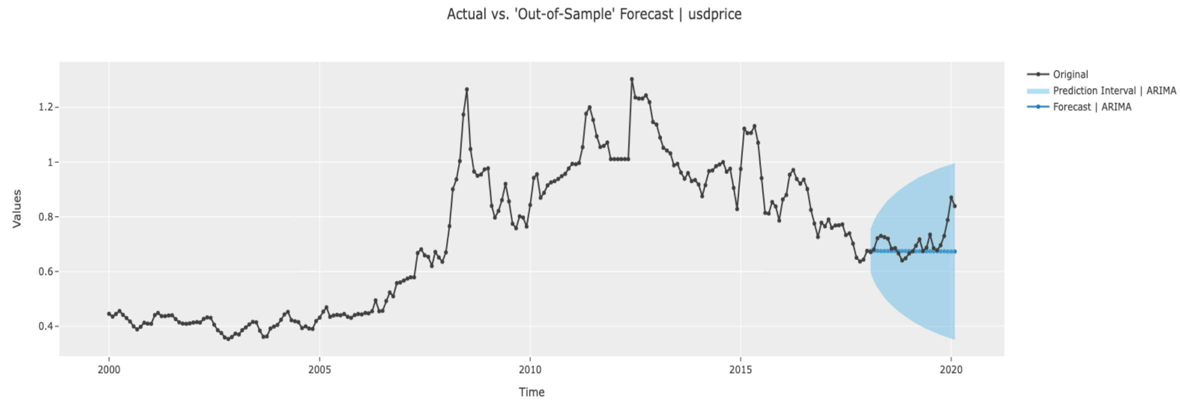

Figure 7. Graph of prediction of prices for rice in Colombia.

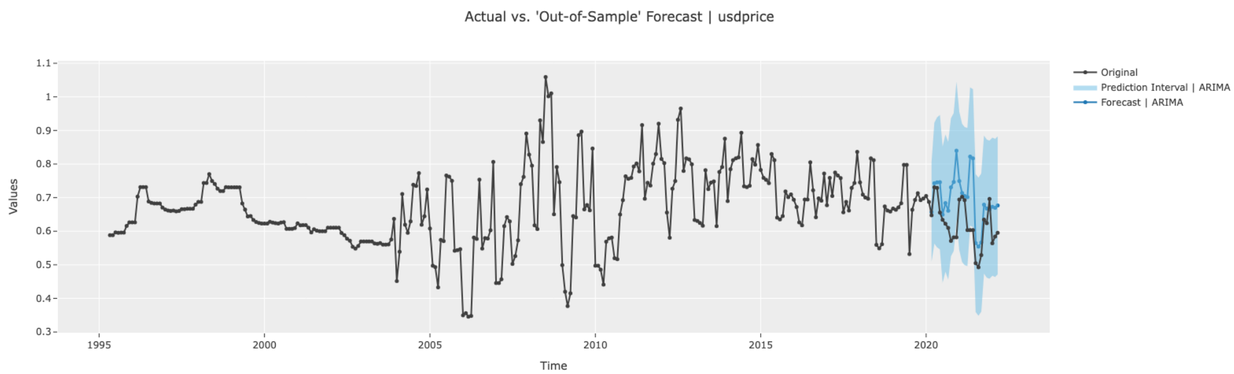

Figure 8. Graph of prediction of prices for rice in Peru.

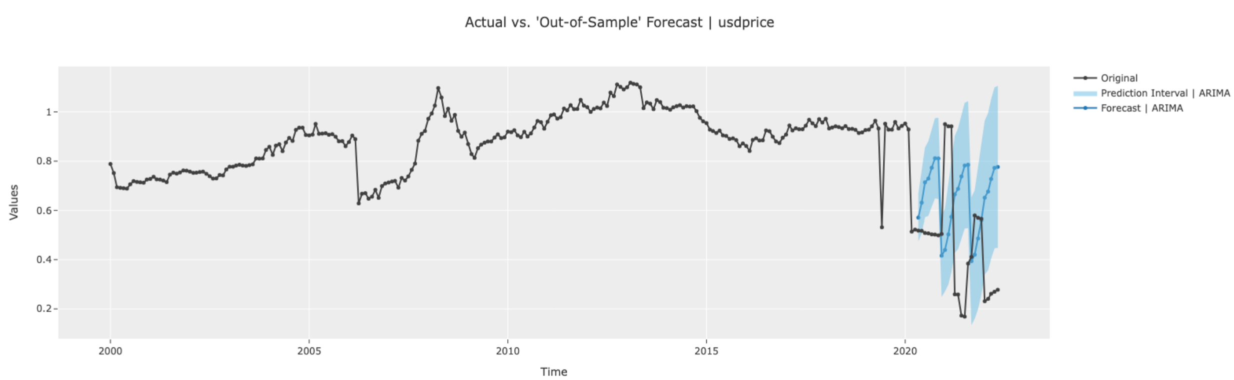

Figure 9. Graph of prediction of price for wheat in Colombia.

In these graphs, we are able to see the actual data which is in black, and the models predicted data which is in blue. In these graphs, we can see that the model was able to predict the food prices really well and how the trends in the graph are pretty close to what actually occurred. This shows us that this model definitely has some capability to accurately predict the prices of food commodities in different countries.

HIGH SCHOOL EDITION
Journal of Student Research

We show three of the inaccurate results in Figures 10, 11, and 12 below.

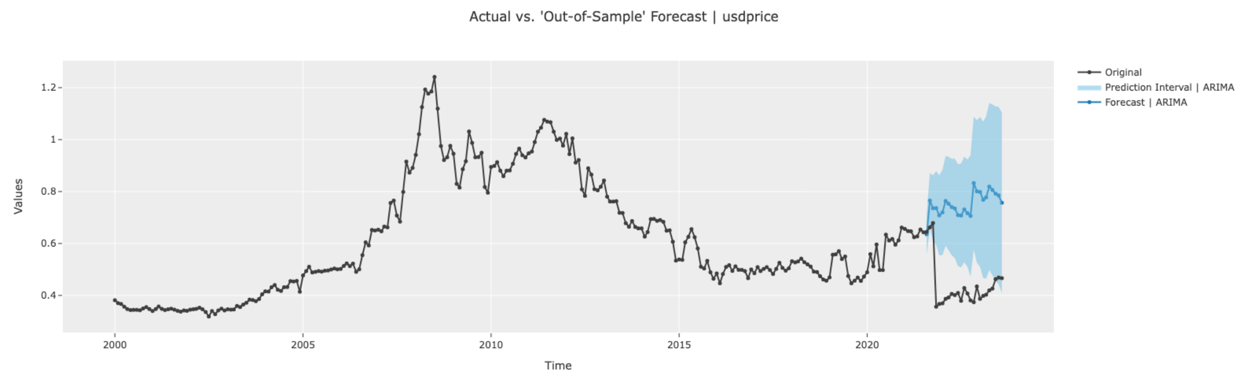

Figure 10. Graph of prediction of price for wheat in Colombia.

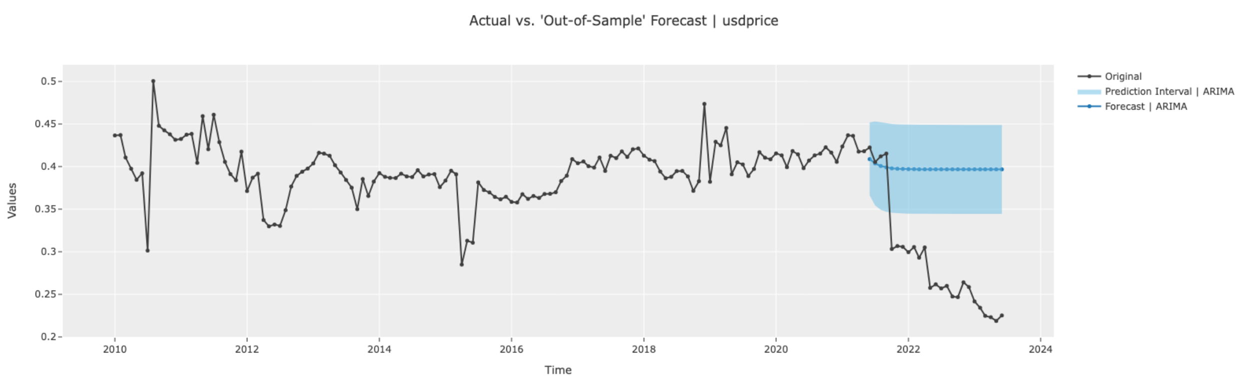

Figure 11. Graph of prediction of price for wheat in India.

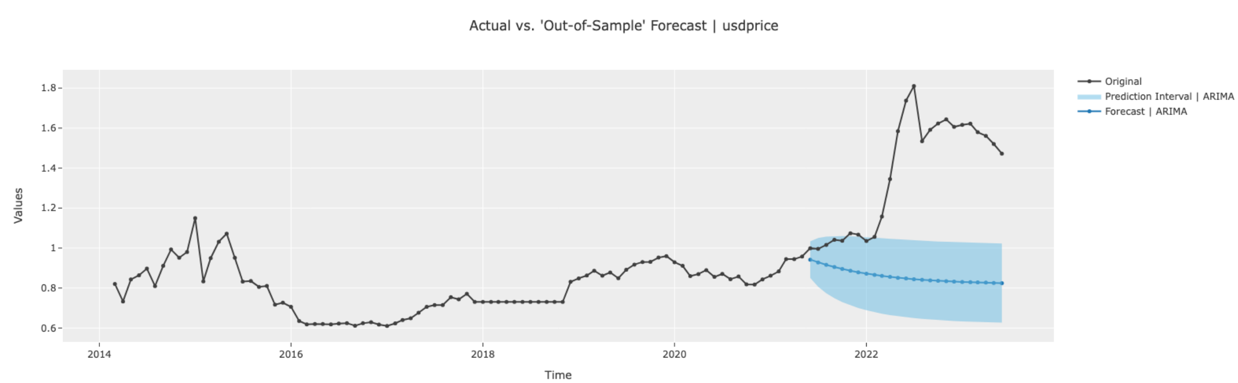

Figure 12. Graph of prediction of price for rice in Ukraine.

HIGH SCHOOL EDITION
Journal of Student Research

These graphs show a different story than the ones previously. The model was not able to predict the trends or the points of these graphs. Upon looking at the graphs further, we realized that the model was not able to predict these trends because from 2020 onwards, the values became very volatile and unpredictable.

We decided to research this more to find out more information. After further research, we realized that one possible explanation for the sudden drop in prices for the first two graphs could be the Covid-19 pandemic. We also realized that for the last graph, the spike in Ukraine could be due to the ongoing Russia and Ukraine war. While these guesses are not confirmed, they are a very good guess on why this was happening.

We also applied the ARIMA model to the agricultural emissions dataset and got some interesting results as well. These are shown below in Figures 13, 14, and 15.

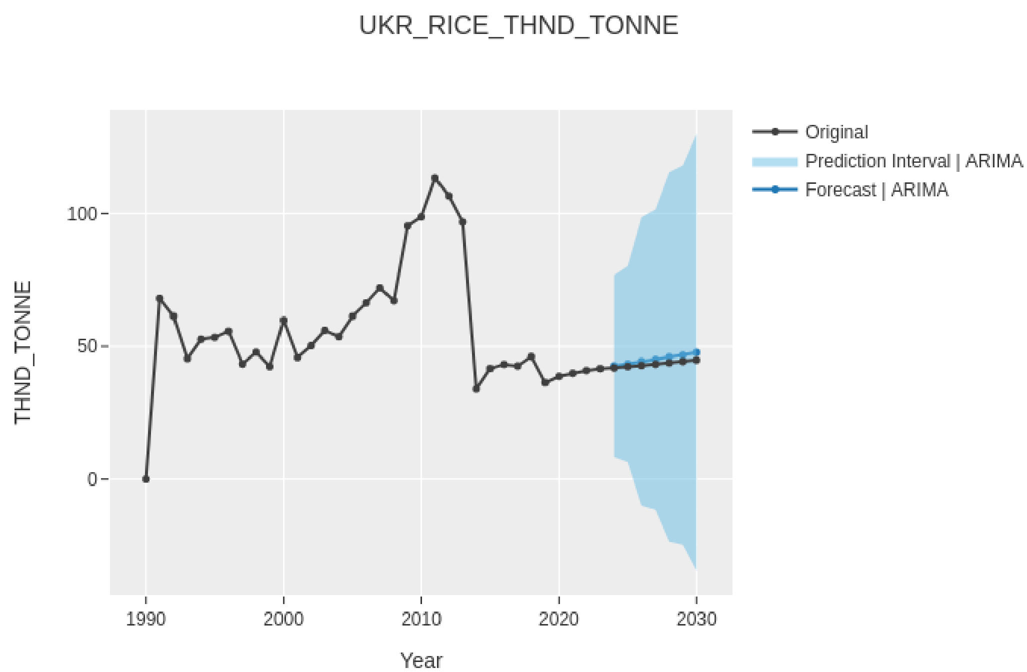

Figure 13. Graph of prediction of weight of crop production for rice in Ukraine.

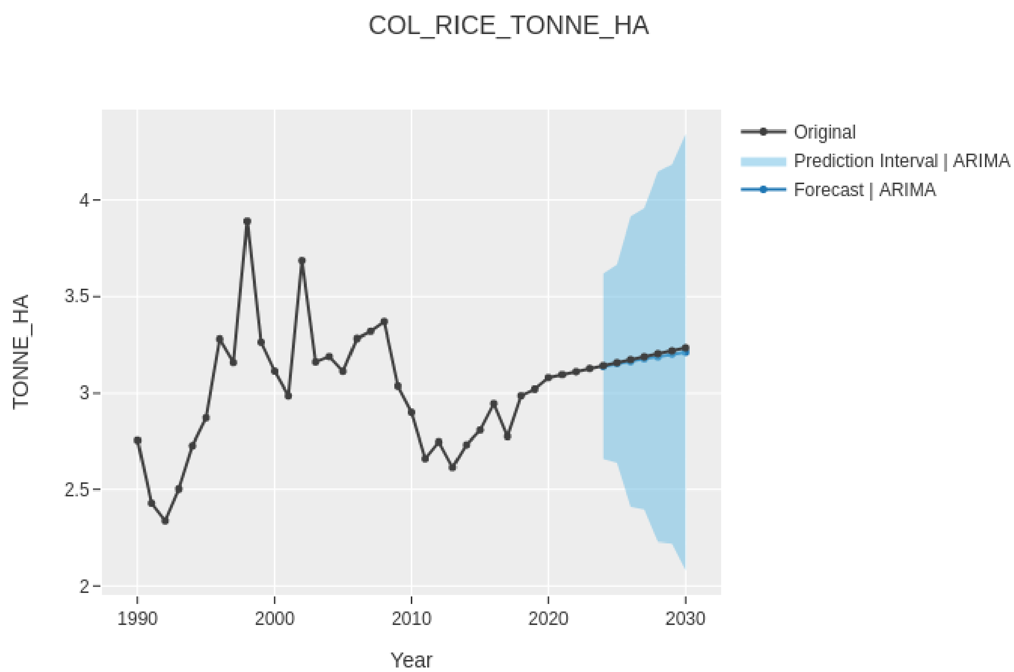

Figure 14. Graph of prediction of density for rice in Colombia.

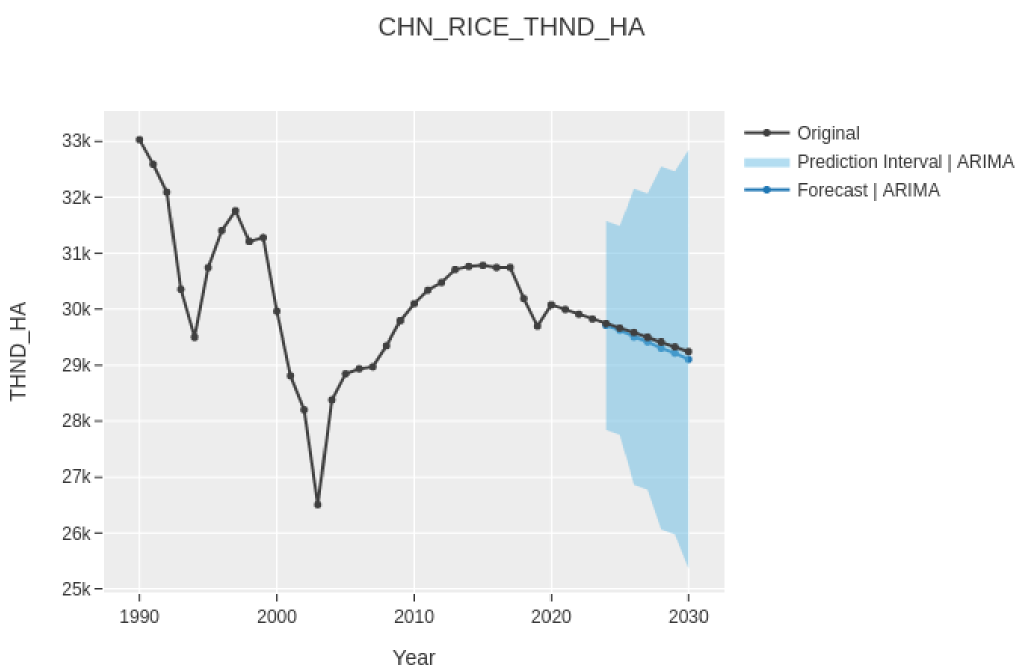

Figure 15. Graph of prediction of land usage for rice in China.

In these graphs, the model was essentially able to predict the trends and the points dead on. There was only some very minor difference between the real and the predicted graphs. This is very good since it means that predicting the food production and land usage is possible and it could be used to help improve agricultural emissions overall.

However, again we found some unfavorable results for other commodity-country pairings, shown below in Figures 16, 17, and 18.

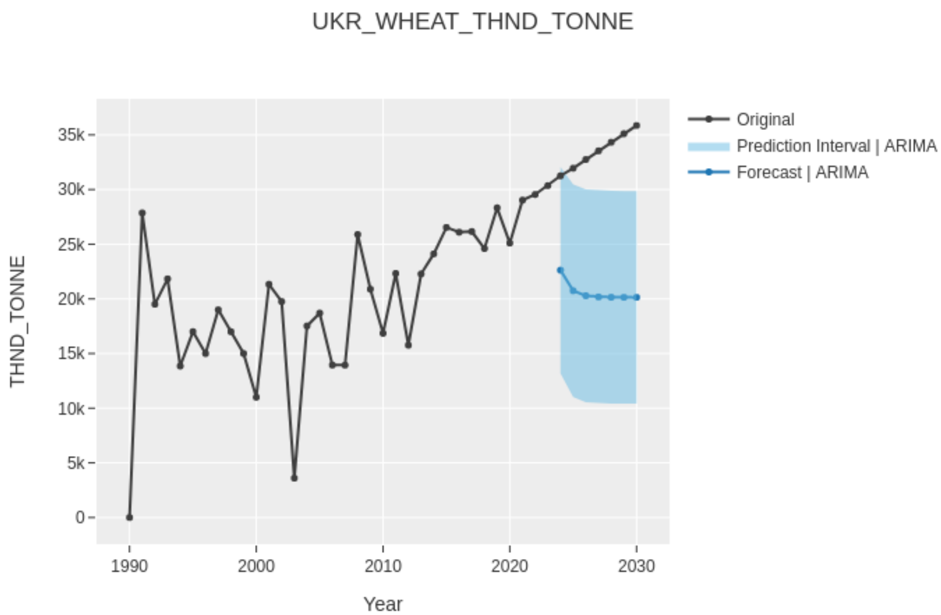

Figure 16. Graph of prediction of land usage for wheat in Ukraine.

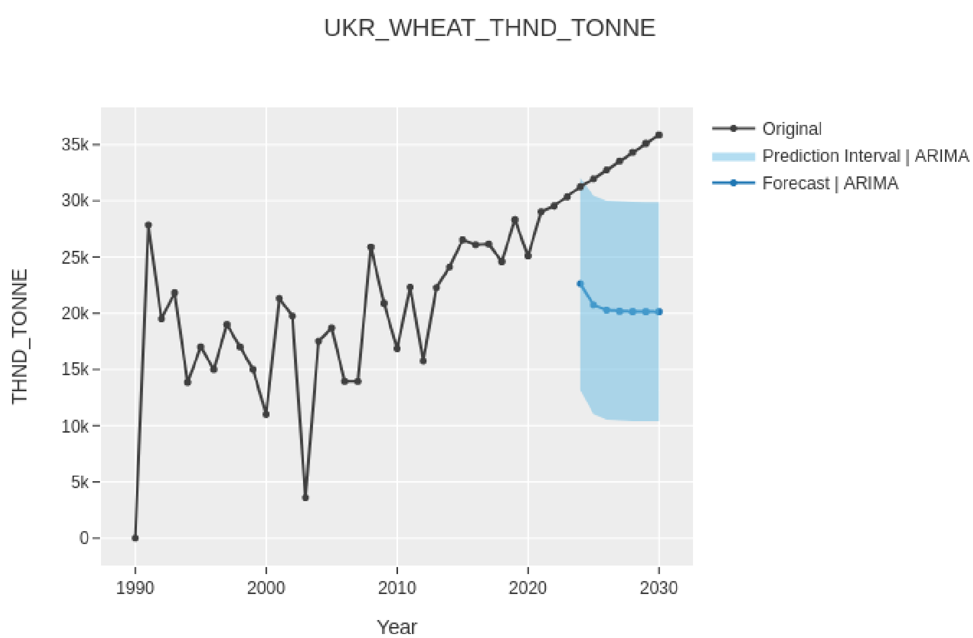

Figure 17. Graph of prediction of weight of crop production for wheat in Ukraine.

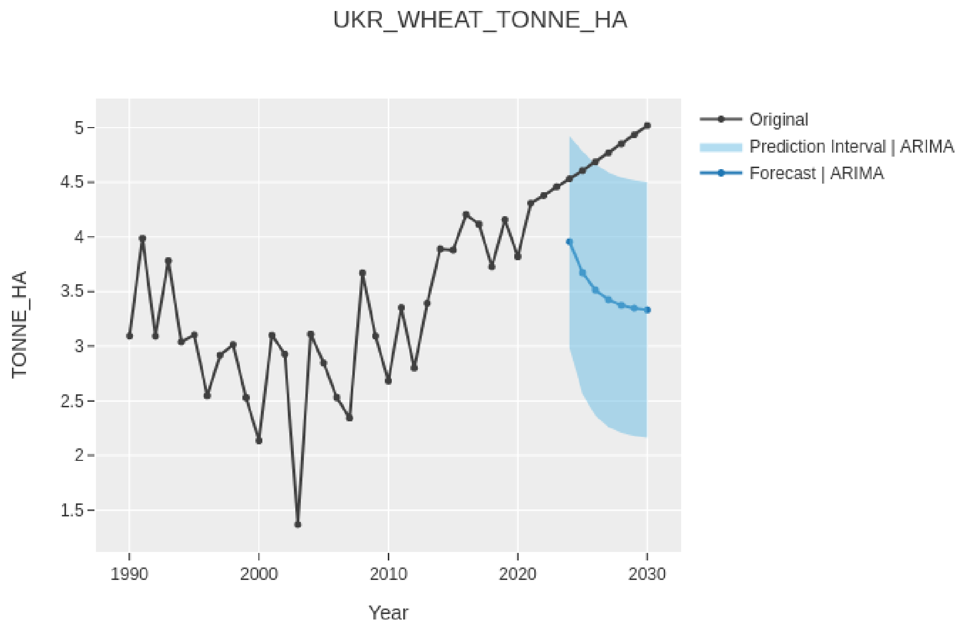

Figure 18. Graph of prediction of density for wheat in Ukraine.

These graphs show that the predictions were extremely off from the real data. After doing some research, we still could not find the exact reason for this. Our guess was that since these graphs were way more volatile, the model had a harder time predicting how the trends would go.

There also could be many external factors affecting these graphs such as inflation, weather, and global pandemics.

Overall, our results show that the ARIMA model definitely has the capability to predict the food prices as well as the agricultural emissions of the data we provided. This means that we could possibly use this in the future. There are numerous ways we could improve this. Through our research we have found that there are many external factors which make these graphs very volatile and hard to predict. If we had the data of these external factors and what is happening in the world, we might be able to get an even more accurate prediction on this data.

Conclusions

In this research, our aim was to predict international commodity prices, production, and land usage, emphasizing the mitigation of agricultural emissions. Utilizing the ARIMA model, we achieved some success in forecasting food prices for certain commodities and countries. While the model demonstrated accuracy, challenges arose in predicting trends influenced by external factors such as the COVID-19 pandemic, weather, and geopolitical events.

The model's performance reflects the complexities of global events and external influences. To enhance its predictive capacity, future steps involve incorporating external factors into the analysis, exploring alternative modeling approaches, and considering diverse datasets. These refinements are crucial for a comprehensive understanding of agriculture, emissions, and food security.

Beyond academic realms, our findings hold practical implications. Policymakers can leverage insights to shape strategies addressing climate change and food insecurity. People can also use our model to buy food when prices are low or for other practical reasons. The predictive models offer a foundation for aid initiatives strategically targeted at countries facing challenges highlighted in our research.

In conclusion, while our research provides significant insights, it lays the groundwork for continued exploration. Ongoing efforts to refine models, expand datasets, and address challenges posed by global dynamics are essential for contributing to a more sustainable and secure future for global agriculture.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

References

[1] "Sources of Greenhouse Gas Emissions," EPA, https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions.

[2] Y. A. Zereyesus and L. Cardell, "Global Food Insecurity Grows in 2022 amid backdrop of higher prices, Black Sea Conflict," USDA ERS - Global Food Insecurity Grows in 2022 Amid Backdrop of Higher Prices, Black Sea Conflict, https://www.ers.usda.gov/amber-waves/2022/november/global-food-insecurity-grows-in-2022-amid-backdrop-of-higher-prices-black-sea-conflict.

[3] "Home," Humanitarian Data Exchange, https://data.humdata.org/dataset/global-wfp-food-prices.

[4] Docxian, "Crop production & climate change eda," Kaggle, https://www.kaggle.com/code/docxian/crop-production-climate-change-eda.

[5] P. Foini, M. Tizzoni, G. Martini, D. Paolotti, and E. Omodei, "On the forecastability of food insecurity," Nature News, https://www.nature.com/articles/s41598-023-29700-y.

[6] G. Martini et al., "Nowcasting food insecurity on a global scale," medRxiv, https://www.medrxiv.org/content/10.1101/2021.06.23.21259419v1.full-text (accessed Jan. 19, 2024).

[7] World Bank Group, "Climate explainer: Food Security and climate change," World Bank, https://www.worldbank.org/en/news/feature/2022/10/17/what-you-need-to-know-about-food-security-and-climate-change.