

Agricultural Commodities and Weather

HITAANSH GAUR, University of Maryland, College Park, United States of America

DAVAYAN SANYAL, University of Maryland, College Park, United States of America

AYUSH SOOD, University of Maryland, College Park, United States of America

Hitaansh Gaur, Davayan Sanyal, and Ayush Sood. 2023. Agricultural Commodities and Weather. 1, 1 (December 2023), 6 pages. 10.1145/nnnnnnnn.nnnnnnnn

1 INTRODUCTION

Climate events can have profound effects on the return of agricultural company stocks. In recent years, the global climate has become increasingly unpredictable, leading to extreme weather conditions such as droughts, floods, and heat waves, all of which significantly impact the agricultural sector. The big picture here is the urgent need for better tools and strategies to mitigate the financial risks associated with climate-induced volatility in agricultural markets. Improved cooperation among stakeholders, including farmers, investors, policymakers, and climate scientists, is essential to address this pressing issue.

Farmers respond to fluctuations in weather patterns affecting their crop output by either keeping prices constant or raising. The problem is predicting which will happen, and how much the subsequent change in prices will change the return of agricultural company stocks over the year. Addressing this problem is critical because it directly impacts the financial stability of agricultural investors and the broader financial market. If we can better understand and predict how farmers will adjust their pricing strategies in response to weather-related challenges, we can improve the accuracy of forecasts regarding agricultural stock returns. This would lead to more informed investment decisions and ultimately contribute to greater financial stability and resilience in the agricultural sector, benefiting both investors and the industry as a whole.

The challenges are that weather is often varied so it is not easy to make broad predictions/have a centralized decision-making process to determine how farmers will respond and further determine how the return of agricultural commodities will be affected. In addition to this variability, the accessibility of crucial data poses substantial challenges. Specific agricultural commodity data, including essential information like wheat futures prices, is frequently locked behind paywalls enforced by institutions such as the Chicago Mercantile Exchange (CME). This restricted our ability to obtain comprehensive, granular data directly from these sources, forcing us to change our original approach. Furthermore, the process of gathering weather data presents was difficult as well. While the National Weather Service (NWS) provides weather data, it often lacks a user-friendly format for convenient download and analysis, necessitating time-consuming manual data copying and transformation. These challenges impeded and slowed down our progress.

We developed linear regression models to find a relationship between weather and agricultural company stock returns. Our contributions were making predictive analyses that can measure and evaluate the impact of weather on agricultural markets. Specifically, we focused on the relationship between temperature and agricultural company stock returns.

Our models are interesting because they identify an important trend between certain weather factors and agricultural company stock returns. They can also be expanded to test more weather

factors and different returns data. Our specific insight is that the relationship between temperature and agricultural stock returns is not as significant as we pose it to be, and stock returns are often more complicated than just single factors. This is interesting because it is counterintuitive - it is expected that weather would dominate the returns of an agricultural company because it controls the core of the business.

This is a problem that has been explored in the past, so the following section explores related works.

2 RELATED WORKS

As we got started we looked at some examples of past works to get an idea of how to approach the problem. One paper we looked at was 'The Impact of El Nino-Southern Oscillation on U.S. Food and Agricultural Stock Returns' by researchers at the University of Clarkson, published in the *Water Resources and Economics* journal. It identified important weather cycles such as El Nino and explained their potential impact on farming and their subsequent impact on farmer's returns. The paper then looks at the real farming output as a result of the fluctuations in weather and how much that possibly impacts commodity prices, looking at time-series dating starting in March 1980 and ending in December 2018. The time series data includes the stock prices of 12 different SP 500 agricultural companies, and each of these companies is broken down by their main products and how they are primarily affected by El Nino. Then a mathematical model was created, and its various properties, such as the variance decomposition, were explained. Overall this paper is very in-depth and informative, as it is easy to understand because of the numerous explanations. Also, the paper questions many of its own stances to check if the correlations and causations match, and shows a lot of visualizations to back up its many claims. It ultimately finds that weather shocks from the El Nino pattern have a relatively short-lived impact on stock returns, of three to six months. This paper only focuses on one region of the world and it looks at short-term commodity pricing, so beyond this scope, we would want to look for a way to predict prices in the longer term. From looking at this paper, we got the idea to find specific companies and look at the locations they were primarily located in. Instead of looking at the country's weather as a whole, we just choose the weather data from that specific region. Finally, we found monthly time series data similar to this paper and created our own model and visualizations.

In our research, we also found a paper by the International Monetary Fund (IMF) on the impact of Climate Shocks on Inflation and Economic Growth. The paper answered the question: does inflation have a long-term effect on the commodity prices we are trying to measure? This paper mainly focuses on the effects of droughts and extreme rainfall on the entire sample of 173 countries. Similar to the previous paper, the researchers came up with a mathematical model, but this time it was to predict the impact of the weather shocks on the country's economy. They were interested in the impact of disastrous weather rather than farming itself, although farming was shown to be something disrupted by weather patterns as well. The paper is very informative with in-depth explanations of the mathematical models and shows various visualizations comparing the impact of adverse weather on developing countries versus developed ones in sectors such as food and finance. The paper also delves into the impacts of administering economic controls to counter the effects of climate shocks. Overall, this paper is interesting but it is related a little less to our own research topic and more broad in scope due to how many countries it generalizes across. Nevertheless, it is important to keep the angle of this paper in mind while we formulate our own research. We ended up using the technique of focusing on specific countries instead of regions. In our case, we based our research on specific US states. Additionally, in our research, we were not able to find any results that pointed to weather having a disastrous impact on the agricultural stocks.

Beyond the papers, we found various Githubs where people have created their own models to predict El Nino weather patterns and commodity pricing. Some repositories, such as the one by user Abhinav Reddy, just focused on the prediction of El Nino and La Nina weather patterns. Other repositories just focused on predicting stock returns through other means. We decided to use some of the same machine learning techniques the authors used in our model, and we discuss them more in-depth below in our approach.

3 APPROACH

We propose a set of linear regression models to evaluate the efficacy of predicting agricultural returns of three major agricultural companies based on averaged weather data across three cities. The three companies we chose to evaluate are Archer-Daniels-Midland Co (NYSE: ADM), Bunge Global SA (NYSE: BG), and Tyson Foods Inc (NYSE: TSN). Further, the three cities whose temperatures we chose to use are Fort Worth, TX, Audubon, IA, and Neosho, MO. We retrieved stock data for these three companies and retrieved the weather data from Wharton Research Data Services, which is a service provided to us by the finance department of the Smith School of Business. From WRDS we looked at historic CRSP data outlining the monthly returns of the stocks. To adjust for inflation over time and get real returns, we also acquired CPI data from the market index historical data in the CRSP database. The reason we chose these companies is because in our research, we found them to be the three largest agricultural companies within the United States, and we thought that getting stock and weather data within the US would be the easiest. Since these companies have farming activities spread across various states, we chose the three cities that we did because they have the largest percentage of farming activity in the US and our companies are located in or near these states. We acquired weather data from the NWS (National Weather Service) website. They had an input form for past temperature data, and it would display a table of the monthly average weather for the location you specified. We had to manually copy this data into an Excel spreadsheet because NWS did not provide a way to download the data into a spreadsheet. This took a bulk of time because we had to copy hundreds of temperatures by hand, and also double-check to make sure there were no typos. Since we only had 20 years of weather data, we only needed to get stock data dating back to 1999. This made the model easier to run, but it also made our training set a lot smaller.

In order to successfully generate a set of linear regression models on this data, a substantial amount of data processing had to be done. In our data retrieval, we stored the stock data in one CSV, the inflation data in another, and the weather data in a third. Additionally, because of the nature of the query we used to retrieve the stock data, there were many columns that we did not actually need. So, the first thing we did was isolate the columns that were of use to us. We did all our data processing through Pandas; after reading the CSV files, we kept only the date, ticker (NYSE), and returns columns from the stock data. The inflation data also had some unnecessary columns, so after merging the stock data table with the inflation data table on date, we dropped the unnecessary columns, so the resulting table contained only the columns for date, ticker, returns, and inflation.

Next, we had to process the weather data. While mostly complete, there were a few entries in each column that were blank, so we decided to average fill these rows. Subsequently, since our stock/inflation data and weather data were, at this point, in separate tables, we needed to merge these together. Our stock/inflation table had a column with specific dates (YYYY-MM-DD), while our weather table's date column was only monthly (YYYY-MM). So, in order to merge the two tables, we removed the date proportion from the stock/inflation table. After joining on the month column, we attempted to add a column with the real returns (adjusted for inflation, so raw returns subtracted by inflation) but ran into an issue where one row in the real returns column,

strangely, did not contain a float like the rest of the rows but instead contained a single alphabetic character. Since this was preventing us from broadcasting the subtraction operation between the two columns, we simply forward-filled this row and proceeded with the subtraction. With this, our data processing was complete, and we wrote the resulting table to a new CSV file for ease of access when developing our linear regression models.

We used Scikit-learn (sklearn) to generate our linear regression models. Since we wanted to generate a model to see how well each specific company’s returns faired when analyzed with respect to the averaged weather data across our cities of interest, we first read the resulting CSV from our data processing and then split it into three tables: one for ADM, one for BG, and one for TSN. Finally, we were able to generate our three linear regression models. We used the averaged temperature column in each table as our X (input) data, and the real returns column in each table as our y (output) data. Using sklearn’s `linear_model.LinearRegression()` function, we were able to fit our three models, plot each one in a combined plot using `matplotlib.pyplot`, and analyze the properties of each model using the `intercept`, `coefficient`, and `score` attributes that `linear_model.LinearRegression()` generates for each model it fits.

4 RESULTS

The linear regression model that we developed for the stock data from ADM had an intercept of about 0.030, a very slight negative slope of about -0.0004, and a score of about 0.005. The model developed for BG had an intercept of about 0.028, a very slight negative slope of about -0.0003, and a score of about 0.004. Finally, the model developed for TSN had an intercept of about 0.058, a slight negative slope of about -0.0008, and a score of about 0.02. Figure 1 summarizes these results.

Ticker	Intercept	Slope	Score
ADM	0.030	-0.0004	0.005
BG	0.027	-0.0003	0.004
TSN	0.058	-0.0008	0.020

Fig. 1. Results of Linear Regression models

As can be seen, none of our models suggest a very strong linear relationship between averaged temperature and returns, but if any such relationship does exist we might presume that the relationship is inverse, since the returns go down as the temperature increases. However, even making that statement is questionable given the results that we achieved, since the score of all 3 models was so low. Below is a graph displaying all 3 linear regression models; the very weak negative correlation can be seen there.

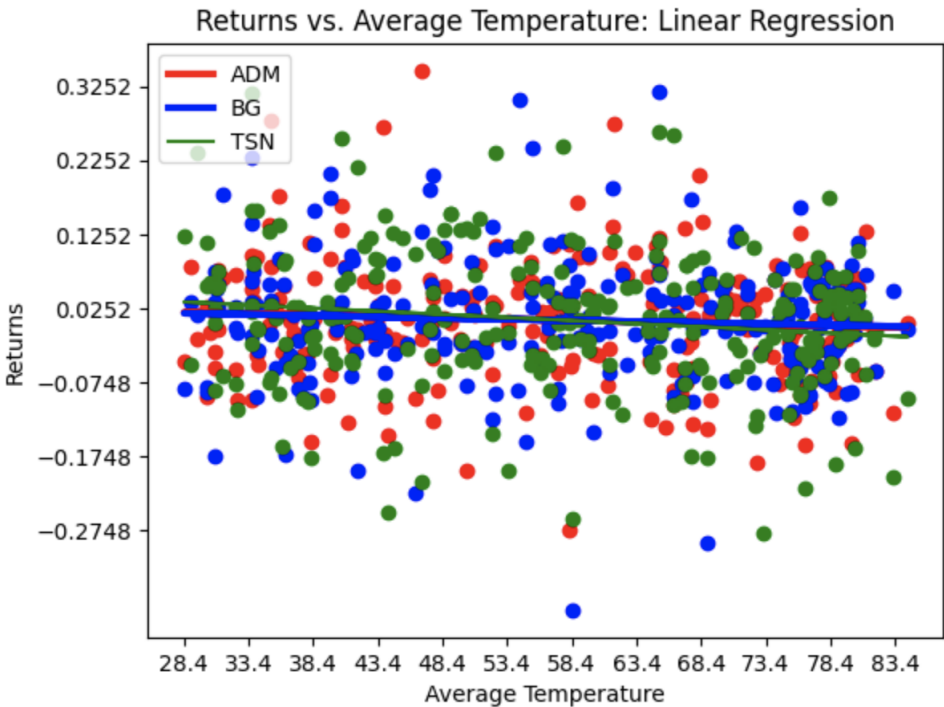


Fig. 2. Returns vs. Average Temperature: Linear Regression

5 CONCLUSION AND RECOMMENDATIONS

Our research project was very short and had a very small number of parameters. We learned that to produce meaningful results, we need to spend quite a bit longer on our data collection stage, and creating and running the model is actually not that hard. The weather data we collected is likely too general to draw a good conclusion about the performance of agricultural stocks. This is for several reasons, the first being that agricultural companies are big and their stocks can change for many different reasons. Month-to-month changes in weather which would negatively impact stock returns are likely mitigated by these companies already to keep shareholders happy. Sometimes, share prices move in relation to the overall market or investor sentiment, and the share prices might have nothing to do with actual agricultural output. Second, our monthly weather forecasts are too broad to be applied directly to companies because we don't know how the weather actually affects the crops, which crops are being grown (different crops react differently to different kinds of weather), and exactly in which state. For the purpose of this project, we were trying to see if we could scout a simple correlation between two broad parameters, but the reality is we actually need data that has some level of correlation first. In conclusion, our project was made from assumptions that were far too broad to produce meaningful results. As a result, it's tough to make meaningful recommendations from our findings. One such recommendation we can make is that there is a slight correlation between TSN stock and weather patterns. This might be because TSN has the most farming activities in the states we picked, or there just might be a slight correlation between weather and TSN returns without exact causation. So while we might recommend buying and selling TSN stock based on the monthly weather forecast, there's not enough data available in this

project to substantially back up this claim. For a future project on this topic, we would spend a lot longer finding data on specific companies, such as what crops they grow, where they grow the crops, and how they work with individual farmers who are the ones raising or lowering crop prices. We might also mitigate a lot of these issues by paying for commodities data, which is better than stock data because it focuses on the day-to-day value of the crop itself instead of the valuation of the entire company.

REFERENCES

- [1] AbhinavReddy-Dev. *El-Nino-La-Nina-Prediction*. 2020. URL: <https://github.com/AbhinavReddy-Dev/El-Nino-La-Nina-Prediction>.
- [2] E. Topalova P. Acevedo S. Mrkaic M. Novta N. Pugacheva. *The effects of weather shocks on economic activity: What are the channels of impact?* 2018. URL: <https://www.imf.org/-/media/Files/Publications/WP/2018/wp18144.ashx>.
- [3] E. Atems B. Maresca M. Ma B. McGraw. "The impact of El Nio-Southern Oscillation on U.S. food and agricultural stock returns." In: *Water Resources and Economics* 32 (). URL: <https://doi.org/10.1016/j.wre.2020.100157>.
- [4] A Naqvi. *Climate-econ*. 2022. URL: <https://github.com/asjadnaqvi/climate-econ>.
- [5] National Weather Service. n.d. URL: <https://www.weather.gov/>.
- [6] Wharton Research Data Services. n.d. URL: <https://wrds-www.wharton.upenn.edu/>.