Will Bartlett and David Whitmer

DATA300 & DATA201 Final Project Report

14 May 2025

Long-Term Wheat Price Prediction Using Time Series and Deep Learning Models

1. Abstract

This study aims to evaluate the long-term predictive potential of time series models compared to machine learning and deep learning models. Accurate price modeling is important for farmers to prepare for price changes and food shortages. The Seven predictive models tested are AR, ARIMA, SARIMAX, Decision Tree, XGBoost, VAR, and LSTM. Exogenous features were chosen to capture changes in other commodity pricing as well as macroeconomic factors. The models were trained on log return of monthly wheat prices from 1990-2018, with forecasting to 2018-2024. Our LSTM model drastically outperformed all other models tested, with an RMSE of 3.97. Despite LSTM's ability to factor non-linear relationships, the errors show significant struggle to capture long-term shocks and movements in the price of wheat.

2. Introduction

Forecasting wheat prices is imperative given wheat's role as a global staple crop, influencing food security, agricultural economics, and commodity markets. This project aims to develop accurate predictive models for long-term wheat prices using historical monthly data. It incorporates key economic features such corn prices to capture supplemental pricing, flour prices to capture the downstream impact of wheat price changes, commodity import and export indices, crude oil pricing to measure energy commodity impact, and consumer price index. Environmental indicators were added, including a drought index, a fertilizer index, and average temperature. The primary goal is to identify the most effective model through error minimization, using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). By leveraging both time-series and machine learning approaches, the project seeks to capture complex temporal and feature-driven patterns in wheat price fluctuations. The motivation for this project stems from the need to support farmers, policymakers, and traders in making informed decisions amidst volatile market conditions. Wheat price volatility, driven by environmental factors like temperature and drought poses significant challenges to agricultural planning. Accurate forecasts can mitigate risks, optimize resource allocation, and enhance market stability. By comparing model performance, the study seeks to provide insights into the suitability of deep learning versus traditional time-series methods for agricultural price prediction. This work is particularly relevant in the context of climate change, which exacerbates environmental impacts on crop prices. Ultimately, the project aspires to deliver reliable forecasting tools that empower stakeholders across the agricultural supply chain. These tools can facilitate better strategic planning and contribute to global food security.

3. Team Member Roles

Will: Primarily handled meeting with Professor Wang, coding our Model, researched prior literature, and collected data.

David: Primarily handled data collection and research. Additionally, he managed the GitHub repository.

Note: Report and poster work was split amongst the two.

4. State-of-the-Art/ Related Work

The prediction of wheat prices has evolved into a multidisciplinary field integrating agronomic, technological, and socioeconomic insights. Traditional time series models (Brandt and Bressler) remain foundational for analyzing historical price trends but often struggle with volatile externalities such as climate shocks or geopolitical disruptions. Advances in software have allowed (Darekar and Reddy) to finetune parameters and forecast with 95 percent accuracy. Moreover, technical indicators, unlike economic indicators, have been able to predict the volatility of wheat prices (Liu and et.). A recent 2023 study (Liu and et.) reviewed agricultural price prediction methods, concluding that future trends involve combined models integrating structured and unstructured data to enhance both value accuracy and trend precision, addressing challenges in achieving sustainable agricultural development. Lastly, Hoffman and Westcott) developed an analytical framework for predicting prices and establishing a mechanism to ensure consistency across USDA's supply, demand, and price forecasts. Further, most of the past research incorporates rolling-windows for short-term prediction.

5. Approach

    a. Data Collection

Using publicly available sources, we collected data for our features. Commodity export index, commodity import index, corn price, crude oil price, fertilizer index data, flour price, consumer price index, and wheat price were all found from FRED. The monthly global average temperature was collected from Our World in Data, and drought index data was collected from drought.gov. The datasets were downloaded directly from the websites listed in the Data Sources section and put into the Original_Data folder.

b. Data Preprocessing

Downloaded cleaned and re-saved in excel to ensure readability from pandas. Cleaned datasets were put into the 'Cleaned Data' folder. The cleaned datasets were merged into one dataset titled Final_Data, having a total of 11 columns with the features we collected and start date of January 1, 1990, and end date of February 1, 2025. Raw price and index data for both wheat price and exogenous features were transformed into log return values, with factors such as cpi and average temperature remaining as original values. For deep learning models, lagged exogenous features were included, as well as sin and cos of month, aiming to capture seasonality in the machine learning and deep learning models. MinMax scaler was implemented to scale data for the LSTM model.

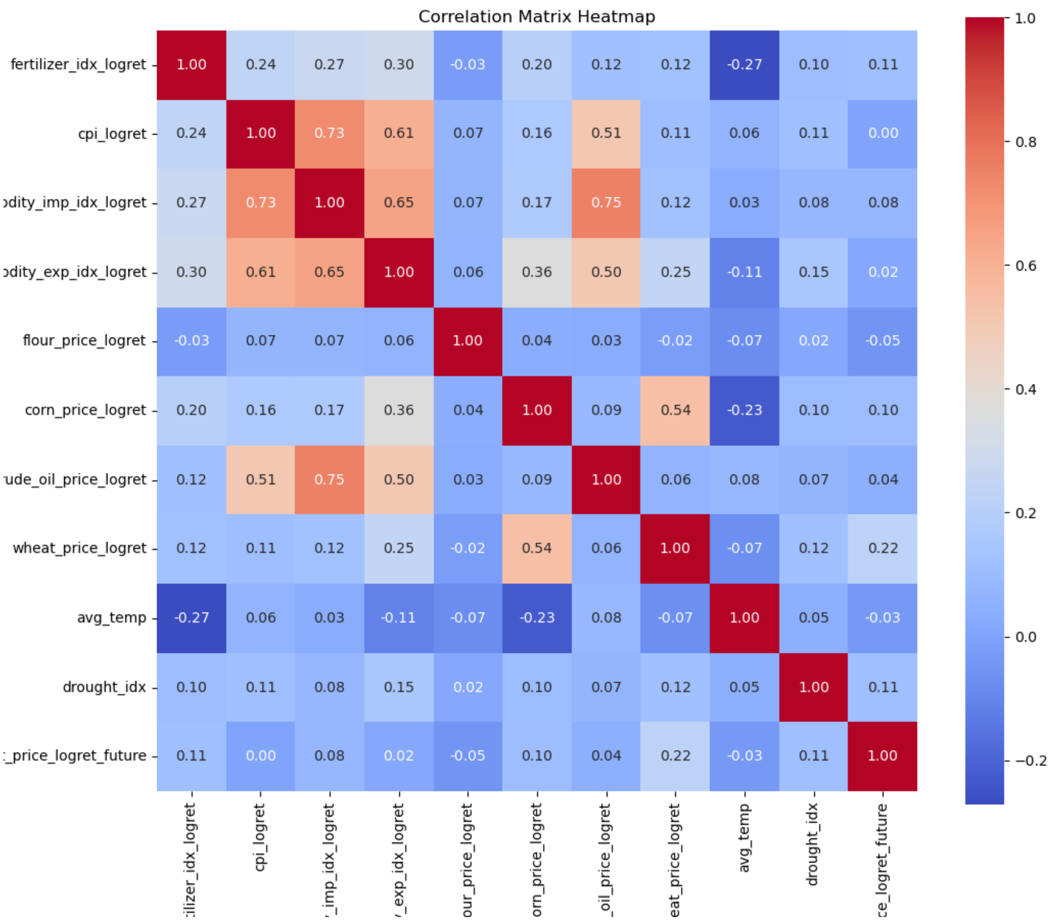

Figure 1: Final Dataset Correlation Heatmap

c. <u>Model Development</u>

Models were trained on data from 1990 to 2018 and tested on data from 2018 to 2024. The target variable was *wheat_price_logret_future,* the next month's wheat price log return. A simple Autoregressive model with exogenous variables was our first step. We then moved to incorporating a moving average factor, with ARIMA (including exogenous variables). Then, we tested SARIMAX, adding a seasonal aspect to the ARIMA model. For ARIMA and SARIMAX, the optimal parameters were selected through maximizing AIC. We also tested a VAR model, finding the best order based on AIC. Following the time series models, we tested a basic Decision Tree model. From there, we hyperparameter tuned and tested an XGBoost model to reduce overfitting present in basic Decision Tree models. Grid search was used to minimize MSE, finding the best model parameters. Finally, we tuned and tested an LSTM model to incorporate deep learning to more adequately predict time series data.

6. Experiments/Models Tested

- AR: used 4 lags and including exogenous variables.
- ARIMA: optimal order was found to be (1, 0, 2), including exogenous variables.
- SARIMAX: optimal order was found to be (1, 0, 2) x (1, 1, 1, 12).
- VAR: optimal order was found to be 1.
- Decision Tree: max_depth = 5.
- XGBoost: grid search on over hyperparameters found best parameters to be: colsample_bytree=0.8, learning_rate=0.01, max_depth=5, n_estimators=50, subsample=0.9
- LSTM: hyperparameter search across LSTM units (50, 75, 100), dropout rates (0.2, 0.3), and learning rates (0.001, 0.0005) found best parameters to be: 50 units, 0.3 dropout, 0.001 learning rate. Sequence length was set to 12 for yearly seasonal changes, and early stopping was implemented to avoid overfitting.

7. Results

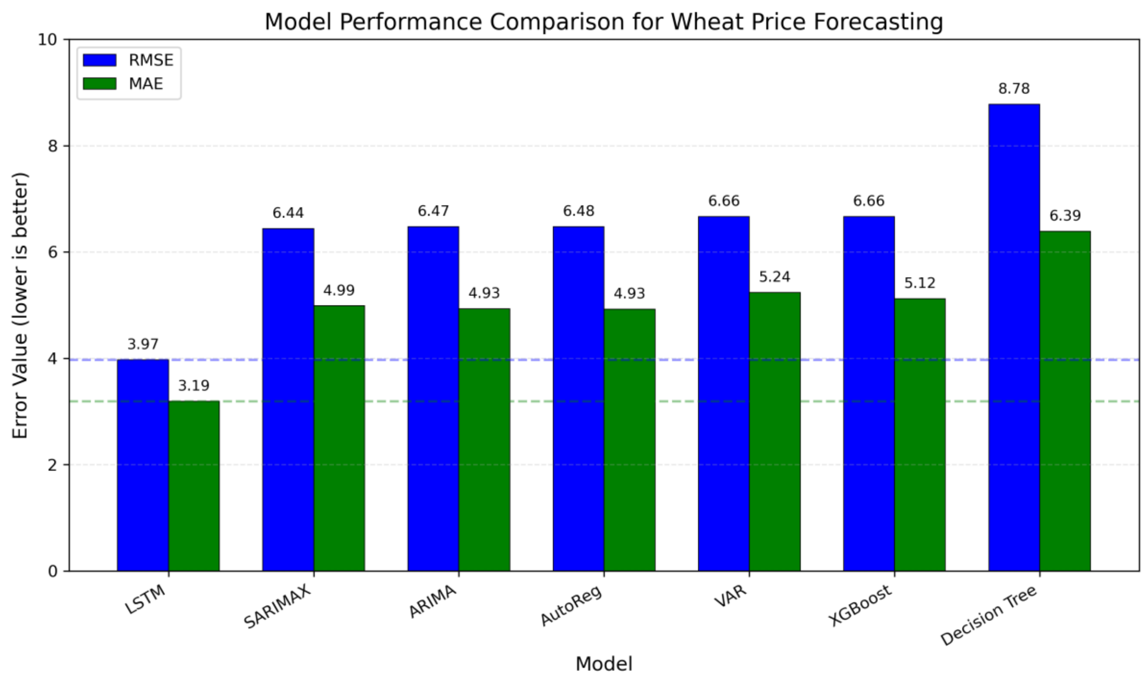

Figure 2: Model Comparison Chart

As seen in Figure 2, the worst performing model was the basic decision tree, with an RMSE of 8.78 and a MAE of 6.39. The basic decision tree fails to model the time series data well. However, when adding XGBoost, an ensemble of many trees, the errors reach similar levels to the time series forecasts, with an RMSE of 6.66 and MAE of 5.12. All four of the time series models tested show nearly identical RMSE and MAE values, suggesting that linear autoregressive models can capture some of the movement in wheat pricing. However, its predictive ability is limited and does not improve greatly with added moving average and seasonal aspects. All the time series models had RMSE values between 6.44 and 6.66 and MAE values between 4.99 and 5.24.

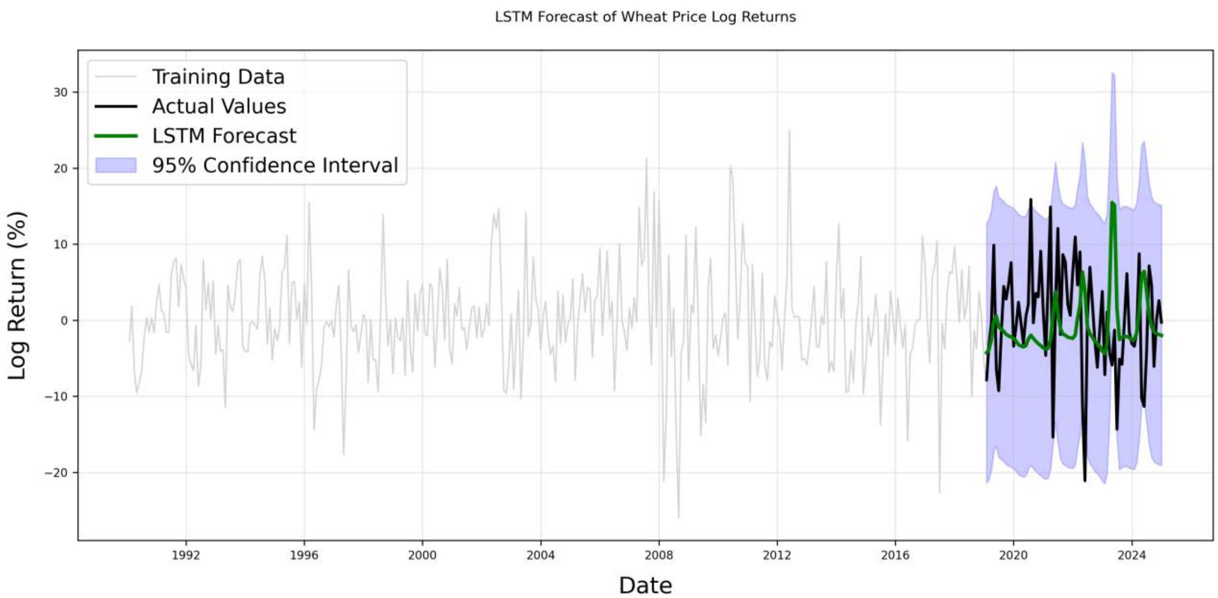

Figure 3: LSTM Model Forecast Results

The LSTM model significantly outperformed the others, achieving the lowest errors, with an RMSE of 3.97 and a MAE of 3.19. The RMSE value means that squared prediction errors are 3.97 percentage points. The MAE value means that predictions differ from the actual log return values by an average of 3.19 percentage points, in either direction. The improved performance is likely driven by non-linear relationships between features and the future log return of wheat price, which LSTM can uniquely account for. As can be seen in Figure 1, the target variable does not have strong linear correlation with the features, further emphasizing existence of more complex, nonlinear relationships. Despite the LSTM model's superiority to others tested, the model struggled to consistently predict the direction and magnitude of price changes. Figure 3 shows the substantial size of the 95% confidence interval, which is approximately between 15 and 20 percentage points. However, the confidence intervals contain nearly all of the actual data points.

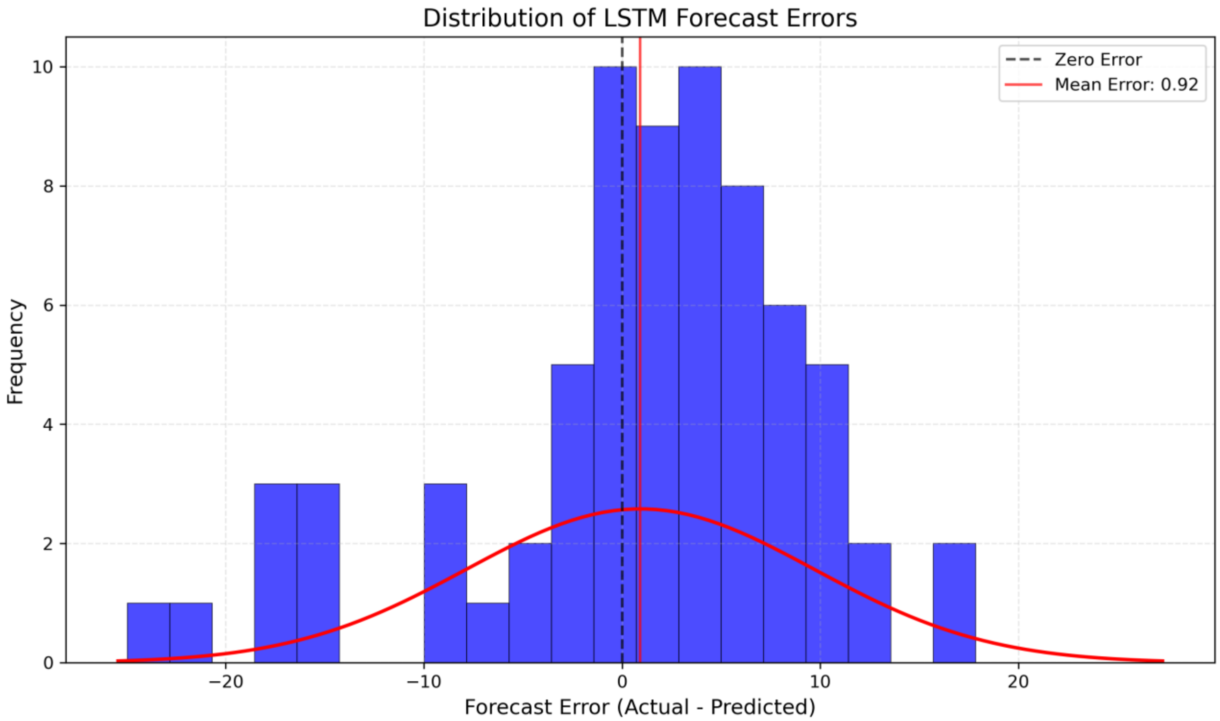

Figure 4: LSTM Error Distribution

Figure 4 shows the error distribution in our LSTM model. The mean error of 0.92 means that the model underestimates the log returns by an average of 0.92 percentage points. The distribution of errors is mostly normal with a slight right skewness. The errors clustering around 0 indicate that most predicted values mirror the actual direction of the log return. There are many outliers, with multiple predictions varying from the actual log return by over 20 percentage points, which further points to the unpredictable volatility present in commodity markets.

8. Lessons Learned

Through the implementation of and research for this project, we learned many lessons about machine learning modeling for time series data. Model selection was one of the main lessons. We learned a lot related to testing, comparing, and analyzing forecasts from autoregressive models for time series data. We also learned how to select machine learning and deep learning models to test, as well as learned how to implement an RNN model. Data preprocessing was utilized quite a bit, with prices being transformed, lagged and seasonal variables being added, and features being selected based on correlation. Hyperparameter tuning for new models was researched and learned for model improvement. Navigation of government datasets also took significant time to understand and interpret.

9. Challenges Faced

The main challenges faced are related to the data and long-term forecasting. Wheat prices vary significantly over a long period of time, making the challenge of multiple-year long predictions difficult. Even with the tuning of seven models and addition of exogenous variables that control for macroeconomic shifts, commodity market changes, and climate variation, the models were still limited in their long-term forecasting. Another issue is the significant variation in the log return of wheat prices during the testing period. Since the models were tested on data between 2018 and 2024, the COVID pandemic likely directly or indirectly caused some of the big spikes in price returns during 2020 and 2021. Although this could explain some of the outlier errors, there is still a struggle to predict the magnitude and direction of log returns. The data limited model performance significantly. The lack of relevant variables that have predictive power on wheat pricing proves very difficult for linear models to deal with and even causes the predictive struggle in deep learning models.

10. References

Darekar A and AA Reddy. 2018. Forecasting wheat prices in India. Wheat and Barley Research 10(1): 33-39. https://sawbar.in/wp-content/uploads/2018/07/75688-202627-1-PB.pdf

Brandt, Jon A., and David A. Bessler. "Price forecasting and evaluation: An application in agriculture." *Journal of Forecasting*, vol. 2, no. 3, July 1983, pp. 237–248, https://doi.org/10.1002/for.3980020306.

Sun, Feihu, et al. "Agricultural product price forecasting methods: A Review." *Agriculture*, vol. 13, no. 9, 24 Aug. 2023, p. 1671, https://doi.org/10.3390/agriculture13091671.

Wang, Yudong, et al. "Forecasting commodity prices out-of-sample: Can technical indicators help?" *International Journal of Forecasting*, vol. 36, no. 2, Apr. 2020, pp. 666–683, https://doi.org/10.1016/j.ijforecast.2019.08.004.

Westcott, Paul C., and Linwood A. Hoffman. "Price Determination for Corn and Wheat: The Role of Market Factors and Government Programs." *AgEcon Search*, 1 Jan. 1999, purl.umn.edu/33581.

11. Links to Download Data Sources

https://fred.stlouisfed.org/series/PWHEAMTUSDM

https://ourworldindata.org/grapher/monthly-average-surface-temperatures-by-year?tab=table

https://fred.stlouisfed.org/series/CPIAUCSL

https://fred.stlouisfed.org/series/IQ

https://fred.stlouisfed.org/series/IR

https://fred.stlouisfed.org/series/MCOILWTICO

https://www.drought.gov/historical-information?dataset=1&selectedDateUSDM=20250422&selectedDateSpi=20250301

https://fred.stlouisfed.org/series/APU0000701111

https://fred.stlouisfed.org/series/PMAIZMTUSDM