Manitoba Crop Prices Exploratory Data Analysis and Forecast Using Machine Learning

1st Ahmadreza Hajaghaie
Department of Computer Science
University of Manitoba
Winnipeg, Canada
Hajaghaa@myumanitoba.ca

2nd Melika Adabinejad
Department of Computer Science
University of Manitoba
Winnipeg, Canada
adabinem@myumanitoba.ca

***Abstract*—Agriculture is a fundamental part of the province of Manitoba's economy and, in general, the economy of Canada. We run an exploratory data analysis on the crop prices from 1988 in the province of Manitoba to obtain key features and correlations between the agricultural economy and other economic measures like CPI and Fertilizer prices. We dig deeper into the correlations and cover the geopolitical events that affect commodity and fertilizer prices, which can heavily affect the agricultural economy. We inevitably come to the knowledge that some macroeconomic metrics can affect the prices and change our forecasting perspective while we know these correlations. In this research, we focus on viewing the crop production and prices from various aspects and forecasting the price for upcoming years using Autoregressive Integrated Moving Average and LSTM. ARIMA is a statistical model that uses past values to predict the future values. LSTM is a deep learning model to learn sequential patterns including time-series patterns.**

***Index Terms*—Forecasting, Machine Learning, ARIMA, LSTM, Price prediction, Food Security**

I. Introduction

Agriculture is a vital part of Canada's economy; it contributes to various sectors in food supply chains. Canada's agriculture and agri-food sector is one of the most important economic drivers, providing employment and contributing to Canada's GDP[1]. In 2022, the agriculture ecosystem in Canada employed 2.3 million people. This sector also generated \$143.8 billion, which is about 7.0% of Canada's total GDP. This is an evident proof of the significance of this sector. Furthermore, the food and beverage processing sector also contributes enormously to Canada's economy. This sector in 2022 employed about 573,100 people and contributed about 3.4% to the total GDP of Canada. Food retailers and wholesalers also provided \$32.8 billion to GDP, which is about 1.6%. According to the official statistics from the government, Canada is home to 189,874 farms, which covers about 62.2 million hectares (6.2%) of Canada's land area. These farms are concentrated in the prairies, Quebec, and southern Ontario [1]. Total crop production in the commodity section of Canada is around \$30.6 billion, and a total of 118,300 jobs. This includes the total 65,135 farms till 2022. Top export markets of principal field crops are China, Japan, and the United States, which are 20.4%, 10.8%, and 9.6%, respectively. In our research, we dig into Manitoba's agriculture and analyze the crop prices from 1988 to 2023, which are based on a monthly price dataset provided by Manitoba's government. The top three crops and livestock commodities from 2018 till 2022 based on farm cash receipts are Canola, Wheat, and Hogs, which area 1.7\$ billion, 1.3\$ billion, and 1.25\$ billion, respectively. .Agriculture is one of the main pillars of the economy in Manitoba province. Agriculture in this province can directly or indirectly affect Manitoba's labor situation, total GDP, and trading. Agriculture in Manitoba directly contributes about 8 percent of the total GDP of the whole province. Based on the official statistics from the provincial government, Food and beverage processing is the largest manufacturing sector in Manitoba. In 2022, the total sales reached 7.3 billion dollars. The government of Manitoba releases statistical analyses annually reporting total sales, total production, created jobs, and some economic measures. Since the agriculture sector in Manitoba is one of the province's economic drivers, it can affect the total GDP heavily. Based on these statistics, the agriculture sector "directly contributes 8.0% of provincial gross domestic product (GDP) and 5.1 percent of provincial jobs or 36,355 direct jobs in 2022" [1]. Farm cash receipts (FCR) is a major metric that measures the gross revenue of farm businesses. This encompasses the sales of crop and livestock products, barring transactions between farms within the same province and program payments. The receipts are documented at the time when farmers receive the payment. In the sector analysis by the government of Manitoba, they mention that "Farm cash receipts (FCR) have risen strongly in recent years, with 2022 showing a 14.8% increase to \$9.75 billion. The largest sectors in Manitoba agriculture by FCRs are canola, wheat, and hogs." [1]. In 2022, the total net farm income reached a record-breaking \$2.81 billion, with net cash income being the second highest ever at \$2.52 billion, only surpassed by 2021. The food manufacturing sector has seen consistent growth over the years, with total sales in 2022 hitting \$8 billion, making up 5.9% of Canada's total food manufacturing sales. Manitoba recognized globally as a reliable source of safe, top-quality grains, oilseeds, livestock, and agri-food products, set a record in 2022 with its agriculture and agri-food exports reaching an all-time high of \$8.82 billion. This represents a 72%

[1]Gross domestic product

increase in agricultural exports since 2013, accounting for 43% of the province's total exports in 2022. The top five major agri-food exports from Manitoba in 2022 were wheat, pork, canola oil, canola seeds, and prepared potatoes, which made up nearly 65% of total agri-food exports. Manitoba province also exports a significant amount of soybeans, oats and oat products. [1]. In this research, we intend to analyze the prices of the main crops being harvested and used in Manitoba. These crops are Flaxseed, Canola, Soymeal, Canola meal, Soybeans, Wheat, Oats, Corn, Barley, Peas, and Rye. Most of the market distribution belongs to Canola, Wheat, and Soybeans. Based on the crop production and harvested area statistics, the highest production of a crop belongs to wheat, 4,758,248 metric tonnes. Price-wise, the list would be different; the top three expensive crops are Canola, Soymeal, and Flaxseed. The prices of most crops did not change much from 1988 to 2020; they moved relatively linearly, and the price fluctuations were not so severe. But in 2020, the pandemic of COVID-19 caused the prices to rise significantly, mostly due to the effects of COVID-19 on the supply chain, the price of fertilizer, and other systematic factors that occurred during the lockdowns and the pandemic in general. In the first part of our work, we try to analyze the price changes with graphs and data analysis techniques so that the changes become more tangible for us, and we can correlate them with other economic metrics, such as the price of other commodities such as industrial fertilizers, and macroeconomic indicators. With these comparisons and correlation analysis, we can conclude that macroeconomic metrics and systematic risks can affect prices and change our forecasting perspective. In the second part of the research, we try to use two machine learning algorithms to analyze price fluctuations and predict them for the coming year. We used ARIMA and LSTM techniques in this research, which are suitable for time series analysis and forecasting.

II. Related Work

Many papers have been published over the years, analyses including crop types, their prices, and critical factors impacting productivity. The origin of these articles is usually the countries that are the leading producers and exporters of agricultural products, such as India and the USA. The data that countries like India have at their disposal are data that, in addition to product prices, have much additional information such as weather, exact harvest rate, soil quality, soil fertility, etc. Agricultural data of the province of Manitoba are only price data and annual or weekly cultivation amount of products, and accurate and clean data of soil and weather are not available and are randomly found in the agriculture section of the Manitoba government website. This lack of data makes it difficult for farmers to make informed decisions about planting and harvesting. Furthermore, it limits the potential of Manitoba's agricultural sector. Suppose more accurate data from the Manitoba government is made available to the public. In that case, it will be possible to perform a more accurate and optimal data analysis, and farmers and analysts in this field can do their work more intelligently. In addition to analyzing and checking price factors, much research can be done in the field of agriculture besides forecasting crop prices. On the other hand, it is important to demonstrate to the public the power of data and encourage governments to report transparent data as much as possible to democratize data. In the following parts, we will discuss how crop prices were affected by some geopolitical events and macroeconomic metrics and study the correlations between them. Recent studies show that a geopolitical even like War in some parts of the world can affect specific industries and economies on the other side of the world. The ongoing conflict between Russia and Ukraine endangered the economic stability of some countries and their economic core, which is based on agriculture. Arndt et al. [2] review that the Russia-Ukraine war has significantly impacted global food, fuel, and fertilizer prices, worsening poverty and food insecurity worldwide. Based on their models, their study shows the crisis's effects on agriculture-based systems in 19 countries. The crisis has pushed an additional 27.2 million people into poverty and 22.3 million into hunger. Rising fuel and fertilizer prices primarily affect agri-food systems and poverty, while higher food prices impact hunger and diet quality. Nasir et al. [3] review of the impact of the Russia-Ukraine conflict on global food security notes that the War has disrupted Ukraine's food production, leading to a significant drop in wheat, soybean, and corn production in 2022-2023. The conflict has also affected the global supply chain and food trade, increasing global food prices from March to May 2022. Their study concludes with a warning that ongoing conflicts may hinder the achievement of a sustainable agricultural ecosystem. Sokhanvar and Bouri [4] examine the impact of commodity price shocks, specifically wheat, crude oil, and natural gas, on the Canadian dollar, Euro, and Japanese yen (EUR/CAD and CAD/JPY) in the context of the War in Ukraine. Their study finds a long-term association between higher commodity prices and the appreciation of the Canadian dollar against both the Euro and yen. The analysis also reveals that commodity price shocks positively impact the Canadian dollar's value. They utilized the quantile autoregressive distributed lag (ARDL) model, which demonstrated a relationship between increasing commodity prices and strengthening Canadian dollar pairs. Kantanantha et al. [5] introduce a weather-based regression model with time-dependent varying coefficients. They devised a solution to overcome the difficulty of predicting results in a correlated space. They reduced the space of predictors to a small number of uncorrelated predictors using Functional Principal Component Analysis (FPCA, they formulated a model based on futures for long-term cash price forecasting. Their model estimates the cash price as a combination of the nearby futures settlement price and the projected commodity basis as a mixture of historical basis data using a functional model-based approach. Cheung et al. [6] introduce a 3D-CNN model for improving crop price prediction. Their model can handle non-stationary data and learn non-linearity. Traditional forecasting methods like ARIMA fall short due to the complex, multi-dimensional factors influencing crop prices, including environmental changes and economic factors. Their study also reviews factors influencing crop production and price changes. Their proposed 3D-CNN model outperforms ARIMA regarding crop price prediction. As mentioned in the research, The data provided by the government of Manitoba needs to include some critical information regarding weather and soil conditions. By these factors, the prediction may be improved. Shao and Dai [7] worked on an ARIMA-based model for food crop price prediction. Their study ”incorporates an ARIMA as the FSM for computational intelligence (CI) models to predict three major food crop (i.e., rice, wheat, and corn) prices” [7]. They also integrated some forecasting models, including support vector regression (SVM) and multivariate adaptive regression splines (MARS). Their integrated model performed better regarding food crop prediction.

Sellam and Poovammal [8] used regression analysis to study the impact of environmental factors like Area under Cultivation, Annual Rainfall, and Food Price Index on crop yield from 1990-2000. Linear Regression establishes a relationship between these factors and crop yield.

Alexander et al. [9] analyzed the impact of increased agricultural input costs and export restrictions from Russia and Ukraine on global food prices, health, and the environment. Their model predicts that these factors could raise food costs by 60-100% in 2023. Recent studies, current data, and the current situation prove this. They also mention that restoring food trade from Ukraine and Russia alone won’t solve the food insecurity problem caused by higher energy and fertilizer prices. Kaur et al. [10] discuss the use of data mining in agricultural crop price analysis. They utilize various data mining techniques like K-Means, K-Nearest Neighbor, Artificial Neural Networks, and Support Vector Machines for crop price prediction. Mulla and Quadri [11] focus on predicting crop prices using past data patterns and consider factors such as rainfall, temperature, market prices, land area, and past yield. They used the Decision Tree algorithm, which is a supervised machine learning method, to make predictions. Ghutake et al. [12] propose an intelligent Crop Price Prediction Model using Machine Learning. They used Random forest regression and decision tree regression algorithms to predict the crop prices. Mohanty et al. [13] review a machine learning framework for crop price prediction, consisting of four blocks: crop yield prediction, supply determination, demand prediction, and crop price prediction. Various time series algorithms are used for forecasting, and the decision tree regressor is found to be the best model for predicting crop prices. Furthermore, Chaitra and Meena [14] used Autoregressive Integrated Moving Average (ARIMA), Decision trees, Long Short-Term Memory (LSTM), K Nearest Neighbor (KNN) etc. for accurate Yield and Crop Price Prediction. Thapaswini and Gunasekara [15] discuss the use of machine learning, specifically decision trees and neuro-evolutionary algorithms, to address agricultural issues like yield recommendations and crop price forecasting. Penone et al. [16] analyze hedging effectiveness in futures markets for Italy’s soybean, corn, and milling wheat producers to reduce income risks. They used (naïve, OLS, GARCH) to estimate the hedge portfolio and assess income risk reduction. Lastly, Cenas [17] reveals the potential to improve the accuracy of time series models in forecasting future rice prices by combining ARIMA and the Kalman filter technique. The combined ARIMA-Kalman filter model showed more accurate and precise estimates than the typical ARIMA model.

III. Background

A. ARIMA

ARIMA, which is the short name for the autoregressive integrated moving average, is a generalized version of the ARMA model. The generalization is pointed out by the word integrated in the name to mention that in ARIMA unlike ARMA, we can model a non-stationary model because the model itself will difference the mean function (or the observed trend) to eliminate the non-stationarity[23]. The differencing is done in the initial phase how many times it is needed. The autoregressive or AR part of the model means that the variable is being modelled or regressed based on its own lagged version. AR captures the dependency between the values of the variable and previous observations. The moving average or MA part of the model is added to account for the short-term dependencies that AR does not capture. This is done by calculating the dependency between the values of the variable and residual errors of the moving average that is applied to previous observations. These different parts of ARIMA can be tuned as ARIMA(p,d,q). p determines the lag order for AR, d determines the order of differencing, and q is the order of MA. The mathematical equation of ARIMA is defined as:

$$\left(1 - \sum_{i=1}^{p} \varphi_i L^i\right) (1 - L)^d X_t = \delta + \left(1 + \sum_{i=1}^{q} \theta_i L^i\right) \varepsilon_t$$

B. LSTM

LSTM is short for long short-term memory which is an extension to RNN models. It can capture long-term dependencies that RNNs fail to do because of the vanishing gradient problem or exploding gradient problem. This essentially means each certain input that is being fed forward through the hidden layers will either have too small or exploding influence when it is circling through RNN[24] (the magnitude of its gradients). LSTM was introduced to address the vanishing gradient problem. LSTM can perform a variety of tasks that involve processing a sequence including time-series modeling. In short, the idea behind LSTM is a design that brings memory cells, gate mechanisms, and cell state into the picture. At each timestep, input is activated in the LSTM cell. Then, the forget gate decides what information is unnecessary in the input based on the last cell’s hidden output and this new input. The input gate decides what new information must be added to the cell’s state based on the same two components. Finally, the output gate calculates an output based on the updated cell state.

IV. Proposed Work

A. Purpose

This research tries to investigate the factors affecting the price of crops from 1988 to 2022. The price trend gives us a view of what has affected the price over the years and how systematic risks can affect the prices and endanger food security. Also, there needs to be a forecasting report about the prices of fertilizers on Manitoba's government website, and the absence of this would be a problem for farmers. If farmers had a broader view of the economic situation ahead, such as the price of chemical fertilizer, they could make better decisions; this action would lead to increased productivity, production, and risk mitigation. As mentioned in the related work section, if the price data of the province of Manitoba had other statistical data such as weather, soil fertility, and soil quality, we could conduct more extensive research. This preliminary study contributes to understanding what is lacking in the public scenes of agricultural data analysis of the province of Manitoba and bases our future works on creating an informative dashboard for the public. It also helps us ask for more data from the government and MASC (Manitoba Agricultural Services Corporation).

B. Proposed solution

In the first phase of our research, we examine the price data of Manitoba's agricultural products (crops) from 1988 until now and see what factors have affected the prices of products. In the first phase, we analyze and visualize the data with the help of Python programming language and libraries such as Pandas, Plotly, and Matplotlib. This action allows us to see how the prices of the products change and how they have changed over the years. Visualizing price fluctuations gives us an idea of how the price trend has changed and what factors have caused price increases or decreases over the years. Furthermore, we review the external factors in detail. In the second research phase, we try to predict the price with the same historical price data with machine learning techniques. We do this for one of the crops and visualize the result. We utilize machine learning and statistical computation in this phase.

C. Dataset and pre-processing

The data set we are using for the research is the historical price dataset of crops in Manitoba province, an open-access dataset on the Government of Manitoba website. This dataset contains historical information about the prices of different harvested crops in the province of Manitoba from 1988 to 2022. The data contains only a few months of 1987, and it is not in the full-year time frame, so we decided to delete those values for cleaner data. We also used different datasets to determine the correlations and macroeconomic metrics. We used the CPI of Manitoba, CPI of Canada, Fertilizers price trend from 2019 till 2023, Urea price, S&P 500 Fertilizers & Agricultural Chemicals Sub-Industry Index, which is the main index for chemical fertilizers, Monthly Bank of Canada commodity price index in Agriculture sector, and Monthly Bank of Canada commodity price index.

	Year	Month	UOM	Crop	Value
0	1988	1	CAD/tonne	Barley, #1CW	51.0000
1	1988	1	CAD/tonne	Canola Meal, 34%, Altona	226.8425
2	1988	1	CAD/tonne	Canola, #1CR	260.3450
3	1988	1	CAD/tonne	Flaxseed, #1CW	186.4750
4	1988	1	CAD/tonne	Oats, #2CW	87.1125

Fig. 1. Manitoba Historical Crop Prices, Data by Government of Manitoba

For the first phase of the research, in order to do an exploratory data analysis on the data set, we need to clean our data first and import the necessary libraries. The environment we are using for EDA is Jupyter Notebook and Python 3 programming language. The libraries we used for visualization are Pandas, Numpy, Plotly, Seaborn, Klib, and Matplotlib. Some of the plots are interactive, and it is available on the GitHub [2]. For the second phase, we chose ARIMA and LSTM, two classic models, to perform less advanced and preliminary forecasts based on the historical crop prices of Manitoba. In other words, our main focus in this preliminary study was to find a univariate model. We want to forecast what the price is for the next time step and how precise would the prediction be without accounting for all the important correlations that we observed closely in the previous sections. The government of Manitoba provides historical monthly and weekly crop prices from late 1987 to the end of 2022. There are 14 different grains in this dataset. Many of the crops in this dataset have unique characteristics of their own. This discrepancy requires the analyst to design a different model and do different analyses to capture what is influencing this time series and what are the components. This is another reason that we only conduct experiments on Canola. Moreover, we are forced to work with monthly data since weekly data has missing values and missing timestamps in several recent years. We performed spline interpolation to replace NaN values. Some crops missed timestamps from one early year, so we deleted those years. However, some missed many timestamps from 2020 to 2022 which is not desirable for forecasting into 2023.

We augmented a dataset that contains other data like temperature, precipitation, food index of CPI of Canada, grain stocks of Canada, seeded and harvested areas of Manitoba, and yields in Manitoba. Temperature and precipitation directly affect the yields each year. In addition, the area seeded and area harvested implicitly indicate how the soil is performing, being other factors that directly affect the yields. Yield and stocks of grains in composition form the supply of the market and CPI represents demand. For CPI we only use the food index, since crops mostly affect this index. For example, Canola is used for producing oil that is healthy and affordable, while it also produces Canola meal which is protein for live stock from which meat is produced. This dataset is solely what we have augmented from public datasets of the government of Canada and we have made the said dataset available on our GitHub

[2]https://github.com/TheSlayerr/Manitoba_Crop_Prices

TABLE I
AUGMENTED DICKEY-FULLER

ADF Statistic	-1.751
p-value	0.405
Critical Values	'1%': -3.446, '5%': -2.869, '10%': -2.570

for anyone who wants to extend our work utilizing our EDA. As Canola has the most share in terms of monetary value in the province of Manitoba, we did the main experiments and analysis for Canola.

D. Training setup

We report the results of our experiments with RMSE and MAPE metrics with walk-forward or rolling validation since in the future the model will be working in the same manner.

1) ARIMA: For all crops, we performed a grid search to find desirable ARIMA configuration:

- p: 0, 1, 2, 10, 20
- d: 0, 1, 2
- q: 0, 1, 2
- Train-Test split is 0.6-0.4

These numbers are based on ACF (Fig. 2.) and PACF (Fig. 3.) plots for Canola. for p we consider at which lag the plot cuts through the confidence surface and for the q we look for the same thing on PACF. Furthermore, we fed data to ARIMA after a log transformation to ameliorate the severe price rise of a few recent years. Doing the augmented Dickey-Fuller test, we observed Canola's time-series is not stationary, and chose to have none-zero values for d.

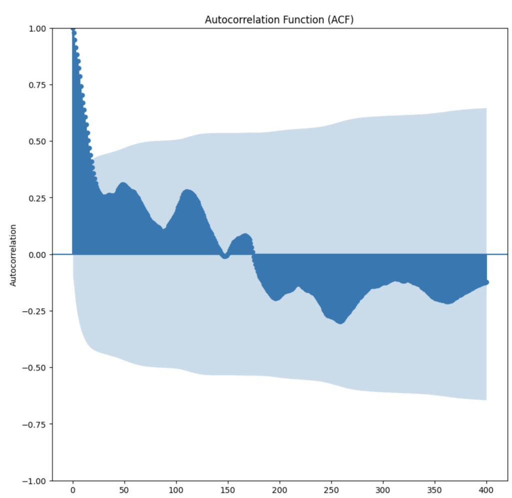

Fig. 2. Autocorrelation Function

2) LSTM: To achieve better results, we used the differencing technique with an order of one on our data before feeding it to the model. We also performed MinMax Scaling.

Our setting for LSTM is:

- 150 epochs

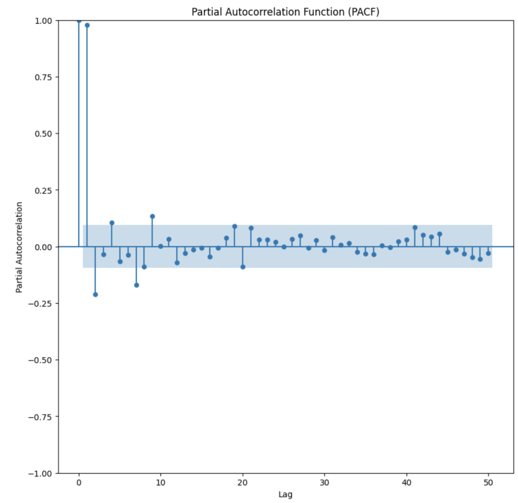

Fig. 3. Partial Autocorrelation Function

TABLE II
ARIMA FOR ALL CROPS

Crop	(p,d,q)	RMSE	Mean Price
Barley	(2,2,1)	11.176	136.845
Canola	(0,1,2)	29.151	401.812
Canola Meal	(10,1,2)	24.584	242.273
Corn	(2,1,1)	12.042	167.589
Flaxseed	(0,1,2)	47.180	406.787
Oat	(10,1,0)	20.698	166.518
Pea	(1,1,0)	21.355	218.120
Rye	(2,1, 0)	26.528	127.983
Soybean	(0,2, 2)	22.448	392.908
Wheat, Northern Hard Red	(0,1, 1)	28.861	278.338
Wheat, Red Winter	(20,1, 0)	24.010	232.738
Wheat, Western Red Spring	(10,2, 0)	23.095	269.807
Wheat, Special Purpose	(0,1, 1)	14.314	157.701

- Batch Size 1
- 6 LSTM Layers each 30 units
- A Dropout Layer after each LSTM with a rate of 20%
- One Dense Layer with 32 units at the end
- Adam optimizer/ MSE Loss
- Train-Test split is 0.8-0.2

As is evident in TABLE II, capturing every temporal and exogenous pattern for each crop has its complicated challenges and requires meticulous analysis. Some of these configurations might indicate there is quite much noise in the observation or an absence of predictable cyclical patterns. For example, while we can achieve a rather good RMSE for Canola with a mean price of 401.812, it is not easy to do the same for Flaxseed. In TABLE III, ARIMA fits the observations better than LSTM, although LSTM is a powerful model. This is because size of data is small (420 records). One way to improve the performance of LSTM is to add other features to have more precise predictions. Predictions are compared to observations in the Fig. 4.

TABLE III
RESULTS

Model	RMSE	MAPE
ARIMA	29.151	3.373
LSTM	35.814	3.459

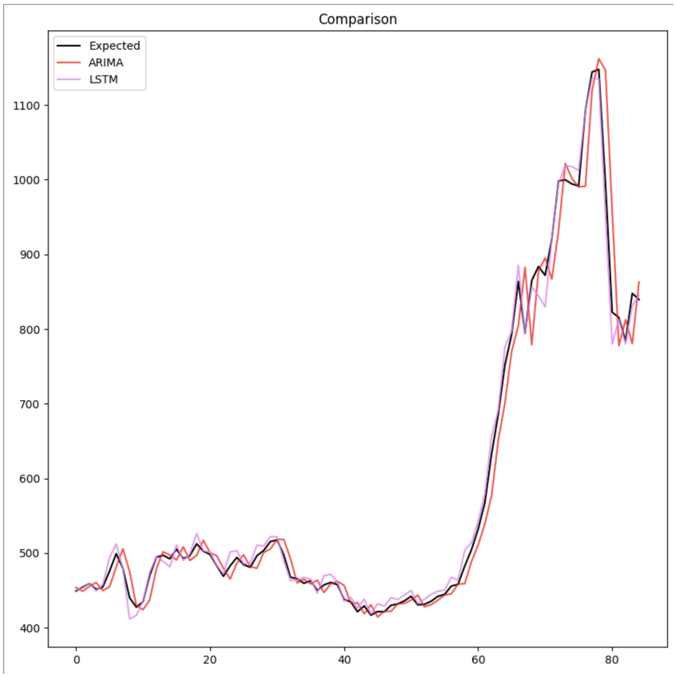

Fig. 4. Predictions of ARIMA and LSTM

V. Results and Discussion

In the first phase of our research, we investigated the price changes from 1988 to 2022. After the food crisis in 2008, till the beginning of 2019, the price was almost following a linear trend without any significant changes. Still, in 2019, after the start of the COVID-19 pandemic, we can clearly see how the prices of crops changed. The rise in prices in the pandemic is due to the damage caused by the Covid lockdowns to the supply chains of different sectors. These sectors include Agriculture, Agri-food processing, transportation, and many more. During the COVID-19 pandemic, the lack of balance between supply and demand caused the price of products to rise. COVID-19 has impacted global food prices due to supply chain disruptions and changes in consumer spending. Outbreaks in food processing facilities led to closures, causing supply constraints and increased prices. These shutdowns also resulted in processing bottlenecks and higher input costs. Specifically, closures of beef processing facilities led to higher beef prices in mid-2020 [14]. Also, weather events have impacted food supply and prices. In 2021, heatwaves and droughts in the Prairie provinces led to higher meat and grain prices. Drought conditions increased the cost of livestock due to a rise in animal feed demand and reduced grain outputs. These conditions also caused the prices of imported food in Canada to be elevated, contributing to higher meat prices in grocery stores. [14] In Fig.5 and Fig.6, we can see the fluctuations in crop prices from 1988 to 2022. Fig.5 is the price trend of canola, which is one of the most important harvested crops in the province of Manitoba.

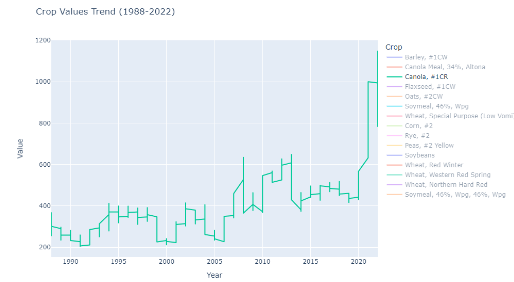

Fig. 5. Canola price trend from 1988 - 2022

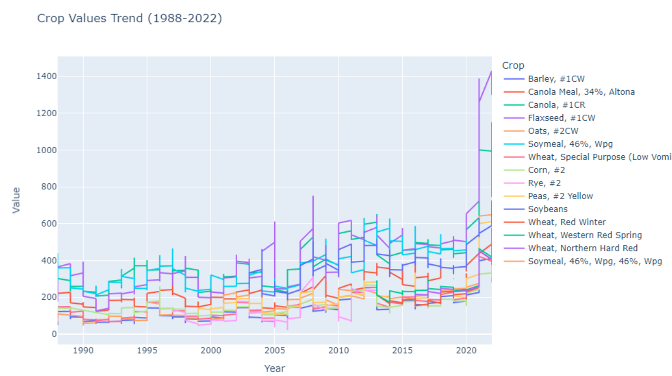

Fig. 6. Crop Values Trend From 1988-2022

For further illustration, we analyzed the price of fertilizers and Urea, which is one of the main chemical fertilizers that are being used for farms worldwide. Also, as you can see in the figure below, we have shown the Fertilizer price index, which clearly shows the correlation between the fertilizer prices and the prices of crops.

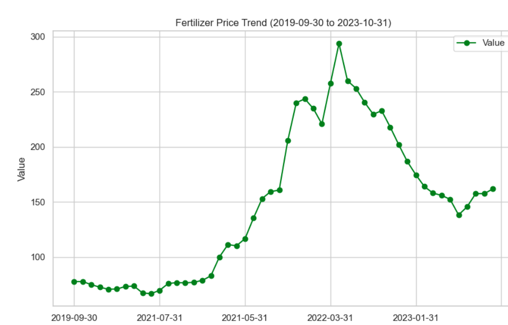

Fig. 7. Fertilizer Price trend from 2019 - 2023, Data: World Bank

The main trading index of fertilizers and chemicals related to agriculture is the S&P 500 Fertilizers & Agricultural Chemicals Sub-Industry Index which is a capitalization-weighted index. The index is designed to measure the performance of companies in the sub-industry group of fertilizers and

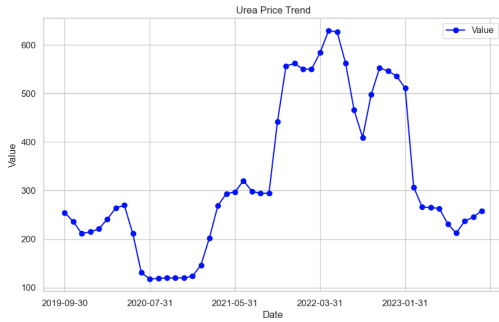

Fig. 8. US Urea Spot Price (Gulf) 2019 - 2023, Data: International Monetary Fund

agricultural chemicals. The index is unmanaged and can not be directly purchased by investors or retailers. Many factors can influence the performance of this index, for instance, the price of fertilizer, the price of crops, food demand, commodity prices, supply chain restrictions, and, most importantly, geopolitical events. In 2021, the price of fertilizer and crops increased during the pandemic and lockdowns. As can be seen on the plots, we see about a 27% rise in the price of fertilizer and around 20% rise in the price of crops in 2021. But the most important factor that most farmers are unaware of is the geopolitical events. The moment that the geopolitical events change, the price of fertilizer and crops will change as well. Because of the Ukraine crisis, the price of fertilizer and crops has increased. This is mainly because Ukraine is one of the largest producers of crops in the world, and the price of crops is directly related to the price of fertilizers. Ukraine is ranked 1st in the world to produce sunflower seeds, 3rd for barley and rapeseed, 4th for rye and soybeans, 5th for wheat and corn, and 8th for oats. Also, Ukraine is the world's largest exporter of sunflower oil, the 3rd largest exporter of corn, the 4th largest wheat exporter, and the 7th largest exporter of soybeans [15].

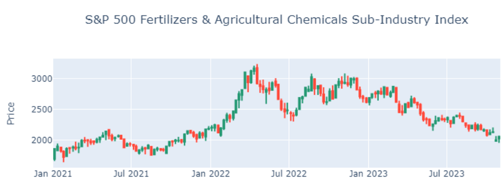

Fig. 9. S&P 500 Fertilizers & Agricultural Chemicals Sub-Industry Index, Data: NASDAQ

The conflict between Russia and Ukraine has escalated global food prices due to their significant roles in agriculture and oil production. Russia's position as a leading producer of key fertilizer components, coupled with sanctions on Belarus, has impacted fertilizer prices. Both countries are major wheat exporters, leading to price volatility in global wheat markets. Wheat prices started rising in December 2021 due to war concerns and have continued to increase, with cereal product prices up by over 11% year over year for eight months. As the conflict continued, cereal prices remained high in October [14].

Lastly, we compared the consumer price index of Canada and Manitoba to the prices of crops and fertilizers. In the results, it is evident that the CPI and prices have a strong correlation. CPI measures the average change in the prices consumers pay for their basket of goods and needs. The price of fertilizers and crops affects the total CPI of Canada and the CPI of Manitoba [16][17]; if you compare the price trend of crops from 1988 till 2022 to Fig.10 and Fig.11, it demonstrates the correlation between the price trend and CPI in the province of Manitoba.

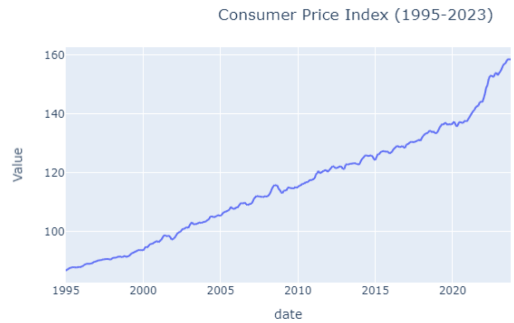

Fig. 10. Total Consumer Price Index of Canada, Data: Bank of Canada

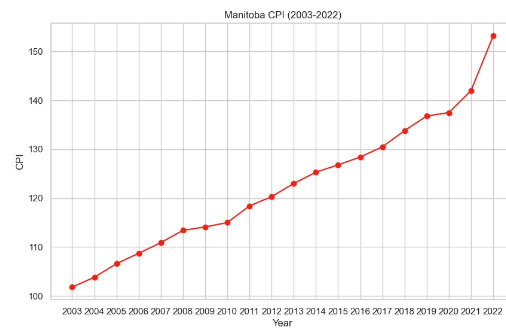

Fig. 11. The Consumer Price Index of the Province of Manitoba 2003-2022, Data: Statistics Canada

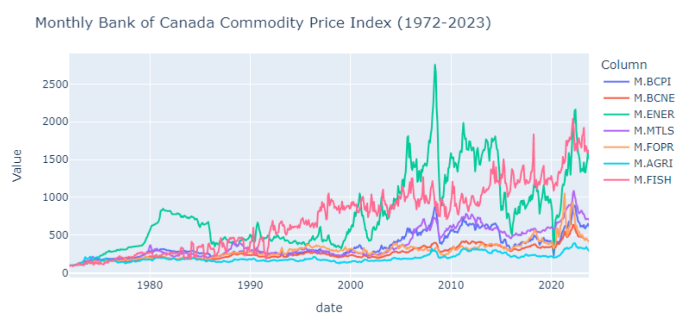

Fig. 12. Bank of Canada Commodity Price Index - Monthly

VI. Limitations and Future work

Given the results of our analysis and experiments, exogenous factors like supply and demand chains can be important to help the models predict precisely. Farmers and governments can also trust a system's forecasting in making informed decisions when these factors are integrated into the models. We hope to inform administrative decisions about public datasets for soil fertility, fertilizer usage, and crop rotations for the province of Manitoba. Our augmented dataset could use more precise approaches and data to augment weather data. We would also like to add oil seed-crushing data and do more analysis on seed supply, canola meal market and its relation to livestock productions of Manitoba in our future works. We did this analysis with hybrid multi-variate models for our future work and modelling more crop items to create an online interactive explainable dashboard on our mind.

REFERENCES

[1] A. and A.-F. Canada, "Overview of Canada's agriculture and agri-food sector." Accessed: Dec. 20, 2023. [Online]. Available: https://agriculture.canada.ca/en/sector/overview

[2] C. Arndt, X. Diao, P. Dorosh, K. Pauw, and J. Thurlow, "The Ukraine war and rising commodity prices: Implications for developing countries," Global Food Security, vol. 36, p. 100680, Mar. 2023, doi: 10.1016/j.gfs.2023.100680.

[3] M. A. Nasir, A. D. Nugroho, and Z. Lakner, "Impact of the Russian–Ukrainian Conflict on Global Food Crops," Foods, vol. 11, no. 19, Art. no. 19, Jan. 2022, doi: 10.3390/foods11192979.

[4] A. Sokhanvar and E. Bouri, "Commodity price shocks related to the war in Ukraine and exchange rates of commodity exporters and importers," Borsa Istanbul Review, vol. 23, no. 1, pp. 44–54, Jan. 2023, doi: 10.1016/j.bir.2022.09.001.

[5] N. Kantanantha, N. Serban, and P. Griffin, "Yield and Price Forecasting for Stochastic Crop Decision Planning," JABES, vol. 15, no. 3, pp. 362–380, Sep. 2010, doi: 10.1007/s13253-010-0025-7.

[6] L. Cheung, Y. Wang, A. S. Lau, and R. M. Chan, "Using a novel clustered 3D-CNN model for improving crop future price prediction," Knowledge-based systems, vol. 260, pp. 110133-, 2023, doi: 10.1016/j.knosys.2022.110133.

[7] Y. E. Shao and J.-T. Dai, "Integrated Feature Selection of ARIMA with Computational Intelligence Approaches for Food Crop Price Prediction," Complexity, vol. 2018, pp. 1–17, Jul. 2018, doi: 10.1155/2018/1910520.

[8] V. Sellam and E. Poovammal, "Prediction of Crop Yield using Regression Analysis," Indian Journal of Science and Technology, vol. 9, no. 38, Oct. 2016, doi: 10.17485/ijst/2016/v9i38/91714.

[9] P. Alexander, A. Arneth, R. Henry, J. Maire, S. Rabin, and M. D. A. Rounsevell, "High energy and fertilizer prices are more damaging than food export curtailment from Ukraine and Russia for food prices, health and the environment," Nat Food, vol. 4, no. 1, Art. no. 1, Jan. 2023, doi: 10.1038/s43016-022-00659-9.

[10] M. Kaur, H. Gulati, and H. Kundra, "Data Mining in Agriculture on Crop Price Prediction: Techniques and Applications," IJCA, vol. 99, no. 12, pp. 1–3, Aug. 2014, doi: 10.5120/17422-8273.

[11] S. Mulla and S. Quadri, "Crop-yield and Price Forecasting using Machine Learning," The International journal of analytical and experimental modal analysis, vol. XII, pp. 1731–1737, Aug. 2020.

[12] I. Ghutake, R. Verma, R. Chaudhari, and V. Amarsinh, "An intelligent Crop Price Prediction using suitable Machine Learning Algorithm," ITM Web of Conferences, vol. 40, pp. 3040-, 2021, doi: 10.1051/itmconf/20214003040.

[13] M. K. Mohanty, P. K. G. Thakurta, and S. Kar, "Agricultural commodity price prediction model: a machine learning framework," Neural Comput & Applic, vol. 35, no. 20, pp. 15109–15128, Jul. 2023, doi: 10.1007/s00521-023-08528-7.

[14] B. Chaitra and K. Meena, "Forecasting Crop Price using various approaches of Machine Learning," in 2023 International Conference on Innovations in Engineering and Technology (ICIET), Jul. 2023, pp. 1–5. doi: 10.1109/ICIET57285.2023.10220616.

[15] G. Thapaswini and M. Gunasekaran, "A Methodology for Crop Price Prediction Using Machine Learning," in 2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC), Dec. 2022, pp. 1–7. doi: 10.1109/ICMNWC56175.2022.10031852.

[16] C. Penone, E. Giampietri, and S. Trestini, "Hedging effectiveness of commodity futures contracts to minimize price risk: Empirical evidence from the Italian field. Crop sector," Risks (Basel), vol. 9, no. 12, pp. 1–14, 2021, doi: 10.3390/risks9120213.

[17] P. V. Cenas, "Forecast of Agricultural Crop Price using Time Series and Kalman Filter Method," vol. 5, no. 4, 2017.

[18] C. Elleby, I. P. Domínguez, M. Adenauer, and G. Genovese, "Impacts of the COVID-19 Pandemic on the Global Agricultural Markets," Environ Resource Econ, vol. 76, no. 4, pp. 1067–1079, Aug. 2020, doi: 10.1007/s10640-020-00473-6.

[19] S. C. Government of Canada, "Behind the Numbers: What's Causing Growth in Food Prices." Accessed: Dec. 20, 2023. [Online]. Available: https://www150.statcan.gc.ca/n1/pub/62f0014m/62f0014m2022014-eng.htm

[20] "How the Ukraine-Russia Conflict Will Raise the Price of Snack Foods," TIME. Accessed: Dec. 20, 2023. [Online]. Available: https://time.com/6155095/sunflower-oil-russia-ukraine/

[21] “Consumer price index, Bank of Canada.” Accessed: Dec. 20, 2023. [Online]. Available: https://www.bankofcanada.ca/rates/price-indexes/cpi/

[22] Consumer Price Indexes - Statistics Canada. Statistics Canada. Accessed [Dec. 20, 2023 [Online].

[23] “Autoregressive integrated moving average,” Wikipedia. Oct. 09, 2023. Accessed: Dec. 18, 2023. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Autoregressive_integrated_moving_average&oldid=1179339959

[24] A. Graves, M. Liwicki, S. Fernández, R. Bertolami, H. Bunke, and J. Schmidhuber, “A Novel Connectionist System for Unconstrained Handwriting Recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 5, pp. 855–868, May 2009, doi: 10.1109/TPAMI.2008.137.