Forecasting the Stock Price of Cal-Maine Foods

MGT 6203 Group Project Final Report - **Team 52**

Sumner Rieland

Andrew Andrey Gamaley

Kelly Khanh-Nguyen Lam

Arun Venkatesh

William Faoro

Table of Contents

1. Introduction

1.1 Background

The United States has seen egg price spikes in late 2022 coming into 2023. Consumers have been paying upwards of $7 for a dozen eggs in California, with an increased average of $4.83 for a dozen eggs in the United States [1]. While consumers are experiencing an inflation in egg costs, as well as shortages, leading egg producer Cal-Maine Foods has seen record earnings. For the last 12 months, Cal-Maine Foods has outperformed with a year-over-growth of 430.86% despite ongoing cases of avian influenza (H5N1) and inflation of feed costs and fuel [2]. H5N1 has a high mortality rate which has affected 58,602,281 poultry in the United States, with the majority being commercial layer hens [3]. Depopulation coupled with significant supply reduction, inelastic demand, egg purchasing seasonality, and increased cost of fuel and feed prices to transport and run farms are possible factors that spiked prices. This along with weather, price set by retailers, and implementation of cage-free practices conversely affects the stock price of Cal-Maine Foods.

1.2 Literature Survey

From industry research, we found that retail egg prices are set by retailers and may have a pricing inflation difference that does not directly reflect the wholesale prices that are paid to wholesalers such as Cal-Maine. Previous avian influenza events occurred in 2008 and 2015 with 2022/23 being the most recent outbreaks. However, Cal-Maine reportedly saw no 2022 H5N1 outbreaks at its facilities [6][12].

General monetary inflation is a considerable factor, including “markup growth” [6][12]. On February 15, 2023, U.S. Congress issued a letter to Cal-Maine about egg price-gouging, introducing the Price Gouging Prevention Act of 2022 to ban excessive price increases, requiring companies to include pricing strategies in their public filings [6][12]. Cal-Maine announced a 65% increase in profits in the third quarter of 2022, and a 600% increase in profits for the 12/2024 quarter [12]. In 2020 NY Attorney General sued 2nd largest egg producer, Hillandale Farms, for illegally price gouging [10][12].

	Cal-Maine Foods Inc CALM (NASDAQ)	**Vital Farms, Inc. VITL (NASDAQ)**
Market Cap	2.40B	579.85M
Employees	2,915	368
P/E Ratio	5.57	494.44
Net Revenue	1.8 Billion	362 Million
2022 Yearly Net Revenue Increase	31.7%	38.7%

Table 1.1 2022 Egg Company Financials Comparison

2. Project Overview

2.1 Problem Statement and Research Questions

This project aims to forecast stock prices of Cal-Maine Foods using factors such as electricity price, egg price, gas price, feed grain price, and outbreak of highly pathogenic avian influenza (H5N1) as a foundation to determine driving variables that allow maximum growth for Cal-Maine Foods. We focused our research on why/how Cal-Maine Foods outperformed the market in the last 12 months in tandem with producing record earnings to select the variables above. Through this project, we strive to pinpoint which factors are the most impactful for Cal-Maine Foods growth and attempt to determine the influence, if any, of H5N1 and its risk to business.

2.2 Business Justification and Who Benefits

From an investment perspective, by forecasting Cal-Maine's stock price, shareholders and potential investors will be able to deduce the ideal period of time to buy and sell Cal-Maine shares. Our models will offer insight to investors with an interest in the egg/produce sector and maximizing returns.

From a business perspective, Cal-Maine Foods will gain insights into the factors that impact the stock price which will allow management to allocate resources as necessary in order to maximize value and minimize risk. This, in turn, increases earnings and benefits shareholders. Our team's study may also be used by the egg producers themselves to make better production decisions that would help maintain or increase revenues.

From a consumer perspective, our group's research can improve consumer comprehension about the variables that affect egg prices, consequently leading to making informed purchasing choices and lessening the financial impact of price increases.

The research team's study could aid egg producers in better cost management and decision-making regarding flock culling. An analysis into key factors in the recent growth of Cal-Maine Foods could help stakeholders decide whether to invest more into implementing cage-free practices as demand for cage-free eggs grows, as well as measures to minimize risk of H5N1 as it persists and proliferates.

2.3 Approach

Before running an exploratory data analysis, we plan to have two datasets. The first would contain variables that are hypothesized to affect egg price in addition to Cal-Maine Foods closing stock price. The second dataset would contain values of Cal-Maine Foods and Vital Farms closing stock price that we could use to observe performance differences. We will join the collected data based on daily dates starting from December 12, 1996 when Cal-Maine Foods went public. We will then check the data for missing values, outliers, autocorrelation, multicollinearity, etc. using descriptive statistics and visualization techniques. If seasonality, trend, or heteroskedasticity is detected, the data will be detrended and logarithmically transformed to better fit the models. The data will then be split into training and testing with 90% being used for training and 10% for testing. The models will continue to be refined, some through the use of hyperparameter tuning in Ensemble and its performance will be compared using the test data's RMSE (root mean squared error).

2.4 Initial Hypothesis

We predict that future stock performance of egg producers may not be as positive as it has been in the past year and that egg price will not be a significant factor in the recent growth of Cal-Maine Foods.

3. Data

3.1 Data Collection

Our dataset consists of 5 different sources which we compiled and interpolated to make 6,557 rows. While historical stock is not difficult to obtain, H5N1 reports from 2014-2015 are not available to the public along with daily prices of each key variable. The key variables in our analysis are Cal-Maine stock closing price, gas price, electricity price, feed grain price, and egg price.

3.1.1 Data Description and Sources:

The first data source is the 2022-2023 Confirmations of H5N1 in Commercial and Backyard Flocks dataset from the website of the United States Department of Agriculture's (USDA) Animal and Plant Health Inspection Service (APHIS). This dataset contains information on confirmed cases of H5N1 in the United States.

The second and third data sources are from Yahoo Finance for the historical stock price data of Cal-Maine Foods Inc. (CALM) and Vital Farms Inc. (VITL). The CALM dataset contains stock price information dating back to December 1996, with approximately 6600 rows of data. The VITL dataset dates back to July 2020 and contains 657 rows of data.

The fourth data source is the Bureau of Labor Statistics (BLS) average prices dataset. Users can access this dataset through the BLS data query tool, which allows them to filter the data based on various criteria. We queried the website to download a specific dataset containing information on average prices for eggs, gas, electricity dating back to 1976. This data source provided us with three separate datasets.

Finally, the fifth data source is the Economic Research Service (ERS) USDA website, which provides a custom query tool for feed grains data. The tool allows users to select specific feed grain commodities, geographic regions, and time periods to generate customized datasets. The data can be accessed by visiting the website and selecting the "FEED-GRAINS-custom-query" link. The specific query we used is: [*Group*: feed grains, *item*: market egg-feed, *geography*: United States, *data attribute*: price ratio, *frequency*: monthly, *year*: all years, *unit*: ratio].

3.2 Data Cleaning and Transformation

Most of the data downloaded came in the format of dates and prices such as electricity, feed, egg, and gas. Data about H5N1, Cal-Maine Foods, and Vital Farms came with additional variables. The H5N1 dataset was used to determine which dates had outbreaks and assigned a binary output with 1 being the days with an outbreak occurrence and 0 as no occurrence. Two datasets were curated, one with all the independent variables along with the closing price of Cal-Maine Foods and the other dataset containing the closing price of both Cal-Maine Foods and Vital Farms for a separate analysis. In order to join the data together, each data object was casted into datetime with the format year-month-day. The H5N1 dataset contained a column named “Flock Type” which had values such as “WOAH Non-Poultry”, “WOAH Poultry", “Commercial Table Egg Layer”, etc. Our analysis involves the impact of factors on eggs, thus, the H5N1 dataset was filtered to contain only data from” Flock Type” equalling “Commercial Table Egg Layer”.

```
        Date elec_price egg_price feed_price gas_price    Close Outbreak
1 1996-12-12      0.092  1.251226   11.31935  1.260355 1.703125        0
2 1996-12-13      0.092  1.246065   11.24839  1.260387 1.781250        0
3 1996-12-16      0.092  1.230581   11.03548  1.260484 1.718750        0
4 1996-12-17      0.092  1.225419   10.96452  1.260516 1.773438        0
5 1996-12-18      0.092  1.220258   10.89355  1.260548 1.812500        0
6 1996-12-19      0.092  1.215097   10.82258  1.260581 1.750000        0
```

Table 3.1 First 5 Rows of Combined Dataset

While the stock price data is aggregated daily, the electricity, feed, egg, and gas price data are aggregated monthly. The latter data was linearly interpolated using the approx() function in R to estimate daily values based on their monthly averages then joined with the Cal-Maine Foods data. Interpolation is a statistical method that can estimate unknown values between two known values. We chose to use linear interpolation to estimate the daily values of these factors based on their monthly averages. By interpolating the monthly data, we were able to match the data frequencies and obtain daily data for all of these factors.

With the combined dataset, a simple multiple linear regression was created: *lm(Close ~ elec_price + egg_price + feed_price + gas_price + Outbreak, data=df)*. The model was used to check for outliers using Cook’s distance and VIF. No outliers were detected, but the VIF showed *egg_price* to have a value of 5.931430 and a strong correlation which can be interpreted in the correlation plot. In addition, *egg_price* was not statistically significant in the multiple linear regression model and confirmed in Stargazer prompting the removal of the variable in the other models. The diagnostic plot of the regression shows normally distributed data, but nonlinear residuals. The Durbin Watson test along with ncvTest was able to detect autocorrelation and heteroskedasticity. In order to counteract these issues, appropriate detrending and differing of time series data needs to be done prior to feeding the variables in the model. Using lagged data could help account for autocorrelation and logarithmic transformed data could help with variance issues in the data. Not only is *egg_price* removed, but variable *Outbreak* as well. Outbreak did not show to be a key variable in our preliminary analysis and because of the lack of data needed to make a significant change in our analysis *Outbreak* was removed.

electricity price	egg price	feed grain price	gas price	Outbreak (0 or 1)
3.906652	5.931430	2.355868	3.738356	1.030320

Table 3.2 VIF of Factors in Multiple Linear Regression Model

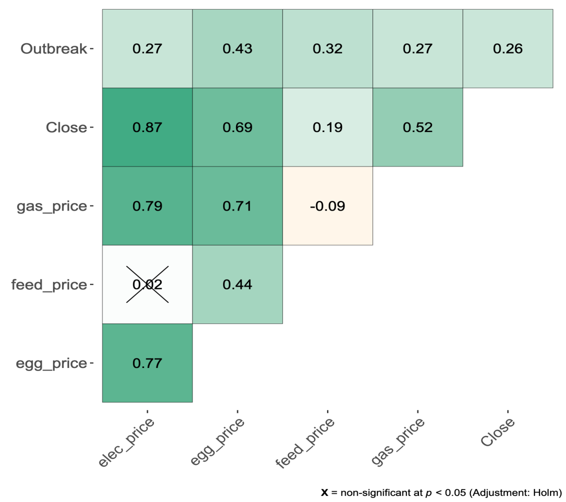


Figure 3.1 Correlation Plot of Variables

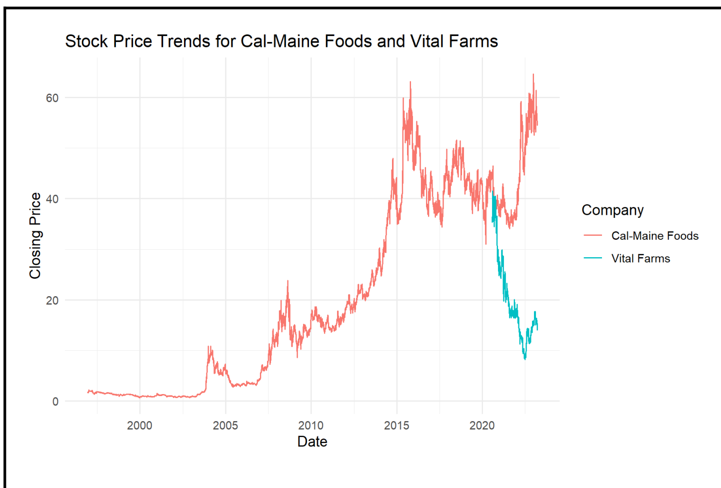
	Dependent variable:
	Close
elec_price	1,007.445***
	(7.862)
egg_price	0.465
	(0.408)
feed_price	0.587***
	(0.035)
gas_price	-8.841***
	(0.190)
Outbreak	1.457**
	(0.699)
Constant	-83.798***
	(0.658)
Observations	6,557
R2	0.855
Adjusted R2	0.855
Residual Std. Error	6.937 (df = 6551)
F Statistic	7,711.423*** (df = 5; 6551)

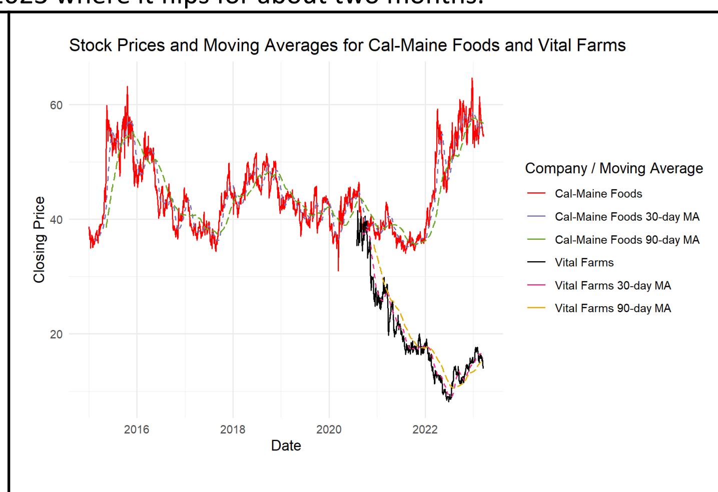
Figure 3.2 Stargazer Output of Predictors

3.3 Exploratory Data Analysis

The stock price patterns and daily percentage changes for Cal-Maine Foods and Vital Farms were visualized along with the summary statistics for these changes as part of the exploratory data analysis. The 30-day and 90-day moving averages were drawn alongside the prices of the stocks, and the volatility of the 30-day stock price was examined. Also calculated was the correlation between the daily percentage changes in stock prices for the two businesses. The same visualizations were made with a focus on the years 2021 to 2023, and monthly returns for both businesses were plotted. Closing prices and daily percentage changes were used to fit linear regression models, and Cook's distance was computed and visualized to look for significant spots. In order to determine whether the data were multicollinear, variance inflation factors were lastly computed.

In the Stock Price Trends for Cal-Maine Foods and Vital Farms visualization, we notice Cal Maine Foods has two large spikes at 2015 and 2022 during outbreaks. Vital Farms Stock steadily decreases until the end of 2022. In the Stock Prices and Moving Averages for Cal-Maine Foods and Vital Farms visualization, we notice Cal Maine Foods' steady price per share increase from 2021 - 2023 ($40 to $55) as well as Vital Farms steady price per share decrease from 2021 - 2022 ($25 to $17) and sharp increase at the start of 2022 to roughly $20. In the Monthly Returns for Cal-Maine Foods and Vital Farms (2021-2023) visualization, we notice roughly equal negative and positive monthly returns for Cal Maine Foods and mainly negative returns for Vital Farms in the year 2021. At the start of 2022, we see large returns for Cal Maine Foods. We see massive monthly returns at the end of 2022 for Vital Farms after being in negative for 6 months. In the 30-day Rolling Volatility for Cal-Maine Foods and Vital Farms (2021-2023) visualization, we see that Vital Farms is about two times more volatile than Cal-Maine Foods until the start of 2023 where it flips for about two months.





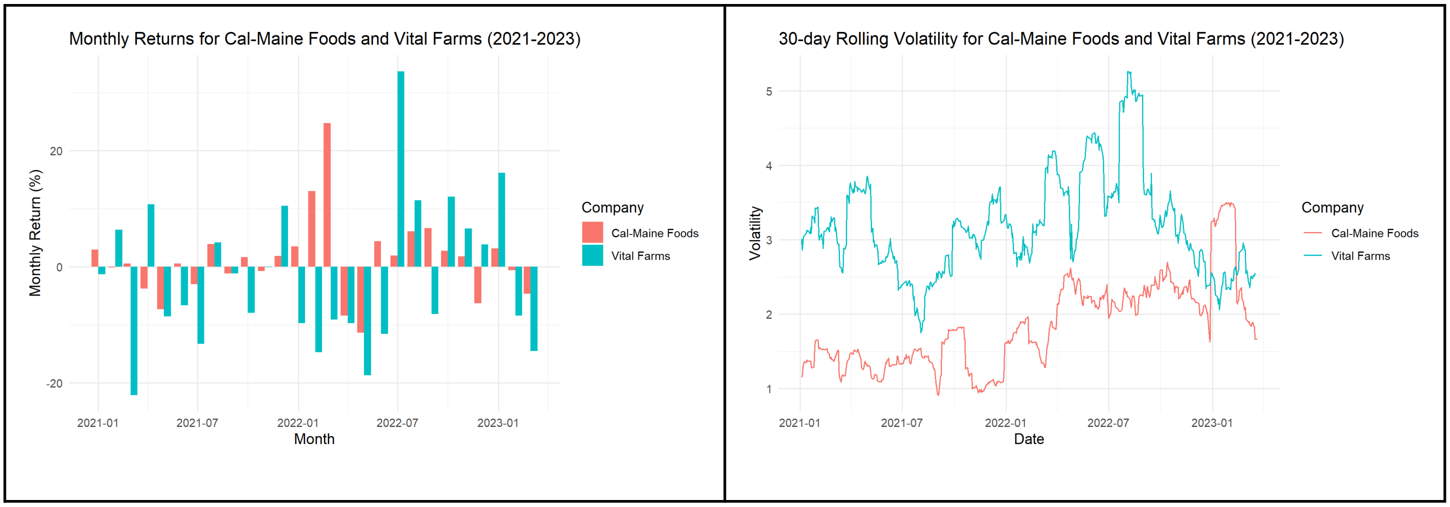


Figure 3.2 Comparisons between CALM and VITL stock prices, moving averages, monthly returns, and rolling volatility

A plot of the predictors were also graphed together to see trends in conjunction with closing price of Cal-Maine. We can see that egg and gas prices follow a similar trend and seasonality to Cal-Maine stock closing price. Feed grain prices seem independent to all patterns and egg price peaks and drops to extremes conversely to closing price.

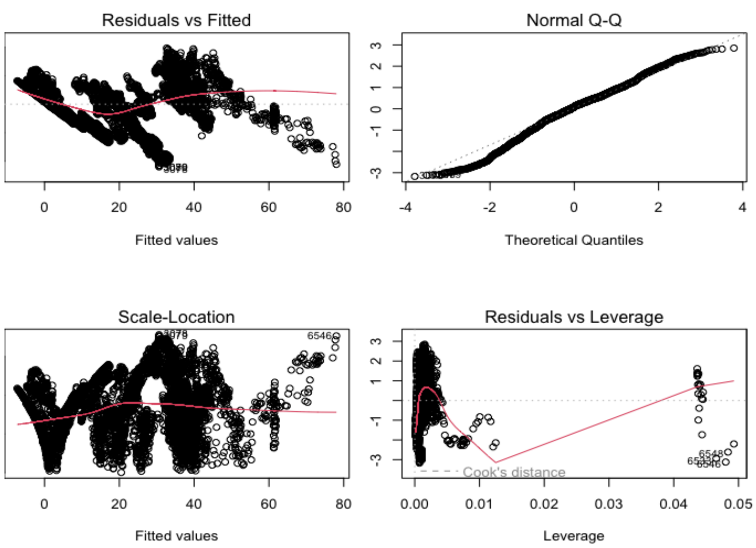


Figure 3.3 Diagnostic Plot of Multiple Linear Regression Model

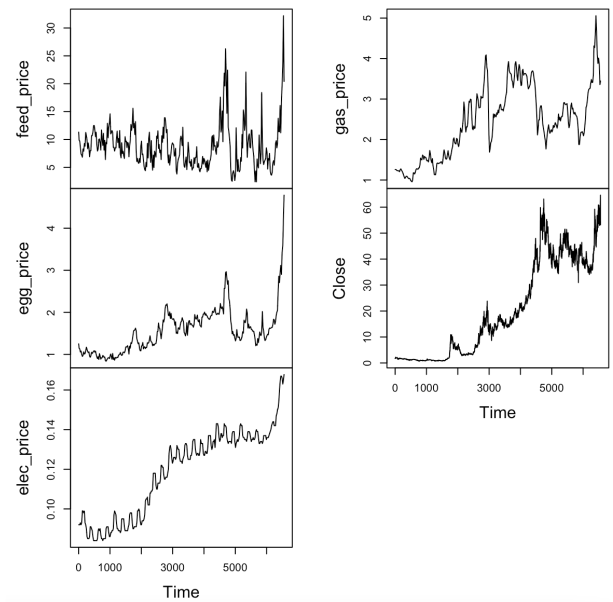


Figure 3.4 Time Series Plot of Independent and Dependent Variables

4. Modeling

4.1 Multiple Linear Regression

We introduce a new predictor which is the lagged dependent variable to account for autocorrelation and arrive at the formula *lm(formula = Close ~ lagged_Close + elec_price + feed_price + gas_price + Outbreak, data = df)*. Prior to this introduction, the dependent variable, *Close*, was log transformed. The Durbin Watson test was used to test for autocorrelation which was still perfectly positive despite the transformation. Therefore, we concluded that a log transformation was not needed for this model. When introducing the lagged dependent, the Durbin Watson test had a value of 1.999 and p-value of 0.9228 which suggests the data is no longer autocorrelated. The lagged predictor causes *gas_price* and *Outbreak* to become insignificant; the adjusted R^2 is 0.9989 which is a strong indicator of overfitting. Thus, we can conclude that multiple linear regression is not a proper model for this data.

4.2 Random Forest

The Random Forest model was built with 500 trees and using three predictors at each split. The variable importance (%IncMSE and IncNodePurity) shows the importance of each predictor in the model:

Random Forest	RMSE: 0.2396396

Coefficient	%IncMSE	IncNodePurity
Date	20.380461	644691.016
elec_price	6.991774	172453.050
feed_price	17.711296	11230.746
gas_price	17.028705	4240.007
lagged_closed	36.861237	1334696.443

Table 4.2 Random Forest Percentage Increase of MSE and Increase of Node Purity

When a variable is permuted, the mean squared error of the model increases (%IncMSE), indicating how important the variable is. IncNodePurity is a measure of how much the variable influences node impurity overall across all trees, i.e., how much the variable influences the accuracy of the model's predictions.

The most crucial predictor in this situation is "lagged_Close," which is followed by "Date". After that “feed_price” and “gas_price” are close on % increase of MSE, but “feed_price” influences node impurity more. Lastly, “elec_price” has the lowest % increase of MSE.

4.3 Lasso, Elastic Net, & Ridge Regression

In our stock price prediction problem, regularization methods like LASSO, Elastic Net, and Ridge Regression aid in reducing multicollinearity and enhancing model performance. The regression is given an L1 penalty via LASSO (Least Absolute Shrinkage and Selection Operator), which can cause some coefficients to be shrunk to zero, thereby performing variable selection. This makes the model easier to understand and lessens multicollinearity. Elastic Net strikes a compromise between variable selection and regularization by combining the penalties of LASSO (L1) and Ridge (L2) regression. When dealing with strongly linked predictors, it is extremely helpful because it can keep key variables while removing less important ones. Ridge Regression reduces multicollinearity without removing variables by adding an L2 penalty to the regression that prevents coefficients from growing too big. When numerous predictors have a moderate impact on the response variable, it is especially helpful. These regularization strategies are useful tools for improving our forecasting model since they aid to increase model interpretability, decrease overfitting, and increase prediction accuracy in our stock price prediction scenario.

The best LASSO, Elastic Net, and Ridge models were obtained using cross-validation to find the ideal lambda values. Cross-validation was used to establish the ideal lambda value, and it was found that it should be around 0.1196522 for Lasso, 0.1035892 for Elastic Net, and 1.82 for Ridge. Lambda regulates the level of model sparsity in LASSO and Elastic Net. A model with more variables is simpler and has fewer coefficients since a greater lambda causes more coefficients to be reduced to zero. The shrinkage of coefficients towards zero in Ridge Regression is controlled by lambda without their total eradication. Our multicollinearity problem is mitigated by a bigger lambda, which leads to a more conservative model with smaller coefficients.

The predictor variables' best fitting of the Lasso, Elastic Net, & Ridge Regression model with the respective lambdas yielded the following RMSE and coefficients:

LASSO	ELASTIC NET	RIDGE
RMSE: 0.5596772	RMSE: 0.5699784	RMSE: 2.242927

elec_price	.	elec_price	10.903056395	elec_price	174.5814731
feed_price	.	feed_price	0.006828629	feed_price	0.1438454
gas_price	.	gas_price	.	gas_price	-1.0697372
lagged_Close	0.9930848	lagged_Close	0.982480907	lagged_Close	0.7965988

Table 4.3 Lasso, Elastic Net, and Ridge RMSE's and Variable Coefficients

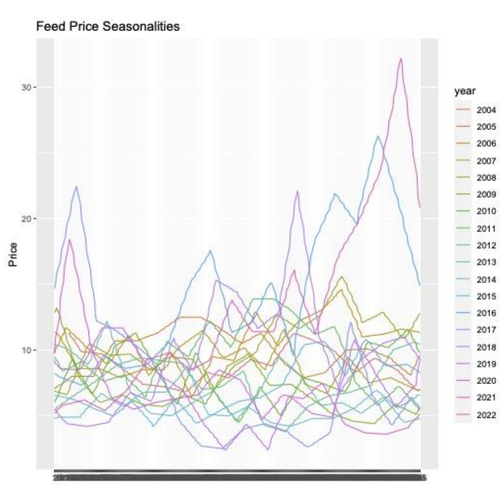
* The "." character in the coefficient outputs denotes a value that has been effectively reduced to zero.

Except for the lagged_Close, all predictor variables for the best LASSO model have coefficients that are practically zero. This suggests that the only meaningful predictor of stock price in this model is the lagged_Close variable. Three predictor variables—elec_price, feed_price, and lagged_Close—have non-zero coefficients for the ideal Elastic Net model. This shows that in this model, these three factors are important predictors of stock price. In a side-by-side comparison, the LASSO model performs slightly better than Elastic Net and much better than Ridge and has the lowest MSE.

4.4 Holt-Winter's

As a preliminary to running a Holt-Winters seasonal smoothing model, ggseasonplot was used to overlay the yearly seasonal trend plots for each factor; Cal-Maine stock price, feed price, electricity price, gas price, and egg price. Comparing these plots, no reliable, distinct seasonal trend was apparent. As a comparative look at individual factors, Holt-Winters was used to smooth additional unnoticed seasonal factors, by minimizing squared error predictions, and then provide a basis for a one-year forecast for each individual factor; within a 95% confidence interval. The model was allowed to autotune the parameters as this seemed to best fit the existing data. The auto-tuned smoothing parameters were: HW CM Stock Price - alpha: *0.9894666* beta: 0 gamma: *1*, HW Feed Price - alpha: *0.9609935* beta: *0.3676927* gamma: *1*, HW Gas Price - alpha: *0.9730097* beta : *0.05469792* gamma: *1*, HW Electricity Price - alpha: *0.8490613* beta : *0.0637162* gamma: *1*, HW Egg Price: alpha: *0.9288765* beta : *0.3030983* gamma: *1.*

The 'additive' seasonal parameter was used as the 'multiplicative' seasonal parameter created a distorted model. The frequency chosen was 365 as per the daily data points.The forecast seemed to be reasonable for all factors, except for feed which showed an unusually steep decline towards the bottom of the confidence interval. The forecast for the Cal-Maine stock price was particularly level seeming to indicate an average of the lows and highs for each of the 2008 & 2022 avian flu event periods.



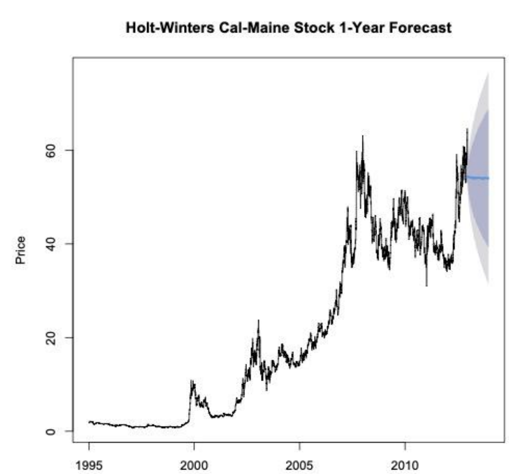


Figure 4.1 Feed Price Seasonalities

Figure 4.2 Holt-Winters Cal-Maine Forecast

4.5 ARIMA

ARIMA is a method that forecasts future outcomes using historical time series data. In order to reduce variance, a log transformation of all the variables was done prior to converting the dataframe into a time series object. The use of log transformation was able to reduce the MSE making the final iteration more accurate. Auto.arima was used as a starting point in order to get *p, d,* and *q* values for Arima. Through trial and error of testing various *p*'s, *d*'s, and *q*'s after obtaining an initial value of *(1,1,0)*, values *(1,1,1)* resulted in the best forecasted values evaluated by the test set's MASE of *0.04088906*. Moreover, *acf()* and *kpss()* for autocorrelation and stationarity were used to test the model and results showed that the model was no longer autocorrelated and stationary. The test set produced an mean error of *0.06588614,* root mean squared error of *0.187418*, and mean absolute error of *0.1456331*. While the accuracy results show low values indicating decent forecasting, the graph of the forecasted values does not do an adequate job of predicting future values. The forecasted values show a slight positive trend of growth for Cal-Maine Foods.

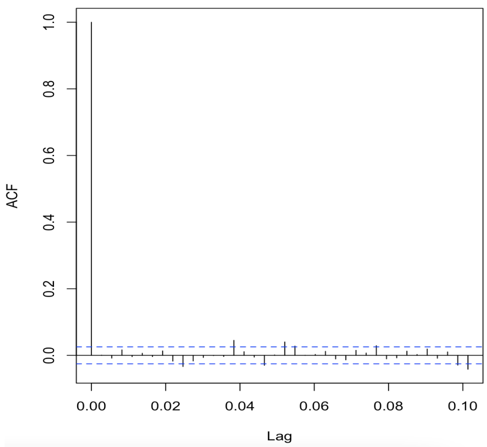


Figure 4.3 Autocorrelation Function (ACF) Plot of Arima Model

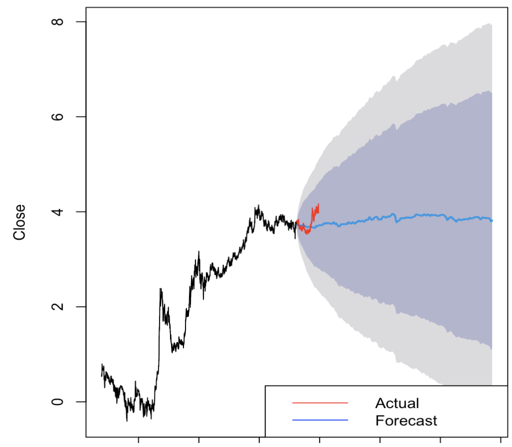


Figure 4.4 ARIMA Forecasted Values

4.6 Ensemble

Stacked Ensemble is a machine learning technique that combines the predictions of multiple base models to improve the accuracy of a prediction. The base models can be trained using different algorithms, and the Stacked Ensemble algorithm learns to optimize the combination of these models based on a validation set. Stacked Ensemble was used to predict the Stock Price Trends for Cal-Maine Foods based on multiple features discussed above. We trained a collection of base models on the dataset using different algorithms such as Random Forest, Gradient Boosting, and Deep Learning. Then, we used the Stacked Ensemble algorithm to combine the predictions of these base models and obtain a more accurate prediction. Cross-validation was used to split the dataset into training and validation sets, and the Stacked Ensemble algorithm learned to optimize the combination of the base models on the validation set. Finally, we evaluated the performance of the Stacked Ensemble model on a test set and found that it outperformed the individual based models in terms of prediction accuracy. Following is the leader board generated using AutoML.

	model_id <chr>	rmse <dbl>	mse <dbl>	mae <dbl>
1	StackedEnsemble_AllModels_1_AutoML_2_20230416_173943	0.6764105	0.4575311	0.3878401
2	StackedEnsemble_BestOfFamily_1_AutoML_2_20230416_173943	0.6994526	0.4892340	0.4028981
3	GBM_2_AutoML_2_20230416_173943	0.7526526	0.5664859	0.4521843
4	GBM_4_AutoML_2_20230416_173943	0.7585724	0.5754321	0.4346140
5	GBM_3_AutoML_2_20230416_173943	0.7641362	0.5839041	0.4463753
6	XGBoost_2_AutoML_2_20230416_173943	0.7907629	0.6253059	0.4239755
7	XGBoost_3_AutoML_2_20230416_173943	0.8717924	0.7600220	0.5111423
8	GBM_1_AutoML_2_20230416_173943	1.0734481	1.1522908	0.6420966
9	XGBoost_1_AutoML_2_20230416_173943	1.1823866	1.3980382	0.6314338
10	DRF_1_AutoML_2_20230416_173943	1.8687674	3.4922915	1.2703560

Table 4.3 AutoML Leaderboard

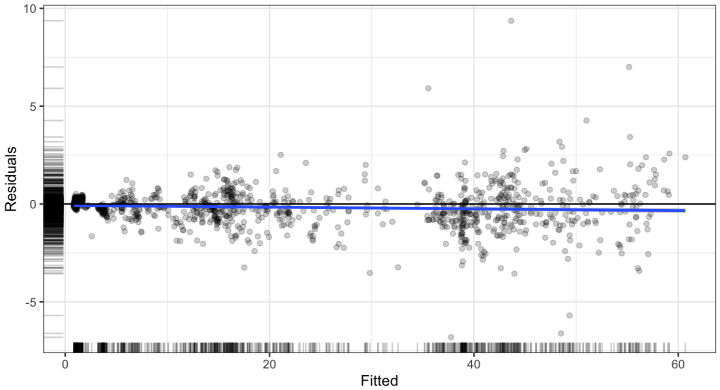


Figure 4.5 Residual Analysis Plot

The plot exhibits a widening of the spread of residuals as the predicted values increase, indicating heteroscedasticity.

4.7 Model Comparison

Model	RMSE
Random Forest	0.2396
LASSO	0.5596
ELASTIC NET	0.5699
RIDGE	2.2429
ARIMA	0.1874
H2O-Stacked Ensemble Model	0.6764

RMSE represents the average distance between the predicted values of the model and the actual values in the dataset. In the case of forecasting, RMSE represents the distance between the forecasted values and the actual values of the dataset of the window date we provided. While RMSE can be a good indicator of how well fitted the model is to the data based on how close the value is to zero, other metrics may be considered alongside to better interpret performance. Although ARIMA performs well with this data, outputting an RSME of 0.1874, we can tell by the forecasted versus actual values graph that ARIMA is not too closely accurate.

5. Conclusion

After examining our data and exploring the issues that could affect a model's performance, we found that our data lacked enough distinct data points to output appropriate forecasting models. Interpolating our data from monthly to daily could have caused perfectly positive autocorrelation, as well as, heteroskedasticity. Considering the background information, it seemed plausible to eliminate the avian flu variable as this black-swan event admittedly did not afflict any Cal-Main production locations in the latest outbreak, and there is pending legislation and vaccines likely to prevent future effects of this factor. While electricity and gas prices were key variables statistically significant in our analysis, predictors such as egg price and outbreak had to be removed in order to improve each model's performance due it being statistically insignificant. More predictors could help to explain variances in Cal-Maine Food stock price which requires more research into our problem statement. Our models, however, show that we can safely assume gas and electricity prices are important factors that follow Cal-Maine Foods stock price trends. Thus, consumers and investors can take those two factors into account when purchasing eggs and stock relative to price spikes.

Improvements in data collection can be made to properly explain variances in Cal-Maine Food stock price. If possible, our models would continue to accept daily data values to refit when available. A further iteration could include more factors such as H5N1 specific to Cal-Maine Food egg suppliers or demand of cage-free eggs.

References

1. Federal Reserve Bank of St. Louis. (2023, March 11). Average Price: Eggs, Grade A, Large (Cost per Dozen) in the U.S. City Average. FRED Economic Data. https://fred.stlouisfed.org/series/APU0000708111

2. Zacks Equity Research. (2023, March 11). Cal-Maine Foods (CALM) Gains As Market Dips: What You Should Know. Yahoo! Lifestyle. https://www.yahoo.com/lifestyle/cal-maine-foods-calm-gains-224510114.html

3. Centers for Disease Control and Prevention. (2023, March 11). H5N1 Bird Flu Detections across the United States (Backyard and Commercial). https://www.cdc.gov/flu/avianflu/data-map-commercial.html

4. Cal-Maine Foods. (2023, January 24). Cal-Maine Statement on Current Egg Prices. https://www.calmainefoods.com/news/cal-maine-statement-on-current-egg-prices/

5. Reuters. (2024, January 20). High egg prices should be investigated, U.S. farm group says. https://finance.yahoo.com/news/high-egg-prices-investigated-u-161846316.html

6. Cal-Maine Foods. (2022, January). Investor Presentation. https://www.calmainefoods.com/media/1402/calm-investor-presentation-192023.pdf

7. U.S. Bureau of Labor Statistics. (2023, February). Table 2. Consumer Price Index for All Urban Consumers (CPI-U): U. S. city average, by detailed expenditure category. https://www.bls.gov/news.release/cpi.t02.htm

8. USDA Animal and Plant Health Inspection Services. (2024, January 18). 2022-2023 Confirmations of Highly Pathogenic Avian Influenza in Commercial and Backyard Flocks. https://www.aphis.usda.gov/aphis/ourfocus/animalhealth/animal-disease-information/avian/avian-influenza/hpai-2022/2022-hpai-commercial-backyard-flocks

9. Journal of Agricultural Science. (2018, October). Analysis of Egg Price Fluctuation and Cause. https://www.researchgate.net/publication/328290901_Analysis_of_Egg_Price_Fluctuation_and_Cause

10. United States Securities and Exchange Commission. (2022, May 28). Form 10-K. https://calmainefoods.gcs-web.com/static-files/986c7031-7d45-4cd0-a626-f3769fc39363

11. USDA Economic Research Service. (2023, January 11). Avian influenza outbreaks reduced egg production, driving prices to record highs in 2022. https://www.ers.usda.gov/data-products/chart-gallery/gallery/chart-detail/?chartId=105576

12. Congress of the United States. (2024, February 16). Letter to Cal-Maine Foods. https://www.warren.senate.gov/imo/media/doc/2023.02.16%20Letter%20to%20Cal-Maine%20Foods%20on%20Egg%20Prices.pdf

Dataset Links

1. https://www.aphis.usda.gov/aphis/ourfocus/animalhealth/animal-disease-information/avian/avian-influenza/hpai-2022/2022-hpai-commercial-backyard-flocks - H5N1 Dataset

2. https://finance.yahoo.com/quote/CALM/history?p=CALM - Cal-Maine Food Stock Price Dataset

3. https://finance.yahoo.com/quote/VITL/history?p=VITL - Vital Farms Stock Price Dataset

4. https://beta.bls.gov/dataQuery/find?fq=survey:[ap]&s=popularity:D - Bureau of Labor Statistics Gas, Electricity,

and Egg Price Dataset

5. https://data.ers.usda.gov/FEED-GRAINS-custom-query.aspx - Feed Grains Price Dataset