Forecasting Egg Prices: A Time Series Analysis with Economic and Market Influences

Author: Heng Wang (2722303)

Daily company supervisors: Joris Slootweg and Frank Ottenhof
First VU supervisor: Robert van der Mei
Second VU supervisor: Shujian Yu

A thesis submitted in fulfillment of the requirements for the
VU Master of Science degree in Business Analytics (Financial Risk Management)

Abstract

This thesis explores the use of time series models and machine learning techniques to forecast egg prices in the Dutch market, incorporating key factors such as feed costs, energy costs, and economic indicators. The models evaluated include ARIMA, SARIMA, Prophet, and LSTM, along with benchmark models such as Naive and Hyperbolic Forecasting. Data from 2016 to 2024 were collected and preprocessed, followed by feature analysis and selection to identify the most relevant variables to predict egg prices. Using an expanding window approaches, models were evaluated over short-term (1 week), medium-term (4 weeks), and long-term (12 weeks) forecasting horizons. The results indicate that ARIMA performs best for short-term forecasts, while LSTM is more suited for medium and long-term prediction. External factors, such as Easter holidays, the Russia-Ukraine war, and the Fipronil crisis, were also tested to assess model performance under market dynamics. The bootstrap method was applied to the generated confidence interval to assess the model's uncertainty. In general, this thesis highlights the importance of the selection of features and the choice of models based on the forecast horizon and market conditions. Further improvement is recommended, including the use of localized data and additional hybrid models, to improve forecast accuracy.

Contents

1 Introduction

1.1 Background

Eggs are a fundamental food item worldwide, providing an essential source of nutrition and representing a significant sector of the agricultural economy. In the European Union(EU), the poultry egg production sector plays a crucial role. The Netherlands is one of the main egg-producing countries in the EU. In 2023, the Netherlands produced 525,000 tonnes of eggs for consumption, accounting for 8% of the total EU egg production, ranking 6th after France, Germany, Spain, Italy, and Poland [14]. The economic importance of this sector is underscored by its contribution to the EU's agricultural economy, with the value of poultry and egg production estimated at 39 billion euros in 2022 [22].

However, this sector's heavy reliance on imported feed materials, especially high-protein sources like soybeans and maize, makes it vulnerable to global market fluctuations. In 2022, 33% of the total industrial compound feed production was used specifically for poultry and egg production. 72% of high-protein feed sources were imported. In 2022, the EU imported 24.8 million tons of feed cereals and 28.1 million tons of soy from non-EU countries [22].

Egg prices are subject to fluctuations caused by a variety of factors, including production costs, supply and demand, and broader economic conditions. The basic production costs account for about 86% of the total poultry costs in the EU in 2021 [56]. Among basic production costs, feed costs represent the most significant portion of the total production costs in poultry farms, about 65% to 75% [6]. As shown in Figure 1, feed costs are the major cost contribution of Dutch laying hen farmers. In the past 10 years, the total feed costs remained relatively stable over the years until 2022, when there was a noticeable increase in both feed costs and total allocated costs. This spike can be attributed to the geopolitical impact of the Russia-Ukraine war, which caused increasing grain prices, the energy crisis, and high inflation. It reflected that the reliance on imported feed materials underscores the vulnerability of EU and NL poultry and egg production. Consequently, the total compound feed production for the poultry category in the EU decreased by 2.3% compared to 2021, with the Netherlands experiencing a 0.7% decline [22].

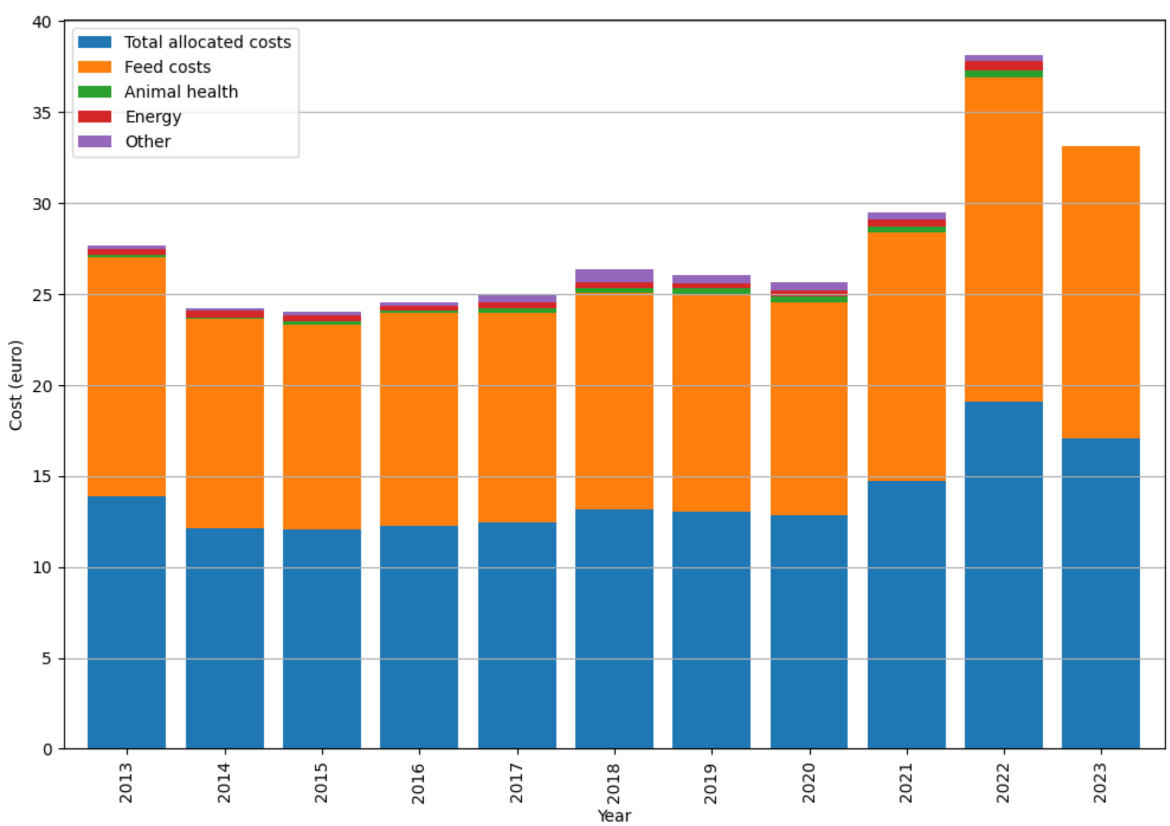


Figure 1: Cost Distribution in Laying Hen Farms (2013-2023). Source: Agrimatie

1.2 Problem Statement

The volatility of egg prices in the global food market poses significant challenges for various stakeholders, including farmers, distributors, retailers, and consumers. Egg price fluctuations can be driven by a multitude of factors such as feed costs, energy costs, inflation rates, and market disruptions such as holidays and geopolitical events. These fluctuations can severely impact production planning, inventory management, and profitability, making it difficult for stakeholders to mitigate risks and optimize decision-making.

Traditional models like AutoRegressive Integrated Moving Average (ARIMA) and Seasonal ARIMA (SARIMA) have been employed for such time series forecasting tasks. While these models have been proven to predict time series effectively, they may fall short when it comes to modeling the complex, nonlinear relationships between various influencing factors. Additionally, it is hard for ARIMA and SARIMA models to detect the increasing demand due to holidays like Easter, which is not a fixed date. Furthermore, economic indicators like the inflation rate often exhibit high multicollinearity with features like feed costs and energy prices, further complicating accurate forecasts.

Given the growing complexity of market dynamics, newer methods like Prophet model and advanced machine learning techniques such as Long Short-Term Memory (LSTM) networks have been employed. Prophet model offers advantages in handling seasonality and special events like the Easter holiday, while LSTM model has the capacity to learn long-term dependencies and capture nonlinear interaction between variables.

However, identifying the most relevant features and their relationship with egg prices is a crucial challenge. It involves selecting the appropriate factors to include in the model while ensuring that high multicollinearity between features is minimized. Additionally, choosing the right model parameters that best capture the dynamics of egg price movements is essential. Beyond feature selection, it is also critical to understand the strengths and limitations of each forecasting model across different prediction horizons to ensure that the chosen model is well-suited to capture both short-term fluctuation and long-term trends effectively.

This thesis aims to explore how different forecasting models, including ARIMA, SARIMA, Prophet, LSTM, as well as simpler models like Naive and Hyperbolic, can be leveraged to improve the accuracy of egg price predictions in the Dutch market. By considering key factors such as feed cost, energy costs, and economic indicators, the thesis seeks to evaluate which models are best suited for short-term, medium-term, and long-term forecasting, while also addressing the impact of uncertainty of decision-making.

1.3 Research Question

The primary research question guiding this thesis is:

How can egg prices be accurately forecasted over short, medium, and long-term horizons using different time series models?

The subquestions include:

- What is the seasonality of egg prices, and how does it affect the overall market dynamics?
- Which features are most relevant in predicting egg prices, and how can multicollinearity among these factors be addressed during feature selection?
- How do traditional forecasting models (ARIMA, SARIMA, Prophet) compare with machine learning models (LSTM) and simpler models (Naive, Hyperbolic) in terms of prediction accuracy of egg prices?
- How do the models handle uncertainty in egg price predictions, and what are the implications for decision-making by stakeholders?

1.4 Outline

This thesis is structured as follows: the relevant literature will be reviewed in Section 2, which provides an overview of existing research on the egg market, focusing on the factors influencing egg prices and the methods used for feature selection in the predictive model. It also reviews various forecasting models, including ARIMA, SARIMA, Prophet, and LSTM, discussing their application in agricultural price forecasting and their strengths and limitations. Section 3 describes the data sources, preprocessing steps, and exploratory data analysis performed to understand trends, seasonalities, and the impact of holidays on eggs. Section 4 details the feature analysis and selection process, including the correlation between the selected features and egg prices.

Section 5 outlines the theoretical background and implementation of the forecasting models used in this thesis, including ARIMA, SARIMA, Prophet, and LSTM. This includes the steps taken for model setup to ensure reproducibility. Special emphasis is placed on the hyperparameter tuning process, detailing how optimal parameters were chosen for each model. In addition, this section introduces the theoretical background of Naive Forecasting and Hyperbolic Forecasting, as two benchmark models to compare models' performance. Furthermore, Chapter 6 discusses the evaluation metrics used to compare model performance and assessment of models over different forecasting horizons, and their ability to handle uncertainty.

Section 7 presents the results of the forecasting models, compares their performance in short-term, medium-term, and long-term predictions, and discusses how they handle uncertainty. The implications of the findings are also analyzed. Section 8 summarizes the key findings, highlights the contributions of this research to the field of agricultural price forecasting, and suggests directions for future research.

2 Literature Review

The literature review analyzes existing research related to agricultural price prediction and methodologies used for time-series forecasting. The review highlights the significant findings, methodologies, and comparative performance of various models, thereby giving insights into the current thesis.

The egg market has been a topic of research, particularly in terms of production cost and price analysis. A study of the global egg market was conducted by RaboResearch, where historical trends in the Rabobank Global Egg Price Index and the Food Price Index (FAO) were examined [35]. From 2013 to 2019, egg prices were slightly above the index's movements. This trend was driven by relatively strong demand for eggs, while feed costs had stabilized. However, post-2021, egg prices closely followed the movement of the global food price index, influenced by factors like rising feed costs, avian influenza outbreaks, market disruptions post-COVID-19, and regulatory changes. This trend, often referred to as "eggflation", shows the interplay between feed prices, global demand, and external market shocks on egg prices.

Further analysis has explored various factors that contribute to the fluctuation of egg prices, particularly in the US market [34]. In 2023, it was found that rising egg prices were not solely due to bird flu, as reported by some newspapers, but were also influenced by post-COVID-19 supply chain disruptions, increased holiday demand, and inflation. Exogenous variables such as feed costs, natural gas prices, and relative loss of birds were integrated into the analysis. The study highlighted that natural gas, being a primary energy input in US egg production, had a significant impact on prices, while electricity was excluded due to its high correlation with natural gas. Using long-term elasticity estimates for the average retail price and the corresponding percentage changes in the determination variables, they demonstrated that a 143.3% increase in natural gas prices significantly affected egg prices.

The feature selections of this thesis were inspired by the above studies about factors contributing to fluctuating egg prices. According to Dutch News, about 200,000 chickens were culled as the bird flu death toll tops four million in 2022 [19]. However, it was not the primary reason for rising egg prices. Similar to the U.S. market, supply chain disruptions, increasing production costs, and high demand are also the contribution of fluctuating markets. Furthermore, recent studies inspired this thesis to explore the features of egg production and possible multicollinearity between features.

In general, the cost of egg production is primarily derived from feed costs, the price of young hens (pullets), housing costs, and labor. The Netherlands has the advantage of low young hens (pullets) and feed costs, which makes relatively lower production costs compared to Denmark, France, Spain, Italy, Germany, and Portugal [56]. Compared to the base years 2015, and 2017, the overall production costs in 2021 increased largely due to the increase in feed prices [55, 54]. These studies highlighted the importance of feed costs in egg production, which was considered a primary feature in model development.

The components of feed were researched by [15, 58]. About 22.4% of compound feed for layer poultry is from soybean meal in Europe. Soybean prices, along with corn prices and overall feed costs, have also been identified as critical factors influencing agricultural markets, particularly pork prices. In addition to feed-related factors, other variables such as pork exports and imports, chicken prices as a related agricultural product, and economic indicators like exchange rates, the Producer Price Index (PPI), and the Consumer Price Index (CPI) have been considered in previous studies. Radial Basis Function (RBF) neural network models have been applied to forecast agricultural price indexes by incorporating these multiple factors. These studies provided inspiration for the feature selection process in this thesis, leading to the inclusion of chicken prices as a related agricultural product and CPI as a measure of inflation during the analysis of egg price prediction models.

Similarly, the relationship between agricultural and energy commodities has also been explored. Recent research analyzed the price connectedness of various agricultural and energy commodities using Quantile Vector Autoregression models [32]. The study collected data from the World Bank commodity market, and focused on agricultural commodities such as wheat, barley, soybean, soybean oil, and sunflower oil, in addition to energy commodities like crude oil and natural gas, and fertilizers such as DAP and Urea. Their study demonstrated that soybeans and their derivatives are the primary transmitters of price signals in the market. In addition to soybean derivatives, barley, and wheat were identified as the next most significant influencers, based on the extent of spillover effects. The third significant influence on soybeans was found to be crude oil. These findings indicate a high degree of interrelation among commodities, with soybeans acting as the market leader.

Moreover, these findings are consistent with earlier studies that highlighted the impact of crude oil on soybean prices [5]. The results proved that fluctuations in agricultural production costs influence the agricultural markets. Also, it provided a choice of public data from the World Bank as the data source of this thesis.

The egg market exhibits notable seasonal patterns that are crucial in forecasting egg prices. A study examining the egg futures market at the Chicago Mercantile Exchange revealed a 12-month seasonal pattern, rather than the previously assumed 30-month cycle [29]. The 12-month cycle highlighted by this research emphasizes the yearly seasonality in predictive models. Furthermore, this study found that the prices of eggs in storage are initially lower than future prices, increasing as the eggs are brought out of storage due to the perishability of the eggs. It provided insights for this thesis about the influence of the limited shelf life of eggs on market pricing, particularly in determining periods leading up to peak demand seasons like the Easter holidays.

Traditional time series forecasting models such as ARIMA (AutoRegressive Integrated Moving Average) and its seasonal counterpart, SARIMA (Seasonal AutoRegressive Integrated Moving Average), have been widely used in agricultural price forecasting, including eggs. ARIMA models, developed by Box and Jenkins in 1976, can handle short-term time series forecasting by capturing dependencies and autocorrelations in the data [7]. However, ARIMA model assumes that the underlying time series is stationary, meaning that it does not exhibit trends or seasonal variations. To address non-stationarity in time series data, ARIMA applies a technique known as differencing, where the previous value is subtracted from the current value to eliminate trends. By using this method, ARIMA model can effectively transform non-stationary data, allowing them to handle long-term trends. This technique is useful for time series data like egg prices, which often exhibit trends and seasonality related to market dynamics.

SARIMA model extends ARIMA model by incorporating seasonality, making it effective in forecasting agricultural prices, where seasonal patterns are significant [16]. Especially, egg prices show seasonal demand fluctuations, particularly around holidays like Easter and Christmas [42]. ARIMA and SARIMA models can be extended to ARIMAX and SARIMAX models with exogenous variables to predict dependent variables by multivariate time series [3]. this extension enables some features like feed costs of poultry production to be added into the model to improve the accuracy.

However, while models like ARIMAX and SARIMAX perform reasonably well under certain conditions, their performances are limited when predicting non-linear or highly volatile data [30]. Both models assume that the relationship between the dependent variable and the exogenous variables is linear [3], which means the changes in the output variable as directly proportional to changes in the input variable. In reality, the relationships are not linear, especially in complex markets like agriculture. For instance, the relationship between feed costs and egg prices are not always be proportional, especially during extreme market conditions like fipronil crisis. The linearity assumption limits the abilities of ARIMAX and SARIMAX models to capture non-linear dynamics.

In addition, ARIMAX and SARIMAX assume that both input variables and output variables follow a Gaussian (normal) distribution [3], which implies that the time series data is symmetrically distributed around the mean. The egg market exhibits skewness and outliers in the data due to unpredictable factors like supply chain distruptions, disease outbreaks and war. The deviation from the normal distribution will introduce high levels of uncertainty. While ARIMAX and SARIMAX can apply differencing to handle non-stationary data, this method is still limited by the assumption of linear relationships between variables.

Furthermore, ARIMAX and SARIMAX are designed to minimize the error values when the lengths of the input and output datasets are similar and their directions are closely aligned [3]. It implies that they perform well when exogenous variables like feed prices have a strong, direct relationship with the target variables like egg prices. However, if the factors have lagged or more complex interactions with the target variables, ARIMAX and SARIMAX models might be struggle to capture the relationships effectively. Moreover, it is hard for ARIMAX and SARIMAX models to know the rising demand for holidays such as Easter, which is an important period in the poultry sector. While seasonal holidays like Easter can typically be captured by seasonal components in SARIMAX models, the challenge arises when these holidays do not fall within the same period each year, causing irregular demand spikes. In such cases, the models may struggle to account for these shifts, as they rely on consistent seasonal patterns to predict holiday-related price fluctuations.

In response to these limitations, more flexible models like Prophet have been developed. Prophet model was developed by Taylor and Letham from Facebook in 2018 to originally handle the daily business time series data forecast with weekly and yearly seasonality and holiday effects [26]. Unlike ARIMA and SARIMA, which require the data to be stationary and often involve complex parameter tuning, Prophet allows users to explore model specifications interactively, which makes it easier for users without deep experience with time series models [52]. Prophet model has the advantage of forecasting time series as it has the ability to interpolate missing values. Studies comparing Prophet to deep learning models have found that while Prophet has slightly less accuracy, it is quick to use because deep learning models with better performances like Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTMs) require longer tuning time [33].

However, Prophet model assumes that the time series components like trend, seasonality, and holidays are additive by default, which means it predicts the value by adding the effect of each component together. The default Prophet model may not capture the multiplicative interactions between components. For example, if the effects of seasonality change with the level of the trend, the default Prophet has difficulty capturing it accurately. Also, Prophet highly relies on detecting patterns in historical data, so it performs well in long-term prediction. For short-term prediction, Prophet may not capture the correlations in the residuals well because Prophet cannot identify the patterns without enough data [21]. In addition, compared to ARIMA and SARIMA, Prophet may not perform well when the data is not stationary. Also, introducing components and seasonalities may cause an overfitting model.

Traditional methods perform well when the variables are independent and simply nonlinear [50]. However, the long-term dependencies of time series are hard to capture. The application of advanced machine learning models in agricultural price predictions allows for the capture of more complex underlying patterns in time series data that traditional methods may overlook. Neural networks have shown considerable promise in improving the accuracy of agricultural price predictions compared to traditional regression analysis [30, 1]. A study compared three neural networks on egg price prediction. Their results showed that the backpropagation neural network (BPNN) recognized the patterns better, whereas the general regression neural network (GRNN) has higher forecast accuracy [1]. Other intelligence prediction methods like Support Vector Machine (SVM), Bayesian networks, and Neural networks are proved to solve high dimensional problems in forecasting agricultural product prices [50].

LSTM, a type of recurrent neural network (RNN), is particularly effective for complex, non-linear relationships and time dependence in time series data. Unlike traditional RNN models, LSTM can selectively forget or remember information to capture long-term dependencies within the data. It was introduced by Hochreiter and Schmidhuber in 1997 and has since been widely used in different fields, including agricultural price forecasting [25]. LSTM models have the advantage of capturing intricate patterns and seasonal variations inherent over longer periods, allowing them to predict agricultural prices more effectively than traditional time series models[12, 44]. For example, in studies focused on forecasting agricultural commodity prices, the LSTM model performed the best with the lowest average error, compared to models like ARIMA, Support Vector Regression (SVR), Prophet, and XGBoost. So, this thesis applied the LSTM model as an advanced machine learning forecasting technique to predict egg prices.

More studies about the application of LSTM forecasting models in agricultural sectors were researched [9, 13, 24]. Machine Learning methods have been proved to have good accuracy in predicting egg production without large volumes of historical data. LSTM with a sliding window size of 2 performs well in forecasting the drop-in production rate. In the agricultural sector, insufficient data often occurs because farmers lack statistical background and may input data inconsistently. This insight is applicable to the focus of this thesis, where models with a sliding window approach were considered during the evaluation section.

Advanced forecasting techniques can predict agricultural prices not only by historical price but also by textual data. For example, deep learning and topic modeling have been applied to predict prices using news sentiment. LSTM models proved effective in predicting market prices, distributor prices, and auction prices, illustrating the potential of combining textual data with traditional time series data for improved forecasting [13].

In addition to incorporating diverse data sources, some research has explored approaches to improve model performance. One strategy has focused on avoiding common pitfalls like overfitting and underfitting in LSTM models, which are widely used in agricultural forecasting. For instance, the introduction of batch normalization layers, reduction in LSTM layers and hidden sizes, and careful grouping of features were effective methods for enhancing the LSTM model's robustness and accuracy

[24]. Moreover, feature selection techniques such as Minimum Redundancy and Maximum Relevance (MRMR) have been implemented to reduce dimensionality and improve model performance, especially for commodities such as beans and pig grain products [62].

Forecasting approaches often struggle to predict extreme or unforeseen events such as wars, diseases, or technological advancements, which introduce significant uncertainty into predictive models. For example, a study from 1970s, developed a mathematical model to forecast egg prices in Los Angeles by stepwise multiple regression techniques. However, an unexpected increase in egg production, driven by the development of a vaccine for Marek's disease, led to a significant deviation from the forecasted prices, highlighting the model's inability to account for such disruptive events [41]. In this thesis, the Fipronil crisis and the Russian-Ukraine war as two unexpected events, influenced the accuracy of the model.

The literature discusses different methodologies for forecasting egg prices from traditional time series model like ARIMA, SARIMA, and Prophet to advance machine learning models like LSTM. While traditional models effectively capture linear trends and seasonal fluctuations but struggle with non-linear relationships and market volatility. The introduction of machine learning models, particularly LSTM, addresses these limitations, offering improved performance by capturing non-linear patterns and long-term dependencies.

The review also highlights the importance of incorporating key external factors such as feed costs, energy prices, and economic indications into forecasting models. These factors have been identified as significant drivers of egg price fluctuations and were crucial in feature selection for this thesis. In particular, soybean prices and energy costs play a major role in the accuracy of agricultural price forecasts.

By combining traditional and machine learning models with key external factors like feed and energy costs, this thesis seeks to enhance the accuracy of egg price forecasts. This approach provides a more comprehensive understanding of price dynamics in the egg market, helping to address the challenges posed by market volatility. The next section outlines how the data was collected and analyzed to develop these forecasting models.

3 Data

This section of Data provides a detailed overview of the data used for egg price forecasting and the steps taken to prepare it for modeling. This section begins with the data collection process in Section 3.1, covering egg prices and relevant exogenous variables such as feed prices, energy prices, CPI, chicken prices, and weather conditions. Following this, data preprocessing techniques in Section 3.2 are applied to clean and format the data for analysis. An exploratory data analysis in Section 3.3 is conducted to examine trends, seasonal patterns, and effects of key events, such as the Fipronil crisis and the Russian-Ukraine war on egg prices. The impact of the Easter holiday on price fluctuations is also analyzed in this section. In the end, this section concludes with descriptive statistics of the final dataset in Section 3.3.2, which describe the relationships of different variables to build accurate forecasting models.

3.1 Data Collection

Egg Prices

The first dataset for this research consists of the weekly Amsterdam Index for free-range egg prices collected over a period from the first week of 2013 to the 14th week of 2024. The data were sourced from Pluimveebeurs [37], a website that collects weekly egg prices across various markets. The prices are expressed in euros per 100 pieces and are announced every Thursday. The database tracks egg prices of sizes S, M, L, and XL for white color and brown color.

Although the Amsterdam Index is a well-known benchmark in the industry and provides pricing information, it was determined that the NOP2.0 index would be suitable for this study after further investigation. The NOP2.0 index is predominantly used by local customers, making it a more practical choice for predicting egg prices. The data were also sourced from Pluimveebeurs [40]. In March 2016, the NOP target price transitioned from version 1.0 to 2.0 by adjusting the pricing standards, and to avoid confusion between different versions, this thesis selected the NOP2.0 index. Similar to the Amsterdam Index, the NOP2.0 index guides prices for free-range and dark-yolk eggs in euros per 100 pieces, with updates announced every Friday. However, unlike the Amsterdam Index, the NOP2.0 index sets prices based on weight categories from 40 grams to 73 grams for both white and brown eggs, rather than by size categories. After consulting with local stakeholders, this thesis defines the size categories as follows: 50/51 grams as small, 58/59 grams as medium, 65/66 grams as large, and 72/73 grams plus an additional 3 euros as extra large.

Exogenous Variables

Feed Prices To explore the impact of feed prices on egg prices, this thesis incorporated monthly commodity prices for soybeans, maize and wheat from March 2016 to June 2024. These commodities are major components of poultry feed, whose prices directly influence the cost of poultry production. The data were obtained from the World Bank Commodity Market Monthly Data [60]. The prices are expressed in USD per milliton (mt). Soybeans [1] were selected as the representative soybean price of Soybean oil and Soybean meal. Wheat HRW [2] was selected as the representative wheat price due to the presence of many missing values in the Wheat SRW data. Although barley is one of the poultry feed components, it was excluded from this analysis due to incomplete data. Maize [3] was also included as a major component of poultry feed.

Energy Prices Energy cost data was also sourced from the World Bank Commodity Market Monthly Data [60]. Energy costs are a critical factor in egg production, especially as energy prices increase in 2022. Average crude oil (\$/bbl) [4], natural gas in Europe (\$/mmbtu) [5], and natural gas in

[1] Soybeans, from January 2021, U.S Gulf Yellow Soybean #2, CIF Rotterdam; December 2007 to December 2020, U.S. No. 2 yellow meal, CIF Rotterdam, according to World Bank data description.

[2] Wheat (U.S.), no. 2 hard red winter Gulf export price; June 2020 backwards, no. 1, hard red winter, ordinary protein, export price delivered at the US Gulf port for prompt or 30 days shipment, according to World Bank data description.

[3] Maize (U.S.), no. 2, yellow, f.o.b. US Gulf ports, according to the World Bank data description.

[4] Crude oil, the average spot price of Brent, Dubai and West Texas Intermediate, equally weighed, according to World Bank data description.

[5] Natural Gas (Europe), from April 2015, Netherlands Title Transfer Facility (TTF); April 2010 to March 2015, average import border price and a spot price component, including UK; during June 2000 - March 2010 prices excludes

the US ($ /mmbtu) [6] were chosen as energy costs from the commodity price data.

CPI Consumer Price Index (CPI) is the price development of a basket of goods and services typically purchased by an average Dutch household. The monthly CPI data used in this thesis were collected from CBS (Statistics Netherlands) [49]. Here, the base year is 2015, which means that the price index 2015 = 100. The CPI is an important indicator of inflation and overall economic health, influencing consumer behavior and purchasing power.

Chicken Prices The monthly commodity price of chicken sourced from the World Bank was analyzed [60]. Chicken ($ /kg)[7] price was selected to analyze as it is an agricultural product that is directly related to eggs. However, the trends analysis and correlation tests in Section 4.1.4 showed no obvious relationship between egg prices and chicken prices from the World Bank. Variable chicken price was excluded from the model development. The possible reasons are that the chicken prices are mostly collected from the American average wholesale price of frozen chicken, which means that chicken prices do not have a direct impact on laying egg production in the Netherlands.

Weather Data To comprehensively understand the factors influencing egg prices, weather data was initially considered as it is the common factor in agricultural production. This thesis collected daily weather data at Station Schiphol sourced from KNML, such as temperature, duration of sunshine, precipitation, and humidity from 2016 to 2024 [43]. However, correlation tests in Section 4.1.5 revealed no significant relationship between egg prices and weather data. Consequently, these variables were excluded from the final model to maintain the focus on more relevant predictors.

3.2 Data Preprocessing

To ensure the accuracy and consistency of different data, several preprocessing steps were applied to address issues such as missing values, duplicate inconsistent week counts, and anomalies.

Amsterdam Index

For the Amsterdam Index, missing values were identified for week 19 and week 39 in 2018. Initially, these missing entries were marked as NaN. Although interpolation was considered to fill in these gaps, it was ultimately not applied as the main egg price data was changed from the Amsterdam Index to NOP2.0. In addition, duplicate entries were discovered for week 40 in 2017. To address it, the duplicate values were averaged, ensuring consistency in the dataset.

In terms of week counts, 2015 and 2020 presented an inconsistency, each containing 53 weeks while all other years had 52 weeks. This discrepancy could impact forecasting models that depend on regular seasonality, such as ARIMA and SARIMA. To ensure consistency across the dataset and avoid introducing irregular seasonal effects, the 53rd week in both years was removed. However, for models like Prophet and LSTM, which are more flexible in handling seasonal patterns, this decision could be revisited but finally kept because the Amsterdam Index was not a data source for model development.

Anomalies in the data were also addressed. For example, in week 26 of 2017, the price for small brown eggs was recorded as 8.9, which exceeded the price of medium-sized eggs. After reviewing surrounding price trends, this value was corrected to 5.9, making it consistent with the general downward trend in egg prices at the time. Similarly, in week 32 of 2020, egg prices for XL and S sizes were swapped incorrectly. Another anomaly was identified in week 14 of 2023, where the price for small brown eggs was mistakenly recorded as 3.15. After reviewing the data, this data was corrected to 13.15 to align with other pricing data from that period.

NOP 2.0 Target Price

For the NOP 2.0 Target Price dataset, preprocessing included transforming date formats and cleaning categorical data. The original dataset used a “week-year” format for the Period column (for example, “19 - 2024”). A new Data column was created for analysis, assigning the corresponding Friday for each week, as NOP 2.0 target prices are always announced on Fridays. Additionally, egg colors were

UK, according to the World Bank data description.

[6] Natural Gas (U.S.), spot price at Henry Hub, Louisiana, according to World Bank data description.

[7] Chicken (Brazil), São Paulo State consolidated average wholesale price of frozen chicken; 2013-August 2021, Urner Barry North East weighted average for broiler/fryer, whole birds, 2.5 to 3.5 pounds, USDA grade ”A”; 1980-2012, Georgia Dock weighted average, 2.5 to 3 pounds, wholesale; previously World Bank estimates, according to World Bank data description.

originally recorded in Dutch ('wit' for white and 'bruin' for brown). These categories were converted to numeric values, with 1 representing white eggs and 2 representing brown eggs, for easier analysis of the models.

Data cleaning also involved handling price entries that included the euro symbol, such as "€ 7.070".These values were converted into plain numerical entries. A missing value in week 38 of 2019 was addressed using forward filling, followed by linear interpolation, ensuring that the data was complete and prepared for model training and evaluation.

Exogenous Variables

The original period column for the monthly data from the World Bank Commodity, which includes Soybeans, Maize, Wheat HRW, Crude Oil, Natural Gas in US, Natural Gas in Europe, and Chicken, was formatted as "2016M01". A new 'Period' column was created to convert it to a standard date format suitable for time series analysis. The monthly data was then transferred to weekly data points by linear interpolation. Each weekly data was assigned a corresponding Friday date to match the weekly frequency of the egg price data. Any missing values identified during the interpolation process were handled using linear interpolation to ensure a complete and continuous dataset.

The preprocessing of monthly CPI data was similar to those applied to World Bank Commodity data. The original CPI data was linearly interpolated to weekly data points and then every value was assigned a corresponding Friday date.

Daily weather data for temperature, sunshine duration, precipitation, and humidity were averaged to compute the weekly average values. Each computed weekly average was assigned to the corresponding Friday date.

3.3 Exploratory Data Analysis

The exploratory data analysis (EDA) section provides an examination of egg price data through data visualization and descriptive statistics. The first part, Data Visualization in Section 3.3.1 focuses on identifying trends and patterns over the year. Key analyses include yearly averages and medians to track egg price changes over time and a detailed analysis of the significant impacts of events like the Fipronil crisis and the Russian-Ukraine war on egg prices. Additionally, the seasonality of egg prices and the effect on the Easter holiday are explored to better understand recurring market fluctuations. Then, the correlation between egg sizes and colors is analyzed, leading to the conclusion that selecting white-medium eggs as the representative category for model development is a reasonable approach.

The second part, Descriptive Statistics in Section 3.3.2, provides a summary of the final dataset, including important variables such as feed costs, energy prices, CPI, and chicken prices. This statistical overview ensures that all relevant factors have been taken into account in preparation for modeling.

3.3.1 Data Visualization

Figure 2 presents the weekly NOP2.0 prices for small (S), medium (M), large (L) and extra-large (XL) eggs, classified by colors (white and brown) from week 13 in 2016 to week 26 in 2024. Each subplot represents the price trends over time for the respective egg sizes. Figure 42 in Appendix 8.3 indicates the weekly Amsterdam Index for different sizes and colors. Both trends have similar observations. In general, all size categories show similar pricing trends over the years, reflecting that market factors influencing egg prices affect the different sizes relatively equally. In addition, eggs with larger sizes have higher prices. Notably, two significant price spikes are observed from mid-2017 to mid-2018 and from early 2022 to mid-2023 for all egg sizes and colors. Moreover, since egg prices do not vary drastically, simple forecasting methods, such as Naive methods, can perform reasonably well in predicting short-term trends.

Brown eggs typically cost more than white eggs in most sizes, except for small eggs. The prices for both small white and brown eggs closely follow each other. To analyze price differences between white small eggs and brown small eggs, absolute and relative differences in Amsterdam Index prices were calculated every year. Using 0.1-point thresholds for absolute differences and 0.05-points for relative differences as criteria, it was found that price differences were consistently small in relative terms each year. However, the absolute differences exceeded the threshold in 2021, 2022, and 2024, which is consistent with the observation of small-size eggs in Figure 3 and 4. Figure 3 shows yearly

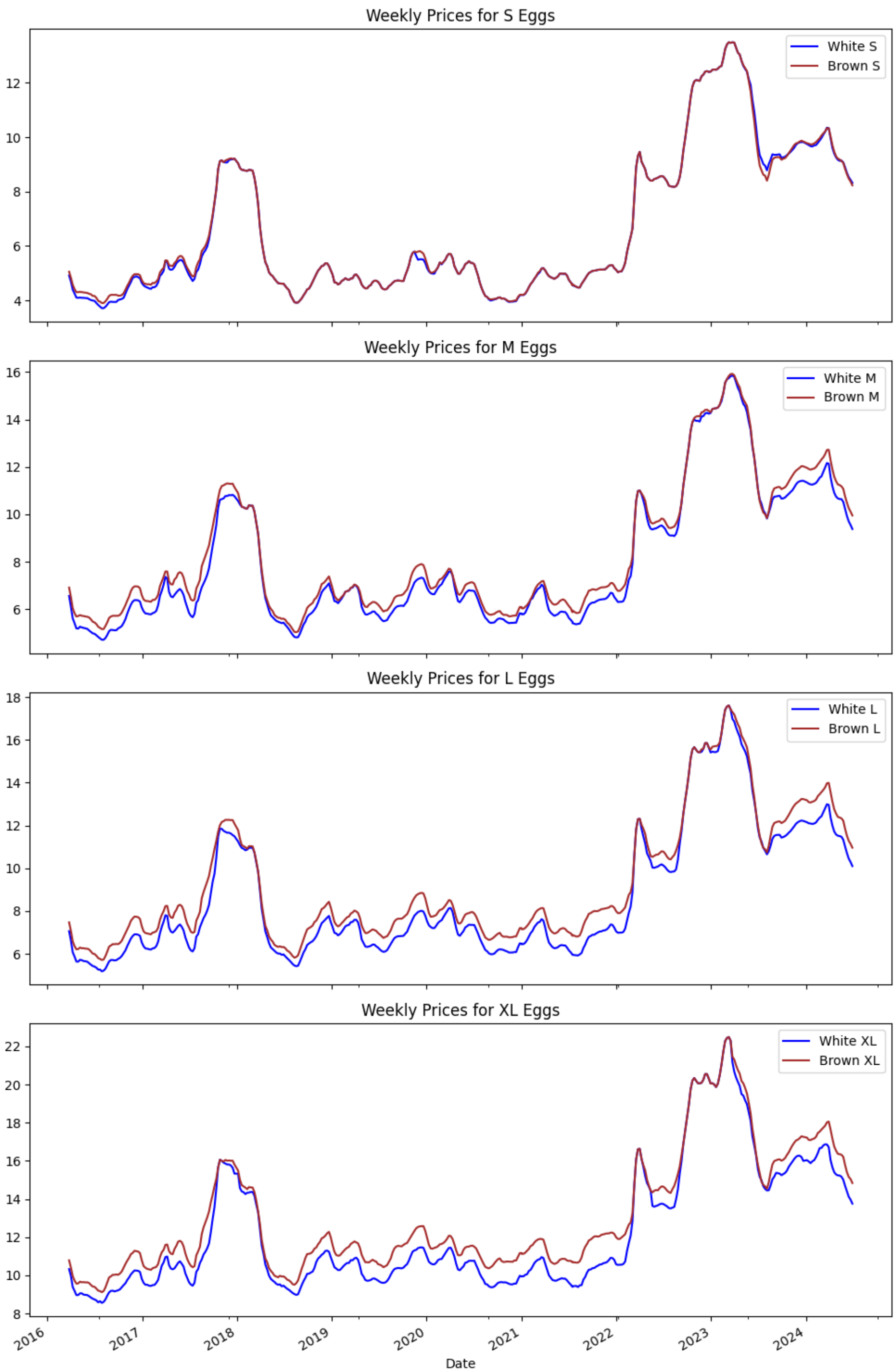


Figure 2: Trends Egg Prices over Time by Color and Size. Source: NOP2.0 from Pluimveebeurs.com

average egg prices in different sizes and colors. Given that egg prices exhibit seasonality and fluctuate throughout the year, relying solely on average prices may not provide an accurate representation of the typical pricing trends because average prices can be influenced by unusually high or low prices. The yearly median prices are shown in Figure 4 shows a central tendency. For small eggs, the prices of brown eggs are generally similar to or higher than those of white eggs. However, from the end of 2019, the price of white eggs has surpassed that of brown eggs. In addition, for other sizes, the price gap between white and brown eggs has also been reduced since 2021. Figures 43 and 44 in Appendix 8.3 illustrate the detailed differences in the average and median prices between white and brown eggs, with each graph representing a different egg size.

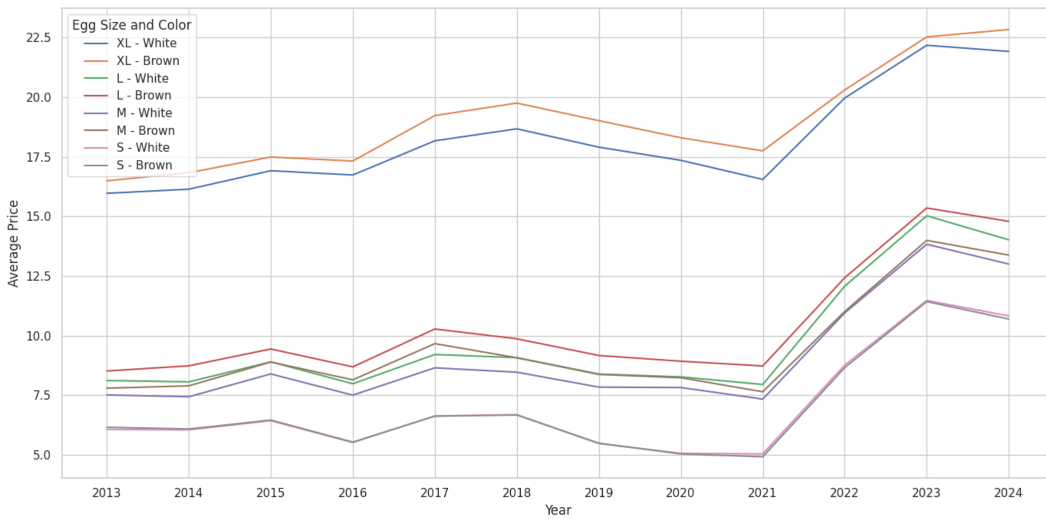


Figure 3: Yearly Average Egg Prices. Source: Amsterdam Index from Pluimveebeurs.com

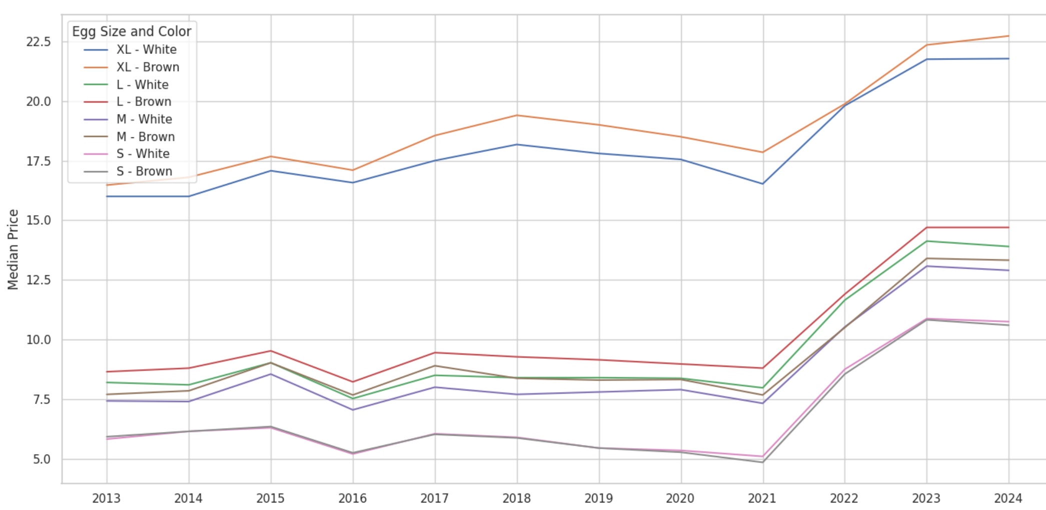


Figure 4: Yearly Median Egg Prices. Source: Amsterdam Index from Pluimveebeurs.com

To investigate the relatively small price differences between white and brown eggs for the small size category, both market and production dynamics were explored. The color of the egg depends on the breed of the hen. Hens with white feathers produce white eggs, whereas hens with brown feathers produce brown eggs. Historically, white eggs have often been produced using caging systems, and brown eggs have commonly been associated with free-range systems. Dutch consumers believe that brown eggs have higher quality due to natural production and better animal welfare [20]. The findings of a study on the acceptance of brown and white eggs by consumers indicated that customers generally have a positive attitude toward brown eggs, perceiving them as more nutritious, better tasting, higher in the content of polyunsaturated fatty acids n-3, and produced by organically fed hens [27]. Although the data pertain to free-range egg prices, eliminating cost differences between free-range and cage production methods, consumer preference for brown eggs still results in higher demand, and consequently higher prices for other sizes. Despite white eggs now offering similar quality to brown eggs, this preference consistently leads to a price premium for brown eggs across most size categories. In addition, hens with white feathers have been bred to be smaller so that more eggs can be laid [10], which leads to lower production costs. For the small egg market, the close price between white and brown eggs could indicate a more balanced supply and demand. Small white eggs in both the Amsterdam Index and the NOP2.0 industry quotation are sometimes more expensive than small brown eggs. Besides, other industry quotations were investigated. Weser Ems [39] in Germany published higher prices in small white eggs, while Kruisem handelsnotering [38] in Belgium has more expensive small brown eggs. These differences show various market dynamics.

The first peak occurred between mid-2017 and mid-2018 due to the fipronil crisis in Dutch laying hen farms. Fipronil, as a broad-spectrum phenylpyrazole pesticide, is used to repel insects such as fleas and ticks for dogs and cats, but not for any food-producing animals [48]. In 2017, a Belgian company was accused of selling the cleaning product, including fipronil, to a Dutch pesticide supply company, which sold it to poultry farms. The Netherlands was the main egg exporter to other EU countries, the fipronil contamination crisis and the following recall period caused huge damage to the laying hen sector. More than 300 farms were temporarily closed until it was confirmed that their facilities were free of fipronil. Additionally, over 100 million eggs were discarded. As forced laying hen moulting proved ineffective, more than 3 million chickens were culled. This results in a net operating loss, estimated at about €387,017 per farm for five months [47]. As shown in Figure 5, egg prices for both sizes and colors saw a significant increase around August 2017. This elevated pricing persisted until March 2018, when prices began to decline. In detail, for medium white eggs, the price was €5.66 on July 14, 2017 (week 29). During the next six months, the price nearly doubled, reaching a peak of €10.82 on December 15, 2017 (week 51).

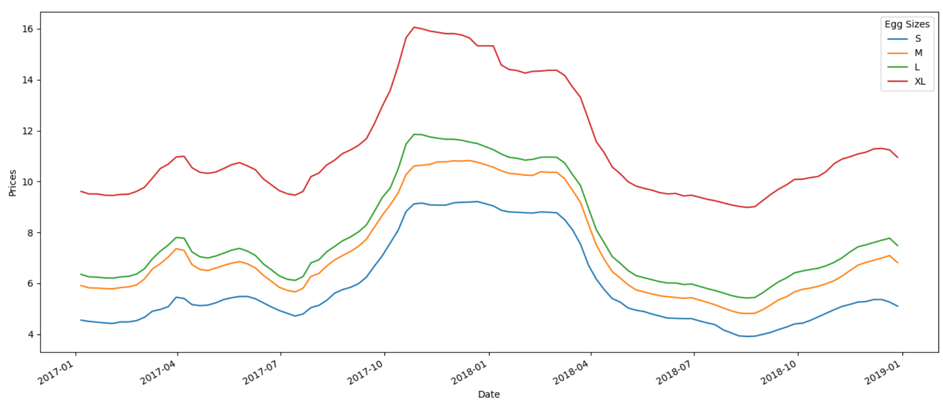


Figure 5: Egg Prices from 2017 to 2018. Source: NOP2.0 from Pluimveebeurs.com

Figure 6 plots the price of eggs from 2022 to 2023. The second peak in egg prices began in February 2022, reaching a high level in about two months. On 24 February 2022, Russia launched an attack on Ukraine, significantly impacting global food prices, including eggs. Ukraine, a major exporter of wheat, corn, barley, sunflower oil, and eggs, faced severe disruptions in its agricultural exports due to the conflict. The initial spike in egg prices can be partly attributed to the disruption in Ukraine's egg exports, which led to short-term supply shortages in the global market. In the longer term, the conflict also increased the prices of essential feed products, raising the cost of poultry and eggs around the world [53]. For instance, the price of medium white eggs was €6.51 on 4 February 2022 (week 6) and rose to €11.0 on 1 April 2022 (week 14). After April 2022, egg prices remained high, but relatively stable until September. As winter approached, the demand for energy, particularly heating, increased significantly, leading to another rise in egg prices. This trend continued throughout the winter, with egg prices peaking at €15.85 on 24 March 2023, which was the historical maximum price.

Figure 7 and 8 show the seasonal trends in egg prices by calculating the average and median prices

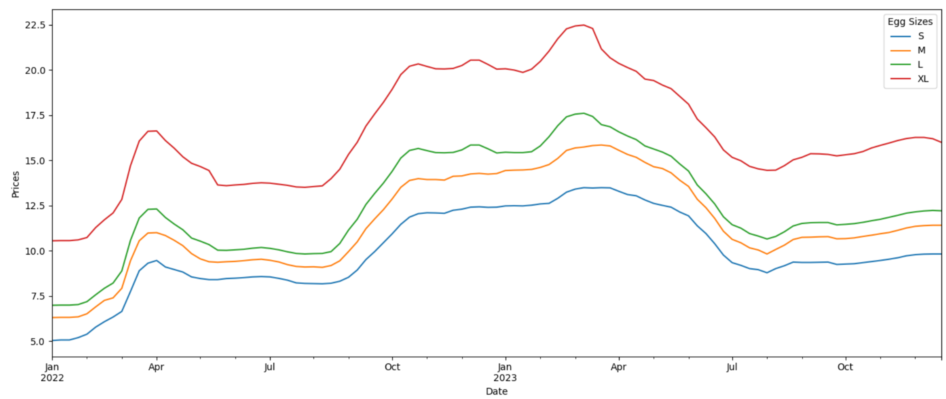


Figure 6: Egg Prices from 2022 to 2023. Source: NOP2.0 from Pluimveebeurs.com

for the corresponding weeks in different years. The average price is the arithmetic mean of all the price values. It can be skewed by extreme outliers. Historical egg prices include some spikes and drops during the period 2016 to 2024. Average prices may be influenced by extreme variations. On the other hand, the median price was then selected by representing the middle value when the prices are sorted in order. It provided a more robust indicator of egg trends with less influenced by outliers. Both graphs show a clear seasonal pattern, with prices peaking around week 12 to 14, corresponding to the Easter holiday, and another smaller rise toward the end of the year. However, the median prices tend to be slightly lower than the averages. This shows that there may be occasional outliers influencing the average, but these do not significantly impact the overall trend represented by the median. Thus, while the average captures occasional price spikes, the median provides a more balanced view, less influenced by outliers.

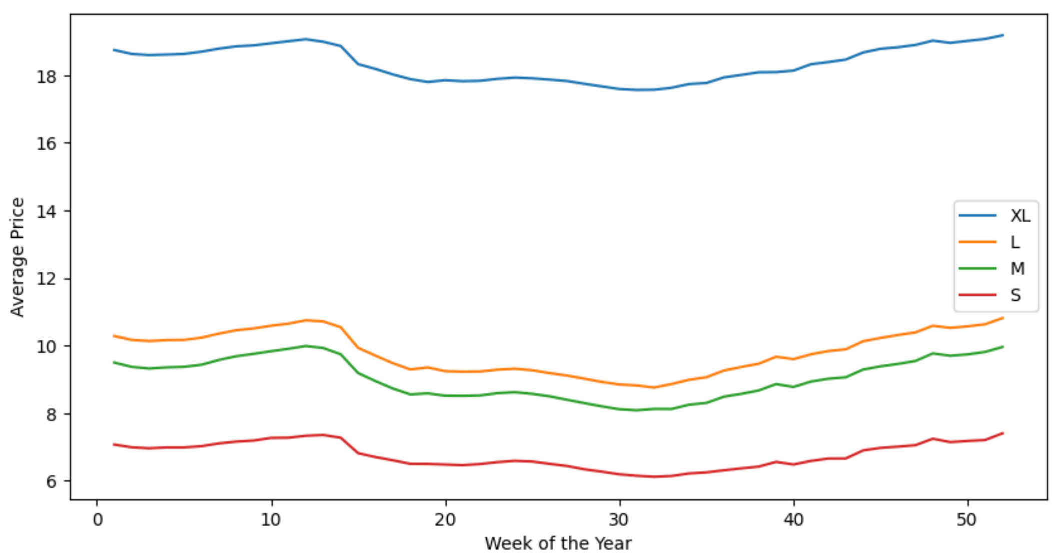


Figure 7: Seasonal Trends in Average Egg Prices. Source: Amsterdam Index from Pluimveebeurs.com

To analyze the impact of increasing demand during the Easter holiday on egg prices, the prices of medium white eggs around Easter were plotted in Figure 9. From 2016 to 2023, the corresponding Easter Sunday date was set as a dashed vertical line. The x-axis represents the number of days before and after Easter Sunday, while the y-axis represents the price of medium-white eggs. Generally, egg prices peak around Easter Sunday, after which they start to decline. However, the egg prices for 2018, shown in the green line, illustrated a decrease about a month before Easter. Easter Sunday in 2018 was on 1 April, which marked a recovery period for the egg market following the fipronil crisis in August 2017. As a result, egg prices leading up to Easter 2018 actually decreased. Moreover, Easter Sunday

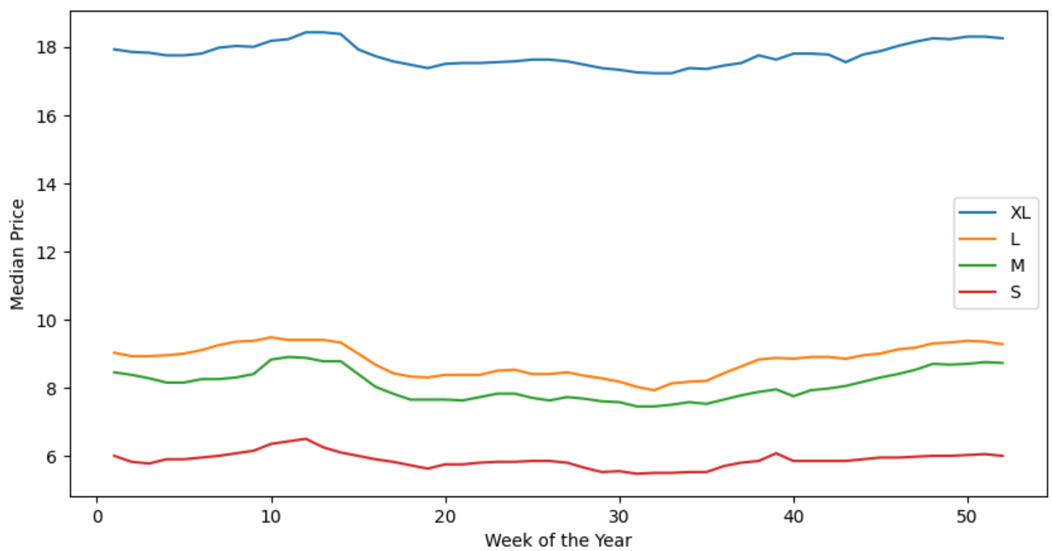


Figure 8: Seasonal Trends in Median Egg Prices. Source: Amsterdam Index from Pluimveebeurs.com

in 2022 was on 17 April, which was two months after the first invasion date. This explains why the pink line started to increase earlier than in previous years. The detailed weekly price change for every year was shown in Figure 45 in Appendix 8.3. The data from 2016 was excluded due to insufficient data coverage starting only from March 2016. With the exception of 2018 and 2022, egg prices in most years began to rise approximately two months before Easter Sunday and then started to decline one or two weeks prior to Easter.

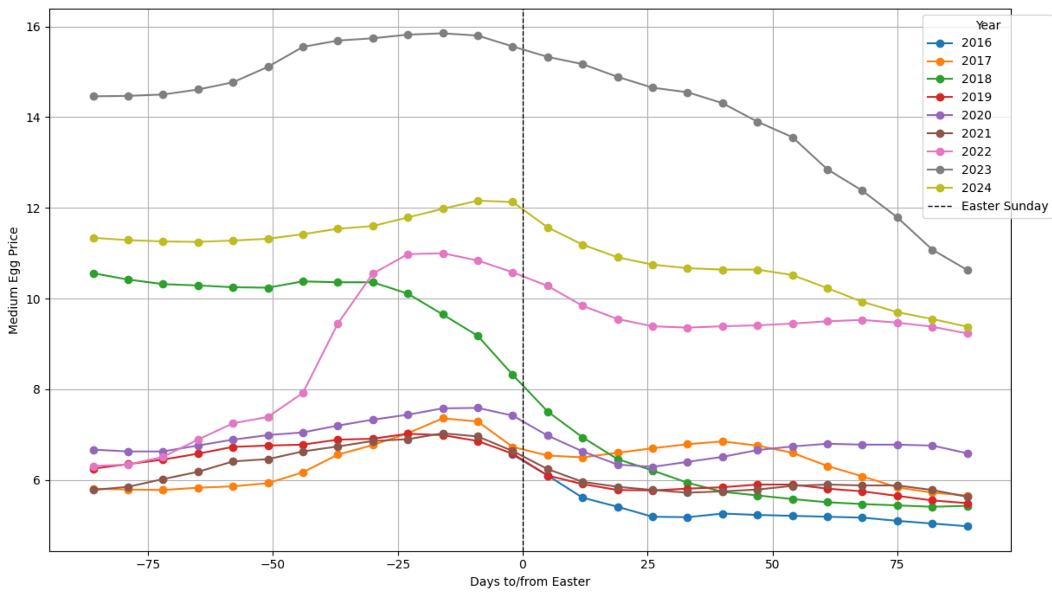


Figure 9: Medium White Egg Price Trends Around Easter. Source: NOP2.0 from Pluimveebeurs.com

The correlation analysis was tested among different egg sizes for white and brown to determine the relationships between their prices. Correlation, specifically Pearson's correlation coefficient, is a statistical measure of the linear relationship between two variables [4]. It is covariance divided by the product of the standard deviations of two variables. The formula is given by:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{1}$$

Where x_i and y_i are the ith value of the variables x and y, $\bar{x}$ and $\bar{y}$ are the mean of all values of the variables x and y, and n is the number of data points.

In general, r always lies between -1 and 1. If $r = 0$, there is no linear relationship between variables x and y. If $r = -1$, there is a perfect negative relationship between x and y. If $r = 1$, x and y have a perfect positive linear relationship. The correlation coefficient provides a simple way to check the strength and direction of the relationship between variables. In this thesis, it helped understand how changes in medium white egg prices might relate to changes in large white egg prices. While the correlation coefficient has limitations, such as potentially being misleading in non-linear relationships and not implying causation, it was still employed in this thesis to explore relationships and compare trends.

Figure 10 revealed that egg prices are highly correlated between different sizes. For white eggs, all correlations exceed 0.9. The prices of medium and large eggs are highly related, whose correlation is 1. Medium and large eggs are usually packaged and sold together in the market. The lowest correlation of egg prices is between small (S) and extra-large (XL), which is reasonable given that they cater to different market segments. For brown eggs, the lowest correlation is also between small (S) and extra-large (XL) sizes, but it is slightly lower at 0.86. The highest correlation for brown eggs is between medium (M) and large (L), which is 0.99. This indicates a more homogeneous market for brown eggs compared to white eggs, reflecting different consumer preferences and market dynamics.

In general, the correlation analysis among egg sizes is consistent with the trends observed in Figure 2, which allows for the selection of medium white eggs for further analysis and model development because the price dynamics for different egg sizes are similar.

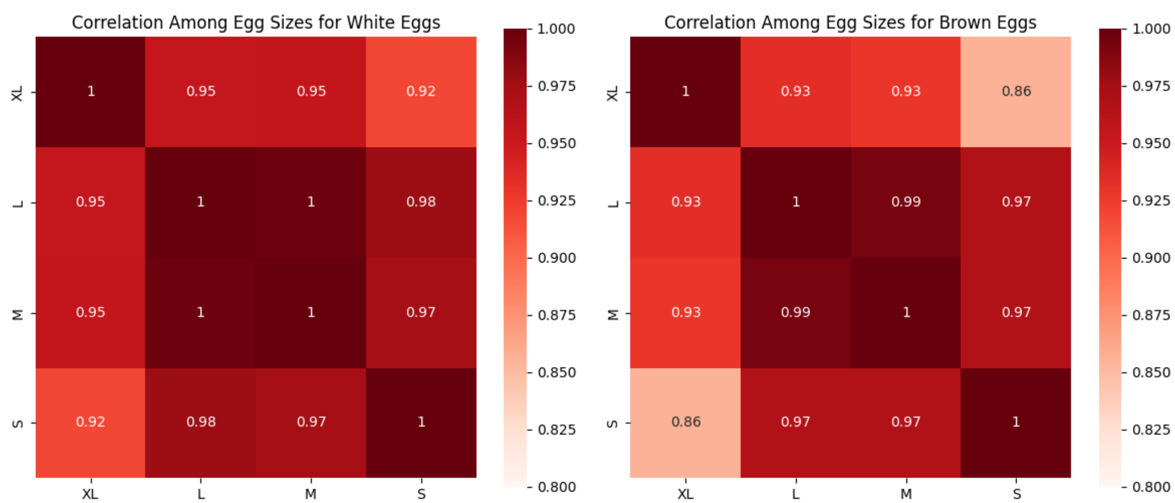


Figure 10: Correlation Among Egg Sizes for Brown and White Eggs. Source: Amsterdam Index from Pluimveebeurs.com

3.3.2 Descriptive Statistics

The dataset includes 431 historical weekly egg prices for medium white eggs and exogenous variables such as feed costs, energy costs, CPI, and chicken prices from March 25, 2016, to June 28, 2024. Table 1 describes the key statistical measures calculated for these variables. The summary helps in the initial exploration and for future analyses and model development.

The mean values indicate the average levels of each variable. The standard deviation values provide information about the variability of the data. The percentage values offer additional information about the data distribution. The Coefficient of Variation (CV) in the additional row was calculated by dividing the standard deviation by the mean. This expresses the relative variability of the dataset in relation to its mean, allowing for comparisons between data with different scales (although most variables are measured in euros, their values vary significantly).

	Medium White Egg	Soybeans	Maize	Wheats	Crude Oil	Natural Gas, US	Natural Gas, EU	CPI	Chicken
mean	7.94	482.92	206.41	262.16	66.27	3.21	12.06	111.18	1.86
std	2.83	113.02	59.68	89.19	18.35	1.43	12.57	9.76	0.31
min	4.71	343.00	143.91	143.59	22.91	1.50	1.58	100.32	1.32
25%	5.81	383.99	159.41	195.67	52.12	2.41	4.96	103.20	1.56
50%	6.67	431.08	175.51	222.28	64.21	2.83	6.99	107.31	1.85
75%	10.26	590.62	250.17	315.02	79.32	3.36	11.59	119.78	2.08
max	15.85	734.47	348.17	516.89	115.46	8.58	67.86	129.69	2.72
CV	0.36	0.23	0.29	0.34	0.28	0.45	1.04	0.09	0.17

Table 1: Descriptive Statistics

The prices of medium-white eggs exhibit moderate variability, with a mean value of 7.84 EUR per 100 pieces and a standard deviation of 2.83 EUR. The minimum price is 4.71, while the maximum reaches 15.85, showing a wide range of prices. The percentiles indicate that 50% of prices fall below 6.67 EUR, with prices in the 25th to 75th order ranging from 5.81 to 10.26 EUR. The CV for medium white egg prices is 0.36, suggesting that the prices are distributed around the mean, there are spikes and drops in the data.

Feed prices, including Soybeans, Maize and Wheats, show noticeable variability, with means of 472.9 USD, 206.41 USD, and 262.16 USD per millitons, respectively. Their standard deviations are 113.02 for Soybeans, 59.68 for Maize, and 89.19 for Wheats. The CVs for these commodities are 0.23 for Soybeans, 0.29 for Maize, and 0.34 for Wheat. These values indicate moderate to high variability, while Wheats shows the highest variability among the three, reflecting its fluctuation in the global markets, especially during the Russian-Ukraine war period.

The prices of Crude Oil and Natural Gas display notable variability. Crude Oil has a mean price of 66.27 USD per barrel with a standard deviation of 18.35 USD, resulting in a CV of 0.28 Natural Gas prices in the US have a mean of 3.21 USD per MMBtu and a standard deviation of 1.43 USD, yielding a CV of 0.45. In contrast, Natural Gas prices in the EU show extreme variability, with a mean of 12.06 USD per MMBtu, a standard deviation of 12.57 USD, and a CV of 1.04. The high CVs of energy prices, particularly in the EU, show significant market instability and variability to external factors like geopolitical events and supply chain disruption.

The Consumer Price Index (CPI) and Chicken prices exhibit relatively low variability. The CPI has a mean of 111.18 with a standard deviation of 9.76, resulting in a CV of 0.09. Chicken prices have a mean of 1.86 USD per kg and a standard deviation of 0.31 USD, with a CV of 0.17. These CVs indicate that both the CPI and Chicken prices are relatively stable with small fluctuations.

In general, descriptive statistics highlight trends and variability in egg prices, energy prices, feed prices, and consumer price indices from March 25, 2016, to June 28, 2024. Both egg and soybean prices have shown a consistent upward trend, with significant increases in the later years, as evidenced by the higher values at the 75th percentile and maximum. Energy prices, particularly crude oil and natural gas in Europe, exhibit substantial variability, reflecting the volatile nature of these markets, as well as influence of geopolitical events. The CPI values indicate a gradual increase, suggesting a rising inflation trend over the period. Although the price of chicken remains relatively stable, it shows slight variability with a standard deviation of 0.31, ranging from 1.32 to 2.72 USD per kg.

These statistics provide a comprehensive overview of the dataset, highlighting both central tendencies and variability. This is crucial for further data analysis, modeling, and forecasting.

4 Feature Analysis and Selection

As in the brief explanation of the exogenous variables in Section 3, this thesis selected soybean, wheat, and maize as three feed characteristics, crude oil, and natural gas in Europe and the US, as three energy features, and the CPI as index of inflation rate, in addition, chicken prices and weather data were also discussed. This section mainly analyzes the trends in the price of features in Section 4.1 and the feature selection process in Section 4.2.

4.1 Feature Analysis

The feature analysis section aims to explore the key variables that can influence egg prices, including feed costs, energy costs, CPI, chicken prices, and weather conditions. This section is divided into subsections for each variable, where conducting trend analyses, correlation assessments, and lag effects studies to determine their relevance as features in forecasting models.

For feed cost in Section 4.1.1, this thesis analyzes Soybean, Maize, and Wheats prices, examining their trends over time and their relationships with egg prices. Through correlation analysis, the importance of each feed component is identified and Soybean is selected as the representative variable to analyze the relationships with egg prices. Then, the lag effect of three feed variables on egg prices is investigated, determining how many weeks of lagged data to include as features for feed costs in the model.

Similarly, Crude Oil, US Natural Gas, and European Natural Gas as energy cost variables are analyzed in Section 4.1.2. These energy prices are subjected to the same trend analysis, correlation, and lag analysis process as feed costs. The same approach as with CPI is employed, assessing its relationship with egg prices in Section 4.1.3.

For chicken prices in Section 4.1.4, the trends along with egg prices are plotted but ultimately decided to drop it as a feature due to the lack of strong correlation. Additionally, a correlation test on four weather variables is performed to assess their relationship with egg prices in Section 4.1.5. Given the low correlation results, the weather is excluded as a feature in the model.

4.1.1 Feed Costs

Figure 11 shows the price trends for three agricultural commodities, Soybeans, Maize, and Wheat (HRW) from 2016 to 2024. Soybean prices were high in 2016, gradually reducing and reaching a low around the end of 2016. There is a small rise with little fluctuation until a sharp peak in 2021, followed by another peak in 2022.

Maize prices are relatively stable compared to soybeans. Prices show small variations from a little sharp in 2016 but remain stable until late 2020. Similarly to soybeans, maize prices experience two significant increases in early 2021 and early 2022. Following two peaks, prices decrease steadily through 2022 and stabilize in 2024.

Wheat prices have similar trends to soybeans and maize. Prices are relatively stable with small seasonal variations until the end of 2020. Different from soybeans and maize, it remains a relatively stable but high price at the beginning of 2021. After that, a more sharp increase is observed at the beginning of 2022.

At the beginning of 2018, the trade war between the United States and China occurred. Although tariffs affected markets, the prices of commodities were unrestrained in a significant way during this period. In addition, the COVID-19 started in early 2020. The commodity markets were stable until the end of 2020. Martignone et al. explain that there were potential lag effects of COVID-19 on global supply chains [32]. Moreover, the conflict between Russia and Ukraine caused a spike in feed commodities in early 2022. Russia and Ukraine are important producers of wheat. The market disruption caused by the war led to a more significant price shock.

Upon examining the price trends of soybean, maize and wheat from 2016 to 2024, a high degree of correlation was observed among these commodities. To quantify it, the correlation test in Figure 12a showed the relationship between the prices of three feed components. The highest correlation is between Soybeans and Maize, which is 0.95, while Soybeans and Wheat have the lowest 0.92 correlation. The results confirm strong positive correlations, which is consistent with results of [32].

The high cross-relatedness between the three feed components might cause overfitting and multi-collinearity problems for model development. Then, the correlation coefficients between the prices of

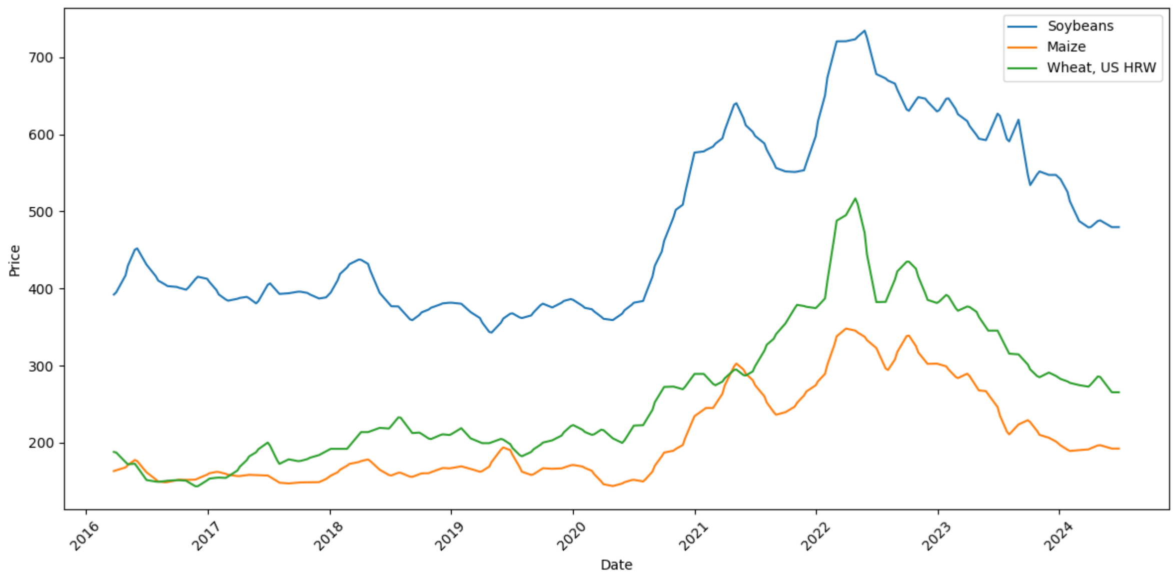


Figure 11: Comparative Trends in Commodity Prices for Soybeans, Maize, and Wheat. Source: World Bank

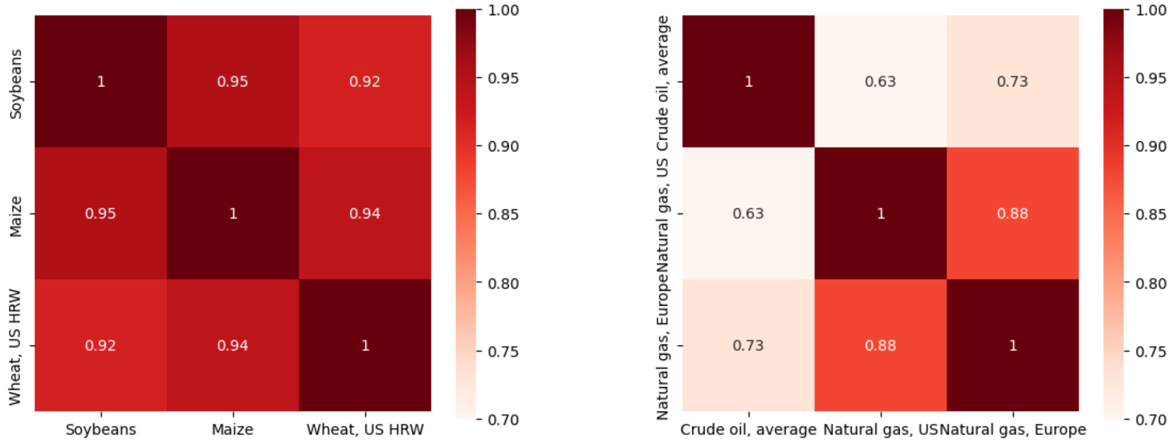
three feed components and the price of medium white eggs were calculated in Figure 13. Despite Wheat exhibiting the highest correlation with medium white egg prices among the three feed components, Soybeans were chosen as a primary feature in the analysis. This decision is supported by the literature review that identifies soybeans as a leading price transmitter in the commodities market [32]. Also, soybean and its products are one of the main ingredients of poultry feed. And correlation coefficient between soybean prices and egg prices was also significantly high. So, incorporating soybean prices as a feature allows the models to capture broader market trends more effectively.

To further investigate the relationship between soybean prices and egg prices, Figure 14 plots the trend of soybean prices and medium white egg prices. It suggested a noticeable lag effect, where changes in soybean prices (orange line) seemed to precede changes in egg prices (blue line). The first peak of egg prices in 2017 was the lag effect of rising soybean prices in mid-2016. After peaking in mid-2016, soybean prices began to decline, followed by a corresponding decrease in egg prices. Then egg prices increased dramatically from July 2017 to April 2018. During this period, there was no obvious relationship with soybean prices due to the Fipronil crisis.

From mid-2018 to mid-2019, egg prices experienced two peaks. The first peak can be partly explained by seasonal patterns, as egg prices are typically higher during winter, as shown in Figure 8. The second peak in mid-2019 was probably due to the increase in the soybean price in mid-2018. From mid-2019 to mid-2020, egg prices remained high and fluctuated level due to the global supply chain disruption of COVID-19. After 2021, there is a clear lagged relationship between soybean prices and egg prices.

The visual inspection of Figure 14 suggests a strong correlation between the prices of medium white eggs and soybeans, with a lag of about one year. To determine exact lag periods, cross-correlation analysis was performed and plotted in Figure 15 by setting 52 weeks as the maximum lag. The line shows that the correlation is constantly increasing until the highest 0.8 in 49 weeks lag. Significant correlations (above the 0.7 threshold) were identified in the range of 28 to 52 weeks.

Cross-correlation analysis and visual inspection suggest that soybean prices significantly influence egg prices with a delay. Based on these findings, a lag of 40 weeks was selected for incorporating soybean prices into the egg price-prediction models. It was chosen because it lies within the range of significant lags, representing a balanced central point. Similarly, the 34-week lag of Wheat prices on egg prices and the 39-week lag of Maize prices on egg prices were incorporated into the model as feed costs.



(a) Correlation Matrix for Feed Prices

(b) Correlation Matrix for Energy Prices

Figure 12: Correlation for Feed Prices and Energy Prices. Source: World Bank

4.1.2 Energy Costs

Figure 16 plots the comparative trends in three energy commodities, Crude Oil, Natural Gas in the US, and Natural Gas in Europe from 2016 to 2024. Crude oil prices exhibit significant fluctuations during the period. From 2016 to mid-2018, crude oil prices showed a steady upward trend, reaching approximately $80 per barrel by the end of 2018. In May 2018, US President Donald Trump announced sanctions against Iran, which led to an increase in oil prices [45]. In response, Saudi Arabia, the largest oil exporter within OPEC, increased its production. Additionally, non-OPEC members like Russia and the United States accelerated their oil production. At the end of 2018, the Trump administration eased sanctions, resulting in a drop in oil prices.

In 2020, COVID-19 disrupted global demand and supply chains. The crude oil markets became more uncertain due to the Russia-Saudi Arabia oil price war in March and April 2020 [31]. It caused crude oil future contract prices to fall below zero on the New York Mercantile Exchange (NYMEX) on April 20, 2020. After that, the prices of crude oil recovered and experienced a significant increase, peaking around $110 per barrel in early 2022. From mid-2022 to 2024, crude oil prices stabilized but are still volatile, which shows the post-influence of conflict in 2022.

Compared to crude oil and European natural gas prices, Natural gas prices in the US remain relatively stable and low. The United States is the largest natural gas-producing country in the world. The market is more stable. When the significant increase in crude oil prices occurred in 2018, 2021, and 2022, only slight increases were observed in the prices of US natural gas. Prices can quickly return to the previous range.

Compared to the US, European natural gas prices show greater volatility. From 2016 to 2020, similar to the US, the prices remain low and stable. From mid-2021, a dramatic increase in natural gas prices is observed in Europe. The prices peaked at around $90 per MMBtu in early 2022. Previously, European pipelines were designed primarily to import from Russia [17]. Regional geopolitical tensions disrupted the EU natural gas supply. It caused the European natural gas market to become more vulnerable and volatile compared to the US market.

Similarly to the selection steps for the feed costs, the correlations between the energy commodities were analyzed and presented in Figure 12b. In comparison to the three feed commodities, the correlations between the three energy commodities are lower. The highest correlation, 0.88, is observed between US natural gas and European natural gas, reflecting the same energy products in different markets. Conversely, natural gas in the US has the lowest correlation with crude oil, at 0.63. This finding is consistent with the trends depicted in Figure 16.

Although there was a relatively lower cross-relatedness between the three energy commodities, the potential overfitting and multicollinearity were also considered in the analysis. Crude oil was selected as the main energy feature to analyze due to its higher correlation coefficient with the price of medium-white eggs, compared to the prices of natural gas, as shown in Figure 13. Additionally, European natural gas prices exhibit a stronger relationship with egg prices than US natural gas prices,

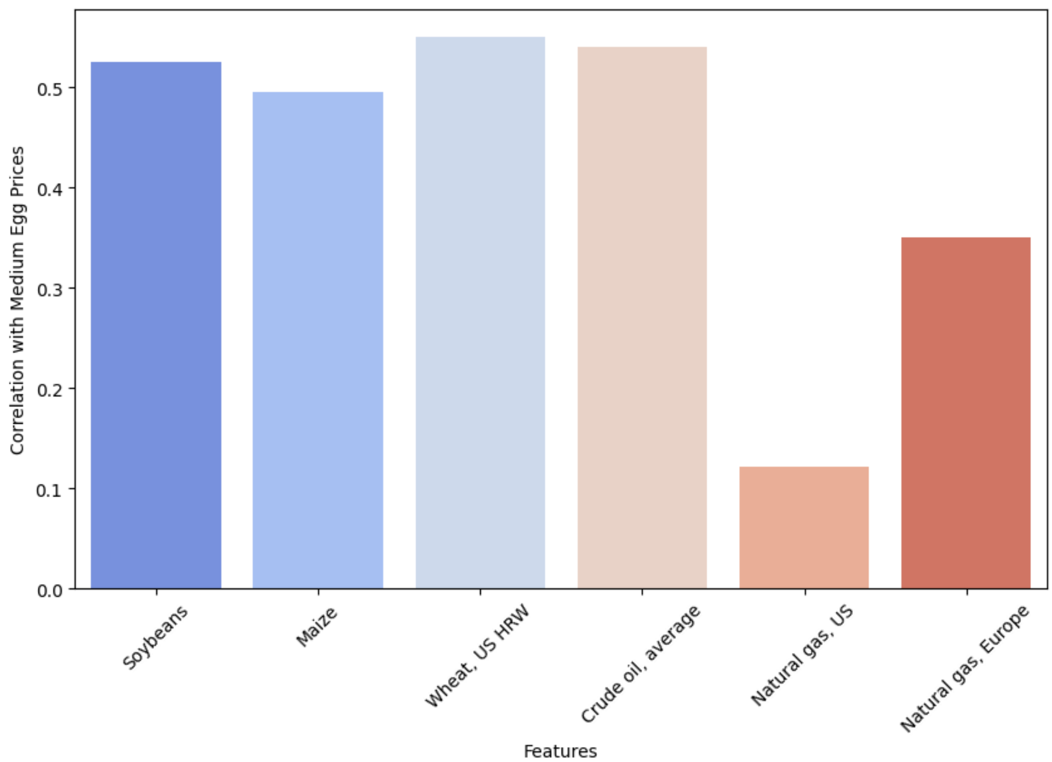


Figure 13: Correlation of Feed and Energy Costs with Medium White Egg Prices.

indicating that egg prices are more impacted by the local energy market. Similarly, the 35-week lag for natural gas in the US and the 36-week lag for natural gas in Europe on egg prices were selected into the energy features.

Figure 17 illustrates the price trends of medium-white eggs and crude oil. In 2018, trends indicate some co-movement between crude oil (orange line) and egg prices (blue line). The significant peak in egg prices around early 2018 coincides with the recovery of the Dutch egg market following the Fipronil crisis. Similarly, the decline in crude oil prices in late 2018 aligns with the easing of US sanctions against Iran. In early 2022, both crude oil and egg prices exhibited a downward trend during the same period. From the end of 2022, egg prices show a clear lag effect of crude oil prices.

The analysis highlights that while energy costs are not the primary cost in egg production, their influence cannot be ignored, especially during periods after 2022. The observed trends support the inclusion of crude oil prices as a feature in egg price prediction models. However, significant lagged correlations in both feed and energy costs likely share the same origin: the disruptions caused by the Russia-Ukraine war. This suggests that the strong correlations are context-specific and should be interpreted with caution, especially for future forecasts that may not face similar geopolitical factors. Furthermore, while energy costs influence egg prices, their lag effect is less pronounced compared to feed costs, indicating that energy plays a secondary role in price fluctuations.

The same strategy was applied to identify the number of weeks of lag between crude oil and egg prices. With the setting of 52 weeks as the maximum lag, the best lag for the prices of medium white eggs and crude oil is 40 weeks with a correlation of 0.765, as shown in Figure 18. Significant correlations above the 0.7 threshold were identified in the range of 26 to 52 weeks. A 39-week lag was selected to incorporate crude oil prices into forecasting models as the median of the significant correlation range.

Similarly, a 35-week lag of feature Natural Gas in the US, and a 36-week lag of feature Natural Gas in the EU were selected as lag correlations of the rest two energy prices.

4.1.3 CPI

Figure 19 presents the price trends for medium white eggs and the general inflation measured by the CPI. CPI (orange line) remains relatively stable before 2017. Between 2017 and 2021, the CPI showed

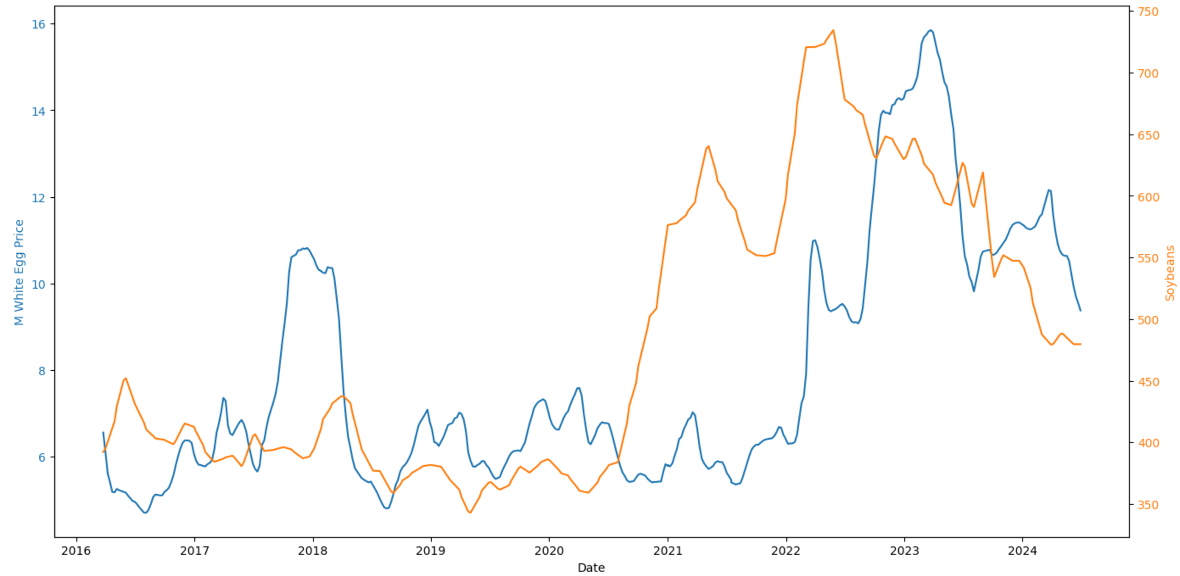


Figure 14: Price Trends of M White Egg with Soybeans.

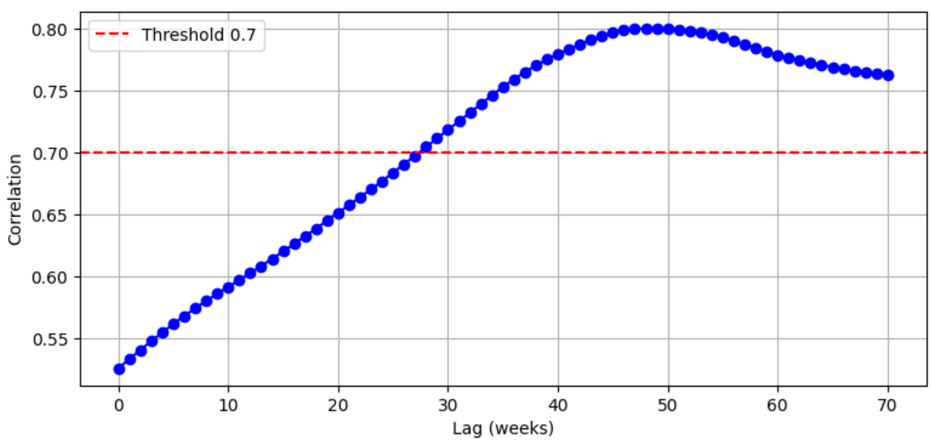


Figure 15: Lagged Correlation between Medium White Egg Prices and Soybeans.

a gradual increase. Egg prices still exhibit periodic spikes, which could be influenced by seasonal factors or supply chain disruptions. The moderate correlation between CPI and egg prices suggests that the inflation rate influences egg prices, but other factors, such as feed costs and energy costs, play a more important role. From early 2021, both the CPI and egg prices show significant upward trends, indicating that inflationary pressures in the entire economy impact egg prices. A delayed effect can be observed in this period. After peaking in 2022, CPI continues to rise at a slower pace, while egg prices generally follow a downward trend after the middle of 2023.

Overall, general inflation has a lag influence on egg prices. Although the CPI is not the most important factor in egg prices, its influence should be considered, particularly during the period of economic fluctuation in 2021-2023. The lagged correlation between the price of medium white eggs and the CPI was calculated, shown in Figure 20. Different with lag correlation with soybean and crude oil, correlations are decreased gradually by increased week number. The best 1-week lag was selected with a correlation of 0.764 to incorporate as a feature in the forecasting models. As the weekly CPI was obtained from linear interpolation of monthly CPI data published by CBS [49], the CPI one week ago here was CPI last month. It is necessary to consider general economic indicators in egg price prediction models.

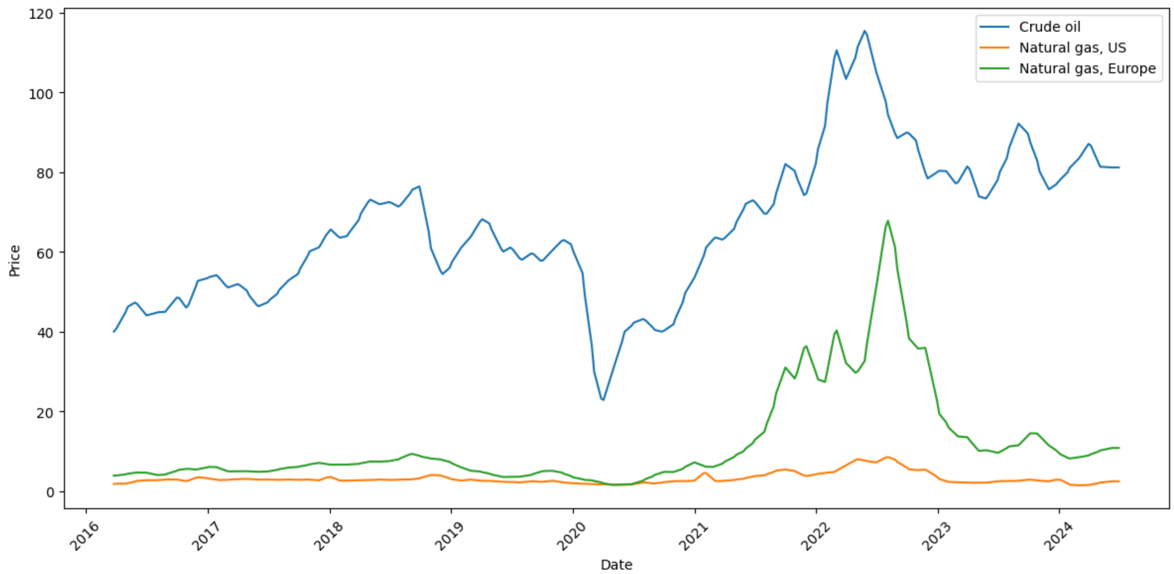


Figure 16: Comparative Trends in Commodity Prices for Crude Oil and Natural Gas. Source: World Bank

4.1.4 Chicken

As the relevant agricultural product with eggs, chicken was considered one of the characteristics. Figure 21 presents the price trends for medium white eggs (blue line) and chicken (orange line). The chicken meat market shows a larger fluctuation compared to egg prices. From 2016 to early 2019, both prices exhibit significant fluctuations. Egg prices show several peaks, particularly in early 2018, coinciding with the recovery of the Dutch egg market following the Fipronil crisis. Chicken prices also fluctuate during this period, but trends do not align consistently with movements in egg prices.

Between mid-2019 and mid-2021, both commodities continue to exhibit variability. Egg prices show periodic peaks and troughs, while chicken prices also fluctuate but follow a different pattern. From 2021 onward, egg prices experienced significant peaks, especially in early 2022, while chicken prices sharply increased in early 2021 and downward to a trough in 2022.

In early 2022, egg prices and chicken prices showed an upward trend due to the war. Chicken prices had an opposite trend to egg prices in mid-2022. From mid-2023, both prices experienced increases and stabilized in 2024.

There appears to be no clear correlation between the prices of eggs and chickens. The results of the lagged correlation test showed that the best lag for chicken price is 1 week with a correlation of -0.466. As a result, chicken prices were not considered a feature in this thesis to avoid introducing extra noise and reducing the accuracy of the model.

4.1.5 Weather

At the beginning of the investigation, weather data, including temperature, humidity, precipitation, and duration of sunshine, were tested in correlation with medium white egg prices from the Amsterdam Index. Figure 22 shows the heat map of the results. The most significant correlation between weather variables and egg prices is with temperature, showing a correlation coefficient of -0.23. This indicates that egg prices tend to be higher when the temperature is lower. This result aligns with the seasonal trend in Figure 8, where egg prices are typically higher during the winter months.

Given the relative correlation between weather variables and egg prices, weather data were excluded from the final dataset and feature set for the model development. However, the seasonality of the price of eggs was important to consider in the models.

Figure 17: Price Trends of M White Egg with Crude Oil.

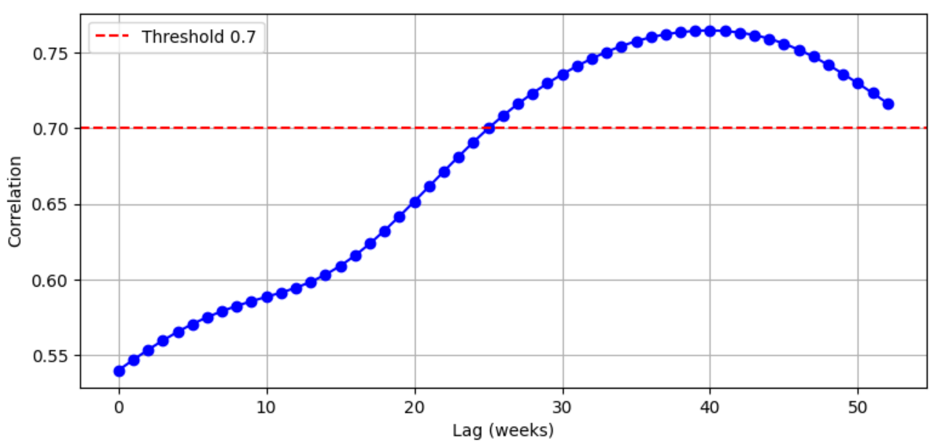


Figure 18: Lagged Correlation between Medium White Egg Prices and Crude Oil.

4.2 Feature Selection

After a thorough analysis of various potential features, the relevant features selected for the egg price prediction models are soybean prices, crude oil prices, and the Consumer Price Index (CPI). They were chosen based on their significant correlations with medium-white egg prices, and they can capture key cost factors influencing the egg market.

Figure 23 illustrates the price trends of egg prices with three selected features. Min-Max rescaling method for direct comparison of changes and trends normalized the prices. From 2016 to 2019, the normalized prices of soybeans, crude oil, and CPI all show upward trends, with soybeans and crude oil exhibiting more volatility and more related trends. From mid-2021, normalized prices of soybeans, crude oil, and CPI show pronounced and similar fluctuations, indicating the cross-relation between the three features.

The analysis reveals that soybeans, crude oil, and CPI all influence egg prices, while they are correlated. As shown in Figure 24, three variables soybeans, crude oil, and CPI have relatively strong positive correlation, which means that there is a greater chance that the other variables will also increase when one of the variables increases, indicating potential multicollinearity.

To ensure the robustness of the forecasting models, the thesis applied the Variance Inflation Factor (VIF) to assess the degree of multicollinearity among the selected features. The VIF is a critical

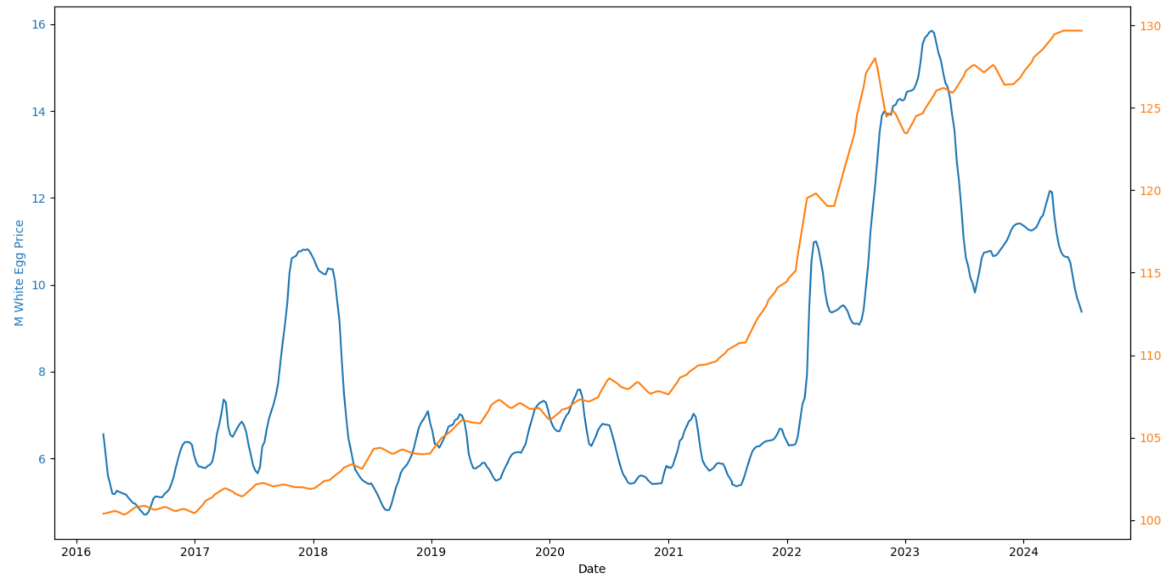


Figure 19: Price Trends of M White Egg Price with CPI.

diagnostic tool to ensure that the model is stable and the regression coefficients are interpretable. The formula of VIF is [2]:

$$VIF_i = \frac{1}{1 - R_i^2} \tag{2}$$

where R_i^2 is the R-squared value from the regression model where the ith predictor is regressed on all the other predictors.

If the predictor is not correlated with the other predictors, VIF equals 1. In this thesis, the values of VIFs of soybeans, crude oil, and CPI are 2.665, 2.657, and 2.444 respectively. Both values are lower than 5, showing that the multicollinearity between them is not strong. Although they have a certain correlation, they will not have a significant negative impact on the stability of the model.

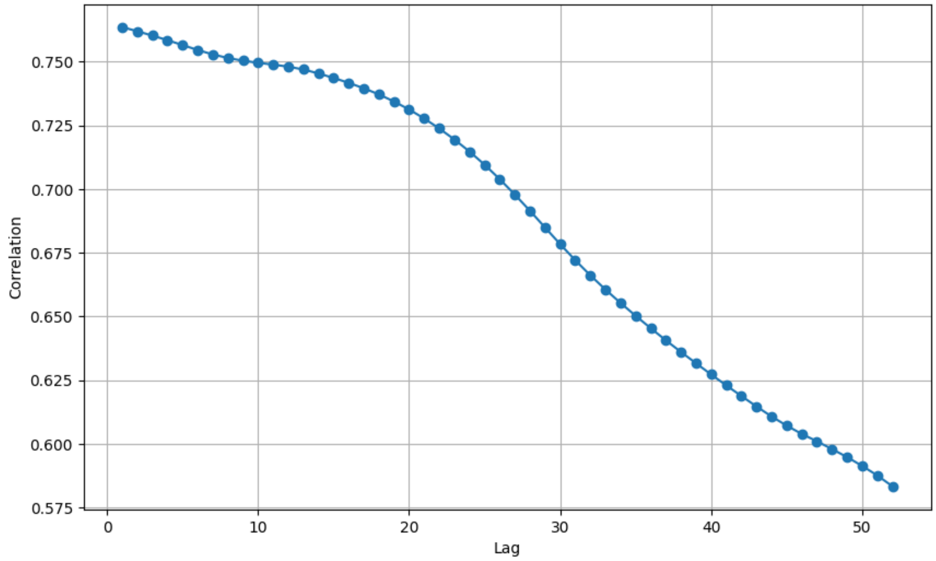


Figure 20: Lagged Correlation between Medium White Egg with CPI.

5 Methodology

The methodology section of this thesis outlines both the theoretical background and the practical implementation of the forecasting models used to predict egg prices. The theoretical background in Section 5.1 delves into the key concepts behind traditional time series models such as ARIMA and SARIMA as well as more complex models like Prophet, LSTM, and simpler benchmark models such as Naive and Hyperbolic forecasting. Furthermore, this section introduces the evaluation metrics in Section 5.1.6 used to assess model performance, including their formulas, and the pros and cons of each metric in the context of time series forecasting. Section 5.2 explains the practical steps for applying ARIMA, SARIMA, Prophet, and LSTM models, including hyperparameter tuning.

5.1 Theoretical Background

5.1.1 ARIMA

AutoRegressive Integrated Moving Average (ARIMA) model is a time series forecasting model that combines AutoRegressive (AR), Integrated (I), and Moving Average (MA) components [26]. AR model assumes that the past values have a linear influence of p-lag observations on the current values. I part involves d as the degree of differencing to make a non-stationary time series stationary. Differencing calculates the differences between consecutive observations to stabilize the mean and reduce the trend of seasonality of a time series. For example, first-order differencing is represented as $\Delta y_t = y_t - y_{t-1}$. MA models use past forecast errors q in the regression model.

The AR(p) model is represented as

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \tag{3}$$

where ϕ_1, ...ϕ_p are the coefficients of the autogressive terms, c is constant, y_{t-1}, ...y_{t-p} are lagged values of the series, and ε_t is white noise.

The MA(q) is written as

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \tag{4}$$

where θ_1, ...θ_q are the estimated parameters, c is constant, ε_t is error term, and ε_{t-1}, ...ε_{t-q} are lagged error terms.

ARIMA (p, q, d) model is combined with three parts together and written as

$$y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \tag{5}$$

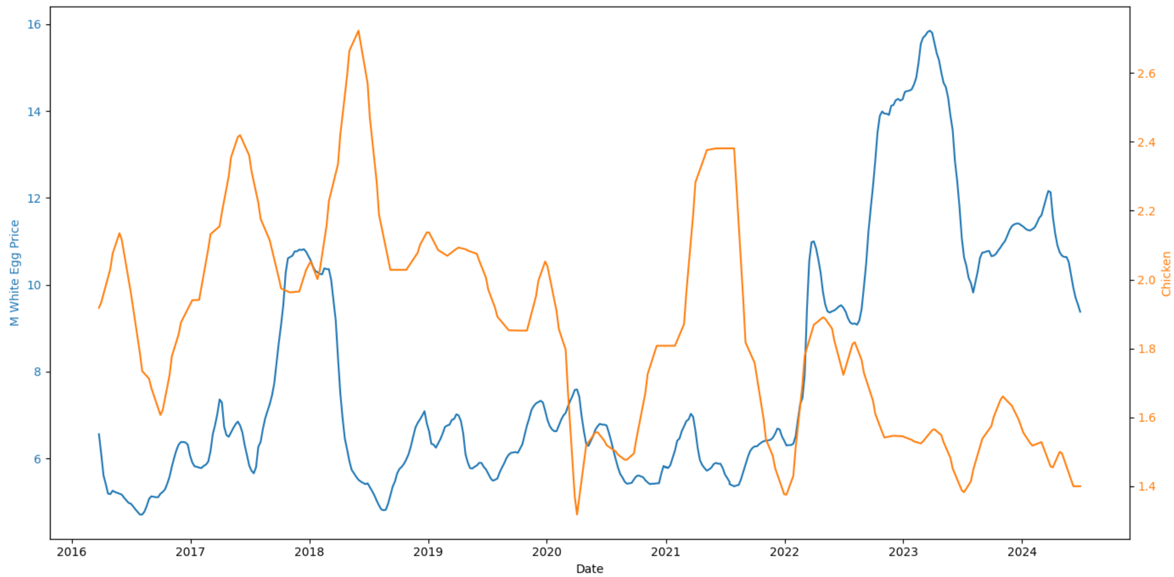


Figure 21: Price Trends of Medium White Egg with Chicken.

Similarly, p and q are the orders of the AR model and MA model separately. y_t' is the differenced series.

Define B as the backshift operator, $\phi(B)$ represents the autoregressive part, and $\theta(B)$ is the moving average part, to the differenced series $Y_t = (1-B)^d X_t$, ARIMA model is written as

$$\phi(B)Y_t = \theta(B)Z_t, \quad \{Z_t\} \sim \mathrm{WN}(0, \sigma^2) \tag{6}$$

where d is the order of differencing and the process X_t is stationary if and only if $d = 0$, which is an ARIMA(p, q) process [8].

Usually, the values of p and q can be determined by Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) respectively. ACF measures the relationship between y_t and y_{t-k} for different lag values k. PACF overcomes the issues of ACF, that mislead correlations between two non-consecutive observations because of their individual connections to the middle observations instead of their direct connections. After removing the effects of intermediate lags $(1, 2, 3, ..., k-1)$, each subsequent partial autocorrelation eliminates the influence of the preceding lags, providing a clearer picture of the direct relationship.

ARIMA model assumes that the time series is stationary. Non-stationary data will cause poor forecasts. This paper performs the Augmented Dickey-Fuller (ADF) Test, which hypothesizes that the series has a unit root and is non-stationary. To make data stationary, methods such as differencing, log transformation, square root, and Box-Cox transformation are applied to stabilize variance.

Parameters (c, ϕ_i, θ_i) of ARIMA models are estimated by maximum likelihood estimation (MLE), which adjusts the parameters iteratively to maximize the probability of observed data. This thesis employs $'statsmodels'$ library in Python to perform MLE.

Akaike's Information Criterion (AIC) is also used to determine the order of an ARIMA model. It is written as

$$AIC = -2\log(L) + 2(p + q + k + 1) \tag{7}$$

where $\log(L)$ is the natural logarithm of the likelihood L, and k is the number of additional parameters estimated in the model.

AIC can be corrected for ARIMA models as

$$AIC_c = AIC + \frac{2(p+q+k+1)(p+q+k+2)}{T-p-q-k-2} \tag{8}$$

where AIC_c is corrected AIC for small sample sizes, and T is the sample size.

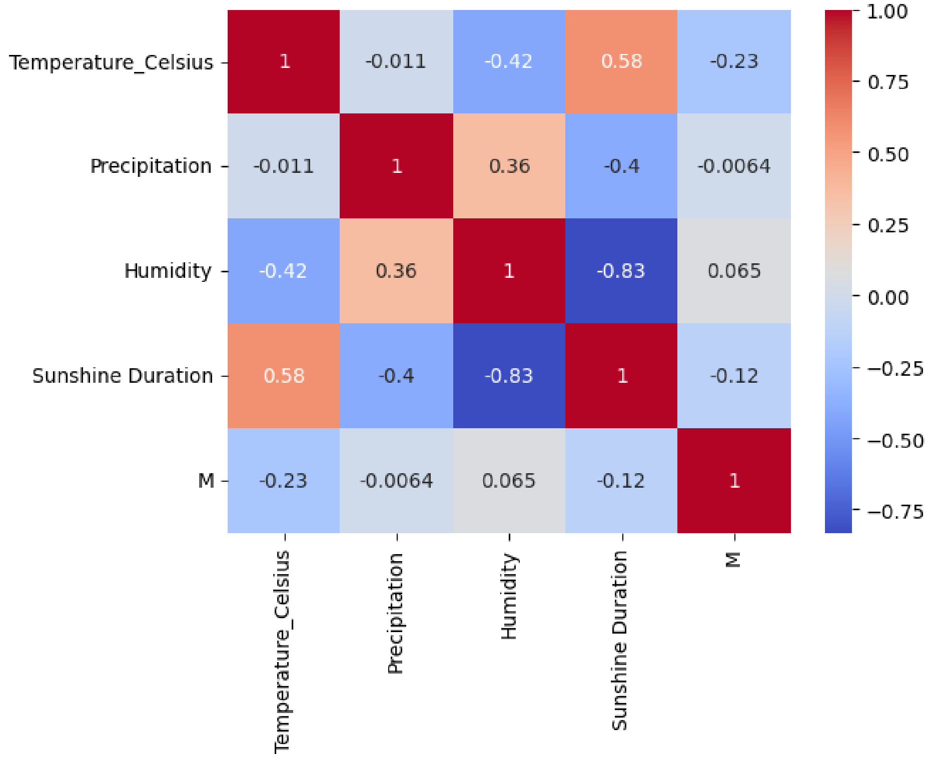


Figure 22: Correlation for Weather and Medium White Egg Prices

Bayesian Information Criterion (BIC) is another criterion to determine the order of an ARIMA model:

$$BIC = AIC + \log(T) - 2(p + q + k + 1) \tag{9}$$

The selection of an optimal model is based on minimizing criteria such as AIC, AIC_c, or BIC. In this thesis, preference is given to using AIC_c.

The standard ARIMA model is designed to handle univariate time series data. To consider other factors influencing the time series, ARIMA model can be extended to an ARIMAX model (AutoRegressive Integrated Moving Average with eXogenous variables) to incorporate some external predictions or covariates. The ARIMAX formula is given by extending Equation 5:

$$y'_t = c + \sum_{i=1}^{p} \phi_i y'_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \sum_{k=1}^{m} \beta_k X_{k,t} + \epsilon_t \tag{10}$$

where $X_{k,t}$ is the additional prediction at time t, m is the number of exogenous variables, and coefficient β_k represents the impact of exogenous variables on the differenced time series y'_t.

So, ARIMAX model as the extension of ARIMA model, is capable of considering both internal patterns through AR and MA components, as well as external predictions through exogenous variables. ARIMAX models can predict time series data influenced by multiple factors.

5.1.2 SARIMA

ARIMA model performs well for non-seasonal data. Seasonal AutoRegressive Integrated Moving Average (SARIMA) model is developed to extend ARIMA models to handle seasonal patterns in time series data [26]. SARIMA model is denoted as ARIMA $(p, d, q)(P, D, Q)_s$, where (p, d, q) is non-seasonal part of the model, (P, D, Q) is seasonal part of the model, and s is the period of seasonality.

For seasonal components of SARIMA model, seasonal differencing (SD), seasonal AutoRegressive (SAR), and Seasonal Moving Average (SMA) explain the relationship between an observation from the same season in the previous cycle.

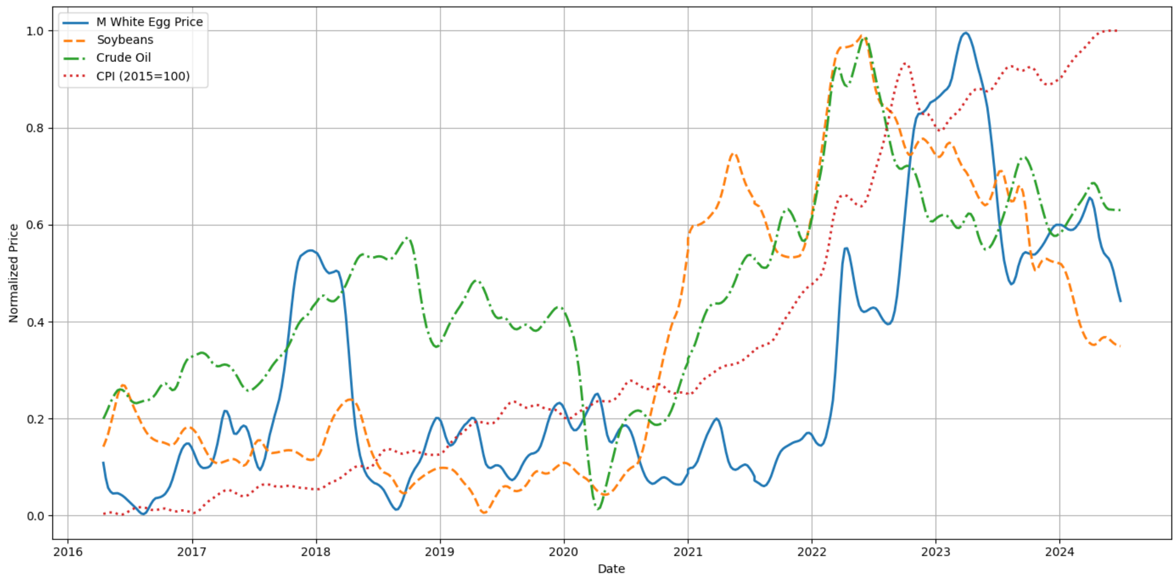


Figure 23: Smoothed Price Trends of Medium White Egg with Soybeans, Crude Oil and CPI

The mathematical representation of SARIMA$(p, d, q)(P, D, Q)_s$ is [8]

$$\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)Z_t, \quad \{Z_t\} \sim \mathrm{WN}(0, \sigma^2) \tag{11}$$

where

$$\begin{aligned}
\phi(z) &= 1 - \phi_1 z - \cdots - \phi_p z^p,\\
\Phi(z) &= 1 - \Phi_1 z - \cdots - \Phi_P z^P,\\
\theta(z) &= 1 + \theta_1 z + \cdots + \theta_q z^q,\\
\Theta(z) &= 1 + \Theta_1 z + \cdots + \Theta_Q z^Q.
\end{aligned}$$

$\phi(z)$ and $\theta(z)$ are non-seasonal components, and $\Phi(z)$ and $\Theta(z)$ are seasonal components with seasonal lags.

The differenced series including both non-seasonal differencing d seasonal differencing D, is given by

$$Y_t = (1 - B)^d (1 - B^s)^D X_t \tag{12}$$

SARIMA model removes seasonality to handle seasonal and non-seasonal data by differencing the series at lag s.

Similarly to ARIMA model, the values of P, D, and Q in SARIMA can be determined by examining the seasonal patterns in the ACF and PACF plots.

Seasonal Autoregressive Integrated Moving Average with Exogenous Regressors (SARIMAX) model is a powerful time series forecasting model that takes into account external variables based on SARIMA model. The general mathematical representation of SARIMAX model extends SARIMA model in Equation 11 by adding the impact of exogenous variables:

$$\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)Z_t + \sum_{k=1}^{m} \beta_k X_{k,t} \tag{13}$$

where the term $\sum_{k=1}^{m} \beta_k X_{k,t}$ represents the linear combination of the exogenous variables and their corresponding coefficients.

So, by adding the exogenous variables, SARIMAX model provides more powerful tools to predict time series data, taking into account both internal patterns and external factors.

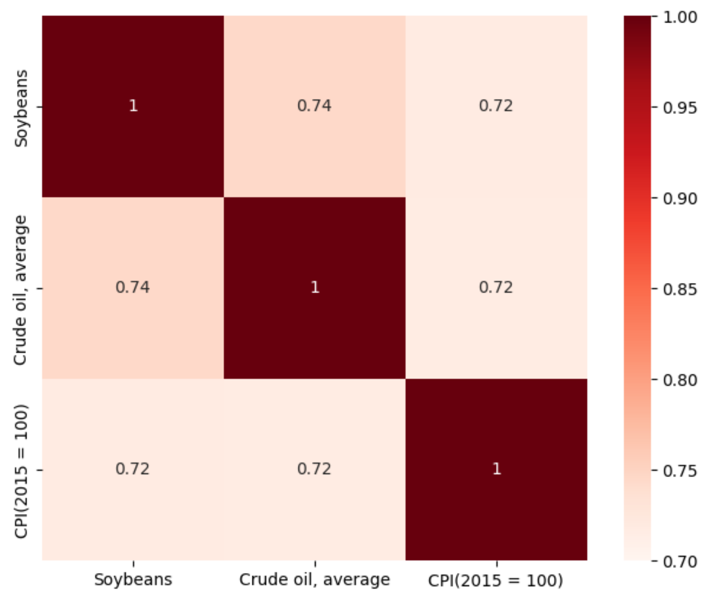


Figure 24: Correlation for Soybeans, Crude Oil and CPI

5.1.3 Prophet

Prophet model is a time series model that is designed to handle common features of business time series such as multiple strong seasonalities, trend changes, outliers, and holiday effects [52]. It decomposes a time series with three main components: trend $g(t)$, seasonality $s(t)$, and holiday $h(t)$ in the following equation:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t \tag{14}$$

where ε_t is the error term assumed to be normally distributed.

The trend component $g(t)$ models non-periodic changes and can be either a linear or non-linear function. Prophet allows for a piecewise linear model as well as a saturating growth model, which distinguishes it from ARIMA, which assumes a purely linear relationship.

For nonlinear growth forecasting that saturates at a carrying capacity, a logistic function can set an upper limit on the trend:

$$g(t) = \frac{C}{1 + \exp(-k(t - m))} \tag{15}$$

where C is the carrying capacity, k is the growth rate, m is an offset parameter. This logistic function introduces non-linearity into the model, allowing for more realistic forecasting of scenarios, an important advantage of Prophet over linear models like ARIMA.

The capacity and growth rate are not constant, so the model needs to fit the data with varying rates. To address it, Prophet model defines change points S at time s_j to allow adjusting growth rates by vector $a(t) \in \{0,1\}^S$, where $a_j(t)$ equals 1 if $t \geq s_j$ and equals 0 otherwise. $\delta_j \sim \text{Laplace}(0, \tau)$ is the rate change that occurs at time s_j. The parameter τ changes the rate δ and controls the flexibility of the model.

For time series with a consistent rate of increase or decrease, the trend component with a piecewise linear model is

$$g(t) = (k + a(t))t + (m + b(t)) \tag{16}$$

where $b(t)$ is the changepoint offset.

Prophet model uses the Fourier series to capture periodic effects.

$$s(t) = \sum_{n=1}^{N} \left(a_n \cos\left(\frac{2\pi n t}{P}\right) + b_n \sin\left(\frac{2\pi n t}{P}\right) \right) \tag{17}$$

where P is the period of seasonality (e.g., $P = 365.25$ for yearly seasonality and $P = 7$ for weekly seasonality), and a_n and b_n are Fourier coefficients. Like ARIMA and SARIMA, grid search or AIC can be applied to select the parameters.

Prophet allows a custom list of holidays and events, which can impact the time series. It assumes that the effects of holidays are independent. The holiday model as a binary variable is

$$h(t) = Z(t)\kappa \tag{18}$$

Here, regressors $Z(t) = [1(t \in D_1), \ldots, 1(t \in D_L)]$, where D_i is the set of past and future dates for holiday i. Each holiday is assigned a parameter $\kappa \sim \text{Normal}(0, \nu^2)$ to correspond to the change.

Prophet uses Stan's L-BFGS to find a maximum a posteriori (MAP) estimation to fit the model, combining the flexibility of Bayesian methods with the computational efficiency of classical optimization techniques [52]. It allows Prophet to provide uncertainty intervals for its forecasts.

Prophet model incorporates several default settings that are designed to handle time series forecasting. The default seasonality components include Yearly Seasonality with a period of 365.25 days and a Fouier order of 10, Weekly Seasonality with a period of 7 days and a Fouier order of 3, and available Daily Seasonality.

Prophet enables fine-tuning through various hyperparameters, such as `changepoint_prior_scale` to manage the trend's flexibility, `seasonality_prior_scale` to regulate the intensity of the seasonal component, and `holidays_prior_scale` to adjust the effects of holidays.

Prophet allows users to incorporate additional external features as regressors into the model to account for factors that might influence the time series and improve the accuracy of the predictions. The effects of extra variables are modeled as an additive component, which is linearly combined with the other default components, such as trend, seasonality, and holiday. Similarly to other components, the model will learn the relationship between the regressor and the target variables to adjust future predictions. Regressors can be continuous or categorical. Continuous regressors are normalized to help the convergence of the model, while categorical regressors are usually one-hot encoded to binary features. The mathematical representation of Prophet model with regressors can be extended from Equation 14 to:

$$y(t) = g(t) + s(t) + h(t) + \beta_1 \cdot r_1(t) + \beta_2 \cdot r_2(t) + \cdots + \beta_k \cdot r_k(t) + \epsilon(t) \tag{19}$$

where $r_1(t), r_2(2), ..., r_k(t)$ are the regressors, and $\beta_1(t), \beta_2(t), ..., \beta_k(t)$ are the coefficients of the regressors.

So, by adding exogenous variables, Prophet model can respond to external influences that might not be captured by the historical time series data alone, leading to more accurate predictions.

5.1.4 LSTM

Traditional Recurrent Neural Networks (RNN) with sigma cells will miss information when the input gap is large [61]. Due to the vanishing gradient problem, RNNs have short memory. Long Short-Term Memory (LSTM) networks are a special type of RNN architecture that is designed to learn long-term dependencies of sequences of data. LSTMs address the issue of RNNs' gradient through a different structure that allows cells to selectively remember or forget important information to save long-term relationships, making it suitable for time series forecasting, such as egg price forecasting.

Each LSTM memory cell contains three gates, which are forged, input, and output gates, to control the flow of information through the cell. The forget gate decides to discard the stored information from the state of the previous cell. The forget gate takes the previous hidden state $\boldsymbol{h}_{t-1}$ and the current input $\boldsymbol{x}_t$, and passes that through a sigmoid function σ to compute a vector $\boldsymbol{f}_t$ with values between 0 and 1, where 0 represents completely forgetting and 1 means completely keeping [28]. The forget vector is shown as:

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_f \cdot [\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_f\right) \tag{20}$$

where $\boldsymbol{W}_f$ represents the weight matrix for the forget gate, and $\boldsymbol{b}_f$ is the bias vetor for the forget gate.

The input gate $\boldsymbol{i}_t$ is a simple sigmoid threshold unit to determine what new information should be updated, and a tanh layer creates a new candidate vector $\tilde{\boldsymbol{C}}_t$ to add to the cell state.

$$\boldsymbol{i}_t = \sigma\left(\boldsymbol{W}_i \cdot [\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_i\right) \tag{21}$$

$$\tilde{\boldsymbol{C}}_t = \tanh\left(\boldsymbol{W}_C \cdot [\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_C\right) \tag{22}$$

By combining the previous cell state, scaled by the forget vector, with the new candidate vector, scaled by the input gate, the current cell state $\boldsymbol{C}_t$ is updated to keep the necessary information.

$$\boldsymbol{C}_t = \boldsymbol{f}_t \odot \boldsymbol{C}_{t-1} + \boldsymbol{i}_t \odot \tilde{\boldsymbol{C}}_t \tag{23}$$

where the operator $\odot$ here is used to represent the element-wise multiplication of two vectors, also known as the Hadamard product.

The output gate $\boldsymbol{o}_t$ decides what information from the cell state is required for the current hidden state $\boldsymbol{h}_t$. A sigmoid function is applied to decide which part of the cell state to be output.

$$\boldsymbol{o}_t = \sigma\left(\boldsymbol{W}_o \cdot [\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_o\right) \tag{24}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \odot \tanh(\boldsymbol{C}_t) \tag{25}$$

In the context of LSTM, some key parameters control the behavior of the training process. The Number of Units (Neurons) determines the number of neurons per layer, where a higher value makes the model more sensitive to complex patterns but might lead to overfitting. The Number of LSTM layers defines how many layers are allowed to learn hierarchical patterns. The dropout rate is used to randomly drop out some neurons in training to avoid overfitting. The default setting for LSTM is 0, which means that there is no dropout during each update cycle. The learning rate determines the learning speed of the model. It affects the convergence of the model by controlling the size of the updates. The batch size decides the number of training samples. The larger batch size makes for more stable updates and reduces training speed. Conversely, a smaller batch size may introduce more noise into the gradient updates. Epochs control the number of complete passes through the entire dataset. The proper number of epochs helps the model learn patterns sufficiently without overfitting.

This thesis implements LSTM model by Keras, a high-level API used for artificial neural networks. Kears simplifies the process of configuring and fine-tuning these hyperparameters. A density layer is added after LSTM layers are processed to the final output. In addition, the model complies with a loss function and an optimizer to guide the learning process. The optimizer is set as Adam in this thesis to adjust the model's weights based on gradients. The model is finally trained by the 'fit method in Keras to include the number of epochs and batch size to minimize the loss function.

Unlike time-series forecasting models, LSTM model is capable of capturing complex, non-linear relationships between the target variables and features without creating linear lag features. LSTM can automatically learn lag effects and non-linear dependencies by processing input sequences step by step. The application of nonlinear activation functions like sigmoid and tanh functions, allows LSTM to learn complex interactions between features and the target variable. In addition, the memory cell in LSTM can remember and use relevant previous information to capture the effect of features on target variables such as egg prices.

Some feature selection methods were used to ensure that the most relevant features were included in LSTM model. Lasso regression, or the least absolute shrinkage and selection operator, which incorporates regularized L1, was used for the features for the first time. Lasso regression is capable of removing the less important features of the model by shrinking their coefficients to zero [36]. The function for Lasso regression is given by:

$$\text{minimize}_\beta \left(\frac{1}{2n} \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right) \tag{26}$$

where n and p are the number of observations and features, respectively. y_i is the actual value of the dependent variable for the ith observation. $x_i j$ is the value of the jth feature for the ith observation. β_j is coefficients for feature j, and λ is the regularization parameter.

The first part of the Lasso regression formula is the loss function to minimize the differences between the observed values and the predicted values. The second part $\lambda \sum_{j=1}^{p} |\beta_j|$ is the L1 regularization term. It adds a penalty for reducing some coefficients to zero, which helps in feature selection.

In addition, a Random Forest algorithm was also applied to find the most important features before training LSTM model. Random Forest is an ensemble learning method that constructs random decision tree classifiers [57]. During training time, some decision trees were built from a bootstrap sample of the training data. At each node of the tree, a random subset of features is selected to find the best split. It is a useful technique for feature selection by ranking features based on their importance in making predictions. This method allows for the evaluation of which features impact the most on the performance of the model, allowing the elimination of redundant features.

However, both Random Forest and Lasso regression have limitations in highly correlated features. Random Forest focuses only on the feature that can minimize the error in each decision tree. A high-collinearity feature will be identified as important if it can reduce the error. Lasso regression might choose one feature and set the others to zero, even if the other features are also relevant. To address this, the Variance Inflation Factor (VIF) was calculated to check the multicollinearity of the features.

5.1.5 Benchmark: Naive and Hyperbolic Models

In this thesis, Naive Forecasting and Hyperbolic Forecasting are used as benchmark models to compare against more advanced forecasting methods. This part will briefly explain the theoretical background of the models and their limitations.

Naive Forecasting Model

Naive model is a simple approach to forecast time-series data. It uses the most recent observation as the prediction for this period, without considering any factors. For a time series y_t Naive forecast for the next period y_{t-1} is [51]:

$$\hat{y}_{t+1} = y_t \tag{27}$$

Naive model assumes that it follows a random walk process, where only some random noise exists between two consecutive values:

$$y_{t+1} = y_t + \epsilon_t \tag{28}$$

Naive model serves as a good benchmark for comparison with more sophisticated forecasting models because it is simple to use. However, it does not consider any trends, seasonality, or cyclical patterns in the time series data, making it less accurate. As it only has the value of the last observation, the noise in the data can not be removed and will be copied in the future.

Hyperbolic Model

Another simple model as a benchmark is hyperbolic forecasting model, which assumes that the rate of change in the time series follows decaying patterns, where the influence of past values diminishes over time. It is an effective model for predicting time series that changes slow down as time increases.

Hyperbolic model can be represented as:

$$y_t = \frac{\alpha}{t+\beta} + \epsilon_t \tag{29}$$

where y_t is the predicted value at time t, α and β are parameters that control the shape of the hyperbola. ϵ_t is the error term.

In this thesis, Hyperbolic model was applied using the most recent three data points to make prediction for future values. It provided a benchmark for short-term forecasting with a more pronounced impact of recent data on future values.

For example, the three most recent data points (t_1, y_1), (t_2, y_2), and (t_3, y_3) were selected to fit the model. The parameters α and β can be estimated by the equations for the three points like $y_1 = \frac{\alpha}{t_1+\beta}$, $y_2 = \frac{\alpha}{t_2+\beta}$, and $y_3 = \frac{\alpha}{t_3+\beta}$, by assuming error terms are zero. So, the predicted value in the time series at time $t+1$ is:

$$y_{t+1} = \frac{\alpha}{(t+1)+\beta} \tag{30}$$

The use of only three recent data points makes the model fit more simply. The model is sensitive to recent trends, which is useful in short-term prediction. However, the limited number of data points cannot capture the underlying trends within the data, leading to the risk of overfitting. It is not a reliable method for long-term predictions. The application of Naive forecasting and Hyperbolic forecasting can provide insights on the performance of other advanced forecasting models.

5.1.6 Evaluation Metrics

In this thesis, the performance of the forecasting models is evaluated using three key metrics, which are Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). The evaluation metrics provide a comprehensive assessment of the model accuracy by calculating different aspects of prediction errors.

Mean Absolute Error (MAE)
MAE measures the natural average magnitude of errors in a set of predictions [59]. It calculates the mean of the absolute differences between the predicted values ($\hat{y}_i$) and the actual values (y_i). It is expressed as

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{31}$$

MAE provides an easy measure of forecast accuracy by directly calculating the mean of residuals. It treats all errors equally, making it less sensitive to outliers. MAE is useful for measuring the magnitude of errors without considering the impact of large deviations. MAE can provide error metrics with the same units as the original data.

Mean Squared Error (MSE) & Root Mean Square Error (RMSE)
MSE measures the average of the squared difference between the predicted and actual values. RMSE is the square root of MSE. Similarly to MAE, RMSE provides error metrics with the same units as the original data. The formulas for MAE and RMSE are given by:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{32}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{33}$$

Compared to MAE, RMSE is suitable for use when model residuals follow a normal distribution [11]. Both MSE and RMSE are sensitive to larger errors because they square the difference of residuals, so they are effective in identifying models with extreme deviations. On the other hand, MSE and RMSE might skew the evaluation if the anomalies are contained. Additionally, MSE and RMSE avoid using absolute values, which is beneficial compared to MAE.

Mean Absolute Percentage Error (MAPE)
MAPE measures the average absolute percentage difference between the predicted values and the actual values. It is given by:

$$\mathrm{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{34}$$

MAPE provides a percentage error, making it easier to compare the models' performance and accuracy across different scales. However, MAPE is sensitive when the actual values are close to zero. For this thesis, the egg prices are always higher than zero, making it suitable for evaluation.

Overall, MAE provides a simple average error, making it less sensitive to outliers, while MSE/RMSE penalizes heavily on larger errors. MAPE normalizes the errors as percentages, allowing for comparison across different scales. In this thesis, the combination of these metrics is well-suited for assessing the forecasting capabilities of ARIMA, SARIMA, Prophet, and LSTM.

5.2 Model Implementation

The model implementation section of this thesis details the practical steps taken to apply ARIMA, SARIMA, Prophet, and LSTM models for egg price predictions. For ARIMA, Section 5.2.1 outlines the diagnostic tests performed to ensure stationarity and model fit, followed by similar diagnostic steps for

SARIMA model in Section 5.2.2, which incorporates seasonality. The implementation of Prophet model in Section 5.2.3 is explored through the decomposition of the time series into its trends, seasonality, and holiday components, with a focus on how these components contribute to the accuracy of the forecast. Each model's implementation process includes hyperparameter tuning and standard modeling steps, such as selecting appropriate parameters to optimize performance.

For LSTM model, which is particularly well-suited to capturing non-linear relationships, Section 5.2.4 begins with adding new interaction terms between variables to explore the potential relationships between features. Then, the feature importance is analyzed and tested, given the high multicollinearity between features. Following it, the hyperparameter tuning process for LSTM is explained, ensuring optimal configuration for layers, dropout rates, learning rate, batch size, and epochs. This section provides a detailed guide to how each model is applied, showing the specific challenges and strengths of different forecasting techniques.

5.2.1 ARIMA

To ensure that the time series data for medium white egg prices were suitable for ARIMA modeling, stationarity was first tested using the augmented Dickey-Fuller (ADF) test. The results of the ADF test indicated that the time series was not stationary, as the p-value was 0.178, larger than 0.06, and the ADF statistic -2.283 was greater than the critical values at the levels of 1%, 5% and 10%, which were -3.446, -2.868 and -2.570, respectively. To address this problem, some transformations were considered, including logarithmic transformation and Box-Cox transformation. Ultimately, differencing was chosen to stabilize the series. After the first differencing, the ADF statistic was -7.805, and the p-value was 7.308×10^{-12}, which indicated that the difference series was stationary. Figure 25 illustrates the comparison of original and differenced data. The differenced egg data appears to be more stationary compared to the original data, with the mean and variance showing a relatively stable over time. However, there are still some notable spikes and possible seasonal effects that have not been completely removed. The spikes in 2017 and 2022 could affect model development.

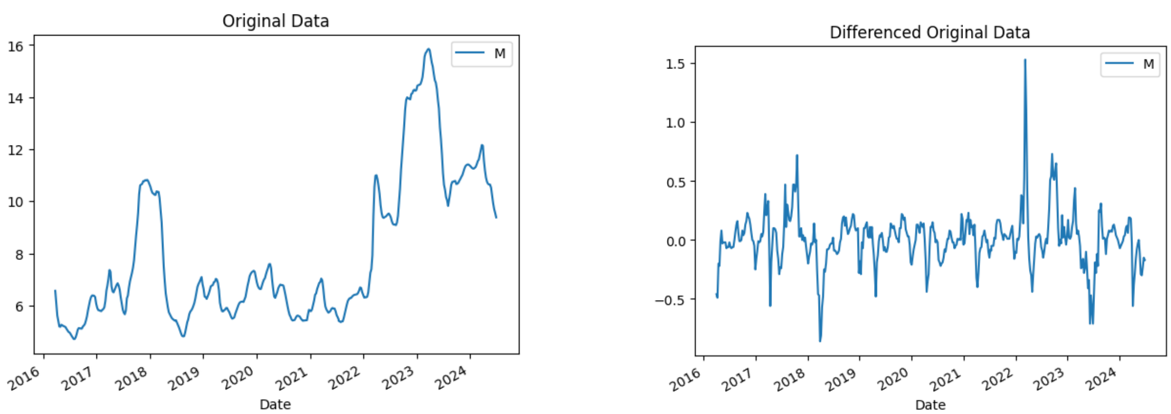


Figure 25: Comparison of Original and Differenced Data

ACF and PACF plots of the differenced series in Figure 26 provide information on more properties of the medium white egg data, which can help determine the order of the AR and MA components. The ACF plot of the differenced data shows significant autocorrelations in the first few lags, indicating that the differenced data still exhibit patterns. There are no clear seasonal spikes, showing that seasonal effects might have been effectively removed. The first lag in the PACF plot is significant, suggesting that an AR(1) model could be suitable. Although there are a few additional significant lags after 6 lags, higher-order AR terms are not necessary because they are not as strong.

This thesis was initially selected $p = 1$ and $q = 0$. Various models were compared using AIC and BIC to choose the one with the lowest criterion value, leading to the final model choice. The parameters of ARIMA model were also estimated by MLE. For the white medium egg price series, ARIMA model configuration was:

$$ARIMA(p, d, q) = ARIMA(1, 1, 1)$$

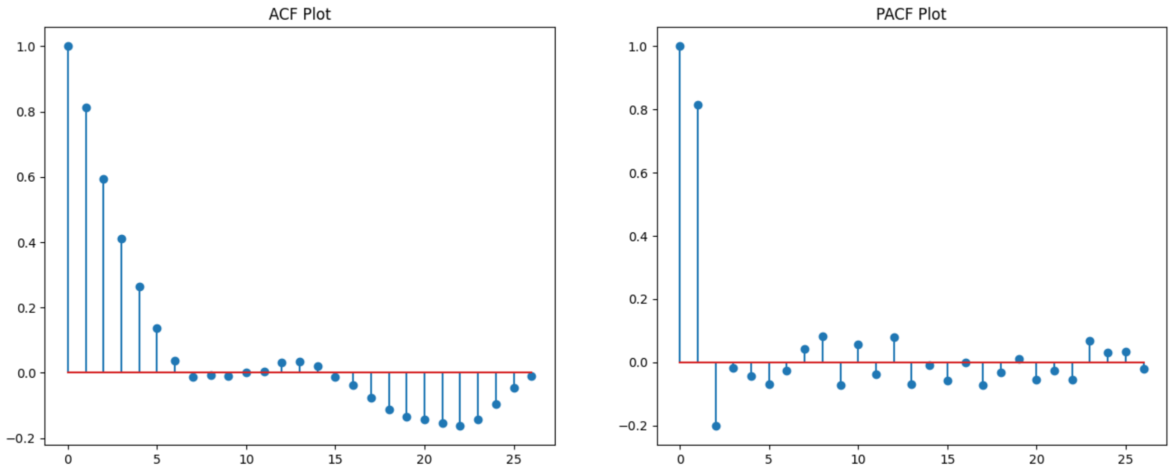


Figure 26: ACF and PACF Plots of Differenced Data

After fitting the ARIMA$(1, 1, 1)$ model, diagnostic checks were performed, visualized in Figure 27. The residuals appear randomly scattered around zero, which shows a good fit. The noticeable spikes indicate volatility in the data. The plot of the historical plus estimated density shows that the distribution of the residuals resembles a normal distribution, but there are some deviations. The Kernel Density Estimation (KDE) curve mostly aligns with the normal distribution curve (N(0,1)), showing that the residuals are approximately normally distributed, though some slight skewness or kurtosis exists. The normal Q-Q plot compared the residual quantiles to the standard normal distribution quantiles. Although normality is indicated by most points around the line, a few points deviate significantly from the line, especially in the upper tail, suggesting potential outliers in the tails.

The plot of the correlogram displays the ACF of the residuals. The most autocorrelations of residuals are located within the confidence bounds, indicating the residuals are not significantly autocorrelated and ARIMA (1,1,1) can capture the underlying patterns in medium-white egg data. However, the first lag shows a slight spike. Additional AR or MA terms were considered to added. Compared to AICs of ARIMA(2,1,1) and ARIMA (1,1,2), ARIMA(2,1,1) has the lowest AIC (-543.696), which is slightly better than ARIMA(1,1,1) (-543.575). ARIMA(1,1,2) has a slightly higher AIC (-542.987), indicating that it is not as good as the other two models. The thesis kept ARIMA (1,1,1) because the difference is minimal and the increase in the model's complexity does not improve the fitting significantly.

The Ljung-Box test was used to check the null hypothesis that the residuals are independently distributed. The p-value is 0.990, which is larger than 0.05 and proves the residuals of ARIMA (1,1,1) model are white noise.

Overall, ARIMA (1,1,1) fits reasonably well with the data of medium-white egg prices based on diagnostics. The residuals appear to be roughly normally distributed and show minimal autocorrelation.

With ARIMA(1,1,1) model fitted, bootstrap simulation was employed to generate predicted prices and confidence intervals by repeatedly resampling the residuals and stimulating new future paths of the time series. The confidence intervals of 25%, 50%, and 95% for the predicted prices were constructed by the bootstrap simulations, providing a way to assess the uncertainty.

5.2.2 SARIMA

In ARIMA model, orders of (1, 1, 1) were selected to consider non-seasonal patterns in egg price data. AR(1) indicates that the prices of eggs are influenced by the price in the last week. MA(1) considers that the prices of eggs are influenced not only by the previous week's prices but also by the error from prices in the last week. Although ARIMA models provided a reasonable fit for egg price data, the residual diagnostics in Figure 27 proved the presence of unaccounted seasonal patterns.

Given the clear evidence of seasonality in the egg price data in Figure 8 and Figure 7, it was necessary to extend ARIMA model to incorporate seasonal components. It led to the development of

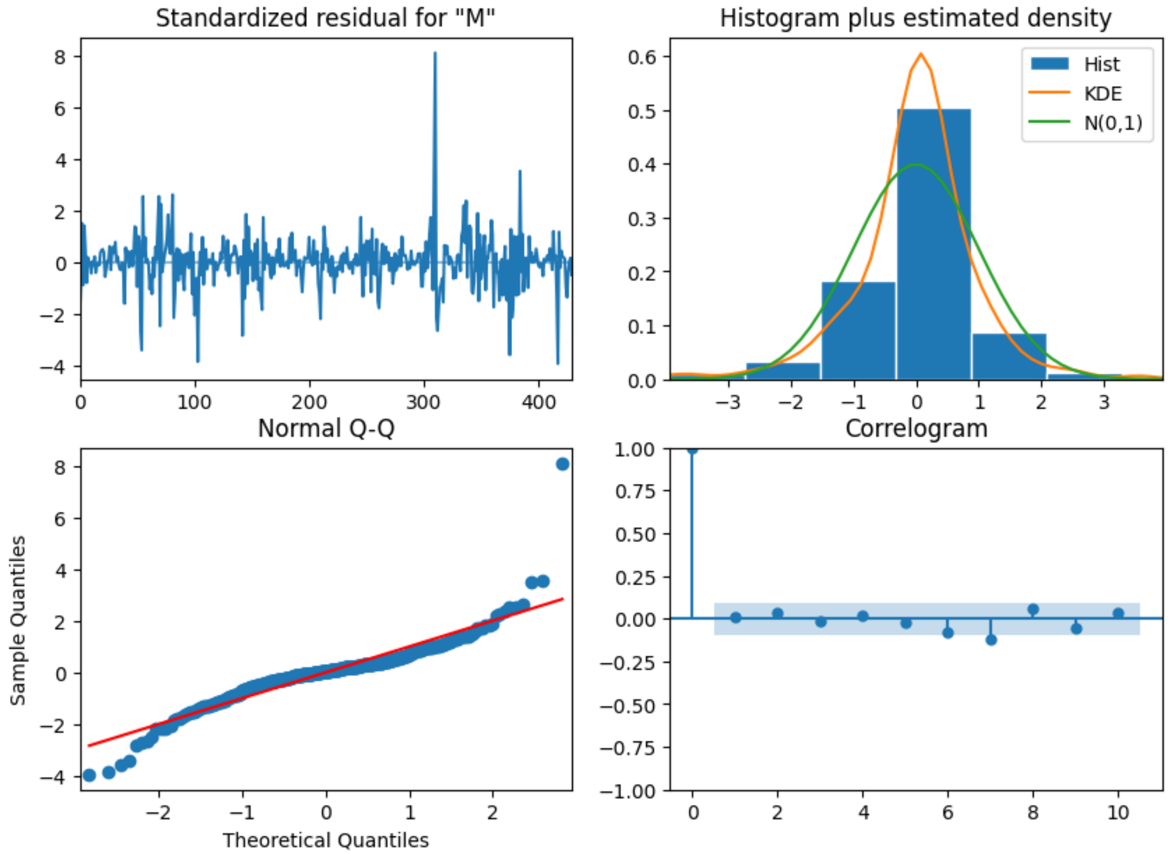


Figure 27: Diagnostic Plots for ARIMA(1,1,1) Model

the Seasonal ARIMA model, with the ability to consider both seasonal and non-seasonal patterns of the data.

To determine the order of SARIMA model, the orders (1, 1, 1) were directly used as the non-seasonal parameters. This decision was based on the ADF test and the ACF / PACF plots, as well as the consideration of model complexity. The grid search was conducted to determine the seasonal orders (P, D, Q, m) of SARIMA by testing different combinations of seasonal parameters. The seasonal differencing order (D) was set to 1 based on the seasonality detected. The seasonal period (m) was fixed at 52 weeks because egg prices are weekly data and show some annual cycles.

The model configuration was evaluated using the Akaike Information Criterion (AIC), with the aim of minimizing the value. The results of the grid search showed that SARIMA (1, 1, 1) (2, 1, 1, 52) model had the lowest AIC value (-413.308) among the tested models, suggesting it was the most suitable configuration.

Seasonal AR (2) means that the model uses the values from the same period in the previous two years to predict the current egg prices. It considers the influence of past seasonal patterns on current value. Seasonal MA (1) indicates that the model considers the seasonal noise, and the current egg prices are influenced by the seasonal lag of forecast errors.

Diagnostic checks for SARIMA models are presented in Figure 28. The standardized residuals for SARIMA model are centered on zero with constant variance. Compared to ARIMA model, the overall pattern is more stable, which means the additional seasonal components stabilize the residual. The spike observed in ARIMA model is also present in it, indicating that outliers in 2022 exist after introducing seasonality.

Similarly to ARIMA model, the histogram and the Kernel Density Estimate (KDE) plot for SARIMA model also indicate approximate normality. Like ARIMA model, the KDE mostly aligns with the N(0,1) curve with small deviations at the tails. The residuals of SARIMA model are more tightly distributed around zero, which indicates that it provides more prediction with high accuracy. But the slight left skew also implies that some predictions have potential underestimation. ARIMA

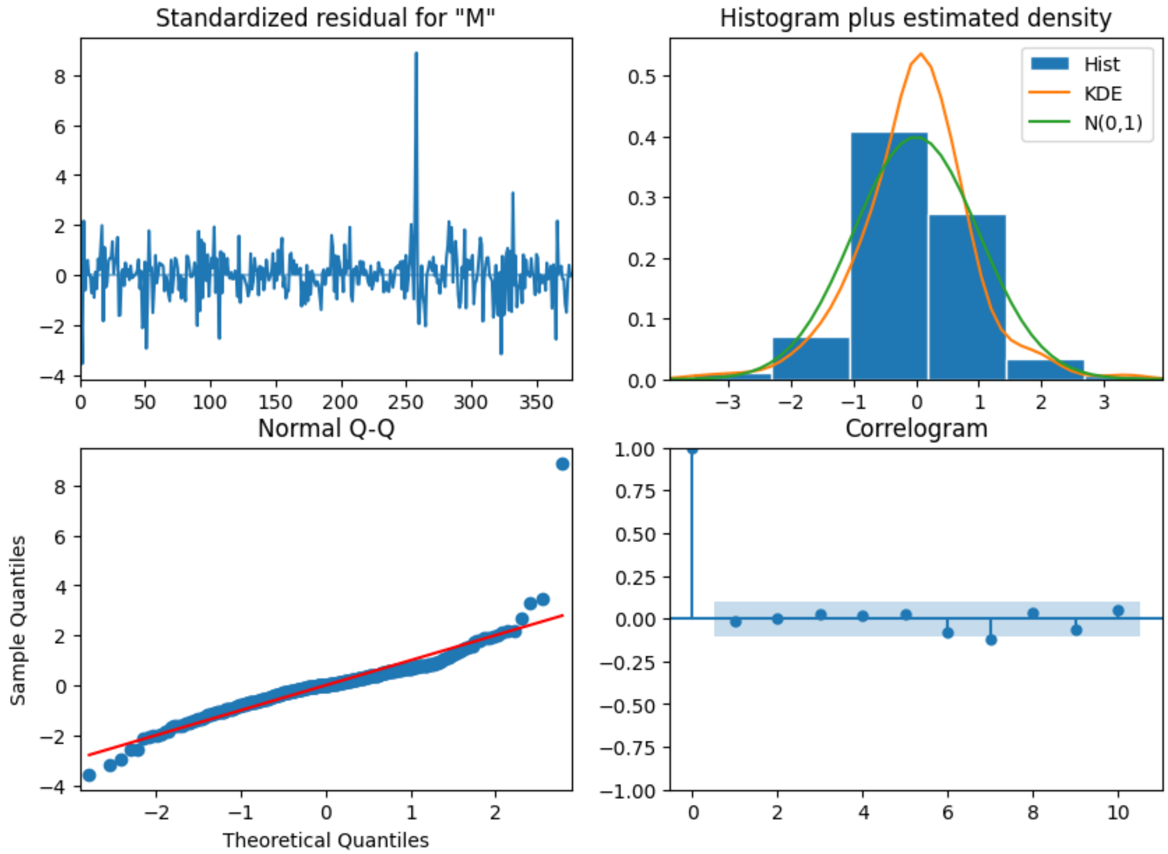


Figure 28: Diagnostic Plots for SARIMA(1,1,1)(2,1,1,52) Model

model shows a more symmetric distribution of residuals but less accurancy. So, the second diagnostic plot shows that although SARIMA model can capture the seasonality of the data, residual skewness needs to be improved. ARIMA model offers a more balanced but low-fitting method.

The Q-Q plot for SARIMA model shows that the residuals generally follow a straight line, as in ARIMA model. Both models have deviations at the tails. Neither model fully solves the extreme values.

The correlogram (ACF of the residuals) for both models shows that most autocorrelations fall within the 95% confidence interval. SARIMA model slightly reduced some minor spikes compared to ARIMA model, indicating that seasonal effects have been captured.

Compared to the AIC and BIC of ARIMA models, SARIMA models have larger AIC (-413.308) and BIC (-389.699), suggesting that the simpler model might provide a better fit to the data. It implies that ARIMA model might be sufficient for capturing the dynamics of the egg prices although it didn't introduce seasonal patterns. The complexity of SARIMA model with more orders caused overfitting. However, ARIMA(1, 1, 1) model has some limitations when used for long-term forecasting. Without seasonal components, ARIMA model may end up repeating the same values after a certain prediction period, failing to consider the cyclical nature of egg prices, which are influenced by factors such as holidays and seasonal costs. In contrast, SARIMA model was still conducted to forecast.

Then, similar with ARIMA and ARIMAX models, SARIMA model was added exogenous variables of feed prices, energy prices and CPI to extend to SARIMAX model to consider more factors that impact on egg prices.

Overall, although ARIMA and SARIMA models have captured the dynamics of medium white egg prices, some more complex models, such as nonlinear models, can be applied to improve prediction accuracy.

5.2.3 Prophet

Prophet, as a piecewise nonlinear model, was considered to predict egg prices because its ability to explicitly incorporate holidays and other special events into the forecasting process. In the context of egg prices, Easter is a particularly important event, as demand for eggs typically increases some weeks before Easter Sunday. By accurately capturing the effect of Easter egg prices, Prophet model can provide more accurate forecasts, particularly around the Easter period.

The dates of the Easter holidays change every year, which makes it hard to identify by models like ARIMA and SARIMA. Prophet model provides custom settings for holiday dates. By setting '2017-04-16', '2018-04-01', '2019-04-21', '2020-04-12', '2021-04-04', '2022-04-17', '2023-04-09' and '2024-03-31' as exact dates of Easter holidays from 2017 to 2024, the parameters for a window of days around Easter were included to identify effects of holidays. As in the price analysis around Easter shown in Figure 9 of Section 3.3, egg prices typically rise two months before Easter Sunday and end one or two weeks before Easter Sunday.

The `'lower_window'` parameter was set to -45 after testing the windows $0, -7, -14, -45$ and -60, reflecting that the model incorporated 45 days leading up to Easter as a holiday effect. The `'upper_window'` was set to 0 because the model requires a nonnegative parameter here. By setting these two parameters, Prophet model captures the entire period when egg prices are influenced by the Easter holidays.

To optimize the performance of the model, some key parameters were adjusted on the basis of the results of the grid search. Changepoint Prior Scale was increased from the default value of 0.05 to 0.5 to allow the model to be more sensitive to detect small changes in the trend. Seasonality Prior Scale was adjusted to 1.0 to enable the model to capture more complex seasonal patterns. A Holiday Prior Scale of 1.0 was added to give moderate weight to the effect of the Easter holiday, ensuring the model captures the influence without overfitting.

The yearly seasonality of the model was added to the model. As historical data were weekly, the model does not have the granularity needed to detect patterns that occur within a single month, such as the change between the beginning, middle, and end of the month. The unnecessary seasonal components could lead to overfitting by fitting the noise instead of the actual underlying patterns. So, the monthly, weekly, and daily seasonalities of Prophet model were disabled in the thesis.

Prophet model provides the default output graph shown in Figure 29 to visualize the summary of the model's forecast against the actual data, with 80% uncertainty intervals. The black dots represent the actual price of eggs. The blue line is the fitted predicted line generated by Prophet model. Prophet model appeared to fit the historical egg price data well, capturing the underlying trends and seasonal variations. In peak 2018 and 2022, they were also well modeled, showing that the model can identify and estimate unusual patterns. The higher Changepoint Prior Scale allows the model to detect more small changes in the trends. Some deviations from the fitted values were observed from 2020 to 2021. It implies that Prophet model responded to the small change in prices too quickly than the actual changes.

The shaded blue area represents the uncertainty intervals that provide the area where true prices are expected to fall. These intervals are plotted under the assumption that the residuals of the model are normal distributions. The assumption can not always hold true in reality, so the residuals from Prophet model were plotted in Figure 30. Overall, the differences between the actual prices and the predicted prices fluctuate around zero. However, the residuals were more volatile in some periods, particularly from 2020 to 2021, and from 2022 to mid-2023, showing that the model had difficulty capturing the true variation in egg prices during these periods. So, the confidence intervals here cannot fully reflect the true uncertainty in the model's predictions.

Figure 31 shows the different components of Prophet forecast. The trend component shows the general trends in the prices of medium-white eggs over the period. It captures the general direction, excluding seasonal effects or holidays. The holiday component represents the impact of the Easter holidays on egg prices. The sharp spikes indicate a strong positive effect. Additionally, a component of yearly seasonality shows a clear pattern where egg prices tend to peak around March and gradually decrease until the end of August; then prices begin to increase again towards the end of the year. The components of Prophet model are consistent with the seasonality analysis in Section 3.3.1.

Similarly to ARIMAX and SARIMAX models, Prophet added feed costs, energy costs, and CPI as three regressors to the models, ensuring that the necessary factors for the forecast were considered.

Overall, Prophet models offered the ability to incorporate Easter holidays as important events by

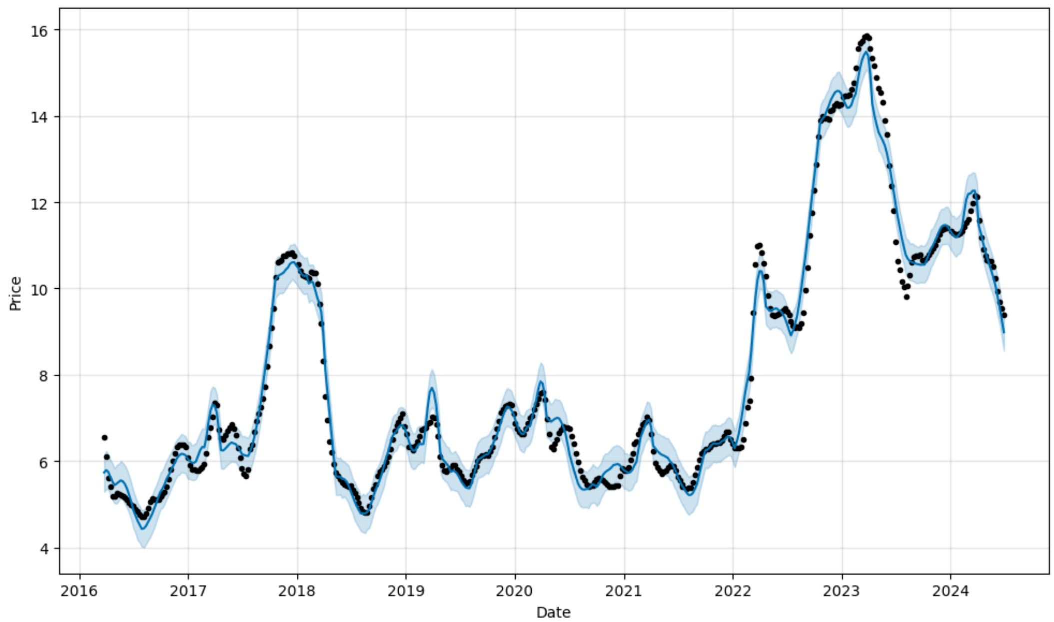


Figure 29: Prophet Model Forecast for Medium White Egg Prices with Uncertainty

introducing exact dates, which allowed the model to consider sudden increasing demand during the impact period of events. Also, Prophet constructed yearly seasonality components that effectively capture complex annual patterns in the data, which might be overlooked by ARIMA and SARIMA models. Prophet model is a powerful tool to predict egg prices.

5.2.4 LSTM

Time-series models such as ARIMA and SARIMA can handle linear relationships and seasonality well. However, some underlying nonlinear patterns were hard to identify by time series models. In the last phase of this thesis, the Long-Short-Term Memory (LSTM) model was employed because of its ability to capture complex, non-linear relationships between features and the target variable. Moreover, none of these traditional time series models is designed to capture the interactions between different exogenous features. They treat each exogenous variable separately and their contributions to the prediction are additive. In contrast, LSTM models can process the non-linear interactions between different features, which made LSTM models tested in the thesis consider more complex market dynamics.

The lag correlation of different features has been analyzed in Section 4 and applied in ARIMA, SARIMA, and Prophet models. However, some subtle patterns might be missed by linear correlation. As shown in Figure 15, the prices of feed and eggs show a clear correlation. For features like energy costs and CPI, the lag correlations were not as strong as with feed costs. Poultry farms spend more money on gas during winter. Energy costs are considered a more significant part of egg production in winter because they are used not only to warm the laying hen but also to be the most expensive part of feed production [6]. A '*Winter*' feature was defined to include the months of December, January, and February. This feature was combined with the Natural Gas in EU feature to create a new interaction term, named '*Winter_NaturalGasEurope*' to capture the impact of gas prices during the colder months. Additionally, given the relationship between crude oil and soybean prices was identified as highly related in [32], a new feature '*CrudeOil_Soybean*' was created to reflect the potential compounded effect of these two variables on the overall production costs. Moreover, the interaction between production costs and inflation was investigated in this thesis. Natural Gas in the EU was selected as it represents the local energy market and is related to the local inflation index CPI. The new feature '*NaturalGasEurope_CPI*' was designed to account for the combined impact of energy prices and inflation on egg prices.

To fully consider linear and nonlinear relationships, all 17 relevant features, including 7 lagged features and 3 interaction features, were included in LSTM model. The multistep feature selection process was employed, combining both Lasso regression and Random Forecast, followed by a multicollinearity

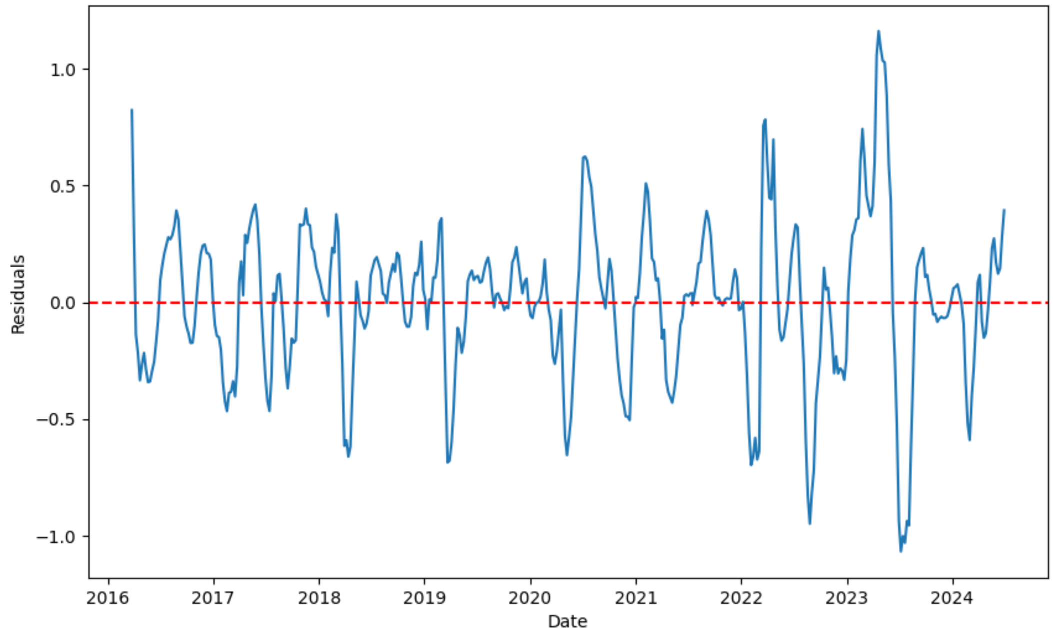


Figure 30: Redisuals of Prophet model

check by Variance Inflation Factor (VIF).

Lasso regression as Equation 26 was first used with a regularization parameter of 0.01 to control the strength of the penalty applied to the coefficients. Only the most relevant features were retained after shrinking the coefficients of the less important features to zero. After utilizing Lasso, the features with non-zero coefficients were kept, while features with fewer contributions were removed, which were Crude Oil, Natural Gas in the US, CPI, interaction between Natural Gas in the EU and CPI, and interaction between Crude Oil and Soybeans.

Feature	**Importance**	**VIF**
CPI_Lag_1	0.7093	265.4692
Soybeans_Lag_40	0.0721	288.2629
CrudeOil_Lag_39	0.0391	134.2594
Wheat_Lag_34	0.0382	118.8494
Natural_gas_US_Lag_35	0.0321	38.9187
Maize	0.0269	358.6818
Natural gas, Europe	0.0265	6.0075
Natural_gas_Europe_Lag_36	0.0200	16.6523
Winter_NaturalGasEurope	0.0145	1.9674
Wheat, US HRW	0.0092	130.4735
Maize_Lag_39	0.0067	281.9493
Soybeans	0.0056	628.4823

Table 2: Rank of Feature Importance by Random Forest with VIF values

With the remaining features, Random Forest was applied to further refine and assess the importance of the features. Table 2 shows the rank of importance of features obtained from the Random Forecast. '*CPI_Lag_1*, '*Soybeans_Lag_40*' and '*CrudeOil_Lag_39*' are the top three important features. Despite their high importance scores, Random Forest is not equipped to handle multicollinearity effectively. When these features were incorporated into LSTM model, the resulting RMSE values were consistently above 6, indicating poor predictive performance. This raised concerns about the effectiveness of the selected features due to potential multicollinearity issues.

To address this, the VIF was calculated for the features. The VIF results in Figure 2 revealed that the top three characteristics had extremely high multicollinearity. The noise was introduced into

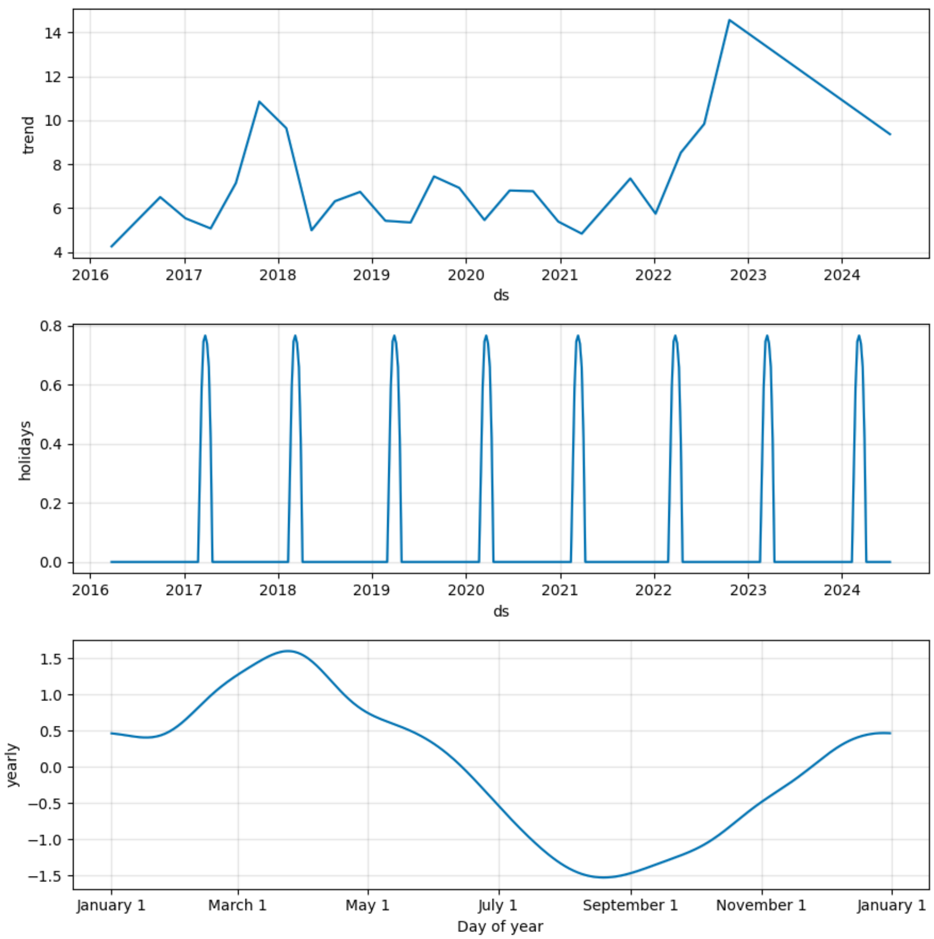


Figure 31: Components of Prophet Forecast

the model, which reduced the model's predictive performance. Only two features, Natural Gas in the EU and Interaction between Winter and Natural Gas in the EU, had VIF values below 10, showing low multicollinearity. However, their ranks in terms of importance in the Random Analysis were not high. Moreover, considering that feed costs are a major component of poultry production costs, '*Soybeans_Lag_40*' was selected as the third feature for LSTM model, despite its high VIF of 288.2629. Although '*CPI_Lag_1*' was the highest rank in the Random Forecast Analysis, it was not considered because the feature of Natural Gas in Europe can partly explain the inflation rate and soybean prices play a more critical role in egg production. So, three features, Natural Gas in the EU, the interaction between Winter and Natural Gas in Europe, and the Lag of Soybeans were added to LSTM model.

After selecting the final set of features, the hyperparameters of LSTM model were fine-tuned to obtain the optimal performance. The entire dataset was split into 80% for training and 20% for validation. Before the training, the data was standardized to ensure that each feature had a mean of zero and a standard deviation of one to improve the model's convergence, as the features had different scales and LSTM networks were sensitive to them. Then, the standardized data was reshaped into a 3D format with dimensions corresponding to '[samples, time steps, features]', as required by LSTM model. It helped the model to capture the dependencies effectively.

The parameter '`Unit`' of 100 was set to the number of memory cells with each LSTM layer. The higher value enables the model to learn more complex patterns, but it also increases the risk of overfitting. '`Dropout Rate`' of 0.2 was chosen to balance the model robustness and model accuracy. 20% of the neurons were randomly dropped during each update cycle in the training period, avoiding the model that relied heavily on specific neurons. In addition, the '`learning rate`' of 0.001 was selected after experimenting with different configurations. It controls the learning speed and ensures the convergence of the model is stable. Moreover, '`batch size`' was set to 16 to define the number of samples used in one iteration of training. It introduced some noise in the gradient updates, which avoided the model with local minima. Then, ‘`epochs`’ of 100 was found to denote the number of complete passes through the entire dataset to learn the underlying patterns. In addition, the penalty parameter ‘`L2 regularization` of 0.0001 was added to avoid overfitting. It can prevent the model from learning large weights because the penalty is proportional to the sum of squared weights. A small L2 regularization parameter balanced the generalization and complexity of the model.

6 Evaluation

The evaluation section of this thesis is designed to assess the performance of various forecasting models, namely ARIMA, SARIMA, Prophet, and LSTM, by examining their effectiveness under different criteria. This chapter is divided into three key areas: Feature Evaluation, Evaluation by prediction length and Evaluation with External Factors.

In Section 6.1, all selected lagged features were incorporated into the ARIMA, SARIMA, and Prophet models and systematically removed one by one and observed the impact on the Root Mean Square Error (RMSE). This step identified the most influential predictors that contribute to accurate forecasting. The features of LSTM model were analyzed and selected in Section 5.2.4 as the model's ability to capture complex and non-linear relationships between features.

Section 6.2 explains how the models perform on different forecasting horizons. By applying an expanding window approach, the models were evaluated in short-term (1 week), medium-term (4 weeks), and long-term (12 weeks) predictions. The performance of these models was benchmarked against naive and hyperbolic forecasting models, using metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE) to assess their precision.

Finally, in Section 6.3, the model's ability to account for significant events like Easter, the Fipronil crisis, and the Russian-Ukraine war. Specifically, this thesis tested how well the models predicted egg prices during these special periods, as well as, whether excluding data during the Fipronil crisis from the training data improved the models' performance. The purpose of this section is to test the robustness of the model in special events and market shocks.

6.1 Feature Evaluation

As shown in Section 4, the lag correlations of features Soybeans, Maize, Wheats, Crude Oil, Natural Gas in the US, Natural Gas in the EU, and CPI were selected as external factors. In this part, the performance of ARIMA, SARIMA, and Prophet models was evaluated by removing one feature at a time to observe its impact on the Root Mean Square Error (RMSE) of the model. The dataset was divided into a training set and a testing set, with 80% of the data for training and the rest for testing. Table 3 summarizes the impact of excluding each feature on the predictive performance of the model.

Feature Removed	**ARIMA RMSE**	**SARIMA RMSE**	**Prophet RMSE**
All Features Included	2.83	2.11	2.83
Without Soybeans_Lag_40	2.85	1.91	1.97
Without Maize_Lag_39	3.08	2.65	2.61
Without Wheat_Lag_34	3.33	3.30	3.12
Without CrudeOil_Lag_39	2.73	2.20	2.69
Without Natural_gas_US_Lag_35	2.65	2.44	3.24
Without Natural_gas_Europe_Lag_36	2.89	2.99	3.08
Without CPI_Lag_1	2.61	2.59	2.42

Table 3: RMSE values for ARIMA, SARIMA, and Prophet models with various feature exclusions

All selected features were initially tested in ARIMAX model, yielding a baseline RMSE of 2.83. More tests were conducted by individually removing each feature to assess its impact on model performance. The removal of wheat lag resulted in a significant increase in RMSE to 3.33, showing that wheat prices contributed to the model's performance. In contrast, the removal of Crude Oil, Natural Gas in the US and CPI led to lower RMSE, with values of 2.73, 2.65, and 2.61, respectively. It shows that these features introduced noise into the model and reduced ARIMAX model's predictive power. The removal of soybean prices increased RMSE slightly to 2.85, while Natural Gas contributed more to the model with 2.89. Therefore, '*Wheat_Lag_34*', '*Maize_Lag_39*', and '*Natural_gas_Europe_Lag_36*' were finally retained in ARIMA model and the rest features were excluded to optimize the performance of the model.

Compared to ARIMA, SARIMA model shows a different sensitivity pattern. SARIMAX model initially reported an RMSE of 2.11 on the test set with all features included. Notably, the exclusion of Soybeans with a 40-week lag resulted in a reduction of RMSE to 1.91, indicating that soybean may be not a beneficial feature for the model. Conversely, the removal of Wheat and Natural Gas in Europe led to notable increases in RMSEs, which were 3.30 and 2.99 respectively. As the exclusion of Crude Oil only made a slight increase in RMSE to 2.20, it was considered to be removed from the model to avoid overfitting. So, lag features of Maize, Wheat, Natural Gas in the US and EU, and CPI were retained in SARIMA models. Soybeans and Crude Oil were removed.

For Prophet model, the baseline of all included features was an RMSE of 2.83. The exclusions of Soybeans, Maize, Crude Oil, and CPI resulted in improvements in models, with a decrease in RMSE. '*Wheat_Lag_34*', '*Natural_gas_US_Lag_35*', and '*Natural_gas_Europe_Lag_36*' were retained due to their significant contribution to the model's accuracy.

For '*Soybeans_Lag_40*', although it shows a high correlation in Figure 13, as a lag correlation feature, it only slightly increased accuracy in ARIMA model. The '*Maize_Lag_39*' contributed to ARIMA and SARIMA models. The '*Wheat_Lag_34*' was beneficial for all three models, showing a strong dependence on this feature, which is consistent with Figure 13.

For energy costs, Crude Oil was removed from all three models because it plays a less or negative role in the model's performance. '*Natural_gas_US_Lag_35*' only contributed negatively to ARIMA model and was retained in SARIMA and Prophet models. '*Natural_gas_Europe_Lag_36*' had a strong influence on three models, which indicates that egg prices are more impacted by the local energy market. CPI was removed from ARIMA and Prophet models, but it was retained in SARIMA models.

In general, features were selected as exogenous variables in ARIMA, SARIMA, and Prophet models based on the above analysis. The feature selection in the final model was slightly different from the analysis in Section 4.2. The explanation is that the time-based split of the data set allows the last 20% of the data to evaluate the features. The last 20% of the data were located from 1 March 2023 to the end of the data set, which was the period after the extreme market in 2022. As analyzed in Section 4.2, the markets of wheat and natural gas in Europe were more volatile and sensitive due to the war. It caused some features like Wheat and Natural Gas in Europe to be evaluated with a greater contribution of the models.

6.2 Evaluation by Prediction Length

The accuracy, reliability, and uncertainty of a forecasting model can vary depending on prediction horizons. The expanding window approach was applied in this part to evaluate the models' performance over the short-term, medium-term, and long-term periods. An expanding window involves incrementally increasing the training dataset's size as the model makes predictions. Compared to the rolling window approach, the expanding window employs all available data to forecast [23]. This method allows the model to learn from the increasing amount of historical data over time, which is close to real business applications with new prices predicted based on updated data points.

Table 4 shows the results of the model performance metrics for three expanding windows. In this thesis, the 1-week window, the 4-week window, and the 12-window were defined as short, medium, and long horizons, respectively. ARIMA, SARIMA, Prophet and LSTM, together with two benchmark models, Naive and Hyperbolic, were evaluated and compared using multiple performance metrics (MAE, MSE, RMSE, MAPE), where 80% of the data was selected as the train set and 20% of the data was selected as the test set.

6.2.1 Short-Term Prediction

Figure 32 presents a comparative analysis using an expanding window approach with a prediction horizon of 1 week. ARIMA model exhibited the best performance across all metrics for a 1-week prediction horizon. The low MAE of 0.0936 indicated a minimal difference between the predicted values and actual values on average. The low values of MSE (0.0198) and RMSE (0.1407) provided that ARIMA can effectively minimize both small and large prediction errors. A MAPE of 0.80% means that the forecasted values are close to the actual values in percentage.

SARIMA model performs nearly as well as ARIMA but with slightly higher error values. The slight degradation in performance could be the model's attempt to capture seasonal patterns, which might not be as important in short-term (1 week) forecasts. In detail, SARIMA's MAE (0.1031) and MAPE (0.88%) are only slightly higher than ARIMA's, suggesting that, on average, SARIMA's predictions are highly accurate, though slightly less precise than ARIMA's. Both MSE (0.0205) and RMSE (0.1431) are also slightly higher than ARIMA's, which means SARIMA might introduce minor additional errors

Window Size	Model	MAE	MSE	RMSE	MAPE (%)
1 Week	ARIMA	0.0936	0.0198	0.1407	0.80
	SARIMA	0.1031	0.0205	0.1431	0.88
	Prophet	0.6635	0.7541	0.8684	5.55
	LSTM	0.2245	0.1062	0.3258	1.88
	Naive	0.1741	0.0574	0.2395	1.49
	Hyperbolic	0.3489	0.2241	0.4734	2.99
4 Weeks	ARIMA	0.5594	0.5702	0.7551	4.79
	SARIMA	0.5594	0.5624	0.7499	4.80
	Prophet	0.9570	1.7608	1.3269	8.16
	LSTM	0.4730	0.3588	0.5990	3.98
	Naive	0.6516	0.7537	0.8681	5.60
	Hyperbolic	0.7896	1.1154	1.0561	6.80
12 Weeks	ARIMA	1.3906	3.8380	1.9591	12.27
	SARIMA	1.4114	3.7863	1.9459	12.54
	Prophet	1.4678	4.3733	2.0913	13.23
	LSTM	0.9442	1.4817	1.2172	8.02
	Naive	1.4360	3.9728	1.9932	12.79
	Hyperbolic	1.5032	4.3526	2.0863	13.44

Table 4: Model Performance Metrics for Different Expanding Windows

However, these values are still low, showing that SARIMA remains effective in minimizing both small and large errors.

Prophet model significantly underperforms compared to other models, including two benchmark models. Prophet is less suited for short-term forecasts. MAE of 0.6635 and MAPE of 5.55% indicate that, on average, Prophet's predictions deviate from the actual values. The MSE (0.7541) and RMSE (0.8684) values highlight that it is difficult for Prophet model to minimize larger errors. Prophet is likely more suitable for capturing seasonal and long-term trends, which explains its underperformance in this short-term prediction.

LSTM model performs moderately well compared to Prophet, though ARIMA and SARIMA still outperform it. The average error of LSTM (0.2245) is higher than ARIMA and SARIMA. MSE of 0.1062 and RMSE of 0.3258 suggest that LSTM's predictions exhibit larger deviations. MAPE of 1.88% is higher than ARIMA and SARIMA but still acceptable. LSTM's strength lies in its ability to handle non-linear relationships and long-term dependencies, which might explain its higher error rates in short-term prediction.

Two benchmark models, Naive model and Hyperbolic model, exhibit good results compared to Prophet model. Especially, Naive model performs surprisingly well compared to LSTM model, which suggests that a simple model can be as effective as the more sophisticated approach for short-term forecasts.

Overall, ARIMA is the most accurate model for short-term egg price predictions, closely followed by SARIMA. Naive model offers a competitive benchmark, while LSTM and Prophet would introduce more noise and need more further tuning.

6.2.2 Medium-Term Prediction

Figure 33 shows the performance of the models over a 4-week forecast horizon. Compared to the 1-week prediction, a longer horizon leads to higher errors. For ARIMA model, the higher MAE (0.5594) indicates that the predictions deviate from actual values by about 0.56 units on average. The MSE (0.5702) and RMSE (0.7551) values suggest an increase in larger errors compared to the 1-week prediction. MAPE (4.79%) shows the acceptable average percentage error for the medium-term forecast.

SARIMA shows similar results to ARIMA for the 4-week forecast chorizo, with almost identical error values across all metrics. MAE of 0.5594 and MAPE of 4.80% suggest that SARIMA's seasonal component does not provide additional accuracy. MSE of 0.5924 and RMSE of 0.7499 indicate that

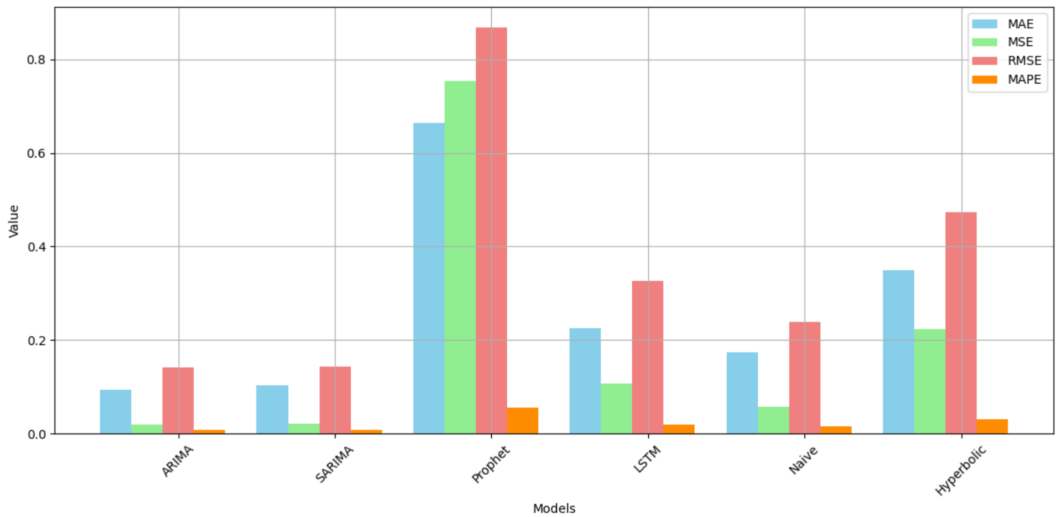


Figure 32: Comparative Analysis of Model Performance Metrics for Short-Term Forecasting

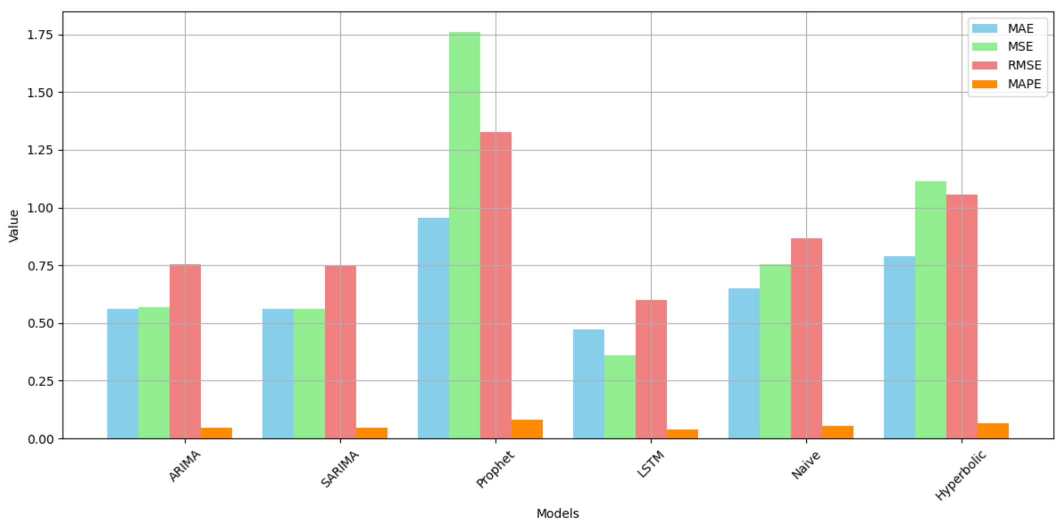


Figure 33: Comparative Analysis of Model Performance Metrics for Medium-Term Forecasting

SARIMA performs marginally better than ARIMA in terms of minimizing large errors, although the difference is small.

For Prophet model, its average error is 0.9570, significantly higher than other models, indicating its less accuracy over the medium-term horizon. Higher values of MSE and RMSE show that Prophet makes larger errors more frequently, with RMSE suggesting an error of about 1.33 units. In addition, the MAPE value above 8% indicates that Prophet is less effective for medium-term forecasting.

LSTM achieves the lowest MAE (0.4730) among all models. The values of MSE (0.3588) and RMSE (0.5990) are the lowest, indicating that LSTM minimizes both small and large derivates more effectively than the other models. The MAPE of 3.98% demonstrates the best excellent percentage-based accuracy. LSTM's ability to capture non-linear relationships seems to give it an advantage in medium-term forecasting.

The performance of Naive and Hyperbolic models diminishes over the 4-week horizon. Hyperbolic model underperformed Naive model. Their inability to capture patterns leads to higher errors, suggesting that they are not suitable for medium-term forecasting of egg prices.

In general, for medium-term forecasts, LSTM model achieves the lowest errors in all metrics. Traditional models like ARIMA and SARIMA, while reliable in the short term, may need more improvement to match the accuracy of advanced machine learning approaches in the medium term. Despite being designed to handle seasonality, holiday and trends, Prophet performs poorly in medium-term forecasts compared to other models, including benchmark models. Its higher error metrics suggest that Prophet may not effectively capture the dynamic of egg prices, as the evaluation periods include the second peak during the Russia-Ukraine war.

6.2.3 Long-Term Prediction

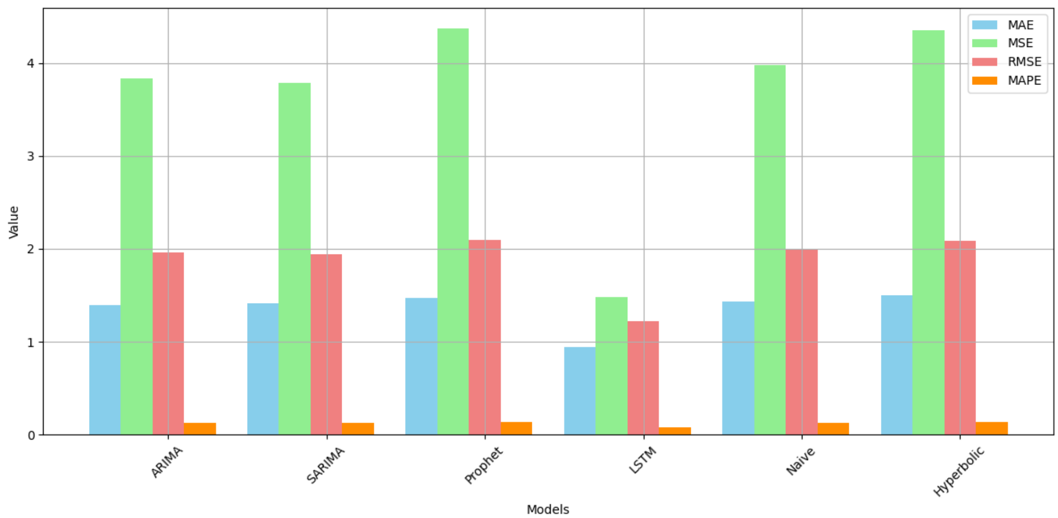


Figure 34: Comparative Analysis of Model Performance Metrics for Long-Term Forecasting

Figure 34 illustrates the model performance metrics for long-term forecasts. Compared to short-term and medium-term forecasts, higher errors were included in longer horizons.

ARIMA and SARIMA perform similarly in long-term forecasts, both showing moderate levels of accuracy. ARIMA's MAE (1.3906) and SARIMA's MAE (1.4114) indicate comparable performance, but both models show limitations in minimizing larger errors, as indicated by their RMSE values (1.9591 and 1.9459, respectively). While SARIMA incorporates seasonality, the benefit in long-term forecasts remains minimal compared to ARIMA.

Prophet performs poorly in the 12-week time frame, with the highest MAE (1.4678), MSE (4.3733), and RMSE (2.0913 among the traditional models. It suggests that Prophet struggles to capture the long-term patterns in egg prices, likely due to its its focus on trend and seasonality without capturing more complex interaction between features. However, Prophet performs relatively well compared to short- and medium-term forecasts, showing its ability to predict the long term.

LSTM significantly outperforms all other models in long-term predictions. With the lowest MAE (0.9442), MSE (1.4817), RMSE (1.2172), and MAPE (8.02%), LSTM shows that its ability on long-term dependencies gives it a distinct advantages for 12-week forecasts.

Naive and Hyperbolic models, used as benchmarks, fall short in long-term forecasting as expected. Naive model performs slightly better than Hyperbolic model but both exhibit high error metrics across all evaluations. Their inability to model complex patterns and interactions results in weaker long-term predictive power.

In conclusion, LSTM remains the model that performs best in the medium- and long-term forecasts. Tradition time series models like ARIMA and SARIMA, while suitable for shorter periods, show limitations in their long-term predictive power. Prophet, Naive, and Hyperbolic models also underperform in 4-week and 12-week horizons, making them less suitable for extended forecasts of egg prices.

6.3 External Factors

The above section summarized the performance of models in different expanding windows. This section evaluates the performance of models during key external events, including the Easter holiday in Section 6.3.1 and the Russia-Ukraine war in Section 6.3.2, using a 4-week expanding window. The

4-week window was chosen for its practical relevance to stakeholders who need forecasts one month ahead for decision-making in the egg market.

Additionally, Section 6.3.3 assesses the impact of the Fipronil crisis, which caused significant local market fluctuations. Since features like feed costs, energy prices, and CPI were less correlated with egg prices during this period, the model performance was evaluated with and without the Fipronil data.

6.3.1 Easter Holiday

The Easter holiday is a significant event that is known to impact egg prices due to increased demand. As analyzed in Section 3.3.1, egg prices usually start to increase about 2 months until the week before Easter Sunday. Also, during the implementation of the model in Section 5.2.3, Prophet model defined 45 days as the impact of the Easter holiday. This section evaluates the performance of the forecasting models, ARIMA, SARIMA, Prophet, and LSTM, as well as two benchmarks, in relation to the Easter holiday.

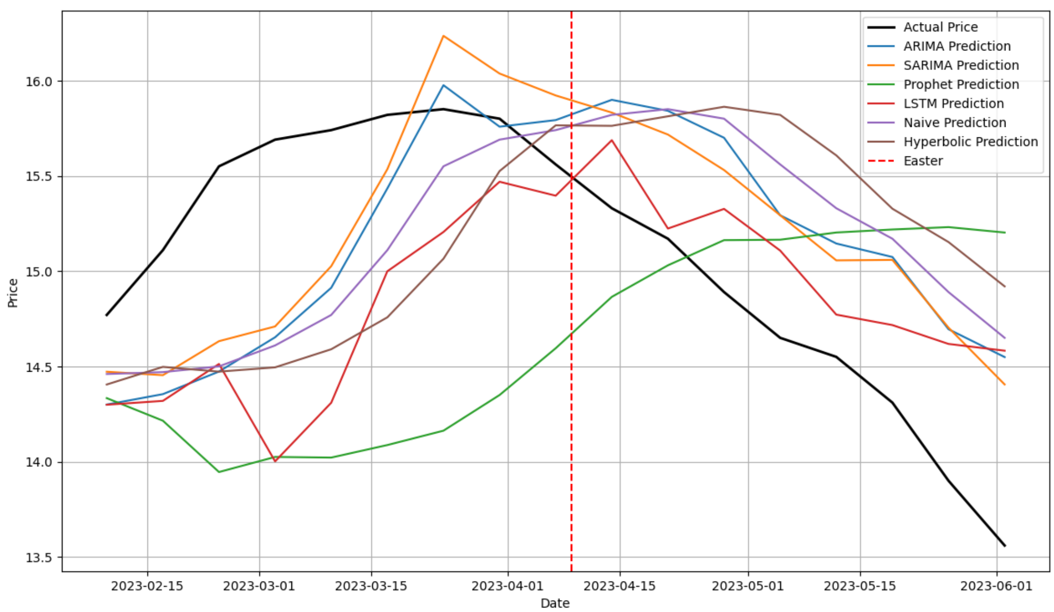


Figure 35: Model Predictions and Actual Prices Around Easter in 2023

Figure 35 compares the model predictions with the actual prices in a 60-day window that surrounds the Easter holiday in 2023. Easter Sunday in 2023 was on 9 April. The prices of medium-white eggs increased significantly on 24 February and remained steady until 31 March, one week before Easter Sunday. In the 2023 predictions, all models show a rise in prices leading up to Easter, which is consistent with actual prices. However, none of the models are close to the actual price. The possible reason could be the models' inability to predict market fluctuations after the war.

ARIMA and SARIMA closely follow the actual price trends up to and after Easter compared to other models. ARIMA and SARIMA underestimate the egg prices pre-Easter. They predict that egg prices started to decrease one week before Easter. Their predictions begin to diverge after the holiday, with both showing an overestimation of prices post-Easter.

Prophet captures the upward trend before Easter but shows a significant divergence from the actual prices. Although it was defined holiday parameters, Prophet model fails to capture the downtrend trend after the holiday. It shows that Prophet might overfit in scenarios where sharp price changes occur.

LSTM appears to struggle to capture the price dynamics during the Easter period, showing a large forecasting fluctuation. It underestimates the price spike around Easter, forecasting a price decline one week after Easter. However, LSTM predicts the prices after Easter to be the most accurate compared to other models.

Naive and Hyperbilic models also fail to closely follow the actual prices. They struggle to capture a downward trend after Easter's impact. Their predictions started to decline one month after Easter, as the new data around Easter were fed into the model. This highlights that it is hard to predict longer time horizons without the use of two benchmark models without new data.

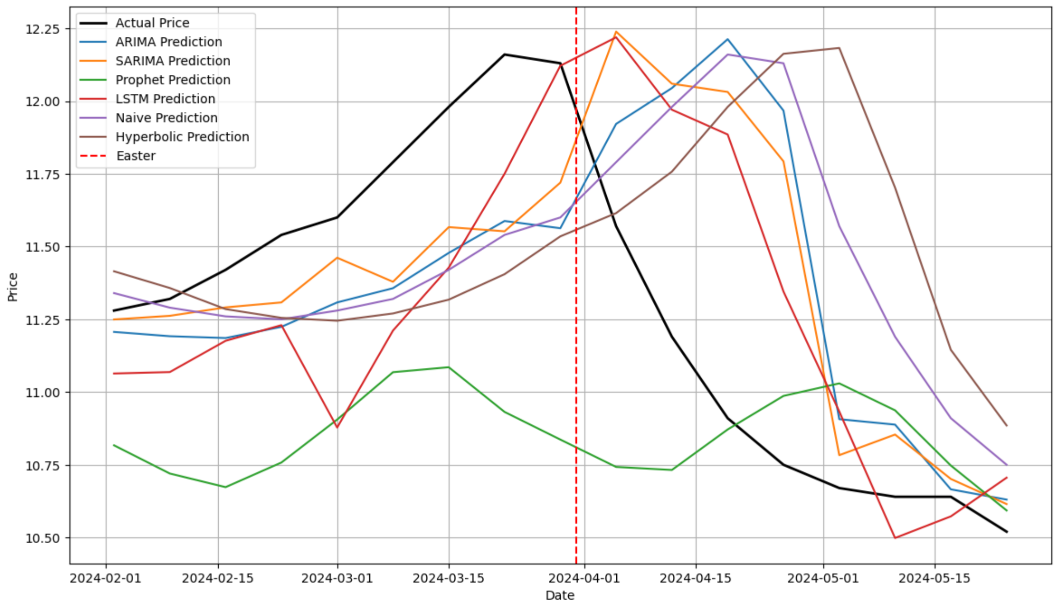


Figure 36: Model Predictions and Actual Prices Around Easter in 2024

In 2023, the egg market fluctuated due to the war. The sharp price changes lead to the overfitting of models. Figure 36 shows the predictions around Easter in 2024. The actual egg prices started to decline at the same week of Easter Sunday, which was 31 March 2024. Like in 2023, the models in 2024 capture a general price increase before Easter, but the accuracy across models shows some variation.

ARIMA and SARIMA perform relatively well, with both models closely following the price increase leading up to Easter Sunday. Similar to 2023 predictions, both models underestimate egg prices before Easter. However, unlike 2023, ARIMA and SARIMA failed to predict the declining egg prices. SARIMA models predict it occurs one week after Easter. ARIMA model showed a one-month delay of declining prices, which performed better than benchmarks.

Prophet continues to overestimate prices, with predictions that either overshoot or undershoot the actual price movements. This behavior suggests that Prophet may struggle with non-linearities and overestimate the impact of external factors such as holidays.

LSTM, while still underestimating the pre-Easter price rise, performed well at tracking the overall trend compared to other models. The accuracy of LSTM predictions was the best across all models. LSTM model responded to the downturn after Easter faster than other models.

Naive and Hyperbolic models again show weaker performance, with a more than one-month lagging response to the observed downturn. Both predictions were far from the actual prices.

Overall, the Easter holiday presents a unique challenge for egg price forecasting models due to the increased demand and subsequent price fluctuations. Across both years, SARIMA and LSTM handled these dynamics well, while ARIMA exhibited limitations in 2023 and Prophet performed badly in capturing abrupt egg price changes. Benchmark models failed to identify holiday impact, highlighting the importance of sophisticated models to handle holiday factors effectively.

6.3.2 Russia-Ukraine War Period

The Russia-Ukraine war, which began in February 2022, caused significant disruptions in global commodity markets, particularly affecting egg prices, as shown in Figure 6. This section evaluates the performance of models over a 4-week expanding window from 2022-02-01 to 2023-02-01 to capture the war's impact on egg prices. The performances of models were assessed using MAE, MSE, RMSE, and MAPE metrics in Table 5, and were visually compared to actual prices in Figure 37.

ARIMA performed reasonably well, with error metrics lower than Prophet and Hyperbolic but higher than LSTM. Although ARIMA was able to model the general trends during the war period but

Model	MAE	MSE	RMSE	MAPE (%)
ARIMA	0.9280	1.6122	1.2697	8.89
SARIMA	0.9497	1.4929	1.2219	9.15
Prophet	1.3140	2.9837	1.7273	12.11
LSTM	0.6812	0.8097	0.8998	6.30
Naive	0.9110	1.7413	1.3196	8.50
Hyperbolic	1.1331	2.5417	1.5943	10.45

Table 5: Model Performance Metrics During Russia-Ukraine War for 4-Week Expanding Window

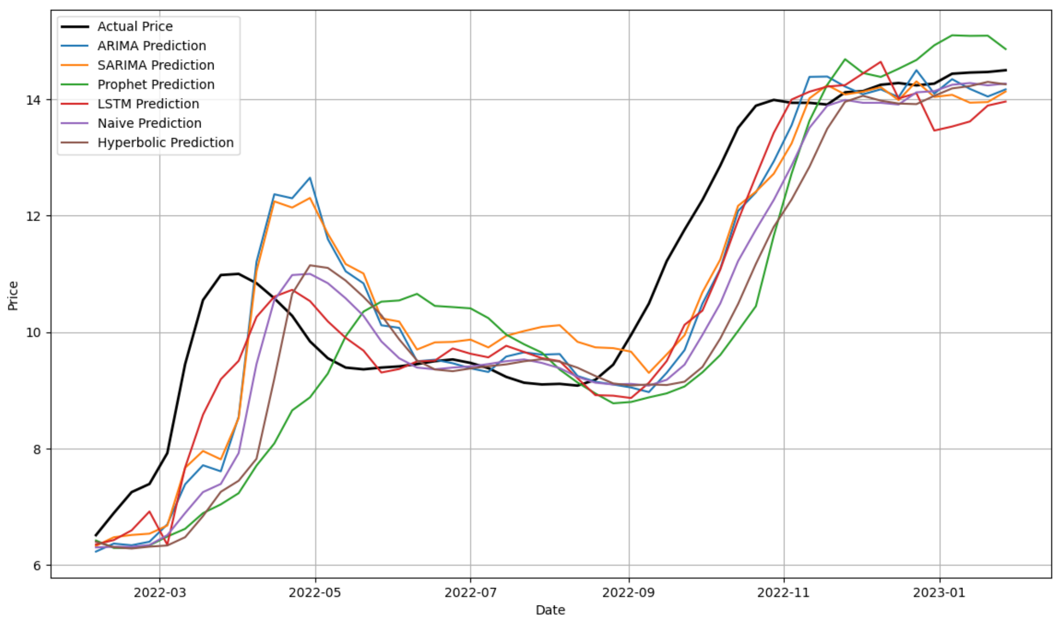


Figure 37: Model Predictions and Actual Prices During Russia-Ukraine War

its linear nature limited its ability to fully capture the sharp price changes caused by the war. ARIMA model overestimated egg prices in April 2022. It reacted slowly to the downtrends after mid-March. The MAPE of 8.89% shows ARIMA's moderate performance when predicting percentage errors during volatile periods.

SARIMA model incorporated seasonality, and performed similarly to ARIMA but had slightly better MSE and RMSE values. SARIMA underperformed ARIMA in terms of MAE and MAPE. These showed that SARIMA made more frequent small errors. The reason could be the seasonality of SARIMA helped it model cyclical patterns better than ARIMA, but the volatility during the war resulted in frequent smaller errors, reflected in MAE and MAPE. In Figure 37, SARIMA followed the actual trend but also struggled during abrupt market shifts, similar to ARIMA. In addition, SARIMA introduced more errors during the period of June to September of 2022, compared to ARIMA.

Prophet had the weakest performance among the models, including two benchmark models, with the highest values in all error metrics. This is largely due to Prophet's reliance on predefined trends and seasonal patterns, which made it less adaptable to the external shocks caused by the war. Prophet was unable to accurately model the non-linear and abrupt changes in egg prices, leading to significant deviations from the actual prices. Its MAPE value of 12.11% indicates that it consistently produced a larger percentage of errors.

LSTM model outperformed the other models, achieving the lowest values for all error metrics. Its ability to capture the non-linear relationships and long-time dependencies allowed it to adjust to the volatile changes during the war period. As shown in Figure 37, LSTM closely monitors the actual price movements, particularly during the early price surge and recovery phase. Its good performance suggests that LSTM is better equipped to handle unexpected market shocks and non-linear fluctuations in data compared to traditional models.

Naive and Hyperbolic as simplistic models performed better than Prophet but underperformed complex models like LSTM, ARIMA, and SARIMA. They captured the general directly of price movements, but lagged in their response to rapid price changes. Compared to ARIMA and SARIMA, Naive, and Hyperbolic were close to the actual price around May 2022. The simplistic nature helps them reduce overfitting after unexpected price movements.

In conclusion, from the 4-week expanding window, it is clear that LSTM was the most effective model during the Russia-Ukraine war period, achieving the lowest errors and closely following the actual price trends. ARIMA and SARIMA performed similarly but struggled with the sudden volatility caused by external shocks. Prophet was the least effective, while Naive and Hyperbolic provided moderate performance but failed to capture the non-linear relationships in the data.

6.3.3 Fipronil Crisis Period

As shown in Figure 5, the Fipronil crisis between July 2017 and April 2018 was an unexpected and extreme event in the egg market that significantly affected prices due to localized market disruptions. This period caused significant fluctuations in egg prices, making it difficult for models to accurately predict prices since the crisis was not related to key external features such as feed costs, energy costs, and the CPI. Figure 38 presents the comparison between actual prices and predictions generated by models during August 2017 to end of April 2018 usring a 4-week expanding window.

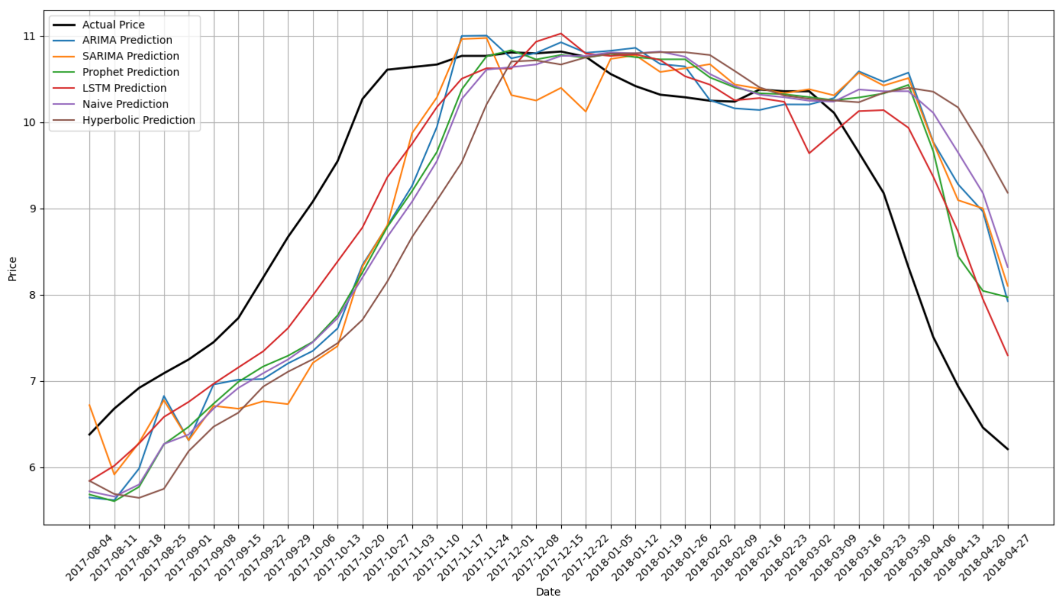


Figure 38: Model Predictions and Actual Prices During Fipronil Crisis Period

Both ARIMA and SARIMA perform similarly during this period. Their linear structure, which relies on autocorrelation and moving average, was not sufficient to handle the suddent price spikes.

In addition, Prophet model also struggled during this period. The crisis was unrelated to any predictable holiday patterns, and Prophet's ability to handle business seasonality failed in the context of unprecedented market event. Unlike other evaluation period like Easter and Russia-Ukraine war, Prophet model performed relatively well during Fipronil period, which did not deviate from actual price largely. It reflected the better tuning of Prophet need to adopt different market events.

LSTM performed marginally better than the other models, especially during period before 2018. LSTM responded to upward trends quickly when the start of crisis and at the end of crisis. Although its ability to capture non-linear relationships allowed it to handle some of crisis-induced dynamics, the unpredictability caused by crisis made LSTM difficult to predict accurate prices.

Both Naive and Hyperbolic models displayed similar behavior during the crisis. Since they are simple model, their performance reflects a general inability to handle the volatile market movements during the Fipronil crisis.

As expected, all models struggled to accurately forecast egg prices during this period, which was

Model	Remove Fipronil Data				Original Data			
	MAE	**MSE**	**RMSE**	**MAPE (%)**	**MAE**	**MSE**	**RMSE**	**MAPE (%)**
ARIMA	0.3639	0.1739	0.4170	6.35	0.3639	0.1739	0.4170	6.35
SARIMA	0.3281	0.1419	0.3766	5.64	0.3281	0.1419	0.3766	5.64
Prophet	0.3731	0.1884	0.4341	6.40	0.3731	0.1884	0.4341	6.40
LSTM	0.2451	0.0742	0.2724	4.24	0.2460	0.0751	0.2740	4.26
Naive	0.3861	0.1883	0.4340	6.67	0.3861	0.1883	0.4340	6.67
Hyperbolic	0.4856	0.2899	0.5384	8.36	0.4856	0.2899	0.5384	8.36

Table 6: Model Performance Metrics With and Without Fipronil Data

dominated by local market events that disrupted the usual price patterns. Since these disruptions were not directly related to the external features included in the models, it is reasonable that none of models could capture the underlying dynamics. Given the poor performance of models during the Fipronil crisis, the model performance was evaluated after removing Fipronil crisis data. The data from 2017-06-30 to 2018-05-01 was removed and the models were re-evaluated over the period from 2018-06-01 to 2019-01-01. The model performance metrics with and without Fipronil data was shown in Table 6.

It is clear that removing the Fipronil crisis data had no significant impact on the performance of most models. For the ARIMA, SARIMA, Prophet, Naive, and Hyperbolic models, the metrics remained unchanged, showing that thesemodelsl did not respond to the removal of the extreme market event data. However, LSTM model displayed a slight improvement in the performance after the Fipronil data was removed. The MAE decrease from 0.2460 to 0.2451, and RMSE improved from 0.2740 to 0.2724, indicating that LSTM is more sensitive to the presence of extreme data in the training set.

ARIMA and SARIMA are linear models that rely on past values and seasonal patterns to make predictions. The removal of Fipronil data had less impact on the patterns that were identified by ARIMA and SARIMA models. Similar to them, Prophet model captures the overall trends and seasonality. The removal of data did not significantly alter the long-term patterns that Prophet is designed to detect, so its results also stayed the same. Two benchmark models did not rely on complex patterns or past events beyond the last observation, removing the data had no effect on their performance.

Unlike traditional models like ARIMA, SARIMA, and Prophet, LSTM can learn and represent complex relationships between the features and target variable. When external events like the Fipronil crisis introduce sharp, irregular fluctuation in the data, LSTM tried to adjust its internal memory to account for these anomalies. However, such market events may not follow the general trends that LSTM is learning, which introduces noise and confusion into the model, leading to a decrease in predictive accuracy. After removing extreme data points, LSTM model can better capture the general trends in the data. The minor improvement of LSTM model suggested that it is better for the more stable period.

Overall, the evaluation of the models during the Fipronil crisis highlights the limitations of all models in dealing with sudden and unpredictable market events. However, when the crisis data was removed, LSTM showed clear improvements, while other models maintained similar performance levels. This suggests that LSTM is more sensitive to extreme events.

7 Results

Based on the evaluation of various models using different expanding windows (1 week, 4 weeks, and 12 weeks) in Section 6.2, ARIMA was selected as the best-performing model for short-term prediction, while LSTM model was chosen for medium-term and long-term predictions. These selections were made after considering the MAE, MSE, RMSE and MAPE metrics across all models, with ARIMA consistently delivering more stable and accurate results for short-term predictions, and LSTM showing the best performance in capturing long-term trends.

This section presents the results of the models in short-term (Section 7.1), medium-term (Section 7.2), and long-term (Section 7.3) predictions. The selected period for analysis is from September 2023 onward, as the egg prices had stabilized after the disruption caused by the Russian-Ukraine war. To evaluate the reliability of the predictions, the confidence intervals of 95%, 50%, and 25% are generated using the bootstrap method. These confidence intervals provide a way to quantify the uncertainty of model forecasts.

The confidence interval provides a range within which the actual values of a predicted variable are expected to fall, with some certainty [46]. Confidence intervals are important to evaluate the robustness and reliability of the model predictions. By showing a range of likely values instead of an exact point, confidence intervals account for the uncertainty of the model's predictions. In this thesis, the confidence intervals 95%, 50%, and 25% represent different levels of certainty. The confidence interval 95% provides the widest range in which almost all actual prices could fall. 50% and 25% confidence intervals offer the narrower range where the price is less likely to fall.

In traditional statistical methods, confidence intervals are often based on the assumption that residuals or errors follow a normal distribution. However, the residuals in this thesis deviate from this assumption, as shown in Figures 27, 28, and 30. Especially, the influence of external factors like feed prices, energy prices, and CPI, made the residuals exhibit a non-normal distribution. The bootstrap method is a statistical technique that involves repeatedly resampling the residuals to create simulated confidence intervals [18]. The bootstrap method does not require the assumption about the distribution of residual, making it was chosen to construct confidence intervals in this thesis.

7.1 Short-Term Prediction using ARIMA

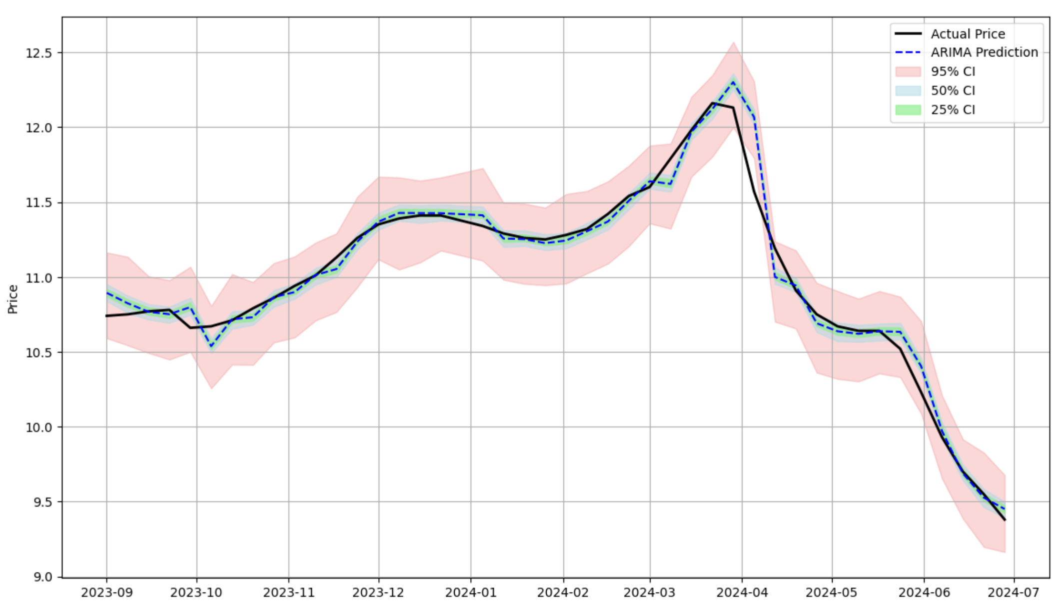


Figure 39: ARIMA Model Predictions with Confidence Intervals (1 Week)

Figure 39 illustrates the performance of ARIMA model for short-term predictions using a 1-week expanding window. The predicted values are shown alongside the actual prices, along with confidence intervals of 95%, 50%, and 25%. ARIMA's predicted values (blue dashed line) closely follow the actual egg price trends (black line), especially during periods of steady price movement. However, the prediction deviates slightly from the actual price between March and May 2024.

ARIMA model handles egg price trends relatively well in periods of stability but it struggles with significant price movements, such as the peak around the Easter holiday in 2024. Although ARIMA is known for handling linear patterns in the short term, it may have difficulty fully capturing the nonlinear market changes, such as increasing demand.

The 95% confidence interval (red-shaded region) provides the widest range, capturing the most of price variations during the prediction period. Only prices of the beginning of April 2024 are out of the range of 95%. This indicates that ARIMA Model maintains a level of uncertainty, especially during the volatility around the holiday season. The 50% confidence interval (blue-shaded region) is narrower, but it still encompasses most of the actual price values, showing that ARIMA Model is generally able to predict egg prices with reasonable accuracy within this range. The 25% confidence interval (green-shaded region) is the tightest, focusing on the most likely price predictions. However, the actual prices move outside this interval.

Overall, ARIMA provides a robust short-term prediction with good accuracy, especially in periods of stable prices. However, during periods of high market volatility such as Easter-related price peaks, the model shows increased uncertainty. In the next section, the results of LSTM of medium-term prediction will be presented.

7.2 Medium-Term Predictions using LSTM

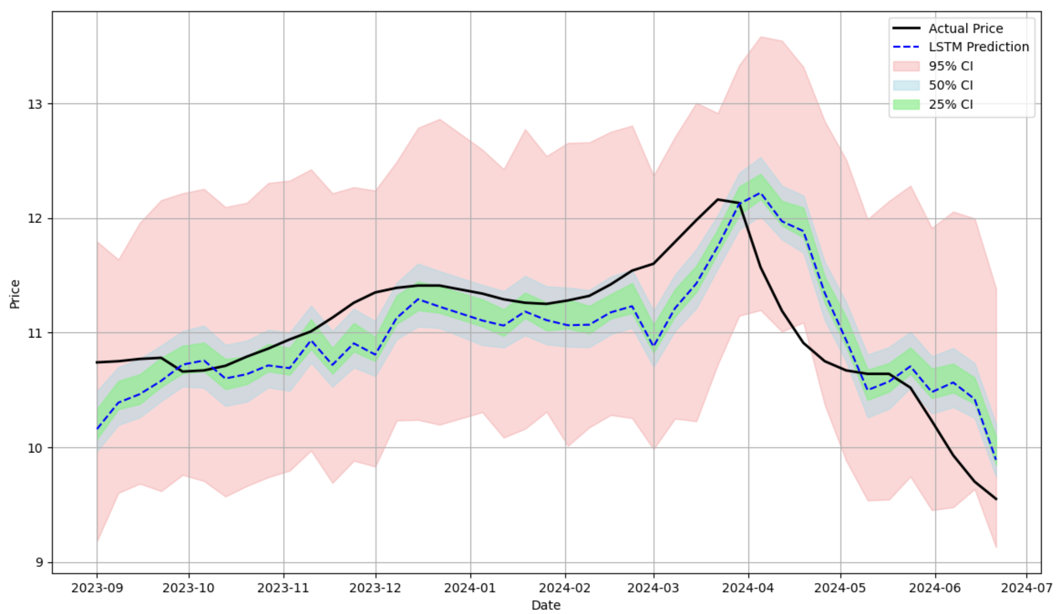


Figure 40: LSTM Model Predictions with Confidence Intervals (4 Weeks)

Figure 40 shows the performance of LSTM model for medium-term predictions using a 4-week expanding window. Compared to short-term predictions, the accuracy of medium-term predictions decreases because a longer horizon introduces more uncertainty.

LSTM Model captures the overall trends of the egg prices based on 4-week expanding window, particularly at the end of 2023 and the beginning of 2024. Its ability to model complex, nonlinear relationships in the data allows it to perform well during periods of steady growth and mild fluctuation in prices. Like ARIMA model in short-term predictions, LSTM captures the general rise and peak in egg prices. However, the predictions lag behind the actual prices, especially during the Easter holiday.

The 95% confidence interval (shaded in red) is broad, particularly in the earlier period, reflecting the LSTM's uncertainty in the first few months. This interval widens further after the price peak of Easter. indicating higher uncertainty around sudden price surges. In addition, the 50% confidence interval (shaded in blue) is narrower and captures the majority of the actual price movement. Most actual prices value fall within this interval, showing LSTM can provide reasonably accurate predictions in the medium term. Nevertheless, the 25% confidence interval (shaded in green) is the tightest. Within this confidence level, few actual prices fall within this interval.

So, LSTM model shows strong performance for medium-term price predictions, offering a reliable balance between accuracy and flexibility. Although there is some level of prediction uncertainty around peak prices, LSTM provides a reliable solution for the trends next month.

7.3 Long-Term Predictions using LSTM

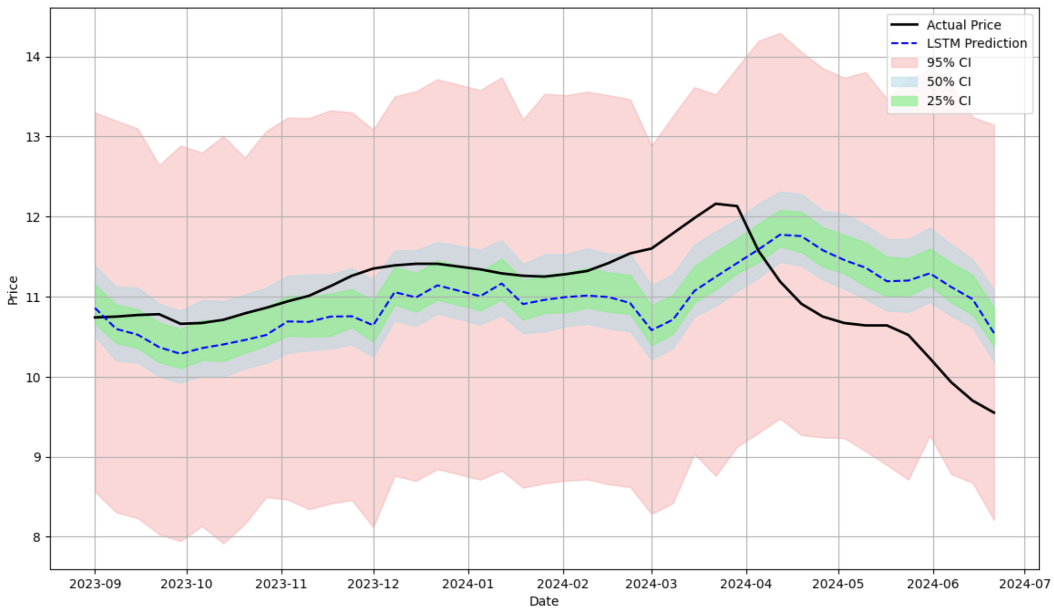


Figure 41: LSTM Model Predictions with Confidence Intervals (12 Weeks)

Figure 41 demonstrates the performance of LSTM model for long-term predictions using a 12-week expanding window. LSTM model shows reasonable accuracy over much of the prediction period. The model follows the overall trend of actual prices. However, due to the extended prediction horizon, LSTM exhibits increasing uncertainty compared to short-term and medium-term predictions, as reflected in widening confidence intervals.

Before March 2023, the model captures the gradual price rise. However, it tends to underestimate the magnitude of the price spike around Easter of 20LSTM predicts a dip at the beginning of March, which is opposite to the actual prices. Similarly, the downward trend after Easter is captured, although the prediction lag behind the actual price, which is reasonable for long-term predictions.

The 95% confidence interval (red-shaded region) is quite wide throughout the prediction horizons, where all actual prices fall within this region. The region widens at the end of the period, indicating that LSTM's uncertainty increased significantly as it attempts to predict longer horizons. Plus, the 50% and 25% confidence intervals (shaded in blue and green, respectively) are narrower, particularly up to February 2024. showing the model has more confidence in its prediction in the earlier stages of the prediction window.

In conclusion, LSTM model performs relatively well, despite the increasing uncertainty inherent in forecasting such a complex market. Its ability to model complex relationship make it more suitable for long-term prediction compared to tradition methods. However, the model struggles with holiday peaks. Despite this, LSTM's flexibility and sequential memory make it a robust choice for long-term forecasting in the egg price market.

8 Conclusion and Discussion

8.1 Conclusion

This thesis aims to develop robust egg price forecasting models in the Dutch market by using various time series and machine learning approaches, including ARIMA, SARIMA, Prophet, and LSTM. The goal was to capture both linear and non-linear relationships between egg prices and key features like feed costs, energy costs, and CPI, and assess the model's performance across different prediction horizons (short, medium, and long-term). Throughout the study, public data sources were utilized, and feature selection process incorporated economic and market-specific variables to improve the accuracy of the forecasts.

The key steps in this thesis include data collection and preprocessing, feature analysis and selection, theoretical background research, model implementation, and extensive model evaluation. Data analysis reveals that trends in egg prices were related to events such as the Fipronil crisis and the Russia-Ukraine war. It also explains the seasonalities and patterns of egg prices. Within one year, egg prices are usually high around Easter holidays and winter.

After analyzing various features such as feed costs (soybeans, maize, and wheats), energy costs (crude oil and natural gas), and broader economic indicators (CPI). Soybean and crude oil prices have been portrayed as critical contributors to the volatility of egg prices. Through correlation analysis and the lag test, it was found that feed costs had a stronger and more consistent lagged effect on egg prices, making it a key feature of the models. Energy costs, though less impactful, showed a notable lag effect, particularly from late 2022, when energy prices contributed to higher egg prices. Furthermore, CPI was included as an inflation indicator, while chicken prices and weather variables were ultimately dropped due to their weak correlation with egg prices. These findings highlight the importance of including external factors in forecasting models.

The thesis employed a range of models, including ARIMA, SARIMA, Prophet, LSTM, Naive, and Hyperbolic, to forecast egg prices. ARIMA and SARIMA models effectively captured linear trends and seasonality in the short term, but their limitations in handling nonlinear relationships and holiday impact became evident, particularly in longer-term forecasts. Prophet, while showing strong performance during weekly predictions, faced challenges in predicting market fluctuation with expanding window method towards the end, particularly around Easter. Despite its ability to incorporate holidays, Prophet failed to capture the full impact of these external events, showing the need for further parameter tuning and refinement.

LSTM, on the other hand, proved to be more capable of capturing complex relationships and a longer dependence between features. This was particularly evident in medium-term and long-term predictions, where LSTM outperformed the traditional models. LSTM also performed well during the Russia-Ukraine war period and the Fipronil crisis period, proving its flexibility and ability to handle complex market dynamics. By incorporating non-linear interaction features and carefully selecting important variables, LSTM model provides an accurate prediction framework.

To simulate the real business scenario, the expanding window approach was applied in the thesis, where the models were re-trained with each new data update. This method allows models to continuously learn from the latest data, increasing their ability to adapt to changing market conditions. Three different forecast horizons were defined as short-term (1 week), medium-term (4 weeks), and long-term (12 weeks), allowing the evaluation of model performance across different time horizons for different business scenarios. The evaluation process involved the use of a set of metrics, including MAE, MSE, RMSE, and MAPE, which helped quantify the precision of each forecasting model. It confirmed that the LSTM model performed best in medium and long term predictions, while ARIMA and SARIMA remained effective in short-term prediction. The use of benchmark models, Naive and Hyperbolic, further validated the results, showing that while these simple models can handle basic trends in short term, more advanced models are necessary for accurate medium-term and long-term predictions. Bootstrap methods were also used to generate confidence intervals, adding reliability to the evaluation process by accounting for uncertainty in the predictions.

The significance of this thesis lies in its contribution to the egg market and to stakeholders involved in production and distribution. By using public data sources such as the World Bank and CBS, the study demonstrates how features can be tracked over time to inform decision-making in egg production. Stakeholders, including farmers, distributors, and retailers, can use these models and analysis to anticipate price trends and make informed business decisions, such as optimizing production schedules or managing inventory based on market conditions. These analysis of different expanding windows further provides the practical application of these models for both short-term planing and long-term strategy.

8.2 Discussion

Although this study presents valuable information and models for egg price forecasting, several limitations have been considered.

First, data for feed and energy costs were sourced from global databases of the World Bank, which may not fully reflect local market conditions. Localized data could provide highly accurate and relevant forecasts, particularly for a specific market such as the Netherlands.

Second, while lagged features such as feed costs, energy prices, and CPI were analyzed and proved, they assume linear correlation between the lagged variables and the target variable. By incorporating lagged features in ARIMA, SARIMA, and Prophet models, the linearity does not fully capture the complexity of real-world agricultural markets. The interactions between variables are oftern non-linear and vary over time. As a result, models can over-simplify these relationships, potentially limiting their ability to accurately predict egg prices.

Third, while Prophet performed well in the one-week prediction in the beginning of this thesis, its performance declined over development. The model also did not predict price increases during Easter, although holiday effects were included in the model. The poor performance of Prophet towards the end of forecast highlights the need for further tuning and potentially more complex methods to handle price spikes like the Fipronil crisis and war.

Fourth, due to time constraints, other machine learning models such as RNN and XGBoost were not tested. Although the literature supports the ability of LSTM in agricultural price forecasting, it is valuable to test more advanced machine learning models.

Lastly, future research could explore hybrid or ensemble approaches by combining models such as ARIMA, SARIMA, and LSTM to use their strengths. For example, the linear trends captured by ARIMA could complement the non-linear dynamics of LSTM, resulting in a more reliable forecasts.

In conclusion, this thesis shows the potential of applying various time series models and machine learning techniques to forecast egg prices in the Dutch market. The findings highlight the importance of selecting appropriate features and models in different prediction horizons and market dynamics. Although it offers practical insights and forecasting tools for stakeholders in the egg industry, there remains space for improvement. Incorporating model-localized data and experimenting with advanced or hybrid models could further improve the prediction of these predictions, making them more applicable to real-world decision-making.

8.3 Recommendations

This thesis, conducted within the framework of The Driving Force (TDF), provides valuable insights that can directly benefit the company's ongoing development of the BuyFresh platform. Based on the findings and results, several constructive recommendations can be made for TDF and its platform, BuyFresh, to increase the value it offers to egg farmers and local markets.

First, one of the key insights from this thesis is the importance of accurate price forecasting for decision making in the poultry sector. TDF could apply predictive analytics tools to the BuyFresh platform to provide egg farmers with real-time price forecasts based on models such as ARIMA for short-term predictions and LSTM for medium and long-term predictions. These forecasts would help farmers make informed decisions about production, pricing, and supply management.

Second, the analysis showed that feed costs, especially soybeans, and energy costs such as crude oil, play a significant role in impacting egg prices. TDF can use this by integrating these external factors, such as trends of feed prices and energy prices, into the BuyFresh platform, allowing farmers to monitor and adjust their pricing strategies accordingly. This would help farmers understand the market dynamics and make proactive decisions in response to changing production costs.

Third, the analysis of seasonal trends, particularly around holidays like Easter showed predictable spikes in egg prices. TDF could recommend that farmers adjust their production levels and marketing efforts based on these seasonal trends. The BuyFresh platform could offer marketing and sales strategies, such as promotions or increased supply before these periods, to use seasonal demand and help farmers optimize their revenue.

Fourth, the results also highlight the impact of external shocks like the Russia-Ukraine war on feed and energy costs, which in turn affect egg prices. Although it is difficult to predict the war, TDF could develop an alert system within BuyFresh to notify farmers of market changes and potential disruption. This system could provide farmers with sufficient time to adjust their operations for market fluctuations in production prices.

Lastly, while the models developed in this thesis relied on global data sources, the inclusion of local market data could improve the accuracy of the prediction. TDF could consider incorporating local feed and energy costs into its forecasting models to provide farmers with a more relevant and precise pricing prediction. This could help BuyFresh become a more reliable and valuable resource for local producers.

By implementing these recommendations, TDF can further improve the BuyFresh platform's capabilities, offering egg farmers advanced tools to navigate market fluctuation, make better pricing strategies, and optimize production in a dynamic environment. These insights within this thesis will help TDF become a more reliable platform that is customized to the specific sales needs of its users. The integration of predictive analytics, seasonal marketing information, and market trends alerts will not only attract more farmers to the platform, but will also provide them with advanced, reliable, and informed advice for both sales and production strategies. This will strengthen TDF's role as a trusted partner for fresh producers, enabling them to make data-driven decisions that enhance business performance and promote long-term sustainability in the agricultural sector.

Appendices

The appendices section includes additional figures that are relevant to the study but were omitted from the main body of the thesis due to space limitations or to avoid overlapping with other key visuals. These supplementary materials provide further insights into the analysis and support the finding discussed in the main chapters.

Figure 42 shows the trends in egg prices from 2013 to 2024 with different sizes and colors. The price of the Amsterdam index provides trends similar to the target price of NOP 2.0, as shown in Figure 2 in the main body. It proves that the Amsterdam Index and NOP 2.0 are influenced by the same market dynamics, some analysis obtained from the Amsterdam Index data sources at the beginning of this thesis can also apply for NOP 2.0 to develop forecasting models.

Figures 43 and 44 present the yearly average and median prices for different size eggs of Amsterdam Index, supplementing Figure 3 and 4 in the main body. Each subplot within these figures illustrates the yearly trends for XL, L, M, and S-size eggs. These figures clearly show that larger eggs tend to have higher prices, and except for S-sized eggs, brown eggs are generally priced higher than white eggs. In particular, the impact of the Russian-Ukraine war in 2022 caused a significant increase in egg prices and reduced the price gap between brown and white eggs in sizes XL, L, and M.

Figure 45 displays weekly price changes for medium eggs around Easter. Each subplot represents the egg price fluctuations for the 12 weeks before and 2 weeks after Easter Sunday for each year from 2017 to 2024. As a supplement to Figure 9 in the main body, this figure provides a clearer view of the changes year-by-year, highlighting the variation in egg prices surrounding the Easter holiday.

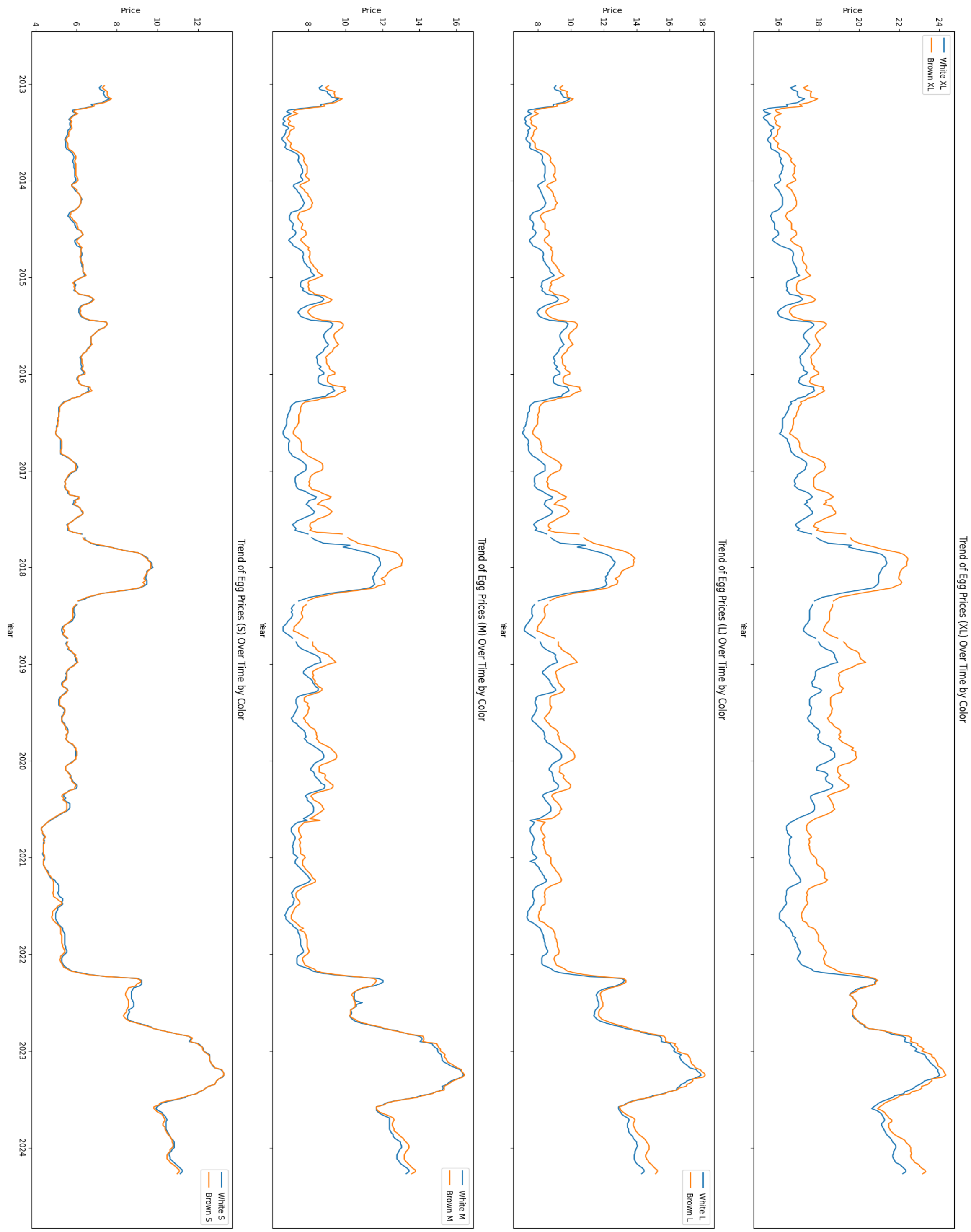


Figure 42: Trends of Egg Prices over Time by Color and Size. Source: Amsterdam Index from Pluimveebeurs.com

Figure 43: Yearly Average Prices for XL, L, M, and S Eggs. Source: Amsterdam Index from Pluimveebeurs.com

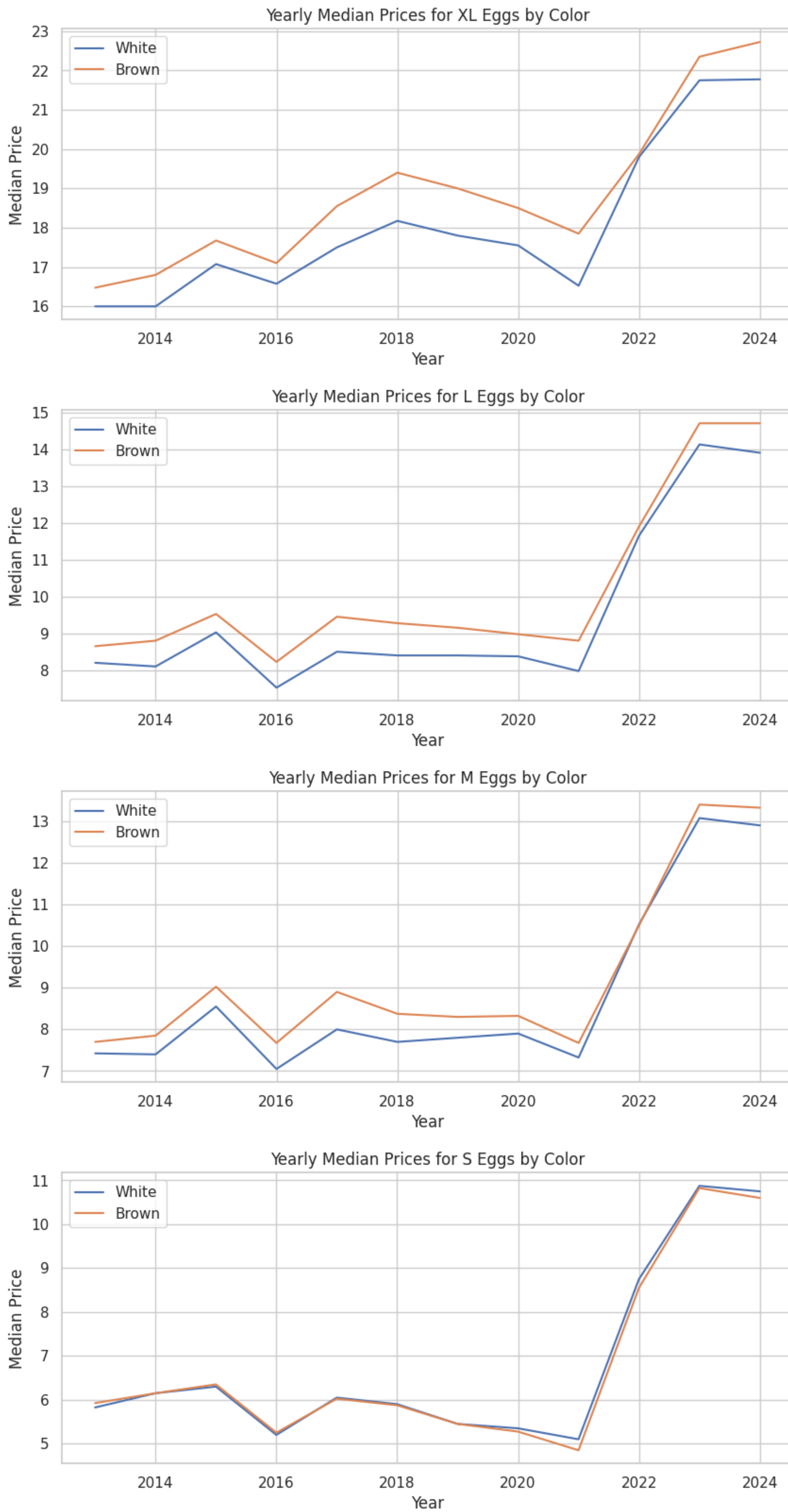
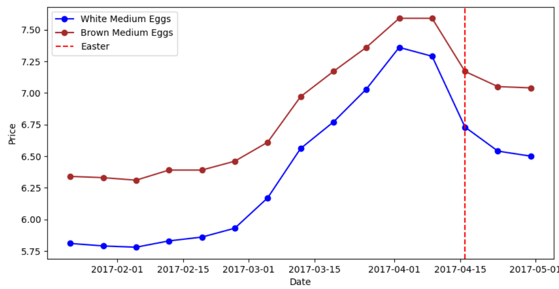
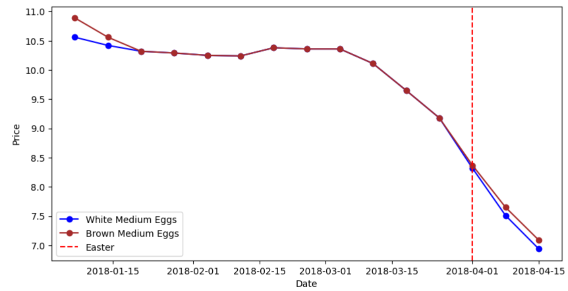


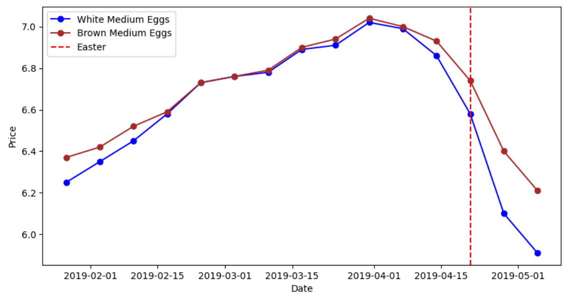
Figure 44: Yearly Median Prices for XL, L, M, and S Eggs. Source: Amsterdam Index from Pluimveebeurs.com



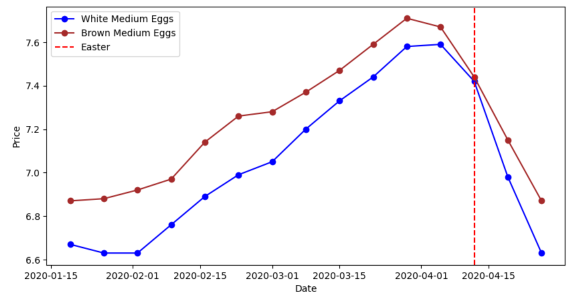
(a) Easter 2017



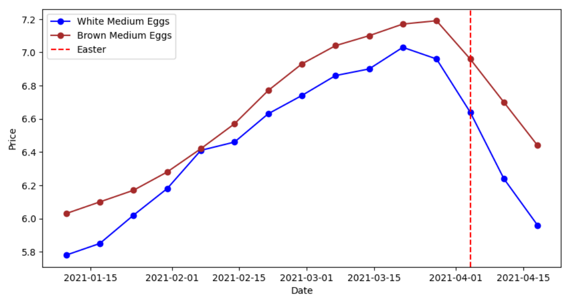
(b) Easter 2018



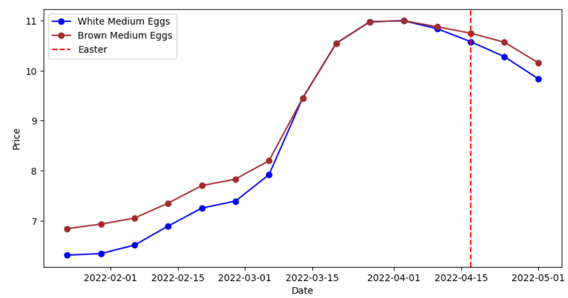
(c) Easter 2019



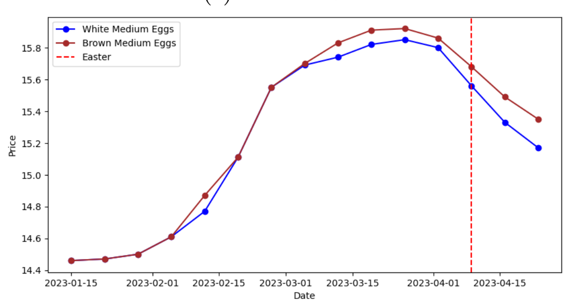
(d) Easter 2020



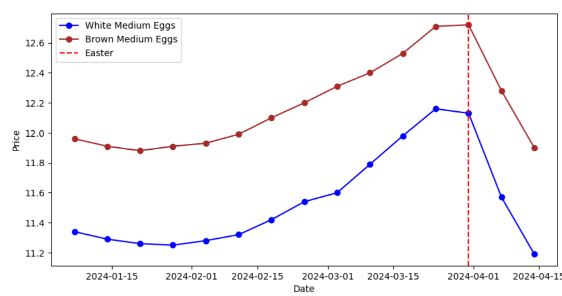
(e) Easter 2021



(f) Easter 2022



(g) Easter 2023



(h) Easter 2024

Figure 45: Weekly Price Changes for Medium Eggs around Easter from 2017 to 2024. Source: NOP2.0 from Pluimveebeurs.com

References

[1] HA Ahmad and M Mariano. Comparison of forecasting methodologies using egg price as a test case. *Poultry science*, 85(4):798–807, 2006.

[2] Michael Olusegun Akinwande, Hussaini Garba Dikko, Agboola Samson, et al. Variance inflation factor: as a condition for the inclusion of suppressor variable (s) in regression analysis. *Open journal of statistics*, 5(07):754, 2015.

[3] Fahad Radhi Alharbi and Denes Csala. A seasonal autoregressive integrated moving average with exogenous factors (sarimax) forecasting model-based time series approach. *Inventions*, 7(4):94, 2022.

[4] Agustin Garcia Asuero, Ana Sayago, and AG González. The correlation coefficient: An overview. *Critical reviews in analytical chemistry*, 36(1):41–59, 2006.

[5] Gustavo Maria Barboza Martignone, Bikramaditya Ghosh, Dimitrios Paparas, and Karl Behrendt. A shift in leadership on the international soybean market; the rise of the triumvirate and the fall of rotterdam. a time-varying investigation. *A Time-Varying Investigation*, 2023.

[6] S Barzegar, S-B Wu, J Noblet, M Choct, and R A Swick. Energy efficiency and net energy prediction of feed in laying hens. *Poultry Science*, 98(11):5746–5758, 2019. ISSN 0032-5791. doi: https://doi.org/10.3382/ps/pez362. URL https://www.sciencedirect.com/science/article/pii/S0032579119457787.

[7] George E. P. Box et al. *Time Series Analysis: Forecasting and Control*. J. Wiley & Sons, 4 edition, 2008. URL http://www.123library.org/book_details/?id=30721.

[8] P. J. Brockwell and R. A. Davis. *Introduction to Time Series and Forecasting*. Springer, 3 edition, 2016. doi: 10.1007/978-3-319-29854-2.

[9] Nikolajs Bumanis, Armands Kviesis, Liga Paura, Irina Arhipova, and Mihails Adjutovs. Hen egg production forecasting: Capabilities of machine learning models in scenarios with limited data sets. *Applied Sciences*, 13(13):7607, 2023.

[10] JL Campo, MG Gil, Davila, and SG. Differences among white-, tinted-and brown-egg laying hens for incidence of eggs laid on the floor and for oviposition time. *Archiv Fur Geflugelkunde*, 71(3): 105–109, 2007.

[11] Tianfeng Chai and Roland R Draxler. Root mean square error (rmse) or mean absolute error (mae)?–arguments against avoiding rmse in the literature. *Geoscientific model development*, 7(3): 1247–1250, 2014.

[12] Zhiyuan Chen, Howe Seng Goh, Kai Ling Sin, Kelly Lim, Nicole Ka Hei Chung, and Xin Yu Liew. Automated agriculture commodity price prediction system with machine learning techniques. *arXiv preprint arXiv:2106.12747*, 2021.

[13] Tserenpurev Chuluunsaikhan, Ga-Ae Ryu, Kwan-Hee Yoo, HyungChul Rah, and Aziz Nasridinov. Incorporating deep learning and news topic modeling for forecasting pork prices: the case of south korea. *Agriculture*, 10(11):513, 2020.

[14] European Commission. Eggs dashboard. https://agriculture.ec.europa.eu/document/download/9bdf9842-1eb6-41a2-8845-49738b812b2b_en, 2024. Accessed: 2024-08-26.

[15] HK Dei. *Soybean as a feed ingredient for livestock and poultry*. IntechOpen London, 2011.

[16] Ramesh Dharavath and Ekaansh Khosla. Seasonal arima to forecast fruits and vegetable agricultural prices. In *2019 IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS)*, pages 47–52, 2019. doi: 10.1109/iSES47678.2019.00023.

[17] Gabriel Di Bella, Mark Flanagan, Karim Foda, Svitlana Maslova, Alex Pienkowski, Martin Stuermer, and Frederik Toscani. Natural gas in europe: the potential impact of disruptions to supply. *Energy Economics*, page 107777, 2024.

[18] Thomas J DiCiccio and Bradley Efron. Bootstrap confidence intervals. *Statistical science*, 11(3): 189–228, 1996.

[19] Dutch News. 200,000 chickens culled as bird flu death toll tops four million, September 2022. URL https://www.dutchnews.nl/2022/09/200000-chickens-culled-as-bird-flu-death-toll-tops-four-million/. Accessed: 2024-07-15.

[20] Dutch News. Brown eggs are disappearing from the supermarket shelves, May 2024. URL https://www.dutchnews.nl/2024/05/brown-eggs-are-disappearing-from-the-supermarket-shelves/. Accessed: 2024-07-15.

[21] Mohammed Elseidi. A hybrid facebook prophet-arima framework for forecasting high-frequency temperature data. *Modeling Earth Systems and Environment*, 10(2):1855–1867, 2024.

[22] FEFAC, the European Compound Feed Manufacturers' Federation. Feed & food statistical yearbook 2023, February 2023. URL https://fefac.eu/statistics/. Published in February 2024. Accessed: 2024-07-15.

[23] Yuqing Feng, Yaojie Zhang, and Yudong Wang. Out-of-sample volatility prediction: Rolling window, expanding window, or both? *Journal of Forecasting*, 43(3):567–582, 2024.

[24] Marsellino Prawiro Halim, Novanto Yudistira, and Candra Dewi. Multicommodity prices prediction using multivariate data-driven modeling: Indonesia case. *IEEE Transactions on Computational Social Systems*, 2022.

[25] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8): 1735–1780, 1997.

[26] R.J. Hyndman and G. Athanasopoulos. *Forecasting: Principles and Practice*. OTexts, Melbourne, Australia, 2 edition, 2018. URL https://otexts.com/fpp2/. Accessed on July 12, 2024.

[27] N.P. Johnston, L.K. Jefferies, B. Rodriguez, and D.E. Johnston. Acceptance of brown-shelled eggs in a white-shelled egg market. *Poultry Science*, 90(5):1074–1079, 2011. ISSN 0032-5791. doi: https://doi.org/10.3382/ps.2010-00914. URL https://www.sciencedirect.com/science/article/pii/S003257911941986X.

[28] Dan Jurafsky and James H. Martin. *Speech and language processing : an introduction to natural language processing, computational linguistics, and speech recognition*. Pearson Prentice Hall, Upper Saddle River, N.J., 2009. ISBN 9780131873216 0131873210. URL http://www.amazon.com/Speech-Language-Processing-2nd-Edition/dp/0131873210/ref=pd_bxgy_b_img_y.

[29] Arnold B Larson. Price prediction on the egg futures market. *Food Research Institute Studies*, 7: 49–64, 1967.

[30] Zhemin Li, Liguo Cui, Shiwei Xu, Lingyun Weng, Xiaoxia Dong, Ganqiong Li, and Haipeng Yu. Prediction model of weekly retail price for eggs based on chaotic neural network. *Journal of integrative agriculture*, 12(12):2292–2299, 2013.

[31] Richie Ruchuan Ma, Tao Xiong, and Yukun Bao. The russia-saudi arabia oil price war during the covid-19 pandemic. *Energy Economics*, 102:105517, 2021.

[32] Gustavo María Barboza Martignone, Bikramaditya Ghosh, Dimitrios Papadas, and Karl Behrendt. The rise of soybean in international commodity markets: A quantile investigation. *Heliyon*, 2024.

[33] Lorenzo Menculini, Andrea Marini, Massimiliano Proietti, Alberto Garinei, Alessio Bozza, Cecilia Moretti, and Marcello Marconi. Comparing prophet and deep learning to arima in forecasting wholesale food prices. *Forecasting*, 3(3):644–662, 2021. ISSN 2571-9394. doi: 10.3390/forecast3030040. URL https://www.mdpi.com/2571-9394/3/3/40.

[34] Andrew Muhammad, Charles Martinez, and Abdelaziz Lawani. Why are eggs so expensive? understanding the recent spike in egg prices. *Choices*, 38(2), 2023.

[35] Nan-Dirk Mulder. What happens after egg prices reach historic highs?, April 2023. URL https://research.rabobank.com/far/en/sectors/animal-protein/what-happens-after-egg-prices-reach-historic-highs.html. Accessed: 2024-07-15.

[36] Ramakrishnan Muthukrishnan and R Rohini. Lasso: A feature selection technique in predictive modeling for machine learning. In *2016 IEEE international conference on advances in computer applications (ICACA)*, pages 18–20. Ieee, 2016.

[37] Pluimveebeurs.com. Amsterdamse index - eierprijzen, 2024. URL https://www.pluimveebeurs.com/prijsinformatie/eierprijzen/amsterdamse-index. Accessed: 2024-04-23.

[38] Pluimveebeurs.com. Kruisem handelsnotering, 2024. URL https://www.pluimveebeurs.com/prijsinformatie/eierprijzen/kruisem-handelsnotering. Accessed: 2024-07-30.

[39] Pluimveebeurs.com. Weser ems, 2024. URL https://www.pluimveebeurs.com/prijsinformatie/eierprijzen/weser-ems. Accessed: 2024-07-30.

[40] Pluimveebeurs.com. Nop richtprijs scharrel 20, 2024. URL https://www.pluimveebeurs.com/prijsinformatie/eierprijzen/nop-richtprijs-scharrel-20. Accessed: 2024-07-15.

[41] David John Post. Forecasting egg prices for the los angeles market. 1972.

[42] Sujit Kumar Roy. *Econometric models for predicting short-run egg prices*. The Pennsylvania State University, 1969.

[43] Royal Netherlands Meteorological Institute (KNMI). Daily weather data in the netherlands, 2024. URL https://www.knmi.nl/nederland-nu/klimatologie/daggegevens. Accessed: 2024-07-15.

[44] Kiran M Sabu and TK Manoj Kumar. Predictive analytics in agriculture: Forecasting prices of arecanuts. 2020.

[45] Sachi Sakanashi, I Kutani, and K Koyama. Scenario analysis on the iran sanctions and its impact on the international oil market. *IEEJ Energy Journal*, 13(4):37–47, 2018.

[46] Ana-Maria Simundic et al. Confidence interval. *Biochemia Medica*, 18(2):154–161, 2008.

[47] J. Sok, P. van Horne, and M. Meuwissen. The impact of the fipronil crisis on the financial performance of dutch laying hen farms. *Parasites & Vectors*, 13:589, 2020. doi: 10.1186/s13071-020-04458-8.

[48] Emma G Stafford, Lisa A Tell, Zhoumeng Lin, Jennifer L Davis, Thomas W Vickroy, Jim E Riviere, and Ronald E Baynes. Consequences of fipronil exposure in egg-laying hens. *Journal of the American Veterinary Medical Association*, 253(1):57–60, 2018.

[49] Statistics Netherlands. Consumer prices; price index 2015=100, 2024. URL https://www.cbs.nl/nl-nl/cijfers/detail/83131NED. Accessed: 2024-07-15.

[50] Feihu Sun, Xianyong Meng, Yan Zhang, Yan Wang, Hongtao Jiang, and Pingzeng Liu. Agricultural product price forecasting methods: A review. *Agriculture*, 13(9):1671, 2023.

[51] Ivan Svetunkov. *Forecasting and analytics with the augmented dynamic adaptive model (ADAM)*. CRC Press, 2023.

[52] S. J. Taylor and B. Letham. Forecasting at scale. *The American Statistician*, 72(1):37–45, 2018. doi: 10.1080/00031305.2017.1380080. URL https://doi-org.vu-nl.idm.oclc.org/10.1080/00031305.2017.1380080.

[53] USDA. Ukraine conflict and other factors contributing to high commodity prices and food insecurity, 2024. URL https://fas.usda.gov/data/ukraine-conflict-and-other-factors-contributing-high-commodity-prices-and-food-insecurity. Accessed: 2024-07-15.

[54] P.L.M. van Horne. *Competitiveness the EU egg sector, base year 2017: international comparison of production costs*. Number 2019-008 in Wageningen Economic Research report. Wageningen Economic Research, Netherlands, 2019. doi: 10.18174/469616. Project number: 2282100296.

[55] P.L.M. van Horne and N. Bondt. *Competitiveness of the EU egg sector, base year 2015: international comparison of production costs*. Number 2017-062 in Wageningen Economic Research report. Wageningen Economic Research, Netherlands, 2017. doi: 10.18174/417151. This research has been commissioned by the EU trade association for egg packers, egg traders and egg processors (EUWEP).

[56] P.L.M. van Horne and N. Bondt. *Competitiveness of the EU egg sector, base year 2021 : international comparison of production costs of eggs and egg products*. Number 2023-006 in Report / Wageningen Economic Research. Wageningen Economic Research, Netherlands, 2023. doi: 10.18174/583668.

[57] Celine Vens and Fabrizio Costa. Random forest based feature induction. In *2011 IEEE 11th international conference on data mining*, pages 744–753. IEEE, 2011.

[58] Yijia Wang. Agricultural products price prediction based on improved rbf neural network model. *Applied artificial intelligence*, 37(1):2204600, 2023.

[59] Cort J Willmott and Kenji Matsuura. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate research*, 30(1):79–82, 2005.

[60] World Bank. Commodity markets, 2024. URL https://www.worldbank.org/en/research/commodity-markets. Accessed on: 2024-07-16.

[61] Yong Yu, Xiaosheng Si, Changhua Hu, and Jianxun Zhang. A review of recurrent neural networks: Lstm cells and network architectures. *Neural computation*, 31(7):1235–1270, 2019.

[62] Dabin Zhang, Shanying Chen, Ling Liwen, and Qiang Xia. Forecasting agricultural commodity prices using model selection framework with time series features and forecast horizons. *IEEE Access*, 8:28197–28209, 2020. doi: 10.1109/ACCESS.2020.2971591.