 https://doi.org/10.5564/jimdt.v6i1.3762

The forecasting for prices of consumer goods using time series methods in Mongolia

Delgermaa Erkhemjargal[1*] Otgonsuvd Badrakh[1]

[1]*Institute of Mathematics and Digital Technology, Mongolian Academy of Sciences, Ulaanbaatar 13330, Mongolia*

**Corresponding author: ddelgermaa74@gmail.com; ORCID:0000-0002-1465-7869*

Article Info: Received: 2024.09.15; Accepted: 2024.10.01; Published: 2024.12.26

Abstract: Time series analysis is important in forecasting price fluctuations of consumer goods, which directly influence economic stability. This study examines the price trends of essential commodities in Ulaanbaatar, Mongolia, from 2012 to 2022. The dataset, sourced from the National Statistical Office of Mongolia, includes prices for 11 key items: flour, bread, rice, beef, milk, yogurt, sugar, eggs, apples, potatoes, and A92 gasoline. A combination of correlation analysis and time series forecasting was used to understand the relationships between goods and to predict future price trends. The results revealed high correlations between certain products, such as eggs and milk, and highlighted the influence of seasonal factors and random movements on prices. Time series models provided accurate predictions for some goods like flour and milk, while others, such as apples and potatoes, showed larger forecast errors. These insights are valuable for policymakers and businesses to anticipate price changes and make informed decisions in economic planning.

Key words: Time series analysis, Correlation analysis, ARIMA

1. Introduction

Time series analysis is one of the most effective methods used for examining data over time, making predictions, and identifying underlying patterns. It is widely applied in various fields such as economics, finance, sales, production, and many others [4]. The main goal of time series methods is to understand trends within data and forecast future behavior. Time series forecasting is essential for economic decision-making, as it aids in predicting revenue flows, product prices, production levels, and more.

Price fluctuations are one of the most important indicators in the economy, directly affecting consumer purchasing behavior, production levels, and trade volume. Moreover, changes in prices have a significant impact on the quality of life, the profitability of businesses, and the overall price stability within a country [8]. Therefore, predicting price changes over time is crucial for maintaining economic stability and for developing policies related to trade, business, and government regulation.

In the context of Mongolia, the price of food and other essential products plays a critical role in economic conditions and directly impacts the livelihood of its citizens. Between 2012 and 2022, there were consistent fluctuations in the prices of essential products in Ulaanbaatar, and these fluctuations are closely related to broader economic indicators [1]. The price increases in consumer goods have a significant effect not only on the standard of living but also on government policy and economic planning.

Food prices, including essential goods like flour, bread, rice, beef, milk, yogurt, sugar, eggs, apples, potatoes, and A92 gasoline, directly influence economic stability and social welfare. Thus, accurately forecasting these prices provides valuable insights into market trends and supports decision-making in both the private and public sectors [10]. The importance of price predictions extends beyond consumer decision-making to strategic policy development at both the corporate and government levels.

The data set used in this study was sourced from the National Statistical Office of Mongolia (NSO) via the 1212.mn website, which provides weekly price information for basic consumer goods. The dataset includes prices for 11 essential commodities, including flour, bread, rice, beef, milk, yogurt, sugar, eggs, apples, potatoes, and A92 gasoline. The primary objective of this study is to examine price changes of essential goods in Ulaanbaatar from 2012 to 2022 using time series analysis and to make price predictions based on these trends. The results of this analysis will be valuable for ensuring economic stability and for implementing effective strategies to control price fluctuations.

2. Methodology

2.1. Time Series Analysis

Time series analysis refers to a set of statistical techniques used to analyze data points collected or recorded at specific time intervals. The primary objective is to identify trends, seasonal variations, and other patterns in data to better understand historical behavior and to predict future outcomes. Time series analysis is widely used across various fields such as economics, finance, and environmental science to forecast future events based on past behavior [3]. Time Series Analysis involves analyzing the past behavior of a variable (or a series of variables) to predict its future values, making it crucial for decision-making in business, policy, and other sectors.

The objectives of time series analysis include:

1. **Descriptive Analysis**: This is the first step in time series analysis, where the goal is to visualize the patterns in data to understand trends, cycles, and irregularities. Graphical tools such as line plots and histograms are often used to observe how the data behaves over time.
2. **Modeling and Forecasting**: The goal of this step is to build statistical models that best capture the observed patterns in the data and use them to forecast future values. Models like ARIMA (Auto-Regressive Integrated Moving Average) and exponential smoothing are popular methods used for forecasting [7].
3. **Validation of Predictions**: After constructing the model and making predictions, it is crucial to validate the forecast accuracy by comparing the predicted values with the observed values. This helps in assessing the model's ability to predict accurately and refining the approach if necessary.

ARIMA (Auto-Regressive Integrated Moving Average)

The ARIMA model, developed by Box and Jenkins (1970), is one of the most widely used models in time series forecasting. It is particularly valuable when analyzing data with trends or seasonality. ARIMA is a combination of three components:

- **AR (Auto-Regressive)**: The AR part of the model represents the relationship between a current value and its past values. For example, the value at time t depends on the values of previous time periods. The general formula for AR(p) is as follows:

$$Y_t = c + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \cdots + \theta_p Y_{t-p} + \varepsilon_t, \tag{2.1}$$

where Y_t is the value at time t, θ_i are the parameters, and ε_t is the error term.

- **I (Integrated)**: The I component is used to make the time series stationary by differencing the series (subtracting the current value from the previous one). The differencing helps remove trends and make the data stable for forecasting. The equation for differencing is:

$$Y_t - Y_{t-1} = \varepsilon_t, \tag{2.2}$$

where t is the error term.

- **MA (Moving Average)**: The MA part models the relationship between the value at time t and the error terms from previous periods. It is represented as:

$$Y_t = c + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \cdots + \theta_q\varepsilon_{t-q} + \varepsilon_t, \tag{2.3}$$

where ε_t represents the residual or error term from the previous time periods.

The combination of these three components, denoted as ARIMA(p, d, q), helps forecast future values by considering the autoregressive component, differencing for stationarity, and the moving average component for error correction.
The parameters of ARIMA are as follows:

- p: The number of autoregressive terms (AR).
- d: The degree of differencing (I).
- q: The number of moving average terms (MA) (Box & Jenkins, 1970).

Exponential Smoothing

Exponential smoothing methods provide an alternative way to model time series data by weighting past observations with exponentially decreasing weights. The key idea is to give more importance to recent observations, as they are assumed to better reflect future trends. Exponential smoothing can be used for forecasting when the data has relatively stable trends or seasonality. The method is based on the following recursive formula:

$$S_t = \alpha Y_t + (1 - \alpha)S_{t-1}, \tag{2.4}$$

where S_t is the smoothed value at time t, Y_t is the actual value at time t, and α is the smoothing parameter [7].

Seasonal Decomposition

Seasonal decomposition involves breaking down time series data into its core components: trend, seasonality, and residuals (irregular variations). This method is helpful for identifying and understanding regular, predictable patterns in the data over time [6].

1. **Trend (T)**: This component reflects the overall direction of the time series data. Trends can be upward, downward, or constant. Identifying the trend allows us to predict future growth or decline. It can be modeled using regression methods or smoothing techniques.
2. **Seasonality (S)**: Seasonality refers to patterns in the data that repeat at fixed, regular intervals. For example, retail sales tend to increase during the holiday season every year. Recognizing these seasonal patterns helps businesses plan for demand surges or drops during certain periods [4].
3. **Cyclic (C)**: Cyclic patterns are similar to seasonal variations, but they occur over longer periods and are not as predictable. These cycles often span multiple years and can be influenced by economic, political, or other long-term factors [4].
4. **Irregular or Random Variations (I)**: These are unpredictable fluctuations caused by unforeseen events like natural disasters, economic shocks, or measurement errors. These variations are not explained by the trend or seasonality and are generally considered as "noise"in the data [6].

Each of these components can be modeled separately to provide a more accurate forecast, and their interaction can reveal deeper insights into the data's behavior. By comparing the observed values with the expected values derived from the trend and seasonality, one can identify anomalies or irregular patterns in the time series data. In summary, time series analysis, through methods like ARIMA, exponential smoothing, and seasonal decomposition, enables researchers and analysts to predict future trends, identify anomalies, and understand the dynamics of time-based data.

2.2. Key Performance Metrics for Time Series Forecasting

Evaluating time series models is a crucial step in determining the accuracy, effectiveness, and stability of a model's predictions. Several performance metrics are commonly used in the literature to assess how well a model performs, each focusing on different aspects of model error. These metrics help to provide a deeper understanding of the model's predictive capability [8].

- Mean Error (ME): Identifies consistent over- or underestimation of values but does not indicate error magnitude [9].
- Root Mean Squared Error (RMSE): Sensitive to large errors, useful for detecting significant deviations but can be skewed by outliers [12].
- Mean Absolute Error (MAE): Provides a balanced view of errors without outlier influence, useful for assessing general model accuracy [11].
- Mean Percentage Error (MPE): Indicates forecast bias as a percentage, but can be misleading when values are small (Bovas, 2015).
- Mean Absolute Percentage Error (MAPE): A widely-used percentage-based metric that's helpful for cross-dataset comparisons, but can be inflated with small actual values [2].
- Mean Absolute Scaled Error (MASE): Compares performance to a simple naïve model, showing if the model adds value [9].
- Autocorrelation at Lag 1 (ACF1): Measures temporal dependencies, helping assess if the model captures time series structure (Box et al., 2015).

These metrics are critical for evaluating time series models, providing different perspectives on forecast accuracy and stability. By utilizing these performance measures, one can assess how well a model fits the data, identify areas for improvement, and make more informed decisions [5].

3. The results

3.1. Correlation Analysis of Consumer Goods Prices

In addition to the time series forecasting, a correlation analysis was conducted to explore the relationships between the prices of various consumer goods. This analysis examines how closely the prices of two products are related, and whether an increase in the price of one product is associated with a similar increase in the price of another.

As shown in Figure 1, the correlation coefficients between various pairs of products were calculated. Higher correlation values indicate a stronger relationship between the prices of the two products. The analysis revealed that products with similar market dynamics, such as those influenced by the same supply chain factors or seasonal demand trends, tend to have higher correlation coefficients. This implies that price fluctuations in one product may be indicative of similar changes in another product.

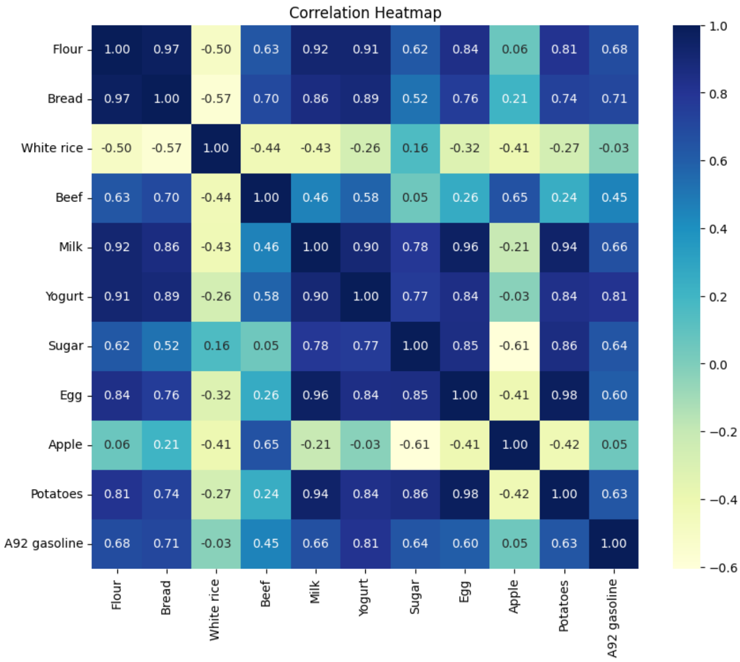

Figure 1: Result of Correlation Analysis

High Correlation Pairs in Consumer Goods Prices

Based on the correlation analysis in the study, several pairs of consumer goods showed a high positive correlation. This suggests that the price changes of one product are closely associated with the price changes of another, likely due to similar production processes, seasonal factors, or market dynamics. The high correlation pairs identified in the analysis include:

1. **Egg and Milk**: The prices of eggs and milk exhibited a strong correlation, likely due to both being staple dairy and poultry products. Their prices are influenced by similar agricultural factors, such as feed costs, farming conditions, and seasonal demand fluctuations.
2. **Milk and Potatoes**: Though these products belong to different categories (dairy vs. vegetables), they demonstrated a significant correlation. This could be attributed to their shared role in consumers' daily diets, meaning price changes in one might reflect broader food price inflation or changes in the cost of agricultural inputs.
3. **Bread and Flour**: Bread and flour are closely related since flour is a primary ingredient in bread. The prices of these two products are strongly linked, with factors like raw material costs, production processes, and seasonal variations influencing both items in similar ways.
4. **Yogurt and Milk**: As dairy products with overlapping production and market dynamics, the prices of yogurt and milk are highly correlated. Increases in milk prices often translate into higher yogurt prices, reflecting the shared production costs and consumer demand for both products.

Implications of High Correlation Pairs:

These high correlation relationships are significant for price forecasting and economic planning. When the price of one item increases, likely, the prices of other closely related goods will also rise. For businesses and policymakers, understanding these relationships can help in anticipating market trends and implementing strategies to manage price inflation or supply chain disruptions effectively.

Low or Non-Correlation in Consumer Goods Prices

The correlation analysis also identified several pairs of consumer goods that exhibited low or non-significant correlation. These goods are likely influenced by different market dynamics, production factors, or consumer demand patterns. Here are the pairs with low or non-correlation:

1. **Yogurt and Apple**: The prices of yogurt and apples showed little to non-correlation. These two products come from entirely different sectors—dairy and agriculture—and are influenced by distinct factors. Yogurt prices are affected by dairy production costs, while apple prices are driven by agricultural conditions, harvest yields, and seasonal fluctuations.
2. **Apple and Flour**: Apple prices and flour prices showed a non-significant correlation. Apples are influenced by agricultural conditions, including weather and harvest yields, while flour prices are typically affected by factors related to grain production, global market conditions, and processing costs. The two products' price changes are therefore largely independent of each other.
3. **Apple and Bread**: Similarly, apple prices and bread prices demonstrated little to non-correlation. While both are essential food items, their prices are influenced by different factors—apple prices are driven by seasonal harvests and agricultural conditions, whereas bread prices are more closely tied to grain and flour production costs, as well as transportation.
4. **Sugar and Beef**: The prices of sugar and beef showed a low correlation. Sugar is a processed agricultural product affected by global supply and demand for sweeteners, while beef prices are driven by livestock production, feed costs, and market conditions in the meat industry. These products do not share similar market forces or seasonal patterns.
5. **Beef and Egg**: Beef and eggs also showed a low correlation. Although both are protein-rich food items, their prices are influenced by different supply chains—beef prices are impacted by cattle farming and meat production costs, while egg prices are more closely tied to poultry farming and feed costs.
6. **Beef and Sugar**: There was a non-significant correlation between beef and sugar prices. Beef is primarily affected by livestock production factors, while sugar prices are driven by the global sugar market, import/export policies, and seasonal agricultural factors.

Implications of Low or Non-Correlation Pairs:

The lack of correlation between these goods implies that their price changes are influenced by largely independent factors. For businesses, policymakers, and analysts, this means that changes in the price of one product will not necessarily signal price movements in another product. Understanding these independent price drivers can help in creating more accurate and isolated forecasting models and allows for more tailored strategies in pricing, supply chain management, and policy development.

In conclusion, the correlation analysis further enriches the forecasting model by demonstrating the interconnectedness of consumer goods prices. A high correlation between the prices of specific goods suggests that price changes in one product may lead to similar changes in others, a factor that can be considered when planning for price adjustments or understanding market trends.

3.2. Time series analysis

This study aimed to predict the prices of commonly used consumer goods in Ulaanbaatar between 2012 and 2022 using time series analysis. The data was sourced from the National Statistical Office of Mongolia's 1212.mn website, which provided the price data for 11 common consumer goods, including flour, bread, rice, beef, milk, yogurt, sugar, eggs, apples, potatoes, and A92 gasoline. These price data points from January 2012 to December 2022 were analyzed using time series analysis, and the prices were forecasted for the period from January 2023 to May 2024.

Time series analysis of beef: Trend: Beef prices have steadily increased, reflecting rising production costs, feed prices, and supply-demand imbalances in the market. Seasonal: Beef prices tend to spike during festive seasons and colder months due to increased consumption and demand.

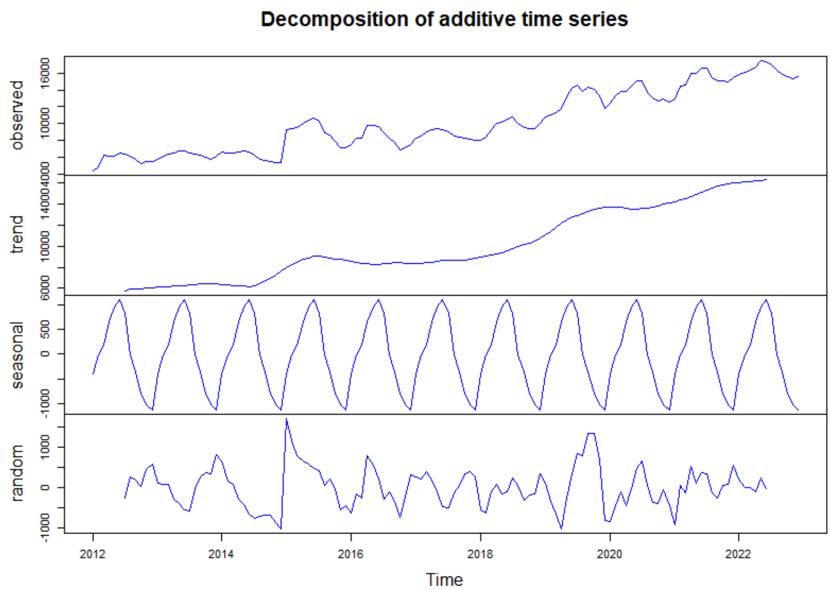

Figure 2: Decomposition of the additive time series of beef

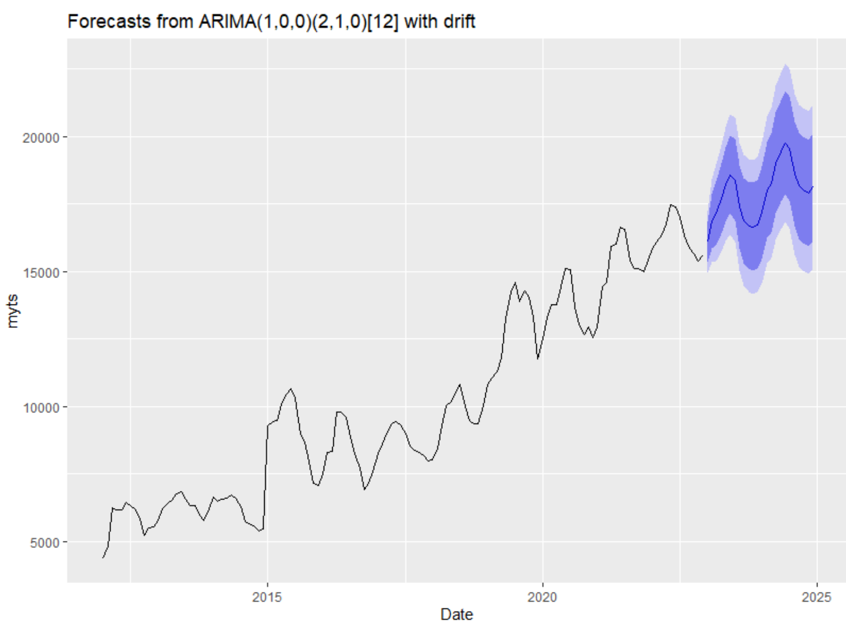

Figure 3: Decomposition of the additive time series of beef

Time series analysis of potatoes: Trend: Potato prices have generally been increasing due to factors like weather conditions, local production challenges, and global market trends. Seasonal: Prices tend to drop after harvest but increase during the winter months when fresh potatoes are less available. Random Movement: Potato prices can experience volatility due to seasonal weather conditions, unexpected supply disruptions, or shifts in domestic production.

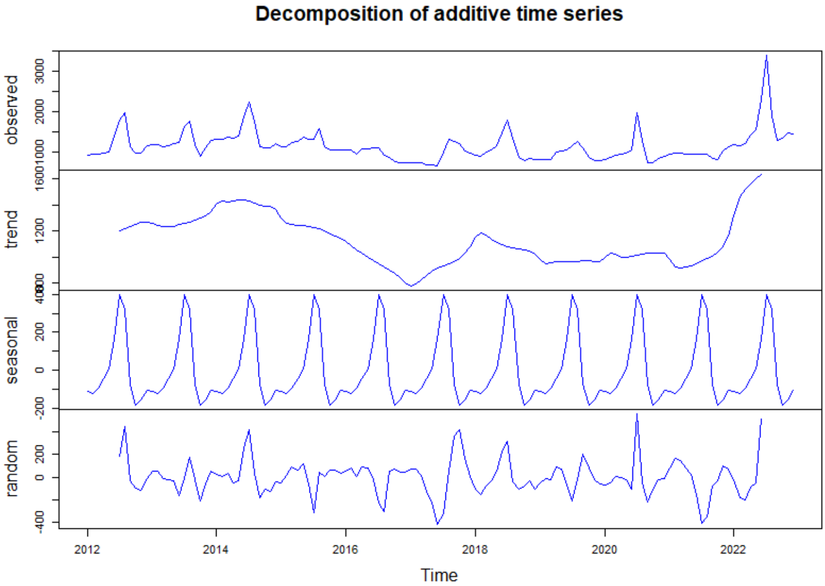

Figure 4: Decomposition of the additive time series of potatoes

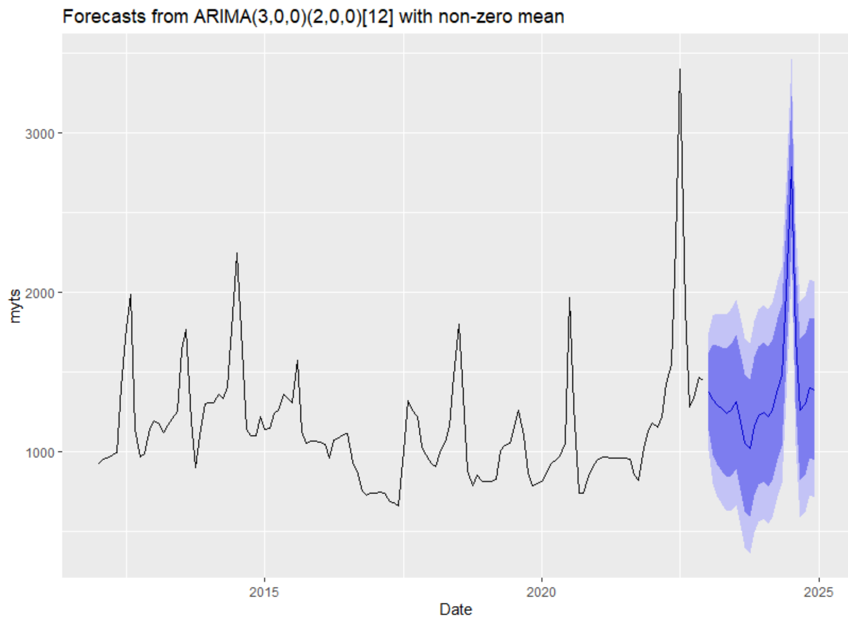

Figure 5: Decomposition of the additive time series of potatoes

The results of this analysis show that time series forecasting provides valuable insights into the price trends of commonly used consumer goods.

The table presents forecasting accuracy metrics for various consumer goods. These metrics include Mean Error (ME), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Percentage Error (MPE), Mean Absolute Percentage Error (MAPE), Mean Absolute Scaled Error (MASE), and Autocorrelation at lag 1 (ACF1), all of which help assess the performance of forecast models.

Key insights from the table include:

Table 1: Forecasting Errors and Accuracy: Consumer Goods

	Consumer Goods	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
1	Flour	-0.0008	0.0167	0.0097	-0.0110	0.1356	1.0814	0.0062
2	Bread	-0.0009	0.0304	0.0135	-0.0125	0.1940	1.4719	0.0009
3	White rice	-0.0008	0.0889	0.0367	-0.0090	0.4530	1.2043	0.0035
4	Beef	-0.0035	0.0598	0.0385	-0.0435	0.4215	0.7925	0.0555
5	Milk	-0.0042	0.0310	0.0209	-0.0556	0.2712	1.0439	0.0389
6	Yogurt	0.0082	0.0422	0.0235	0.1046	0.3025	1.2705	-0.0337
7	Sugar	-0.0031	0.0367	0.0210	-0.0411	0.2771	1.2523	-0.0066
8	Egg	-0.0031	0.0484	0.0300	-0.0545	0.5145	1.0208	-0.0060
9	Apple	0.0002	0.1395	0.0580	0.0007	0.7168	1.4400	-0.0164
10	Potatoes	-0.0026	0.1183	0.0807	-0.0415	1.1522	0.7638	-0.0345
11	A92 gasoline	-0.0004	0.0351	0.0192	-0.0057	0.2575	1.2194	-0.0473

- **Flour and Milk** demonstrate the best forecasting accuracy with low RMSE, MAPE, and MAE values, indicating minimal errors in prediction.
- **Egg and White Rice** show relatively higher errors, especially Egg, which has a higher MAPE (0.5145%) and RMSE (0.0484), suggesting less precise forecasts.
- **Apple and Potatoes** have the highest RMSE and MAPE values, reflecting weaker forecasting performance. Apple, in particular, exhibits the highest MAPE of 0.7168%, indicating a larger forecast error relative to the actual values.
- **A92 Gasoline** stands out as another product with relatively low error metrics, showing a better forecasting performance compared to other goods with low MAPE (0.2575%) and RMSE (0.0351).

Overall, the table highlights which consumer goods have more accurate forecast models, with Flour and Milk showing the best performance and Apple and Potatoes requiring improvements in prediction accuracy.

The correlation between the predicted values from the time series analysis and the actual values obtained from the statistics office website was examined. The time series analysis performed best for predicting the price of premium flour, with a 94.9% correlation between the predicted and actual values for the period from January 1, 2023, to May 1, 2024. Additionally, the predicted prices for bread, beef, and milk were correlated with the actual prices by 50-68%. However, for other commodities such as white rice, yogurt, sugar, eggs, apples, potatoes, and A92 gasoline, the predicted prices were almost unrelated to the actual values, with correlations ranging from 10 to 20%.

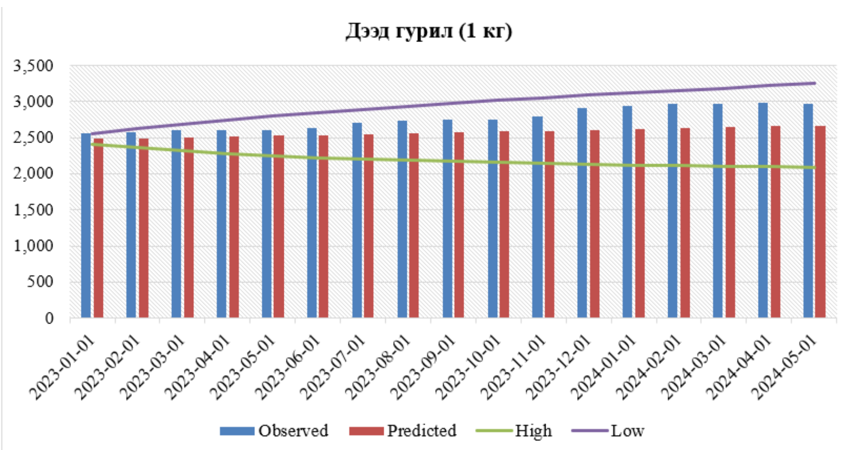

Figure 6: Predicted Price Range for Premium Flour

Among these, the model was most suitable for premium flour, and the forecasted price of premium flour was analyzed (Figure 6). The figure shows the highest and lowest predicted price values derived from the results of the price time series model, with the columns representing the actual and predicted values.

4. Discussion

The primary objective of this study was to analyze the price trends of essential consumer goods in Ulaanbaatar, Mongolia, between 2012 and 2022, and to forecast their future prices using time series analysis. This research used data from the National Statistical Office of Mongolia and employed various time series forecasting methods to examine trends, seasonality, and random movements in prices for 11 essential goods, including flour, bread, rice, beef, milk, yogurt, sugar, eggs, apples, potatoes, and A92 gasoline.

The analysis also uncovered notable seasonal variations, with prices generally falling during harvest seasons for agricultural products and rising during festive or high-demand periods. Random fluctuations, driven by unpredictable events like supply chain disruptions or geopolitical factors, were observed in several products, particularly A92 gasoline.

Correlation analysis provided further insights into the relationships between the prices of different consumer goods. Strong correlations were found between items that are often consumed together, such as eggs and milk, or products within the same category, such as bread and flour. These correlations suggest that price changes in one product can serve as an indicator of similar changes in another, offering valuable information for businesses and policymakers. On the other hand, weak or non correlation was found between products that are influenced by different market forces, like beef and sugar, highlighting the need for independent forecasting models for different sectors.

One key finding was the accuracy of forecasting for certain goods. Flour and milk exhibited the best forecasting performance, with low error values across various metrics, indicating that their prices can be predicted with greater precision. On the other hand, items like apples and potatoes showed higher forecasting errors, suggesting the need for refined models to improve prediction accuracy. This discrepancy in forecasting performance underscores the challenges of predicting prices for agricultural and perishable products, which are often subject to more volatile market conditions.

In terms of implications, the results emphasize the importance of accurate price forecasting for economic stability, policy development, and business planning. By predicting future price trends, policymakers can better anticipate inflationary pressures and take proactive measures to ensure that essential goods remain affordable for consumers. For businesses, forecasting price fluctuations allows for more efficient inventory management, pricing strategies, and cost control. Furthermore, understanding correlations between goods can assist in planning for supply chain disruptions and adjusting pricing strategies in response to changes in the cost of raw materials or production inputs.

5. Conclusions

This study demonstrates the value of time series analysis in forecasting the prices of essential consumer goods in Ulaanbaatar, Mongolia. The analysis revealed important trends, seasonal patterns, and random movements that affect the prices of goods such as flour, bread, beef, and milk. Correlation analysis further enhanced the study by highlighting the relationships between different goods, which can aid in more informed forecasting and policy-making.

The forecasting results indicate that while some goods, like flour and milk, can be predicted with relatively high accuracy, others, such as apples and potatoes, pose more challenges due to higher levels of volatility and weaker model performance. This suggests that further refinement of forecasting models, including the incorporation of more sophisticated techniques and external factors, is necessary to improve the accuracy of price predictions for certain goods.

Overall, this research underscores the importance of accurate price forecasting in ensuring economic stability, guiding policy decisions, and supporting business operations. Future research can expand on these findings by exploring the effects of external factors, such as climate change or global trade disruptions, on the prices of essential goods and incorporating more advanced machine learning models for improved prediction accuracy.

References

[1] D. Batbold and O. Lkhagvasuren, “Economic factors influencing the price fluctuations of basic goods in Mongolia,” *Mongolian Economic Journal*, Vol. 6, no. 1, pp. 45–59, 2021.

[2] S. Bovas, “Forecasting time series with MAPE and MPE,” *Journal of Time Series Analysis*, Vol. 36, no. 2, pp. 100–112, 2015, https://doi.org/10.1016/j.intfin.2014.12.008 .

[3] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-Day, 1970.

[4] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, Hoboken, NJ: Wiley, 2015.

[5] E. Chong, C. Han, and F. C. Park, *Predictive Models and Time Series Analysis*, Singapore: Springer, 2017.

[6] R. B. Cleveland, W. S. Cleveland, J. E. McRae, and I. Terpenning, “STL: A seasonal-trend decomposition procedure based on Loess,” *Journal of Official Statistics*, Vol. 6, no. 1, pp. 3–73, 1990.

[7] C. C. Holt, “Forecasting seasonal and trend data,” *Journal of the American Statistical Association*, Vol. 52, no. 318, pp. 253–270, 1957, https://doi.org/10.1016/S0140-6736(57)90613-X .

[8] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, Melbourne: OTexts, 2018.

[9] R. J. Hyndman *et al.*, *Forecasting with Exponential Smoothing: The State Space Approach*, Berlin: Springer-Verlag, 2008.

[10] S. Khan, I. Awan, and S. Zafar, “Forecasting of agricultural commodity prices using time series modeling,” *International Journal of Agriculture and Crop Sciences*, Vol. 12, no. 3, pp. 110–115, 2020.

[11] S. Makridakis *et al.*, “M4 Competition: Results, findings, conclusion,” *International Journal of Forecasting*, Vol. 34, no. 4, pp. 802–808, 2018, https://doi.org/10.1016/j.ijforecast.2018.06.001 .

[12] C. J. Willmott and K. Matsuura, “Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance,” *Climate Research*, Vol. 30, pp. 79–82, 2005, https://doi.org/10.2307/40158798 .