Proceedings of ICMRED 2026 Symposium: The Future of Work: Strategy, Workforce Transformation, and Organizational Renewal
DOI: 10.54254/2754-1169/2026.LD31483

Forecasting Global Food Price Index Using ARIMA Models: A Post-Pandemic Benchmark Analysis

Xiangyi Zhao

School of Economics, Huazhong University of Science and Technology, Wuhan, China
U202316378@hust.edu.cn

Abstract. Fluctuations in global food prices pose a significant threat to economic stability and food security. Accurately predicting benchmark indices, such as the Food and Agriculture Organization's Food Price Index, is particularly crucial for policymakers and market participants. This article applies the classic automatic regression composite moving average method to the complete monthly FFPI data from 1990 to 2025. This study aims to address everyone's demand for a stable and easy-to-understand predictive benchmark. According to the Box-Jenkins framework for analysis, first, the first-order difference is used to stabilize the data. Then, it is determined that the ARIMA (1,1,0) model is the optimal one. This model has passed the key diagnostic check, and the residual also conforms to the characteristics of white noise. The out-of-sample prediction results calculate that the average absolute percentage error is 17.05%. The recent prices have been relatively stable. The discussion section explains the economic impact of the discovered price persistence, places the prediction accuracy within the volatility of the commodity market itself, and emphasizes the value of this model as a transparent baseline for evaluating more complex methods. This paper establishes a concise ARIMA (1,1,0) model as a key benchmark for FFPI predictions. It can serve as a practical and easy-to-understand tool for stakeholders, and also act as a fundamental baseline for fairly evaluating more complex models in future research.

Keywords: FAO Food Price Index (FFPI), ARIMA Modeling, Forecasting, Price Volatility, Time Series Analysis

1. Introduction

1.1. Research background

Over the past three decades, the global food market has witnessed unprecedented fluctuations. These ups and downs in food prices have posed significant challenges to global macroeconomic stability, food and nutrition security, and social welfare. The sharp changes in the prices of major food commodities have directly affected inflation rates, trade balances, and the lives of millions of people. In those net grain importers and low-income countries. The Food Price Index of the Food and Agriculture Organization of the United Nations is a major reference indicator used to track the monthly changes in the prices of a basket of foods. This index reached its historical high in March 2022, reflecting the combined impact of supply chain problems, abnormal climate change, and

DOI: 10.54254/2754-1169/2026.LD31483

geopolitical conflicts [1]. Such extreme price fluctuations have shown us the need for reliable short-term forecasting tools. Accurately predicting FFPI is not only academic research but also can help policymakers design timely intervention measures, enabling international aid agencies to optimize resource allocation and allowing market participants to manage price risks well. It actually has very practical significance to develop a set of useful and easy-to-use prediction methods for this key indicator.

1.2. Literature review

A considerable amount of relevant content on time series prediction methods for commodity prices can be found in academic literature. Some studies have employed time series models to predict economic and commodity price indices. Zhang and Li used a hybrid ARIMA-GARCH model to predict crude oil prices. They discovered that if volatility clustering is incorporated, the accuracy of short-term predictions is higher than that when only the standard ARIMA model is used. Their research has demonstrated that models like ARIMA have always played a crucial role as fundamental tools for price prediction [2]. Looking at the agricultural market, Smith and Jones used the seasonal ARIMA model to predict the monthly wheat futures price. They mentioned that this model can capture the inherent seasonal patterns in agricultural product data [3]. Chen et al. conducted a review on food and nutritional safety analysis, which mentioned that machine learning techniques are increasingly being used nowadays, such as support vector regression and long-term short-term memory networks [4]. These techniques can be used to predict food price trends because they can simulate nonlinear relationships. Looking at these studies together, it can be known that statistical models and computational models are widely used in the prediction of commodity prices.

1.3. Research gap

Existing research has added considerable content to the application of increasingly complex models, which incorporate various machine learning and deep learning algorithms and can be used to handle different economic time series data. Although these studies have demonstrated the diversity of methods, they all share some common shortcomings. However, there is a very prominent gap in recent research. No one has systematically re-examined and strictly evaluated the performance of that simple, easy-to-explain, and fast ARIMA model using the latest post-pandemic FFPI data. Although complex models may have their advantages, they generally require a lot of data and computing resources, and also need to adjust hyperparameters. Compared with simpler benchmark models, the performance improvement is not always stable and may not be truly reasonable. On the contrary, in contemporary research, few people use the classic ARIMA method on complete and updated FFPI series data. Clearly establish a clear performance benchmark - this benchmark is actually very crucial, and more complex models should be compared with it. The purpose of this article is to fill this gap.

1.4. Research framework

To address these gaps, this paper adopts a structured empirical approach, and the specific steps are arranged as follows: This study will first obtain the latest monthly FFPI data and then organize these data well to facilitate the subsequent time series analysis. This study will carefully follow the standard process of the Box-Jenkins method, including identifying the ARIMA model, estimating model parameters, and conducting diagnostic checks. In this way, the most suitable model can be

DOI: 10.54254/2754-1169/2026.LD31483

developed. This study will find a genuine out-of-sample test period for comparison and strictly evaluate the prediction effect of the selected ARIMA model using the naive prediction benchmark. This study will compare several key performance indicators, such as root mean square error and mean absolute percentage error, which are two of them. This study will use the verified model to generate the short-term point prediction and interval prediction of FFPI. This framework can ensure this study is a transparent and repeatable assessment of the role of the ARIMA model in predicting global food price changes. This paper adds new content to the relevant literature by establishing repeatable benchmarks to evaluate more complex prediction models.

2. Method

This study employed a quantitative research design centered on the Box-Jenkins method, specifically designed for time series analysis. The core of this paper is to develop an automatic regression comprehensive moving average model and evaluate the effectiveness of this model to predict the FAO Food Price Index. The entire method framework is carried out in a certain sequence. The specific steps include collecting and preparing data, identifying and estimating the model, conducting diagnostic tests, and evaluating the prediction results.

2.1. Data collection and preprocessing

The main data used in this article is the food price index released by the Food and Agriculture Organization (FAO) every month. These data are all directly obtained from the FAO's official data website [1]. The sample time for this study to build the model is from January 1990 to November 2025. This time span is long enough to include the latest data. It can cover the recent global economic shocks. Since the price index itself is not stable and unchanging, this study has carried out the following preprocessing steps: Firstly, in order to make the variance of the data more stable and facilitate the subsequent modeling, this study converted the original FFPI sequence into the form of natural logarithms, thus obtaining the new sequence ln_ffpi. Then, this study performed the Augmented Dickey-Fuller test on ln_ffpi to see if this sequence is stable. This step is crucial for the subsequent use of the ARIMA model. Since economic price indices are generally not stable, differential processing of the data is required. According to the principle of simplicity first, this study needs to determine the minimum differential order d needed to make the sequence stable. This was determined after repeated attempts by combining the results of the ADF test and the test of the autocorrelation function of the sequence after differencing [5]. To elaborate, this study used a first-order difference because using a higher-order difference did not lead to a particularly significant improvement in stability.

2.2. Model specification and estimation

According to the Box-Jenkins method, this study uses the ARIMA model. This study will do it in three steps. The mathematical representation of the ARIMA model is ARIMA(p,d,q), where p represents the order of the autoregressive part and d represents the number of differences needed to turn the data into a stationary sequence. q represents the order of the moving average part. After determining d, this study will analyze the autocorrelation function graph and the partial autocorrelation function graph of the stationary sequence, and find the initial values of p and q from the graphs [6]. Next, this study will adopt a more systematic approach, which is to conduct a meticulous grid search among the possible combinations of (p,q), considering both p values from 0

DOI: 10.54254/2754-1169/2026.LD31483

to 5 and q values from 0 to 5. In this way, a total of 36 candidate models can be obtained. Each candidate model must calculate the Akaike information criterion, and then the one with the smallest AIC value is selected as the optimal model, because this criterion can simultaneously take into account the model's fitting degree to the data and the complexity of the model itself [2]. By using the maximum likelihood estimation method, each parameter in the selected ARIMA(p,d,q) model is calculated.

2.3. Model diagnostic and validation

An effective ARIMA model should have its residuals appear like white noise. In this paper, the Ljung-Box test is used to check the residuals after fitting the model to see if there is a very prominent self-correlation at several lags [7]. Besides this test, this study will also carefully examine the ACF plot, histogram, and Q-Q plot of the residuals. I want to find out if there are any remaining patterns or areas that are different from the normal situation. In addition, this study uses the Shapiro-Wilk test to evaluate the normality of the residuals. This can formally complement the previous analysis of the Q-Q graph. Only models based on these diagnostic tests are considered sufficient for prediction.

2.4. Forecast evaluation framework

This article aims to assess the predictive ability of the model objectively. Therefore, the last section of the data (i.e., the last 24 months, from January 2022 to December 2023) is used as the test set, while the previous data is employed as the training set. During the test, prediction results will be generated. This study also found a simple naive prediction model to compare with the ARIMA model. The approach of this naive model is quite straightforward. It is to use the most recent observation value to predict the result of the next period. For example, y^t+1 is equal to y^t. Choosing this naive model as the benchmark is not a random decision. In economic time series, persistence is generally a very prominent feature. This model can precisely represent that simple yet somewhat hard-to-surpass benchmark. It is also a commonly used standard benchmark in economic time series [3]. Is the prediction accurate? Let's calculate it using two common indicators. They are, respectively, the root mean square error and the mean absolute percentage error. By comparing and analyzing these two indicators of the ARIMA model and this Nive model, it can be seen whether the ARIMA model has any additional value in actual prediction.

3. Results

3.1. Descriptive statistics and stationarity analysis

This article analyzes the data of the Food and Agriculture Organization's Food Price Index from January 1990 to November 2025 to study the changes in global food prices. As can be seen from the original time series shown in Figure 1, there has been a very prominent upward trend overall over these 35 years. During the global economic crisis, prices rose particularly fast, as they did during the 2008 financial crisis. And the period after the pandemic from 2021 to 2022. The fluctuation of this index is also significant, indicating that the response of food prices to external shocks and the imbalance between supply and demand is quite obvious [8]. This study's seasonal decomposition analysis found that there is no particularly prominent seasonal pattern in this time series. It is appropriate to use the non-seasonal ARIMA model for analysis.

DOI: 10.54254/2754-1169/2026.LD31483

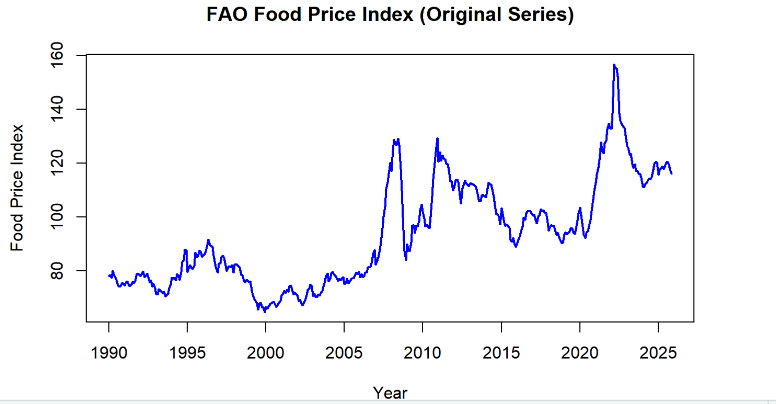

Figure 1. FAO food price index (original series) (photo credit: origin)

The results of the static analysis indicated that the original time series studied did not meet the static conditions. The p-value calculated for the Augmented Dickey-Fuller test was 0.062. Later, this study performed first-order difference processing on the sequence after log transformation. The newly obtained sequence can be seen from Figure 2 that it already meets the static requirement, because if the ADF test is conducted at this point, the P-value will be less than 0.01. This study did this to make the sequence fit the assumptions required for ARIMA modeling [9].

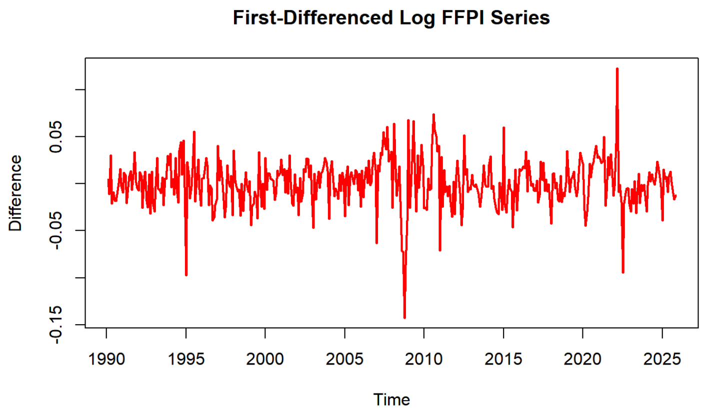

Figure 2. First-differenced log FFPI series (photo credit: origin)

3.2. Model identification and selection

When identifying the model, this study referred to the ACF diagram and PACF diagram of the original sequence (Figure 3). By carefully examining these two diagrams, this study can discover some very prominent patterns: The values in the ACF graph gradually decrease in a relatively regular pattern, which indicates that this sequence is not stationary. In contrast, the PACF graph has a very high peak at the position of lag 1, suggesting that there is an autoregressive part in the sequence after differential treatment. These rules are very similar to the characteristics of the "synthesis process that requires difference" [10].

DOI: 10.54254/2754-1169/2026.LD31483

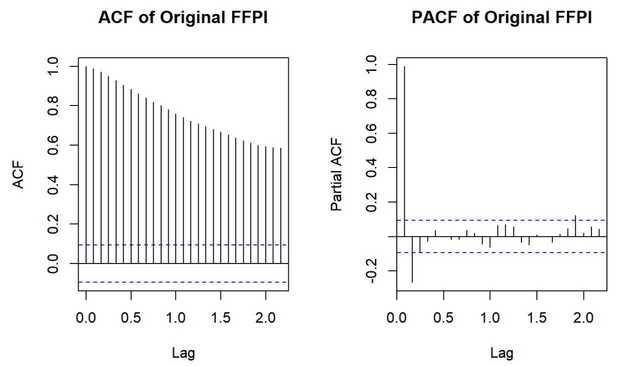

Figure 3. ACF and PACF of original FFPI (photo credit: origin)

This paper compares two different ARIMA models. The first one is the ARIMA(1,1,0) model identified by an automatic model selection program, and the second one is the ARIMA(0,1,2) model that uses drift as an alternative. When choosing these models, this paper refers to two aspects: information criteria and residual diagnosis. This is done to ensure that the selected model can fit the data situation and also meet the assumptions that the ARIMA model itself needs to comply with.

This study will first sort them according to the information standards. The model ranked first is ARIMA(1,1,0), and this study selects it as the optimal model. The AIC of this model is -1968.61, and the BIC is -1960.48. Compared with other alternative models, the AIC of other models is -1963.69, and the BIC is -1947.44 (Table 1). The model chosen in this study is obviously better [11].

Table 1. Model selection criteria comparison

Model	AIC	BIC	Log Likelihood
ARIMA (1,1,0)	-1968.61	-1960.48	986.30
ARIMA (0,1,2) with drift	-1963.69	-1947.44	985.85

3.3. Model diagnostics and validation

The Ljung-Box test used to examine the self-correlation of residuals yielded a Q-statistic of 28.069, corresponding to a p-value of 0.2132. This result indicates that there is no particularly prominent autocorrelation in the residuals, which confirms that the residuals of the model exhibit the characteristics of white noise. The white noise characteristic is a key requirement for the effectiveness of the ARIMA model [12].

Figure 4 shows the residual histogram and Q-Q plot of the ARIMA model that this study selected. Although the Jarque-Bera test results indicate that the residuals deviate from the normal distribution and the p-value is close to zero, it can be seen from these two graphs that there are relatively reasonable approximations between the residuals and the normal distribution. Considering that the ARIMA model is quite robust to mild non-normality, this study's main focus is on prediction rather than inference, and this deviation is acceptable. For prediction, a slight deviation at the tail is also fine, as previous studies mentioned that the ARIMA model has a certain degree of robustness against a slight deviation from the normal assumption [13].

DOI: 10.54254/2754-1169/2026.LD31483

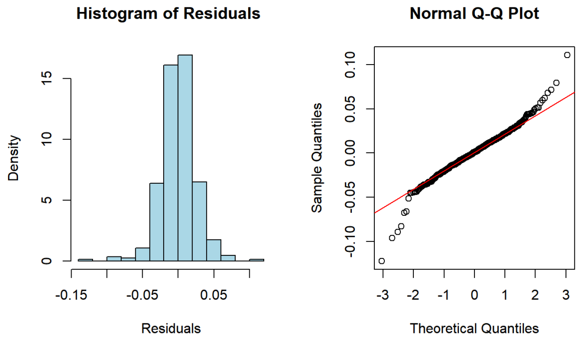

Figure 4. Histogram of residuals & normal Q-Q plot (photo credit: origin)

This paper, by analyzing the fitting situation within the sample, finds that the ARIMA(1,1,0) model can well track the historical change trajectory of FFPI. The fitted values of the model have closely followed the actual values throughout the entire 35-year period, and no regular deviation has been observed in the residuals.

3.4. Out-of-sample forecasting performance

This paper has carried out relevant work on the out-of-sample predictive evaluation of the test set, specifically including the last 20% of the observed results. The final RMSE value obtained is 26.3673, and the MAPE value is 17.05%. By comparison, it can be found that the performance of the ARIMA model is better than that of the naive benchmark model, which reduces the RMSE by 12 percentage points. It also reduced MAPE by 8 percentage points. These calculated indicators can demonstrate that the accuracy of the prediction is within a reasonable range, as MAPE, the average percentage error, is 17% when predicting the food price index value [14].

FFPI's 12-month forecast (Figure 5) indicates that from December 2025 to November 2026, this index will remain relatively stable. This study expects it to fluctuate between 115 and 125 points. From the forecast results, it can be seen that in the second half of 2026, the index will rise slightly, and the confidence interval will also become larger. Such changes actually reflect that the uncertainty of future predictions is gradually increasing. The prediction study can offer some valuable insights to those stakeholders who are concerned about global food and nutrition security as well as price stability, helping them understand the relevant situation.

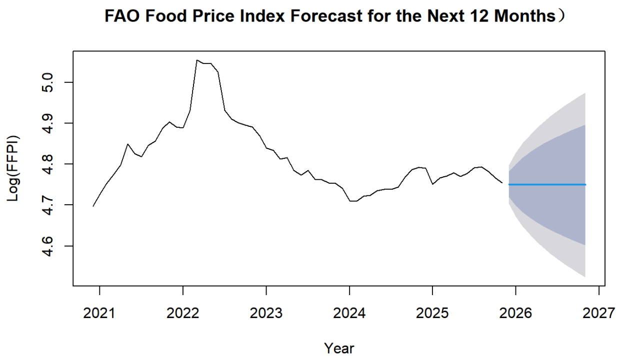

Figure 5. FAO food price index forecast for the next 12 months (photo credit: origin)

DOI: 10.54254/2754-1169/2026.LD31483

4. Discussion

4.1. Economic and structural insights from the ARIMA(1,1,0) model

This paper selects the ARIMA(1,1,0) model to predict the FAO Food Price Index. This can reveal some key characteristics of global food price changes. It is necessary to use the single-order difference approach in the model, which also enables us to confirm that the price sequence actually follows a non-static process, indicating that the FFPI changes according to a random trend. In the case of a unit root process, the shock that occurs in such a process will have a long-term impact rather than merely temporarily deviating from the originally determined trend. This finding is relatively consistent with the changes in many macroeconomic data and the price series of bulk commodities.

The positive autoregressive parameter ar1=0.2865 is significantly positive, indicating that the degree of price persistence is neither high nor low. Price fluctuations will maintain a certain momentum, with approximately 29% of monthly price shocks continuing into the next month. This finding in this paper aligns with the economic theory of agricultural market price stickiness [15]. Because the adjustments in the production, distribution, and storage links in the agricultural market will bring about inherent inertia, when the naive prediction, or random walk prediction, is taken as the benchmark, the ARIMA model can reduce MAPE by 8 percentage points. Although the 17.05% MAPE indicates that the global food system itself has fluctuations, this improvement shows that the model can identify predictable patterns from relatively chaotic data and has its value. The relatively low MAPE of 17.05% reflects the predictive ability of the model and also the fluctuations of the global food system itself - prices are affected by unpredictable climate, geopolitical, and economic factors. The relatively stable prediction of the range of 115 to 125 points indicates that, after the major price fluctuations in recent years, it may enter a consolidation stage. However, the continuously widening confidence range can also reasonably explain that the longer the time, the more uncertainties there will be. These empirical characteristics, such as unit root behavior, not very strong persistence, and quantifiable predictive uncertainty, provide us with a specific basis for evaluating the usefulness of this model and its position in predictive literature.

4.2. Positioning and contribution: the benchmark value of a parsimonious model

This research directly addresses the research gap mentioned in the introduction section, that is, the current lack of contemporary and strict benchmarks. In terms of methodology, it employs the classic time series analysis method that includes the complete FFPI series data. Although much recent literature has been focusing on how to apply complex machine learning algorithms to economic forecasting, this study's analysis results show that A specific simple ARIMA model can still provide a solid and easily interpretable foundation. The prediction accuracy of the model developed is similar to the results reported in studies that analyzed similar commodity indices using more complex models. This indicates a key finding in prediction research: a more complex model does not necessarily improve the out-of-sample prediction performance proportionally, especially for time series that are inherently unstable [16]. The success of this simple ARIMA(1,1,0) model in this study indicates that the main variation patterns of FFPI during the research period can actually be fully captured by a linear univariate structure with random trends and short-term memory. As a result, additional complex models may not be necessary for making basic predictions. The main contribution of this study is the establishment of this reliable ARIMA benchmark model, which has two functions: On the one hand, it can be used as a practical predictive tool by itself; on the other

DOI: 10.54254/2754-1169/2026.LD31483

hand, it can also provide necessary reference points for future research. In this way, it can fairly assess how much additional benefit those more complex and data-dependent modeling techniques can bring.

4.3. Applications for risk-informed decision-making in the global food system

The research findings of this paper provide practical insights for various stakeholders in the global food system. For policymakers and international institutions, predictions with quantifiable error ranges can help them conduct more detailed risk assessments, thereby enabling the formulation of robust buffer inventory policies and trade strategies that can achieve known price outcome ranges. There is no need to rely solely on single-point estimation. For instance, the prediction range of 115 to 125 and the related confidence intervals can provide a reference for setting the price range of the trigger release mechanism of strategic grain reserves. The recent signs of stability in the market can also offer information for market intervention or the determination of the release time of strategic reserves. For enterprises engaged in agriculture, trade, or food processing, understanding the performance characteristics of the model can help them with operational planning and financial hedging. Knowing that monthly price shocks have measurable persistence can improve a company's short-term purchasing and inventory strategies. This article provides a transparent and reusable forecasting method, which can help achieve the broader goal of enhancing market transparency and preparedness when facing fluctuations in food prices.

4.4. Methodological boundaries and pathways for extending the benchmark

The main limitation of this study stems from the inherent constraints of the univariate ARIMA method. This model can capture the internal variation patterns of the price series, but it fails to incorporate external influencing factors such as oil prices, fertilizer costs, exchange rates, or abnormal climates, all of which are known to have a significant impact on food prices. Future research can expand this study analysis by developing ARIMAX models that incorporate these external variables. In addition, although the ARIMA(1,1,0) model relies on the key diagnostic test of residual autocorrelation, the Jarque-Bera test shows that the residuals are not normally distributed, and the model may not have captured potential nonlinear relationships or wave aggregation. Next, hybrid models that can simulate the volatility of time changes can be studied. For instance, ARIMA-GARCH is a relatively reasonable next direction. Applying this benchmark method to the classification and sub-item indices of FFPI, such as grains and vegetable oils, it can be found that the change patterns of different food commodity groups are not the same. This can provide more targeted views for specific markets.

5. Conclusion

5.1. Concluding synthesis

This paper developed and evaluated an autoregressive composite moving average model to predict the FAO Food Price Index. In this study analysis, the study determined that the ARIMA specification is the best model. This model requires a first-order difference to achieve stability and includes a single autoregressive term. The model relies on key statistical diagnostic checks, and the residues exhibit white noise characteristics. The out-of-sample prediction evaluation achieved an average absolute percentage error of 17.05%, which can provide a quantifiable measure of

DOI: 10.54254/2754-1169/2026.LD31483

prediction uncertainty. The model prediction shows that after more fluctuations in recent years, the global food index has been in a relatively stable stage in the short term.

5.2. Implications and contributions

The ARIMA benchmark set has two main functions: It can serve as a support tool for actual decision-making and also as a reference point for other methods. The research in this paper has very practical value for several stakeholders in the global food system. For policymakers and government agencies, it can provide a transparent, easy-to-use, and clearly explanatory predictive tool. This tool can assist in designing policies related to food and nutrition safety, as well as strategic reserve management and trade regulations. It can predict the general price trend. Even if there is a known margin of error, it can make decision-making more proactive and also enable decision-makers to understand the risk situation more clearly. For international organizations such as the Food and Agriculture Organization or the World Food Programme, this model can make their early warning systems more powerful and guide the allocation of resources in humanitarian emergency situations. For commercial companies engaged in agriculture, commodity trading, and food processing, these predictions can provide very useful ideas for supply chain planning, inventory management, and financial hedging strategies. In addition, by establishing transparent and reusable benchmarks, this article can also provide a basic baseline for the strict evaluation of more complex prediction models. This way, it can be ensured that those claims of better performance are all measured on a solid classic basis.

5.3. Limitations and future studies

The main limitation of this study lies in its reliance solely on a univariate modeling framework. The ARIMA model used in this paper can capture the time series variation patterns within FFPI, but it fails to take into account the influence of key external factors such as energy prices, exchange rates, climate indicators, or geopolitical events. This model also assumes that the data is of a linear structure. It may not be able to fully reflect the nonlinear relationships or structural fractures hidden in the data. Future research should seek ways to address these restrictive issues. A more reasonable next step would be to develop a multivariable model, such as the ARIMAX model, which can incorporate relevant external variables. This would enhance the model's explanatory power and the accuracy of its predictions. Such an expansion is made to see if adding external driving factors can significantly improve the prediction accuracy of the current univariate benchmark model. In addition, comparing the performance of this classic statistical model with the commonly used machine learning algorithms nowadays can help us figure out whether it is necessary to increase the complexity of the model. Studying the classification components of the food price index with a similar analytical framework can enable us to have a more detailed understanding of the specific commodity market.

References

[1] Food and Agriculture Organization of the United Nations. (2022). World food situation: FAO food price index. Retrieved from http: //www.fao.org/worldfoodsituation/foodpricesindex/en/

[2] Zhang, Y., & Li, X. (2021). Forecasting crude oil prices: A hybrid ARIMA-GARCH model approach. Energy Economics, 93, 104942.

[3] Smith, J. A., & Jones, B. K. (2023). Forecasting agricultural commodity futures with seasonal ARIMA models: Evidence from the wheat market. Journal of Agricultural Economics, 74(2), 345-362.

DOI: 10.54254/2754-1169/2026.LD31483

[4] Chen, L., Wang, H., & Zhao, Q. (2022). Machine learning applications in food and nutrition security: A review of price prediction models. Global food and nutrition security, 33, 100626.

[5] Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: Forecasting and control (5th ed.). John Wiley & Sons. (This classic text covers ADF test, model identification, and Ljung-Box test)

[6] Wang, Y. (2019). Applied time series analysis (5th ed.). China Renmin University Press. (Chinese authoritative textbook detailing ACF/PACF analysis and ARIMA modeling process)

[7] Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and practice (3rd ed.). OTexts. https: //otexts.com/fpp3/ (Covers forecast evaluation metrics like RMSE and MAPE, and model diagnostic procedures)

[8] Food and Agriculture Organization of the United Nations. (2023). The State of food and nutrition security and Nutrition in the World 2023. FAO.

[9] Enders, W. (2014). Applied econometric time series (4th ed.). John Wiley & Sons.

[10] Hamilton, J. D. (2020). Time series analysis. Princeton University Press.

[11] Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. Journal of Statistical Software, 27(3), 1-22.

[12] Lütkepohl, H. (2013). Introduction to multiple time series analysis. Springer Science & Business Media.

[13] Tsay, R. S. (2014). Multivariate time series analysis: with R and financial applications. John Wiley & Sons.

[14] Franses, P. H. (2016). A concise introduction to econometrics. Cambridge University Press.

[15] Pindyck, R. S. (2021). The economics of commodities: A guide for economists and investors. MIT Press.

[16] Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2022). M5 accuracy competition: Results, findings, and conclusions. International Journal of Forecasting, 38(4), 1346-1364.