International Journal of Statistics and Applied Mathematics 2024; 9(3): 40-47

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2024; 9(3): 40-47

https://www.mathsjournal.com
Received: 12-03-2024
Accepted: 16-04-2024

Phung Duy Quang
Faculty of Technology and Data Science, Foreign Trade University, Hanoi, Vietnam

Pham Ngoc Mai
Faculty of Technology and Data Science, Foreign Trade University, Hanoi, Vietnam

Hoang Nam Quyen
Student K61- School of Economics and International Business, Foreign Trade University, Hanoi, Vietnam

Corresponding Author:
Phung Duy Quang
Faculty of Technology and Data Science,Foreign Trade University, Hanoi, Vietnam

Application of ARIMA model in forecasting Vietnam's black pepper export price to the U.S. market

Phung Duy Quang, Pham Ngoc Mai and Hoang Nam Quyen

DOI: https://doi.org/10.22271/maths.2024.v9.i3a.1729

Abstract
Research into the application of ARIMA model to analyze and forecast the export price of Vietnamese black pepper to the U.S. market for the period from March to August 2024, based on monthly black pepper price data from January 2019 to February 2024, aiming to assist producers and exporters in developing appropriate export strategies and to support policymakers in making sustainable development strategies for the Vietnamese black pepper industry.

Keywords: ARIMA, time series, forecast, black pepper price, data analysis

Introduction

Pepper is Vietnam's main perennial crop, totaling over 131,8 thousand hectares of cultivated land, 95% of which is situated mainly throughout the Southeast and Central Vietnam. With high economic value, pepper is Vietnam's main agricultural exports, reaching 267 thousand tons, worth 912 million USD in 2023 (Nguyễn Hạnh, 2024) [2]. At present, Vietnam has exported black pepper to over 105 countries and territories (General Statistic Office of Vietnam, 2021), with U.S.A, China, Germany, Netherlands, India and Saudi Arabia having been Vietnam's main and biggest markets. In which, exports to the U.S. In which, exports to the US have always accounted for about 20% of Vietnam's total black pepper export volume to the world. Furthermore, Vietnam is able to utilize opportunities from global trade agreements such as CPTPP and EVFTA, which promote exports to new potential markets (Nông Hữu Tùng and Phạm Công Toản, 2022) [3].

Vietnam is the world's largest producer and exporter of black pepper, accounting for 40% of global production and 60% of market share (followed by Brazil and Indonesia).However, Vietnam Import-Export Report by the Ministry of Industry and Trade (Ministry of Industry and Trade, 2023) indicates that pepper exports are facing a sharp decline in output. More specifically, in 2022, Vietnam's black pepper output decrease by 12.4% compared to 2021, to 228,7 thousand tons, but the average export value recorded a growth of 3.5% compared to the previous year. In 2022, the U.S. remained the largest export market for Vietnamese pepper, with an export volume of 57,809 tons, down 2.5% from the previous year. While China, Vietnam's pepper second largest export market, recorded the sharpest decline, with a decrease of 46.4% (equivalent to 17,761 tons), down to 20,498 tons compared to 2021. And in several major export markets, black pepper export turnover had been declining, including: Pakistan (-45.4%); France (-23.6%); Egypt (-45.4%); Turkey (-26.1%); and Saudi Arabia (-83.1%). Additionally, Vietnam has been losing black pepper market share to Brazil and Indonesia that have much more competitive pricing, a lower export price as these markets apply a lower shipping rates compared to Vietnam. At the same time, political volatility, especially the tensions between Russia and Ukraine, are continuing to affect pepper consumption demand. Furthermore, geopolitical tensions are driving up inflation, leading to a global trend of consumer tightened spending. These are all negatively impacting import and export of many goods, including that of pepper. Therefore, global pepper prices will continue to fluctuate in the coming period as pepper demand continues to rise and fall erratically (Phan Thị Xuân Huệ, 2023) [4]. Consequently, forecasting Vietnam's black pepper price to the U.S. market is of great importance for producers, exporters throughout the country, and to the Vietnamese pepper industry as a whole.

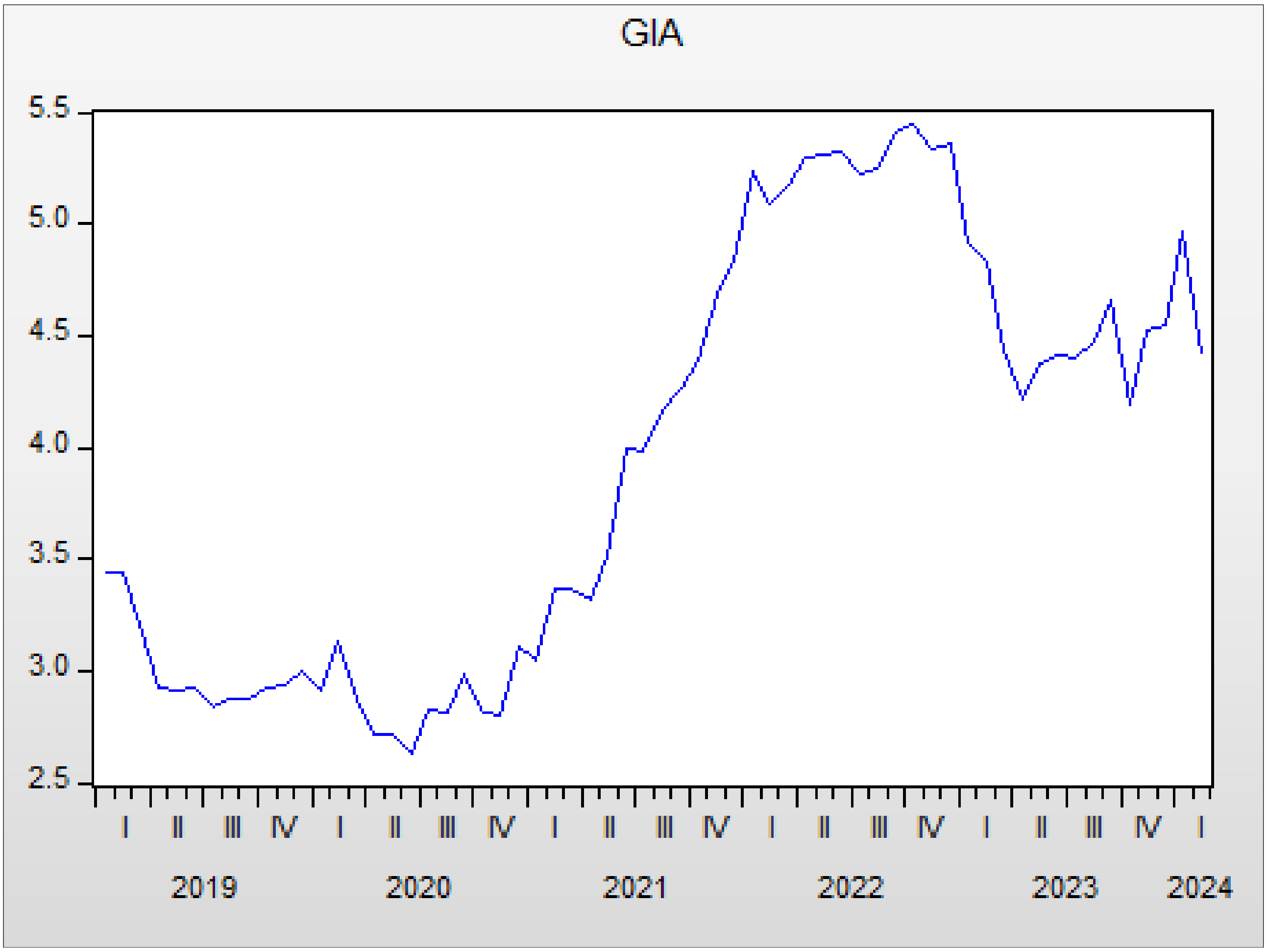
ARIMA (Autoregressive Integrated Moving Average) model is widely known and widely used in the field of forecasting time series for its outstanding strengths. This model, pioneered by Box and Jenkins (1970), include Autoregressive model (AR), Moving-average model (MA) and Autoregressive Integrated Moving Average model (ARIMA). The application of this model became successful in multiple fields, such as in economic, agriculture, climatology, natural science and technology. For instance, the study done by Agus Dwi Nugroho and Imade Yoga Prasada (2020), applying ARIMA model in analyzing data of Indonesian pepper export to Italy from 1989 to 2018 to forecast pepper export volume by 2030. Research done by Weerasinghe W.P.M.C.N. and Jayasundara D.D.M. (2021) build a decomposition model and autoregressive integrated moving average model with season effect (SARIMA) that used data of monthly pepper export volume in Sri Lanka from January 2000 to December 2018 to forecast future outcomes. A study by Naha Hanifah Putri and partners (2023) use ARIMA model with statistics of Lampung's pepper export to global market from 2002 to 2022 to forecast the value for 2023 to 2033. In Vietnam, there are many studies applying ARIMA model in forecasting but there has not been a study about its uses in forecasting exported Vietnamese black pepper price to the U.S. market.

The first part of the article has provided an overview of time series in forecasting domestically and internationally, highlighting the urgency of researching export forecasts for pepper from Vietnam. The remaining sections of the article are structured as follows: Section 2. Data and Research Methods, Section 3. Estimation Results and Discussion, Section 4. Conclusion.

Data and research methodology

Data

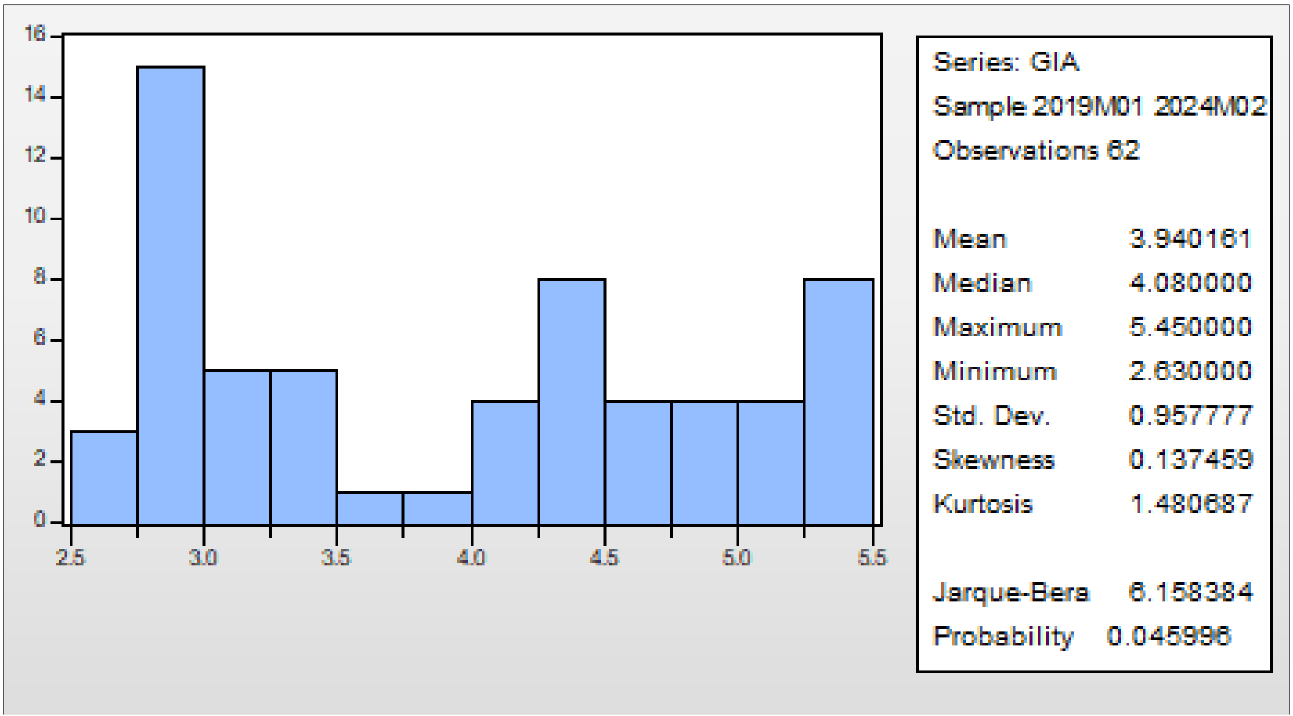
The research is based on secondary data on black pepper export volume collected from reports by the General Department of Customs of Vietnam and the Ministry of Industry and Trade. The data is synthesized by month and processed on Excel and Eviews 12 software. The data range is from January 2019 to February 2024.



Source: Calculated with Eviews 12 by authors with statistics

Fig 1: Changes to Vietnam's black pepper export price to the U.S. market from January 2019 to February 2024

Statistics of a data series in Figure 2.



Source: Calculated with Eviews 12 by authors with statistics

Fig 2: Statistics of Vietnamese pepper export price to U.S. market from January 2019 to February

Methodology

George Box and Gwilym Jenkins (1976) studied an autoregressive integrated moving average model (ARIMA). Their names (Box-Jenkins) is associated with the general ARIMA processes applied to the analysis and forecasting of time series. Autoregressive model of order p (AR(p)) is a linear dependence process of lagged values and random errors, which can be expressed as:

$$Y_t = j_1 Y_{t-1} + j_2 Y_{t-2} + ... + j_p Y_{t-p} + d + e_t \quad (1)$$

Moving-average model of order q (MA(q)), is a process describe through linear weighted equation of random errors and its lagged values. The model is written as follows.

$$Y_t = m + e_t - q_1 e_{t-1} - q_2 e_{t-2} - ... - q_q e_{t-q} \quad (2)$$

Autoregressive integrated moving average model (ARIMA (p,d,q)) is built upon the integration of two processes: (1) and (2). The general equation for the ARIMA (p, d, q) model is.

$$Y_t = j_1 Y_{t-1} + d + e_t - q_1 e_{t-1} - q_2 e_{t-2} - ... - q_q e_{t-q} \quad (3)$$

The Box-Jenkins method consists of four iterative steps: (i) Model identification, (ii) Estimation, (iii) Diagnostic and checking, and (iv) Forecast, as presented below:

Step 1: Model identification

Identifying ARIMA (p, d, q) model involves finding the appropriate values for p, d, and q. Where d represents the differencing order of the time series, p represents the autoregressive order, and q represents the moving average order. Determining p and q depends on the SPAC = f(t) and SAC = f(t) plots, with SAC being the sample autocorrelation function (SACF) and SPAC being the sample partial autocorrelation function. Selecting an AR(p) model relies on the SPAC plot if it exhibits high values at lags 1, 2, ., p and declines abruptly afterward, while the SAC gradually dies out. Similarly, choosing an MA(q) model is based on the SAC plot if it shows high values at lags 1, 2, ., q and drops sharply after q, while the SPAC gradually dies out.

Step 2: Estimate the parameters of ARIMA (p, d, q) model: Parameters of the ARIMA model will be estimated using least squares method.

Step 3: Diagnostic and checking

After identifying the parameters of the ARIMA model, then a diagnostic need to be done to test for white noise characteristic in error term e_t, homoscedasticity of residuals, normality of residuals and stationarity of the ARIMA model.

Step 4: Forecast

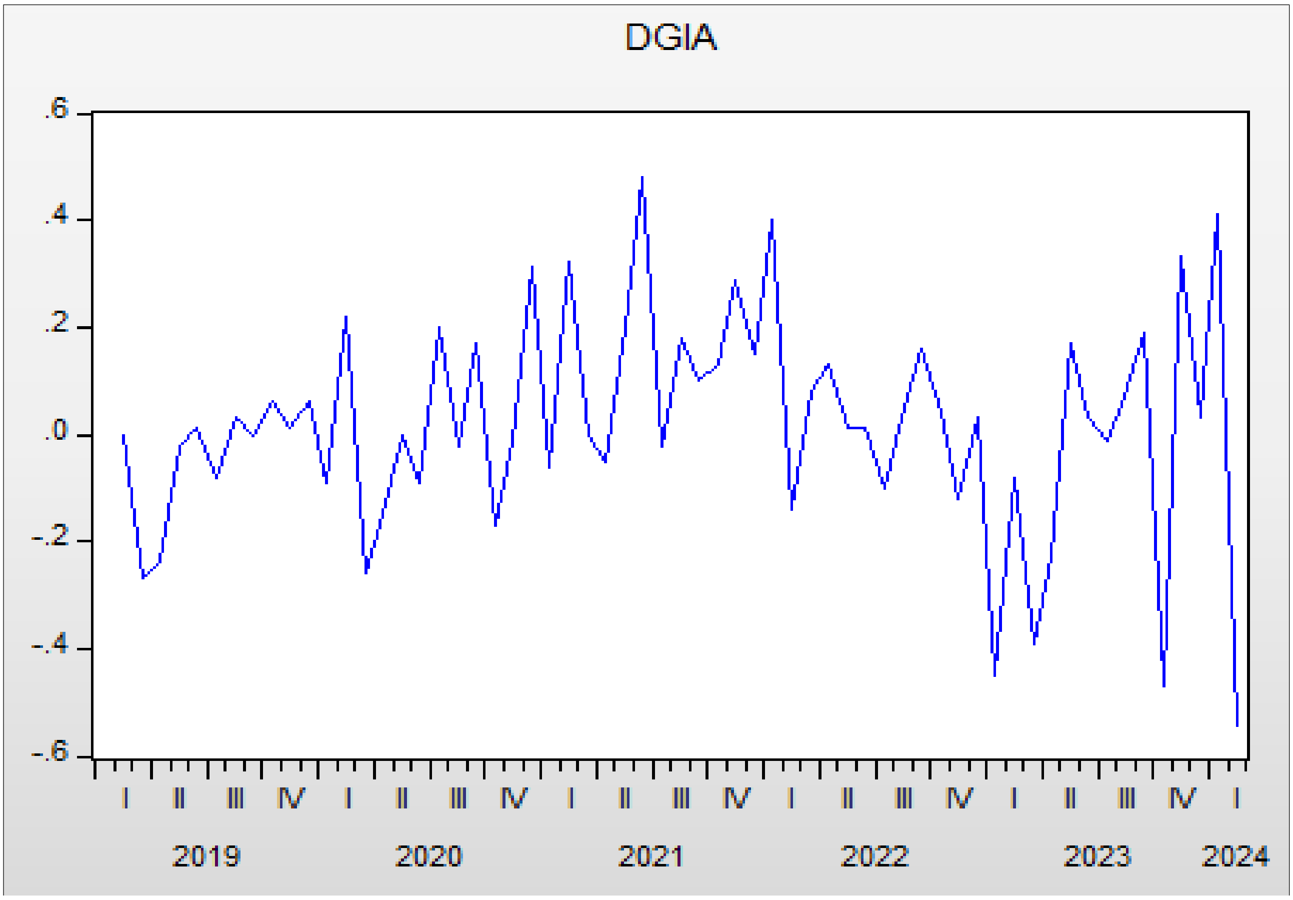
Based on the equation of ARIMA model, analyze point forecast and forecast confidence interval in ARIMA model.

Estimated results and discussion

Stationarity test: In mathematics, stationarity is used as a tool in time series analysis. In time series analysis, stationarity testing is a crucial step to ensure the statistical validity to form a model. A stationary process is a random process where the mean and variance of the errors remain constant over time. In reality, most economic time series (original series) are non-stationary. This implies that the mean and variance of these time series change over time. However, differencing data often transforms non-stationary time series into stationary ones.

Data series used in ARIMA model is assumed to be stationary. Therefore, to forecast the amount of international travelers to Vietnam with ARIMA model, we have to identify whether the input data are stationary. To confirm this, it is possible to first rely on direct observation of the graph of the time series, and then test it. We utilize a widely used testing method, Augmented Dickey-Fuller (ADF), a method, which in econometrics is known as unit root test for root and differenced time series.

Figure 1 shows that the monthly export price of black pepper from Vietnam to the US (denoted as GIA) is unstable and has an increasing trend. Specifically, its mean tends to increase over time periods. This suggests that the export price of black pepper from Vietnam to the US is non-stationary. However, when the first-order difference of this series is taken, the new series, the change in the export price of black pepper from Vietnam to the US (abbreviated as dgia), does not have a clear trend and fluctuates around a certain average value (Figure 3). This is considered a characteristic of a stationary series. To confirm these speculations, we use the ADF and PP test methods to see if the results are consistent.



Source: Calculated with Eviews 12 by authors with statistics

Fig 3: Changes of the order-one differenced series of black pepper export price from Vietnam to the U.S. from January 2019 to February 2024

Stationary characteristic test results using ADF and PP is given below in table 1.

Table 1: ADF and PP test results with root series

	Time series symbol	Testing statistics
		ADF
Time series	gia	-0.838805
Notation: The values at the 1%, 5%, and 10% significance levels are -3.542097, -2.910019, and -2.592645, respectively.		

Source: Calculated with Eviews 12 by authors with statistics.

From table 1, based on values at the 1%, 5% and10% all conclude that gia series is non-stationary.

Results from testing stationarity of dgia series through ADF and PP is show in table 2 below:

Table 2: ADF and PP test results for differenced series order 1

	Time series symbol	Testing statistics
		ADF
Time series	dgia	-7,889260
Notation: The values at the 1%, 5%, and 10% significance levels are -3,544063, -2,910860 and -2,593090.		

Source: Calculated with Eviews 12 by authors with statistics

From table 1, values at the 1%, 5% and 10% conclude dgia is stationary series. Therefore, both of the testing methods are uniformed and we can conclude that dgia is stationary series.

Building an ARIMA Model for the Time Series of Vietnamese Black Pepper Export Price Volatility to the US Market: To construct the ARIMA model, we utilize a time series dataset comprising 62 observations from January

2019 to February 2024. The original data is represented by the gia series, and the first-order differenced data is represented by the dgia series.

Step 1: Identification (identify the values of p, d and q): The gia series test above indicates that the series is stationary at first-order differencing, implying that *d=1*.
By using autocorrelation and partial autocorrelation function plots of dgia time series, we determined the order of the ARIMA model *(p,1,q)* to be *p = 2* and *q = 0.*

Step 2: Estimate ARIMA (2,1,0) model: Estimating the model using maximum likelihood method and test to retain regression coefficients that were statistically significant at 5% significance level. The results are as follows.

$$dgia_t = 0,285257 dgia_{t-2} + \varepsilon_t \tag{4}$$

Source: Calculated with Eviews 12 by authors with statistics.

Step 3: Model testing and diagnostics
***Test that the residual is stationary series**
Results from testing residual series (e) with ADF and PP testing methods are shown in table 3 below.

Table 3: ADF and PP test results on residuals series

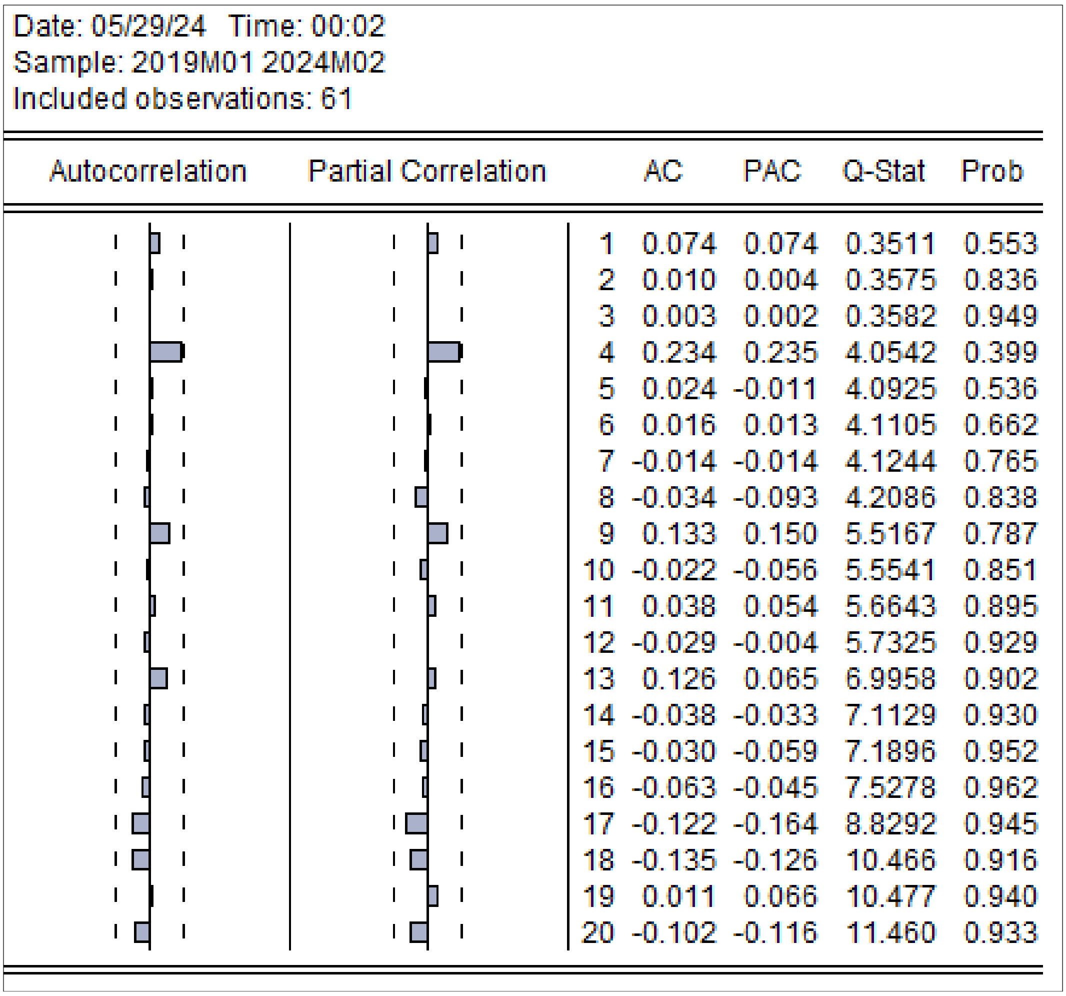
	Time series symbol	Testing statistics
		ADF
Time series	e	-8,222255
Notation: The values at the 1%, 5%, and 10% significance levels are -3,548208, -2,912631 and -2,594027.		

Source: Calculated with Eviews 12 by authors with statistics

From table 1, with values at 1%, 5% at 10% all conclude that residual series of model (4) is stationary.

*** Test for white noise characteristic in residuals**
White noise characteristic in residuals is done using Ljung Box Q test.

Null hypothesis H_o: The model's residual is white noise.
Alternative hypothesis H_1: The model's residual is not white noise.
With α = 5%, test done with Eviews 12 give results shown in table 2.



Date: 05/29/24 Time: 00:02
Sample: 2019M01 2024M02
Included observations: 61

Lag	AC	PAC	Q-Stat	Prob
1	0.074	0.074	0.3511	0.553
2	0.010	0.004	0.3575	0.836
3	0.003	0.002	0.3582	0.949
4	0.234	0.235	4.0542	0.399
5	0.024	-0.011	4.0925	0.536
6	0.016	0.013	4.1105	0.662
7	-0.014	-0.014	4.1244	0.765
8	-0.034	-0.093	4.2086	0.838
9	0.133	0.150	5.5167	0.787
10	-0.022	-0.056	5.5541	0.851
11	0.038	0.054	5.6643	0.895
12	-0.029	-0.004	5.7325	0.929
13	0.126	0.065	6.9958	0.902
14	-0.038	-0.033	7.1129	0.930
15	-0.030	-0.059	7.1896	0.952
16	-0.063	-0.045	7.5278	0.962
17	-0.122	-0.164	8.8292	0.945
18	-0.135	-0.126	10.466	0.916
19	0.011	0.066	10.477	0.940
20	-0.102	-0.116	11.460	0.933

Source: Calculated with Eviews 12 by authors with statistics

Fig 4: Ljung Box Q test on model (4)'s residuals

From Table 4, we have all Prob (probability) larger than 5%, so we accept H_o, the residuals in model (4) have white noise characteristics.

Homoscedasticity test
We use ARCH test to identify model's homoscedasticity.

Null hypothesis H_o: the model does not have homoscedasticity.

Alternative hypothesis H_1: The model has homoscedasticity.
With α = 5%, test done with Eviews 12 give the results show in table 4.

Table 4: ARCH test for model (4)'s residual heteroscedasticity

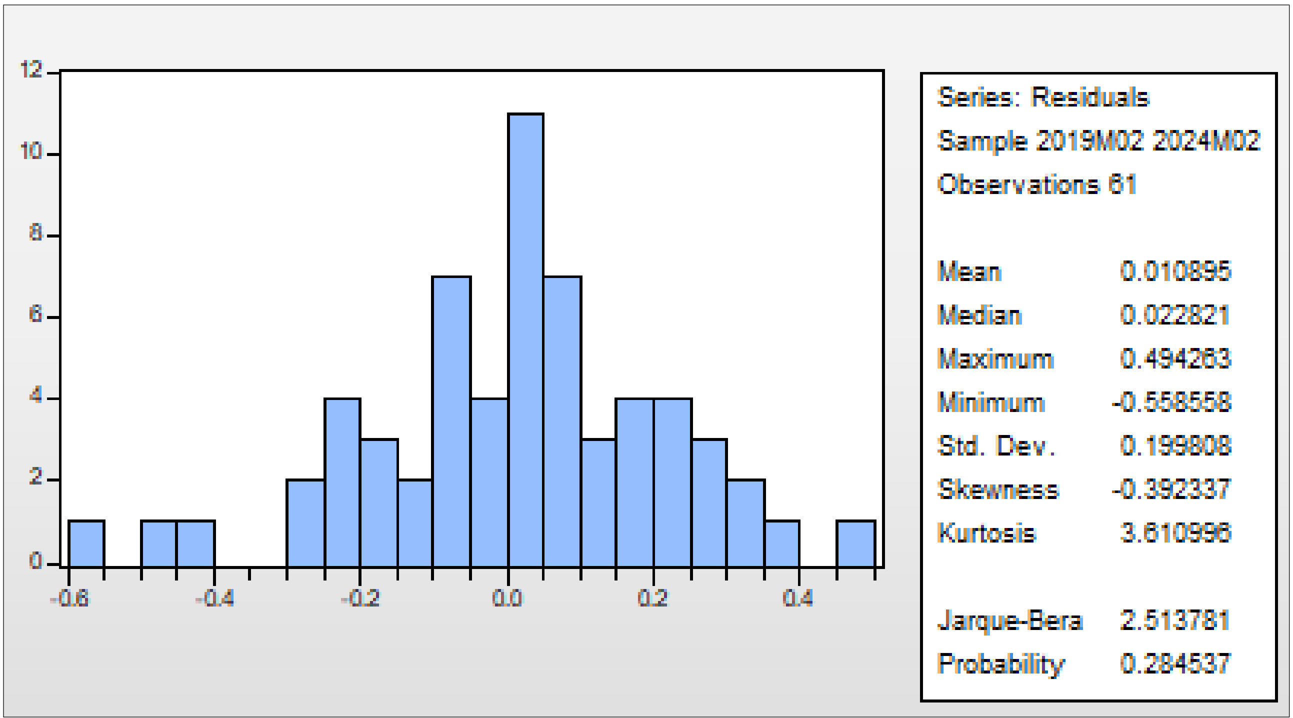
Heteroskedasticity Test: ARCH			
F-statistic	0.481864	Prob. F(1,58)	0.4904
Obs*R-squared	0.494373	Prob. Chi-Square(1)	0.4820

Source: Calculated with Eviews 12 by authors with statistics.

From Table 4, we have all Prob (probability) larger than 5%, so we accept H_o, the model does not have homoscedasticity.

***Residual's distribution normality test**
We use Jarque-Bera test to assess the normality of residuals.

Null hypothesis H_o: normal residuals distribution.
Alternative hypothesis H_1: abnormal residuals distribution.
With α = 5%, making test with Eviews 12 give the results in fig 5.



Source: Calculated with Eviews 12 by authors with statistics

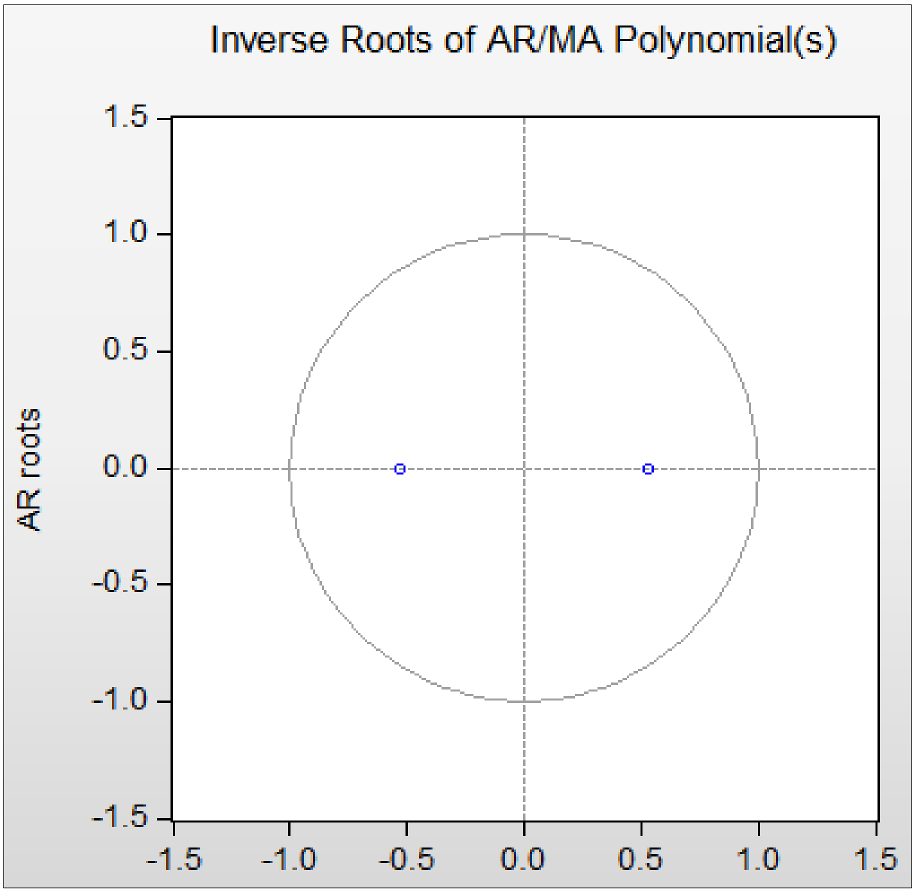
Fig 5: JB test and plot of model residuals graph (4)

From fig 5, we have Prob (probability) larger than 5%, so we accept H_0, the residuals have a normal distribution.

*** Testing stationarity in ARIMA model**
When testing, stationarity of AR and MA process, 2 details must be ensured.

- Stationary autoregressive model: inverse AR roots must lie within the unit circle.
- Invertible moving average model: inverse MA roots must lie within the unit circle.

The stability test conducted using Eviews 12 software yielded the results shown in Figure 6.



Source: Calculated with Eviews 12 by authors with statistics

Fig 6: Testing the stationarity of AR and MA processes

From figure 6, we can see that all the inverse roots lie within the unit circle. From the tests above, the ARIMA (2,1,0) model meets the stability conditions and the residual term is white noise, with a constant variance of the error.

Step 4: Forecast: From model (4), we identify.

$$dgia_t = 0,285257 dgia_{t-2} + \varepsilon_t \Leftrightarrow gia_t - gia_{t-1} = 0,285257(gia_{t-2} - gia_{t-3}) + \varepsilon_t$$

$$gia_t = gia_{t-1} + 0,285257(gia_{t-2} - gia_{t-3}) + \varepsilon_t \qquad (5)$$

Using model (5), we can identify the forcasted value of Vietnam's peppper export price to the U.S. market for the next 6 months, from March 2024 to August 2024, shown in table 5.

Table 5: Forcasting Vietnam's black pepper export price to the U.S.

Month	3/2024	4/2024	5/2024	6/2024	7/2024	8/2024
Price forecast US (Dollar/Kilograms)	4,53	4,37	4,40	4,35	4,36	4,35
Real price	4,75	4,47				
Error	0,22	0,10				

Source: Calculated with Eviews 12 by authors with statistics

Table 5 shows that the forecasted values for Vietnamese black pepper export price to the U.S from March to April 2024 is close to reality. This states that ARIMA (2,1,0) model is able to explain fluctuations of black pepper prices However, future forecast points involve larger errors. Therefore, it is necessary to update the data regularly to provide more realistic forecasts.

Conclusion

The fluctuations in the price of Vietnamese pepper exports to the US market is a time series follow an ARIMA model (2,1,0). Base on this model we can give short-term forecast Vietnam's black pepper export price to the U.S. market. Results for months from March 2024 to August 2024, indicates little error compared to reality, with up to 0,22 units, which is considerable in a volatile global economy at present. Nevertheless, the ARIMA model can be used to forecast, while not being optimal as the dependencies in the model are assumed to be linear.

Therefore, forecast results from ARIMA model (2,1,0) export price of Vietnamese black pepper to the US market can serve as a reference for management agencies, policymakers and businesses to implement guidelines, strategies, and propose solutions and recommendations for developing black pepper cultivated areas across the country, especially for applying technology in production to improve productivity and sustainability in the future.

References

1. Bộ Công Thương. Báo cáo xuất nhập khẩu Việt Nam 2022. Hà Nội: Nhà xuất Bản Hồng Đức; c2023.
2. Nguyễn Hạnh. Xuất khẩu hồ tiêu năm 2023 chỉ thu về 912 triệu USD. Công Thương; c2024 Jan 3.
3. Nông Hữu Tùng, Phạm Công Toản. Xuất khẩu hồ tiêu của Việt Nam và giải pháp phát triển. Tạp chí Nông nghiệp và Phát triển. 2022;(7):16-19.
4. Phan Thị Xuân Huệ. Cơ hội và thách thức trong xuất khẩu hồ tiêu của Việt Nam ra thị trường thế giới. Tạp chí Công thương. 2023;(5):1-9.
5. Tổng cục Thống kê. Hướng tới phát triển bền vững cây hồ tiêu: dữ liệu và số liệu thống kê năm 2021; c2021 [cited 2023 Jun]. Available from: https://www.gso.gov.vn/du-lieu-va-so-lieu-thong-ke/2021/06.
6. Nugroho AD, Prasada IY. Performance and forecast of Indonesian pepper exports to Italy. Bulgarian Journal of Agricultural Science. 2020;26(5):927-934.
7. Putri NH, Abidin Z, Situmorang S. The prospects of Lampung's pepper export to the global market: An analysis using the ARIMA model. HABITAT. 2023;34(3):289-298.
8. Weerasinghe WPMCN, Jayasundara DD. Modelling pepper export income in Sri Lanka using deterministic decomposition and seasonal ARIMA models. Statistics and Applications. 2021;19(2):89-100.