Pak. j. life soc. Sci. (2024), 22(1): 5914-5923 E-ISSN: 2221-7630;P-ISSN: 1727-4915

Pakistan Journal of Life and Social Sciences

www.pjlss.edu.pk

https://doi.org/10.57239/PJLSS-2024-22.1.00437

RESEARCH ARTICLE

Application of ARIMA Model in Forecasting Vietnam's Cashew Nut Export Volume

Phung Duy Quang[1*], Tran Le Phong[2], Nguyen Duy Phuong[3], Do My Chi[4], Trinh Ha Uyen[5], Bui Duc Duong[6]

[1] Faculty of Technology and Data Science,Foreign Trade University, Hanoi, Vietnam
[2]Student of K61 Advanced Program, Faculty of Banking and Finance, Foreign Trade University, Hanoi, Vietnam
[3] People's Security Academy
[4] Students of the Faculty of International Trade Law - Hanoi Law University, Vietnam
[5]Student K62 - Advanced Japanese Style International Business Program, Foreign Trade University
[6]Luong The Vinh Secondary and High School – Hanoi, Viet Nam

ARTICLE INFO

Received: Apr 24, 2024

Accepted: Jun 24, 2024

Keywords

ARIMA
Time series
Forecast
Cashew nut export volume
Data analysis

***Corresponding Author:**

quangpd@ftu.edu.vn

ABSTRACT

The study uses the ARIMA model to analyze and forecast Vietnam's cashew nut export output from July 2024 to December 2024 based on Vietnam's cashew nut export output data by month from January 2020 to June 2024. Research shows that the ARIMA (14, 0, 0) model is the most suitable model for the collected data. The results show that Vietnam's pepper export output in the last 9 months of 2024 tends to decrease slightly, the lowest export output is forecast to range from 57452,12 tons to 57817,34 in November and December. December 2024. According to the forecast results, Vietnam's cashew nut output tends to decrease slightly, so Vietnam needs to invest in improving techniques and technology to improve the quality of exported cashew nuts and find new markets for cashew export. At the same time, the Vietnamese government needs to support cashew producers so they can access new techniques and technology as well as access new export markets.

INTRODUCTION

Vietnam, with its tropical monsoon climate and fertile soil, is an ideal environment for cashew nut cultivation. As a result, Vietnam has become one of the world's leading producers of raw cashew nuts. In addition, Vietnamese cashew nuts are highly regarded for their quality and flavor. Cashew nuts are a significant perennial industrial crop with a large cultivation area and generate substantial economic benefits for cashew nut farmers. According to General Statistics Office of Vietnam, from 2008 to 2019, Vietnam emerged as the world's cashew nut processing hub, consistently leading in cashew nut export value. During this period, the total cashew nut harvest volume in Vietnam's cashew nut growing regions increased significantly, reaching an average of 298,000 tons per year, almost 100,000 tons more than the previous period. Due to ongoing research and improvements in cashew nut varieties, cultivation techniques, and plant care practices, the average yield in Vietnam's cashew nut regions during this period increased to nearly 0.98 tons per hectare.

According to ITC statistics, Vietnam and India remain the top two cashew nut exporters globally. During the 2016-2020 period, the average global export volume of processed cashew nuts reached 583,960 tons, equivalent to USD 4.257 billion. Vietnam contributed 360,438 tons, India contributed 71,626 tons, and the remaining quantity was exported by other countries. While raw cashew nuts are primarily consumed in major producing countries like Vietnam and India, the market for processed cashew nuts is concentrated in high-income and developed countries and territories. Currently, over 90 countries and territories worldwide have imported processed cashew nuts. The top five cashew nut importers in 2020 were the United States, Germany, the Netherlands, China, and the United Kingdom. Among these, the United States, Germany, and the Netherlands consistently ranked among the three largest markets for processed cashew nuts during the 2016-2020 period. In 2020, global processed cashew nut imports reached USD 4.025 billion, equivalent to 596,291 tons of cashew nuts. The top three importing countries were the United States (27.4%), Germany (12.2%), and the Netherlands (8.5%), accounting for a combined global market share of 48.1%.

There are foreign studies employing time series models to forecast cashew nut production. Prabakaran et al. (2015) developed ARIMA (1, 1, 0) and ARIMA (2, 1, 0) models using data from 1982 to 2010 to forecast cashew nut area and production in India during 2011-2015. The results indicated that the forecasted cashew nut area in 2015 would be around 100.656 million hectares, with upper and lower bounds of 1107.51 and 905.6 million hectares, respectively. The model also forecasted cashew nut production in 2015 to be around 768.71 thousand tons, with upper and lower bounds of 866.80 and 670.62 thousand tons, respectively.

Chaithra M. et al. (2019) employed ARIMA and Holt's exponential smoothing models to forecast the area and production of cashew nuts in Dakshina Kannada to assist policymakers in developing cashew nut-related policies. Holt's exponential smoothing model was found to outperform other forecasting models, with the lowest RMSE value (1386.13). Based on this model, the forecasted cashew nut area for 2018, 2019, and 2020 was 34,492.10 ha, 34,974.81 ha, and 35,474.87 ha, respectively. For cashew nut production, Brown's linear trend model, with an RMSE value of 10020.19, was identified as the superior forecasting model among those tested. The forecasted cashew nut production for 2018, 2019 and 2020 was 10230,20 tons, 10996,81 tons and 11833,00 tons, respectively.

The study by Alfred Nkubito Ngendahayo et al. (2024) provides a comprehensive analysis of the future prospects of Tanzania's cashew nut export industry using the ARIMA (1,1,1) model based on data from 59 years from 1965 to 2023 to forecast for the years from 2024 to 2033.

In Vietnam, author Vo Van Tai (2012) used various mathematical models such as regression and time series ARIMA(1,2,1) model to forecast Vietnam's rice production in the period 2015-2018. The study by Nguyen Dinh Thuan and Ho Cong Hoai (2018) studied and showed that the ARIMA and Support Vector Machine combined model provides more accurate forecasting results than other models. The study by Bui Thi Minh Nguyet et al. (2019) used the ARIMA (1,1,16) model to forecast Vietnam's export value for the last 6 months of 2018. The study by Tran Quang Canh and Phan Thi Dong Hoai (2021) used the ARIMA(1,1,1) model to forecast Vietnam's economic structure in 2025. The study by Nguyen Thi Hien et al. (2023) used the ARIMA (1, 1,12) model to analyze the data collected from July 2009 to January 2023 to forecast inflation in Vietnam in the first half of 2023. The study by Ho Trong Phuc and Pham Xuan Hung (2023) applied the ARIMA (1, 1, 0) model to analyze the time series data from 1995-2022 to forecast rice production, while proving the ARIMA (0,1,0) model is the optimal model for forecasting rice area and yield in the region. The study by Phung Duy Quang et al. (2024) used the ARIMA model to forecast Vietnam's coffee export volume based on data from 2000 to 2019 to forecast export volume for 2020 to 2023. The ARIMA model to forecast the export price of Vietnamese coffee to the US from March 2024 to August 2024 based on price data from January 2019 to February 2024 (Phung Duy Quang et al., 2024; Kanval et al., 2024). However, currently in Vietnam, there is no study using the ARIMA model in the analysis and forecasting of Vietnam's cashew nut export volume. Therefore, the research using the ARIMA model in the analysis and forecasting of Vietnam's cashew nut export volume is of scientific and practical significance.

The first part of the paper introduces an overview of time series research in forecasting in Vietnam and abroad, the urgency of forecasting research of Vietnam's cashew nut export volume. The rest of the paper is organized as follows: Section 2. Data and research methodology, Section 3. Estimation results and discussion, Section 4. Conclusion and Section 5 is Recommendations.

2. DATA AND RESEARCH METHODOLOGY

2.1. Data

The study is based on secondary data on cashew nut export volume synthesized from reports of the General Statistics Office and ITC. The data is synthesized by month and processed on Excel and Eviews 12 software. The data range starts from January 2020 to May 2024.

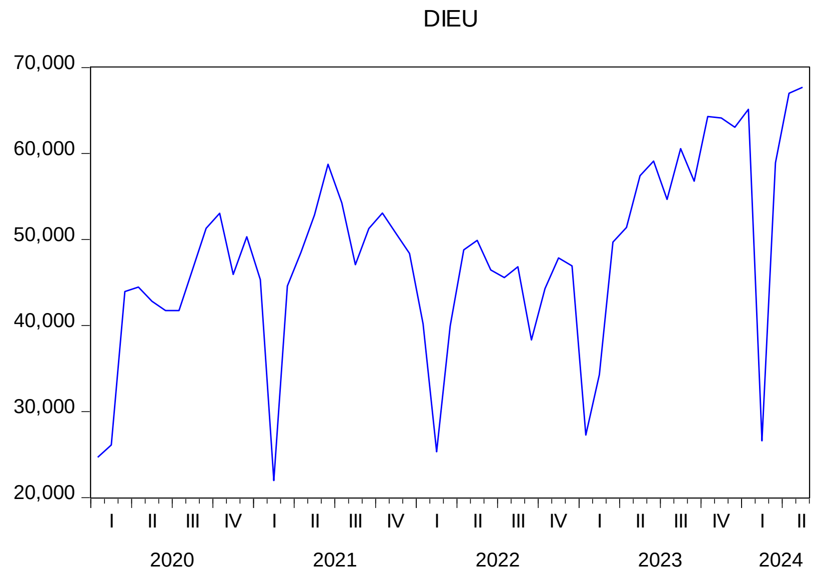

Figure 1. Changes in cashew nut export volume of Vietnam from January 2020 to May 2024

Source: Calculated by Eviews 12 by the authors from the data

The simple descriptive statistics of the data series are shown in Figure 2.

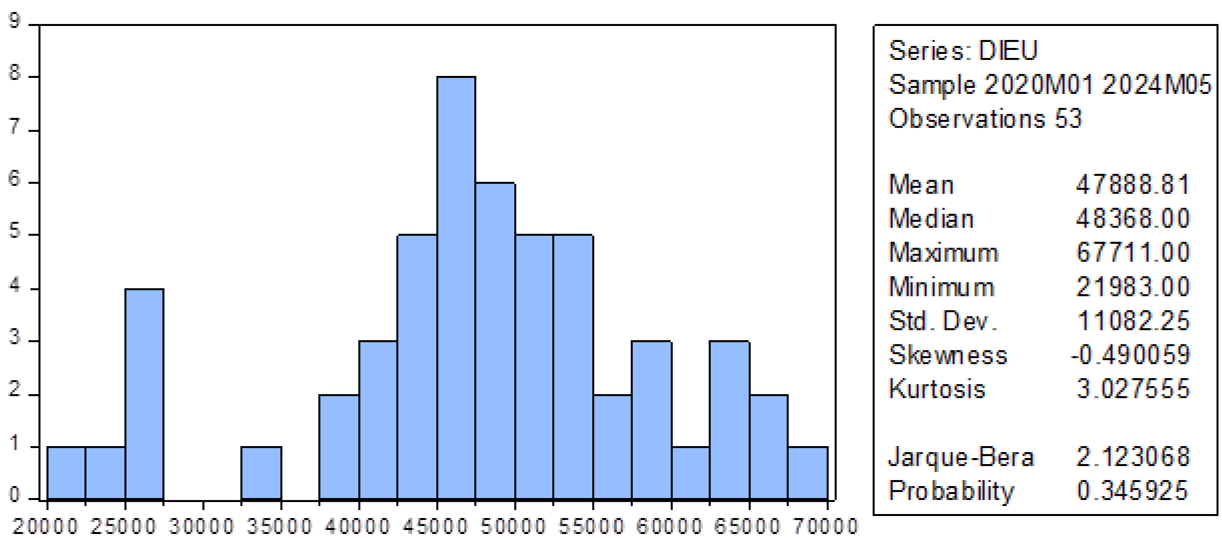

Figure 2. Descriptive statistics of cashew nut export volume of Vietnam from January 2020 to May 2024. *Source:* Calculated by Eviews 12 by the authors from the data

2.2. Research Methodology

George Box and Gwilym Jenkins developed the Autoregressive Integrated Moving Average (ARIMA) in 1976. Their names (Box-Jenkins) are used to refer to general ARIMA processes applied to time series analysis and forecasting. The autoregressive model of order p (AR(p)) is a linear dependence process of lagged values and random errors, expressed as follows:

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \ldots + \varphi_p Y_{t-p} + \delta + \varepsilon_t \quad (1)$$

The moving average model of order q (MA(q)) is a process completely described by a linear weighted equation of current random errors and their lagged values. The model is written as follows:

$$Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \quad (2)$$

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

$$Y_t = \varphi_1 Y_{t-1} + \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \quad (3)$$

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

Step 1: Model Identification

Identifying the ARIMA (p, d, q) model involves finding the appropriate values of p, d, and q. With d being the order of differencing of the time series under investigation, p is the order of autoregression, and q is the order of moving average. Determining p and q will depend on the SPAC = f(t) and SAC = f(t) graphs, where SAC is the sample autocorrelation function (SACF) and SPAC is the sample partial autocorrelation function (SPACF). The choice of the AR(p) model depends on the SPAC graph if it has high values at lags 1, 2,..., p and then drops sharply, as the SAC function fades out. Similarly, the choice of the MA(q) model is based on the SAC graph if it has high values at lags 1, 2,..., q and drops sharply after q, as the SPAC function fades out.

Step 2: Estimation of ARIMA (p, d, q) Model Parameters

The parameters of the ARIMA model will be estimated using the least squares method.

Step 3: Diagnostic Checking of the Model

After determining the parameters of the ARIMA process, it is necessary to check whether the error term e_t of the model is white noise, to check that the residuals have no constant variance of the error, to check that the residuals have a normal distribution, and to check the stability of the ARIMA model.

Step 4: Forecasting

Based on the equation of the ARIMA model, the point forecast value and the forecast confidence interval are determined.

3. ESTIMATION RESULTS AND DISCUSSION

3.1. Stationarity Test

In mathematics, stationarity is used as a tool in time series analysis. To form a meaningful statistical model, the time series must first be tested for stationarity. A stationary process is a random process that is represented by a constant mean and variance of the error over time. In reality, most economic time series (raw series) are non-stationary. This means that these time series have a mean and variance that change over time. However, when we take the difference, the time series often become stationary (Do Q. Giam & cs., 2009).

The time series used in the ARMA model is assumed to be stationary. Therefore, to forecast the number of international tourists to Vietnam using the ARMA model, we need to consider whether these series are stationary. To confirm this, we can first rely on direct observation of the time series graph, and then test it. We use the very common Augmented Dickey-Fuller (ADF) test, which econometrics calls a unit root test for the raw time series and the differenced time series.

The results of the stationarity test of the DIEU series using the ADF test method are shown in Table 1 below:

Table 1. Stationarity test results for the DIEU series

Null Hypothesis: DIEU has a unit root			
Exogenous: Constant			
Lag Length: 0 (Automatic - based on SIC, maxlag=10)			
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.144820	0.0019
Test critical values:	1% level	-3.562669	
	5% level	-2.918778	
	10% level	-2.597285	
*MacKinnon (1996) one-sided p-values.			

Source*:* Calculated by Eviews 12 by the authors from the data

3.2. Building an ARIMA Model to Forecast Vietnam's Cashew Nut Export Output

To build the ARIMA model, we use a data series of 53 observations from January 2020 to May 2024. The original data is the DIEU series.

Step 1: Identification (Determine the values of p, d, q)

The DIEU data series checked above shows that this series is stationary, so *d = 0*.

Using the autocorrelation and partial autocorrelation plots of the DIEU series in combination with the model selection criteria, we determine the best ARIMA (p, 0, q) model: *p = 14* and *q = 0*.

Step 2: Estimate the ARIMA(14,0,0) model

The model is estimated using the maximum likelihood method and testing to retain statistically significant regression coefficients at the 5% significance level, we obtain:

$$DIEU_t = 47305,62 + 0,51416 DIEU_{t-1} - 0,335033 DIEU_{t-14} + \varepsilon_t \qquad (4)$$

Source: Calculated using Eviews 12 by the authors from the data

Step 3: Model Diagnostic Check

* Check if the residuals are a stationary series

The results of the stationarity test of the residual series (denoted by e) using the ADF test are shown in Table 2 below:

Table 2: Stationarity test results for residual series

Null Hypothesis: E has a unit root			
Exogenous: Constant			
Lag Length: 0 (Automatic - based on SIC, maxlag=10)			
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-7.447118	0.0000
Test critical values:	1% level	-3.562669	
	5% level	-2.918778	
	10% level	-2.597285	
*MacKinnon (1996) one-sided p-values.			

Source: Calculated using Eviews 12 by the authors from the data

From Table 2, with the significance levels of 1%, 5%, and 10%, we conclude that the residual series of model (4) is a stationary series.

* Check if the residuals are white noise

To test whether the model has white noise residuals, we use the Ljung Box Q statistical test.

Hypothesis H_o: The model has white noise residuals.

Alternative hypothesis H_1: The model does not have white noise residuals.

With α = 5%, perform the test using Eviews 12 for the results in Table 3:

Table 3. Ljung Box Q Test of Residuals of Model (4).

Date: 07/23/24 Time: 18:34				
Sample: 2020M01 2024M05				
Included observations: 53				
	AC	PAC	Q-Stat	Prob
1	-0.035	-0.035	0.0668	0.796
2	-0.047	-0.048	0.1904	0.909
3	0.016	0.013	0.2062	0.977
4	-0.064	-0.065	0.4470	0.978
5	0.076	0.074	0.8012	0.977
6	0.008	0.006	0.8050	0.992
7	0.072	0.083	1.1327	0.992
8	-0.090	-0.093	1.6590	0.990

9	-0.026	-0.014	1.7041	0.995
10	0.067	0.049	2.0057	0.996
11	-0.039	-0.026	2.1099	0.998
12	0.259	0.249	6.8876	0.865
13	0.189	0.223	9.4865	0.735
14	-0.020	0.038	9.5178	0.797
15	-0.060	-0.051	9.7928	0.833
16	-0.173	-0.188	12.137	0.734
17	0.044	-0.021	12.294	0.782
18	-0.108	-0.167	13.258	0.776
19	0.013	-0.041	13.274	0.824
20	-0.047	-0.062	13.473	0.856
21	-0.052	0.035	13.717	0.881
22	-0.040	-0.079	13.868	0.906
23	0.006	0.022	13.872	0.930
24	0.248	0.198	20.043	0.694

Source: Calculated using Eviews 12 by the authors from the data

According to Table 3, we have all Prob (probability) values greater than 5%, so we accept H_o, model (4) has white noise residuals.

* Check if the residuals have constant variance of the error

To test whether the model has constant variance of the error, we use the ARCH test.

Hypothesis H_o: The model has constant variance of the error.

Alternative hypothesis H_1: The model has changing variance of the error

With $\alpha = 5\%$, perform the test using Eviews 12 for the results in Table 4:

Table 4. ARCH Test of Residuals of Model (4).

Heteroskedasticity Test: ARCH			
F-statistic	0.137028	Prob. F(24,4)	0.9995
Obs*R-squared	13.08490	Prob. Chi-Square(24)	0.9647

Source: Calculated using Eviews 12 by the authors from the data

According to Table 4, we have all Prob (probability) values greater than 5%, so we accept H_o, the model has constant variance of the error.

* Check the stationarity of the ARIMA model

When checking the stationarity conditions of the AR and MA processes, two factors need to be ensured:

- The estimated model is stationary in covariance: the inverse AR roots must lie within the unit circle.
- The estimated process is invertible: the inverse MA roots must lie within the unit circle.

The stationarity test was performed using Eviews 12 software and the results are shown in Figure 3.

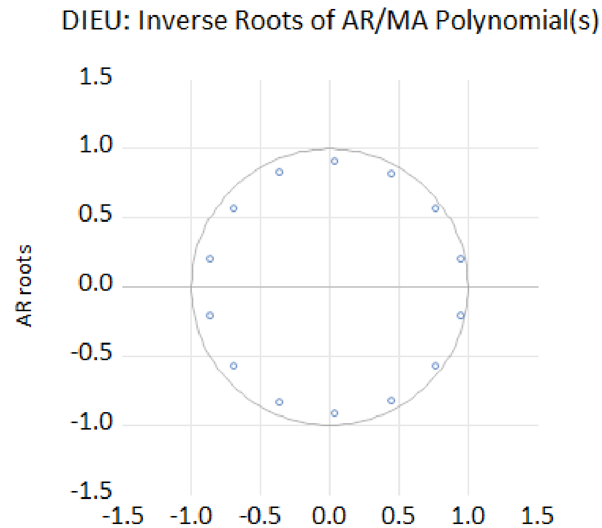

Figure 3. Test for Stationarity of the AR and MA Processes

Source: Authors' calculations using Eviews 12 from data

As can be seen from Figure 3, all the inverse roots lie within the unit circle.

From the above tests, the ARIMA (14,0,0) model meets the stationarity conditions and the residual term is white noise, with constant error variance.

Step 4: Forecasting

Using model (4):

$$DIEU_t = 47305,62 + 0,51416 DIEU_{t-1} - 0,335033 DIEU_{t-14} + \varepsilon_t$$

we can determine the forecast values for the next 7 months from June 2024 to December 2024 for Vietnam's cashew nut export volume, as shown in Table 5.

Table 5. Forecast of Vietnam's Cashew Nut Export Volume from June 2024 to December 2024

Month	**Forecast Price**
6/2024	65468,77
7/2024	63748,69
8/2024	60843,38
9/2024	58782,70
10/2024	59211,41

11/2024	57452,12
12/2024	57817,34

Source: Authors' calculations using Excel from data

Table 5 shows that the forecast for Vietnam's cashew nut export volume to the world from June 2024 to the end of 2024 is on a downward trend.

4. CONCLUSION

This study utilizes time series analysis and monthly data collected from January 2020 to May 2024 to forecast Vietnam's cashew nut export volume from June 2024 to December 2024. Based on the AIC, BIC, HQC, R^2 criteria, and Log likelihood function values, the study indicates that the ARIMA (14, 0, 0) model is the most suitable model with the smallest AIC, BIC, HQC values and the largest R^2 and Log likelihood values. The results suggest that Vietnam's cashew nut exports will exhibit a slight downward trend in the last nine months of 2024, with the lowest export volume forecasted to range from 57,452.12 tons to 57,817.34 tons in November and December 2024.

5. Recommendations

Based on the forecast data for Vietnam's cashew nut exports, policymakers and management agencies can develop appropriate export strategies and solutions in line with the slight downward trend in export volume during the last seven months of 2024. Some suggested policies include:

Investing in technologies, cultivation, and production processes to enhance the quality of cashew nut products. Improving quality will boost the competitiveness of Vietnamese cashews in the international market and facilitate exports to demanding markets like the US, EU, and Japan.

Encouraging access to new markets and promoting trade. Diversifying export markets can mitigate risks and enhance the stability of cashew nut exports in particular and the cashew nut industry as a whole, especially during the current period of global economic and political instability.

Supporting cashew nut farmers and businesses in overcoming the current volatile market conditions. This includes providing technical advice, financial assistance, and risk management support to help them adapt to changes and strengthen their competitiveness.

REFERENCES

Alfred Nkubito Ngendahayo, Dr. Jayavantha Nayak (2024), Examining the dynamics of cashew kernel exports in Tanzania: a time series forecasting approach, *Journal of Management*, Vol.11, No. 2, Pages: 1-8.

Bui Thi Minh Nguyet, Nguyen Thi Quynh Cham, Nguyen Thi Quynh Nga (2019), Using the ARIMA model in forecasting Vietnam's export value, *Journal of Research in Finance and Accounting (Vietnam)*, Vol. 186, No. 1, Pages: 58 – 65.

Chaithra M. Pramit Pandit, Bishvajit Bakshi (2019), Forecasting of area and production of cashew nut in Dakshina Kannada using ARIMA and Exponential Smoothing Models, *Journal of Reliability and Statistical Studies*, Vol. 12, No. 2, Pages: 61-76.

Ho Trong Phuc, Pham Xuan Hung (2023), Forecasting Vietnam's rice area, yield and output: Applying the ARIMA model, *Hue University Science Journal: Economics and Development (Vietnam)*, Vol.132, No. 5C, 2023, Pages: 85–104.

Nguyen Dinh Thuan, Ho Cong Hoai (2018). Combining the Arima model and Support vector machine for forecasting at a Vietnamese community online service company, *Proceedings of the 11th*

National Science and Technology Conference on Basic Research and Application of Information Technology, Vietnam. doi: 10.15625/vap.2018.00030.

Nguyen Thi Hien, Le Mai Trang, Pham Long Vu, Phan Van Duy Hoang, Le Quang Huy (2023), Application of the ARIMA model to forecast Vietnam's inflation and some recommendations, *Banking Journal (Vietnam)*, Vol. 2023, No. 1, Pages: 65- 76.

Phung Duy Quang, Trinh Quoc Thang, Do Quang Truong, Nguyen Ngan Giang, Nguyen Van Ha,Ngo Gia Khiem, Tran Thi Minh Ngoc (2024), Building the ARIMA Model for Forecasting the Production of Vietnam's Coffee Export, *Journal of Applied Mathematics and Physics*, Vol.12, No.4, Pages: 1237-1246.

Phung Duy Quang, Pham Ngoc Mai and Hoang Nam Quyen (2024), Application of ARIMA model in forecasting Vietnam's black pepper export price to the U.S. market, *International Journal of Statistics and Applied Mathematics*; Vo.9, No. 3, Pages: 40-47

Prabakaran, K., Sivapragasam, C., Arulanandu, U. (2015), Forecasting cultivated areas and production of cashew nut in India using ARIMA model, *Acta Horticulturae* , Vol.1080,Pages:281-289.