JURNAL STATISTIKA DAN APLIKASINYA

e-ISSN 2620 – 8369
VOLUME 8 (ISSUE 2) DECEMBER 2024: 164-170
https://doi.org/10.21009/JSA.08203

FORECASTING HORTICULTURAL PRODUCTION BASED ON RAINFALL WITH VECTOR AUTOREGRESSIVE EXOGENOUS (VARX) METHOD

Ayu Septiani[1], Hardianti Fatmi [2*], Ristu Haiban Hirzi[3], Muhammad Gazali[4]

[1,2,3,4] Department of Statistics, Faculty of Mathematics and Natural Science Statistics, Universitas Hamzanwadi Jln. TGKH.M. Zainuddin Abdul Madjid, Pancor Subdistrict, East Lombok, NTB, 83612, Indonesia

Corresponding author's e-mail: ***hardiantifatmi6@gmail.com***

Article History:
Received: 27, October 2024
Revised: 24, December 2024
Accepted: 27, December 2024
Published: 31, December 2024
Available online.

Keywords:
Horticultural Production, Rainfall, VARX Method, Forecasting

ABSTRACT

Changes in the production yield of three leading commodities in NTB Province continue to fluctuate and may harm society. To examine the condition of the production of cayenne pepper, red onion, and tomatoes based on rainfall in the coming period, a model is needed that can predict multivariate time sequence data. There are several models of multivariate time sequence analysis, one of which is the autoregressive exogenous vector (VARX). The VARX model is a multivariate time series model consisting of several endogenous variables (p) and supplemented with exogenous variables (q). The purpose of this study is to obtain a suitable VARX model and an estimation of cayenne pepper, onion, and tomato data. Using the VARX method, the optimum lag was obtained with the smallest Akaike Information Criterion (AIC) value, namely at lag 5 with a value of 66.5160. Based on the overfit carried out, the appropriate and best model to be estimated is VARX (1.1) with a total value of 66.42364, which meets the assumption of white noise and multivariate normal distribution that satisfies the optimum forecast amount.

How to cite this article:
A. Septiani, H. Fatmi, R.H. Hirzi, M. Gazali, "FORECASTING HORTICULTURAL PRODUCTION BASED ON RAINFALL WITH VECTOR AUTOREGRESSIVE EXOGENOUS (VARX) METHOD", Journal Statistika dan Aplikasinya, vol. 8, iss. 2, pp. 164 – 170, December 2024

1. INTRODUCTION

Regional development is considered very strategic within the framework of implementing national development. Not only is regional development recognized as an integral part of national development, but it is also recognized as successful in encouraging increased equity, stability, growth, and community welfare as the main actors of development [1]. Regional economic growth is basically influenced by the comparative advantage of a region, regional specialization, and the economic potential of the region. Therefore, the utilization and development of all economic potential is a top priority that must be explored and developed in carrying out regional economic development in a sustainable manner [2].

The Agriculture Sector is an engine of growth, but in terms of providing raw materials, employment opportunities, food, and a focusing power for the products produced by another sector. The horticulture subsector is an essential subsector in Agriculture development that continuously to grow and develop from time to time. According to the horticultural fixed number data, cayenne pepper occupies the second position as the commodity with the most significant production in NTB Province. Cayenne pepper production in 2019 amounted to 164.77 thousand tons, a decrease of 21.73 percent compared to 2018, which reached 210.53 thousand tons. Furthermore, the shallot production in 2015 amounted to 1,229,189 tons, then increased yearly, and in 2019 amounted to 1,580,247 tons. This shows an increase of 22.2 percent. Meanwhile, tomato production in NTB Province experienced fluctuations. Tomato production reached its highest peak in 2022 at 40,742 tons, and the lowest production occurred in 2018 at 20,871 tons; then, tomato production rose again in 2019 to 29,215 tons [5].

Based on the data above, it can be explained that the phenomenon that can occur is the increase in prices of three leading commodities; these commodities are also included in the five primary commodities that provide the largest share of inflation in December 2016 [6]. The production of cayenne pepper, red onions, and tomatoes needs to be handled to minimize fluctuations in the prices of cayenne pepper, red onions, and tomatoes that cause inflation, and the provincial government quickly handles and makes policies in managing national food stocks and controlling the shipment of goods out of the region to meet the needs of the area.

Based on the increase in the price of cayenne pepper, red onions, and tomatoes that occur, to minimize price fluctuations, know how much the supply of cayenne pepper, red onions, and tomatoes is so that it can anticipate when price fluctuations will occur from the amount of production that will be produced. Production data of three leading commodities in NTB Province includes time series data. Time series data is defined as a sequence of observation data based on the same time interval where the observations have a correlation or independence [7]. Therefore, a model is needed to predict the amount of production of the three leading commodities in the future through forecasting. The Vector Autoregressive Exogenous (VARX) model is a time series model to model several endogenous variables that are interconnected and influenced by the previous time, and there are exogenous variables that affect these endogenous variables [8].

2. METHODOLOGY

Data

The data used in this study are secondary data. In this study, the secondary data used is data on the amount of cayenne pepper production per month (Y1), the amount of red onion production per month (Y1), the amount of tomato production per month (Y3), and the amount of rainfall per month (X) with the period January 2018 to December 2023. This research data was obtained from the NTB Provincial Statistics Agency and the official annual publication of the District Statistics Agency in NTB Province.

Research Methods

In the analysis of horticultural forecasting based on rainfall in NTB Province, researchers used the Vector Autoregressive Exogenous (VARX) method. The selection of the VARX method is used because it can model several endogenous variables that are interconnected and influenced by the previous time, and there are exogenous variables that affect the endogenous variables. The Vector Autoregressive Exogenous (VARX) model is a development of the Vector Autoregressive (VAR) model that uses exogenous variables in the equation system [9]. The Vector Autoregressive (VAR) model is a multivariate form of the Autoregressive (AR) model. In the VAR model, all variables are considered endogenous and interconnected variables. The VAR model extends the AR (autoregressive) model on univariate time series [10]. Exogenous variables (independent variables) in VARX are determined outside the model and affect endogenous variables in equations. Meanwhile, the endogenous variable (dependent variable) in VARX is defined in the model and can be influenced by exogenous variables.

The previous research [11] had the title VarX Modeling for Inflation Forecasting by Commodity Subgroup in Jakarta with the TSClust Approach as Preprocessing. Based on the research conducted, this research contains the creation of a forecasting model for 35 inflation rates according to commodity sub-groups. TSClust will cluster 35 inflation commodity subgroups into k clusters. From each cluster, the ARIMA model is identified. The difference with previous researchers lies in the approach and forecasting methods used, namely subgroups and ARIMA identification only. Furthermore, the second research [12] analyzed the relationship between education spending and economic growth using a research approach to this relationship. After analyzing and processing the data, the results show that the two research variables have a causal relationship, meaning that the two variables of education spending and economic growth affect each other. The difference with previous researchers is that they only see the causality relationship.

Time Series

Time series analysis is one of the statistical procedures applied to forecast the probabilistic structure of future states to make decisions [13]. Time series models attempt to predict future conditions using historical data and exploit the pattern into the future. A sequence of observations has a time series model if it satisfies the following [14].

a. The time interval between time indices t can be expressed in the same (identical) unit of time.
b. There is a dependency between observations Z_t and Z_{t+k} separated by a distance in the form of a time multiple of at k times (expressed as *lag k*).

VARX Method

The Vector Autoregressive Exogenous (VARX) model is a Vector Autoregressive (VAR) model development that uses exogenous variables in the equation system. Exogenous variables (independent variables) in VARX are determined outside the model and affect endogenous variables in equations. The stages in conducting the research are as follows [15]:

1. Conduct descriptive statistical analysis systematically describing factual and accurate data.
2. Tested the stationarity of all variables using the Augmented Dickey-Fuller (ADF) unit root test.
3. Test the causal relationship between one variable and another using the Granger causality test.
4. Determine the optimal *lag* length to limit the *lags* in the VARX model.
5. Determine the order of the VARX model by looking at the minimum value of the frequently used information criterion, Akaike's Information Criterion (AIC).
6. Perform parameter estimation of the VARX model with the *Least Squares Method* estimation.
7. Test the significance of the VARX model parameters using the statistical *t-test*.

8. Conduct a residual feasibility test of the VARX model, which includes the following assumptions:
    a. Assumption of white noise residuals using the *Ljung Box test*.
    b. Assumption of multivariate normality using the *Jarque-Bera test.*
9. Perform forecasting for future periods with the selected VARX model.

3. RESULT AND DISCUSSION

Descriptive Analysis

The study used data on the production of cayenne pepper, red onions, tomatoes, and rainfall in West Nusa Tenggara Province from January 2018 to December 2023. The following is a summary of the data used in the study.

Table 1. Data summary

Variables	Maximum Value	Minimum Value	Standard Deviation	Average
Cayenne Pepper	322183.0	17914.580	73779.71	91023.13
Red Onions	517374.620	9972.000	135392.3	164461.5
Tomatoes	43458.820	11258.00	7953.054	24792.23
Rainfall	308.940	2.67000	90.93995	122.1528

Table 1 above shows one of the summaries used in the study. **Figure 1** shows that the amount of production of these leading commodities fluctuates every year.

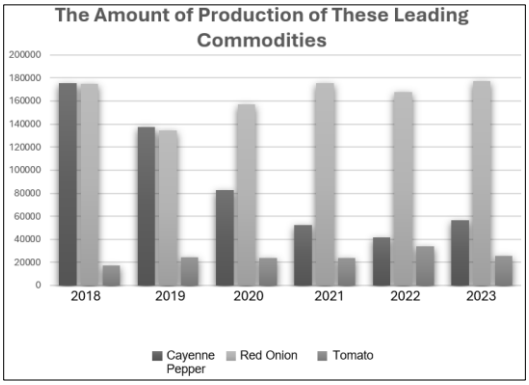

Figure 1. Total Production of Leading Commodities

Vector Autoregressive Exogenous (VARX) Analysis

Stationarity Test

In forming the VARX model, it is necessary to test the data variables to see their stationarity. Therefore, it is essential to test the stationary data.

Table 2. Summary of Data Stationary Test

Variables	Probability	Description
Cayenne Pepper	0.0840	Non – stationary
Red Onion	0.0041	Stationary
Tomato	0.5172	Non – stationary
Rainfall	0.0155	Stationary

From **Table 2**, it can be explained that variables have not been stationary in the average, so differencing one is carried out. Then, stationarity testing is repeated so that all data is stationary.

Tabel 2. Stationarity test of Data Differencing 1

Variables	Probability	Description
Cayenne Pepper	0.01	Stationary
Red Onion	0.004	Stationary
Tomatoes	0.01	Stationary
Rainfall	0.01	Stationary

Lag Optimum

Table 3. Optimum Lag Value

Lag optimum	AIC
0	67.0318
1	66.6678
2	66.7647
3	66.7001
4	66.6644
5	66.5160
6	66.6904
7	66.5974
8	66.7207
9	66.8394
10	66.7694

Based on Table 4, the smallest Akaike Information Criterion (AIC) value is at lag 5 of 66.5160, so the model is limited to the optimum lag length, namely 5.

Model Identification

Based on table 4 with the smallest AIC value obtained from VAR analysis using chili, onion and tomato variables, the AIC value at lag 5 is 66.5160, so the best model is VARX (1, x). Then after overfit from VARX (5,1) to VARX (1,1), the AIC value of each model is obtained as follows:

Tabel 4. Overfitting Model VARX

Model VARX	AIC
(5,1)	66.62564
(5,2)	66.64831
(4,1)	66.67079
(4,2)	66.71155
(3,1)	66.59371
(3,2)	66.65055
(2,1)	66.55409
(2,2)	66.61443
(1,1)	66.42364
(1,2)	66.51769

After overfitting, the best model formed is VARX (1,1), with an AIC value of 66.42364. Then, the residual test for the White Noise test obtained a p-value of 0.127, 0.14, and 0.2586, respectively;

this shows that the p-value of the three variables is more significant than 0.05, meaning that the data meets the White Noise assumption. The multivariate regular residual test obtained a p-value of 0.208, 0.3625, and 0.7911, respectively; this shows that the p-value of the three variables is more significant than 0.05, meaning that the residual data is multivariate normally distributed. So, it is concluded that the best model meets the diagnostic testing assumptions.

Model VARX

The estimated equation for the VARX (1,1) model for cayenne pepper, red onions, and tomato variables as endogenous variables and rainfall as an exogenous variable is:

$$\begin{pmatrix} \Delta Cayenne\ Pepper_t \\ \Delta Red\ Onion_t \\ \Delta Tomato_t \end{pmatrix} = (-345.179\ \ -228339)\begin{pmatrix} \Delta Y_{1_{t-1}} \\ \Delta Y_{2_{t-1}} \\ \Delta Y_{3_{t-1}} \end{pmatrix} + \begin{pmatrix} -0.112 & -0.654 & 8.488 \\ -0.006 & 0.500 & 0.784 \\ 0.001 & -0.004 & -0.119 \end{pmatrix}\begin{pmatrix} \Delta Y_{1_{t-1}} \\ \Delta Y_{2_{t-1}} \\ \Delta Y_{3_{t-1}} \end{pmatrix}$$

Forecasting

The results of forecasting the amount of production of chilies, onions, and tomatoes from January 2024 to December 2024 can be seen in Table 6 below:

Table 5. Results of Forecasting the Amount of Production

Period	Forecast		
	Cayenne Pepper	Red Onions	Tomato
January	102677.4	138352.84	35876.19
February	113922.17	122997.24	36110.17
March	124596.33	108971.37	36369.96
April	133809.08	96739.56	36583.09
May	142118.86	85775.81	36778.57
June	149455.18	76093.18	36949.42
July	155998.44	67478.31	37102.01
August	161804.28	59842.92	37236.9
September	166967.91	53063.63	37356.71
October	171554.2	47050.51	37462.87
November	175629.83	41715.01	37557.04
December	179250.01	36982.22	37640.51

4. CONCLUSIONS

Based on the results and discussion of the general description that occurs in the production of cayenne pepper, shallots, and tomatoes experienced an increase and decrease, which is influenced by the amount of rainfall. The best model estimate formed based on the smallest AIC value is VARX (1,1) with the forecast results, namely the amount of production of these three commodities experiencing differences in the amount of output where an increase over the next year occurs in two types of plants, namely cayenne pepper and tomatoes while onions experience a decrease over the next year.

5. ACKNOWLEDGMENTS

Thank you to the Learning Development and Quality Assurance Institute (LP3M), Hamzanwadi University, for funding this research through the 2024 Internal Grant Scheme.

6. REFERENCES

[1] Faisal, A. (2022). "Analysis Of The Economic Potential Of Rice, Corn, Shallot, And Chili Commodities In Mojokerto Regency''upn Veteran East Java.

[2] Rahayu, S. (2021). "Determination of Leading Agribusiness of Agricultural Commodities Based on Production Value in Kerinci Regency". J-Mas (Journal of Management and Science), 6(1), 154-159.

[3] Alifia, H. (2024). Analysis of the Leading Sector of the Economy and Development Strategies Based on Leading Commodities in West Pesisir Regency as a Disadvantaged Regency. University of Lampung.

[4] Widya, H. (2023). Analysis of the Work Ethic of Rubber Farmers in Meeting Family Needs from an Islamic Economic Perspective (Study in Srikaton Village, Tanjung Bintang District, South Lampung Regency). Uin Raden Intan Lampung.

[5] Agriculture and Plantation Office of NTB Province. (2020). Recapitulation of Harvest Area, Productivity, and Tomato Production in NTB.

[6] Zainal, A., Faisal Rizal, Z., Ahmad, F., & Ikawidjaja, I. (2023). The Role of Horticultural Commodities in Inflation Control in Makassar City.

[7] Firdausi, N. T., & Kodoatie, J. M. (2010). Projections of Poverty Levels in Indonesia (Case Study: 30 Provinces). Diponegoro University.

[8] El Vinie, R. A., Martha, S., & Rizki, S. W. (2019). Vector Autoregressive Exogenous (VARX) Model in Predicting Rubber Production of PTPN XIII West Kalimantan Province. Bimaster: Scientific Bulletin of Mathematics, Statistics and Applied, 8(4).

[9] Fajriah, Z. I. (2022). Exogenous Autoregressive Vector (VARX) in Forecasting Covid-19 Cases in Indonesia.

[10]]Septiani, A., Sumertajaya, I. M., & Aidi, M. N. (2019). Vector autoregressive X (VARX) modeling for Indonesian macroeconomic indicators and handling different time variations with cubic spline interpolation. Repositories-Dept. of Statistics, IPB University, 175-180.

[11] Utami, B. (2018). VARX Model for Inflation Forecasting by Commodity Subgroup in Jakarta with TSClust Approach as Preprocessing (Doctoral dissertation, Bogor Agricultural University (IPB)).

[12] Ariyan, H. (2022). Policy Analysis of the Use of Foreign Debt in Indonesia's Economic Development Viewed from the Perspective of Sharia Economics (Doctoral Dissertation of Raden Intan Lampung State Islamic University).

[13]] Purnama, J., & Juliana, A. (2020). Prediction analysis of the composite stock price index using the ARIMA method. Cakrawala Management Business Journal, 2(2), 454-468.

[14] Idrus, R. A., Ruliana, R., & Aswi, A. (2022). Application of Singular Spectrum Analysis Method in Forecasting the Quantity of Rice Production in Gowa Regency. Variance: Journal of Statistics and Its Application on Teaching and Research, 4(2), 49-58.

[15] Irma, D. (2023). "Application of Hybrid Vector Autoregressive Exogenous-Convolutional Neural Network (VARX-CNN) Method for Forecasting Indonesian Import Exports".