# Cracking the Code of Crop Growth: Illuminating the Future of Philippines' Onion Production for a Resilient Filipino Diet with the ARMA Forecasting Model

**Caryl Vivien C. Capiral, Ruth Jane T. Lotrinia, Romie C. Mabborang, Josephine R. Macasieb**

Pamantasan ng Lungsod ng Maynila (University of the City of Manila)

doi: https://doi.org/10.37745/ejcsit.2013/vol11n3123 Published June 17 2023

**Citation**: Capiral C.V.C., Lotrinia R.J.T., Mabborang R.C., Macasieb J.R. (2023) Cracking the Code of Crop Growth: Illuminating the Future of Philippines' Onion Production for a Resilient Filipino Diet with the ARMA Forecasting Model, *European Journal of Computer Science and Information Technology*, Vol.11, No.3, pp.1-23

**ABSTRACT:** *This study employed the Box-Jenkins methodology and the Autoregressive Moving Average (ARMA) model to forecast onion production in the Philippines. By utilizing historical data from the Philippine Statistics Authority, an optimal forecasting solution was achieved through the selection of the ARMA (4,2) model. The model demonstrated a favorable fit, passing diagnostic tests and exhibiting a mean absolute percentage error (MAPE) of 10.406%. Projections for onion production in 2023 and 2024 were provided, highlighting expected yields for each quarter. The analysis of historical data revealed periodic fluctuations in onion supply driven by factors such as weather patterns, market demand, agricultural practices, and imports or exports. The study's implications emphasize the value of accurate forecasting models for decision-making in production planning, resource allocation, pricing, and market positioning. Policymakers, farmers, and stakeholders can utilize the findings to optimize onion production sustainably and enhance the agricultural sector's performance in the Philippines.*

**KEYWORDS:** autoregressive moving average (ARMA) model, Box-Jenkins analysis, log transformation, onion production, onion supply, agriculture, variance, variance stabilization, forecast, python

## INTRODUCTION

Filipino cuisine is a mouth-watering tapestry of flavors, each with a unique blend of taste and aroma. The onion is at the heart of this culinary symphony, adding depth and character to every meal. Apart from their culinary prowess, onions offer a treasure trove of nutritional benefits, making them an essential component of the Filipino diet. Nevertheless, recent events have raised concerns about the nation's food security as the availability and affordability of this indispensable vegetable have been thrust into the spotlight.

Supported by the enlightening findings of the National Onion Association (2020), onions emerge as a nutritional powerhouse, brimming with dietary fiber, vitamins C and B6, folate, and potassium. These nutrients enhance the flavor profile of dishes and promote overall health and well-being, underscoring their significance in the Filipino gastronomic landscape.

Fascinatingly, the Philippines proudly assumes the role of an onion aficionado, consuming an astonishing 17,000 metric tons every month, as unveiled by Hutchinson (2022) in the esteemed Plantwise Blog. However, a troubling pattern emerges from the annals of recent history. The final quarter of 2022 witnessed a disconcerting undersupply of onions, leading to an unprecedented price surge. According to the Department of Agriculture's (2023) price monitoring, the cost per kilo of onions skyrocketed to a range of ₱500.00 to ₱700.00 ($9-$13) during the first week of 2023. This alarming shortage highlights the inadequacy of onion production, leaving the nation grappling with an insufficient supply to meet its soaring demand.

The implications of this crisis extend far beyond the realm of culinary inconvenience. In a country where the daily minimum wage for laborers stands at ₱570.00, as affirmed by the Department of Labor and Employment (2022), unreasonable onion prices place an immense burden on the livelihoods of hardworking individuals. Consider the math: to provide three meals a day, each including one to two onions, and factoring in the current onion price of 500.00 pesos per kilo, with an approximate yield of 13-15 pieces per kilo, the cost of a single onion amounts to a staggering 33.3 pesos. With such financial constraints, laborers find themselves hard-pressed to afford nutritious meals for their families, exacerbating the plight of food insecurity across the nation.

Moreover, if left unchecked, this protracted scarcity of onions could have dire consequences for the country's food security. The potential disruption looms large, posing a catastrophic crisis that demands immediate attention and intervention.

In response to these pressing circumstances, a team of dedicated researchers has embarked on a pioneering endeavor—a forecasting study designed to unravel the mysteries of onion production. Armed with the Box-Jenkins Analysis method, they seek to construct an Autoregressive Moving Average Model (ARIMA Model) based on historical data on onion production in the Philippines. Their mission: to predict onion production for the next two years, divided into quarters. By illuminating the future landscape of onion production, this forecast will empower the government to proactively implement preemptive measures, averting the impending crisis and ensuring the nation's resilience.

As we delve into the depths of onion production dynamics, this groundbreaking study aims to unlock the hidden potential within the agricultural landscape. By charting a path toward sustainable onion production, we can fortify our food security, secure affordable access to essential nutrients, and pave the way for a brighter future for all Filipinos.

European Journal of Computer Science and Information Technology 11 (3), 1-23, 2023
Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

*A. Conceptual Framework*

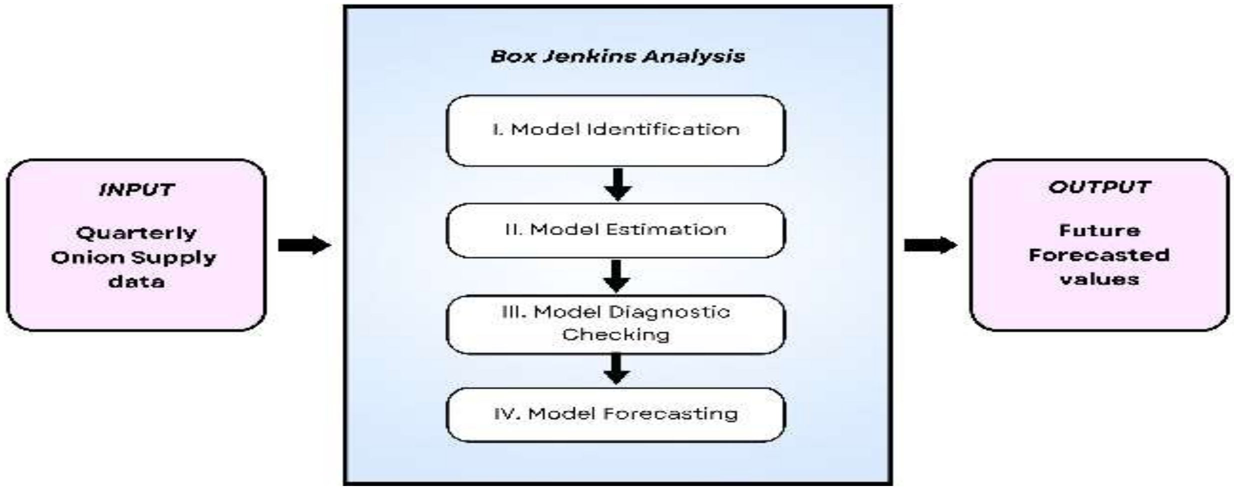


Figure 1.1 : Conceptual Framework of the Study

This study utilized historical data obtained from the Philippine Statistics Authority website to conduct an in-depth analysis and forecasting of quarterly onion production in the Philippines. The Box-Jenkins methodology was employed to identify and develop the most appropriate Autoregressive Moving Average (ARMA) model necessary for accurate prediction of the data. The selection of the best-fitting model was based on selection criteria to ensure its suitability for forecasting purposes. The forecasted onion production data was subjected to an accuracy test, specifically the Mean Absolute Percentage Error (MAPE), to assess the model's performance in predicting future values. This evaluation serves as a validation process to determine the model's reliability and its ability to provide accurate forecasts of onion production in the Philippines.

*B. Statement of the Problem*

This paper aims to address the following research questions:

(1) What is the optimal ARMA model that demonstrates the most favorable fit in forecasting the onion production in the Philippines?

(2) What is the effectiveness of the chosen ARMA model in forecasting onion production in the Philippines?

(3) Based on the selected ARMA model, what are the projected values for the following periods in 2023 and 2024:

a. the first quarter (January to March),
b. the second quarter (April to June),

c. the third quarter (July to September), and
d. the fourth quarter (October to December)?

(4) What is the behavior of the onion supply in the Philippines based on historical data? If the behavior of the data exhibits periodic fluctuations, what are the drivers behind it?

(5) What are the implications of the study's results in the Philippine agricultural industry?

(6) How can onion production in the Philippines be optimized to meet market demands sustainably?

(7) How can the findings of the research be utilized by stakeholders in the agricultural sector of the Philippines?

## RELATED LITERATURE

In this review of related literature, we explore various studies that provide valuable insights into onion production, time series analysis, and forecasting. By examining these studies, we aim to identify gaps and differences that highlight the significance of the present study. The literature encompasses the critical role of onions in cooking, the challenges faced in onion production, elements of time series analysis, time series forecasting, and the autoregressive moving average (ARMA) model.

A. Onion

Onion is a crucial ingredient in cooking, adding flavor and aroma to dishes. Beckett (2019) and Munson (2022) emphasize the significance of onions in enhancing the taste of sauces, garnishes, stocks, and stews. These articles highlight the critical role of onions in the cooking process and their contribution to the overall flavor profile.

B. Onion Production

The Philippines is the third-largest onion producer in ASEAN, with significant land dedicated to onion cultivation. However, the country faces challenges related to onion supply and the economy. Yu (2023) discusses the controversial issue of onion supply, including the market's support for imported onions and the impact on local farmers. Panti (2023) raises concerns about the safety of imported onions and the need for storage facilities to improve production.

C. Elements of Time Series

Time series analysis involves understanding the dependence mechanism and forecasting future values based on deterministic and random components. Guerrier et al. (2019) explain the decomposition of time series data using Wold's Decomposition Theorem. The analysis considers

European Journal of Computer Science and Information Technology 11 (3), 1-23, 2023

Print ISSN: 2054-0957 (Print),

Online ISSN: 2054-0965 (Online)

Website: https://www.eajournals.org/

Publication of the European Centre for Research Training and Development -UK

mean changes, variance, state changes, outliers, long-run cycles, and abrupt changes in the series. Time series analysis aims to identify trends and patterns for forecasting, without necessarily uncovering the underlying causes.

*D. Time Series Forecasting and its Properties*

Forecasting time series involves analyzing trends, seasonality, and cyclicality. Athanasopoulos and Hyndman (2018) discuss the importance of adjustments and transformations to enhance forecast accuracy. Autocorrelation function (ACF) and partial autocorrelation function (PACF) plots help detect patterns and randomness in the series. Residual diagnostics, such as the Ljung-Box test, ensure the accuracy and reliability of time series models.

*E. Autoregressive Moving Average (ARMA) Model*

The ARMA model is a popular choice for time series analysis, involving autoregressive (AR) and moving average (MA) components. It assumes stationarity in the data and requires appropriate differencing. ARIMA models incorporate the integrated (I) component to handle non-stationary data. SARIMA models account for seasonal patterns in the data. Stationarity tests, such as the Augmented Dickey-Fuller (ADF) test, help determine the appropriate model for forecasting.

*F. Related Studies*

Studies related to onion production, crop forecasting, and Philippine agriculture provide valuable insights. Gavino Jr. (2020) explores land suitability for agriculture in Occidental Mindoro, offering potential solutions to the land resource problem. Pascual et al. (2018) propose vertical farming using a hydroponic system to increase onion production. Balilla (2023) forecasts retail native onion prices, providing insights for the coming years. Amin et al. (2021) forecast onion production in Pakistan using ARIMA models. Diaz, Mingo, and Urrutia (2017) and Handa et al. (2023) employ Box-Jenkins approaches for crop forecasting in the Philippines.

*G. Synthesis*

The literature review reveals the critical role of onions in cooking and the challenges faced by the agricultural sector in onion production. Limited research specifically focuses on onion supply in the Philippines. The present study aims to contribute to the understanding of onion supply forecasting by employing time series analysis and considering the factors highlighted in the literature. By addressing the gaps and utilizing the relevant methodologies, this research strives to find solutions to challenges such as price fluctuations, supply shortages, and the need for storage facilities.

## METHODOLOGY

The Box-Jenkins model, depicted in Figure 3.1, is a powerful and widely used methodology in

Print ISSN: 2054-0957 (Print),

Online ISSN: 2054-0965 (Online)

Website: https://www.eajournals.org/

Publication of the European Centre for Research Training and Development -UK

time series analysis and forecasting. This dynamic framework, known as Autoregressive Moving Average (ARMA) model, captures the intricate interplay between past observations and stochastic errors to unlock hidden patterns and make accurate predictions. With its statistical rigor and adaptability, the Box-Jenkins model has become an indispensable tool for researchers and analysts seeking to unravel the complexities of temporal data and glimpse into the future.

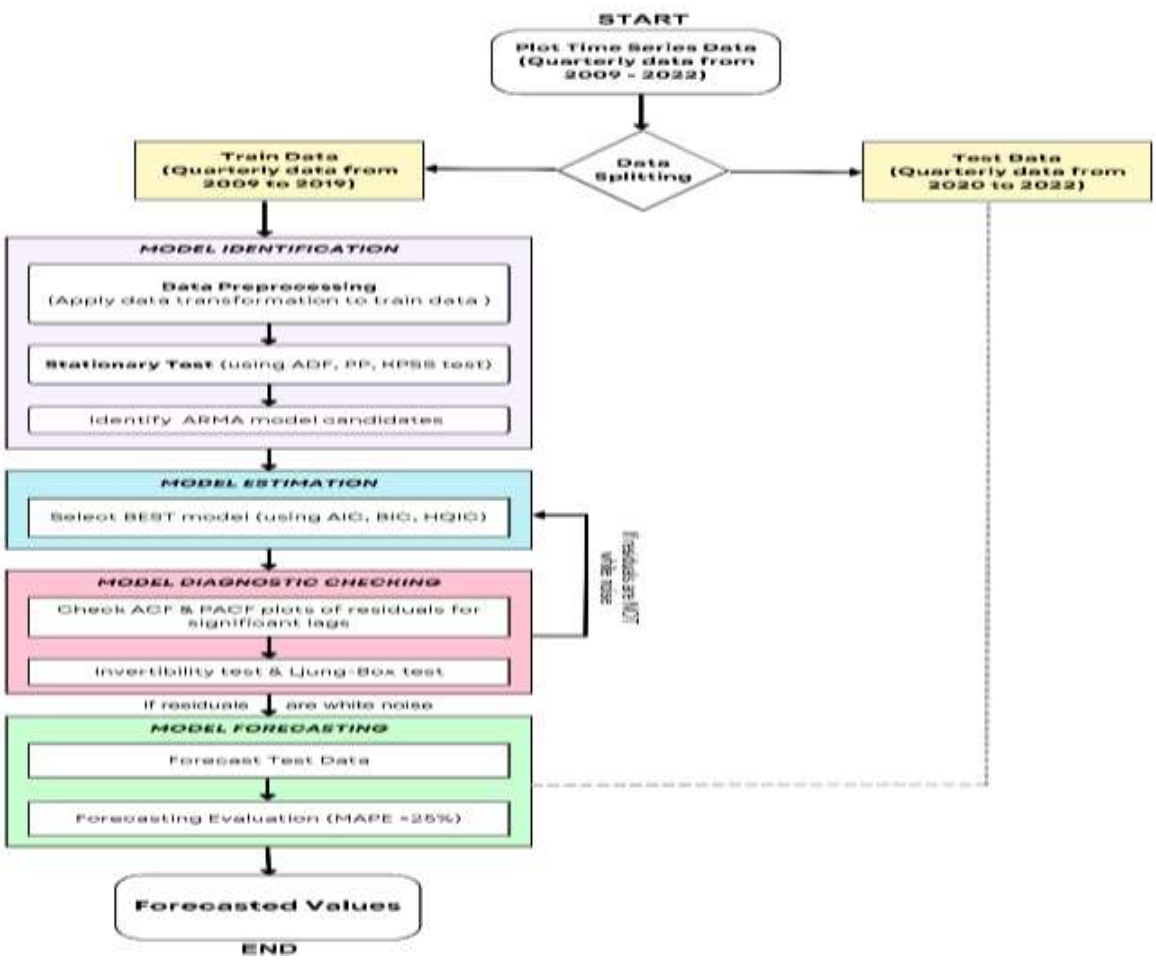


*Figure 3.1: Box-Jenkins Methodology – ARMA model*

*A. Research Design*

This study employed the Box-Jenkins method, specifically the Autoregressive Moving Average (ARMA) model, to forecast the quarterly onion supply in the Philippines. The research process consisted of four main phases: model identification, model estimation, model diagnostic checking, and model forecasting.

*Model Identification:*

The time series data was carefully examined and plotted to identify any noticeable patterns or trends. A logarithmic (log) transformation was applied to address variance issues and promote stationarity. Stationary tests, including the Augmented Dickey Fuller (ADF), Phillips-Perron (PP), and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests, were conducted to assess stationarity. Based on significant spikes in the Partial Autocorrelation Function (PACF) and Autocorrelation Function (ACF) plots, ARMA candidate models were identified.

*Model Estimation:*

The optimal model was determined by comparing the values of statistical measures such as the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan-Quinn Information Criterion (HQIC). The model with the lowest values of these criteria was considered the best choice, indicating a balance between model fit and complexity.

*Model Diagnostic Checking:*

The presence of overfitting was assessed by examining the residuals' autocorrelation function (ACF) and partial autocorrelation (PACF) plots. Significant lags in these plots would indicate autocorrelation and suggest the need for modifications to address overfitting. Invertibility tests and the Ljung-Box test were conducted to ensure the stability of the model's components and assess the presence of autocorrelation in the residuals.

*Model Forecasting:*

The selected optimal model was applied to forecast the test data. The forecast accuracy was evaluated using the Mean Absolute Percentage Error (MAPE), which compares the test forecast to the actual values. A MAPE of less than 25% indicates a higher level of accuracy within an acceptable range. Once the model achieved an acceptable MAPE, it was considered ready for forecasting future values.

*B. Data Collection and Procedure*

The dataset used in this study was obtained from the official website of the Philippine Statistics Authority (PSA), specifically from the "Major Vegetable and Rootcrops Quarterly Bulletin" available in the PSA's publication catalog. A total of 56 data points covering the period from 2009 to 2022 were gathered and organized into an Excel file, which was then exported to a CSV file format for analysis.

Preprocessing steps were executed to ensure data quality and consistency. These steps involved handling missing values, resolving data inconsistencies, and formatting the data appropriately. The Visual Code software tool was used for data analysis and programming.

Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

Python was chosen as the programming language for its extensive libraries for time series analysis and forecasting. Libraries such as Pandas, NumPy, Scikit-learn, Arch, Matplotlib, and Seaborn were utilized for data manipulation, statistical analysis, mathematical calculations, and data visualization. These libraries played a crucial role in preprocessing the time series data and applying the Box-Jenkins methodology for forecasting the onion supply in the Philippines.

By implementing this methodology, the study aimed to identify the optimal ARMA model for forecasting onion production, assess its effectiveness, project values for specific periods, analyze the behavior of onion supply, determine implications for the agricultural industry, optimize onion production sustainably, and provide practical utilization of the research findings by stakeholders in the Philippine agricultural sector.

**RESULTS AND DISCUSSION**

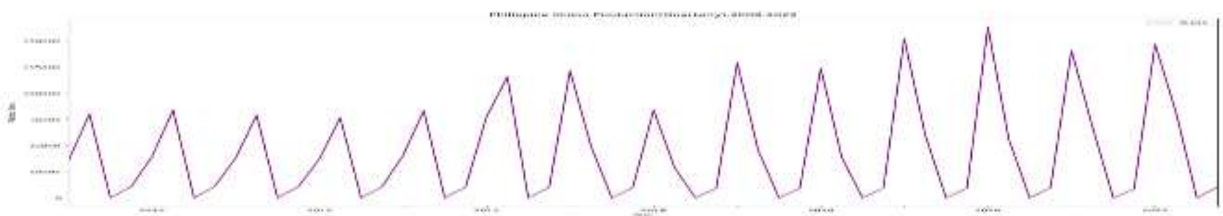

**Figure 4.1** : *Philippine Onion Production Time Series Plot (2009-2022)*

The onion production time series data was first plotted to observe the characteristics of the data. Based on figure 4.1, the plot reveals a possible absence of an upward or downward trend, indicating a relatively stable or constant mean. However, the high and low fluctuations in the figure showed changing variances over time. In addition, discernible periodic patterns can be observed from the plot, implying the presence of seasonality.

**PHASE 1: Model Identification**

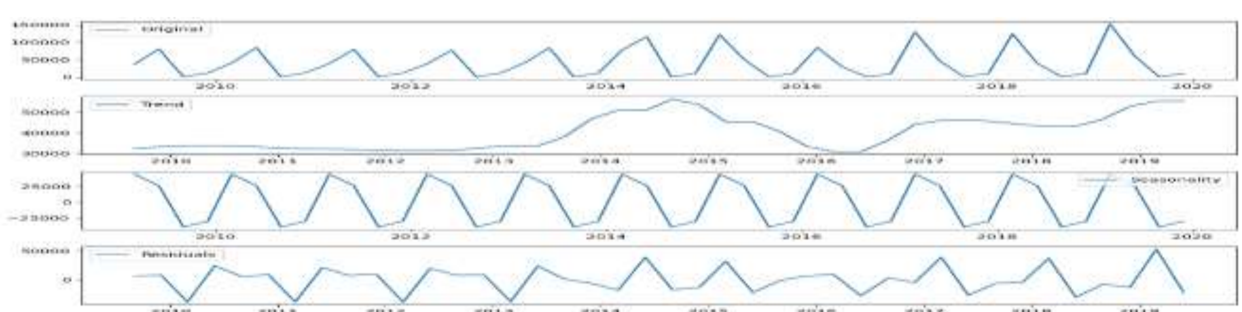


Figure 4.1.1: Decomposition of Philippine Onion Production Time Series into trend and seasonal components.

European Journal of Computer Science and Information Technology 11 (3), 1-23, 2023

Print ISSN: 2054-0957 (Print),

Online ISSN: 2054-0965 (Online)

Website: https://www.eajournals.org/

Publication of the European Centre for Research Training and Development -UK

The top figure in the illustration (figure 4.2) presents the original time series at level. From the trend component, it can be observed that there is no continuous increase or decrease in the data, indicating the absence of a trend from the data. From the third component – the seasonal component, periodic patterns confirm the existence of seasonality in the data. The last component represents the residuals of the time series after removing the trend and seasonal components.

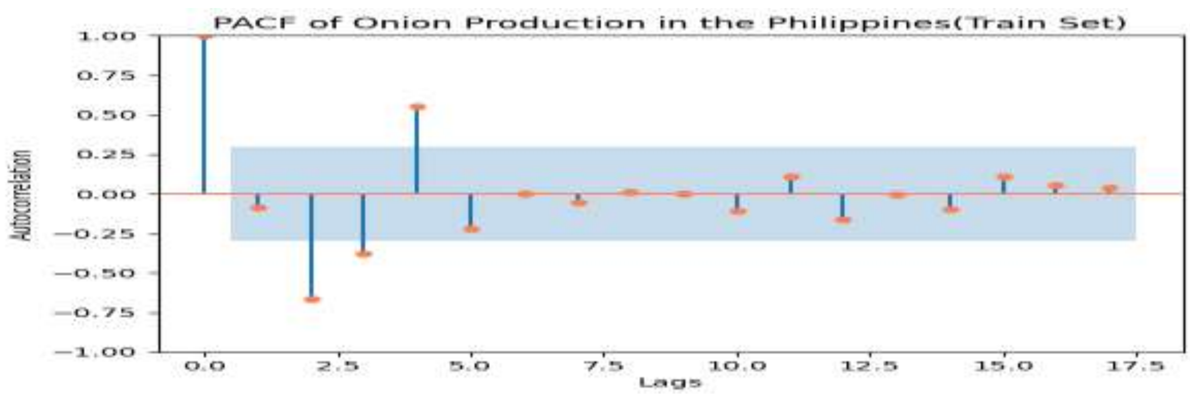


Figure 4.1.2: PACF plot of Time Series Data (Train Set)

The blue-shaded area in the ACF and PACF plots illustrates the 95% confidence interval or the error bounds; lags within the area are statistically close to zero (Monigatti, 2022). The spike at lag 0 shows a correlation value of 1. It indicates a high correlation between the time series data and itself. In Figure 4.1.2, which represents the partial autocorrelation function (PACF) plot, it can be observed that lags 2, 3, and 4 are the significant lags that extend outside error bounds. At the same time, significance cuts off starting from lag 5 to 17.

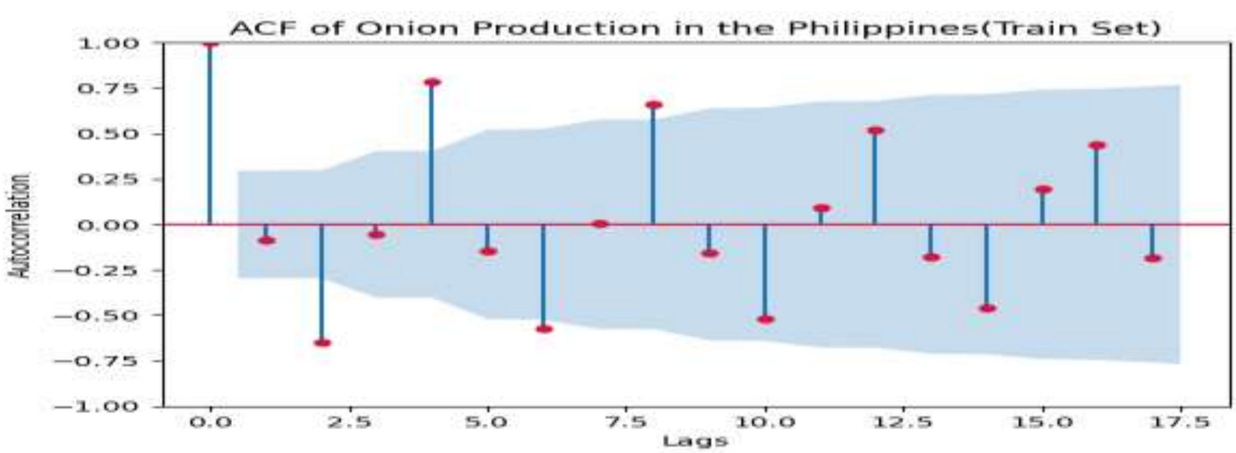


**Figure 4.1.3:** *ACF plot of Time Series Data (Train Set)*

In Figure 4.1.3, it can be observed that the autocorrelation plot exhibits significant autocorrelation values at lags 2, 4, 6, and 8, which extend beyond the 95% confidence interval. Furthermore, the significant lags of the autocorrelation demonstrate periodic oscillations every 2nd lag. These oscillations indicate the presence of seasonality in the data, implying that the observations exhibit a recurring pattern at fixed intervals. The periodic nature of the oscillations strengthens the notion of seasonal behavior in the dataset, which violates one of the stationarity conditions – having no seasonal characteristic. Therefore, based on the visual inspection of the PACF and ACF plots, the time series data is non-stationary.

**Stationary Tests (ADF test, PP test, KPSS test)**

**Table 4.1.4**

*ADF test of Onion Production Time Series (Train Set)*

```
   Augmented Dickey-Fuller Results
=====================================
Test Statistic                 -1.316
P-value                         0.622
Lags                                3
-------------------------------------

Trend: Constant
Critical Values: -3.61 (1%), -2.94 (5%), -2.61 (10%)
Null Hypothesis: The process contains a unit root.
Alternative Hypothesis: The process is weakly stationary.
```

The first statistical test employed on the train data was the Augmented Dickey-Fuller (ADF) test. Upon conducting the test (table 4.1.4), the obtained p-value of 0.622 exceeded the predetermined significance level of 0.05. With this, there is no strong evidence to reject the null hypothesis associated with the ADF test, thereby suggesting the presence of a unit root. Hence, it is deduced that the series exhibits non-stationarity.

**Table 4.1.5**

*PP test for Onion Production Time Series (train set)*

```
     Phillips-Perron Test (Z-tau)
=====================================
Test Statistic                 -8.980
P-value                         0.000
Lags                               10
-------------------------------------

Trend: Constant
Critical Values: -3.59 (1%), -2.93 (5%), -2.60 (10%)
Null Hypothesis: The process contains a unit root.
Alternative Hypothesis: The process is weakly stationary.
```

Contrary to the ADF test result, the Phillips-Peron (PP) test obtained a p-value of zero which is less than the significance level of 0.05. Therefore, there is strong evidence to reject the null hypothesis that a unit root exists, indicating that the series is stationary.

**Table 4.1.6**

*KPSS test for Onion Production Time Series (train set)*

```
1. KPSS Test Statistitics :   0.20817122126606497
2. P-Value:   0.1
3. Num Of Lags :   13
4. Critical Values :
         10% : 0.347
         5% : 0.463
         2.5% : 0.574
         1% : 0.739
```

In contrast to the ADF and PP tests, the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test is a statistical test in which the null hypothesis in the KPSS test suggests that if the p-value exceeds 0.05, the series can be considered stationary. In the results of the conducted KPSS test (figure 4.5), the series is stationary since the p-value = 0.10 > 0.05.

Due to the non-stationary result obtained from the Augmented Dickey-Fuller (ADF) test, while the Phillips-Perron (PP) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests yielded stationary outcomes, it is essential to enhance the stationarity of the time series through data transformation. Another keynote, the inspection of the time series plot reveals the presence of unstable variance, requiring the need for log transformation. This type of transformation serves the purpose of stabilizing the time series' variance and improving the stationary test results. Notably, non-constant variance in a time series indicates the presence of heteroskedasticity, which violates the assumption of constant variance. Such violations can affect the accuracy of forecasts conducted in time series analysis. Thus, log transformation plays a crucial role in addressing unstable variance and mitigating the impact of heteroskedasticity in the time series data.

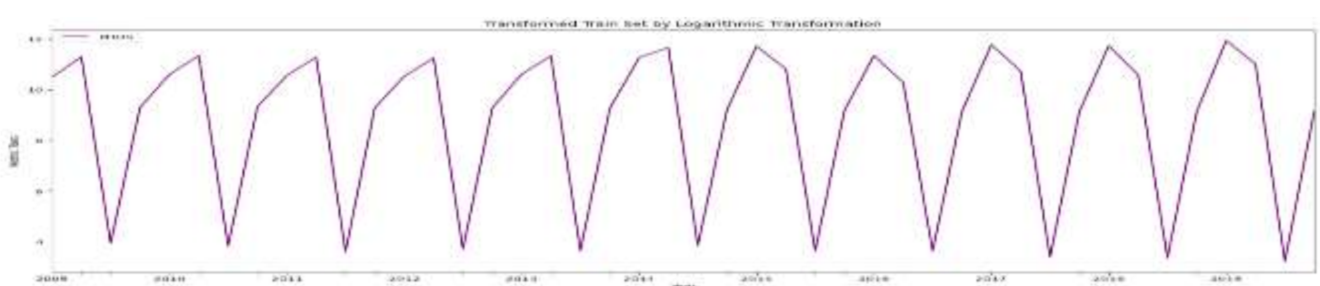


Figure 4.1.7 Philippine Onion Production Time Serie Data (Transformed)

European Journal of Computer Science and Information Technology 11 (3), 1-23, 2023

Print ISSN: 2054-0957 (Print),

Online ISSN: 2054-0965 (Online)

Website: https://www.eajournals.org/

Publication of the European Centre for Research Training and Development -UK

Figure 4.1.7 displays the resulting plot after log transformation on the initial time series data. The time series variance has stabilized, indicating that the series satisfies the assumption of stationarity by exhibiting constant variance.

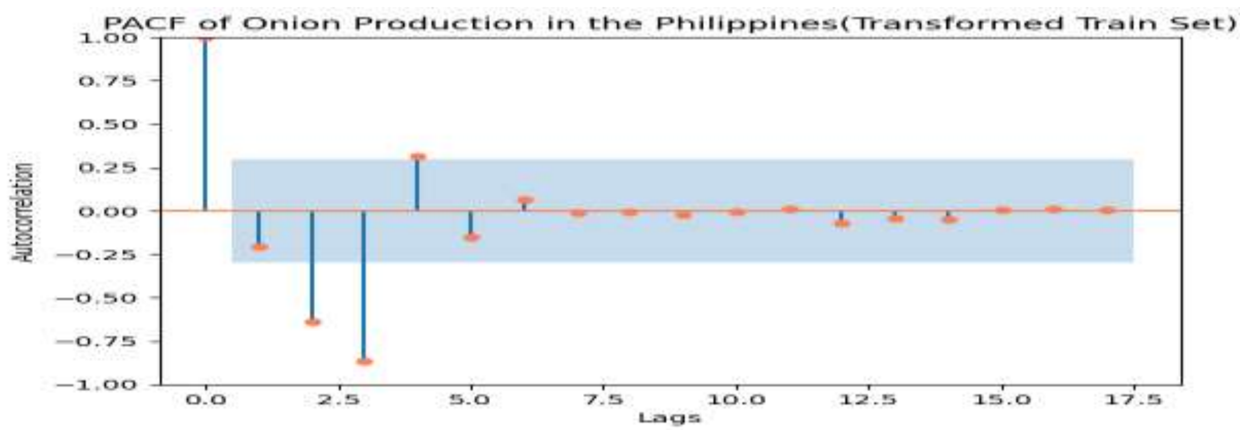


Figure 4.1.8 PACF Plot of Transformed Philippine Onion Production

Inspecting the PACF plot of the transformed data, figure 4.1.5, the partial autocorrelation function has three significant lags – lag 2, 3, and 4. These lags were also significant prior to transforming the data.

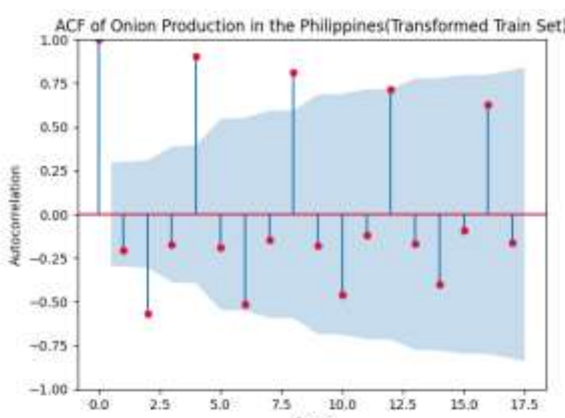


Figure 4.1.9: ACF Plot of Transformed Philippine Onion Production

After transforming the data, the number of significant lags on the ACF plot (figure 4.1.9) has decreased to three. These significant lags are now lags 2, 4, and 8. However, periodic fluctuations can still be observed from the plot, indicating the seasonal characteristic of the transformed data.

**Table 4.1.10**

*ADF test of Onion Production Time Series (Transformed Train Set)*

```
   Augmented Dickey-Fuller Results
=====================================
Test Statistic                 -3.172
P-value                         0.022
Lags                                6
-------------------------------------

Trend: Constant
Critical Values: -3.62 (1%), -2.94 (5%), -2.61 (10%)
Null Hypothesis: The process contains a unit root.
Alternative Hypothesis: The process is weakly stationary.
```

The ADF test on the transformed data yielded a p-value of 0.022, which falls below the significance level of 0.05. As a result, sufficient evidence was found to reject the null hypothesis that a unit root exists, indicating that the time series data is stationary.

**Table 4.1.11**

*PP test for Onion Production Time Series (Transformed Train Set)*

```
     Phillips-Perron Test (Z-tau)
=====================================
Test Statistic                -15.494
P-value                         0.000
Lags                               10
-------------------------------------

Trend: Constant
Critical Values: -3.59 (1%), -2.93 (5%), -2.60 (10%)
Null Hypothesis: The process contains a unit root.
Alternative Hypothesis: The process is weakly stationary.
```

The PP test was conducted on the transformed data, yielding a p-value of zero, which falls below the significance level of 0.05. Therefore, sufficient evidence was found to reject the null hypothesis that a unit root exists, indicating that the time series data is stationary.

**Table 4.1.12**

*KPSS test for Onion Production Time Series (Transformed Train Set)*

```
1. KPSS Test Statistics:  0.44411897905978065
2. P-Value:  0.05813837109492214
3. Num Of Lags :  11
4. Critical Values :
         10% : 0.347
         5% : 0.463
         2.5% : 0.574
         1% : 0.739
```

The KPSS test conducted on the transformed data yielded a p-value of 0.0581, which falls beyond the significance level of 0.05. Therefore, sufficient evidence was found to accept the null hypothesis; thus, the series is stationary.

The stationary tests on the transformed series confirmed that the data now demonstrates stationarity. Hence, the need for differencing procedures was deemed unnecessary to conduct.

**Model Candidate Identification**

Following the stationary tests above, potential ARIMA models were selected based on significant lag orders observed in the partial autocorrelation function (PACF) plot and the autocorrelation function (ACF) of the transformed train data. The autoregressive (AR) components were determined by considering significant 2, 3, and 4 lag orders from the PACF plot (Figure 4.1.8). Conversely, the ACF plot's lags 2, 4, and 8 (Figure 4.1.9) exceeded the error bounds, indicating potential lag components for the moving average (MA) part of the ARIMA model combinations. Since differencing was not performed, the ARIMA models' integrated component ('I') is set to zero.

The following ARIMA model combinations are considered potential representations for onion production in the Philippines: ARIMA (2, 0, 2), ARIMA (2, 0, 4), ARIMA (2, 0, 8), ARIMA (3, 0, 2), ARIMA (3, 0, 4), ARIMA (3, 0, 8), ARIMA (4, 0, 2), ARIMA (4, 0, 4), ARIMA (4, 0, 8).

**PHASE 2: Model Estimation**

**Table 4.2.1**

*Candidate ARIMA Models with corresponding Selection Criteria*

| ARIMA MODELS | LOGLIKEHOOD | AIC | BIC | HQIC |
|---|---|---|---|---|
| ARIMA (2,0,2) | -507.99527 | 1027.990534 | 1038.69568 | 1031.96052 |
| ARIMA (2,0,4) | -505.12131 | 1026.24262 | 1040.51614 | 1031.53593 |
| ARIMA (2,0,8) | -508.54290 | 1041.08580 | 1062.49608 | 1049.02576 |
| ARIMA (3,0,2) | -503.67448 | 1021.34895 | 1033.83828 | 1025.98060 |
| ARIMA (3,0,4) | -494.65372 | 1007.30743 | 1023.36514 | 1013.26240 |
| ARIMA (3,0,8) | -492.37961 | 1010.75922 | 1033.95368 | 1019.36084 |
| ARIMA (4,0,2) | -495.85579 | 1007.71159 | 1021.98511 | 1013.00490 |
| ARIMA (4,0,4) | -493.99366 | 1007.98733 | 1025.82922 | 1014.60396 |
| ARIMA (4,0,8) | -491.92799 | 1011.85597 | 1036.83463 | 1021.11926 |

Figure 4.2.1 presents the results of the model selection process, indicating that the optimal model for the time series data is ARIMA (4,0,2), as it demonstrates the lowest values of the Bayesian Information Criterion (BIC) and the Hannan-Quinn Information Criterion (HQIC). Additionally, ARIMA (3, 0, 4) and ARIMA (4,0,8) were models which only had lowest value of AIC and highest log-likelihood value, respectively. Before forecasting using the optimal model, it is crucial to conduct a diagnostic check on the model to ensure whether any significant remaining lags need to be captured from the optimal model's residuals.

Print ISSN: 2054-0957 (Print),

Online ISSN: 2054-0965 (Online)

Website: https://www.eajournals.org/

Publication of the European Centre for Research Training and Development -UK

**PHASE 3: Model Diagnostic Checking**

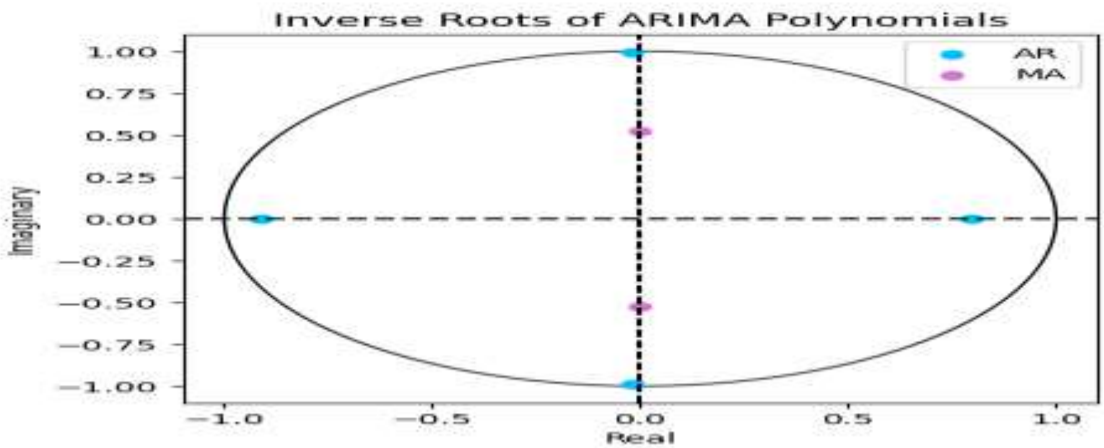


Figure 4.3.1 Inverse roots of ARIMA (4,0,2)

The optimal model ARMA (4,0,2) is considered stable because the inverse roots of the autoregressive (AR) components are within the unit circle. In other words, the absolute values of these roots are less than 1. This stability condition ensures that the model satisfies diagnostic checks and is reliable for forecasting.

**Ljung-Box Q Test**

**Table 4.3.2**

*ARIMA (4.0,2) Ljung – Box Q Test*

| | lb_stat | lb_pvalue |
|---|---|---|
| 8 | 5.29871 | 0.725224 |

As can be seen in table 4.3.2, the p-value is 0.725224 which is more than the rejection level of 0.05. Hence, from the conducted Ljung-Box test it is concluded that the residuals of ARIMA (4,0,2) are white noise, indicating that the optimal model fits the data. Thus, it is determined that the most suitable model for accurately forecasting the onion supply in the Philippines is ARMA (4,2).

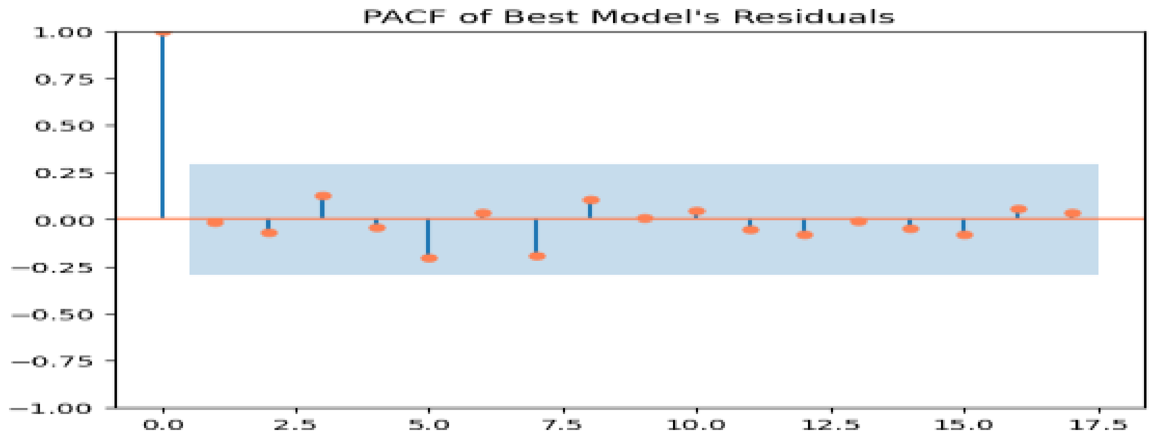


Figure 4.3.3 : PACF plot of ARIMA (4.0,2)'s Residual

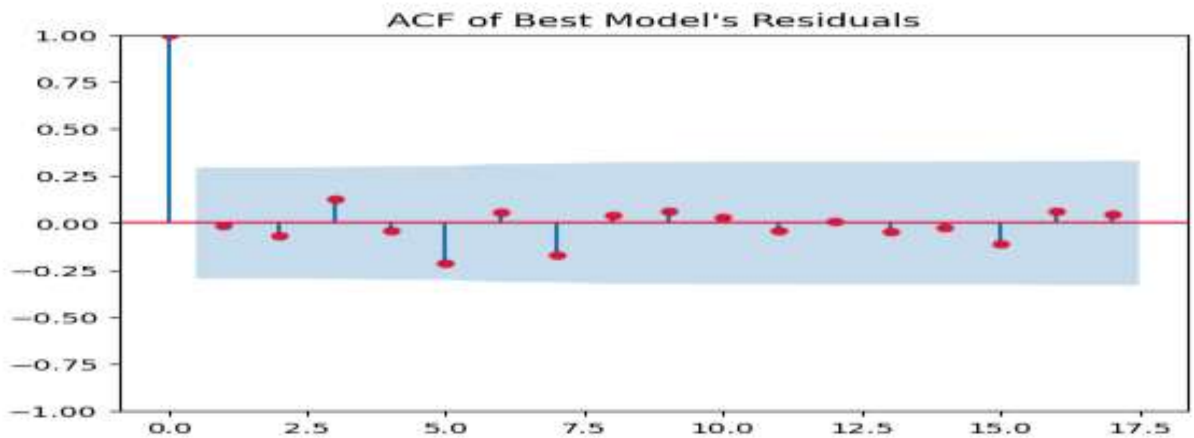


Figure 4.3.4 ACF plot of ARIMA (4.0,2)'s Residual

After examining the correlograms (Figure 4.3.3 and Figure 4.3.4) of the partial autocorrelation function (PACF) and autocorrelation function (ACF) of the selected ARIMA (4,0,2) model, no significant lags were observed. Hence, it can be inferred that the model's residuals exhibit white noise behavior, indicating the absence of autocorrelation within the model's residuals. Overall, diagnostic checking has been satisfactory. Thus, the ARIMA (4,0,2) can now be used for forecasting.

Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

## PHASE 4: Model Forecasting

### Forecasted Test Set

**Table 4.4.1** *Values of Forecasted Test Set (2020-2022)*

| | |
|---|---|
| 2020-03-01 | 153118.466579 |
| 2020-06-01 | 56395.998877 |
| 2020-09-01 | 25.317175 |
| 2020-12-01 | 9733.343711 |
| 2021-03-01 | 153131.548341 |
| 2021-06-01 | 53347.777920 |
| 2021-09-01 | 25.562755 |
| 2021-12-01 | 10028.640485 |
| 2022-03-01 | 153669.983254 |
| 2022-06-01 | 50623.588620 |
| 2022-09-01 | 25.876084 |
| 2022-12-01 | 10349.652249 |

Freq: QS-DEC, Name: predicted_mean, dtype: float64

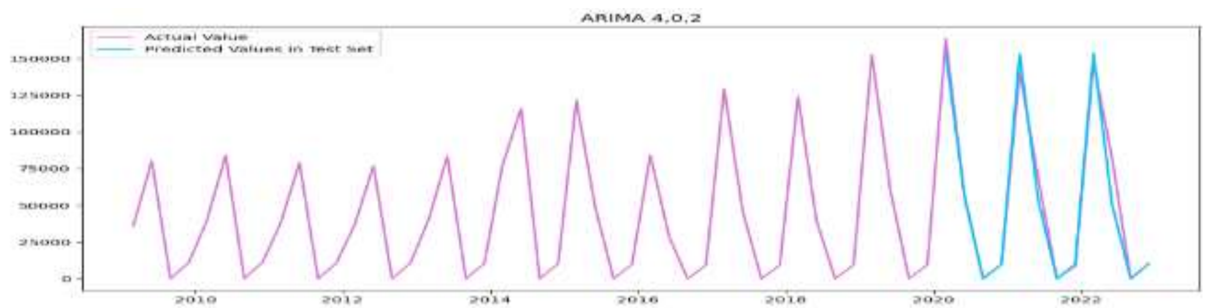


Figure 4.4.2 Plot of Actual data set vs. Forecasted (test set) Onion Production in the Philippines

Table 4.4.1 shows the forecasted onion supply data from the 1st quarter of 2020 until the 4$^{th}$ quarter of 2022 using the model ARIMA (4,0,2). In Figure 4.4.2, the plot shows the comparison between the actual data and forecasted data on onion production in the Philippines.

### Forecast Accuracy

**Table 4.4.3**

*Forecasted Accuracy of ARIMA (4,0,2)*

```
MAPE:0.10406426752764893
MAE:6470.807324401046
RMSE:125929188.10334933
```

European Journal of Computer Science and Information Technology 11 (3), 1-23, 2023

Print ISSN: 2054-0957 (Print),

Online ISSN: 2054-0965 (Online)

Website: https://www.eajournals.org/

Publication of the European Centre for Research Training and Development -UK

A model with a Mean Absolute Percentage Error (MAPE) of less than 25% is deemed to be acceptable. in contrast, a MAPE exceeding 25% indicates that the generated forecast is inaccurate, hence, unacceptable (Swanson, 2015). Table 4.4.3 reveals that the optimum model has a Mean Absolute Percentage Error (MAPE) of 10.41% which falls under the accepted range of MAPE.

**Sample Forecast (1st quarter of 2023 to 4th quarter of 2024)**

**Table 4.4.4**

*Sample Forecast from 2023 to 2024*

```
2023-03-01     154341.424509
2023-06-01      48095.967800
2023-09-01         26.220504
2023-12-01      10686.136892
2024-03-01     155014.858118
2024-06-01      45719.476003
2024-09-01         26.583922
2024-12-01      11034.735686
Freq: QS-DEC, Name: predicted_mean, dtype: float64
```

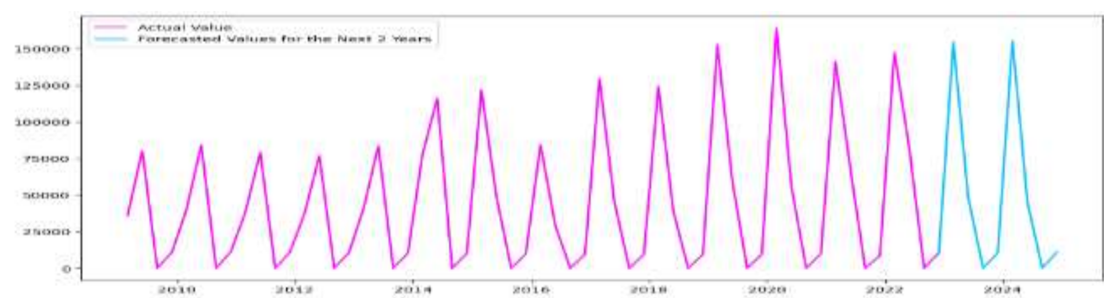


Figure 4.4.5 Actual values for 2009 to 2022 and Forecasted values for 2023 to 2024

The data in Table 4.4.4 are the generated out-of-sample (or future) forecasts using the optimal model, ARIMA (4,0,2). The model captured the unique character of the data, specifically the discernible periodic decrease in onion supply every third quarter of each year. Overall, ARIMA (4,0,2) yields a proper forecast of the quarterly onion production in the Philippines.

## CONCLUSION

This paper implemented the Box-Jenkins methodology to forecast onion production in the Philippines. The study generated a two-year forecast of eight quarters using the Autoregressive Moving Average (ARMA) model. The data were subjected to logarithmic (log) transformation to stabilize its variance characteristics.

Among the examined ARMA models, the ARMA (4,2) model demonstrated the best fit for forecasting. It was selected based on the lowest values of the Bayesian Information Criterion (BIC) and Hannan-Quinn Information Criterion (HQIC). The model passed diagnostic tests, including the invertibility test, the Ljung-Box test, and the inspection of PACF and ACF residual plots, confirming its suitability for forecasting purposes. The model captured and forecasted onion supply at a mean absolute percentage error (MAPE) of 10.406%, which is within the acceptable MAPE range of below 25%.

The projected values for the first, second, third, and fourth quarters of 2023 would be 154 341 metric tons, 48 096 metric tons, 26.22 metric tons, and 10 686.14 metric tons, respectively. For the first to the fourth quarter of 2024, the onion supply would be as follows: 155,015 metric tons, 45,719 metric tons, 26.58 metric tons, and 11,034.74 metric tons, respectively.

**Implications**

The findings of this research study have significant implications for the onion production sector in the Philippines, providing valuable insights and introducing the optimal model as a reliable forecasting tool. The accurate forecasts generated by the ARMA (4,2) model enable policymakers, farmers, and stakeholders to make informed decisions and allocate resources effectively. With a mean absolute percentage error (MAPE) of 10.406%, the model demonstrates its credibility in capturing and predicting onion supply trends within an acceptable range.

The study's projections for onion production in the upcoming quarters and years offer strategic planning insights. Stakeholders can use this information to align production strategies with market demands and make well-informed choices regarding crop planning, infrastructure investments, and marketing strategies. By analyzing historical data, the study also uncovers the underlying factors influencing onion supply fluctuations, such as seasonal weather patterns, market demand, agricultural practices, and the impact of imports and exports. This understanding enables sustainable management and stabilization of onion production to meet dynamic market demands.

Policymakers can use the findings to shape policies supporting onion production, fostering market stability, and addressing supply-demand imbalances. Farmers and agricultural organizations can leverage the insights to make informed decisions about crop planning, infrastructure and technology investments, and transformative marketing strategies. Overall, this research contributes to optimizing and growing the onion industry in the Philippines.

**Recommendations**

The study recommends several areas for future research to enhance the understanding and application of onion supply trends in the Philippines. The recommendations are as follows:

*Model Refinement*: Although the ARMA (4,2) model demonstrated exceptional accuracy, exploring alternative time series models such as SARIMA or other advanced forecasting techniques could provide additional insights. Comparing and evaluating different models may reveal superior forecasting capabilities or uncover new patterns in onion production data.

*Incorporating Exogenous Variables*: Future research could explore the integration of exogenous variables, such as socioeconomic indicators, climate change factors, or policy changes, to enhance the accuracy and robustness of forecasting models.

*Conducting Regional-Level Analyses*: The Philippines comprises diverse regions with varying climates, soil conditions, and market dynamics. Regional-level analyses and forecasting could provide region-specific insights for localized decision-making in onion production. Exploring the variations in factors affecting onion supply across regions would contribute to a more comprehensive understanding of the industry.

*Long-Term Forecasting*: Extending the forecasting horizon beyond the studied period can provide valuable insights for long-term planning and strategic decision-making. Investigating the impact of climate change, emerging technologies, and evolving market dynamics on onion production in the Philippines would enable stakeholders to prepare for future challenges and opportunities.

*Comparative Analysis*: Conducting comparative studies between onion production in the Philippines and other countries can offer valuable benchmarks and insights. Analyzing the similarities and differences in production trends, forecasting models, and influencing factors could lead to cross-country knowledge exchange and best practices adoption.

*Decision Support Systems*: Developing decision support systems based on the forecasted model and other forecasting approaches can assist policymakers, farmers, and stakeholders in real-time decision-making. Integrating these systems with data visualization, market analysis, and risk assessment tools would enable proactive management and optimization of onion production.

By pursuing these future research directions, future researchers can further refine the understanding of onion production dynamics, enhance forecasting capabilities, and contribute to the sustainable development and growth of the onion industry in the Philippines.

## REFERENCES

Amin, M., Nazir, H., & Wasim Amir, M. (2021). STATISTICAL MODELING AND FORECASTING FOR ONION PRODUCTION OF PAKISTAN. 59. 171-176.

Athanasopoulos, G., & Hyndman, R. J. (2018). Forecasting: Principles and Practice. https://otexts.com/fpp2/index.html

Balilla, J., Bondoc, M., Castro, K. A., & Padua, A. (2023). A 6-YEAR FORECAST OF EGG, RICE, AND ONION RETAIL PRICES IN THE PHILIPPINES: AN APPLICATION OF ARIMA AND. . . ResearchGate. https://www.researchgate.net/publication/370910461_A_6-YEAR_FORECAST_OF_EGG_RICE_AND_ONION_RETAIL_PRICES_IN_THE_PHILIPPINES_AN_APPLICATION_OF_ARIMA_AND_SARIMA_MODELS

Beckett, F. (2022). Super-ingredient: onions – and three imaginative ways to cook with them. Club Oenologique. https://cluboenologique.com/story/super-ingredient-onions/

Department of Agriculture. (2023). Price Monitoring. Price-Monitoring-January-2-2023.pdf (da.gov.ph)

Department of Labor and Employment. (2022). DAILY MINIMUM WAGE RATES. National Capital Region | National Wages Productivity Commission (dole.gov.ph)

Diaz, J., Mingo, F., & Urrutia, Jackie. (2017). Forecasting the Quarterly Production of Rice and Corn in the Philippines: A Time Series Analysis. Journal of Physics: Conference Series. 820. 012007. 10.1088/1742-6596/820/1/012007.

Diaz, JL. B., Urrutia, J. D., & Mingo, FL., T. (2017). Forecasting the Quarterly Production of Rice and Corn in the Philippines: A Time Series Analysis. Journal of Physics: Conference Series. 820. 012007. 10.1088/1742-6596/820/1/012007.

Finance Train. (2023). Financial time series analysis with R. https://financetrain.com/transforming-a-series-to-stationary

Gavino Jr., J. T., Ramos, E. E. & Alberto, R. T. (2020). SITE SUITABILITY MAPPING OF ONION IN THE PROVINCE OF OCCIDENTAL MINDORO USING GEOGRAPHIC INFORMATION SYSTEM (GIS) MODELBUILDER. Science Asia Review, 2(1).

Gholamy, A., Kosheleva, O., & Kreinovich, V. (2018). Why 70/30 or 80/20 Relation Between Training and Testing Sets: A Pedagogical Explanation. https://scholarworks.utep.edu/cgi/viewcontent.cgi?article=2202&context=cs_techrep

Guerrier, S., Smith, S., & Tremblay, V. (2019). Introduction to Time Series Analysis. https://smac-group.github.io/ts/introtimeseries.html

Handa, K. C., Mila, M. M., & Sabalberino AJ. A. (2023). Forecasting value of production of palay and retail price of rice in the Philippines using ARIMA modelling. World Journal of Advanced Research and Reviews. 17. 035-054. 10.30574/wjarr.2023.17.3.0345.

Hutchinson, D. (2023). The continuing struggle for onion farmers in the Philippines. PlantwisePlus Blog. https://blog.plantwise.org/2023/02/14/the-continuing-struggle-for-onion-farmers-in-the-philippines/

Hyndman, R. J. (2014). Thoughts on Ljung-Box test. https://robjhyndman.com/hyndsight/ljung-box-test/

Katchova, A. (2021). Time Series ARIMA Models. https://sites.google.com/site/econometricsacademy/econometrics-models/time-series-arima-models

Kern, C. (2017). Regression Analysis vs Time Series Analysis. [Video] https://www.youtube.com/watch?v=Prpu_U5tKkE

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? Journal of Econometrics, 54(1-3), 159–178. 10.1016/0304-4076(92)90104-Y

Lütkepohl, H., & Xu, F. (2012). The role of the log transformation in forecasting economic variables. Empirical Economics, 42(3), 619–638. 10.1007/s00181-010-0440-1.

Monigatti, G. (2022). Interpreting ACF and PACF Plots for Time Series Forecasting. Medium. https://towardsdatascience.com/interpreting-acf-and-pacf-plots-for-time-series-forecasting-af0d6db4061c?gi=4d9a2df2dec5

Munson, Kristen. (2020). DIY: You Can't Cook Without Onions - Utah State Magazine. Utah State Magazine. https://utahstatemagazine.usu.edu/diy-life/diy-you-cant-cook-without-onions/

National Institute of Standards and Technology. (2012). Time Series Analysis: Autocorrelation (ACF) and Partial Autocorrelation (PACF) Plots. https://www.itl.nist.gov/div898/handbook/pmc/section4/pmc4481.htm

Onion Nutrition - National Onion Association. (2020). National Onion Association. https://www.onions-usa.org/all-about-onions/onion-nutrition/

Panti, Llanesca T. (2023). Bureau of Plant Industry: Onion prices rose amid entry of smuggled onions | Money |. GMA News Online. https://www.gmanetwork.com/news/money/economy/858587/bureau-of-plant-industry-onion-prices-rose-amid-entry-of-smuggled-onions/story/

Parreño, S.J. (2022). Application of Time Series Analysis in Forecasting Coal Production and Consumption in the Philippines. ICTACT Journal on Soft Computing. 13. 2798-2804. 10.21917/ijsc.2022.0388

Pascual, M. P., Lorenzo, G. A., & Gabriel, A. G. (2018). Vertical Farming Using Hydroponic System: Toward a Sustainable Onion Production in Nueva Ecija, Philippines. Open Journal of Ecology, 08(01), 25–41. https://doi.org/10.4236/oje.2018.81003

Philippine Atmospheric, Geophysical and Astronomical Services Administration. (2023). Tropical Cyclone Information. https://www.pagasa.dost.gov.ph/climate/tropical-cyclone-information#:~:text=The%20peak%20of%20the%20typhoon,70%25%20of%20al%20typhoon%20develop

Philippine Statistics Authority. (2023). Major Vegetables and Rootcrops Quarterly Bulletin. (2094-618X). https://psa.gov.ph/content/major-vegetables-and-rootcrops-quarterly-bulletin

Rahul, N. (n.d.). ARIMA Model Selection Guidelines. https://people.duke.edu/~rnau/arimrule.htm

Simeon, L. M. (2016). Philippines adopts Vietnam technology to improve onion cultivation. Philstar.com. https://www.philstar.com/business/agriculture/2016/08/14/1613112/philippines-adopts-vietnam-technology-improve-onion-cultivation

Sjösten, L. (2022). A Comparative Study of the KPSS and ADF Tests in terms of Size and Power. https://www.diva-portal.org/smash/get/diva2:1668033/FULLTEXT01.pdf

Statista. (2022). Land area used for onion cultivation Philippines 2016-2021. [Graph] https://www.statista.com/statistics/1046183/land-area-used-for-onion-cultivation-philippines/#:~:text=In%202021%2C%20approximately%2019.3%20thousand,onion%20cultivation%20in%20the%20Philippines

Swanson, D. A. (2015). On the Relationship among Values of the Same Summary Measure of Error when it is used across Multiple Characteristics at the Same Point in Time: An Examination of MALPE and MAPE. Revie of Economic & Finance. 1923-7529-2015-03-01-14.

The Pennsylvania State University. (2023). STAT 510: Applied Time Series Analysis. https://online.stat.psu.edu/stat510/

Yu, L. S. (2023). EXPLAINER: The rise in onion prices – and why late imports don't help. RAPPLER. https://www.rappler.com/business/explainer-rise-onion-prices-why-late-imports-do-not-help/