

FORECASTING NATIONAL RICE PRODUCTION USING AUTOREGRESSIVE DISTRIBUTED LAG AND K-MEANS CLUSTERING

Dara Sakina¹⁾

Alfi Hidayatullah²⁾

Yuliana Kartika Permadani³⁾

Nasrudin⁴⁾

^{1,2,3,4)} Politeknik Statistika STIS, Jakarta, Indonesia

e-mail: 212212927@stis.ac.id

ABSTRACT

Indonesia as an agricultural country faces various obstacles in its development. The increasing population, decreasing agricultural land area, and differences in potential in each province are the biggest problems in rice production. The limitations of research in general in presenting differences in rice production characteristics between provinces are the urgency discussed in this study. The purpose of this study is to develop a rice production forecasting model in Indonesia using the ARDL approach combined with the K-Means clustering technique. The variables used in this study are rice production, harvested area, and farmers' terms of trade (NTP). Forecasting is carried out for the period 2024 with the aim of obtaining an accurate estimate of rice production in Indonesia. The results show that the ARDL model integrated with K-Means clustering provides highly accurate rice production forecasts, as indicated by low MAPE and RMSE values. In particular, Cluster 2 achieves the best performance with a MAPE of 0.36% and an RMSE of 3,860.40, followed by Cluster 1 (MAPE 2.17%; RMSE 33,192.31) and Cluster 3 (MAPE 5.88%; RMSE 40,577.10). In contrast, the national ARDL model without clustering records much larger errors (MAPE 14.71%; RMSE 582,062.00), confirming that clustering substantially improves forecasting accuracy and produces prediction patterns that closely match actual rice production.

Keywords: ARDL, clustering, forecasting, rice production, time series.

INTRODUCTION

Food is one of the most essential human needs. As a fundamental necessity for human survival, food availability is not only a prerequisite for maintaining physical life but also serves as the foundation for the sustainability of various human activities such as working, learning, and socializing. Without adequate food, individuals would be unable to carry out daily activities, maintain their health, or build social relationships within society (Beyene, 2023). The Food Law No. 7 of 1996 defines food security as "a condition in which food needs for households are met, as reflected in the availability of sufficient, safe, nutritious, evenly distributed, and affordable food."

As an agrarian country, Indonesia continues to rely heavily on rice as its primary staple food. To this day, rice remains difficult to substitute with other staple foods, and dependency on rice has continued to grow. This trend is evident from the shifting consumption patterns in eastern Indonesia, where communities have increasingly adopted rice as their main staple food. Consequently, national rice consumption has continued to rise in line with population growth (Septiadi & Joka, 2019).

According to BPS (2024), rice production in 2023 reached 53.98 million tons of Dry Milled Grain (GKG), representing a decline of 767.98 thousand tons GKG or about 1.4 percent compared with 2022. After conversion, rice production in 2023 amounted to 31.10 million tons, reflecting a decrease of 439.24 thousand tons or 1.39 percent from the previous year. This decline indicates that domestic rice production has not been able to fully meet the increasing national demand, which continues to grow in tandem with population expansion and changing consumption patterns.

At the same time, Indonesians remain highly dependent on rice as their primary staple food. Demand for rice tends to increase each year, while domestic production often fluctuates and slows. One of the main contributing factors is the decline in harvested paddy area, largely due to land-use conversion for industrial and residential purposes. Land conversion has become a serious issue, not only reducing the extent of productive agricultural land but also posing a potential threat to national food security. The widespread conversion of fertile agricultural land has led to declining rice production capacity as areas previously cultivated for rice continue to shrink (Prasada & Rosa, 2018).

Beyond land availability, farmers' welfare also plays a crucial role in rice production. The Farmers' Terms of Trade (NTP) is an important indicator in assessing farmers' welfare, as it reflects their purchasing power relative to production costs. An increase in NTP suggests that farmers' income is rising faster than agricultural production costs, indicating improvements in welfare (Keumala & Zainuddin, 2018). In the context of rice production, a higher NTP can directly influence farmers' decisions to increase production. When NTP is favorable, farmers are more confident in investing in agriculture, such as purchasing superior seeds, fertilizers, and modern farming equipment. Conversely, a low NTP often constrains farmers' production capacity, leading to lower rice output (Ayub et al., 2021).

Rice production in Indonesia also varies significantly across provinces, influenced by differences in land availability, productivity, water resources, and geographical conditions. According to BPS (2024), provinces such as East Java, Central Java, and West Java remain the largest contributors to national rice production, while several provinces outside Java record lower production levels. Although many previous studies have examined rice production at the national level, studies that explicitly account for inter-provincial differences in production patterns remain limited. These inter-provincial differences make the K-Means method an appropriate clustering approach to group provinces based on rice production trends (Zahra et al., 2024). The resulting clusters can then be utilized to apply the Autoregressive Distributed Lag (ARDL) model in order to identify the most suitable forecasting method. By integrating clustering with forecasting, policymakers will not only be able to predict rice production trends but also design more effective policies to strengthen food security and promote sustainable rice production.

METHOD

The data used in this study are monthly data covering the period from January 2020 to December 2024, consisting of 60 observations. These secondary data were obtained from Statistics Indonesia (BPS). The variables used in the analysis are rice production (Y), measured in tons; harvested area (X1), measured in hectares; and Farmers' Terms of Trade (X2), measured in percentage.

The analytical framework employed combines both descriptive and inferential approaches. The descriptive analysis involves graphical presentation and clustering using the K-Means method, while the inferential analysis relies on the Autoregressive Distributed Lag (ARDL) model to investigate both short-run and long-run relationships among the variables. To ensure robustness, the dataset is partitioned into training and testing subsets, with the first 48 observations used for model estimation and the last 12 observations reserved for evaluation.

In the descriptive stage, exploratory analysis of rice production across 34 provinces from the data train is conducted. The provinces are further grouped based on their annual production trends using K-Means clustering. Prior to clustering, the data are standardized using the Z-score method to ensure comparability across series. The similarity among provinces is then assessed through the Euclidean distance:

$$d_{ab} = \sqrt{\sum_{t=1}^n (x_{at} - x_{bt})^2} \tag{1}$$

where d_{ab} denotes the distance between provinces a and b , x_{at} denotes the standardized rice production value of province a in period t , x_{bt} denotes the standardized rice production value of province b in period t , a and b are province indices, t is the periods of the rice production, and n is the total number of observation periods. Provinces with the smallest distances are merged iteratively until all observations are grouped into clusters, allowing for the identification of regional production patterns.

The clustering results serve as the basis for the subsequent ARDL modeling stage. Provinces within the same cluster are assumed to share similar structural characteristics and production dynamics. Accordingly, ARDL models are estimated separately for each cluster as well as at the national level, allowing a direct comparison between cluster-based and aggregate (national) specifications. By explicitly accounting for these differences, the cluster-level ARDL models are expected to provide more accurate and reliable estimates than a single national ARDL specification. To implement this, cluster-level time series are constructed by aggregating rice production and harvested area through summation across provinces within each cluster, while Farmers' Terms of Trade are aggregated using averages to preserve their index-based interpretation. These aggregated series are then used as inputs for the ARDL estimation.

Building on these aggregated cluster-level and national series, the inference stage is carried out using the Autoregressive Distributed Lag (ARDL) model. The first step involves testing the stationarity of the data using the Augmented Dickey-Fuller (ADF) test. This test determines whether the series exhibit constant mean, variance, and covariance over time, with the decision rule based on the comparison between the t-statistic and the critical values at the 5% significance levels. The general form of the ADF regression can be expressed as:

$$\Delta Y_t = \alpha + \beta_t + \gamma Y_{t-1} + \sum_{i=1}^p \delta_i \Delta Y_{t-i} + \epsilon_t \tag{2}$$

where ΔY_t denotes the first difference of the series, α is the constant term, β captures the deterministic trend, γ is the coefficient of the lagged level of the series, δ_i denotes the coefficient of the i -th lag of the first-differenced series, Y_{t-1} is the lagged level of the variable, i is the lag index, t is the time index, p represents the maximum lag length included to account for serial correlation, and ϵ_t is the error term. The null hypothesis of the ADF test states that the series contains a unit root (non-stationary), while the alternative hypothesis states that the series is stationary. If the t-statistic is lower than the critical value, the series is considered stationary, meaning it does not exhibit systematic trends over time (Rizki Julianto & Atiqi Rohmawati, 2021). Once, the stationarity properties are established, the optimal lag length is selected, and the ARDL model is estimated.

The Autoregressive Distributed Lag (ARDL) model is an econometric approach designed to analyze the dynamic relationships among economic variables. This model enables the simultaneous estimation of both short-run and long-run relationships between independent and dependent variables. The ARDL framework combines two key components: the Autoregressive (AR) model, which incorporates past values of the dependent variable into the analysis, and the Distributed Lag (DL) model, which evaluates the lagged effects of independent variables on the

dependent variable. This combination provides considerable flexibility, as the ARDL model can be applied to data with different levels of stationarity, without requiring uniform stationarity across all variables (Gujarati & Porter, 2009). The ARDL specification can be expressed as follows.

$$Y_t = C_0 + \sum_{i=1}^p \beta_{0i} Y_{t-i} + \sum_{i=1}^q \beta_{1i} X_{1t-i} + \sum_{i=1}^r \beta_{2i} X_{2t-i} + \epsilon_t \quad (3)$$

where Y_t represents rice production, X_1 denotes harvested area which serves as a proxy for the capital production factor, and X_2 refers to the Farmers' Terms of Trade (NTP) which reflects the labor production factor. The parameter C_0 is the constant term. The coefficients $\beta_{0i}, \beta_{1i}, \beta_{2i}$ measure the lagged effects of past rice production, harvested area, and NTP, respectively, on current rice production, thereby capturing the short-run dynamics of the model. Meanwhile, i is the lag index, t is the time index, p, q, r represent the maximum lag lengths for $Y_t, X_1,$ and $X_2,$ respectively, while ϵ_t is the error term representing random shocks or other factors outside the model that may affect rice production. Accordingly, the model explains how past rice production, along with capital (land area) and labor (farmers' welfare), influence current rice production in both the short run and the long run.

Since ARDL estimation is conducted using Ordinary Least Squares (OLS), the classical regression assumptions must be satisfied. These include the absence of multicollinearity, autocorrelation, and heteroskedasticity, as well as the normal distribution of residuals (Gujarati & Porter, 2009). Model stability is further evaluated using the Cumulative Sum of Recursive Residuals (CUSUM) test. A model is considered stable if the blue CUSUM line remains within the 5% significance bounds represented by the red lines. If the line crosses the critical boundaries, the model is deemed unstable (Ahmad Ridha et al., 2021). To assess whether the variables share a long-run equilibrium relationship, the ARDL bounds test is employed. The conditional error correction representation of the ARDL model is expressed as follows.

$$\Delta Y_t = C_0 + \sum_{i=1}^{p-1} \beta_{0i} Y_{t-i} + \sum_{i=1}^{q-1} \beta_{1i} X_{1t-i} + \sum_{i=1}^{r-1} \beta_{2i} X_{2t-i} + \theta_0 Y_{t-1} + \theta_1 X_1 + \theta_2 X_2 + \epsilon_t \quad (4)$$

where the coefficient θ_0 reflects the speed of adjustment toward the long-run equilibrium after a short-run shock, whereas θ_1 and θ_2 capture the long-run equilibrium relationships of harvested area and Farmers' Terms of Trade (NTP) with rice production, while the remaining notations are defined as in the previous equations. The null hypothesis for the bounds test is no cointegration, while the alternative hypothesis is cointegration exist or the long-run relationship present. If the computed F-statistic exceeds the upper bound critical value, the null hypothesis is rejected, implying that the variables are cointegrated (Ahmad Ridha et al., 2021).

The performance of the models, both with and without clustering, is evaluated using forecast accuracy measures, namely the Mean Absolute Percentage Error (MAPE) and the Root Mean Square Error (RMSE). MAPE quantifies the average percentage deviation of the predicted values from the actual values, while RMSE measures the square root of the average squared prediction errors, reflecting the magnitude of forecast deviations. The formula for MAPE is given by Makridakis et al. (1998):

$$\text{MAPE} = \left(\frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - y_t}{Y_t} \right| \right) \times 100\% \quad (5)$$

and RMSE is calculated as (Afrianto et al., 2022):

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - y_t)^2} \tag{6}$$

where Y_t represents the actual value, y_t denotes the predicted value, $Y_t - y_t$ corresponds to the prediction error, t indicates the time index, and n refers to the total number of observation periods. The model that yields the smallest values of MAPE and RMSE considered to have the best predictive performance and is subsequently applied to forecast rice production using the reserved test dataset.

Figure 1 summarizes this modeling process, illustrating how clustering precedes data aggregation and ARDL estimation.

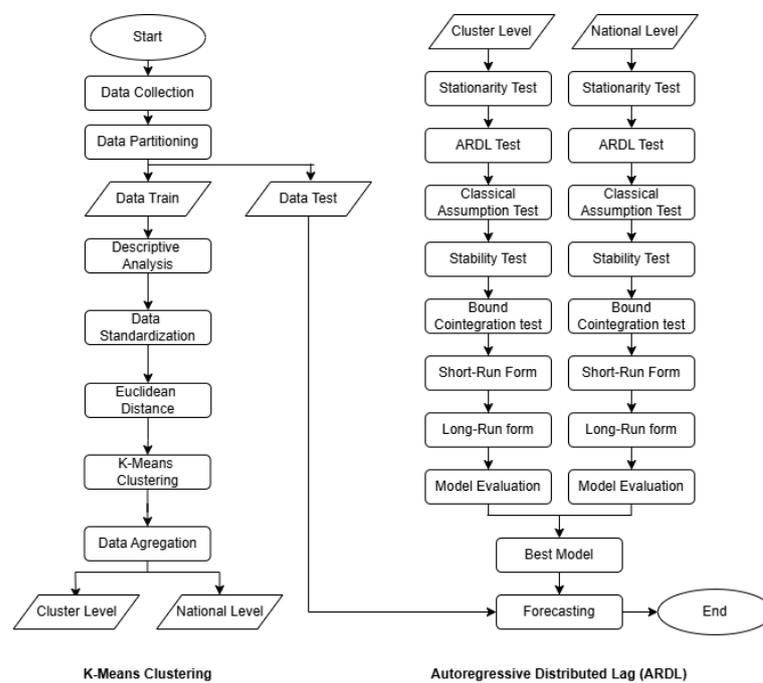


Figure 1. Research Methodology Flowchart

RESULTS AND DISCUSSION

Graphical Exploration

Graphical exploration was conducted by examining the time series patterns of rice production in Indonesian provinces during the period 2020–2023, as follows.

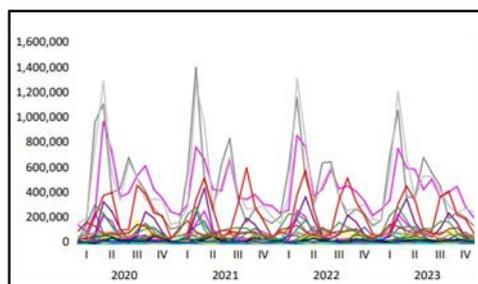


Figure 2. Rice Production across 34 Provinces in Indonesia during 2020-2023

Based on Figure 2, it can be observed that each province exhibits different trends and levels of rice production. According to Rahman et al. (2021), grouping Indonesian provinces can serve as an alternative approach to improve the efficiency of national-scale forecasting, as it tends to produce relatively lower error rates. Therefore, in this study, in addition to conducting a national time series analysis, clustering was applied to group provinces with similar characteristics in order to optimize the forecasting of national rice production.

Clustering of Provinces Using the K-Means Algorithm

The provinces were grouped using the K-Means clustering algorithm based on rice production trends. The similarity of time series data across provinces was measured using Euclidean distance. Provinces with the shortest distance to the cluster center (centroid) were assigned to the same cluster. This process was carried out iteratively, where each province was mapped to the nearest cluster and the cluster centers were updated until convergence was achieved. The optimal number of clusters was determined using the Elbow method, where the elbow point on the curve indicates the number of clusters that minimizes the error.

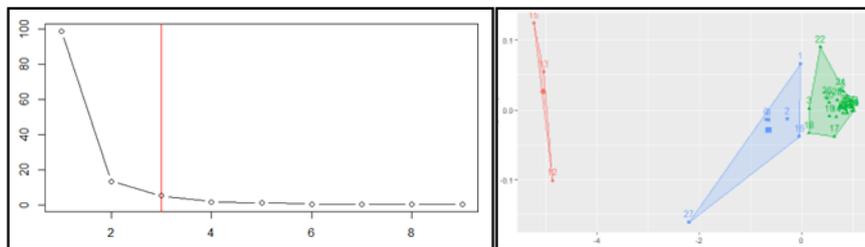


Figure 3. Elbow Plot and Clustering Result Plot

Based on the elbow plot in Figure 3, the optimal number of clusters was found to be three. The first cluster consists of West Java, Central Java, and East Java Provinces, characterized by high average rice production with production patterns concentrated in the first semester. The second cluster includes Aceh, North Sumatra, South Sumatra, Lampung, Banten, and South Sulawesi Provinces, characterized by moderate average rice production with production patterns that are relatively balanced between the first and second semesters. Meanwhile, the third cluster consists of the remaining 25 provinces, characterized by low average rice production and more heterogeneous production patterns. The graphical exploration results of each cluster are presented in the following figure (from left to right: clusters 1, 2, and 3).

Based on Figure 4, it can be seen that clustering is able to capture and group rice production patterns in Indonesian provinces quite effectively. This is evidenced by the similarity in trends, patterns, and average production volumes among provinces within the same cluster.

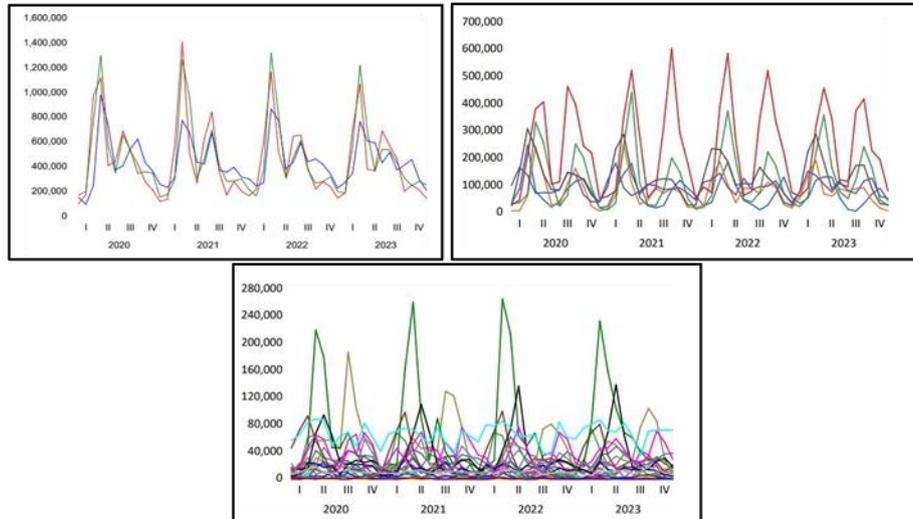


Figure 4. Rice Production Based on Clustering Results, 2020-2023

Stationarity Test

The coefficient of a variable in an autoregressive model can be tested through a stationarity test. The stationarity test was carried out using the Augmented Dickey-Fuller (ADF) Unit Root Test. This test aims to ensure that the variables analyzed are at the same order (Rusdi, 2011). At the level test stage, the first step is to examine whether the data are stationary. If the data are non-stationary, further testing is required, namely the first difference and the second difference.

Table 1. Stationarity Test Using the Augmented-Dickey Fuller (ADF) Unit Root Test

Region	Variable	Level		First Difference	
		Prob.	Remark	Prob.	Remark
Cluster 1	Y	0.0000	Stationary	0.0000	Stationary
	X1	0.0000	Stationary	0.0000	Stationary
	X2	0.9917	Non-stationary	0.0000	Stationary
Cluster 2	Y	0.0000	Stationary	0.0000	Stationary
	X1	0.0357	Stationary	0.0000	Stationary
	X2	0.9865	Non-stationary	0.0000	Stationary
Cluster 3	Y	0.0005	Stationary	0.0000	Stationary
	X1	0.0000	Stationary	0.0000	Stationary
	X2	0.8637	Non-stationary	0.0001	Stationary
Indonesia	Y	0.0000	Stationary	0.0000	Stationary
	X1	0.0003	Stationary	0.0000	Stationary
	X2	0.9990	Non-stationary	0.0499	Stationary

Based on Table 1, it can be concluded that all variables have achieved stationarity, either at the level or the first difference (no variable is stationary beyond the first difference). Therefore, the ARDL method can be applied and is considered an appropriate approach for this study.

ARDL Modeling

The selection of the best ARDL model was carried out using the optimal lag length criterion based on the AIC value. The modeling was performed at both the cluster and national levels to allow for comparison of the results.

Table 2. ARDL Model at the National Level Based on the AIC Criterion

ARDL Model	AIC
ARDL (2,0,0)	29.5034

Based on Table 2, the best ARDL model for national-level modeling is ARDL(2,0,0). This indicates that, at the national level, rice production is influenced by its past values (up to two previous periods) as well as by the contemporaneous values of Farmers' Terms of Trade (NTP) and harvested area (without lag).

Cluster-level modeling was conducted in the same manner as at the national level, except that prototypes were used, represented by the sum of time series within a cluster (for rice production and harvested area variables) and the average of time series within a cluster (for the farmers' terms of trade variable). These prototypes were then modeled using the ARDL approach, and the best model for each cluster was obtained as follows.

Table 3. ARDL Models at the Cluster Level Based on the AIC Criterion

Cluster	ARDL Model	AIC
1	ARDL (4,4,0)	24.0873
2	ARDL (4,0,0)	19.6716
3	ARDL (1,2,0)	24.3134

The ARDL modeling results indicate that rice production in each cluster is influenced by harvested area and the farmers' terms of trade, both in the short run and the long run, with varying dynamics across clusters. Clusters 1 and 2 exhibit models with longer lags, suggesting significant historical effects, whereas Cluster 3 is more sensitive to short-term changes. This reflects differences in characteristics and dependencies across clusters with respect to these variables.

Assumption Tests for the ARDL Model

Testing of the ARDL model is necessary to ensure that the research model does not violate the principles of econometrics. These tests include diagnostic checks for classical assumptions (normality, absence of autocorrelation, homoskedasticity, and non-multicollinearity), model stability tests, and overall model diagnostics.

Table 4. Classical Assumption Testing Results at the National Level

ARDL Model	Assumption Test	Method (Value)	Std. Error
ARDL (2,0,0)	Normality	KS (0.63)	Satisfied
	No Autocorrelation	Breusch-Godfrey (0.46)	Satisfied
	Homoskedasticity	ARCH (0.97)	Satisfied
	No Multicollinearity	VIF (1.21; 1.44; 1.14)	Satisfied

Table 5. Classical Assumption Tests by Cluster

Cluster	ARDL Model	Assumption Test	Method (Value)	Std. Error
1	ARDL (4,4,0)	Normality	KS (0.45)	Satisfied
		No Autocorrelation	Breusch-Godfrey (0.41)	Satisfied
		Homoskedasticity	ARCH (0.53)	Satisfied
		No Multicollinearity	VIF (1.31; 1.33; 1.24)	Satisfied
2	ARDL (4,0,0)	Normality	KS (0.89)	Satisfied
		No Autocorrelation	Breusch-Godfrey (0.81)	Satisfied
		Homoskedasticity	ARCH (0.28)	Satisfied
		No Multicollinearity	VIF (1.21; 1.94; 1.13)	Satisfied
3	ARDL (1,2,0)	Normality	KS (0.07)	Satisfied
		No Autocorrelation	Breusch-Godfrey (0.63)	Satisfied
		Homoskedasticity	ARCH (0.38)	Satisfied
		No Multicollinearity	VIF (3.13; 1.77; 1.10)	Satisfied

In general, the classical assumption tests on the ARDL model, both at the national level (without clustering) and in Clusters 1, 2, and 3, indicate that the model does not violate the classical assumptions and is appropriate for further analysis.

Stability Test of the ARDL Model

In this study, model stability was tested using the CUSUM test. The CUSUM test was conducted to examine whether the model parameters remained stable throughout the observation period or whether structural changes occurred.



Figure 5. CUSUM Test for Indonesia

Based on Figure 5, it can be observed that for Indonesia as a whole, without clustering, the CUSUM line lies between the two significance lines indicated by the dashed lines. This suggests that the model parameters are stable, implying that the model can be reliably used for prediction.

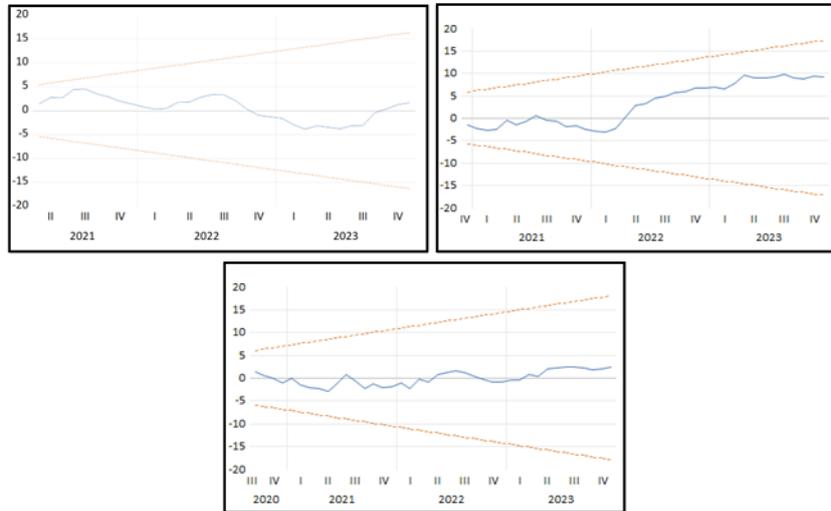


Figure 6. CUSUM Test by Cluster

Based on Figure 6, the CUSUM lines for all three clusters (Clusters 1, 2, and 3) lie between the two significance lines. This indicates that the model parameters did not undergo significant changes during the observation period. This provides confidence that the resulting models are reliable for further analysis, both in examining the relationships among variables and in making predictions.

Bounds Cointegration Test

The bounds test was conducted to identify the existence of long-run relationships in the ARDL models selected for this study. The test results focus on the F-statistic value. In EViews 12, the F-statistic is compared against the critical values at the 5% significance level.

Table 6. Results of the Bounds Cointegration Test for Indonesia

Model	Statistic Test	Value	Sig.	I(0)	I(1)
ARDL(2,0,0)	29.5034	36.7842	5%	3.10	3.87

Based on Table 6, the F-statistic value for the national-level model is greater than the upper bound (I(1)) at the 5% significance level. This indicates the presence of a significant long-run relationship among the variables in the study for the whole of Indonesia, without accounting for clustering.

Table 7. Results of the Bounds Cointegration Test by Cluster

Cluster	ARDL Model	F-Bound Test				
		Statistic Test	Value	Sig.	I(0)	I(1)
1	ARDL(4,4,0)	F-statistic	14.4902	5%	3.10	3.87
2	ARDL(4,0,0)	F-statistic	33437.63	5%	3.10	3.87

Cluster	ARDL Model	F-Bound Test				
		Statistic Test	Value	Sig.	I(0)	I(1)
3	ARDL(1,2,0)	F-statistic	20.5835	5%	3.10	3.87

Based on Table 7, the F-statistic values for Clusters 1, 2, and 3 are greater than the upper bound (I(1)) at the 5% significance level. This indicates the presence of cointegration in the models, meaning that the variables in this study share a significant long-run relationship. With co-integration present, the ARDL models are not only able to capture the short-run dynamics among rice production, harvested area, and the farmers' terms of trade, but also able to model their long-run linkages, both with and without clustering. Moreover, previous studies have shown that combining clustering with forecasting models substantially improves forecasting accuracy, because clustering groups time series with similar patterns into more homogeneous subsets, thereby reducing aggregation bias and allowing the model to better capture local dynamics (Liao et al., 2022). This implies that the clustered ARDL approach is more effective in representing regional heterogeneity in rice production and in producing more reliable forecasts than a single national model.

Short Run Form

After establishing the existence of a long-run relationship through the bounds cointegration test, the next step is to examine the short-run dynamics of the ARDL model. The short-run form allows the identification of how changes in the explanatory variables affect rice production in the immediate period, before the system converges to its long-run equilibrium. The short-run model estimation based on the selected best model is presented as follows.

Table 8. ARDL Model Estimation for Indonesia

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y(-1)	0.124	0.119	1.034	0.307
Y(-2)	-0.285	0.093	-3.067	0.004
X1	2.246	0.329	6.830	0.000
D(X2)	-188956.2	98317.76	-1.922	0.062
C	1266888.	324729.1	3.901	0.000
CointEq(-1)	-1.162	0.092	-12.566	0.000
F-Statistic				0.000

Based on Table 8, the variable CointEq(-1), also known as the error correction term, indicates a negative and significant error from the previous period. This suggests that the ARDL model in this study is valid and confirms the existence of cointegration among the variables. In other words, an adjustment of 116.2 percent will occur in the following period if fluctuations arise in the current year.

In the short run, the independent variables in the model collectively have a significant effect on Indonesia's rice production. Partially, the variables that significantly influence rice production in Indonesia are the second lag of rice production and harvested area.

Table 9. ARDL Model Estimation by Cluster

Cluster	Variable	Coefficient	Std. Error	t-Statistic	Prob.	
1	Y(-1)	0.412	0.144	2.860	0.007	
	Y(-2)	-0.457	0.149	-3.063	0.004	
	Y(-3)	0.000	0.152	0.004	0.997	
	Y(-4)	-0.405	0.134	-3.028	0.005	
	X1	3.261	0.038	85.909	0.000	
	X1(-1)	-1.412	0.467	-3.022	0.005	
	X1(-2)	1.424	0.488	2.922	0.006	
	X1(-3)	-0.060	0.495	-0.122	0.905	
	X1(-4)	1.224	0.433	2.827	0.008	
	DX2	-7672.273	8002.408	-0.960	0.345	
	C	121973.7	29928.80	4.076	0.000	
	CointEq(-1)*					0.000
F-Statistic					0.000	
2	Y(-1)	0.002	0.004	0.499	0.621	
	Y(-2)	-0.009	0.034	-1.832	0.075	
	Y(-3)	0.001	0.031	0.284	0.778	
	Y(-4)	-0.008	0.025	-2.363	0.024	
	X1	0.398	0.070	293.978	0.000	
	D(X2)	1004.378	761.440	1.577	0.123	
	C	13101.74	6054.240	2.514	0.016	
	CointEq(-1)*					0.000
	F-Statistic					0.000
	3	Y(-1)	-0.036	0.124	-0.288	0.775
X1		1.959	0.135	14.528	0.000	
X1(-1)		0.906	0.316	2.865	0.007	
X1(-2)		-0.687	0.142	-4.819	0.000	
D(X2)		-4936.669	5224.242	-0.945	0.350	
C		90174.88	31829.96	2.833	0.007	
CointEq(-1)*					0.000	
F-Statistic					0.000	

Based on Table 9, the variable CointEq(-1) is negative and significant for all models in each cluster. This indicates that the ARDL models in this study are valid and confirm the existence of cointegration among the variables. In other words, adjustments of 144.9 percent for Cluster 1, 109.8 percent for Cluster 2, and 103.6 percent for Cluster 3 will occur in the following period if fluctuations arise in the current year.

In Cluster 1, the independent variables in the model collectively have a significant effect on rice production. Partially, the variables influencing rice production are the first, second, and fourth lags of rice production, as well as harvested area, together with its first, second, and fourth lags.

In Cluster 2, the independent variables in the model collectively have a significant effect on rice production. Partially, the variables influencing rice production are the second and fourth lags of rice production and harvested area.

In Cluster 3, the independent variables in the model collectively have a significant effect on rice production. Partially, the variables influencing rice production are harvested area, along with its first and second lags.

Long Run Form

In addition to the short-run dynamics, the ARDL framework also provides estimates of the long-run equilibrium relationships among the variables. The long-run form illustrates how changes in the explanatory variables, such as harvested area and farmers’ terms of trade, affect rice production once the system has fully adjusted to equilibrium. The long-run model estimation based on the selected best model is presented as follows.

Table 10. Long-Run Equation of the ARDL Model for Indonesia

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	1.9332	0.2686	7.1974	0.0000
D(X2)	-162644.7	82254.19	-1.9773	0.0548
C	1090478.	245393.2	4.4438	0.0001

Based on Table 10, in the long run, the harvest area is found to significantly influence rice production, whereas the farmers’ terms of trade (NTP) do not show a significant long-run effect. This may occur because NTP tends to exert a more immediate influence, such as encouraging farmers to expand their planting activities when prices are favorable. However, in the long term, price fluctuations may diminish the sustained impact of NTP on rice production.

Table 11. Long-Run Equations of the ARDL Model by Cluster

Cluster	Variable	Coefficient	Std. Error	t-Statistic	Prob.
1	X1	3.0622	0.0426	71.9303	0.0000
	D(X2)	-5293.207	5385.426	-0.9843	0.3328
	C	84151.34	19061.06	4.4153	0.0001
2	X1	0.3926	0.003	150.2352	0.0000
	D(X2)	991.0424	629.678	1.5738	0.1240
	C	12927.78	5081.248	2.5442	0.0153
3	X1	2.1033	0.1571	13.3924	0.0000
	D(X2)	-4766.818	5015.778	-0.950	0.3476
	C	87072.33	32840.59	2.6510	0.0114

Based on Table 11, the clustering of provinces according to rice production yields conclusions consistent with the results for Indonesia as a whole without clustering. Across all clusters, the findings indicate that the harvest area significantly influences rice production, while the farmers' terms of trade (NTP) do not exert a significant effect. These results are in line with Sumaryanto et al. (2023), who found that the harvest area plays a crucial role in determining rice production.

Model Evaluation

To assess the performance of the estimated ARDL models, an evaluation was conducted by comparing the models with and without clustering. This comparison provides insights into whether clustering improves the accuracy and explanatory power of the models in capturing both the short-run and long-run dynamics of rice production. The results are summarized in Table 12.

Table 12. Comparison of ARDL Model Evaluation with and without Clustering

Region	ARDL		
	MAPE	AIC	RMSE
Cluster 1	2.17	24.0873	33,192.31
Cluster 2	0.36	19.672	3,860.40
Cluster 3	5.88	24.3134	40,577.10
Indonesia	14.71	29.5034	582,062.00

Overall, the ARDL model evaluation at the cluster level demonstrates better performance compared to the non-clustered model. The clustered models, particularly those for Cluster 2 and Cluster 1, exhibit higher accuracy (as indicated by lower MAPE and RMSE values) and more optimal complexity (as reflected by lower AIC values). In contrast, the non-clustered model (Indonesia) tends to be less accurate and more complex than the cluster-level models. Subsequently, the ARDL models at the cluster level will be employed to forecast rice production for 2024 to further assess the accuracy of these models in predicting the data.

Forecasting

Forecasting was conducted to project rice production in Indonesia for 2024 using the previously obtained ARDL models at the cluster level. The forecasting process was carried out for each cluster, and the estimated values from all clusters were then aggregated to obtain the total estimated rice production in Indonesia. To evaluate the accuracy of the forecasting results, the Mean Absolute Percentage Error (MAPE) was calculated by comparing the total estimated values with the actual rice production data. The forecasting results are presented in the following table and figure.

Table 13. Forecasted Rice Production in Indonesia for 2024

Month	Estimated				Actual Data	MAPE
	Cluster 1	Cluster 2	Cluster 3	Total		
Jan	438,250.15	489,870.58	214,851.25	1,142,971.97	870,000	2.61
Feb	511,900.12	632,786.10	589,896.86	1,734,583.08	1,390,000	2.07
Mar	1,900,701.55	1,067,708.27	951,056.67	3,919,466.50	3,430,000	1.19

Month	Estimated				Actual Data	MAPE
	Cluster 1	Cluster 2	Cluster 3	Total		
Apr	3,097,924.48	1,217,215.49	1,427,751.85	5,742,891.83	5,380,000	0.56
May	1,850,823.62	791,321.02	1,213,915.40	3,856,060.04	3,710,000	0.33
Jun	1,045,712.24	934,788.08	482,832.38	2,463,332.70	2,100,000	1.44
Jul	1,282,435.31	765,390.53	320,795.19	2,368,621.03	2,050,000	1.30
Aug	1,432,389.77	515,445.44	938,395.53	2,886,230.73	2,950,000	0.18
Sep	1,112,733.03	799,589.94	1,255,086.72	3,167,409.68	3,050,000	0.32
Oct	1,077,030.23	810,582.21	924,187.71	2,811,800.16	2,480,000	1.11
Nov	886,828.96	696,340.67	493,962.63	2,077,132.26	1,730,000	1.67
Dec	638,406.63	638,153.64	262,265.83	1,538,826.10	1,190,000	2.44
Mean	1,272,928.01	779,932.66	756,249.84	2,809,110.51	2,527,500	1.27

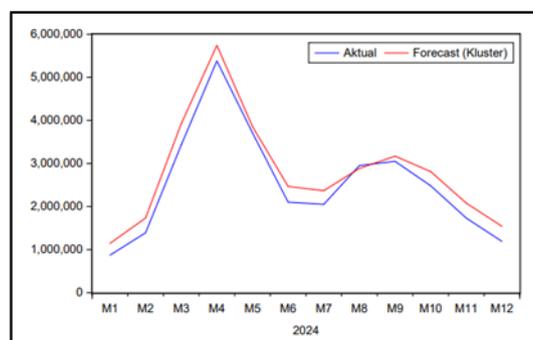


Figure 6. Comparison of Forecasted and Actual Rice Production in 2024

Based on Table 13 and Figure 6, forecasting using the ARDL model with the k-means clustering approach produces results that are fairly accurate and closely aligned with the actual data, as evidenced by a relatively low MAPE value of 1.27%. This is in line with the findings of Rahman et al. (2021b), who demonstrated that integrating clustering with time-series models effectively captures the heterogeneous production patterns across Indonesian provinces. Furthermore, from the comparison between the forecasted and actual values in the graph, it can be observed that the estimated rice production pattern closely mirrors the actual rice production trend.

Overall, this study finds that integrating ARDL with K-Means clustering substantially improves the accuracy of national rice production forecasting in Indonesia by capturing regional heterogeneity that is overlooked in a single national model. The clustered ARDL models produce significantly lower forecasting errors, with harvested area consistently emerging as the most important determinant of rice production in both the short and long run, while farmers' terms of trade mainly affect production in the short run. These findings imply that national rice policy should adopt a more region-specific approach, prioritizing the protection and optimization of harvested areas while maintaining short-term price and input stabilization policies to support farmers' production decisions. However, this study is limited by the use of a relatively small set of explanatory variables and a short forecasting horizon, which may not fully account for climatic, technological, and institutional factors influencing rice production. Future research is therefore encouraged to incorporate additional variables, explore alternative clustering techniques, and

extend the forecasting period to strengthen the robustness of policy-relevant rice production projections.

CONCLUSION

The results of this study demonstrate that the ARDL model integrated with K-Means clustering provides highly accurate rice production forecasts, significantly outperforming the national model. This is evidenced by the hybrid model's superior performance across all groups, particularly in Cluster 2, which achieved the highest precision with a MAPE of 0.36% and an RMSE of 3,860.40, followed by Cluster 1 (MAPE 2.17%; RMSE 33,192.31) and Cluster 3 (MAPE 5.88%; RMSE 40,577.10). In contrast, the national ARDL model without clustering recorded substantially larger errors (MAPE 14.71%; RMSE 582,062.00), confirming that capturing regional heterogeneity through clustering is essential for producing prediction patterns that closely match actual production.

Beyond its technical accuracy, this study identifies harvested area as the primary long-run and short-run determinant of production across clusters, while farmers' terms of trade (NTP) tend to provide a short-term signal for production decisions, although its statistical significance is less robust than that of harvested area. These insights imply that the government should transition toward region-specific policies that prioritize the protection of harvested areas alongside short-term price stabilization. Despite these strengths, the research is subject to certain methodological limitations, particularly its reliance on a limited set of explanatory variables and a short forecasting horizon, which may not fully account for climatic anomalies or technological shifts. Consequently, future research is encouraged to incorporate broader environmental factors and explore alternative clustering techniques to further enhance the robustness of national food security projections.

REFERENCE

- Afrianto, N., Fudholi, D. H., & Rani, S. (2022). Stock Price Prediction using BiLSTM with Public Sentiment Factor. *Jurnal RESTI*, 6(1), 41–46. <https://doi.org/10.29207/resti.v6i1.3676>
- Ahmad Ridha, Nurjannah, & Ratna Mutia. (2021). Analisis Permintaan Uang di Indonesia: Pendekatan Autoegressive Distributed lag (Ardl). *Jurnal Samudra Ekonomika*, 5(2), 152–160. <https://doi.org/10.33059/jse.v5i2.4273>
- Ayub, A., Noorachmat, B. P., & Yanuar Jarwadi Purwanto, M. (2021). Analisis alih fungsi lahan sawah dan keterkaitannya dengan nilai tukar petani (NTP) di Kabupaten Bantul. *Jurnal Ilmiah Rekayasa Pertanian Dan Biosistem*, 9(1), 57–65. <https://doi.org/10.29303/jrpb.v9i1.215>
- Beyene, S. D. (2023). The impact of food insecurity on health outcomes: empirical evidence from sub-Saharan African countries. *BMC Public Health*, 23(1), 338. <https://doi.org/10.1186/s12889-023-15244-3>
- BPS. (2024). Luas panen dan produksi padi di Indonesia 2023.
- Gujarati, D. N. ., & Porter, D. C. . (2009). *Basic econometrics*. McGraw-Hill Irwin.
- Keumala, C. M., & Zainuddin, Z. (2018). Indikator Kesejahteraan Petani melalui Nilai Tukar Petani (NTP) dan Pembiayaan Syariah sebagai Solusi. *Economica: Jurnal Ekonomi Islam*, 9(1), 129–149. <https://doi.org/10.21580/economica.2018.9.1.2108>
- Liao, Q. Z., Xue, L., Lei, G., Liu, X., Sun, S. Y., & Patil, S. (2022). Statistical prediction of waterflooding performance by K-means clustering and empirical modeling. *Petroleum Science*, 19(3), 1139–1152. <https://doi.org/10.1016/j.petsci.2021.12.032>
- Makridakis, S. G., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: methods and applications* (3rd ed.). John Wiley & Sons.
- Prasada, I. M. Y., & Rosa, T. A. (2018). Dampak alih fungsi lahan sawah terhadap ketahanan pangan di Daerah Istimewa Yogyakarta. *Jurnal Sosial Ekonomi Pertanian*, 14(3), 210.

- <https://doi.org/10.20956/jsep.v14i3.4805>
- Rahman, R. A., Afendi, F. M., Nugraheni, W., Sadik, K., & Rizki, A. (2021). Pengelompokan dan Peramalan Deret Waktu pada Produksi Bawang Merah Tingkat Provinsi di Indonesia. *Seminar Nasional Official Statistics*, 2021(1), 457–464. <https://doi.org/10.34123/semnasoffstat.v2021i1.910>
- Rizki Julianto, I., & Atiqi Rohmawati, A. (2021). Prediksi Jumlah Kunjungan Wisatawan Di Jawa Barat Dengan Model ARIMAX Dan SARIMAX Menggunakan Data Google Trends. *EProceedings of Engineering*.
- Rusdi. (2011). Uji Akar-Akar Unit dalam Model Runtun Waktu Autoregresif . *Statistika*, 11(2), 67–78.
- Septiadi, D., & Joka, U. (2019). Analisis Respon dan Faktor-Faktor yang Mempengaruhi Permintaan Beras Indonesia. *AGRIMOR*, 4(3), 42–44. <https://doi.org/10.32938/ag.v4i3.843>
- Sumaryanto, S., Susilowati, S. H., Saptana, S., Sayaka, B., Suryani, E., Agustian, A., Ashari, A., Purba, H. J., Sumedi, S., Dermoredjo, S. K., Purwantini, T. B., Yofa, R. D., & Pasaribu, S. M. (2023). Technical efficiency changes of rice farming in the favorable irrigated areas of Indonesia. *Open Agriculture*, 8(1). <https://doi.org/10.1515/opag-2022-0207>
- Zahra, A. L., Tiara, S., Ada, R., & Ardini, A. F. (2024). Implementasi Clustering Algoritma K-Means Pada Produksi Beras di Provinsi Jawa Timur Tahun 2022. *Journal of Computer and Information Systems Ampera*, 5(3), 2775–2496. <https://journal-computing.org/index.php/journal-cisa/index>