UNIVERSITI TEKNOLOGI MARA

MATHEMATICS AND STATISTICS

UNDERGRADUATE RESEARCH PROCEEDINGS 2025

UiTM CAWANGAN NEGERI SEMBILAN

PADDY PRODUCTION ANALYSIS IN NON-GRANARY AREAS

Nora Binti Mohd Basir[1*], Nur Aliah Najwa Binti Ahmad Mahyudin[1]
[1]*College of Computing, Informatics and Mathematics, Universiti Teknologi MARA (UiTM) Cawangan Negeri Sembilan, Kampus Seremban, Negeri Sembilan*
[*]noram661@uitm.edu.com

Abstract

Currently, it has been shown that Malaysia has a low production of paddy. That is one of the reasons why Malaysia has been importing rice from other countries. To help Malaysia overcome the rice output shortage, this study has decided to determine whether planted area and paddy yield affect paddy production in a long-run scenario. This study analyses the paddy production in non-granary areas in Johor, Melaka, Negeri Sembilan, Pahang, Pulau Pinang, and Terengganu from 2011 until 2020, during the main season in Malaysia. In order to determine the presence of the long-run relationship, this study employed the Pedroni cointegration test. It was found that most of the statistical tests from the Pedroni cointegration test are significant at a 5% significance level. It can be concluded that there is a long-run relationship between planted area, paddy yield, and paddy production. Thus, it is recommended that future researchers investigate these effects in detail to help maximise paddy production, especially in non-granary areas.

Keywords*:* Cointegration test, Paddy production, Panel data, Stationary test

Introduction

The world's three most important food crops are rice, wheat, and maise. When taken as a whole, rice makes up almost half of the calories people consume worldwide [9]. Besides, rice has also been one of the staple foods in many countries around the world. Countries located at the Association of Southeast Asian Nations (ASEAN), also known as the Southeast Asian regions, have been categorised as the countries that consume rice the most. In 2023, a study found that Malaysia only consumes 106 kilograms of rice per person yearly, which is lower than some Southeast Asian countries [14]. Table 1 shows Malaysia's total rice consumption from 2019 to 2024 [10].

Table 1: Total Consumption of Rice in Malaysia

Year	Rice Consumption (million metric tons)
2019	2.75
2020	2.85
2021	2.90
2022	2.90
2023**	2.91
2024**	2.91

** indicates the estimation of rice consumption in the future.

In 2021, Malaysia only produced roughly 1.68 million metric tons of rice, 180 thousand metric tons more than the production in 2020 [11]. It shows that rice production does not align with rice consumption as the current production is still unable to meet the demands of the Malaysian market. Naturally, as the demand for rice increases, production should also rise to correspond to the demand. While Malaysia produces 67% of its rice, the remaining is imported from Pakistan, Vietnam, and Thailand [12]. Thus, the Malaysian government is preparing to increase local rice production by 2025 [11].

Essentially, Malaysia needs to produce more rice. To maximise production, Malaysia should focus on land and climate change. Malaysia has two seasons: main season and off-season, during which the activities of paddy cultivation, from the preparation of land to harvesting, start [6]. Table 2 shows that the non-granary areas in Johor, Melaka, Perak, and Terengganu have a decreasing productivity status. Meanwhile, Kedah, Negeri Sembilan, and Perlis have optimum productivity, while productivity in Kelantan and Pahang have an increasing pattern.

Table 2: Summary of Productivity Status

Non-Granary States	Productivity Status	Non-Granary States	Productivity Status
Johor	Decreasing	Pahang	Increasing
Kedah	Optimum	Perak	Decreasing
Kelantan	Increasing	Perlis	Optimum
Melaka	Decreasing	Terengganu	Decreasing
Negeri Sembilan	Optimum		

Many aspects should be considered when discussing the causes of paddy production. A study found that rice production can be affected by human capital, labour, wages, wetlands, urban population, and rice prices [3]. However, the technology has barely any effect on the production. Exploring both short- and long-run factors affecting paddy production is important. A study on rice production in Nepal revealed that in both short- and long-term, rice production is positively influenced by cultivated area, fertiliser consumption, and agricultural credit [5].

The average yield of cleaned paddy is the mean crop after drying and cleaning and when the moisture content is precisely 14% [6]. There is no significant difference in the yield during the dry season, regardless of the types of water management [13]. On the contrary, the yield differs with alternate wetting and drying systems having higher paddy yields than continuous flooding [13]. Next, [6] stated that the paddy plantation is divided into granary and outside granary areas. The granary areas refer to the dominant irrigation schemes, which are about 4,000 hectares minimum, while the outside granary areas are those with minor or outside irrigation systems.

A study by [13] used both water management systems, continuous flooding (CF) and alternate wetting and drying (AWD), to manage water in maintaining rice production. During the wet season, CF needed more irrigation than AWD, as CF irrigates 1,835 mm of water while AWD irrigates 1,522 mm [13].

Nevertheless, the current paddy production is insufficient to cover Malaysia's population. Since the granary areas are likely to have optimum and increasing rice productivity, this study focuses on the non-granary areas during the main season to help increase productivity. Hence, the study will determine whether a long relationship exists between planted area, paddy yield, and paddy production.

Methodology

Source of Data

This study used data from the Paddy Production Survey Report 2020/2021 by the Department of Agriculture, Peninsular Malaysia. The data provided is categorised as panel data, a combination of cross-sectional and time series data [15]. The annual data utilises information from six states in Malaysia, specifically non-granary states, from 2011 until 2020, with 60 observations and two independent variables. The six non-granary states are Johor, Melaka, Negeri Sembilan, Pahang, Pulau Pinang, and Terengganu, and the independent variables that are being investigated are the planted area and paddy yield. Table 3 shows the variable of interest in this study.

Table 3: List of Variables

Variable	Description	Measurement
AREA	Planted area in hectare	Ratio
YIELD	Average yield of cleaned paddy in kilogram/hectare	Ratio
PRODUCTION	Wetland paddy production in metric ton	Ratio

Stationary Test

Since panel data combines time-series and cross-sectional data, the study needs to test for model stationary using a unit root test. The study decided to use the Hadri Lagrange Multiplier test proposed by Hadri in 2000. Hadri has a null hypothesis of stationary, while in contrast, the alternative hypothesis of at least one series has a unit root [7]. The Hadri test's components are represented by Equation 1 [1].

$$y_{it} = z'_{it}\gamma + r_{it} + \varepsilon_{it} \tag{1}$$

Where z_{it} is the deterministic component, while r_{it} is a random walk,

$$r_{it} = r_{it} + u_{it} \tag{2}$$

And ε_{it} is known as a stationary process, while $u_{it} \sim iid(0, \sigma_u^2)$. The partial sum of the residuals, S_{it} is written as in Equation 3

$$S_{it} = \sum_{j=1}^{t} \hat{e}_{ij} \tag{3}$$

While the LM statistic represented in Equation 4

$$LM = \frac{\frac{1}{n}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T} S_{it}^2}{\hat{\sigma}_{\varepsilon}^2} \tag{4}$$

Hadri (2000) believes that the standardised statistics are as in Equation 5 and Equation 6 [7]:

$$Z_{\mu} = \frac{\sqrt{N}(LM_{\mu} - \xi_{\mu}}{\zeta_{\mu}} \Rightarrow N(0,1) \tag{5}$$

and

$$Z_{\tau} = \frac{\sqrt{N}(LM_{\tau} - \xi_{\tau}}{\zeta_{\tau}} \Rightarrow N(0,1) \tag{6}$$

As $T \to \infty$ followed by $N \to \infty$.

Cointegration Test

To verify the stationary variables, a unit root test is run [2]. It is crucial to determine whether the stationary first-difference variables, $I(1)$, are cointegrated. The cointegration test also checks whether a long-run relationship among the variables exists or not [4]. The study uses Pedroni's (1999) residual-based cointegration test to test the cointegration. It is constructed based on Engle-Granger (1987), as the test statistic is from the residuals of the panel static regression [2].

The test statistics and p-values by Pedroni (1999) are a panel and weighted panel of v-statistic, ρ-statistics, PP-statistic, and ADF-statistic. Besides, there are also group mean approaches of ρ-statistics, PP-statistic, and ADF-statistic [8]. The panel and weighted panel statistics are from pooled data along the within-dimension, while the group statistics are from pooled data along the between-dimension [2]. Pedroni had designed the following test statistics [2]:

1. Panel v-statistic:

$$Z_{\hat{v}_{NT}} = \frac{1}{\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{L}_{11i}^{-2}\ \hat{e}_{it-1}^{2}\right)} \tag{7}$$

2. Panel ρ-statistic:

$$Z_{\hat{\rho}_{NT}-1} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{L}_{11i}^{-2}\ \left(\hat{e}_{it-1}\Delta\hat{e}_{it} - \hat{\lambda}_i\right)}{\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{L}_{11i}^{-2}\ \hat{e}_{it-1}^{2}\right)} \tag{8}$$

3. Panel t-statistic (non-parametric):

$$Z_{t_{NT}} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{L}_{11i}^{-2}\ \left(\hat{e}_{it-1}\Delta\hat{e}_{it} - \hat{\lambda}_i\right)}{\sqrt{\tilde{\sigma}_{NT}^{2}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{L}_{11i}^{-2}\ \hat{e}_{it-1}^{2}\right)}} \tag{9}$$

4. Panel t-statistic (parametric):

$$Z_{t_{NT}}^{*} = \frac{\sum_{i=1}^{N}\sum_{t=2}^{T}\hat{L}_{11i}^{-2}\ \hat{e}_{it-1}^{*}\Delta\hat{e}_{it}^{*}}{\sqrt{\tilde{s}_{NT}^{*2}\left(\sum_{i=1}^{N}\sum_{t=2}^{T}\hat{L}_{11i}^{-2}\ \hat{e}_{it-1}^{*2}\right)}} \tag{10}$$

5. Group ρ-statistic:

$$\tilde{Z}_{\hat{\rho}_{NT}-1} = \sum_{i=1}^{N}\frac{\sum_{t=1}^{T}\left(\hat{e}_{it-1}\Delta\hat{e}_{it} - \hat{\lambda}_i\right)}{\left(\sum_{t=1}^{T}\hat{e}_{it-1}^{2}\right)} \tag{11}$$

6. Group t-statistic (non-parametric):

$$\tilde{Z}_{t_{NT}} = \sum_{i=1}^{N}\frac{\sum_{t=1}^{T}\left(\hat{e}_{it-1}\Delta\hat{e}_{it} - \hat{\lambda}_i\right)}{\sqrt{\hat{\sigma}_i^{2}\left(\sum_{t=1}^{T}\hat{e}_{it-1}^{2}\right)}} \tag{12}$$

7. Group t-statistic (parametric):

$$Z_{t_{NT}}^{*} = \sum_{i=1}^{N}\frac{\sum_{t=1}^{T}\hat{e}_{it-1}^{*}\Delta\hat{e}_{it}^{*}}{\sqrt{\sum_{t=2}^{T}\tilde{s}_i^{*2}\hat{e}_{it-1}^{*2}}} \tag{13}$$

As mentioned before, the null hypothesis of the Pedroni residual cointegration test is that there is no cointegration relationship between the non-stationary variables. So, to estimate a model in a long-run relationship, the study must reject the null hypothesis. To reject the null hypothesis, the study must ensure that the majority of the test statistics and their p-value are significant. Only then can it be proven that there is a long-run relationship between planted area, paddy yield and paddy production.

Results and Discussion

Stationary Test

Since panel data combines cross-sectional and time series data, stationary assumptions must be checked. This study uses the Hadri Lagrange Multiplier (LM) Stationary test, as stated in Table 4.

Table 4: Hadri LM test

Variables	Level	First Difference
AREA	1.443e-03 *	0.5037
YIELD	1.203e-06 *	0.7009
PRODUCTION	1.035e-04 *	0.6687

* significant at 0.05.

Table 4 displays that at level, all p-values are significant at 5%. The study can conclude that area, yield, and production are not stationary at this level. Hence, the variables must be tested again at the first difference level. Next, it is proven that area, yield, and production are stationary at first difference as the p-values are insignificant. It can be concluded that all variables are stationary at first difference and integrated at first order, $I(1)$. Thus, the study can check the cointegration between the variables to determine the existing long-run relationship between them.

Cointegration Test

A cointegration test is performed since the study wants to determine the long-run relationship between planted area, paddy yield, and paddy production. If cointegration exists, it can be proven that there is a long-run relationship.

Table 5: Pedroni Residual Cointegration test

	Test	Statistic	Probility
Alternative hypothesis: common AR coefs. (within-dimension)	Panel v-statistic	-1.2141	0.8877
	Panel rho-statistic	1.1603	0.8770
	Panel PP-statistic	-1.7310	0.0417 *
	Panel ADF-statistic	-1.7291	0.0419 *
	Weighted Panel v-statistic	-1.5085	0.9343
	Weighted Panel rho-statistic	0.7018	0.7586
	Weighted Panel PP-statistic	-6.9495	0.0000 *
	Weighted Panel ADF-statistic	-5.4273	0.0000 *
Alternative hypothesis: common AR coefs. (between-dimension)	Group rho-statistic	1.6467	0.9502
	Group PP-statistic	-7.6406	0.0000 *
	Group ADF-statistic	-4.8891	0.0000 *

* significant at 0.05.

Six out of eleven p-values from Table 5 are less than 0.05. Hence, it can be concluded that the variables are cointegrated. Thus, the study can conclude that there is a long-run relationship between planted area, paddy yield, and paddy production.

Conclusion and Recommendation

Rice is the staple food in Malaysia, and it is common for the demand for rice to exceed that of other foods. When the demand for rice increases, paddy production should ideally increase proportionately to meet this demand. However, since Malaysia's market needs are not being fulfilled, this study explores the factors that could enhance local paddy production. The dataset of paddy production, planted area, and paddy yield during the main season from 2011 until 2020 for a few non-granary states in Malaysia was analysed. Hadri LM test is applied to test the stationary, and it is found that planted area, paddy yield and paddy production are not stationary at level but stationary at first difference. Hence, the Pedroni residual cointegration test is run. From the cointegration test, the finding reveals that all the variables are cointegrated with each other. It is concluded that there is a long-run relationship between planted area, paddy yield, and paddy production.

Hence, the study believes that there is a long-run effect between planted area and paddy yield with paddy production. Therefore, future researchers should investigate in detail these effects and how the planted area and paddy yield affect the paddy production output in the long run. Besides, it is recommended for future studies to take into account the data around the whole year, which includes both main and off seasons, since this research only investigates during the main season.

Acknowledgement

We sincerely thank each member for their unwavering support and commitment, which ensured that this article was published successfully.

References

[1] Baltagi, B. H., & Kao, C. (2000). Non-stationary panels, cointegration in panels and dynamic panels: A survey. *Advances in Econometrics*

[2] Barbieri, L. (2008). Panel Cointegration Tests: A Survey. *Rivista Internazionale Di Scienze Sociali, 116*(1), 3–36. https://www.jstor.org/stable/41625199

[3] Bashir, A., & Yuliana, S. (2019). Identifying factors influencing rice production and consumption in Indonesia. *Jurnal Ekonomi Pembangunan: Kajian Masalah Ekonomi Dan Pembangunan, 19*(2) https://doi.org/10.23917/jep.v19i2.5939

[4] Bouchoucha, Yahyaoui, I., & Bouchoucha, N. (2019). The long-run relationship between ODA, growth and governance: An application of FMOLS and DOLS Approaches

[5] Chandio, A. A., Jiang, Y., Ahmad, F., Adhikari, S., & Ain, Q. U. (2021). Assessing the impacts of climatic and technological factors on rice production: Empirical evidence from Nepal. *Technology in Society, 66* 101607. https://doi.org/10.1016/j.techsoc.2021.101607

[6] Department of Agriculture Peninsular Malaysia (2021). Paddy Production Survey Report MALAYSIA, Main Season 2020/2021.

[7] Giulietti, M., Otero, J., & Smith, J. (2009). Testing for stationarity in heterogeneous panel data in the presence of cross-section dependence. *Journal of Statistical Computation and Simulation, 79*(2), 195–203. https://doi.org/10.1080/00949650701719136

[8] Örsal, D. D. K. (2007). Comparison of panel cointegration tests. Www.econstor.eu. http://hdl.handle.net/10419/25201

[9] Pariona, A. (2019). What are the World's Most Important Staple Foods? WorldAtlas. https://www.worldatlas.com/articles/most-important-staple-foods-in-the-world.html

[10] Statista. (2023a). Annual Rice Consumption in Malaysia 2019-2024. https://worldpopulationreview.com/country-rankings/rice-consumption-by-country

[11] Statista. (2023b). Production of Rice in Malaysia 2013-2021. https://www.statista.com/statistics/794700/rice-production-volume-malaysia/

[12] Syahmi Makhtar, Irwan Shah Zainal Abidin, & Rabiul Islam. (2022). REVIEWING FOOD SECURITY ON PADDY PRODUCTION: A CONCEPTUAL PAPER. *International Journal of Industrial Management, 16(*1), 51–58. https://doi.org/10.15282/ijim.16.1.2022.9012

[13] Win, E. P., Win, K. K., Bellingrath-Kimura, S. D., & Oo, A. Z. (2020). Greenhouse gas emissions, grain yield and water productivity: a paddy rice field case study based in Myanmar. *Greenhouse Gases: Science and Technology, 10*(5), 884–897. https://doi.org/10.1002/ghg.2011

[14] World Population Review. (2024). Rice Consumption by Country 2023. Worldpopulationreview.com. https://worldpopulationreview.com/country-rankings/rice-consumption-by-country

[15] Yohandoko, S., & Supriyanto, S. (2023). Panel Data Analysis on Rice (Paddy) Production in Indonesia 2018-2021. *International Journal of Mathematics, Statistics, and Computing, 1*(3), 44–53. https://doi.org/10.46336/ijmsc.v1i3.27