Jurnal Statistika Dan Komputasi (STATKOM), Vol. 5 No. 1 (June 2026)
ISSN : 2963-038X (print), ISSN : 2963-0398
DOI : https://doi.org/10.32665/statkom.v5i1.6453

Spatial Dependence and Spillover Effects on Very Low-Income Agricultural Enterprises at the Provincial Level in Indonesia

Nur Kamilah Sa'diyah[1], Meilina Retno Hapsari[2]

[1,2]Department of Statistics, Faculty of Mathematics and Natural Sciences, Brawijaya University
Email: nurkamilahs24@ub.ac.id

Submitted: April 14, 2026 ***Revised:*** June 14, 2026 ***Accepted:*** June 19, 2026

Abstract

Background: *Agricultural welfare disparities across Indonesian provinces remain substantial, as reflected in the high percentage of agricultural enterprises with very low income. These disparities are influenced by regional characteristics and spatial interactions among neighboring provinces.*
Objective: *This study analyzes the determinants of the percentage of agricultural enterprises with very low income at the provincial level while accounting for spatial dependence and spillover effects.*
Methods: *Cross-sectoral data from 38 provinces in Indonesia were obtained from the 2024 Agricultural Economic Survey. The analysis was conducted using spatial regression models, the Spatial Lag Model (SLM) and the Spatial Error Model (SEM). The spatial weight matrix was constructed using the KNN approach with k = 5. Moran's I test was used to detect spatial autocorrelation. The LM test and AIC were used to select the best model.*
Results: *Significant positive spatial autocorrelation was detected in the dependent variable and OLS residuals. The SLM was selected as the best model. The explanatory variables were shown to have both direct and indirect effects through spatial spillover effects.*
Conclusion: *Incorporating spatial dependence is essential for effective agricultural policymaking. Coordinated interprovincial policies on credit access, production facilities, and agricultural input supply are needed to improve agricultural welfare.*

Keywords : *Spatial Regression, Spatial Lag Model, Spatial Autocorrelation, Spillover Effects, Agricultural Welfare.*

Abstrak

Latar Belakang: Kesenjangan kesejahteraan pertanian di seluruh provinsi Indonesia masih cukup besar, sebagaimana tercermin dalam tingginya persentase usaha pertanian dengan pendapatan sangat rendah. Kesenjangan ini dipengaruhi oleh karakteristik regional dan interaksi spasial antar provinsi tetangga.
Tujuan: Penelitian ini bertujuan untuk menganalisis variabel-variabel penentu usaha pertanian berpendapatan sangat rendah pada tingkat provinsi dengan mempertimbangkan ketergantungan spasial dan efek spillover.
Metode: Data cross-section 38 provinsi di Indonesia diperoleh dari Survei Ekonomi Pertanian (SEP) 2024. Analisis dilakukan menggunakan model regresi spasial, yaitu Spatial Lag Model (SLM) dan Spatial Error Model (SEM). Matriks bobot spasial dibangun menggunakan pendekatan KNN dengan k = 5. Uji Moran's I digunakan untuk mendeteksi autokorelasi spasial. Uji LM dan AIC untuk memilih model terbaik.
Hasil: Autokorelasi spasial positif yang signifikan terdeteksi pada variabel dependen dan residual OLS. SLM terpilih sebagai model terbaik. Variabel penjelas terbukti berpengaruh baik secara langsung maupun tidak langsung melalui efek spillover spasial.
Kesimpulan: Mempertimbangkan ketergantungan spasial sangat penting untuk pembuatan kebijakan pertanian yang efektif. Kebijakan antarprovinsi yang terkoordinasi mengenai akses kredit, penyediaan sarana produksi, dan pasokan input pertanian diperlukan untuk meningkatkan kesejahteraan pertanian.

Kata kunci: Regresi Spasial, Model Lag Spasial, Autokorelasi Spasial, Efek Spillover, Kesejahteraan Pertanian.

INTRODUCTION

The agricultural sector continues to play a strategic role in the Indonesian economy, particularly in employment generation, poverty reduction, and rural development. Despite its importance, substantial disparities in the welfare of agricultural enterprises persist across regions. According to the *Hasil Survei Ekonomi Pertanian (SEP)* 2024 published by the *Badan Pusat Statistik* (Badan Pusat Statistik, 2025), the average percentage of agricultural enterprises with very low income across Indonesian provinces stands at 5.22%, ranging from as low as 0% to as high as 16.6%. These figures illustrate that income inequality among agricultural enterprises is closely associated with regional characteristics and structural constraints, such as differences in access to resources, infrastructure, and institutional support (Meena et al., 2024). In Indonesia, these disparities are reflected in the persistently high percentage of agricultural enterprises classified as having very low income in several provinces, highlighting deep-rooted structural challenges in agricultural production systems and rural economic development.

Prior empirical studies in Indonesia have applied regression-based approaches to analyze agricultural and regional economic indicators, such as rice availability using panel regression models (Yuliana, 2022). However, these studies have often assumed that agricultural economic challenges are spatially independent across regions. More recent literature highlights the pivotal role of spatial dependence in shaping regional agricultural outcomes (Kopczewska & Elhorst, 2024; Meena et al., 2024). Income inequality and low productivity in the agricultural sector remain critical issues in developing countries, including Indonesia. Recent studies have highlighted the significance of spatial dependency in understanding these economic disparities. For instance, (Aqiilah et al., 2025) found that the reduction of agricultural land in East Java significantly exacerbates income inequality, with spatial interactions playing a crucial role. Similarly, (Theresia et al., 2025) demonstrated that food producer price volatility in Indonesia exhibits strong spillover effects across regions, necessitating spatially coordinated policy interventions. (Tran et al., 2025) report that although improved market access is positively associated with agricultural income, its effects tend to be spatially heterogeneous across regions. Consequently, neglecting spatial interdependencies in empirical analysis may result in biased estimations and suboptimal policy interventions.

The application of spatial econometrics has become increasingly relevant to capture these regional interdependencies. While Anselin (2022) and Ruttenauer (2019) established the foundational framework for spatial lag and error models, recent applications have expanded these methods to complex agricultural issues. (Saska, 2025), for example, utilized geographically weighted regression to model unemployment rates in Indonesia, a method analogous to analyzing labor absorption in agriculture. Sugiyanto et al. (2025) applied spatial analysis to examine labor absorption in the agricultural sector, revealing distinct spatial patterns in youth employment. Furthermore, empirical evidence from China by Chen et al. (2025) and Feng et al. (2025) suggests that factors such as environmental efficiency and digital finance also propagate spatially, influencing agricultural outcomes in neighboring regions. This study aims to contribute to this growing body of literature by explicitly modeling the spatial dependence of low-income agricultural enterprises in Indonesia using the Spatial Lag Model (SLM).

In response to these challenges, spatial econometric approaches have been increasingly employed in studies of agriculture and regional development, as they are well-suited to capturing geographic dependence and interregional interactions. Meena et al. (2024) emphasize that agricultural data commonly exhibit spatial autocorrelation due to similarities in agroclimatic conditions, infrastructure availability, and market connectivity among neighboring regions. Similarly, agricultural productivity and farmers' income tend to cluster spatially, providing strong empirical evidence of spatial clustering inherent in the agricultural sector.

Beyond spatial clustering, recent studies highlight the importance of spatial spillover effects in shaping agricultural welfare. Chen et al. (2025) show that farmers' income growth exhibits significant spatial autocorrelation, where improvements in one region generate indirect effects on neighboring regions. Similarly, Cui et al. (2025) find that agricultural agglomeration not only raises local income levels but also produces positive spillover effects across regions. Feng et al. (2025) further argue that in modern agricultural systems, indirect spatial effects may be as important as, or even more important than, direct effects. These findings suggest that policies targeting agricultural development in a single region may generate broader interregional impacts through spatial interaction mechanisms.

Although the spatial dimension of farmers' income has received increasing attention in the literature, most existing studies focus on average income levels, income growth, or overall agricultural productivity. Limited attention has been paid to agricultural enterprises with very low income, which represent the most economically vulnerable segment of the agricultural sector and play a crucial role in rural poverty alleviation. Moreover, empirical studies that explicitly examine spatial dependence and spillover effects of production-related constraints among very low-income agricultural enterprises, particularly in the Indonesian context, remain scarce.

From a statistical and spatial modeling perspective, the majority of empirical studies in the Indonesian agricultural context have relied on conventional regression approaches, including ordinary panel regression and simple cross-sectional models without formally testing for, or rigorously accounting for, spatial structure in the data. The systematic application of spatial diagnostic tools such as Moran's I, Lagrange Multiplier tests, and Robust LM tests, followed by formal model comparison and spatial impact decomposition, has remained underutilized in studies of agricultural enterprise welfare at the provincial level in Indonesia.

Addressing this research gap, the present study aims to analyze spatial dependence and spillover effects of production-related variables on the percentage of agricultural enterprises with very low income at the provincial level in Indonesia. Specifically, this study examines the roles of limited access to credit information, difficulties in obtaining production inputs, restricted market access, limited availability of production facilities, and small-scale capital constraints. The methodological contribution of this study lies in the systematic application of spatial econometric specification testing, including Moran's I for spatial autocorrelation detection, Lagrange Multiplier and Robust LM tests for model selection, and spatial impact decomposition for quantifying direct and indirect spillover effects within the context of agricultural enterprise welfare analysis in Indonesia. By employing a Spatial Lag Model, this study identifies both direct and indirect interprovincial effects, thereby providing empirical evidence to support the formulation of more spatially coordinated and regionally integrated agricultural development policies.

METHOD

Research Design

This section describes the data structure, variable definitions, and analytical framework employed in this study. To account for potential spatial interactions among provinces, the methodology integrates classical linear regression with spatial econometric techniques. The analysis begins with an examination of spatial dependence using global spatial autocorrelation measures, followed by the specification and estimation of spatial regression models. Model selection is conducted using Lagrange Multiplier diagnostics to identify the most appropriate form of spatial dependence.

Population and Sample

The population of this study consists of all agricultural enterprises at the provincial level in Indonesia as recorded in the SEP 2024. The unit of analysis is the province, and the data cover all 38 provinces that have been officially recognized following the administrative expansion of Indonesian provinces as of 2024. Since the study includes all units in the target population, the data represent a complete enumeration rather than a probabilistic sample.

Sampling Techniques

Given that data are available for all 38 provinces in Indonesia, this study employs a total sampling (census) approach. All provinces are included in the analysis without exclusion, ensuring that the spatial weight matrix constructed for this study captures the complete interprovincial neighborhood structure across the entire national territory. No probabilistic or purposive sampling was applied, as the secondary data source (SEP 2024) provides comprehensive provincial-level coverage.

Research Subjects

This study uses secondary data obtained from the official publication *Hasil Survei Ekonomi Pertanian* (SEP) 2024, published by the *Badan Pusat Statistik* (Badan Pusat Statistik, 2025). The data are collected at the provincial level and cover all 38 provinces in Indonesia as of the 2024 administrative division. As secondary data, the variables are derived from aggregated survey results without direct involvement of the authors in the data collection process.

The study involves one response variable and five explanatory variables. The response variable (Y) is the percentage of agricultural enterprises with very low income at the provincial level, used as a proxy for economic vulnerability in the agricultural sector. The five explanatory variables (X_1 through X_5) represent production-related constraints faced by agricultural enterprises, including limited knowledge of credit access, difficulties in obtaining production inputs, restricted market access, limited access to production facilities, and small-scale capital constraints. Complete definitions and descriptions of all variables are presented in Table 1.

Table 1. Definition and Description of Research Variables

Variable	Description
Y (Percentage of Very Low-Income Agricultural Enterprises)	This variable represents the percentage of agricultural enterprises in a province that fall into the very low-income category. It is used as an indicator of economic vulnerability and welfare conditions of agricultural enterprises.
X_1 (Percentage of Agricultural Enterprises with No Knowledge of Credit Access)	This variable measures the proportion of agricultural enterprises that lack knowledge of or information about access to formal credit. It reflects limitations in financial literacy and access to capital.
X_2 (Percentage of Agricultural Enterprises Facing Difficulties in Obtaining Production	This variable captures the proportion of agricultural enterprises experiencing constraints

Variable	Description
Inputs)	in obtaining production inputs, such as seeds, fertilizers, and pesticides.
X_3 (Percentage of Agricultural Enterprises with Limited Market Access)	This variable reflects the degree of difficulty faced by agricultural enterprises in accessing markets in terms of distribution channels, transportation, and price information.
X_4 (Percentage of Agricultural Enterprises with Limited Access to Production Facilities)	This variable indicates constraints in access to production facilities, including agricultural machinery, irrigation systems, and other supporting infrastructure.
X_5 (Percentage of Agricultural Enterprises with Small-Scale Capital Constraints)	This variable represents the proportion of agricultural enterprises experiencing limitations in small-scale working capital.

Each province is represented by its geographical coordinates (latitude and longitude), which are used to construct the spatial structure required for spatial econometric analysis.

The analytical framework integrates classical Ordinary Least Squares (OLS) regression with spatial econometric methods, specifically the Spatial Lag Model (SLM) and Spatial Error Model (SEM), to examine both direct and indirect (spillover) effects of production-related constraints on agricultural welfare across provinces.

Data Analysis Techniques

The data analysis proceeded in four stages. First, descriptive statistical analysis was conducted to characterize the distribution of all variables. Second, spatial autocorrelation was assessed using Moran's I statistic applied to both the response variable and the OLS residuals. Third, classical linear regression (OLS) was estimated as a baseline model. Fourth, spatial regression models, namely the Spatial Lag Model (SLM) and Spatial Error Model (SEM), were estimated and compared using the Akaike Information Criterion (AIC) and Lagrange Multiplier (LM) diagnostics. Spatial impact decomposition (direct, indirect, and total effects) was subsequently performed on the selected model. The analytical procedures are described in detail below.

In regional data analysis, it is common to observe that the value of a variable in one location is influenced by the values of the same variable in neighboring locations. This phenomenon is known as spatial dependence. As discussed in the spatial econometrics literature (Kopczewska & Elhorst, 2024), spatial dependence arises when observations across regions are not geographically independent, thereby violating the classical regression assumption of independently distributed observations.

Spatial dependence is mathematically represented through a spatial weight matrix, denoted by **W**, which describes the neighborhood structure among regions:

$$\boldsymbol{W} = \begin{bmatrix} 0 & w_{12} & \cdots & w_{1n} \\ w_{21} & 0 & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & 0 \end{bmatrix} \quad (1)$$

where:

- w_{ij} denotes the spatial relationship between region i and region j,
- $w_{ii} = 0$, since a region is not considered a neighbor of itself.

In spatial analysis, the weight matrix is commonly row-standardized, such that:

$$\sum_{j=1}^{n} w_{ij} = 1 \quad (2)$$

Spatial weights can be defined based on contiguity (rook or queen), geographical distance, or k-nearest neighbors. This study employs the k-nearest neighbors (KNN) approach with $k = 5$, meaning each province is assigned five nearest neighboring provinces based on Euclidean distances between provincial centroids. The value $k = 5$ was selected based on preliminary sensitivity analyses using alternative specifications of k = 4, 5, and 6. The results demonstrated consistent spatial patterns and stable parameter estimates across these specifications, as further confirmed in the robustness check presented in the Results section. Furthermore, $k = 5$ provides an adequate number of neighbors to capture spatial interdependencies among Indonesia's 38 provinces while avoiding excessive smoothing that may arise with larger k values. The choice of spatial weight matrix is crucial, as it directly affects the detection and estimation of spatial dependence (Anselin, 2022; Ruttenauer, 2019).

Before estimating spatial regression models, it is necessary to test for the presence of global spatial autocorrelation. One of the most widely used statistics for this purpose is Moran's I, defined as:

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2} \quad 3)$$

where:

- y_i is the value of the variable in region i,
- $\bar{y}$ is the sample mean,
- w_{ij} is an element of the spatial weight matrix.

The interpretation of Moran's I is as follows:

- $I > 0$: positive spatial autocorrelation (clustering of similar values),
- $I < 0$: negative spatial autocorrelation,
- $I \approx 0$: absence of spatial pattern.

A significant Moran's I indicates that the independence assumption of OLS residuals is violated, thereby justifying the use of spatial regression models.

As an initial benchmark, the classical linear regression model is specified as:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \quad (4)$$

Where:

- $\boldsymbol{y}$ is the dependent variable vector,
- $\boldsymbol{X}$ is the matrix of explanatory variables,
- $\boldsymbol{\beta}$: is the parameter vector,
- $\boldsymbol{\varepsilon} \sim N(0, \boldsymbol{\sigma}^2 \boldsymbol{I})$.

This model assumes the absence of spatial dependence in both the dependent variable and the error term.

In the presence of spatial dependence, spatial regression models provide a more appropriate analytical framework than conventional OLS models. The Spatial Lag Model explicitly incorporates interregional interaction through the dependent variable (Anselin, 2022; Ruttenauer, 2019), while recent developments further strengthen the econometric foundations of spatial autoregressive models (Kelejian & Piras, 2017). The general form of the SLM is given by:

$$\boldsymbol{y} = \boldsymbol{\rho W y} + \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \quad (5)$$

where:

- $\boldsymbol{\rho}$ is the spatial autoregressive coefficient,
- $\boldsymbol{Wy}$ is the spatially lagged dependent variable,

A significant ρ indicates that changes in the percentage of very low-income agricultural enterprises in one province are influenced by changes in neighboring provinces.

Unlike conventional regression models, coefficient interpretation in the SLM cannot be conducted directly. Following Anselin (2022) and Ruttenauer (2019), the impacts of explanatory variables must be decomposed into direct effects, indirect (spillover) effects, and total effects. The SLM can be rewritten as:

$$\boldsymbol{y} = (\boldsymbol{I} - \boldsymbol{\rho W})^{-1} \boldsymbol{X\beta} + (\boldsymbol{I} - \boldsymbol{\rho W})^{-1} \boldsymbol{\varepsilon} \quad (6)$$

From this expression:

1. Direct effects measure the impact of changes in X on y within the same region,
2. Indirect effects (spillovers) capture the impact transmitted to other regions,
3. Total effects are the sum of direct and indirect effects.

This impact analysis is essential for assessing how regional policies or improvements in production factors in one province may affect agricultural welfare in other provinces.

An alternative spatial regression specification is the Spatial Error Model (SEM), which accounts for spatial dependence in the error term rather than in the dependent variable itself. The SEM is appropriate (Kopczewska & Elhorst, 2024) when spatial dependence arises when spatial dependence arises from unobserved spatially correlated factors, such as institutional characteristics, local policies, or geographical conditions. The SEM is specified as:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{7}$$

with the error structure:

$$\boldsymbol{\varepsilon} = \boldsymbol{\lambda W \varepsilon} + \boldsymbol{u}, \quad u \sim N(0, \sigma^2 I) \tag{8}$$

or equivalently:

$$\boldsymbol{y} = \boldsymbol{X\beta} + (\boldsymbol{I} - \boldsymbol{\lambda W})^{-1}\boldsymbol{u} \tag{9}$$

where $\boldsymbol{\lambda}$ denotes the spatial error coefficient. A significant λ indicates the presence of spatial autocorrelation in the residuals, implying that unobserved influences tend to cluster geographically.

To determine the most appropriate spatial model specification (SLM or SEM), Lagrange Multiplier (LM) tests were conducted following the framework reviewed by Elhorst et al (2021) and Kopczewska & Elhorst (2024). These tests assess the significance of spatial lag and spatial error dependencies in the model residuals. Once spatial autocorrelation is detected in the dependent variable and/or OLS residuals, the next step is to determine the dominant form of spatial dependence, whether it arises from:

1. Spatial interaction in the dependent variable (spatial lag dependence), or
2. Spatial correlation in the error term (spatial error dependence).

For this purpose, Lagrange Multiplier (LM) tests as outlined in Elhorst et al. (2021) and Kopczewska & Elhorst (2024), are employed.

The LM Lag test examines whether spatial dependence occurs through a spatially lagged dependent variable.

Hypothesis:

$H_0: \rho = 0$ (no spatial lag dependence)

$H_1: \rho \neq 0$ (spatial lag dependence)

The test statistic is given by:

$$LM_{lag} = \frac{(e^{\top} W y / \sigma^2)^2}{\mathrm{tr}(W^{\top} W + W^2)} \tag{10}$$

where $\boldsymbol{e}$ denotes OLS residuals. The statistic follows a chi-square distribution with one degree of freedom. If the LM Lag is significant, then there is an indication that the value of the dependent variable in a region is influenced by the value of the dependent variable in neighboring regions.

The LM Error test evaluates whether spatial dependence occurs in the error term.

Hypotheses:

$H_0: \lambda = 0$ (no spatial error autocorrelation)

$H_1: \lambda \neq 0$ (spatial error autocorrelation)

The test statistic is defined as:

$$LM_{error} = \frac{(e^{\top} W e / \sigma^2)^2}{\mathrm{tr}(W^2 + W^{\top} W)} \tag{11}$$

Like the LM Lag, the LM Error statistic also follows a chi-square distribution with one degree of freedom.

In empirical applications, it is common for both the LM Lag and LM Error tests to be statistically significant simultaneously. This situation may lead to ambiguity in identifying the dominant form of spatial dependence. To address this issue, Robust Lagrange Multiplier (Robust LM) tests were developed to adjust each LM statistic by accounting for the presence of alternative forms of spatial dependence.

The Robust LM Lag test examines the significance of spatial lag dependence while accounting for potential spatial autocorrelation in the error term.

Hypothesis:

$H_0: \rho = 0$

The Robust LM Lag statistic is given by:

$$RLM_{lag} = \frac{\left(e^T Wy/\sigma^2 - \frac{tr(W^T W)}{tr(W^T W + W^2)} e^T We/\sigma^2\right)^2}{tr(W^T W + W^2) - \frac{[tr(W^T W)]^2}{tr(W^T W + W^2)}} \quad (12)$$

where:

- $\boldsymbol{e}$ denotes the OLS residual vector,
- $\boldsymbol{\sigma^2}$ is the estimated variance of the OLS residuals,
- $\boldsymbol{W}$ is the spatial weight matrix.

Under the null hypothesis, the statistic follows an asymptotic chi-square distribution with one degree of freedom. A significant Robust LM Lag statistic indicates that the Spatial Lag Model (SLM) is more appropriate, even after controlling for possible spatial error dependence.

The Robust LM Error test evaluates the presence of spatial autocorrelation in the error term while controlling for potential spatial lag dependence in the dependent variable. The hypotheses are:

$H_0: \lambda = 0$

The Robust LM Error statistic is defined as:

$$RLM_{error} = \frac{\left(e^T We/\sigma^2 - \frac{tr(W^T W)}{tr(W^T W + W^2)} e^T Wy/\sigma^2\right)^2}{tr(W^T W + W^2) - \frac{[tr(W^T W)]^2}{tr(W^T W + W^2)}} \quad (13)$$

This statistic also follows a chi-square distribution with one degree of freedom under the null hypothesis. A significant Robust LM Error statistic suggests that the Spatial Error Model (SEM) is the preferred specification.

RESULTS AND DISCUSSION

Conceptual Distinction

Before presenting the empirical results, it is important to clarify the conceptual distinction between spatial dependence and spatial spillover effects, as both terms appear prominently in the title and research objectives of this study. Spatial dependence refers to the statistical property whereby observations at one location are correlated with observations at neighboring locations, as formalized by Moran's I and the spatial autoregressive parameter ρ in the SLM. In contrast, spatial spillover effects refer to the substantive economic transmission mechanism by which changes in an explanatory variable in one province generate ripple effects on the dependent variable in neighboring provinces. While spatial dependence is a statistical phenomenon detected through diagnostic tests (Moran's I, LM tests), spillover effects are the policy-relevant economic impacts quantified through the spatial impact decomposition analysis (Table 4). In this study, the detection of statistically significant spatial dependence (ρ) provides the empirical basis for the estimation and interpretation of spillover effects.

Descriptive statistic

Descriptive statistical analysis was conducted for 38 provinces in Indonesia to provide an initial overview of the economic welfare of agricultural enterprises and the structural constraints they face. The summary statistics of the response and explanatory variables are presented in Table 2.

Table 2. Descriptive Statistics of Research Variables

Variable	Min	Max	Mean	Median	SD
Y	0.00	16.60	5.22	4.90	3.29
X_1	6.03	94.60	28.10	22.90	17.70
X_2	1.69	89.40	42.40	39.90	22.60
X_3	0.36	57.90	13.90	9.36	12.80
X_4	4.15	81.90	28.10	23.30	16.90
X_5	6.70	65.80	35.10	34.70	12.30

Descriptively, variable Y, representing the percentage of agricultural enterprises with very low income, has a mean value of 5.22 percent with a standard deviation of 3.29 and a range from 0.00 to 16.60 percent. This indicates a moderate degree of variation in the welfare conditions of agricultural enterprises across provinces. Variable X_1 (percentage of agricultural enterprises with no knowledge of credit access) has an average value of 28.10

percent and exhibits very high variability (SD = 17.70), with values ranging from 6.03 percent to as high as 94.60 percent. The median value is lower than the mean, suggesting a right-skewed distribution and substantial inequality in financial literacy and access to formal credit across regions.

Variable X_2 (percentage of agricultural enterprises facing difficulties in obtaining production inputs) records the highest mean among the explanatory variables, at 42.40 percent, with a standard deviation of 22.60 and a wide range (1.69–89.40 percent). This reflects considerable interprovincial disparities in the availability of production inputs such as seeds, fertilizers, and pesticides.

Variable X_3 (percentage of agricultural enterprises with limited market access) has a relatively low mean value of 13.90 percent, but a nearly comparable standard deviation (12.80). This indicates extreme disparity, where most provinces experience relatively minor market access constraints, while a small number of provinces face very severe marketing barriers.

Variable X_4 (percentage of agricultural enterprises with limited access to production facilities) has an average value of 28.10 percent and a standard deviation of 16.90, with a maximum value reaching 81.90 percent. This suggests substantial inequality in the availability of agricultural infrastructure, such as machinery, irrigation, and supporting facilities, across provinces.

Meanwhile, variable X_5 (percentage of agricultural enterprises with small-scale capital constraints) has a mean of 35.10 percent and a standard deviation of 12.30, with a median close to the mean value. This pattern indicates a relatively homogeneous distribution compared to other variables, implying that small-scale capital constraints constitute a common challenge faced by agricultural enterprises in nearly all provinces.

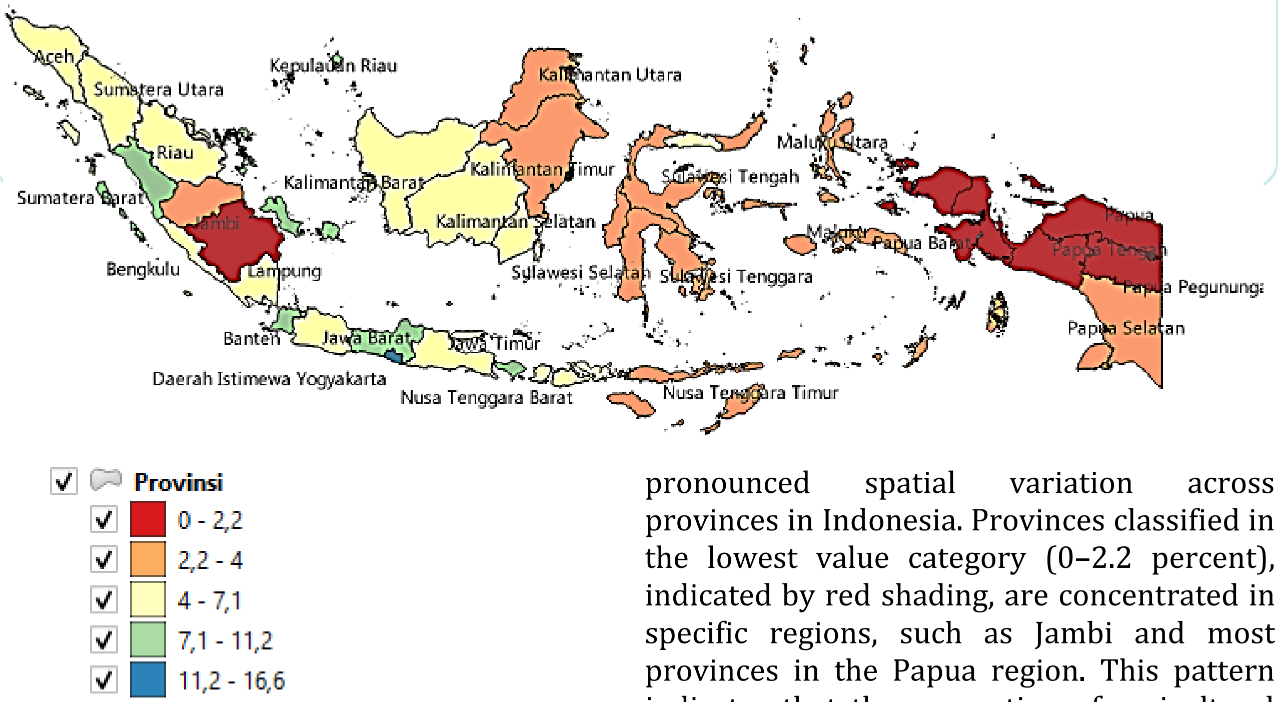

Figure 1. Geographic Distribution of Very Low-Income Agricultural Enterprises in Indonesia

The spatial distribution map of variable **Y**, which represents the percentage of agricultural enterprises with very low income, reveals pronounced spatial variation across provinces in Indonesia. Provinces classified in the lowest value category (0–2.2 percent), indicated by red shading, are concentrated in specific regions, such as Jambi and most provinces in the Papua region. This pattern indicates that the proportion of agricultural enterprises with very low income in these areas is relatively small compared to other provinces.

In contrast, provinces with higher values of **Y**, particularly those in the ranges of 7.1–

11.2 percent and 11.2–16.6 percent, tend to be concentrated in western Java, Bali, and several provinces in Sumatra and Kalimantan. This spatial pattern suggests that although Java is generally recognized for relatively higher agricultural productivity, structural pressures such as land fragmentation and the predominance of small-scale farming continue to contribute to a high proportion of very low-income agricultural enterprises in certain provinces.

Furthermore, the presence of clusters of provinces with similar values of **Y** indicates potential spatial dependence, implying that the welfare condition of agricultural enterprises in a given province does not evolve independently but is influenced by the characteristics of neighboring regions. This visual evidence provides preliminary support for the existence of spatial autocorrelation and justifies the application of spatial regression models to capture interprovincial spillover effects in explaining variations in the percentage of agricultural enterprises with very low income.

Spatial Autocorrelation

The results of Moran's I test yield an index value of 0.528 with a very small p-value of 3.8×10^{-11} ($p < 0.01$), indicating the presence of strong positive spatial autocorrelation in variable Y. This implies that geographically adjacent provinces tend to exhibit similar values of the percentage of agricultural enterprises with very low income.

Moran's I test applied to the residuals of the OLS model also reveals significant spatial autocorrelation, with a p-value of 3.567×10^{-6}, indicating a violation of the residual independence assumption. Consequently, the OLS model is not adequate as an inferential model, and a spatial regression framework is required to properly account for the underlying spatial structure in the data.

Lagrange Multiplier Test Results

The results of the spatial dependence diagnostics using the Lagrange Multiplier (LM) tests show that both the LM Lag statistic (RSlag = 18.213; p-value = 1.98×10^{-5}) and the LM Error statistic (RSerr = 13.031; p-value = 0.0003) are statistically significant at the 5 percent significance level. These findings indicate the presence of spatial dependence in both the dependent variable (spatial lag dependence) and the error term (spatial error dependence) within the OLS framework. Therefore, the classical linear regression model is insufficient to fully capture the spatial structure inherent in the data.

To identify the dominant form of spatial dependence, the Robust LM tests are subsequently employed. The results indicate that the Robust LM Lag statistic remains significant (adjRSlag = 5.370; p-value = 0.0205), whereas the Robust LM Error statistic is not significant (adjRSerr = 0.189; p-value = 0.664). This outcome suggests that the observed spatial dependence primarily arises from interregional interaction mechanisms through the dependent variable, rather than from spatial correlation in the error component.

Accordingly, the Spatial Lag Model (SLM) is identified as the most appropriate specification for analyzing the factors influencing the percentage of agricultural enterprises with very low income. Substantively, these results imply that the proportion of very low-income agricultural enterprises in a given province is influenced not only by its internal characteristics but also by the welfare conditions of agricultural enterprises in neighboring provinces. The presence of such spillover effects highlights the importance of adopting cross-regional policy approaches, as interventions implemented in one province may generate indirect impacts on surrounding regions.

Estimation Result

To examine the presence of spatial effects and assess the robustness of the estimated relationships, this study compares the results obtained from the Ordinary Least Squares (OLS) model with those from spatial econometric models, namely the Spatial Lag Model (SLM) and the Spatial Error Model (SEM). Table 3 presents the estimation results and goodness-of-fit measures for each model.

Table 3. Comparison of OLS and Spatial Regression Results

Variable	OLS Coefficient	SLM Coefficient	SEM Coefficient
Intercept	8.674***	4.673**	8.746***
X_1	-0.05**	-0.033*	-0.041**
X_2	-0.012	-0.018	-0.015
X_3	-0.011	0.003	-0.016
X_4	-0.048**	-0.022*	-0.011
X_5	0	-0.008	-0.02
Spatial Parameter	–	ρ = 0.622***	λ = 0.669***
AIC	199.45	187.87	189.22

*** p < 0.01; ** p < 0.05; * p < 0.10

The estimation results reveal notable differences in both coefficient estimates and model performance between the OLS regression and the spatial regression models. In the OLS model, most explanatory variables exhibit negative coefficients, although not all are statistically significant. Variables representing lack of knowledge of credit access (X_1) and limited access to production facilities (X_4) have statistically significant effects on the percentage of very low-income agricultural enterprises, while the remaining variables are not significant. However, given the presence of significant spatial dependence as indicated by the Lagrange Multiplier tests, the OLS estimates may be biased and inefficient due to the omission of interregional interactions.

In the Spatial Lag Model (SLM), the spatial autoregressive parameter ρ is positive and statistically significant, confirming the existence of strong spatial dependence in the dependent variable. The results confirm the presence of positive spatial autocorrelation, consistent with findings by Ananda et al. (2025), who observed similar spatial clustering in Indonesia's middle-income trap phenomenon. This suggests that low-income agricultural enterprises are not randomly distributed but are influenced by the economic performance of neighboring provinces. This finding aligns with the spillover effects of high-value agriculture adoption observed by Guo et al. (2026), where regional integration helps narrow the income gap. This result implies that an increase in the percentage of very low-income agricultural enterprises in one province significantly raises the corresponding proportion in neighboring provinces. After accounting for spatial interaction effects, variables X_1 and X_4 remain statistically significant with negative coefficients, although their magnitudes are smaller than those obtained from the OLS model. This finding suggests that part of the effects of limited credit knowledge and restricted access to production facilities is transmitted through spatial spillover mechanisms. Meanwhile, variables X_2 and X_3 are not statistically significant, indicating that their impacts on agricultural welfare are likely to be indirect and context-dependent across regions rather than purely local.

In the Spatial Error Model (SEM), the spatial error parameter λ is statistically significant, indicating the presence of spatial correlation in the error term. This result suggests that unobserved factors with spatial patterns, such as institutional characteristics, local policies, or geographical conditions not explicitly included in the model, also influence the percentage of very low-income agricultural enterprises. The regression coefficients in the SEM are relatively close to those obtained from the OLS model and are largely insignificant, reflecting that the SEM primarily serves to correct the error structure rather than to capture direct interaction mechanisms across the dependent variable.

Based on the Akaike Information Criterion (AIC), the Spatial Lag Model yields the lowest AIC value compared to the OLS and SEM models, indicating that it provides the best balance between goodness-of-fit and model complexity. This result is consistent with the Robust Lagrange Multiplier tests, which point to the dominance of spatial lag dependence, thereby reinforcing the selection of the SLM as the most appropriate model specification.

To further assess the robustness of the spatial regression results, the Spatial Lag Model was re-estimated using alternative k-nearest neighbors spatial weight matrices with k = 4, 5, and 6. The results show that the key explanatory variables, particularly X_1 and X_4, consistently retain their negative signs across all spatial weight specifications. Moreover, the spatial autoregressive parameter ρ remains positive, with estimated values ranging from 0.56 to 0.67, indicating stable and substantial spatial spillover effects. In addition, the Akaike Information Criterion varies only slightly, decreasing from 188.45 (k = 4) to 186.99 (k = 6). These findings demonstrate that the main empirical results are not driven by a specific choice of spatial weight matrix and are robust to alternative definitions of spatial interaction.

Spatial Impact Analysis

The direct and indirect (spillover) effects reported in Table 4 are derived from the full impact matrix $\boldsymbol{S}(\boldsymbol{W}) = (\boldsymbol{I} - \boldsymbol{\rho W})^{-1}$. The diagonal elements of the partial derivative matrix $\partial y/\partial x_k$ are averaged to yield the direct effect for each province, while the off-diagonal elements are averaged to yield the indirect (spillover) effect transmitted to other provinces. These summary measures account for the circular feedback effects inherent in the spatial autoregressive process, which is why direct effects in Table 4 may differ slightly from the point estimates in Table 3.

Table 4. The Results of The Impact Analysis

Variable	Direct Effect	Indirect Effect	Total Effect
X_1	-0.036	-0.050	-0.087
X_2	-0.019	-0.027	-0.046
X_3	0.003	0.004	0.007
X_4	-0.024	-0.034	-0.058
X_5	-0.009	-0.012	-0.021

The impact analysis results indicate that the influence of the explanatory variables on the percentage of agricultural enterprises with very low income occurs not only directly within a province but also spreads to other provinces through spatial dependence mechanisms. Variable X_1 (percentage of agricultural enterprises with no knowledge of credit access) exhibits the largest total effect (−0.087), with a direct effect of −0.036 and an indirect spillover effect of −0.050. The dominance of the indirect over the direct effect underscores the cross-provincial transmission of financial literacy conditions.

Variable X_2 (percentage of agricultural enterprises facing difficulties in obtaining production inputs) shows a direct effect of −0.019 and an indirect effect of −0.027. yielding a total effect of −0.046. The fact that the indirect effect exceeds the direct effect highlights the cross-regional nature of input supply constraints. implying that disruptions in input availability in one province can increase the proportion of very low-income agricultural enterprises in surrounding provinces.

In contrast. variable X_3 (percentage of agricultural enterprises with limited market access) has very small positive direct and indirect effects. with a total effect of 0.007. This result indicates that market access constraints tend to be localized and segmented. and therefore do not generate strong spillover effects at the interprovincial level.

Variable X_4 (percentage of agricultural enterprises with limited access to production facilities) also exhibits a relatively large total effect (−0.058). consisting of a direct effect of

−0.024 and an indirect effect of −0.034. This pattern suggests that limited access to production facilities not only affects local agricultural efficiency but also reduces agricultural performance in other regions through interprovincial spatial linkages.

Variable X_5 (percentage of agricultural enterprises with small-scale capital constraints) displays a relatively small total effect (−0.021). reflecting that capital constraints are relatively uniform across regions and do not serve as a primary driver of spatial variation in the percentage of very low-income agricultural enterprises.

Overall, the dominance of indirect effects for most explanatory variables underscores the importance of regionally coordinated agricultural policies, as interventions implemented in one province have the potential to generate broader welfare benefits for surrounding regions through spatial spillover mechanisms.

Policy Implications and Discussion

The estimation results and spatial impact analysis of the Spatial Lag Model (SLM) provide several important policy implications. First, variable X_1 (lack of knowledge of credit access) exhibits the largest negative total effect on variable Y. Interestingly, the result shows a negative relationship between lack of credit knowledge and low income. This implies that farmers with higher financial literacy or credit access might be exposed to higher risks, such as high-interest loans or debt traps, which negatively affect their net income. Conversely, farmers who are unaware of credit access might rely on self-funding or traditional methods that, while limited, prevent them from significant financial liabilities.

Second, variables X_4 (limited access to production facilities) and X_2 (difficulties in obtaining production inputs) also demonstrate relatively large negative total effects. These results emphasize that improvements in production infrastructure and the smooth distribution of agricultural inputs should be designed in a coordinated interprovincial manner. Given the strong spatial linkages present in agricultural production systems.

Third, variable X_3 (limited market access) shows a relatively small and positive total effect. This suggests that market access constraints tend to be more localized in nature and generate weaker spillover effects compared to financial access and production facility constraints.

Fourth, variable X_5 (small-scale capital constraints) has negative direct and indirect effects. although their magnitudes are smaller than those of X_1 and X_4. This indicates that small-scale capital policies remain important; however. their effectiveness is likely to be enhanced when integrated with broader policies aimed at improving credit access and expanding production facilities.

Overall, these findings underscore that policies aimed at reducing income vulnerability in the agricultural sector cannot be implemented in a partial or purely administrative manner. Given the strong spatial spillover effects identified. an area-based policy approach that emphasizes cross-provincial coordination particularly in financing. provision of production facilities. and strengthening of agricultural input systems is essential for achieving more effective and equitable agricultural development outcomes.

CONCLUSION

Conclusion

This study makes two principal contributions. First, it provides empirical evidence that the percentage of agricultural enterprises with very low income at the provincial level in Indonesia exhibits strong positive spatial dependence with a Moran's I of 0.528 and that this dependence is best captured by the Spatial Lag Model (SLM), as confirmed through systematic application of

Lagrange Multiplier and Robust LM specification tests. Second, through spatial impact decomposition, this study demonstrates that the effects of production-related constraints operate not only directly within individual provinces but also indirectly through spatial spillover mechanisms, with indirect effects often exceeding direct effects for key variables.

This distinction between spatial dependence, a statistical property of the data confirmed by significant and spatial spillover effects, and the substantive interprovincial transmission of production constraint impacts constitutes a key analytical contribution of this study. From a methodological standpoint, the systematic application of spatial diagnostic tools and formal model comparison within the context of agricultural enterprise welfare in Indonesia fills an important gap in the existing literature, where conventional regression approaches have predominated.

These findings imply that agricultural development policies should be designed in a spatially coordinated and regionally integrated manner. Policies focusing on credit access, production facilities, and input availability are likely to be more effective when formulated through cross-provincial collaboration, given the presence of substantial spatial spillovers. Specifically, the findings support the implementation of: (1) spatially targeted financial literacy programs and subsidized credit schemes (such as *Kredit Usaha Rakyat*) coordinated across provincial borders; (2) integrated supply chain programs for agricultural inputs (fertilizers, seeds, pesticides) that account for interprovincial distribution linkages; and (3) cross-provincial agricultural infrastructure development programs that prioritize provinces with high spatial spillover potential, as identified in the impact decomposition results.

Suggestion

For future research, several extensions can be considered. First, the analysis may be expanded by incorporating spatio-temporal data to capture dynamic spatial interactions over time. Second, alternative spatial specifications such as the Spatial Durbin Model (SDM) could be employed to examine more complex transmission mechanisms involving both dependent and independent variables. Finally, the inclusion of more granular data, such as district- or farm-level observations, may provide deeper insights into local heterogeneity and strengthen the empirical basis for targeted agricultural policy interventions.

BIBLIOGRAPHY

Ananda, W. K. & Nuryadin, D. (2025). Spatial Panel Regression Modelling of Middle-Income Trap in Indonesia. Elastisitas : Jurnal Ekonomi Pembangunan, 7(1), 9–18. https://elastisitas.unram.ac.id/index.php/elastisitas/article/view/178

Anselin, L. (2022). Spatial econometrics. In Handbook of Spatial Analysis in the Social Sciences: (pp. 101–122). Edward Elgar Publishing. https://ideas.repec.org/h/elg/eechap/19110_6.html

Aqiilah, H. S., Sujarwo, S., & Toiba, H. (2025). Analysis of Agricultural Land area decrease on income inequality in East Java, Indonesia. *Agro Bali Agricultural Journal*, 8(3), 873–884. https://doi.org/10.37637/ab.v8i3.2162

Badan Pusat Statistik. (2025). Hasil Survei Ekonomi Pertanian (SEP) 2024. https://www.bps.go.id/id/publication/2024/12/20/90b5aa3fa4244575063ba0e7/hasil-survei-ekonomi-pertanian--sep--2024.html

Chen, Y., Deng, X., & Sun, B. (2025). Research on the spatial spillover effect of increasing farmers' income on environmental efficiency: evidence from China. *Frontiers in Sustainable Food Systems*, 9. https://doi.org/10.3389/fsufs.2025.1611018

Cui, Y., Ma, X., & Zhang, Q. (2025). Agricultural geographical indications, industrial agglomeration, and farmers' income: evidence from China. *Agricultural and Food Economics*, 13(1). https://doi.org/10.1186/s40100-025-00410-5

Elhorst, J. P., Gross, M., & Tereanu, E. (2020). Cross-Sectional Dependence And Spillovers In Space And Time: Where Spatial Econometrics And Global Var Models Meet. *Journal of Economic Surveys*, 35(1), 192–226. https://doi.org/10.1111/joes.12391

Feng, S., Long, Y., Yasmeen, R., & Yang, H. (2025). The Spatial Spillover Impact of Digital Finance on Agricultural Carbon Emission Intensity: Evidence from China. *Sustainability*, 17(19), 8583. https://doi.org/10.3390/su17198583

Guo, Z., Lu, M., & Su, L. (2026). Can high-value agriculture narrow the urban-rural income gap? A policy-based evidence from China. *China Agricultural Economic Review*, 1–19. https://doi.org/10.1108/caer-06-2025-0299

Kelejian, H., & Piras, G. B. T.-S. E. (2017). Spatial Econometrics. Academic Press. https://www.researchgate.net/publication/321062228_Spatial_Econometrics

Kopczewska, K., & Elhorst, P. (2023). New developments in spatial econometric modelling. *Spatial Economic Analysis*, 19(1), 1–7. https://doi.org/10.1080/17421772.2023.2281173

Meena, A., Manjunatha, B., Kumar, A. A., Sunil, J., Shankar, S. V., Manjubala, M., D, A., & Behera, H. S. (2024). Spatial Econometrics in agriculture: Modelling spatial dependencies in data. *Archives of Current Research International*, 24(6), 488–506. https://doi.org/10.9734/acri/2024/v24i6807

Rüttenauer, T. (2019). Spatial Regression models: A systematic comparison of different model specifications using Monte Carlo experiments. *Sociological Methods & Research*, 51(2), 728–759. https://doi.org/10.1177/0049124119882467

Saska, I., & Sifriyani, S. (2025). Implementation Of Geographically Weighted Panel Regression With Gaussian Kernel Weighting Function In The Open Unemployment Rate Model. *Barekeng Jurnal Ilmu Matematika Dan Terapan*, 19(2), 733–742. https://doi.org/10.30598/barekengvol19iss2pp733-742

Sugiyanto, A., Gunarto, T., Lampung, U., Taher, A. R. Y., & Lampung, U. (2025). Spatial Analysis of Agricultural Sector labor absorption among Generation Z in Indonesia 2024. *Journal of Economics Finance and Management Studies*, 08(07). https://doi.org/10.47191/jefms/v8-i7-21

Theresia, A., Ikhsan, M., Kacaribu, F. N., & Sumarto, S. (2025). Spillover effect of food producer price volatility in Indonesia. *Economies*, 13(9), 256. https://doi.org/10.3390/economies13090256

Tran, D. L., Nguyen, V. C., Tran, H., Illukpitiya, P., Nguyen, M. T., & Tran, M. T. (2025). Market accessibility and its impact on the agricultural income of farmers. *Journal of Agribusiness in Developing and Emerging Economies*, 1–17. https://doi.org/10.1108/jadee-12-2024-0431

Yuliana, U. A. (2022). Pemodelan regresi data panel untuk memprediksi ketersediaan beras di Kabupaten Bojonegoro. *Jurnal Statistika Dan Komputasi*, 1(1), 1–11. https://doi.org/10.32665/statkom.v1i1.447