JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI e-ISSN: 2614-8811
Published by Departement of Mathematics, Hasanuddin University, Indonesia
https://journal.unhas.ac.id/index.php/jmsk/index
Vol. 23, No. 1, September 2026, pp. 111-125
DOI: 10.20956/1tmzsm24

Integrating Bibliometric Analysis Using VosViewer and Path Analysis to Model Farmer Loyalty

Viky Iqbal Azizul Alim[1*], Candra Rezzining Wulat Sariro Weni Utomo[2], Arifin[3], Firdha Mahrifatul Zana[4]

[1,3,4]Departement Mathematics, State University of Malang
[2]Departement Statistics, Brawijaya University

Email:[1] *vikyiqbal.fmipa@um.ac.id*, [2]*rezziningcandra@gmail.com*, [3]*arifin.fmipa@um.ac.id*, [4]*firdhamz.fmipa@um.ac.id*
**Corresponding author*
Received: 5 May 2026, revised: 27 July 2026, accepted: 3 August 2026

Abstract

This study aims to develop a comprehensive model of farmer loyalty by integrating bibliometric mapping and path analysis. A Systematic Literature Review (SLR) combined with bibliometric analysis using VOSviewer was conducted on 434 Scopus-indexed documents to identify key variables. The results revealed that farmer needs is a dominant variable with strong relevance in the research network. Based on these findings, a path analysis model was constructed to examine the relationships among farmer needs, farmer success, and farmer loyalty. The results indicate that farmer needs and farmer success have significant direct effects on farmer loyalty. Furthermore, farmer success partially mediates the relationship between farmer needs and farmer loyalty, indicating both direct and indirect influences. The model achieved a high coefficient of total determination (R_m^2=0.9340), reflecting the strong theoretical relationships among the selected constructs identified through bibliometric analysis rather than data-driven model optimization. This study contributes methodologically by integrating bibliometric mapping with statistical modeling and provides practical insights for improving farmer loyalty through the fulfillment of farmer needs and support for agricultural success.

Keywords: Bibliometric Mapping; Bootstrap Resampling; Path Analysis; Systematic Literature Review.

1. INTRODUCTION AND PRELIMINARIES

Statistics is a scientific approach that plays a crucial role in analyzing data to generate accurate information and support data-driven decision-making [20]. Over time, statistical analysis methods have expanded beyond simple relationships between variables to accommodate more complex relationships. One method widely used to model causal relationships is path analysis, which is an extension of regression analysis and is capable of representing simultaneous relationships through several structural equations [6]. Path analysis can identify direct and indirect effects between variables by involving exogenous variables, endogenous variables, and intervening variables [13].

In the agricultural sector, particularly in Indonesia, farmer loyalty has become an important issue following changes in government policies regarding subsidized fertilizer distribution. The implementation of Minister of Agriculture Regulation No. 10 of 2022 limits subsidized fertilizers to urea and NPK, influencing farmers' purchasing behavior and preferences for agricultural inputs. Such policy changes may indirectly affect farmers' loyalty toward fertilizer providers and agricultural support programs. Farmer loyalty reflects a long-term commitment to continue using, purchasing, and recommending specific products or services and is influenced by various factors related to perceived benefits, satisfaction, and farming outcomes [16].

The relationship between farmer success and farmer loyalty can be explained through Social Exchange Theory, which proposes that individuals tend to maintain relationships when the perceived benefits obtained from an exchange exceed the associated costs [4]. In the agricultural context, successful farming outcomes, including increased productivity, improved income, and better farm performance, generate positive evaluations of agricultural products, services, and institutional support. These positive experiences strengthen farmers' trust and commitment, thereby encouraging them to continue using the same products or services and maintain long-term relationships with agricultural stakeholders. This perspective is also consistent with the Expectation-Confirmation Theory, which explains that positive performance outcomes strengthen satisfaction and continuance intentions [3]. Therefore, farmer success is theoretically expected to be an important antecedent of farmer loyalty.

Recent studies have demonstrated that farmer loyalty is influenced by multiple behavioral, relational, and managerial factors. Hien and Kim [14] found that relationship quality significantly affects farmer loyalty in agribusiness partnerships, emphasizing the importance of trust and long-term collaboration between farmers and agribusiness companies. Similarly, Rutsaert et al. [19] reported that brand loyalty, product information, and price incentives significantly influence farmers' decisions in purchasing agricultural inputs. These findings indicate that loyalty is not only determined by product performance but also by relational and informational aspects. Furthermore, Asriadi et al. [2] showed that socio-economic and managerial factors significantly influence farmer participation and agricultural development through a Structural Equation Modeling (SEM) framework. Their findings highlight the importance of understanding complex causal relationships among factors affecting agricultural performance. These empirical findings support the theoretical proposition that successful farming experiences create positive evaluations, strengthen farmers' commitment, and ultimately increase their long-term loyalty toward agricultural products, services, and institutional support [16][24].

Alongside the development of structural modeling approaches, bibliometric analysis has emerged as a valuable method for identifying research trends, dominant themes, and knowledge structures within a scientific field. Bibliometric mapping supported by software such as VOSviewer enables researchers to quantitatively analyze scientific publications and objectively identify influential variables and research clusters [27]. Donthu et al. [7] further emphasized that bibliometric analysis provides a systematic approach for exploring intellectual structures and emerging research directions. Recent studies by Alamsyah et al. [1] and Xu et al. [25] demonstrated that bibliometric mapping is effective for identifying dominant themes and future research opportunities in agricultural innovation and sustainable agriculture. These studies confirm the usefulness of bibliometric techniques in extracting relevant variables from large bodies of scientific literature.

Despite the increasing number of studies on farmer behavior and agricultural performance, several important limitations remain. First, previous studies generally investigate individual determinants of farmer loyalty, such as relationship quality [11], purchasing behavior [15], or socioeconomic and managerial factors [2], without simultaneously examining the structural relationships among farmer needs, farmer success, and farmer loyalty within a unified causal framework. Consequently, the mechanisms through which these variables interact remain insufficiently explained. Second, the selection of variables in conventional path analysis is commonly based on theoretical assumptions or findings from a limited number of previous studies. While this approach is widely accepted, it may overlook emerging concepts that have recently gained attention in the scientific literature. Third, although bibliometric analysis has been widely applied to identify research trends, thematic clusters, and influential topics [1][7][25][27], most studies use it only as a descriptive mapping technique. The identified themes are rarely transformed into empirically testable constructs or validated through statistical causal modeling. As a result, the potential of bibliometric analysis to support objective model development remains underutilized.

To address these limitations, this study integrates bibliometric mapping using VOSviewer with path analysis to develop a more objective and theoretically grounded model of farmer loyalty. Unlike conventional path analysis, where research constructs are determined primarily through literature review and researcher judgment, the proposed framework first identifies influential and emerging concepts using bibliometric evidence derived from keyword co-occurrence networks and thematic clusters. These concepts are subsequently synthesized through a systematic literature review before being incorporated into the structural model. This integrated approach offers several methodological advantages. First, it reduces subjectivity in construct selection by incorporating quantitative bibliometric evidence. Second, it enables the inclusion of emerging concepts that may not yet be widely represented in conventional structural models, thereby strengthening the novelty of the proposed framework. Third, it extends the role of bibliometric analysis beyond descriptive knowledge mapping by statistically validating the direct and indirect relationships among the identified constructs through path analysis.

Accordingly, this study contributes methodologically by demonstrating how bibliometric analysis and path analysis can be integrated into a single evidence-based framework for model development. This integration not only improves the transparency and objectivity of construct selection but also provides a more comprehensive understanding of the causal mechanisms underlying farmer loyalty. Practically, the proposed model provides useful insights for agricultural companies and policymakers in designing strategies that strengthen farmer loyalty and promote sustainable agricultural development.

2. MAIN RESULTS

2.1 Data Collection

This study integrates bibliometric analysis using VOSviewer with statistical modeling to analyze farmer loyalty. In line with the research objectives, the methodology consisted of two sequential stages: bibliometric analysis and empirical statistical modeling.

The first stage employed a Systematic Literature Review (SLR) using the Scopus database to identify scientific publications related to farmer loyalty and its associated determinants. The literature search was restricted to recent publications and screened based on predefined inclusion and exclusion criteria, including publication year, subject area, document type, language, and relevance to the research topic. The selected publications were subsequently analyzed using VOSviewer to identify research themes and determine the research constructs used in the proposed path analysis model. The detailed search strategy, screening procedure, and bibliometric results are presented in Section 3.1.

The questionnaire employed a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree) to measure respondents' perceptions of the research constructs.

The proposed model consisted of three latent variables, namely SLR Result (X_1), Farmer Success (Y_1), and Farmer Loyalty (Y_2). The study population comprised farmers in West Nusa Tenggara Province who used subsidized fertilizer distributed by one of the Indonesian State-Owned Enterprise (SOE) fertilizer companies, particularly SP-36 fertilizer users.

A total of 100 respondents were selected using purposive sampling. Respondents were required to be active farmers, users of subsidized SP-36 fertilizer, and willing to participate in the survey. The sample size was determined based on the recommendation of Hair et al. (2019), which suggests a minimum of five to ten observations for each measurement indicator in path analysis and structural equation modeling. Since the questionnaire consisted of 11 indicators, the recommended minimum sample ranged from 55 to 110 respondents. Therefore, the sample of 100 respondents was considered sufficient to estimate the proposed path model and produce stable parameter estimates.

2.2 Reseacrh Variables

This study uses variables of SLR Result (X_1), Farmer Success (Y_1), and Farmer Loyalty (Y_2). The research model can be seen in **Figure 1**.

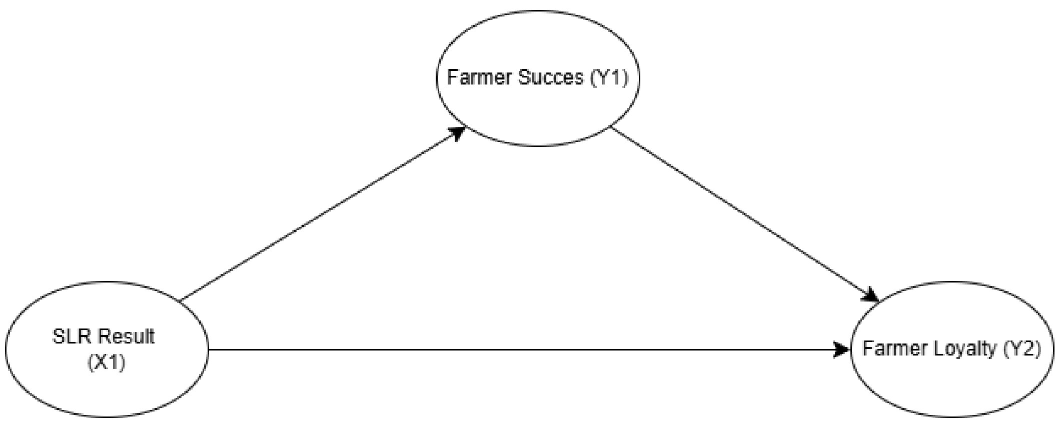

Figure 1**. Research Model**

The variables SLR Result (X_1), Farmer Success (Y_1), and Farmer Loyalty (Y_2) were measured using Likert-scale indicators, resulting in ordinal data. The responses were transformed using the Summated Rating Scale (SRS) to obtain composite scores with interval-scale properties. SRS was selected because it provides a straightforward and consistent transformation of ordinal responses while preserving the cumulative contribution of all indicators within each construct [23]. Previous studies have shown that scaling methods, including SRS, improve measurement accuracy and influence the reliability, construct validity, and parameter estimation in subsequent multivariate analyses compared with the original ordinal scores [9]. The transformed interval scores were then used as the observed variables in the subsequent statistical analysis. The indicators of the analyzed variables are presented in **Table 2.1**.

Table 2.1. Operational Definition

Variable	Operational Definition	Indicator
Farmer Needs (X_1)	Farmer Needs refer to the extent to which agricultural products and services fulfill farmers' expectations and farming requirements. This construct was identified through the Systematic Literature Review (SLR) and bibliometric analysis as an important antecedent of farmer loyalty.	Tangibles ($X_{1.1}$) Perceived ($X_{1.2}$) Performance ($X_{1.3}$) Convenience of Use ($X_{1.4}$)
Farmer Success (Y_1)	Farmer Success represents farmers' perceptions of the achievement of desirable farming outcomes resulting from the use of agricultural products and services.	Farm Business Efficiency ($Y_{1.1}$) Selling Price ($Y_{1.2}$) Income ($Y_{1.3}$) Productivity ($Y_{1.4}$)
Farmer Loyalty (Y_2)	Farmer Loyalty refers to farmers' commitment to continue using, repurchasing, and recommending agricultural products or services over time.	Repurchase ($Y_{2.1}$) Retention ($Y_{2.2}$) Recommend Product ($Y_{2.3}$)

2.3 Bibliometrix Analysis

Bibliometric analysis is a widely recognized and rigorous quantitative approach for exploring and analyzing large volumes of scientific data [5]. It provides a systematic method for evaluating an article's contribution to the advancement of knowledge and is frequently employed to assess research trends and performance within a specific field. In addition, bibliometric studies offer valuable insights into both the scientific impact and the social relevance of a particular discipline or research domain. This approach is also useful for identifying patterns in scientific publications, including authorship, collaboration networks, and thematic evolution [6].

2.4 Summated Rating Scale (SRS)

The Summated Rating Scale (SRS) is a scaling method commonly used to transform responses collected through Likert-scale instruments into composite scores with interval-scale properties. This transformation is appropriate when ordinal responses are analyzed using statistical methods that assume interval-scale measurements. In this study, SRS was employed because the variables SLR Result (X_1), Farmer Success (Y_1), and Farmer Loyalty (Y_2) were measured using Likert-scale indicators. Transforming these ordinal responses into interval-scale scores provides more representative measurements while preserving the cumulative contribution of all indicators within each construct [22]. Furthermore, previous studies have demonstrated that SRS improves measurement accuracy and produces composite scores suitable for subsequent statistical analyses [9].

The SRS transformation was carried out in four steps. First, the cumulative proportion for each response category was calculated based on the observed frequency distribution [22]. Second, the midpoint of each cumulative proportion was determined. Third, each midpoint was transformed into its corresponding standard normal (Z) value using the inverse cumulative distribution function. Finally, a constant was added to all Z-values to eliminate negative values and generate positive interval-scale scores. The transformed interval scores obtained from this procedure were subsequently used in the statistical analysis.

2.5 Path Analysis

Path analysis is an extension of multiple regression analysis developed to overcome its limitations in handling complex data structures, particularly those involving multiple dependent (response) variables. The concept was first introduced by Sewall Wright in 1934 [11]. as a methodological framework to analyze both direct and indirect relationships among variables. In this approach, variables are explicitly classified based on their roles, allowing researchers to distinguish between causal (explanatory) and outcome (response) variables. In addition to estimating the direct effects of exogenous variables on endogenous variables, path analysis also enables the assessment of indirect effects mediated through intervening variables [13]. Therefore, this method provides a more comprehensive understanding of causal mechanisms compared to conventional regression approaches and is often used to examine the causal relationships between extrinsic and intrinsic factors [15].

Structurally, path analysis consists of a system of multiple equations represented in a causal model framework. Each structural equation typically includes at least one exogenous variable, one intervening (mediating) variable, and one or more endogenous variables. Exogenous variables are those that influence other variables within the model but are not influenced by any variables in the system. Intervening variables play a dual role, as they both affect and are affected by other variables, serving as key components that distinguish path analysis from traditional regression models, which do not explicitly incorporate such mediating structures [22]. Meanwhile, endogenous variables are variables whose values are determined entirely by other variables within the model. The relationships specified in path analysis are assumed to be linear and additive [21]. Therefore, when these assumptions are satisfied, parametric path analysis can be appropriately applied [12]. Based on Figure 1, the model is written in Equation (2.1) and Equation (2.2).

$$Y_{1i} = \beta_{01} + \beta_{X_1Y_1}X_{1i} + \varepsilon_{Y_{1i}} \qquad (2.1)$$

$$Y_{2i} = \beta_{02} + \beta_{X_1Y_1}X_{1i} + \beta_{Y_1Y_2}Y_{1i} + \varepsilon_{Y_{2i}} \qquad (2.2)$$

Equation (2.1) and equation (2.2) that have been standardized can be seen in equation (2.3) and equation (2.4).

$$Z_{Y_{1i}} = \beta_{X_1Y_1}Z_{X_{1i}} + \varepsilon_{Y_{1i}} \qquad (2.3)$$

$$Z_{Y_{2i}} = \beta_{X_1Y_1}Z_{X_{1i}} + \beta_{Y_1Y_2}Z_{Y_{1i}} + \varepsilon_{Y_{2i}} \qquad (2.4)$$

where:

Y_{ji} : jth endogenous variable of i-th observation
X_{pi} : pth exogenous variable of i-th observation
i : number of observations; 1, 2, …, n
j : number of exogenous variables; 1, 2, …, p
l : of endogenous variables; 1, 2, …, q
β_1, β_2 : path coefficient
$\varepsilon_{y_{ji}}$: the residuals on the jth endogenous variable of the i-th observation

Equations (2.3) and (2.4) are written in matrix form as follows.

$$\begin{bmatrix} Z_{y11} \\ Z_{y12} \\ \vdots \\ Z_{y1n} \\ Z_{y21} \\ Z_{y22} \\ \vdots \\ Z_{y2n} \end{bmatrix} = \begin{bmatrix} Z_{x11} & 0 & 0 \\ Z_{x12} & 0 & 0 \\ \vdots & \vdots & \vdots \\ Z_{x1n} & 0 & 0 \\ 0 & Z_{x11} & Z_{y11} \\ \vdots & \vdots & \vdots \\ 0 & Z_{x1n} & Z_{y1n} \end{bmatrix} \begin{bmatrix} \beta_{x1y1} \\ \beta_{y1y2} \\ \beta_{x1y2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{y11} \\ \varepsilon_{y12} \\ \vdots \\ \varepsilon_{y1n} \\ \varepsilon_{y21} \\ \varepsilon_{y22} \\ \vdots \\ \varepsilon_{y2n} \end{bmatrix}$$

$$\begin{bmatrix} Z_{y11} \\ Z_{y12} \\ \vdots \\ Z_{y1n} \\ Z_{y21} \\ Z_{y22} \\ \vdots \\ Z_{y2n} \end{bmatrix} = \begin{bmatrix} Z_X & \underset{\sim}{0}_{nx2} \\ \underset{\sim}{0}_{nx1} & Z_{XY} \end{bmatrix} \begin{bmatrix} \beta_{x1y1} \\ \beta_{y1y2} \\ \beta_{x1y2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{y11} \\ \varepsilon_{y12} \\ \vdots \\ \varepsilon_{y1n} \\ \varepsilon_{y21} \\ \varepsilon_{y22} \\ \vdots \\ \varepsilon_{y2n} \end{bmatrix}$$

2.6 Model Validity

The coefficient of determination is a fundamental measure used to assess the extent to which variability in the data can be explained by a specified model [5]. Model validity can be evaluated through several approaches, one of which is the coefficient of total determination. This coefficient is employed to quantify the overall proportion of variance in the observed data that is accounted for by the structural model [18]. Thus, the total explanatory power of the model can be represented by the coefficient of total determination, which is calculated using the following formulation.

$$R_m^2 = 1 - P_{e1}^2 P_{e2}^2 \dots P_{ep}^2 \qquad (2.5)$$

Calculation of the effect of the residual (error) can be done using the following formula.

$$P_{ei} = \sqrt{1 - R_i^2} \qquad (2.6)$$

where:

P_{ei} : the effect of error on each equation

R_i^2 : coefficient of determination in each equation

R_m^2 : total determination coefficient

i : $1, 2, 3, \dots, p$

The coefficient of total determination ranges from 0% to 100%, indicating the proportion of variance explained by the model. A value approaching 100% suggests that the model has strong explanatory power and is considered to have a good fit to the data.

3. RESULTS AND DISCUSSION

3.1 Bibliometrix Analysis Using VosViewer

This section presents the research findings obtained through bibliometric analysis and statistical modeling to support the development of a conceptual framework for modeling farmer loyalty. The analysis was conducted sequentially, beginning with a Systematic Literature Review (SLR), followed by bibliometric mapping using VOSviewer, and concluding with the identification of research constructs for path analysis.

The literature selection process is illustrated in Figure 1. An initial search of the Scopus database identified 44,981 publications related to farmer loyalty and its associated determinants. After applying the predefined screening criteria, including publication year, subject area, document type, language, and relevance to the research objectives, a total of 434 publications were retained for bibliometric analysis.

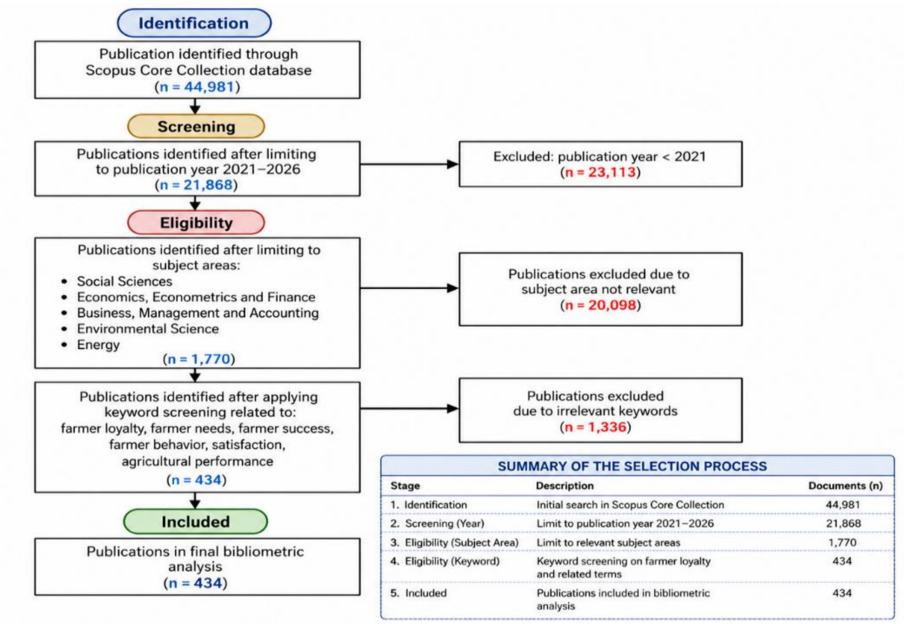

Figure 3.1. Literature screening process

Bibliographic information, including article titles, abstracts, author keywords, publication years, citation data, authors, affiliations, and source titles, was extracted from the selected publications and imported into VOSviewer. A keyword co-occurrence analysis was then performed to explore the conceptual structure of the literature and identify relationships among research topics associated with farmer loyalty.

The bibliometric network generated by VOSviewer is presented in Figure 3.2. The visualization groups related keywords into thematic clusters based on their co-occurrence relationships, illustrating the conceptual structure of previous studies on farmer loyalty.

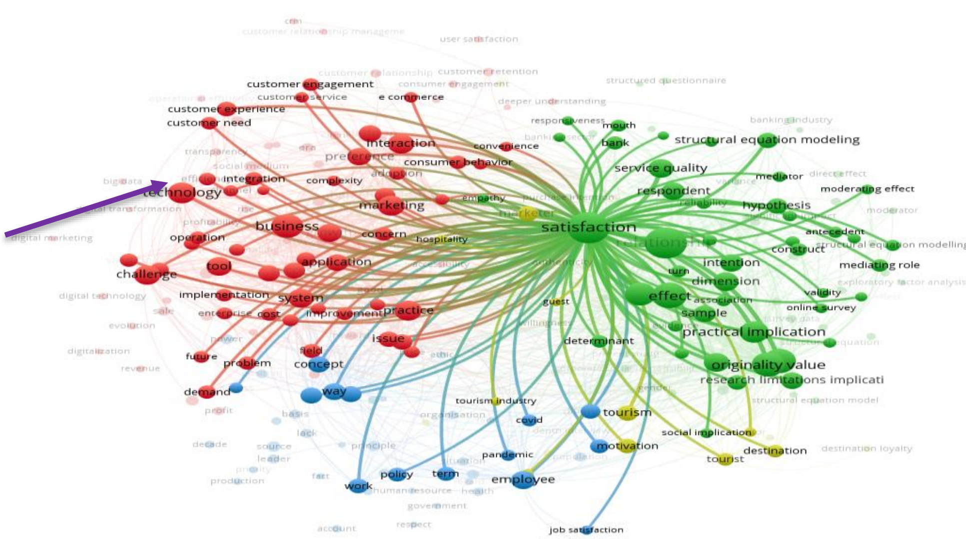

Figure 3.2. Visualization of SLR Results

The bibliometric analysis was employed as an exploratory approach to identify concepts that frequently appeared together in the scientific literature. Rather than directly selecting variables from the visualization, the identified thematic clusters were further examined through a Systematic Literature Review to evaluate the conceptual relevance, theoretical definitions, and empirical evidence supporting each concept. Consequently, the transformation from bibliometric findings to research constructs was based on the integration of bibliometric mapping and systematic literature synthesis, thereby reducing subjectivity in variable selection and providing a stronger theoretical foundation for the proposed model.

The keyword co-occurrence analysis generated several thematic clusters representing different research domains related to farmer behavior and agricultural development. Each cluster reflects groups of concepts that frequently co-occur in the literature and illustrates the conceptual relationships among research topics.

Although several publications use the term *customer need* to describe users or purchasers of agricultural products and services, this study adopts the term *Farmer Needs* because the unit of analysis specifically focuses on farmers. In this context, Farmer Needs refers to farmers' needs for agricultural inputs, agricultural services, and institutional support that influence farming decisions and long-term behavioral commitment. Therefore, the terminology was adapted to reflect the characteristics of the research population while maintaining the same underlying theoretical concept identified in the reviewed literature.

The thematic clusters identified in Figure 3.2 were subsequently examined through the SLR to determine concepts that consistently appeared in the literature and were theoretically relevant to the proposed research model. Based on this synthesis, Farmer Needs was identified as the most appropriate exogenous construct because previous studies consistently recognize it as an antecedent influencing farmers' behavioral responses. Furthermore, the reviewed literature indicates that the fulfillment of farmers' needs contributes to Farmer Success, which subsequently strengthens Farmer Loyalty. Accordingly, Farmer Needs was selected as the exogenous variable, Farmer Success as the mediating variable, and Farmer Loyalty as the endogenous variable in the proposed path analysis model.

This procedure demonstrates that bibliometric analysis was not used as the sole basis for variable selection. Instead, it functioned as an objective knowledge-mapping tool to identify potential concepts, while the final research constructs were established through a comprehensive synthesis of the selected literature. This integrated approach ensures that the proposed path analysis model is supported by both bibliometric evidence and theoretical justification.

3.2 Path Analysis

3.2.1 Parameter Estimation

In this study, the variables included in the structural model were treated as observed composite scores rather than latent variables. The original Likert-scale responses were first transformed into interval-scale composite scores using the Summated Rating Scale (SRS). Consequently, the structural relationships among variables were analyzed using path analysis estimated by the Ordinary Least Squares (OLS) method. This approach is appropriate because path analysis operates on observed variables and aims to estimate direct and indirect effects among constructs represented by composite scores [22]. The estimation results are presented in **Table 3.1**.

Table 3.1. Path Analysis Coefficient

Parameter	Path Coefficient
$\beta_{x_1 y_1}$	0,766

Parameter	Path Coefficient
$\beta_{x_1 y_2}$	0,141
$\beta_{y_1 y_2}$	0,459

Table 2 indicates the presence of direct effects among the variables. Based on the estimated path coefficients, the structural equations of the path analysis model can be formulated as follows.

$$Z_{Y_{1i}} = 0{,}766 Z_{X_{1i}} + \varepsilon_{Y_{1i}}$$
$$Z_{Y_{2i}} = 0{,}141 Z_{X_{1i}} + 0{,}459 Z_{Y_{1i}} + \varepsilon_{Y_{2i}}$$

3.2.2 Assumption Testing

The following are the results of assumption testing on path analysis which consists of linearity and normality assumption tests.

1. he linearity assumption is evaluated using the Ramsey RESET test and the results can be seen in **Table 3.2**.

Table 3.2. Linearity Assumption Test Results

Variable	p-value	Relationship
$X_1 \rightarrow Y_1$	0,114	Linier
$X_1 \rightarrow Y_2$	0,840	Linier
$Y_1 \rightarrow Y_2$	0,421	Linier

Based on Table 3, the relationship between the exogenous and endogenous variables yields a p-value greater than α (0.05), leading to the acceptance of H_0. This result indicates that the relationship among the variables can be considered linear.

The assumption of additivity is evaluated by examining the specified model structure, as represented in the following equation.

$$Z_{Y_{1i}} = \beta_{X_1 Y_1} Z_{X_{1i}} + \varepsilon_{Y_{1i}}$$
$$Z_{Y_{2i}} = \beta_{X_1 Y_2} Z_{X_{1i}} + \beta_{Y_1 Y_2} Z_{Y_{1i}} + \varepsilon_{Y_{2i}}$$

Based on the specified model, it is clear that an incremental form is used, and there are no interaction effects among the variables; therefore, the model satisfies the additive assumption.

2. Assumption of Normality of Residuals

The normality assumption of the residuals is assessed using the Jarque–Bera test, and the corresponding results are presented in **Table 3.3**.

Table 3.3. Results of Sisaan Normality Assumption Test

Endogenous Variable	p-value	Conclusion
Farmer Success (Y_1)	< 0,001	Not Normal
Farmer Loyalty (Y_2)	< 0,001	Not Normal

Based on **Table 3.3**, it can be concluded that all endogenous variables violate the normality assumption, indicating that the residuals are not normally distributed. This violation may affect the validity of statistical inference when using conventional parametric methods. Therefore, a non-parametric bootstrap procedure was employed to estimate the sampling distribution of the path coefficients. The bootstrap analysis was performed in R by repeatedly drawing 500 bootstrap samples with replacement from the original dataset. For each bootstrap sample, the path coefficients were re-estimated to obtain bootstrap standard errors, confidence intervals, and p-values. Statistical significance was evaluated at the 5% significance level ($p < 0.05$). Since this resampling procedure does not rely on the normality assumption, it provides more reliable statistical inference and improves the robustness of the estimated path coefficients.

3.2.3 Direct Effect Test

1. Direct Effect Test

Hypothesis testing is done with a bootstrap approach to determine the significance of the model formed. The results of hypothesis testing can be seen in **Table 3.4**.

Table 3.4. Hypothesis Testing of Direct Influence

Variable	OLS Path Coefficient	Bootsrap Path Coefficient	*Standard Error*	*p-value*
$X_1 \rightarrow Y_1$	0,766	0,763	0,002	<0,001*
$X_1 \rightarrow Y_2$	0,141	0,141	0,01	<0,001*
$Y_1 \rightarrow Y_2$	0,459	0,462	0,008	<0,001*

Based on **Table 3.4,** the bootstrap estimates are highly consistent with the corresponding OLS estimates, as indicated by the minimal differences between the path coefficients. The bootstrap standard errors are relatively small, suggesting low sampling variability and stable parameter estimates across the 500 resamples. Furthermore, all path coefficients remain statistically significant at the 5% significance level ($p < 0.001$), indicating that the observed relationships are robust to repeated resampling. The close agreement between the OLS and bootstrap estimates indicates that the estimated path coefficients are stable and are not substantially influenced by sampling variation. Therefore, the bootstrap procedure provides reliable statistical inference despite the violation of the normality assumption identified in the residual analysis.

2. Indirect Influence Test

The indirect effect test is a hypothesis test that aims to determine whether a mediating variable has a significant effect on the model. The results of testing indirect effects using the Sobel Test can be seen in **Table 3.5**.

Table 3.5. Hypothesis Testing of Indirect Influence Variable P-value Conclusion

Variable	Indirect Effect	P-value	Conclusion
$X_1 \rightarrow Y_2$	-	<0,001*	Partially Mediated
$X_1 \rightarrow Y_1 \rightarrow Y_2$	0,352	0,004*	

Based on the results presented in **Table 3.5**, the indirect effect testing using the Sobel Test indicates that the relationships among variables are statistically significant. Specifically, the indirect effect of X_1 on Y_2 shows a p-value of < 0.001, indicating a significant effect and suggesting that the relationship is partially mediated. Furthermore, the indirect path from X_1 through Y_1 to Y_2also demonstrates a significant effect, with a p-value of 0.004. The indirect effect (0.352) is considerably larger than the direct effect (0.141), indicating that most of the influence of X_1 on Y_2 is transmitted through Y_1. Although the direct effect remains statistically significant, confirming partial mediation, the results suggest that Y_1 serves as the primary mechanism linking X_1 to Y_2. These findings confirm that the mediating variable (Y_1) plays a significant role in transmitting the effect of X_1 on Y_2 within the proposed model.

3.2.4 Research Result Model

Based on the path significance testing using the bootstrap approach, all variables exhibit statistically significant effects, so that the research model formed is as in **Figure 3**.

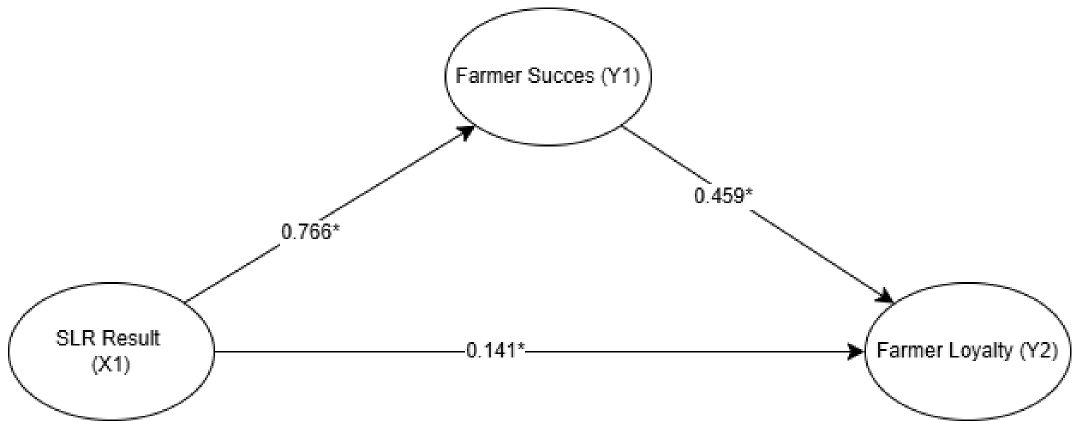

Figure 3.2. Path Diagram of Research Results

3.2.5 Model Validity

The results of the calculation of the coefficient of determination for each equation are presented in **Table 3.6** below.

Table 3.6. Coefficient of Determination (R_i^2)

Endogenous Variable	Coefficient of Determination (R_i^2)	Coefficient of Total Determination (R_m^2)
Farmer Success (Y_1)	0,7317	0,9340
Farmer Loyalty (Y_2)	0,7541	

Based on the coefficient of determination for each structural equation presented in **Table 7**, the coefficient of total determination (R^2) of 0.9340, indicating that the proposed model is capable of explaining 93.40% of the total variance in the observed data. This result demonstrates that the path analysis model possesses a high level of explanatory power. The remaining 6.60% of the variance is attributable to factors not incorporated within the model, which may include omitted variables, measurement errors, or other unobserved influences.

CONCLUSION

The conclusions drawn based on the results of the analysis that has been carried out from this research are as follows.

1. The results of the bibliometric analysis using VOSviewer on 434 Scopus-indexed documents identified **farmer needs** as a key determinant variable with strong relevance and connectivity in the research network, forming the basis for the proposed model.
2. The path analysis results show that **farmer needs** and **farmer success** have a significant direct effect on **farmer loyalty**, and **farmer success partially mediates** the relationship between farmer needs and farmer loyalty, indicating both direct and indirect effects.

3. The proposed model demonstrates a very high level of explanatory power, as indicated by the coefficient of total determination (R^2) of 0.9732, meaning that the model is highly effective in explaining variations in farmer loyalty.

However, this study has several limitations. The proposed model includes only farmer needs and farmer success as the primary determinants of farmer loyalty. Other factors, such as institutional support, market access, government policies, farming experience, socioeconomic characteristics, and environmental conditions, were not incorporated and may also influence farmer loyalty. In addition, the analysis was conducted using cross-sectional survey data, which limits the ability to capture changes in farmer behavior over time. Future research is recommended to incorporate additional explanatory variables and employ longitudinal data or alternative analytical approaches to obtain a more comprehensive understanding of the factors influencing farmer loyalty across different agricultural contexts.

ACKNOWLEDGEMENT (optional)

This research received no specific grants from funding agencies in the public, commercial or non-profit sectors. Non-profit. The authors would like to thank the agency that provided the data for this study, all those who assisted in this study, and the anonymous reviewers for this study.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest to report in connection with this study.

REFERENCES

[1] Alamsyah, R. T. P., Wulandari, E., Saidah, Z. & Hapsari, H., 2024. Discovering sustainable finance models for smallholder farmers: A bibliometric approach to agricultural innovation adoption. *Discover Sustainability*, Vol. 5, Article No. 107. doi:10.1007/s43621-024-00277-4.

[2] Asriadi, A. A., Salam, M., Nadja, R. A., Fudjaja, L., Rukmana, D., Jamil, M. H., Arsyad, M., Rahmadanih & Maulidiyah, R., 2024. Determinants of Farmer Participation and Development of Shallot Farming in Search of Effective Farm Management Practices: Evidence Grounded in Structural Equation Modeling Results. *Sustainability*, Vol. 16, No. 15, Article No. 6332. doi:10.3390/su16156332.

[3] Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. MIS Quarterly, 25(3), 351-370. https://doi.org/10.2307/3250921

[4] Blau, P. M. (1964). Exchange and Power in Social Life. New York: John Wiley & Sons.

[5] Chicco, D., Warrens, M. J. & Jurman, G., 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, Vol. 7, Article No. e623. doi:10.7717/peerj-cs.623.

[6] Dash, G. & Paul, J., 2021. CB-SEM vs PLS-SEM methods for research in social sciences. *Technological Forecasting and Social Change*, Vol. 173, Article No. 121092. doi:10.1016/j.techfore.2021.121092.

[7] Donthu, N., Kumar, S., Mukherjee, D., Pandey, N. & Lim, W. M., 2021. How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, Vol. 133, 285–296. doi:10.1016/j.jbusres.2021.04.070.

[8] Ezama-Nicolás, R., 2018. A bibliometric method for assessing technological maturity: The case of additive manufacturing. *Scientometrics*, Vol. 117, No. 3, 1425–1452. doi:10.1007/s11192-018-2916-6.

[9] Febriana, B. W., & Setiawati, F. A. (2024). *Increasing measurement accuracy: Scaling effect on academic resilience instrument using Method of Successive Interval (MSI) and Method of Summated Rating Scale (MSRS).* Jurnal Penelitian dan Evaluasi Pendidikan, 28(1), 32-42. https://doi.org/10.21831/jpep.v28i1.69334.

[10] Fernandes, A. A. R. & Solimun, 2017. Moderating effects orientation and innovation strategy on the effect of uncertainty on the performance of business environment. *International Journal of Law and Management*, Vol. 59, No. 6, 1211–1219. doi:10.1108/IJLMA-09-2016-0085.

[11] Fernandes, A. A. R., Solimun & Asmara, E., 2017. Estimation of spline function in nonparametric path analysis based on penalized weighted least square (PWLS). *AIP Conference Proceedings*, Vol. 1913, No. 1, Article No. 020030. doi:10.1063/1.5016660.

[12] Fernandes, A. A. R., Solimun & Hadi, B., 2020. Comparison of use of linkage in integrated cluster with discriminant analysis approach. *International Journal of Advanced Science and Technology*, Vol. 29, No. 7, 5654–5668.

[13] Henseler, J., 2021. *Composite-Based Structural Equation Modeling: Analyzing Latent and Emergent Variables*. New York: Guilford Press.

[14] Hien, L. T. D. & Kim, J., 2024. An Analysis of Relationship Quality and Loyalty Between Farmers and Agribusiness Companies in the Rice Industry: Using Multi-Group Analysis. *Agriculture*, Vol. 14, No. 12, Article No. 2197. doi:10.3390/agriculture14122197.

[15] Iriany, A. & Fernandes, A. A. R., 2023. Hybrid Fourier series and smoothing spline path non-parametrics estimation model. *Frontiers in Applied Mathematics and Statistics*, Vol. 8, Article No. 1045098. doi:10.3389/fams.2022.1045098.

[16] Kotler, P. & Keller, K. L., 2011. *Marketing Management*. 14th Edition. Upper Saddle River, NJ: Prentice Hall.

[17] Oliver, R. L. (1999). Whence consumer loyalty? Journal of Marketing, 63(Special Issue), 33-44. https://doi.org/10.1177/00222429990634s105

[18] Rodríguez Sánchez, A., Salmerón Gómez, R. & García, C., 2022. The coefficient of determination in the ridge regression. *Communications in Statistics – Simulation and Computation*, Vol. 51, No. 1, 201–219. doi:10.1080/03610918.2019.1589510.

[19] Rutsaert, P., Donovan, J., Murphy, M. & Hoffmann, V., 2024. Farmer decision making for hybrid maize seed purchases: Effects of brand loyalty, price discounts and product information. *Agricultural Systems*, Vol. 218, Article No. 104002. doi:10.1016/j.agsy.2024.104002.

[20] Sarstedt, M., Ringle, C. M. & Hair, J. F., 2021. Partial Least Squares Structural Equation Modeling. In: Homburg, C., Klarmann, M. & Vomberg, A., eds. *Handbook of Market Research*. Cham: Springer International Publishing, 587–632. doi:10.1007/978-3-319-57413- 4_15.

[21] Solimun, 2002. *Multivariate Analysis–Structural Equation Modelling (SEM) Lisrel dan Amos*. Malang: UM Press.

[22] Solimun, Fernandes, A. A. R. & Nurjannah, 2017. *Metode Statistika Multivariat Pemodelan Persamaan Struktural (SEM) Pendekatan WarpPLS*. Malang: UB Press.

[23] Solimun, Fernandes, A. A. R., & Nurjannah. (2022). Metode Statistika Multivariat: Generalized Structured Component Analysis (GSCA). Malang: UB Press.

[24] Suhandoko, D. & Pakarti, P., 2023. Analisis Faktor-Faktor yang Mempengaruhi Keberhasilan Peserta Program Perkebunan Inti Rakyat Pola PIR-Trans Desa Sungai Melayu Jaya Kalimantan Barat Tahun 1995–2021. *Jurnal Ekonomi dan Bisnis*, Vol. 2, No. 1, 62–76. doi:10.33633/jekobs.v2i1.7948.

[25] Xu, J., Li, Y., Zhang, M. & Zhang, S., 2024. Sustainable agriculture in the digital era: Past, present, and future trends by bibliometric analysis. *Heliyon*, Vol. 10, No. 14, Article No. e34612. doi:10.1016/j.heliyon.2024.e34612.

[26] Yi, Y., & La, S. (2004). What influences the relationship between customer satisfaction and repurchase intention? Investigating the effects of adjusted expectations and customer loyalty. Psychology & Marketing, 21(5), 351-373. https://doi.org/10.1002/mar.20009

[27] Yu, Y., Li, Y., Zhang, Z., Gu, Z., Zhong, H., Zha, Q., Yang, L., Zhu, C. & Chen, E., 2020. A bibliometric analysis using VOSviewer of publications on COVID-19. *Annals of Translational Medicine*, Vol. 8, No. 13, Article No. 816. doi:10.21037/atm-20-4235.