Green Intelligent Systems and Applications

Research Article
Volume 6(2), 2026 207−216
https://doi.org/10.53623/gisa.v6i2.1294

Performance Analysis of the Multivariate Multiple Linear Regression Algorithm for Economic Growth Based on Oil Palm Plantation Land Expansion in North Sumatra Province

Afridayani, Erna Budhiarti Nababan*, Baihaqi Siregar

Master of Data Science and Artificial Intelligence, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Indonesia

*Correspondence: ernabrn@usu.ac.id

SUBMITTED: 8 July 2026; REVISED: 23 July 2026; ACCEPTED: 29 July 2026

ABSTRACT: Oil palm plantations are one of the main sectors contributing to regional economic development in North Sumatra Province. This study aimed to analyze the performance of the Multivariate Multiple Linear Regression (MMLR) algorithm in modeling the relationship between oil palm plantation expansion and regional economic indicators. Secondary data from five oil palm-producing regencies covering the period 2013–2023 were obtained from the Central Statistics Agency (BPS). The independent variables consisted of plantation area and oil palm production, whereas the dependent variables were Gross Regional Domestic Product (GRDP) and per capita income. The dataset was divided into training (2013–2019) and testing (2020–2023) subsets. Model performance was evaluated using the coefficient of determination (R^2), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). The results showed that the MMLR model was able to capture the relationship between oil palm plantation expansion and regional economic indicators, although its predictive performance varied across regencies. These findings indicated that the model provided useful insights into the contribution of the oil palm plantation sector to regional economic development in North Sumatra.

KEYWORDS: Multivariate multiple linear regression; oil palm plantation; GDRP; per capita income; North Sumatra.

1. Introduction

The plantation sector played an important role in Indonesia's economic development by contributing to Gross Domestic Product (GDP), employment, and community welfare [1, 2]. Among plantation commodities, oil palm was one of the largest contributors to the national economy. As the world's leading palm oil producer, Indonesia depended on the sustainable development of this sector to support regional economic growth. North Sumatra Province was one of the country's major oil palm-producing regions, with plantation areas exceeding 1.3 million hectares in 2023 [3]. Consequently, the expansion of oil palm plantations became an

important driver of regional economic activities, particularly by increasing Gross Regional Domestic Product (GRDP) and improving community income [4–6].

The development of the oil palm sector was commonly represented by plantation area and palm oil production. Plantation area reflected the development capacity of the sector, whereas production indicated the actual output generated from land utilization. As these indicators described different aspects of plantation performance, considering both variables provided a more comprehensive representation of oil palm development. In regional economic analysis, GRDP and per capita income were widely used to measure economic performance and community welfare [7]. Therefore, understanding the simultaneous relationship between plantation development and these economic indicators was essential [8].

Previous studies reported that oil palm plantations positively contributed to regional economic development. However, most studies examined either plantation area or production separately and focused on a single economic indicator [9]. Moreover, these studies primarily emphasized statistical relationships rather than predictive modeling, whereas the predictive performance of regression algorithms was rarely evaluated using objective performance metrics [10]. Consequently, the capability of regression models to predict regional economic performance remained insufficiently explored [11].

To address these gaps, this study employed the Multivariate Multiple Linear Regression (MMLR) algorithm to simultaneously analyze the effects of plantation area and palm oil production on GRDP and per capita income in North Sumatra Province [12]. Unlike conventional regression models that estimated only one dependent variable, MMLR enabled the simultaneous analysis of multiple correlated economic indicators within a single model [13]. The novelty of this study lay in integrating two indicators of oil palm sector development, namely plantation area and production, with two regional economic indicators, namely GRDP and per capita income, using the MMLR algorithm [14]. Furthermore, the proposed model was evaluated using the coefficient of determination (R^2), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) on separate training and testing datasets to assess its predictive performance [15]. The findings were expected to provide a more comprehensive understanding of the contribution of the oil palm sector to regional economic development and to support evidence-based policymaking for sustainable plantation development in North Sumatra Province [16, 17].

2. Materials and Methods

2.1. Research Data.

This study employed a quantitative approach using secondary data obtained from Statistics Indonesia (BPS) of North Sumatra Province. The dataset covered the period from 2013 to 2023 and included two major oil palm-producing regencies in North Sumatra Province, namely Asahan and Langkat. Two independent variables representing the development of the oil palm sector were included, namely plantation area (X_1) and palm oil production (X_2). Regional economic performance was represented by two dependent variables, namely Gross Regional Domestic Product (GRDP) (Y_1) and per capita income (Y_2). Table 1 presents the variables used in this study.

Table 1. Research Variables.

Variable	Type	Unit
Plantation Area (X_1)	Independent	Hectares (ha)
Palm Oil Production (X_2)	Independent	Tons
GRDP (Y_1)	Dependent	Million Rupiah (Constant Prices)
Per Capita Income (Y_2)	Dependent	Rupiah per person

The research dataset consisted of annual observations collected from the two regencies during the 2013–2023 period. All data were obtained from official publications of Statistics Indonesia (BPS) to ensure data consistency and reliability (Table 2).

Table 2. Research Dataset.

Description	Information
Study Area	North Sumatra Province
Observation Units	Two regencies
Study Period	2013–2023
Independent Variables	Plantation Area, Palm Oil Production
Dependent Variables	GRDP, Per Capita Income
Data Source	Statistics Indonesia (BPS)

2.2. Research Procedure.

The research procedure consisted of four stages: data preprocessing, model development, model evaluation, and result interpretation. First, the collected data were cleaned to remove inconsistencies and missing values. Subsequently, Min-Max normalization was applied to reduce differences in variable scales before model development. The dataset was then divided chronologically into training data (2013–2019) and testing data (2020–2023), representing approximately 77% and 23% of the observations, respectively. A chronological split was selected to preserve the temporal sequence of the data and to evaluate the model's ability to predict future observations using historical information.

2.3. Multivariate Multiple Linear Regression.

This study employed the Multivariate Multiple Linear Regression (MMLR) algorithm to simultaneously analyze the effects of plantation area and palm oil production on GRDP and per capita income. Unlike conventional multiple linear regression, which estimates only a single dependent variable, MMLR simultaneously estimates multiple correlated dependent variables within a single regression framework. The MMLR model was expressed as follows:

$$Y_1 = a_1 + b_{11}X_1 + b_{12}X_2 + \cdots + b_{1n}X_n$$
$$Y_2 = a_2 + b_{21}X_1 + b_{22}X_2 + \cdots + b_{2n}X_n$$
$$Y_m = a_m + b_{m1}X_1 + b_{m2}X_2 + \cdots + b_{mn}X_n$$

where: $Y_1, Y_2, \ldots, Y_m$represent the dependent (response) variables; $a_1, a_2, \ldots, a_m$denote the intercepts; $b_{11}, b_{12}, \ldots, b_{mn}$represent the regression coefficients; and $X_1, X_2, \ldots, X_n$denote the independent (predictor) variables.

In matrix form, the MMLR model can be expressed as:

$$\mathrm{Y} = \mathrm{XB} + \mathrm{E}$$

where Y represents the matrix of dependent variables, X denotes the matrix of independent variables, B is the regression coefficient matrix, and E represents the error matrix.

2.4. Regression Assumption Testing.

Prior to model estimation, the classical regression assumptions were evaluated to ensure the validity of the MMLR model. The diagnostic tests included normality, linearity, homoscedasticity, and independence tests. These tests were performed to verify that the regression model satisfied the assumptions required for reliable parameter estimation and valid statistical inference [18].

2.5. Model Performance Evaluation.

The predictive performance of the proposed model was evaluated using the coefficient of determination (R^2), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). The coefficient of determination (R^2) measured the proportion of variance explained by the model, whereas MSE and RMSE quantified prediction errors. Lower MSE and RMSE values indicated better predictive performance, whereas higher R^2 values indicated stronger agreement between the observed and predicted values. Model performance was evaluated using both the training and testing datasets to assess predictive accuracy and generalization capability [19].

3. Results and Discussion

3.1. Correlation Analysis.

To examine the relationships among the research variables, a correlation analysis was conducted. The resulting correlation matrix is presented in Figure 1. the correlation analysis revealed varying strengths and directions of association among the research variables. Plantation area exhibited a very strong positive correlation with palm oil production ($r = 0.84$), indicating that larger plantation areas are generally associated with higher palm oil output. This pattern is consistent with earlier findings that expansion of oil palm plantation area is a primary driver of increased production volume [2]. Plantation area also showed a strong positive relationship with per capita income ($r = 0.75$, Figure 1), while palm oil production showed a moderately strong positive correlation with per capita income ($r = 0.67$). These results suggest that the development of plantation land contributes more directly to household-level income than production volume alone, a relationship also reported in studies of the plantation sector's contribution to regional welfare and per-capita income growth [3, 5].

Conversely, Figure 1 shows that Gross Regional Domestic Product (GRDP) had weak negative correlations with plantation area ($r = -0.35$) and palm oil production ($r = -0.34$). This indicates that plantation area and production alone are insufficient to explain regional economic output, since GRDP reflects the combined contribution of multiple economic sectors, including manufacturing, trade, services, and construction, rather than the plantation sector in isolation [12]. Furthermore, the correlation between GRDP and per capita income shown in Figure 1 was relatively weak ($r = 0.32$), implying that growth in aggregate regional output does not necessarily translate into proportional gains in average household income, a discrepancy also noted in analyses of employment structure and income distribution in the region [4, 9]. Overall,

Figure 1 indicates that plantation area and palm oil production exhibited the strongest pairwise relationship among the research variables, whereas GRDP showed comparatively weaker associations with both plantation-related variables. This pattern suggests that, although the oil palm plantation sector plays an important role in the regional economy, it remains only one component of a broader and more diversified economic structure [8].

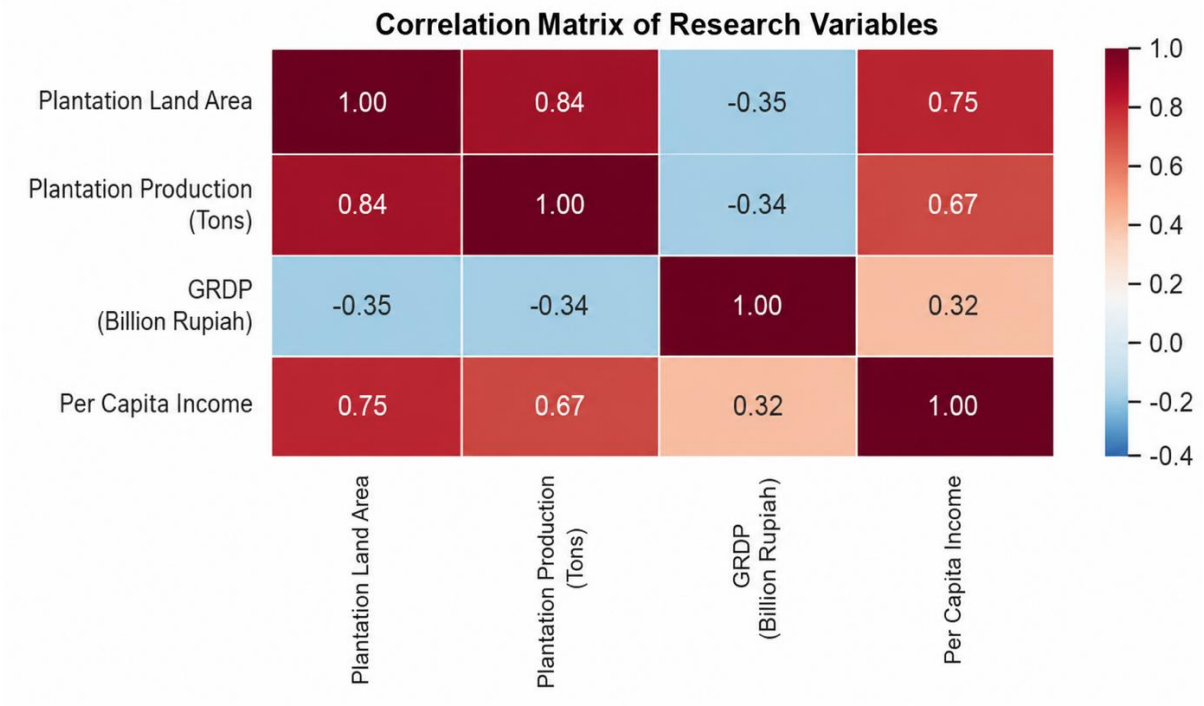

Figure 1. Correlation Matrix of the Research Variables.

3.2. Model Performance Evaluation.

The predictive performance of the Multivariate Multiple Linear Regression (MMLR) model was evaluated using the coefficient of determination (R^2) and the Root Mean Square Error (RMSE). Model performance was assessed separately using the training and testing datasets to evaluate both model fitting and generalization capability [19]. Table 3 presents the performance of the MMLR model for predicting Gross Regional Domestic Product (GRDP). The model achieved high training performance in both regencies, with R^2 values of 0.9469 for Asahan and 0.8341 for Langkat, indicating that 94.69% and 83.41% of the variation in GRDP could be explained by plantation area and palm oil production, respectively. The corresponding training RMSE values were relatively low (561.52 for Asahan and 1,026.79 for Langkat), suggesting that the model closely fitted the training observations.

Table 3. Performance Evaluation Results for the GRDP (Y_1) Model.

Regencies	R^2 Train	RMSE Train	R^2 Test	RMSE Test
Asahan	0.9469	561.52	-59.8113	10,664.77
Langkat	0.8341	1,026.79	-259.5708	23,734.11

Table 4 summarizes the performance of the model for predicting per capita income. Similar to the GRDP model, the training results indicated excellent goodness-of-fit, with R^2 values of 0.9460 in Asahan and 0.8133 in Langkat. The corresponding RMSE values of 668,196.08 and 874,149.00 further indicated relatively small prediction errors within the

training dataset. Overall, the training results demonstrated that the MMLR model effectively captured the relationships between the plantation variables and both regional economic indicators.

Table 4. Performance Evaluation Results for the Per Capita Income (Y_2) Model.

Regencies	R^2 Train	RMSE Train	R^2 Test	RMSE Test
Asahan	0.9460	668,196.08	-73.5478	10,211,963.59
Langkat	0.8133	874,149.00	-322.2263	18,577,699.42

In contrast, the testing performance deteriorated substantially for both dependent variables. As shown in Table 3 and Table 4, all testing R^2 values were negative, indicating that the predictive accuracy of the model declined considerably when applied to unseen data. This result is summarized in Table 5, where all testing models produced negative R^2 values. Specifically, the GRDP model yielded testing R^2 values of −59.8113 for Asahan and −259.5708 for Langkat, whereas the per capita income model produced values of −73.5478 and −322.2263, respectively. A negative R^2 indicates that the regression model performed worse than a baseline model that simply predicts the mean of the observed values, suggesting poor predictive capability and severe overfitting [19].

Table 5. Summary of Model Performance on the Testing Dataset.

Target	Regencies	R^2 Test	RMSE Test	Information
GRDP	Asahan	-59.8113	10,664.77	Negative testing R^2
GRDP	Langkat	-259.5708	23,734.11	Negative testing R^2
Per Capita Income	Asahan	-73.5478	10,211,963.59	Negative testing R^2
Per Capita Income	Langkat	-322.2263	18,577,699.42	Negative testing R^2

The testing RMSE values also increased markedly compared with those obtained during training. For example, as shown in Table 3, the RMSE for GRDP prediction in Langkat increased from 1,026.79 during training to 23,734.11 during testing, representing an increase of more than twenty-three times. Similarly, Table 4 shows that the RMSE for per capita income prediction in Langkat increased from 874,149.00 to 18,577,699.42, indicating a substantial decline in prediction accuracy. Comparable patterns were also observed in Asahan for both dependent variables. These findings indicate that although the proposed MMLR model successfully fitted the historical data, it failed to generalize to new observations.

The large discrepancy between the training and testing results suggests that the model experienced overfitting. This limitation was likely attributable to the relatively small sample size (11 annual observations per regency) and the inclusion of only two explanatory variables. Regional economic performance is influenced by numerous socioeconomic factors, including investment, labor productivity, industrial development, commodity prices, government expenditure, and demographic changes, which were not incorporated into the present model. Consequently, plantation area and palm oil production alone were insufficient to accurately predict variations in GRDP and per capita income. Similar limitations in the predictive performance of regression-based economic models with limited explanatory variables have been reported in previous studies [14, 16].

3.3 Regression Significance Analysis.

The statistical significance of the regression model was evaluated using both the F-test and the t-test at a significance level of α = 0.05, with results summarized in Table 6. the F-test results indicate that all four regression models were statistically significant ($p < 0.05$). For Asahan Regency, the *p*-values were 0.002822 for GRDP and 0.002916 for per capita income, while Langkat Regency produced *p*-values of 0.027518 and 0.034846, respectively. These results confirm that plantation area and palm oil production jointly exert a statistically significant influence on both regional economic indicators, in line with earlier evidence that plantation-related variables are jointly significant predictors of regional economic outcomes [3, 17].

Table 6. Significance Test Results (α = 0.05).

Regency	Target	*p-value F*	*p-value* t X_1 / X_2	Arah X1 / X2
Asahan	GRDP (Y_1)	0.002822	X_1 0.009551; X_2 0.165909	X_1 +; X_2 -
Asahan	Income per capita (Y_2)	0.002916	X_1 0.009807; X_2 0.167648	X_1 +; X_2 -
Langkat	GRDP (Y_1)	0.027518	X_1 0.019284; X_2 0.411946	X_1 +; X_2 -
Langkat	Income per capita (Y_2)	0.034846	X_1 0.025826; X_2 0.386073	X_1 +; X_2 -

The t-test results further reveal that plantation area (X_1) consistently had a statistically significant positive effect on both GRDP and per capita income, with all associated *p*-values below the 0.05 threshold (ranging from 0.009551 to 0.025826). This confirms that expansion of oil palm plantation area contributes positively to regional economic development [2, 5]. In contrast, palm oil production (X_2) was not statistically significant in any of the four models, with *p*-values ranging from 0.165909 to 0.411946, well above the 0.05 threshold. Although the estimated regression coefficients for production were negative, these coefficients cannot be interpreted as meaningful effects because they are not statistically significant. Therefore, after controlling for plantation area, palm oil production alone did not significantly influence either GRDP or per capita income, suggesting that production-related gains are more sensitive to external factors such as productivity, weather, and market prices than to production volume itself [11, 14].

3.4 Discussion

The findings summarized in Figure 1 and Table 6 indicate that plantation area is the most influential and consistent predictor of regional economic performance among the two variables examined. Larger plantation areas were consistently associated with significant improvements in both GRDP and per capita income. This result suggests that land expansion contributes to economic development by increasing production capacity, generating employment opportunities, and stimulating economic activity within the plantation sector [1, 4, 9]. n contrast, palm oil production did not exhibit a statistically significant effect once plantation area was included in the regression model (Table 6). This implies that production volume alone does not adequately explain regional economic performance, since production is influenced by numerous external factors, including productivity, weather conditions, plantation management practices, commodity prices, and harvesting efficiency [11, 14]. This finding is broadly

consistent with studies showing that export value and downstream commodity prices, rather than raw production volume, tend to drive the plantation sector's contribution to regional output [10, 15]. Although the proposed MMLR model demonstrated strong performance on the training dataset (Table 3 and Table 4), its predictive performance deteriorated substantially on the testing dataset, as evidenced by the negative R^2 values and considerably higher RMSE values reported in Table 5. These findings indicate that the model suffers from overfitting and has limited generalization capability. One likely explanation is the small number of observations, which restricts the model's ability to learn representative patterns [19]. Furthermore, regional economic indicators such as GRDP and per capita income are shaped by many additional factors, including investment, industrial activity, population growth, labor market conditions, and service-sector development, none of which were incorporated into the current model [4, 12].

A comparison between the two study areas, shown across Table 3, Table 4, and Table 6, also reveals differences in model performance and significance. Asahan achieved higher training R^2 values than Langkat for both dependent variables (Table 3 and Table 4), indicating a better in-sample fit, and also produced smaller *p*-values in the significance tests (Table 6), indicating a stronger and more robust relationship between plantation variables and economic indicators in that regency. Nevertheless, both regencies produced negative testing R^2 values (Table 5), demonstrating that the model was unable to accurately predict unseen data in either region. These results suggest that regional economic characteristics differ substantially between regencies and should be taken into account when developing predictive models for the plantation sector [7, 13]. Overall, this study demonstrates that while plantation area significantly contributes to regional economic performance, the current MMLR model is not yet sufficiently robust for predictive applications. Future studies should incorporate additional explanatory variables, such as investment, commodity prices, industrial output, employment, and population, together with larger datasets and longer observation periods, to improve model generalization and predictive accuracy [6, 18].

4. Conclusions

This study analyzed the influence of oil palm plantation area and palm oil production on Gross Regional Domestic Product (GRDP) and per capita income in Asahan and Langkat Regencies using the Multivariate Multiple Linear Regression (MMLR) algorithm. The results indicate that plantation area consistently has a positive and statistically significant effect on both GRDP and per capita income, whereas palm oil production does not show a significant partial effect despite being included in the model. At the same time, both independent variables significantly influence the two economic indicators, as confirmed by the F-test. The model achieved high predictive performance on the training dataset, reflected in high R^2 values; however, the negative R^2 values and relatively large RMSE values obtained on the testing dataset indicate poor generalization and limited predictive capability for unseen data. The correlation analysis further revealed a strong positive relationship between plantation area and per capita income, whereas the relationship between the plantation variables and GRDP was relatively weak. Overall, the findings suggest that plantation area is the most consistent factor explaining regional economic performance, while changes in GRDP and per capita income are also influenced by other socioeconomic factors beyond those included in this study. Future research should therefore incorporate additional explanatory variables, such as investment, commodity

prices, population, employment, and industrial activity, as well as larger datasets and longer observation periods, to improve model robustness and predictive accuracy.

Author Contributions

All authors contributed equally to the conception and design of the study, data collection, data analysis and interpretation, methodology development, manuscript drafting, critical revision of the manuscript, and approval of the final version. All authors have read and agreed to the published version of the manuscript.

Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The data that support the findings of this study are publicly available from Statistics Indonesia (Badan Pusat Statistik, BPS) North Sumatra Province through its official publications and statistical databases. The processed data used in this study are available from the corresponding author upon reasonable request.

References

[1] Nainggolan, H.L.; Gulo, C.K.; Waruwu, W.S.S.; Egentina, T.; Manalu, T.P. (2021). Strategi pengelolaan usahatani kelapa sawit rakyat masa pandemi Covid-19 di Kecamatan STM Hilir Kabupaten Deli Serdang, Provinsi Sumatera Utara, Indonesia. *Agro Bali: Agricultural Journal*, *4*, 260–275. https://doi.org/10.37637/ab.v4i2.724.

[2] Xin, Y.; Sun, L.; Hansen, M.C. (2021). Biophysical and socioeconomic drivers of oil palm expansion in Indonesia. *Environmental Research Letters*, *16*, 034048. https://doi.org/10.1088/1748-9326/abce83.

[3] Suardi, T.F.; Sulistyowati, L.; Noor, T.I.; Setiawan, I. (2022). Analysis of the Sustainability Level of Smallholder Oil Palm Agribusiness in Labuhanbatu Regency, North Sumatra. *Agriculture, 12,* 1469. https://doi.org/10.3390/agriculture12091469.

[4] Ilmi, A.F.; Laoli, S.; Suhendi.; Nasution, L.N. (2025). Employment development strategies to support economic growth in North Sumatra. *International Journal of Economics and Management Sciences*, *2*, 110–126. https://doi.org/10.61132/ijems.v2i3.889.

[5] Feni, R.; Marwan, E. (2023). Perkembangan luas areal dan produksi sawit serta pengaruhnya terhadap PDRB sub sektor perkebunan Provinsi Bengkulu. *Jurnal Sosial Ekonomi Pertanian*, *19*, 139–152. https://doi.org/10.20956/jsep.v19i2.21718.

[6] Rusu, M.; Filip, M.; Cara, I.G.; Ţopa, D.; Jităreanu, G. (2025). Soil Nutrient Dynamics and Farming Sustainability Under Different Plum Orchard Management Practices in the Pedoclimatical Conditions of Moldavian Plateau. *Agriculture*, *15*, 509. https://doi.org/10.3390/agriculture15050509.

[7] Lumban Tobing, F.A.B.; Marliyah. (2023). Kajian kontribusi sektor pertanian terhadap Produk Domestik Regional Bruto di Kabupaten Labuhan Batu Tahun 2018–2022. *Journal Economic Excellence Ibnu Sina*, *2*, 234–245. https://doi.org/10.59841/excellence.v2i1.1015.

[8] Hasibuan, M.; Rahmanta, R.; Ayu, S.F. (2022). Analisis faktor-faktor yang mempengaruhi pertumbuhan sektor pertanian di Provinsi Sumatera Utara. *Jurnal Agrica*, *15*, 23–34. https://doi.org/10.31289/agrica.v15i1.5065.

[9] Faoziyah, U.; Rosyaridho, M.F.; Panggabean, R. (2024). Unearthing Agricultural Land Use Dynamics in Indonesia: Between Food Security and Policy Interventions. *Land*, *13*, 2030. https://doi.org/10.3390/land13122030.

[10] Purba, B.; Zahra, D.A.; Almas, K.U.; Sihotang, S.E.A.; Mardiah, S. (2025). Pengaruh ekspor minyak kelapa sawit terhadap pertumbuhan ekonomi Indonesia Tahun 2014–2023. *Economicus: Jurnal Ekonomi dan Manajemen*, *15*, 1–10. https://doi.org/10.47860/economicus.v15i2.1.

[11] Hariyanti, F.; Syahza.; Zulkarnain.; Nofrizal. (2024). Economic transformation based on leading commodities through sustainable development of the oil palm industry. *Heliyon*, *10*, e25674. https://doi.org/10.1016/j.heliyon.2024.e25674.

[12] Panggabean, G.F.; Mansyur, A. (2023). Penerapan multivariate adaptive regression splines pada laju Produk Domestik Regional Bruto menurut lapangan usaha di Provinsi Sumatera Utara. *Jurnal Sains dan Teknologi*, *2*, 159–171. https://doi.org/10.58169/saintek.v2i1.160.

[13] Siringoringo, Y.B.P. (2025). Analysis of economic factors on food and non-food expenditure consumption in North Sumatra using multiple linear regression test. *Formosa Journal of Applied Sciences*, *4*, 2316–2326. https://doi.org/10.55927/fjas.v4i7.282.

[14] Sari, I.K.; Liyantono.; Solahudin, M. (2024). Comparative study of oil palm productivity estimation models: Multilinear regression and Holt–Winters approach. *IOP Conference Series: Earth and Environmental Science*, *1386*, 012031. https://doi.org/10.1088/1755-1315/1386/1/012031.

[15] Mughniyati, N.; Sa'roni, C. (2025). Pengaruh produksi kelapa sawit dan nilai ekspor minyak kelapa sawit terhadap Produk Domestik Regional Bruto (PDRB) sektor pertanian Provinsi Kalimantan Selatan. *JIEP: Jurnal Ilmu Ekonomi dan Pembangunan*, *6,* 136–144.

[16] Nurani, V.; Sudarno.; Rahmawati, R. (2021). Penerapan regresi linier multivariat pada distribusi ujian nasional. *Jurnal Gaussian*, *4*. 697−704. https://doi.org/10.14710/j.gauss.4.3.697-704

[17] Oktamianti, P.; Kusuma, D.; Amir, V.; Tjandrarini, D.H.; Paramita, A. District-Level Inequalities in Hypertension among Adults in Indonesia: A Cross-Sectional Analysis by Sex and Age Group. *International Journal of Environmental Research and Public Health*, *19*, 13268. https://doi.org/10.3390/ijerph192013268.

[18] Fiola, E.; Yulius, F.; Presilia, P.; Risani, D.M.; Alvionita, M.; Irawati, F.D. (2024). Metode seleksi variabel dalam pemodelan regresi linear data curah hujan Provinsi Lampung. *Prosiding Seminar Nasional Sains Data*, *4*, 351–366. https://doi.org/10.33005/senada.v4i1.213.

[19] Fernandes, A.A.R. (2022). Metode Analisis Data Penelitian: Pendekatan Regresi; Universitas Brawijaya Press: Malang, Indonesia.

[20] Nurani, V.; Rahmawati, R. (2020). Penerapan regresi linier multivariat pada distribusi ujian nasional 2014 (pada studi kasus nilai ujian nasional 2014 SMP Negeri 1 Sayung). *Jurnal Gaussian*, *4*, 697–704.