



# Agrochemicals, GHG Emissions, and GDP in Southeast Asia: A Machine Learning Approach with Hierarchical Clustering

Qalbin Salim Fazli<sup>1</sup>, Ghalieb Mutig Idroes<sup>2</sup>, Iin Shabrina Hilal<sup>3</sup>, Iffah Hafizah<sup>4</sup>, Irsan Hardi<sup>2</sup> and Teuku Rizky Noviandy<sup>5,\*</sup>

<sup>1</sup>Department of Plant Protection, Faculty of Agriculture, Universitas Syiah Kuala, Banda Aceh, 23111, Indonesia; <sup>2</sup>Economics and Sustainable Development Unit, Konstanta Utama, Aceh Besar 23371, Indonesia; <sup>3</sup>Department of Civil Engineering, Faculty of Engineering, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia; <sup>4</sup>The Bartlett School of Environment, Energy and Resources, University College London, London WC1E 6BT, United Kingdom; <sup>5</sup>Department of Information Systems, Faculty of Engineering, Universitas Abulyatama, Aceh Besar 23372, Indonesia

\* Correspondence: [rizky\\_si@abulyatama.ac.id](mailto:rizky_si@abulyatama.ac.id)

## Article History

Received  
6 May 2025

Accepted  
15 July 2025

Available Online  
23 July 2025

## Keywords

Agrochemicals  
Pesticide  
Nitrogen fertilizer  
GHG emissions  
Southeast Asia  
Hierarchical clustering analysis

## Abstract

Agrochemical use, GHG emissions, and gross domestic product (GDP) vary widely across Southeast Asia, making the region suitable for cluster-based sustainability analysis. This study applies hierarchical clustering analysis (HCA) to classify nine Southeast Asian countries using four standardized indicators: pesticide use, nitrogen fertilizer use, GHG emissions, and GDP. Exploratory data analysis reveals significant disparities, with Brunei and Indonesia emerging as outliers due to exceptionally high input intensity and emissions, respectively. HCA identifies four distinct clusters: (1) low-input, low-emission economies (Cambodia, Laos, Myanmar); (2) moderately intensive systems (Malaysia, Thailand, the Philippines, Vietnam); (3) a high-pesticide profile (Brunei); and (4) a high-emission, high-output outlier (Indonesia). Principal Component Analysis confirms the cluster structure and highlights variation in emission efficiency. The findings show that similar agroecological contexts can yield divergent environmental outcomes, emphasizing the role of policy and technology. This study provides the first region-wide, data-driven typology of agricultural sustainability in Southeast Asia using HCA.

## Introduction

Southeast Asia (SEA), as one of the most dynamic and diverse regions in the world, offers a compelling case for analyzing the intersection of agricultural development, environmental sustainability, and economic performance [1,2]. With over 655 million people and significant contributions to global food supply chains, particularly rice, palm oil, and aquaculture, the region faces growing challenges in balancing productivity with environmental stewardship [3,4]. Its varied geography, socio-economic structures, and institutional capacities result in highly diverse agricultural practices and environmental outcomes across countries [5–7].

Agriculture in Southeast Asia contributes substantially to national gross domestic product (GDP) and rural employment, particularly in lower-income countries [8]. However, the sector is also a significant source of greenhouse gas (GHG) emissions, particularly methane (CH<sub>4</sub>) from rice cultivation and nitrous oxide (N<sub>2</sub>O) from livestock and synthetic fertilizer use [9–11]. The increasing use of agrochemicals, including pesticides and nitrogen-based fertilizers, has further intensified environmental concerns related to input-driven emissions. Growing attention to climate change and food system sustainability has brought urgency to understanding how agricultural inputs and outputs affect environmental outcomes [12,13]. Yet, regional assessments remain limited and are often constrained to single-country studies or national averages that obscure important cross-country patterns [14,15].

In this context, this study conducts a comparative analysis of Southeast Asian countries using four key agricultural indicators: pesticide use, nitrogen fertilizer application, GHG emissions, and GDP. These variables were selected to reflect both environmental pressures (inputs and emissions) and economic outputs, providing a multidimensional view of agricultural sustainability. Pesticide and fertilizer use indicate the intensity of agricultural input use, which has implications for environmental degradation and indirect emissions [16,17]. GHG emissions directly reflect contributions to climate change, while GDP serves as a proxy for the sector's economic performance and productivity [18].

By integrating these indicators, this research moves beyond simple reporting of emissions or productivity to explore how different countries balance agricultural output with environmental impacts. The concept of emissions intensity, often defined as emissions per unit of economic output, is extended here to include a broader mix of input-output variables, offering a more nuanced assessment of sustainability [19–22].

To identify patterns and groupings among countries, we use machine learning, a subset of artificial intelligence that enables computers to identify structures or make predictions based on data [23,24]. Machine learning methods are broadly categorized into supervised and unsupervised techniques [25–27]. In this study, we apply hierarchical clustering analysis (HCA), an unsupervised machine learning technique well suited for small to medium-sized datasets. HCA does not require a predetermined number of groups and produces a nested structure of clusters, allowing for the visualization of relationships at multiple scales through dendrograms [28–30]. This method has proven useful in regionalization and environmental classification studies [31].

We hypothesize that countries sharing similar agricultural profiles, such as intensive rice cultivation, high fertilizer use, or low GDP, may cluster together and reveal shared structural challenges or opportunities for collaboration. For instance, countries with high input use and low GDP might benefit from targeted interventions to improve efficiency, while those achieving higher output with lower environmental intensity may offer models for sustainable agricultural practices [32–34].

Cluster memberships may also reflect surprising alignments. Countries with similar geographies may differ in cluster placement due to policy divergence or differences in technology adoption, while others with distinct ecosystems may cluster due to similar agricultural systems or development strategies [35,36]. These patterns can inform more tailored policy approaches rather than one size fits all solutions.

Ultimately, this study aims to fill a critical gap in regional agricultural sustainability analysis by offering a data driven classification of Southeast Asian countries based on four complementary indicators. The results can support regionally differentiated strategies, promote knowledge exchange, and guide policy actions aligned with climate targets under frameworks such as the Paris Agreement [37]. To our knowledge, this study use hierarchical clustering to classify countries in Southeast Asia based on a combined assessment of pesticide use, fertilizer use, GHG emissions, and GDP. This approach offers a holistic view of agriculture's dual role in food security and climate mitigation.

## Materials and Methods

---

### *Data*

The dataset comprises values from nine Southeast Asian (SEA) countries: Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar, the Philippines, Thailand, and Vietnam. Four key indicators used for the clustering analysis are presented in Table 1. The data are from the year 2021, selected to ensure temporal consistency across SEA countries and to reflect the most recent pre-pandemic trends in agricultural and environmental performance. Other SEA countries, such as Singapore and Timor-Leste, are excluded due to the unavailability or incompleteness of

consistent data across the selected indicators. To ensure the robustness and comparability of the analysis, only countries with complete and reliable data were included.

**Table 1.** Variable description.

Variable	Unit	Description	Source
Pesticide Use	Kilograms per hectare	Measures the intensity of chemical inputs in agricultural production. Higher values may indicate more intensive farming practices.	OWID [38]
Nitrogen Fertilizer Use	Kilograms per hectare	Represents the level of nitrogen-based nutrient inputs, which are critical for crop yield but may also contribute to environmental degradation.	OWID [39]
Greenhouse Gas (GHG) Emissions	Metric tons	Denotes total emissions from agricultural sources, serving as a proxy for the sector's environmental footprint and sustainability performance.	OWID [40]
Gross Domestic Product (GDP)	Constant 2015 US dollars	Reflects the overall economic output and development level of a country, adjusted for inflation to ensure temporal comparability.	WDI [41]

The selection of pesticide use, nitrogen fertilizer use, GHG emissions, and GDP for the cluster analysis reflects their importance in representing key aspects of agricultural practices, environmental impact, and economic development in Southeast Asia. Many countries in the region rely heavily on agriculture, with rising input use indicating intensification. GHG emissions capture the environmental effects of these practices, while GDP provides insight into each country's development level and ability to invest in sustainability. These variables together allow for a balanced assessment of regional differences and support informed policy planning.

All data were standardized prior to clustering by removing the mean and scaling to unit variance (z-score normalization). This preprocessing ensures that each feature contributes equally to the distance calculations in the clustering algorithm, regardless of its original scale or unit. Without standardization, variables with larger numerical ranges such as GDP or GHG emissions could dominate the clustering outcome, biasing the analysis. Standardizing the features allows the algorithm to identify clusters based on relative patterns across all variables rather than absolute magnitudes.

#### *Exploratory Data Analysis*

Exploratory data analysis (EDA) was performed to gain an initial understanding of the dataset's structure, distribution, and inter-variable relationships. The analysis began with the computation of descriptive statistics for each of the four variables: pesticide use, nitrogen fertilizer use, GHG emissions, and GDP. Measures such as the mean, standard deviation, minimum, and maximum provided insight into the central tendencies and spread of each variable across the nine Southeast Asian countries. These statistics revealed significant variation, particularly in GHG emissions and GDP, where countries like Indonesia and Brunei exhibited notably higher values compared to others.

To investigate potential relationships between variables, a correlation matrix was generated [42,43]. This matrix visually represented the strength and direction of pairwise correlations among the features. For instance, a moderate positive correlation was observed between fertilizer use and GHG emissions, suggesting that increased agricultural input may be associated with higher environmental impact. Conversely, GDP showed weaker correlations with input-related variables, indicating that economic performance in the region may not be directly aligned with agricultural intensity. Identifying such patterns is crucial, as it informs the interpretation of clustering results and supports the hypothesis that countries with similar agricultural or economic profiles may cluster together.

In addition to statistical summaries and correlation analysis, the distributions of each variable were visualized using histograms overlaid with kernel density estimation (KDE) curves [44]. These plots highlighted non-normal distributions and the presence of skewness in several variables, particularly GHG emissions and GDP, which showed right-skewed behavior due to a few high-performing outliers. Boxplots were also employed to further explore the distribution

and detect outliers. This visualization confirmed the disparities observed in the numerical summaries, reinforcing the importance of standardization prior to clustering. The combination of these EDA techniques provided a comprehensive view of the data landscape, laying the groundwork for meaningful and unbiased clustering analysis [45].

#### *Hierarchical Clustering Analysis*

To identify groups of countries with similar profiles in terms of agricultural inputs, greenhouse gas emissions, and economic output, hierarchical clustering analysis (HCA) was employed. Specifically, an agglomerative approach was used, where each country initially represents its own cluster, and clusters are iteratively merged based on similarity until a single cluster encompassing all observations is formed [46].

The similarity between countries was measured using Euclidean distance, which captures the straight-line distance between points in multidimensional space. Euclidean distance is particularly appropriate when variables are standardized, as in this study, to ensure equal weight across different feature scales [47].

The Ward linkage method was chosen to determine how clusters are merged at each step. Ward's method minimizes the total within-cluster variance, resulting in compact and spherical clusters. This method tends to produce more balanced and interpretable cluster structures compared to other linkage strategies such as single or complete linkage, which may be sensitive to outliers or generate elongated clusters [48].

The output of the clustering process was visualized using a dendrogram, which displays the hierarchical arrangement of the clusters and assists in determining an appropriate number of final groupings based on the linkage distances [49]. By cutting the dendrogram at various heights, different cluster solutions can be explored and compared, allowing for flexibility in interpretation and selection of the most meaningful segmentation of countries.

## **Results and Discussion**

---

#### *Exploratory Data Analysis Results*

Table 2 presents the descriptive statistics for the four variables analyzed across nine Southeast Asian countries. The data reveal substantial variability in all indicators, underscoring the diverse agricultural practices, environmental impacts, and economic scales within the region.

Pesticide use shows the highest relative dispersion, with a mean of 6.44 kg/ha and a standard deviation of 10.49 kg/ha. The wide range, from as low as 0.14 kg/ha to as high as 33.91 kg/ha, indicates vastly different agricultural chemical intensities, likely reflecting variations in crop types, farming technologies, and regulatory practices. Similarly, fertilizer use varies notably, though it is more centered around the mean of 53.95 kg/ha, with an interquartile range (IQR) spanning from 26.47 to 64.02 kg/ha.

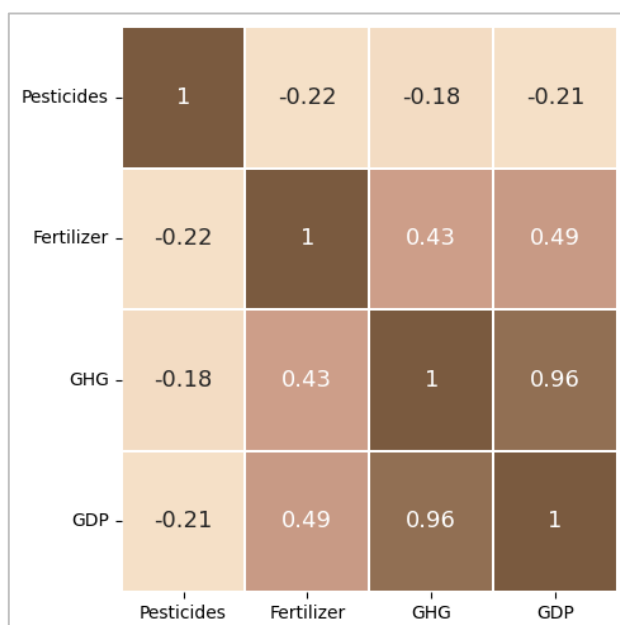
Greenhouse gas (GHG) emissions also show a highly skewed distribution, with a mean of approximately 401 million tons but a standard deviation that exceeds this average. The minimum value is around 17.6 million tons, while the maximum reaches nearly 1.65 billion tons, indicating that countries like Indonesia dominate regional emissions, likely due to land use change and extensive agricultural activity.

GDP values exhibit a similarly wide spread, ranging from \$13.2 billion to over \$1 trillion, with a mean of nearly \$300 billion. This disparity reflects the heterogeneity in national economic structures and development stages across the region. The high standard deviation suggests that economic powerhouses like Indonesia and Malaysia heavily influence regional averages.

**Table 2.** Descriptive statistics of key variables.

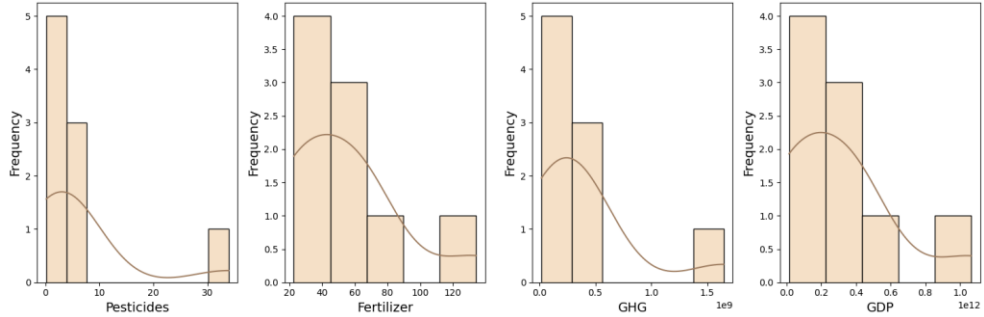
Statistic	Pesticides	Fertilizer	GHG	GDP
Mean	6.436	53.947	401,463,444	299,790,900,000
Std	10.488	35.603	499,453,470	336,504,400,000
Min.	0.14	22.68	17,570,880	13,215,630,000
25%	0.94	26.47	86,967,704	32,909,180,000
50%	3.64	55.38	269,455,600	332,245,600,000
75%	5.29	64.02	421,066,850	378,998,600,000
Max.	33.91	134.22	1,648,610,000	1,065,709,000,000

The Pearson correlation matrix among the four variables is presented in Figure 1. The strongest positive correlation is observed between GHG emissions and GDP ( $r = 0.96$ ), indicating a near-linear relationship. This suggests that countries with higher economic output also tend to have significantly higher greenhouse gas emissions, likely due to more intensive and industrialized agricultural practices. A moderate positive correlation is also found between fertilizer use and GDP ( $r = 0.49$ ), and between fertilizer use and GHG emissions ( $r = 0.43$ ). These relationships imply that nitrogen input is moderately associated with both environmental and economic outcomes, possibly reflecting increased crop productivity and greater environmental impact from synthetic input use. In contrast, pesticide use shows weak and negative correlations with the other variables. Its correlation with GDP ( $r = -0.21$ ) and GHG emissions ( $r = -0.18$ ) suggests that higher pesticide input is not a clear indicator of economic scale or environmental burden in this dataset. This may reflect variation in crop types, pest management strategies, or reporting accuracy across the sampled countries.

**Figure 1.** Correlation matrix.

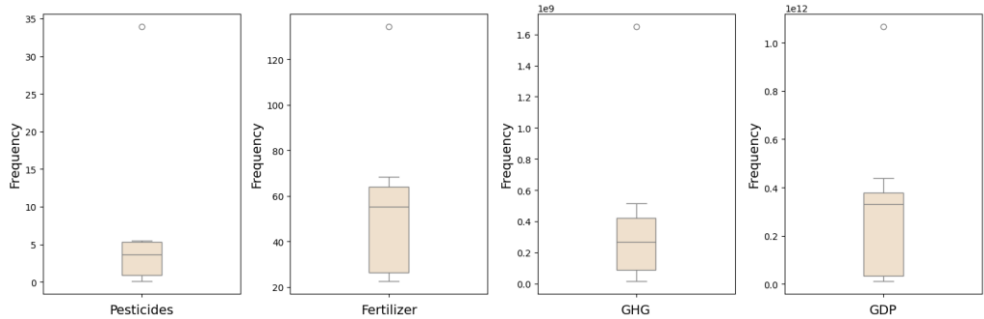
Histograms overlaid with kernel density estimation (KDE) curves are used in Figure 2 to illustrate the distributions of the four variables. These plots highlight the presence of right-skewed distributions in all variables, indicating that most countries cluster at lower values, with a few high-value outliers. Pesticide use is concentrated below 10 kg/ha, with a single extreme value above 30 kg/ha pulling the distribution sharply to the right. This suggests that most Southeast Asian countries apply relatively modest levels of pesticides, but at least one country, likely Brunei, uses a significantly higher amount. Fertilizer use also displays right skewness, though to a lesser extent. Most values lie between 20 and 70 kg/ha, but one country exceeds 130 kg/ha, contributing to the long tail. This indicates heterogeneous nitrogen application intensity across the region. GHG emissions also exhibit a pronounced long tail, with most countries emitting fewer than 500 million tons and one or two countries far exceeding this level. The distribution confirms the presence of strong outliers, such as Indonesia, which dominates

regional emissions due to its large-scale agricultural sector and land use changes. GDP exhibits a similar pattern, with most countries falling well below \$400 billion, while a few reach beyond that, including one with GDP over \$1 trillion. This long-tail distribution underscores the wide economic disparity among the countries analyzed. These non-normal distributions further justify the need for feature standardization prior to clustering, as raw scale differences would otherwise dominate distance calculations.



**Figure 2.** Histograms with kernel density estimation curves for pesticide use, fertilizer use, GHG emissions.

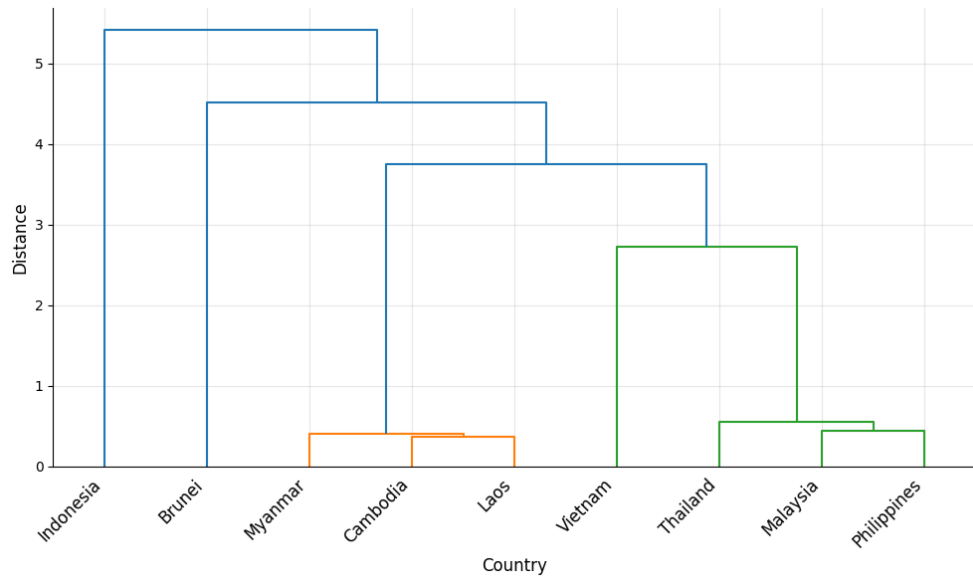
Figure 3 presents boxplots for Pesticides, Fertilizer, GHG emissions, and GDP, visually emphasizing the skewed distributions and presence of outliers observed in the previous figure. All variables exhibit right-skewness, with medians clustered near the lower quartiles and one or more extreme values extending the upper whiskers. Notably, pesticide and fertilizer use show compact interquartile ranges, while single high outliers indicate disproportionately high usage in certain countries. Similarly, GHG emissions and GDP display substantial outliers, reflecting the dominance of a few countries, such as Indonesia and Singapore in regional emissions and economic output. These patterns reinforce the need for normalization before clustering to prevent outliers and scale disparities from biasing the analysis.



**Figure 3.** Boxplots of pesticide use, fertilizer use, greenhouse gas (GHG) emissions, and GDP across nine Southeast Asian countries.

*Hierarchical Clustering Results*

The results of the hierarchical clustering analysis are visualized in Figure 4, which displays the dendrogram generated using Ward's linkage method and Euclidean distance on the standardized dataset. The dendrogram illustrates the step-by-step merging of clusters based on similarity and reveals the overall structure of relationships among the nine Southeast Asian countries.



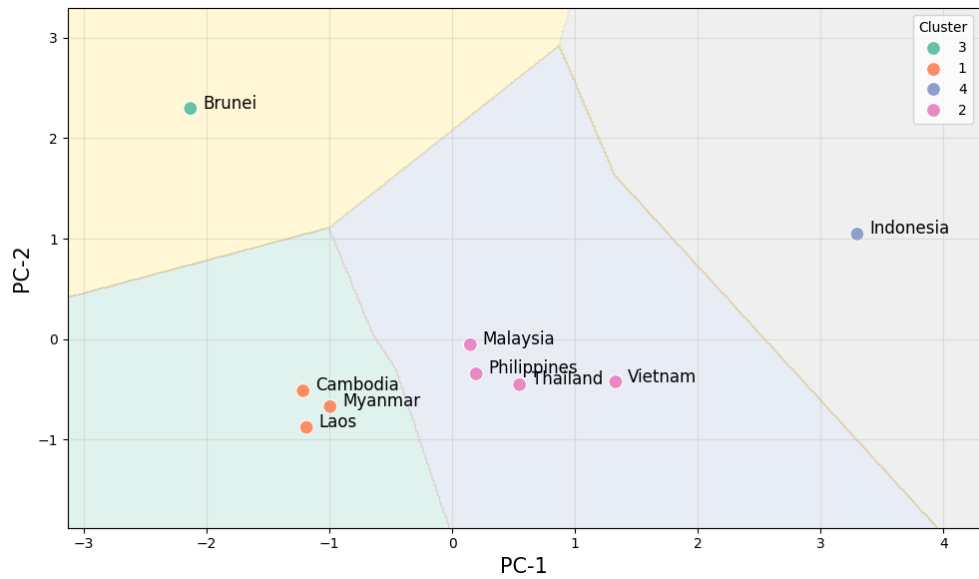
**Figure 4.** Dendrogram illustrating the hierarchical clustering of nine Southeast Asian countries.

At the lowest levels of the dendrogram, the Philippines, Malaysia, and Thailand form a tight cluster, indicating a high degree of similarity in terms of pesticide and fertilizer use, GHG emissions, and GDP. These countries are likely to share common agricultural intensification patterns and similar stages of economic development. Closely linked to this group is Vietnam, suggesting a moderate similarity, perhaps due to its rising economic status and evolving agricultural sector.

A second cluster comprises Cambodia, Laos, and Myanmar, which merge at a very low linkage distance, reflecting near-identical profiles across the analyzed indicators. This grouping likely represents countries with lower GDPs, modest agricultural input use, and relatively low GHG emissions, suggesting more traditional and less industrialized farming systems.

At a higher linkage distance, Brunei and Indonesia emerge as distinct from the rest of the countries. Brunei, while economically affluent, appears as a clear outlier due to its extremely high pesticide use despite its small scale. Indonesia, on the other hand, is separated from all other countries early in the clustering process, reflecting its dominant values across several variables, particularly GHG emissions and GDP. This outlier status aligns with Indonesia's large agricultural footprint and economic scale relative to its neighbors.

To better visualize the clustering structure and validate the results of the hierarchical analysis, a Principal Component Analysis (PCA) was conducted and plotted in Figure 5. This scatter plot projects the nine Southeast Asian countries onto the first two principal components (PC-1 and PC-2), which together capture the majority of variance in the standardized dataset. Each country is colored according to its assigned cluster from the hierarchical analysis, and decision boundaries between clusters are overlaid for clarity.

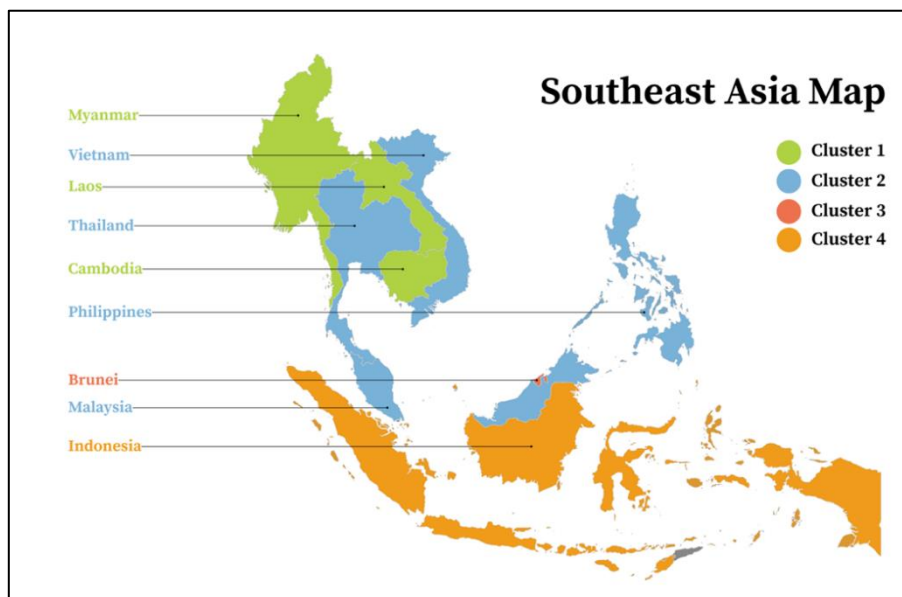


**Figure 5.** Principal Component Analysis (PCA) plot showing the distribution of Southeast Asian countries based on the first two principal components.

The PCA visualization confirms the distinctions suggested by the dendrogram. Cluster 1, comprising Cambodia, Laos, and Myanmar, appears tightly grouped in the lower-left quadrant, indicating highly similar profiles with lower agricultural inputs, emissions, and GDP. Cluster 2, which includes Malaysia, the Philippines, Thailand, and Vietnam, forms a more spread-out group in the center-right region of the plot. This suggests shared yet somewhat varied characteristics typical of emerging economies with moderate to high agricultural and economic activity.

Notably, Brunei and Indonesia are located far from the central clusters, reinforcing their status as outliers. Brunei, placed in Cluster 3, is positioned in the upper-left quadrant, likely due to its disproportionately high pesticide use relative to its small size and moderate emissions. In contrast, Indonesia, which forms Cluster 4 by itself, is positioned far to the right along PC-1, reflecting its exceptionally high GHG emissions and GDP. These extreme values cause it to stand apart from all other countries.

Figure 6 presents a geographic visualization of the clustering results, mapping the nine Southeast Asian countries according to their assigned clusters. This spatial representation complements the statistical findings and provides a clearer context for interpreting regional similarities and differences. Cluster 1, consisting of Cambodia, Laos, and Myanmar, includes countries that are geographically contiguous in mainland Southeast Asia. Their close proximity likely contributes to shared socioeconomic and environmental characteristics, such as low GDP, limited agricultural input use, and minimal greenhouse gas emissions, which reflect traditional and less industrialized farming systems. Cluster 2 encompasses Malaysia, the Philippines, Thailand, and Vietnam, countries that are more geographically dispersed but maintain relatively close economic ties and similar levels of agricultural intensification. Their positioning suggests moderate to high input use and economic activity typical of emerging economies undergoing rapid transformation. Brunei, assigned to Cluster 3, is geographically and statistically distinct, located on the northern coast of Borneo and characterized by its small population size and disproportionately high pesticide use. Indonesia, forming Cluster 4 alone, is the largest and most diverse country in the region, both geographically and economically. Its separation from other countries reflects its dominant values across several indicators, particularly greenhouse gas emissions and GDP. Overall, the spatial distribution of clusters highlights the role of geographic and economic context in shaping agricultural and environmental profiles across Southeast Asia, reinforcing the patterns observed in both the dendrogram and the principal component analysis.



**Figure 6.** Cluster visualization based on the findings.

## Conclusions

This study reveals significant diversity in the environmental and economic dimensions of agriculture across Southeast Asia. Through hierarchical clustering analysis of nine countries, four distinct groupings emerged: (1) low-input, low-emission economies; (2) moderately intensive agricultural systems; (3) unique high-input outliers; and (4) high-output, high-emission systems. Indonesia and Brunei stood out due to their exceptional emission levels and pesticide use, respectively, while countries like Cambodia, Laos, and Myanmar reflected more traditional, lower-impact agricultural profiles.

The findings emphasize that agricultural sustainability in the region is highly uneven, shaped by both structural and policy-related factors. Countries with similar geographies do not always share similar emission intensities or input profiles, highlighting the importance of context-specific strategies. Rather than relying on generalized policies, targeted interventions tailored to the characteristics of each cluster are needed to enhance productivity without compromising environmental goals. This analysis provides a clearer basis for regional cooperation and policy alignment in advancing low-emission, climate-resilient agricultural systems in Southeast Asia.

**Funding:** This study does not receive external funding.

**Data Availability Statement:** The data are publicly available.

**Acknowledgments:** The authors would like to express their gratitude to their respective institutions.

**Conflicts of Interest:** All the authors declare that there are no conflicts of interest.

## References

- [1] Bhatti UA, Bhatti MA, Tang H, Syam MS, Awwad EM, Sharaf M, et al. Global Production Patterns: Understanding the Relationship between Greenhouse Gas Emissions, Agriculture Greening and Climate Variability. *Environmental Research* 2024;245:118049. <https://doi.org/10.1016/j.envres.2023.118049>.
- [2] Hardi I, Ray S, Attari MUQ, Ali N, Idroes GM. Innovation and Economic Growth in the Top Five Southeast Asian Economies: A Decomposition Analysis. *Ekonomikalia Journal of Economics* 2024;2:1–14. <https://doi.org/10.60084/eje.v2i1.145>.

- [3] Modisa, Jaikaew T. Sustainability in Food Supply Chains: Case of Southeast Asia. In: Mor RS, Kamble SS, Sangwan KS, editors. *Oper. Supply Chain Manag. Food Ind.*, Singapore: Springer Nature Singapore; 2022, p. 53–72.
- [4] Hardi I, Idroes GM, Hamaguchi Y, Can M, Noviandy TR, Idroes R. Business Confidence in the Shift to Renewable Energy: A Country-Specific Assessment in Major Asian Economies. *Journal of Economy and Technology* 2025;3:44–68. <https://doi.org/10.1016/j.ject.2024.08.002>.
- [5] Blesh J, Mehrabi Z, Wittman H, Kerr RB, James D, Madsen S, et al. Against the Odds: Network and Institutional Pathways Enabling Agricultural Diversification. *One Earth* 2023;6:479–91. <https://doi.org/10.1016/j.oneear.2023.03.004>.
- [6] Mouratiadou I, Wezel A, Kamilia K, Marchetti A, Paracchini ML, Bàrberi P. The Socio-Economic Performance of Agroecology. A Review. *Agronomy for Sustainable Development* 2024;44:19. <https://doi.org/10.1007/s13593-024-00945-9>.
- [7] Surna MI, Fazli QS, Chamzurni T, Susanna S, Idroes GM. Influence of Elevational and Environmental Factors on Parasitic Nematode Distribution in Arabica Coffee in the Gayo Highlands, Indonesia. *Leuser Journal of Environmental Studies* 2025;3:56–66. <https://doi.org/10.60084/ljes.v3i2.293>.
- [8] Chandio AA, Gokmenoglu KK, Dash DP, Khan I, Ahmad F, Jiang Y. Exploring the Energy-Climate-Agriculture (ECA) Nexus: A Roadmap toward Agricultural Sustainability in Asian Countries. *Environment, Development and Sustainability* 2024. <https://doi.org/10.1007/s10668-023-04418-z>.
- [9] Qian H, Zhu X, Huang S, Linqvist B, Kuzyakov Y, Wassmann R, et al. Greenhouse Gas Emissions and Mitigation in Rice Agriculture. *Nature Reviews Earth & Environment* 2023;4:716–32. <https://doi.org/10.1038/s43017-023-00482-1>.
- [10] Xing Y, Wang X. Impact of Agricultural Activities on Climate Change: A Review of Greenhouse Gas Emission Patterns in Field Crop Systems. *Plants* 2024;13:2285. <https://doi.org/10.3390/plants13162285>.
- [11] Fazli QS, Pramayudi N, Sayuthi M. Uji Keefektifan Ekstrak Daun Gegarang (*Mentha sp.*) untuk Mengendalikan Hama *Plutella xylostella* L. di Laboratorium. *Jurnal Ilmiah Mahasiswa Pertanian* 2023;8:1004–14.
- [12] Wijerathna-Yapa A, Pathirana R. Sustainable Agro-Food Systems for Addressing Climate Change and Food Security. *Agriculture* 2022;12:1554. <https://doi.org/10.3390/agriculture12101554>.
- [13] Lin H-I, Yu Y-Y, Wen F-I, Liu P-T. Status of Food Security in East and Southeast Asia and Challenges of Climate Change. *Climate* 2022;10:40. <https://doi.org/10.3390/cli10030040>.
- [14] Ren C, He L, Ma Y, Reis S, Van Grinsven H, Lam SK, et al. Trade-Offs in Agricultural Outcomes across Farm Sizes. *Earth Critical Zone* 2024;1:100007. <https://doi.org/10.1016/j.ecz.2024.100007>.
- [15] Xiang T, Malik TH, Hou JW, Ma J. The Impact of Climate Change on Agricultural Total Factor Productivity: A Cross-Country Panel Data Analysis, 1961–2013. *Agriculture* 2022;12:2123. <https://doi.org/10.3390/agriculture12122123>.
- [16] Ma B, Karimi MS, Mohammed KS, Shahzadi I, Dai J. Nexus between Climate Change, Agricultural Output, Fertilizer Use, Agriculture Soil Emissions: Novel Implications in the Context of Environmental Management. *Journal of Cleaner Production* 2024;450:141801. <https://doi.org/10.1016/j.jclepro.2024.141801>.
- [17] Gerssen-Gondelach SJ, Wicke B, Faaij APC. GHG Emissions and Other Environmental Impacts of Indirect Land Use Change Mitigation. *GCB Bioenergy* 2017;9:725–42. <https://doi.org/10.1111/gcbb.12394>.
- [18] Idroes GM, Hardi I, Hilal IS, Utami RT, Noviandy TR, Idroes R. Economic Growth and Environmental Impact: Assessing the Role of Geothermal Energy in Developing and Developed Countries. *Innovation and Green Development* 2024;3:100144. <https://doi.org/10.1016/j.igd.2024.100144>.
- [19] Coderoni S, Vanino S. The Farm-by-Farm Relationship among Carbon Productivity and Economic Performance of Agriculture. *Science of The Total Environment* 2022;819:153103. <https://doi.org/10.1016/j.scitotenv.2022.153103>.
- [20] Kamyab H, SaberiKamarposhti M, Hashim H, Yusuf M. Carbon Dynamics in Agricultural Greenhouse Gas Emissions and Removals: A Comprehensive Review. *Carbon Letters* 2024;34:265–89. <https://doi.org/10.1007/s42823-023-00647-4>.
- [21] Hardi I, Idroes GM, Zulham T, Suriani S, Saputra J. Economic Growth, Agriculture, Capital Formation and Greenhouse Gas Emissions in Indonesia: FMOLS, DOLS and CCR Applications. *Ekonomikalia Journal of Economics* 2023;1:82–91. <https://doi.org/10.60084/eje.v1i2.109>.
- [22] Maulidar P, Fitriyani F, Sasmita NR, Hardi I, Idroes GM. Exploring Indonesia's CO2 Emissions: The Impact of Agriculture, Economic Growth, Capital and Labor. *Grimsa Journal of Business and Economics Studies* 2024;1:43–55. <https://doi.org/10.61975/gjbes.v1i1.22>.

- [23] Noviandy TR, Maulana A, Emran TB, Idroes GM, Idroes R. QSAR Classification of Beta-Secretase 1 Inhibitor Activity in Alzheimer's Disease Using Ensemble Machine Learning Algorithms. *Heca Journal of Applied Sciences* 2023;1:1–7. <https://doi.org/10.60084/hjas.v1i1.12>.
- [24] Noviandy TR, Maulana A, Idroes GM, Mau Lydia NB, Patwekar M, Suhendra R, et al. Integrating Genetic Algorithm and LightGBM for QSAR Modeling of Acetylcholinesterase Inhibitors in Alzheimer's Disease Drug Discovery. *Malacca Pharmaceutics* 2023;1:48–54. <https://doi.org/10.60084/mp.v1i2.60>.
- [25] Noviandy TR, Idroes GM, Hardi I. Integrating Explainable Artificial Intelligence and Light Gradient Boosting Machine for Glioma Grading. *Informatics and Health* 2025;2:1–8. <https://doi.org/10.1016/j.infoh.2024.12.001>.
- [26] Noviandy TR, Maulana A, Irvanizam I, Idroes GM, Mau Lydia NB, Tallei TE, et al. Interpretable Machine Learning Approach to Predict Hepatitis C Virus NS5B Inhibitor Activity Using Voting-Based LightGBM and SHAP. *Intelligent Systems with Applications* 2025;25:200481. <https://doi.org/10.1016/j.iswa.2025.200481>.
- [27] Noviandy TR, Maulana A, Idroes GM, Suhendra R, Afidh RPF, Idroes R. An Explainable Multi-Model Stacked Classifier Approach for Predicting Hepatitis C Drug Candidates. *Sci* 2024;6:81. <https://doi.org/10.3390/sci6040081>.
- [28] Zang L, Xue Q, Wang X, Tian B, Zhu S. An Optimized Product Model Based on the K-Means++ Algorithm. CAIBDA 2024 2024 4th Int. Conf. Artif. Intell. Big Data Algorithms, ACM; 2024, p. 677–84. <https://doi.org/10.1145/3690407.3690522>.
- [29] Tarnutzer AA, Weber KP. Pattern Analysis of Peripheral-Vestibular Deficits with Machine Learning Using Hierarchical Clustering. *Journal of the Neurological Sciences* 2022;434:120159. <https://doi.org/10.1016/j.jns.2022.120159>.
- [30] Boegl L, Kounadi O. Introducing Spatial Heterogeneity via Regionalization Methods in Machine Learning Models for Geographical Prediction: A Spatially Conscious Paradigm. *European Journal of Geography* 2024;15:244–55. <https://doi.org/10.48088/ejg.l.boe.15.4.244.255>.
- [31] Shui B, Cai Z, Luo X. Towards Customized Mitigation Strategy in the Transportation Sector: An Integrated Analysis Framework Combining LMDI and Hierarchical Clustering Method. *Sustainable Cities and Society* 2024;107:105340. <https://doi.org/10.1016/j.scs.2024.105340>.
- [32] Yang X, Liu Y, Bezama A, Thrän D. Agricultural Carbon Emission Efficiency and Agricultural Practices: Implications for Balancing Carbon Emissions Reduction and Agricultural Productivity Increment. *Environmental Development* 2024;50:101004. <https://doi.org/10.1016/j.envdev.2024.101004>.
- [33] Liu B, Guan Y, Shan Y, Cui C, Hubacek K. Emission Growth and Drivers in Mainland Southeast Asian Countries. *Journal of Environmental Management* 2023;329:117034. <https://doi.org/10.1016/j.jenvman.2022.117034>.
- [34] Qayyum M, Zhang Y, Wang M, Yu Y, Li S, Ahmad W, et al. Advancements in Technology and Innovation for Sustainable Agriculture: Understanding and Mitigating Greenhouse Gas Emissions from Agricultural Soils. *Journal of Environmental Management* 2023;347:119147. <https://doi.org/10.1016/j.jenvman.2023.119147>.
- [35] D'Orazio P. Mapping the Emergence and Diffusion of Climate-Related Financial Policies: Evidence from a Cluster Analysis on G20 Countries. *International Economics* 2022;169:135–47. <https://doi.org/10.1016/j.inteco.2021.11.005>.
- [36] Grillitsch M, Sotarauta M. Trinity of Change Agency, Regional Development Paths and Opportunity Spaces. *Progress in Human Geography* 2020;44:704–23. <https://doi.org/10.1177/0309132519853870>.
- [37] Islam MS, Kieu E. Tackling Regional Climate Change Impacts and Food Security Issues: A Critical Analysis across ASEAN, PIF, and SAARC. *Sustainability* 2020;12:883. <https://doi.org/10.3390/su12030883>.
- [38] Our World in Data (OWID). Pesticides 2023. <https://ourworldindata.org/pesticides>.
- [39] Our World in Data (OWID). Fertilizers 2023. <https://ourworldindata.org/fertilizers>.
- [40] Our World in Data (OWID). CO<sub>2</sub> and Greenhouse Gas Emissions 2024. <https://ourworldindata.org/#entries>.
- [41] World Bank. WorldBank Development Indicator Databased 2023. <https://databank.worldbank.org/source/world-development-indicators>.
- [42] Idroes GM, Syahnur S, Majid SA, Sasmita NR, Idroes R. Provincial Economic Level Analysis in Indonesia Based on the Geothermal Energy Potential and Growth Regional Domestic Products Using Cluster Analysis. *IOP Conference Series: Materials Science and Engineering* 2021;1087:012079. <https://doi.org/10.1088/1757-899X/1087/1/012079>.
- [43] Noviandy TR, Nisa K, Idroes GM, Hardi I, Sasmita NR. Classifying Beta-Secretase 1 Inhibitor Activity for Alzheimer's Drug Discovery with LightGBM. *Journal of Computing Theories and Applications* 2024;1:358–67. <https://doi.org/10.62411/jcta.10129>.
- [44] Sasmita NR, Phonna RA, Fikri MK, Khairul M, Apriliansyah F, Idroes GM, et al. Statistical Assessment of Human Development Index Variations and Their Correlates: A Case Study of Aceh Province, Indonesia. *Grimsa Journal of Business and Economics Studies* 2023;1:12–24. <https://doi.org/10.61975/gjbes.v1i1.14>.

- [45] Noviandy TR, Hardi I, Zahriah Z, Sofyan R, Sasmita NR, Hilal IS, et al. Environmental and Economic Clustering of Indonesian Provinces: Insights from K-Means Analysis. *Leuser Journal of Environmental Studies* 2024;2:41–51. <https://doi.org/10.60084/ljes.v2i1.181>.
- [46] Gil-Garcia RJ, Badia-Contelles JM, Pons-Porrata A. A General Framework for Agglomerative Hierarchical Clustering Algorithms. 18th Int. Conf. Pattern Recognit., IEEE; 2006, p. 569–72. <https://doi.org/10.1109/ICPR.2006.69>.
- [47] Sasmita NR, Khairul M, Sofyan H, Kruba R, Mardalena S, Dahlawy A, et al. Statistical Clustering Approach: Mapping Population Indicators Through Probabilistic Analysis in Aceh Province, Indonesia. *Infolitika Journal of Data Science* 2023;1:63–72. <https://doi.org/10.60084/ijds.v1i2.130>.
- [48] Vijaya, Sharma S, Batra N. Comparative Study of Single Linkage, Complete Linkage, and Ward Method of Agglomerative Clustering. 2019 Int. Conf. Mach. Learn. Big Data, Cloud Parallel Comput., IEEE; 2019, p. 568–73. <https://doi.org/10.1109/COMITCon.2019.8862232>.
- [49] Granato D, Santos JS, Escher GB, Ferreira BL, Maggio RM. Use of Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) for Multivariate Association between Bioactive Compounds and Functional Properties in Foods: A Critical Perspective. *Trends in Food Science & Technology* 2018;72:83–90. <https://doi.org/10.1016/j.tifs.2017.12.006>.