Journal of
Information Systems Engineering
and Business Intelligence

Vol.11, No.3, October 2025
Available online at: http://e-journal.unair.ac.id/index.php/JISEBI

Clustering and Mixture Modeling of Schooling Expectancy Trends in Papua Province: A Spatial Analysis Using the Mapping Toolbox

Jonathan Kiwasi Wororomi [1)*], Felix Reba [2)], Frans A. Asmuruf [3)]

[1),2)] *Department of Mathematics, Faculty of Mathematics and Natural Sciences, Cenderawasih University,* Jayapura, Indonesia
[1)] *jkwororomi@gmail.com*, [2)]*felix.reba85@gmail.com*

[3)] *Chemistry Department, Faculty of Mathematics and Natural Sciences, Universitas Cenderawasih, Jayapura, Indonesia*
[3)] frans.asmuruf@fmipa.uncen.ac.id

Abstract

Background: Persistent educational inequality in Papua Province, particularly in remote highland districts, is driven by limited infrastructure and accessibility. Although Schooling Expectancy (Harapan Lama Sekolah, HLS) is widely recognized as a forward-looking educational metric, existing studies rarely incorporate probabilistic modeling with spatial analysis to examine regional disparities.
ObjectiveThis study aimed to identify spatial and statistical patterns of schooling expectancy across 29 districts in Papua from 2010 to 2023 by combining probabilistic clustering with spatial visualization methods.
Methods: The analysis applied Gaussian Mixture Model (GMM) clustering, which was validated using the Silhouette Index and Davies–Bouldin Index (DBI), to group districts based on HLS trends. Fourteen candidate probability distributions were evaluated using Kolmogorov–Smirnov and Anderson–Darling tests. In addition, five model selection criteria (AIC, BIC, AICc, CAIC, HQC) were applied to refine the fit. Cluster-wise mixture model was constructed, and spatial interpretation was improved through MATLAB's Mapping Toolbox as well as wind rose diagrams.
Results: During the process of the analysis, four statistically distinct clusters were identified. Cluster 3 (coastal districts) showed the highest and most stable HLS (12.1–14.0 years), while Cluster 4 (remote highlands) signified the lowest (2.4–5.6 years) with high dispersion. Right-skewed distributions (e.g., Weibull, Gamma) modeled high-performing districts, and heavy-tailed, left-skewed ones (e.g., Stable, Inverse Gaussian) modeled marginalized regions. Spatial visualization confirmed a clear coastal–highland divide in educational attainment.
Conclusion: The proposed incorporation of probabilistic modeling and spatial clustering offered a robust analytical tool for capturing intra-regional educational disparities. This framework provided empirical evidence to support geographically differentiated policy interventions in Papua and could be adapted to similar underserved regions in future studies.

Keywords: Schooling Expectancy, Gaussian Mixture Model, Probabilistic Modeling, Silhouette Index, Davies–Bouldin Index, Spatial Clustering, Education Inequality, Papua Province.

Article history: Received 25 April 2025, first decision 14 July 2025, accepted 06 October 2025, available online 28 October 20XX

I. INTRODUCTION

Schooling expectancy, also known as Harapan Lama Sekolah (HLS), is a process that represents the projected number of years a child is expected to spend in formal education. As a forward-looking educational indicator, HLS plays a crucial role in evaluating long-term investment in education and informing equitable policy decisions [1]. In Papua Province, Indonesia, HLS remains significantly beneath the national average due to geographic isolation, infrastructural limitations, and socio-economic disparities, particularly in the central mountainous regions.

In contrast to aggregate indicators such as the Human Development Index (HDI), which often obscure disparities in regions, HLS offers a more focused method to examine educational development. Even though various studies have analyzed educational inequality using HDI or average years of schooling [2], [3], relatively few have used probabilistic clustering methods or spatial visualization to study HLS specifically in marginalized regions such as Papua. Literature concerning health disparity [4], digital infrastructure inequality [5], and regional development [6], as well as spatial

* Corresponding author

ISSN 2443-2555 (online) 2598-6333 (print)
doi: http://dx.doi.org/10.20473/jisebi.11.3.459-472

inequality using exploratory spatial data analysis and spatial econometrics [7], prioritizes the need for sub-national spatial analysis in understanding regional disparities.

This study positions HLS analysis in the broader domain of Information Systems and Business Intelligence outside the statistical interpretation. By using clustering methods such as Gaussian Mixture Model (GMM) combined with spatial visualization, the analysis shows how raw educational data can be transformed into meaningful visions for decision support. This perspective signifies the role of information systems in guiding data-driven policies and geographically targeted educational planning, which is associated with the core focus of the journal on decision-support as well as business intelligence systems.

This study addresses the gap by combining GMM clustering with spatial mapping to analyze HLS data from 29 districts in Papua between 2010 and 2023. GMM enables soft classification, assigning membership probabilities for each cluster rather than forcing hard boundaries. This method is particularly suitable for regions such as Papua, where socioeconomic and geographic heterogeneity is pronounced. Spatial mapping via MATLAB's Mapping Toolbox further improves interpretability by visualizing patterns across geography. The methodology is also informed by clustering ensemble frameworks [8] and comparative evaluations in educational contexts where ensemble model such as GMM have shown adaptability in handling complex datasets [9].

The analysis builds upon previous work conducted by the same study team, which applied GMM to examine human development disparities in Papua based on the Mean Years of Schooling (MYS) indicator from 2010 to 2023 [10]. As both studies used a probabilistic modeling framework, the earlier investigation focused solely on four probability distributions. These divisions include Inverse Gaussian, Rician, Weibull, and Nakagami, without incorporating explicit spatial analysis. The present study significantly broadens the methodological scope by evaluating 14 candidate probability distributions, including less commonly used features in educational studies, such as Stable and t Location-Scale. Furthermore, the study incorporates spatial mapping and directional pattern analysis (wind rose) to effectively capture geographic variation in schooling expectancy**,** advancing both the analytical depth as well as policy relevance of the results. This is the first study to incorporate GMM clustering with spatial analysis and probabilistic modeling using 14 candidate distributions to assess schooling expectancy at the district level in Papua.

The analysis offers practical value by providing empirical evidence for targeted intervention. Cluster-based understanding enables policymakers to identify districts most in need of support. This supports recent academic demands for data-driven, spatially explicit methods to educational policy-making [11], [12], [13].

II. METHODS

This study adopted a spatial-statistical framework to analyze trends in schooling expectancy across 29 districts in Papua Province from 2010 to 2023. The HLS data was clustered using GMM, with cluster quality validated through two evaluation metrics, namely, Silhouette Score and Davies–Bouldin Index (DBI).

Probability distribution parameters were estimated using Maximum Likelihood Estimation (MLE) and evaluated through Anderson–Darling (AD) as well as Kolmogorov–Smirnov (KS) goodness-of-fit tests. During the process, fourteen candidate distributions were assessed using AIC, BIC, CAIC, and HQC to determine the best-fitting model for each district.

Figure 1 showed the total methodological workflow, from data collection to the interpretation of clustering results. Spatial visualization was conducted using MATLAB's Mapping Toolbox, which enabled the identification of educational disparities across regions. In addition, wind rose diagrams were used to represent directional clustering trends. The final outputs were analyzed in relation to geographical, infrastructural, and policy-related factors, supporting evidence-based recommendations for targeted educational interventions.

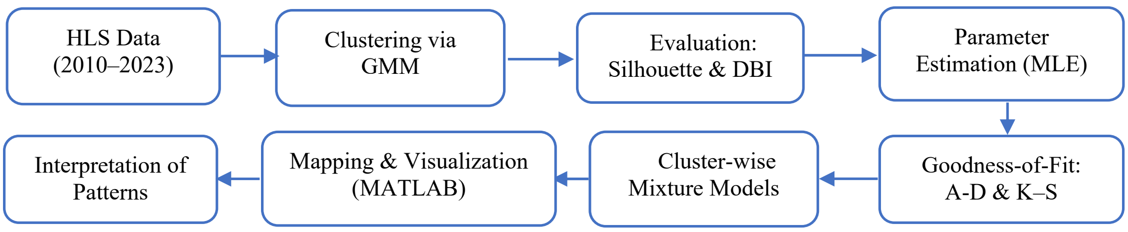


Fig. 1 Overview of statistical and spatial methods applied to schooling expectancy data (2010–2023)

A. Schooling Expectancy (HLS): Educational Significance, Data Source, and Preparation

Schooling expectancy was a forward-looking educational indicator that estimated the total number of years a child was expected to spend in formal education, based on current enrolment rates and the structure of the national education system [1]. It played a critical role in monitoring long-term educational development, as the process reflected the capacity of a region to provide sustained access to education across all levels.

Differing from retrospective measures such as Mean Years of Schooling (RLS), which assessed past educational attainment, HLS prioritized opportunities and planning for the future. This made the process specifically valuable for evaluating policy impacts and projecting outcomes in underdeveloped or geographically isolated regions such as Papua [14]. Regions with low HLS typically faced barriers such as poor school accessibility, limited infrastructure, teacher shortages, or sociocultural constraints that hindered participation [15]. Consequently, districts with higher HLS tended to benefit from stronger governance, targeted investment, and better policy implementation. Analyzing HLS patterns spatially enabled the identification of localized educational inequalities and supported evidence-based policy design to ensure equitable as well as inclusive education.

The dataset used in this study was obtained from the Papua Provincial Statistics Office (*Badan Pusat Statistik Papua*) through the official open-data platform (https://papua.bps.go.id), accessed on January 15, 2024. The data contained annual HLS data for 29 districts from 2010 to 2023, compiled according to the national methodology introduced by BPS in 2020 [11]. Before analysis, the dataset passed through standardization of district names, removal of duplicate records, and linear interpolation for minor missing values (<5%). The data were then transformed into MATLAB-compatible formats for spatial and statistical modeling [16].

Internal verification was conducted by cross-referencing yearly values with official government publications, including *Papua in Figures* and *Social and Economic Indicators of Papua Province*, ensuring both temporal consistency as well as inter-source reliability. The official provenance and validation of the dataset by the central statistical agency of Indonesia provided high credibility. Moreover, the selection of HLS over traditional retrospective indicators such as RLS [1] offered added conceptual value for forward-oriented educational planning.

The incorporation of high-quality, geographically specific, and temporally rich data provided a strong foundation for probabilistic modeling, spatial clustering, and for developing information system tools that transformed raw educational data into actionable perceptions. This dual contribution strengthened both the methodological consistency and the practical value of the study, supporting data-driven policy design for addressing educational inequality in Papua.

B. Clustering Methods and GMM for Educational Data Analysis

To examine regional disparities in schooling expectancy across Papua, this study adopted GMM as a probabilistic clustering method. GMM accommodated overlapping distributions and assigned soft membership probabilities different from traditional clustering methods (e.g., K-Means), making it particularly suitable for educational data with heterogeneous patterns [8], [9]. Mathematically, the GMM was defined as:

$$p(x) = \sum_{i=1}^{K} \pi_i \mathcal{N}(x|\mu_i, \Sigma_i) \qquad (1)$$

Where π_i represented the mixing proportions, and $\mathcal{N}(x|\mu_i, \Sigma_i)$ signified the multivariate normal distribution for each component i with mean μ_i and covariance matrix Σ_i.

The clustering was applied to HLS data from 29 districts (2010–2023) using the Expectation-Maximization (EM) algorithm [17], with random initialization, a maximum of 100 iterations, and a convergence threshold of 1×10^{-6}. The optimal number of clusters (K) was determined using the Bayesian Information Criterion (BIC) for model parsimony and validated using the Silhouette Index [18] as well as DBI [19], ensuring robust and interpretable clusters reflecting regional schooling inequality.

Following clustering, the statistical distribution of HLS in each cluster was modeled. Fourteen candidate probability distributions were tested to reflect data characteristics such as asymmetry, heavy tails, and socio-economic skewness [20]. During the process, goodness-of-fit was assessed using the KS and AD tests ($p > 0.05$), after which a mixture model was built by combining the best-fitting distributions for each district through a weighted linear combination based on the inverse of KS statistics.

$$f_{mixture}(x) = \sum_{i=1}^{N} w_i f_i(x) \qquad (2)$$

Where $f_i(x)$ represented the probability density function (PDF) of the i −th distribution, and w_i was the normalized weight such that $\sum w_i = 1$. Distributions with better goodness-of-fit (i.e., higher p-values) received larger weights.

The incorporated method, which combined GMM clustering with weighted mixture modeling, improved the interpretability of both spatial and statistical trends in schooling expectancy. It was specifically valuable for capturing complex regional variation and distributional changes in HLS, while offering improved flexibility for data-driven, model-based educational analysis [21].

C. Spatial Visualization Using MATLAB's Mapping Toolbox

This study used MATLAB's Mapping Toolbox, which provided comprehensive tools for importing, analyzing, and visualizing geospatial data to improve the spatial interpretation of schooling expectancy trends in Papua Province. The process was particularly valuable for regional education studies, including administrative boundaries and spatial disparities.

In this study, HLS data from 29 districts were converted into georeferenced polygon layers representing administrative regions. The clustering results from GMM were then spatially visualized using the geoshow() and mapshow() functions. These functions generated thematic maps with distinct color classifications for each cluster, enabling easier identification of regional disparities in educational development.

Spatial visualization played a critical role in helping policymakers and stakeholders detect clusters of low schooling expectancy and prioritize targeted interventions. By incorporating time-series HLS data, this study animated temporal changes and showed year-by-year shifts in educational outcomes.

The application of MATLAB's Mapping Toolbox had been widely validated in various socio-economic and geodemographic contexts. For instance, Saarenpää et al. [22] showed its effectiveness in spatial clustering and visualization related to early electric vehicle adoption, signifying the strength of the toolbox in managing spatial data with socio-economic attributes.

Several validation steps were assumed to ensure the effectiveness of the spatial visualization in this study. First, the clarity of spatial patterns, particularly the coastal–highland divide, was evaluated by cross-referencing the cluster map with the wind rose diagram (Figure 4), ensuring statistical consistency between spatial orientations and cluster assignments. Second, the position between spatial groupings and actual HLS values was verified using Table 1b, which confirmed the contextual relevance of the cluster divisions. Third, visual clarity and readability were evaluated by assessing label overlap, marker visibility, and map precision using MATLAB's geobasemap rendering. These steps ensured that the spatial display preserved both analytical rigor and accessibility, while also showing the potential as a component of decision-support systems, transforming raw educational data into actionable visions for policy-making.

D. Probability Distribution Selection and Validation for Schooling Expectancy Analysis

In analyzing educational data, particularly schooling expectancy, it was essential to identify suitable probability distributions that accurately represented the statistical characteristics of each cluster. This study evaluated HLS distribution patterns from GMM clustering by testing 14 candidate distributions, selected for individual ability to capture common features in socio-educational datasets such as skewness, heavy tails, and kurtosis [23], [24].

The selection was guided by two goodness-of-fit tests, namely the KS and the AD. The examinations measured the configuration between theoretical and observed distributions during the process. The KS test was recognized for its sensitivity to discrepancies in the distribution tails, while AD test, as discussed by Pestman [25], effectively complemented the initial assessment. Pobočíková et al. [26] showed the suitability of these tests for selecting best-fitting distributions under asymmetric and heavy-tailed conditions, relevant for modeling HLS data.

This study applied five widely used information criteria to address model complexity. These criteria included Akaike Information Criterion (AIC), BIC, corrected AIC (AICc), consistent AIC (CAIC), and the Hannan–Quinn Criterion (HQC) [27], [24]. The criteria provided a framework for model evaluation by balancing goodness-of-fit and model simplicity. During the analysis, AIC and AICc prioritized flexibility in modeling, BIC, as well as CAIC, focused on parsimony, while HQC offered a middle ground.

Based on the combined outcomes of the goodness-of-fit tests and minimized information criteria, the best-fitting distributions were incorporated into cluster-level mixture model that captured intra-cluster variability. Further model validation included assessment of log-likelihood values, residuals, and cumulative distribution function (CDF) plots.

Rather than getting a single 'best' distribution, this study adopted a mixture modeling method to represent the complex variation in HLS across Papua. As shown in [23], [28], and [24], the probabilistic selection of distributions was relevant in socio-educational data, where interpretability was crucial. As the statistical consistency ensured an accurate representation of data structures, the greater contribution was in transforming these results into information system components for decision support. By incorporating probabilistic selection with spatial clustering and visualization, this study bridged statistical modeling with actionable tools that supported data-driven educational policy as well as regional planning.

III. RESULTS

The results were structured into clustering analysis, mixture distribution modeling, and spatial interpretation using Mapping Toolbox to address the study aim specified in the title.

A. Cluster-Based Analysis of Schooling Expectancy Distribution Using GMM

The schooling expectancy data across 29 districts in Papua Province from 2010 to 2023 were clustered using the GMM, a probabilistic method that accommodated overlapping regional characteristics in educational data. The optimal number of clusters (k) was determined using two internal validity indices, namely, the Silhouette Coefficient and DBI, ensuring both cohesion as well as separation. Table 1(a) showed that the configuration with four clusters produced the highest Silhouette score (0.713) and the lowest DBI score (0.580), signifying a suitable balance between intra-cluster compactness and inter-cluster separation.

TABLE 1
EVALUATION OF OPTIMAL CLUSTER NUMBER (K) USING SILHOUETTE INDEX AND DBI

k Optimum	2	3	4	5	6	7	8	9
Silhouette	0.677	0.639	**0.713**	0.653	0.491	0.578	0.467	0.413
DBI	0.593	0.618	**0.580**	0.607	0.664	0.623	0.736	0.739

Following the selection of four clusters, the annual schooling expectancy trends were aggregated and reviewed. Table 1(b) showed that the average HLS values from 2010 to 2023 presented distinctive temporal patterns across clusters. For example, Cluster 3 consistently showed the highest schooling expectancy (12.1–14.0 years), while Cluster 4 remained the lowest (2.4–5.6 years) throughout the 14 years. These values reflected the temporal evolution of educational conditions across spatial groupings.

TABLE 2
AVERAGE SCHOOLING EXPECTANCY BY CLUSTER FROM 2010 TO 2023

Year													
2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Cluster 1													
5.7	6.0	6.4	6.8	7.1	7.2	7.3	7.54	7.8	8.09	8.3	8.51	8.6	8.81
Cluster 2													
9.8	10.0	10.2	10.3	10.5	10.6	10.7	10.95	11.1	11.41	11.5	11.6	11.7	11.7
Cluster 3													
12.1	12.3	12.5	12.7	12.9	13.1	13.4	13.56	13.6	13.71	13.7	13.9	13.9	14.0
Cluster 4													
2.4	2.6	3.0	3.3	3.5	3.8	3.9	4.27	4.5	4.87	5.0	5.25	5.4	5.57

To statistically assess the differences between clusters, a one-way ANOVA was conducted. The descriptive statistics, including mean**,** standard deviation (SD)**,** and confidence interval (CI), with the ANOVA results, were shown in Table 1(c)**.** The results presented a statistically significant difference in HLS across clusters ($F = 319.41$, $p < 0.001$), validating that the grouping reflected meaningful educational stratifications. Fig. 2 showed these inter-cluster differences using a boxplot during the process. This graphical representation reinforced the statistical significance and provided a view of the variance as well as overlap between clusters.

TABLE 3
DESCRIPTIVE STATISTICS AND ONE-WAY ANOVA RESULTS ACROSS CLUSTERS

ANOVA	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Mean	7.47	10.90	13.29	4.14
Standard Deviation	0.96	0.65	0.63	1.03
Confidence Interval	0.50	0.34	0.33	0.54
F-Statistic	319.41	319.41	319.41	319.41
p-Value	0.00	0.00	0.00	0.00

Fig. 2 showed the distribution of schooling expectancy across the four clusters. Clusters 2 and 3 presented narrower interquartile ranges, signifying more consistent educational attainment across member districts. Cluster 3 showed the highest median and lowest variance, consistent with its composition of urban and coastal districts, including Kota Jayapura, Biak Numfor, as well as Merauke.

Cluster 4 showed the lowest median and widest spread, consistent with its composition of remote and mountainous districts, including Nduga as well as Pegunungan Bintang, which faced severe educational challenges. Moreover, Clusters 1 and 2 were in between, with Cluster 1 representing underdeveloped yet improving regions. This visual interpretation supported the statistical evidence shown by the ANOVA test and further signified the spatial disparities in educational attainment across Papua.

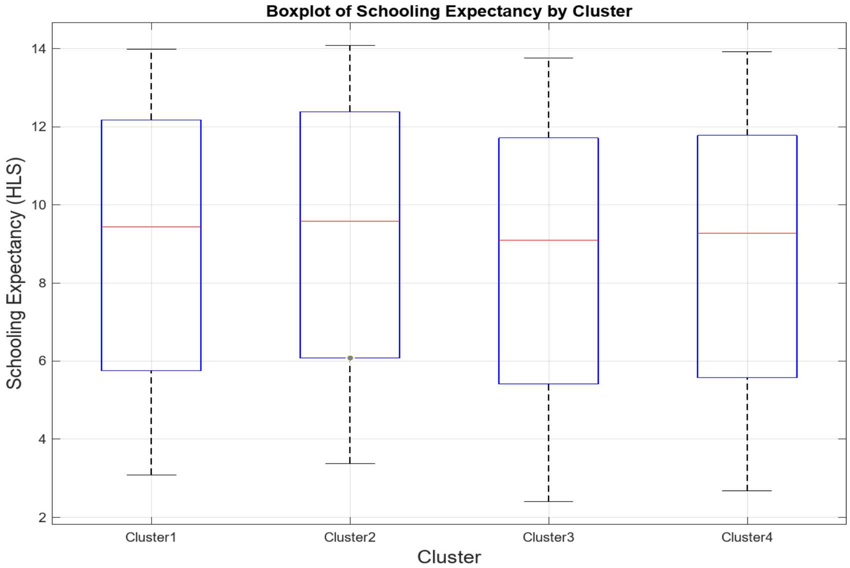


Fig. 2 Boxplot of schooling expectancy across identified clusters

B. Goodness-of-Fit Results for Clustered Schooling Expectancy Distributions

Following clustering process, a goodness-of-fit analysis was conducted to identify which probability distributions best represent the schooling expectancy patterns in each cluster. Several candidate distributions were evaluated using KS and AD tests.

The analysis in Table 2 showed that the Normal distribution provided the best fit for Clusters 1 and 4. Meanwhile, Log-Normal as well as Generalized Extreme Value (GEV) best characterized Clusters 2 and 3, respectively. The results confirmed that the clustered data were statistically consistent with well-established probability model. By reporting only the best-fitting distributions, the focus was shifted from statistical detail to practical interpretability, associating clustering results with subsequent spatial analysis and policy-oriented perceptions.

TABLE 2
HYPOTHESIS TESTING RESULTS FOR CLUSTER 1-4

Cluster	Distribution	Kolmogorov-Smirnov (KS)				Anderson-Darling (AD)			
		H_0/H_1	p-Value	KSStat	CV	H_0/H_1	p-Value	ADStat	CV
1	Normal	H_0	0.9932	0.1052	0.3489	H_0	0.9932	0.1899	2.5061
2	Log-Normal	H_0	0.8955	0.1437	0.3489	H_0	0.9628	0.2630	2.5061
3	(GEV)	H_0	0.9682	0.1223	0.3489	H_0	0.9609	0.2662	2.5061
4	Normal	H_0	0.9763	0.1184	0.3489	H_0	0.9875	0.2106	2.5061

A total of five model selection criteria were used to assess goodness-of-fit across all candidate distributions to strengthen the distribution selection process. The model included AIC, CAIC, BIC, and HQC. Among the model, HQC was selected as the most reliable metric due to its favorable balance between parsimony and explanatory power, particularly for small or moderate sample sizes. Table 3 showed the HQC values for each distribution across clusters.

The results in Table 3 showed that distributions with high p-values in KS and AD tests generally corresponded to lower HQC values, signifying strong model performance. For instance, distributions including Log-Normal and Inverse Gaussian consistently showed both statistical significance in goodness-of-fit as well as minimal HQC scores. This consistency strengthened the credibility of the selected distributions and supported the general reliability of the probabilistic modeling method. The inclusion of HQC-based evaluation ensured the interpretability, accuracy, and generalizability of results across the four educational clusters.

TABLE 3
HQC VALUES OF PROBABILITY DISTRIBUTIONS BY CLUSTER

Distribution	Hannan–Quinn Criterion (HQC)			
	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Gamma	41.8575	30.5348	29.7058	43.9007
Log Normal	42.1422	30.6137	29.8424	44.4731
Logistic	42.5440	31.8499	30.7472	44.6408
Log Logistic	42.9411	31.9080	30.9794	45.3281
Normal	41.5508	30.5162	29.5613	43.4334
GEV	41.5183	30.3931	27.6239	43.4590
Weibull	41.0083	30.5021	28.1518	42.8880
Rician	41.5149	30.4787	29.5239	43.3906
Birnbaum Saunders	42.0894	30.5713	29.8021	44.3718
Extreme Value	41.2616	30.6766	28.1052	43.3357
Inverse Gaussian	42.0916	30.5714	29.8022	44.3932
Nakagami	41.6500	30.4991	29.6103	43.5056
Stable	45.3950	34.3604	33.4055	47.2776
t Location-Scale	43.4542	32.4195	31.4646	45.3367

C. Cluster-Based Distribution and Mixture Model Fitting

This study initially calculated the goodness-of-fit of each candidate distribution using KS test to construct a statistically sound mixture model for each cluster. Table 4 showed the KS statistics for all 14 probability distributions across the four identified clusters. These statistics served as the basis for deriving the weights used in the cluster-wise mixture model, with lower KS values signifying better distributional fit.

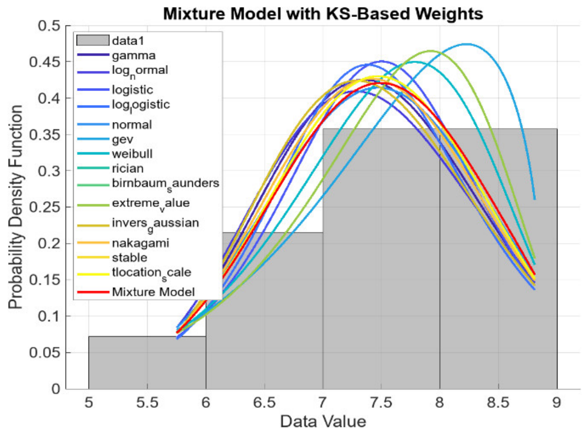
TABLE 4
DISTRIBUTION FIT EVALUATION USING KS GOODNESS-OF-FIT STATISTICS BY CLUSTER

Distribution	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Gamma	0.0720	0.0692	0.0635	0.0688
Log Normal	0.0782	0.0730	0.0653	0.0727
Logistic	0.0746	0.0720	0.0744	0.0745
Log Logistic	0.0798	0.0720	0.0730	0.0764
Normal	0.0767	0.0740	0.0673	0.0764
GEV	0.0568	0.0626	0.0921	0.0648
Weibull	0.0713	0.0815	0.0878	0.0764
Rician	0.0707	0.0696	0.0649	0.0709
Birnbaum Saunders	0.0726	0.0692	0.0631	0.0677
Extreme Value	0.0620	0.0790	0.0885	0.0721
Inverse Gaussian	0.0726	0.0692	0.0661	0.0677
Nakagami	0.0713	0.0696	0.0642	0.0693
Stable	0.0707	0.0696	0.0649	0.0715
t Location-Scale	0.0707	0.0696	0.0649	0.0709

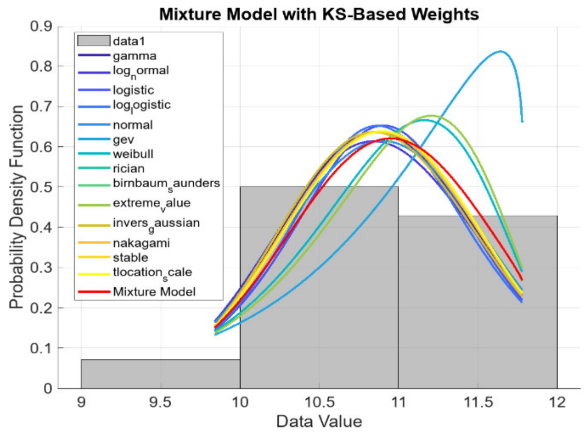
Based on the KS statistics shown in Table 4, weights were assigned to each distribution in a cluster, inversely proportional to individual KS values. These weights were then used to construct the cluster-specific mixture model. Fig. 3 showed the resulting mixture distributions, enabling comparative interpretation of the curve shapes and distributional behaviors across clusters.

Each subplot in Fig. 3 represented the mixture model corresponding to one of the four identified clusters. These curves were generated by combining the PDFs of selected statistical distributions, each weighted based on personal KS goodness-of-fit score. The resulting model captured the empirical distribution of schooling expectancy data in each cluster, offering robust approximations of regional educational disparities.

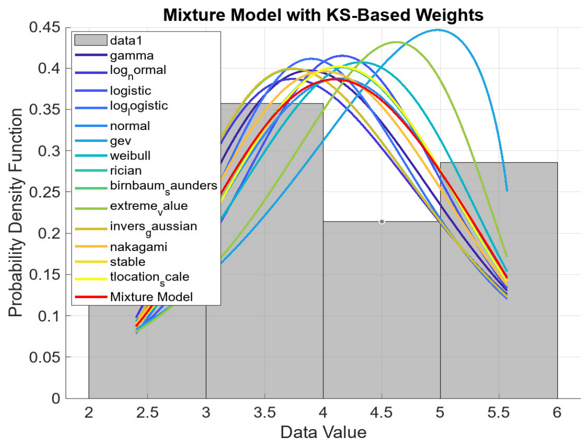
Cluster 1 showed a moderately left-skewed distribution, predominantly shaped by Gamma and Normal components. Consequently, Cluster 2 signified a more symmetric distribution pattern, implying lower intra-cluster disparity. Cluster 3 presented a distinct right-skewed distribution, typical of saturated, high-performing regions with extended HLS values. Meanwhile, Cluster 4, consisting of underdeveloped districts, showed broad and strongly left-skewed distributions with high variance, best fitted by Stable as well as Inverse Gaussian models. These distinct distributional patterns further supported the interpretive validity of the cluster assignments and showed differing educational development dynamics across regions.



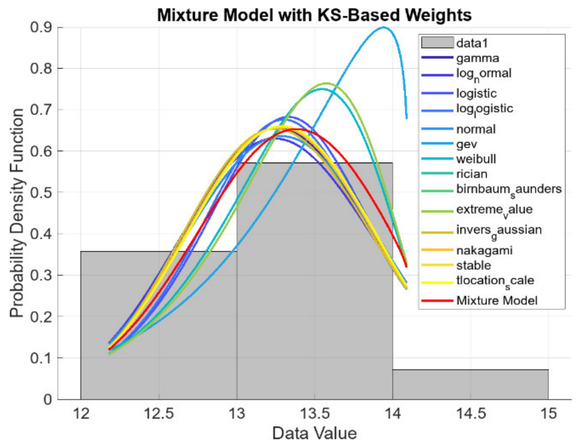
a. KS-Based Mixture Model – Cluster 1



b. KS-Based Mixture Model – Cluster 2



c. KS-Based Mixture Model – Cluster 3



d. KS-Based Mixture Model – Cluster 4

Fig. 3 (a-d) KS-based distribution fitting and mixture modeling of clustered HLS data

D. Wind Rose Visualization of Cluster-Based Educational Distribution

A wind rose diagram was constructed to improve the spatial interpretation of the clustered schooling expectancy distributions, as shown in Fig. 4. This visualization provided an intuitive way to compare districts by the geographic orientation and cluster membership. Each ray in the diagram represented a district, while the color-coded circular lines showed the four cluster groups based on the best-fitting probability distributions. This method offered a compact representation of spatial tendencies in each cluster, facilitating the identification of directional patterns in regional educational outcomes.

The wind rose visualization showed distinct spatial orientations across clusters during the process. Clusters 2 (red) and 3 (green), representing districts with moderate-to-high HLS values, signified a visible concentration along the northern as well as coastal regions of Papua. The position showed stronger accessibility to educational services, better infrastructure, and improved learning outcomes in these regions.

Clusters 1 (blue) and 4 (purple), associated with lower HLS values, showed more diffuse and irregular patterns, predominantly across the central highlands and southern mountainous regions. These regions were characterized by geographic isolation and limited access to public services, reflecting systemic challenges in sustaining equitable education.

The observed angular dispersion in Cluster 4 showed spatial heterogeneity, while the more uniform patterns in Cluster 3 signified relatively balanced educational progress in coastal districts. When combined with the results of Fig. 5, the wind rose further reinforced the interpretation that educational inequality in Papua followed a coastal–highland divide, where geography played a crucial role in shaping opportunities.

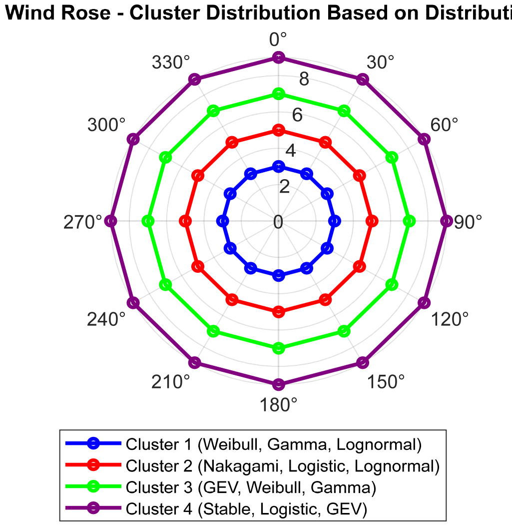


Fig. 4 Spatial clustering of schooling expectancy in Papua Province

The statistical distribution patterns in Fig. 3 reinforced the directional spatial patterns shown in Fig. 4. The combination of statistical, directional, and geographic perspectives across Figs. 3–5 supported a computer-science-based decision-support method to understand educational inequality. From a computer-science perspective, this visualization showed how spatial clustering methods were incorporated into decision-support systems. The tools allowed policymakers to identify priority regions at a glance and to design geographically targeted interventions by focusing infrastructure and teacher allocation programs in highland regions while optimizing governance efficiency in coastal regions.

E. Spatial Stratification of Schooling Expectancy: Coastal–Highland Divide and Its Policy Implications

A spatial map was developed using MATLAB's Mapping Toolbox to provide a comprehensive interpretation of clustering outcomes (Fig. 5). The map visualized the spatial distribution of districts across Papua Province according to the assigned clusters. Each cluster was distinctly color-coded and mapped according to geographic boundaries.

The visualization showed a visible spatial divide in schooling expectancy, particularly between coastal and highland districts. The observed spatial stratification signified that educational development in Papua was not random, but geographically structured, consistent with clustering results.

Fig. 5 showed that coastal and urban districts (mainly in Clusters 2 and 3) tended to achieve higher or moderate HLS values, reflecting stronger accessibility to schools, better infrastructure, and improved governance. Consequently, highland and remote districts (Clusters 1 and 4) signified consistently lower HLS values as well as more irregular spatial patterns. The process implied persistent educational challenges, including isolation, limited resources, and uneven service delivery. The results obtained during the process of this analysis followed previous expert-based studies. For example, [29] showed the systemic failure of education in the highlands of Papua, where extremely low completion rates and teacher absenteeism created a cycle of inequality. A study by [30] explained that despite increased educational investment under special autonomy, policy implementation remained weak in remote Papua, leading to persistent disparities. A government-commissioned investigation [31] further reported very high absenteeism rates among teachers and principals in highland districts, reinforcing the stratification this study identified.

When combined with the directional patterns shown in Fig. 4, the results confirmed that educational inequality in Papua followed a visible coastal–highland divide. From a computer-science perspective, the incorporation of clustering algorithms (Fig. 3), spatial visualization, and expert-validated stratification (Fig. 5) provided a multidimensional perspective that strengthened policy implications. Figs. 3–5 showed that statistical disparities, visual confirmation, and literature-based validation joined to signify the urgent need for geographically targeted interventions. Therefore, the spatial stratification analysis, validated through statistical clustering and supported by existing expert perspectives in the literature, provided a strong foundation for developing policy-oriented information systems to promote equitable as well as geographically differentiated educational development across Papua.

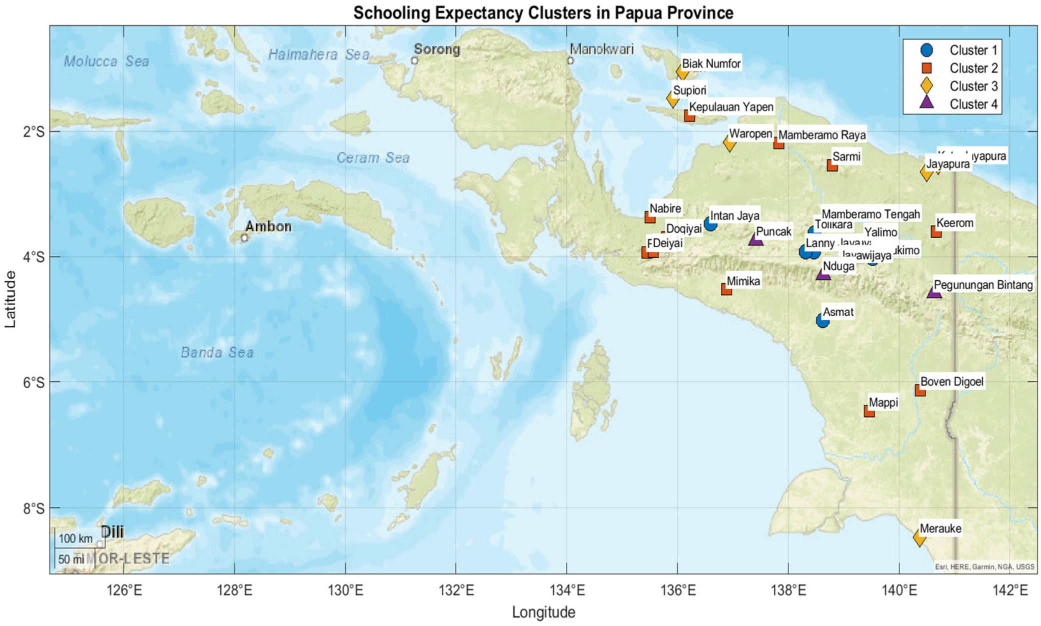


Fig. 5 Spatial clustering of schooling expectancy in Papua Province

IV. Discussion

This section presents an in-depth discussion of the study's main findings by integrating statistical clustering outcomes, probabilistic distributional patterns, and spatial dimensions of Schooling Expectancy (HLS) across Papua Province. The discussion is organized into five sub-sections, each focusing respectively on the interpretation of clustering and educational stratification (A), statistical distributional behaviors and their implications (B), spatial disparities and directional trends (C), regionally differentiated policy recommendations (D), and study limitations with suggestions for future research (E). This analytical structure not only captures the empirical realities of educational inequality but also provides a conceptual and practical foundation for more adaptive and context-sensitive policy interventions.

A. Interpretation of Clustering and Distributional Patterns in Educational Stratification

The Gaussian Mixture Model (GMM) successfully classified the 29 districts in Papua Province into four clusters, each reflecting different stages of educational development. This clustering was validated through the Silhouette Index and Davies–Bouldin Index, confirming that four clusters were optimal. The results highlight a clear educational hierarchy, with Cluster 3 and Cluster 4 representing the two extremes.

Cluster 3, which includes Kota Jayapura, Jayapura (district), and Merauke, shows the highest Schooling Expectancy (HLS) values and relatively stable outcomes, indicating equitable access and stronger governance support. In contrast, Cluster 4 comprises disadvantaged highland districts such as Nduga, Pegunungan Bintang, and Puncak, with the lowest HLS values and wide disparities, reflecting persistent systemic challenges such as isolation, limited resources, and service inequality.

Cluster 2 represents transitional districts (e.g., Mimika, Jayawijaya, and Kepulauan Yapen), showing moderate HLS values and gradual improvement, while Cluster 1 covers lower-middle HLS areas (e.g., Yahukimo and Tolikara) that still face uneven progress. These distinctions suggest that geographic and governance factors remain central in shaping educational outcomes across the province.

Taken together, the clustering outcomes reveal a multidimensional stratification of educational development in Papua. Rather than emphasizing statistical models, the focus here is on how clustering results can support computer-science-oriented applications, such as decision-support systems for policy design, resource allocation, and monitoring educational inequalities (cf. Silhouette Index; Davies–Bouldin Index). This integrated approach provides a strong empirical foundation for identifying hidden disparities and guiding targeted interventions.

B. Spatial Implications: Directional Trends and Geographic Disparities

The integration of spatial visualization techniques, particularly the wind rose diagram (Fig. 4) and the geospatial cluster map (Fig. 5), reveals pronounced directional and locational stratification in schooling expectancy (HLS) outcomes across Papua Province. These visual tools contextualize educational disparities within Papua's geographically complex terrain.

Figure 4 shows that Clusters 2 (red) and 3 (green), associated with moderate to high HLS, are predominantly aligned along the northern and coastal axis, including districts such as Kota Jayapura, Merauke, Biak Numfor, Supiori, Waropen, and Mappi. This pattern suggests that proximity to infrastructure and urban centers strongly enhances educational attainment. In contrast, Clusters 1 (blue) and 4 (purple), dominated by remote highland districts such as Nduga, Pegunungan Bintang, and Puncak, display irregular and diffuse orientations, reflecting systemic barriers such as isolation and limited service delivery.

Figure 5 further emphasizes this coastal–highland divide: Cluster 3, concentrated in coastal and urban districts, demonstrates high HLS values and cohesion, while Cluster 4, in mountainous areas, reflects the lowest outcomes and the greatest internal disparity. Cluster 2 spans transitional districts with gradual progress, whereas Cluster 1 remains scattered with lagging development.

Together, Figures 4 and 5 provide a clear visualization of how geography—particularly remoteness, infrastructure, and urban proximity—shapes educational inequality in Papua. Instead of focusing on detailed statistical models, this section emphasizes how spatial clustering outcomes can support computer-science-oriented decision-support systems. Such integration strengthens the analytical framework and informs targeted interventions, offering a practical basis for addressing regional disparities and guiding policy planning.

C. Policy Implications for Regionally Differentiated Educational Interventions

The clustering results reveal that educational inequality in Papua is spatially embedded and highly context-dependent. A uniform or centralized policy approach risks being ineffective and may worsen disparities. Instead, a regionally differentiated strategy is essential, tailored to the characteristics of each cluster.

Cluster 3 (Kota Jayapura, Jayapura District, Merauke) demonstrates strong educational attainment and equity. Here, policies should focus on quality improvements such as advanced teacher training, curriculum enrichment, and innovation hubs that can serve as models for neighboring districts.

Cluster 2 (e.g., Mimika, Jayawijaya, Kepulauan Yapen) represents transitional areas that require institutional strengthening, teacher retention incentives, and equitable resource allocation to sustain positive educational progress.

Cluster 1 (e.g., Puncak Jaya, Yahukimo, Tolikara) shows stagnation and high variability. Community-based models, temporary learning facilities, transportation support, and targeted scholarship schemes can help improve access for underserved populations.

Cluster 4 (e.g., Nduga, Pegunungan Bintang) faces the most severe challenges due to remoteness. Priority interventions include mobile schools, satellite classrooms, and *flying teacher* programs, supported by cross-sector collaboration in education, health, and infrastructure.

To remain effective, this cluster-based policy framework should be dynamic, supported by longitudinal monitoring. A spatial education dashboard, integrated with decision-support systems, can provide real-time evaluation and planning.

In conclusion, this geographically and statistically informed framework provides a practical basis for regionally adaptive education policies in Papua, enabling targeted resource allocation and reducing structural inequalities.

D. Study Limitations and Recommendations for Future Research

This study builds upon previous work conducted by the same research team, which applied the Gaussian Mixture Model (GMM) to analyze human development disparities in Papua based on educational data from 2010 to 2023 [10]. While both studies employed a similar probabilistic modeling approach, the earlier research examined only four probability distributions—Inverse Gaussian, Rician, Weibull, and Nakagami—and did not incorporate explicit spatial analysis. In contrast, the present study significantly expands the methodological framework by evaluating 14 candidate probability distributions, including less commonly used ones in educational contexts such as Stable and t Location-Scale, and by integrating spatial mapping and directional pattern analysis (wind rose) to capture geographic variations in schooling expectancy.

This broader approach enhances the flexibility in representing the shapes of HLS distributions and enables more precise modeling of educational inequality according to the specific characteristics of each cluster. Consequently, the study contributes both a richer distributional framework and a more contextually grounded spatial interpretation. Future research is encouraged to combine in-depth probabilistic modeling with multivariate spatial frameworks to develop more holistic and policy-relevant analyses of educational disparities across regions.

Second, the use of GMM inherently assumes normality in cluster formation, which may limit its ability to detect patterns in highly skewed or non-elliptical data. Moreover, the mixture modeling in this study remains descriptive and does not account for spatial interactions among districts. For example, Miranti and Mendez [6] demonstrated through spatial econometric modeling (Spatial Durbin Model) that social and economic convergence in Indonesia is significantly influenced by neighboring regions. These findings underscore two methodological limitations of the current study: (1) the restrictive distributional assumptions of GMM, and (2) the absence of spatial effects in the statistical modeling. Future studies are encouraged to apply spatial econometrics to statistically assess inter-regional dependencies in educational development.

Third, the study relies on aggregated data at the district level, which may obscure meaningful intra-district variations. Given Papua's vast geographic size and ethnolinguistic diversity, using higher-resolution data (e.g., sub-district or village level) would yield more detailed insights. Additionally, data bias may arise from underrepresentation in remote regions. Caraka et al. [4] addressed this issue by integrating household survey data (SUSENAS, RISKESDAS) with remote sensing indicators to uncover health disparities in underserved areas. Their approach highlights the importance of utilizing alternative data sources to bridge spatial information gaps.

Fourth, the current study does not incorporate broader development indicators such as digital infrastructure, income levels, or labor market participation. In this context, Kartiasih et al. [5] found that digital inequality across Indonesia is closely tied to educational attainment and socioeconomic characteristics. Their multivariate spatial framework underscores the importance of integrating intersectoral variables in modeling regional disparities. Future research should therefore consider combining education data with digital and economic indicators within a unified spatial-analytic framework.

Finally, although spatial maps and wind rose diagrams in this study improve interpretability, they remain descriptive in nature. Future investigations should adopt spatial regression models or spatial panel techniques to rigorously assess geographic spillover effects on educational outcomes.

In summary, while this study provides a foundational spatial-statistical mapping of educational inequality in Papua, it also opens multiple avenues for methodological refinement, expanded indicator integration, and cross-sectoral development modeling.

V. CONCLUSIONS

This study developed an integrative spatial-statistical framework to analyze regional disparities in schooling expectancy (HLS) across 29 districts in Papua Province between 2010 and 2023. By combining Gaussian Mixture Model (GMM) clustering, probabilistic distribution modeling, and directional spatial visualization, the approach captured complex educational patterns that traditional descriptive methods often overlook. The identification of four distinct clusters—each with unique statistical and geographic profiles—reveals a pronounced coastal–highland divide in educational access and outcomes.

Methodologically, the use of GMM enabled soft clustering with probabilistic assignments, accommodating overlapping group boundaries common in socio-geographic data. Mixture modeling allowed for a better fit to the varied distributional shapes observed across clusters, while wind rose diagrams introduced a novel way to interpret the directional spread of educational inequality. Together, these components addressed the research objective by offering a replicable model that is both statistically rigorous and spatially interpretable.

The findings carry significant policy relevance. High-performing clusters require targeted innovations and quality enhancement strategies, while transitional zones need support through institutional strengthening and equitable resource allocation. Lagging districts demand basic infrastructure expansion and teacher deployment, and the most marginalized highland areas necessitate urgent interventions including mobile education units, integrated social services, and flexible funding mechanisms such as special autonomy grants.

Looking ahead, this framework can be further extended through spatial econometric models to capture inter-regional dependencies, and enhanced with village-level data for finer-grained analysis. Future research should also incorporate cross-sectoral indicators—such as income inequality, digital access, and health infrastructure—to develop a more comprehensive diagnostic tool for educational equity planning. Overall, this study contributes both theoretically and practically by providing a transferable methodology for addressing educational inequality in geographically diverse and data-limited contexts.

Author Contributions: *Jonathan Kiwasi Wororomi*: Conceptualization, Methodology, Writing - Original Draft, Writing - Review & Editing, Supervision. *Felix Reba*: Software, Investigation, Data Curation, Writing - Original Draft. *Frans A. Asmuruf*: Investigation, Data Curation.

Funding: This research was financially supported by the Institute for Research and Community Service (LPPM) of Cenderawasih University through the 2025 PNBP BLU Flagship Research Grant. This grant played a vital role in facilitating the successful completion of this study.

Acknowledgments: The authors gratefully acknowledge the Institute for Research and Community Service (LPPM) of Cenderawasih University for providing the 2025 PNBP BLU Flagship Research Grant. Their support significantly contributed to the research process and outcomes.

Conflicts of Interest: The authors declare no conflict of interest. No other potential conflict of interest relevant to this article was reported.

Data Availability: The schooling expectancy dataset used in this study was obtained from Badan Pusat Statistik Papua (https://papua.bps.go.id) and is available upon reasonable request from the corresponding author due to regional data-sharing policies.

Informed Consent: There were no human subjects.

Institutional Review Board Statement: Not applicable.

Animal Subjects: There were no animal subjects. This study did not involve animal testing.

ORCID:
Jonathan Kiwasi Wororomi: https://orcid.org/0000-0003-3879-3731
Felix Reba: https://orcid.org/0000-0001-8798-2366
Frans A. Asmuruf: https://orcid.org/0000-0002-3234-134

REFERENCES

[1] R. Read and A. Benavot, “Global education monitoring report,” in *International Encyclopedia of Education: 4th ed.*, 2022. doi: 10.1016/B978-0-12-818630-5.01026-5.
[2] R. Karagiannis and G. Karagiannis, “Constructing composite indicators with Shannon entropy: The case of human development index,” *Socio-Economic Planning Sciences*, vol. 70, p. 100701, Jun. 2020, doi: 10.1016/j.seps.2019.03.007.
[3] C. Türe and Y. Türe, “A model for the sustainability assessment based on the human development index in districts of megacity Istanbul (Turkey),” *Environment, Development and Sustainability*, vol. 23, pp. 13966–13990, Aug. 2021, doi: 10.1007/s10668-020-00735-9.
[4] R. E. Caraka *et al.*, “Understanding pediatric health trends in Papua: Insights from SUSENAS, RISKESDAS, remote sensing, and its relevance to Prabowo and Gibran’s free lunch and milk program,” *IEEE Access*, vol. 12, pp. 33977–33989, 2024, doi: 10.1109/ACCESS.2024.3380018.
[5] F. Kartiasih, N. Djalal Nachrowi, I. D. G. K. Wisana, and D. Handayani, “Inequalities of Indonesia’s regional digital development and its association with socioeconomic characteristics: a spatial and multivariate analysis,” *Information Technology for Development*, vol. 29, no. 1, pp. 1–24, 2023, doi: 10.1080/02681102.2022.2110556.
[6] R. C. Miranti and C. Mendez, “Social and economic convergence across districts in indonesia: A spatial econometric approach,” *Bulletin of Indonesian Economic Studies*, vol. 59, no. 1, pp. 69–93, 2023, doi: 10.1080/00074918.2022.2071415.
[7] P. A. Widyastaman and D. Hartono, “Geographic distribution of economic inequality and crime in Indonesia: Exploratory spatial data analysis and spatial econometrics approach,” *Applied Spatial Analysis and Policy*, 2024, doi: 10.1007/s12061-023-09556-5.
[8] T. Alqurashi and W. Wang, “Clustering ensemble method,” *International Journal of Machine Learning and Cybernetics*, vol. 10, pp. 1225–1246, 2019, doi: 10.1007/s13042-017-0756-7.
[9] D. Hooshyar, Y. Yang, M. Pedaste, and Y. M. Huang, “Clustering algorithms in an educational context: An automatic comparative approach,” *IEEE Access*, vol. 8, pp. 146407–146417, 2020, doi: 10.1109/ACCESS.2020.3014948.
[10] A. Sroyer, H. Morin, F. Reba, J. Wororomi, and A. Languwuyo, “Clustering and mixture distribution analysis of average years of schooling in Papua (2010 – 2023),” vol. 10, no. 2, pp. 495–506, 2025, doi: 10.18860/cauchy.v10i2.32988.
[11] V. A. Sari and S. Tiwari, “The geography of human capital: Insights from the subnational human capital index in Indonesia,” *Social Indicators Research*, vol. 170, pp. 673–692, 2024, doi: 10.1007/s11205-024-03322-x.
[12] P. A. Widyastaman and D. Hartono, “Economic inequality decomposition and spatial pattern of crime in Indonesia,” *Papers in Applied Geography*, vol. 8, no. 1, pp. 22–37, 2022, doi: 10.1080/23754931.2021.1991842.
[13] H. Kurniawan, H. L. F. de Groot, and P. Mulder, “Are poor provinces catching-up the rich provinces in Indonesia?,” *Regional Science Policy and Practice*, vol. 11, no. 2, pp. 341–356, 2019, doi: 10.1111/rsp3.12160.
[14] F. Astika Saputra, A. Barakbah, and P. Riza Rokhmawati, “Data Analytics of human development index (HDI) with features descriptive and predictive mining,” in *IES 2020 - International Electronics Symposium: The Role of Autonomous and Intelligent Systems for Human Life and Comfort*, Sep. 2020, pp. 126–131. doi: 10.1109/IES50839.2020.9231661.
[15] S. A. Takyi, O. Amponsah, M. O. Asibey, and R. A. Ayambire, “An overview of Ghana’s educational system and its implication for educational equity,” *International Journal of Leadership in Education*, vol. 24, no. 2, pp. 150–168, 2021, doi: 10.1080/13603124.2019.1613565.

[16] X. Liu and J. LeSage, "Arc_Mat: A matlab-based spatial data analysis toolbox," *Journal of Geographical Systems*, vol. 12, no. 4, pp. 357–385, Dec. 2010, doi: 10.1007/s10109-009-0096-6.
[17] J. Kim, Changho; Do, Thang N.; Kim, "Spatially explicit supply chain for nationwide CO_2-to-fuel infrastructure: Data-driven optimization with gaussian mixture model-based region screening and clustering," *Journal of Cleaner Production*, vol. 445, p. 141313, 2024, doi: 10.1016/j.jclepro.2024.141313.
[18] G. Barzizza, E.; Biasetton, N.; Ceccato, R.; Maculan, N.; Masiero, A.; Tardivo, "Evaluating optimal number of clusters: A permutation-based approach," *SSRN Electronic Journal*, 2023, doi: 10.2139/ssrn.5028456.
[19] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 2, pp. 224–227, 1979, doi: 10.1109/TPAMI.1979.4766909.
[20] Y. Huang, P. Sripathanallur Murali, and G. Vejarano, "Distributed monitoring of moving thermal targets using unmanned aerial vehicles and gaussian mixture models," *Robotics*, vol. 14, no. 7, p. 85, 2025, doi: 10.3390/robotics14070085.
[21] L. P. da Silva, E. J. P. Santiago, F. Gomes-Silva, A. S. A. da Silva, and R. S. C. Menezes, "Mixture models of probability distributions applied to rainfall in the state of Pernambuco, Brazil," *Acta Scientiarum - Technology*, vol. 45, e60621, 2023, doi: 10.4025/actascitechnol.v45i1.60621.
[22] J. Saarenpää, M. Kolehmainen, and H. Niska, "Geodemographic analysis and estimation of early plug-in hybrid electric vehicle adoption," *Applied Energy*, 2013, doi: 10.1016/j.apenergy.2013.02.066.
[23] V. Kumar, Shanu, and Jahangeer, "Statistical distribution of rainfall in Uttarakhand, India," *Applied Water Science*, vol. 7, no. 8, pp. 4765–4776, 2017, doi: 10.1007/s13201-017-0586-5.
[24] N. Balakrishnan, E. Chimitova, and M. Vedernikova, "An empirical analysis of some nonparametric goodness-of-fit tests for censored data," *Communications in Statistics: Simulation and Computation*, vol. 44, no. 4, pp. 1101–1115, 2015, doi: 10.1080/03610918.2013.796982.
[25] R. A. Fox and P. G. Hoel, "Introduction to mathematical statistics.," *The Statistician*, 1964, doi: 10.2307/2987351.
[26] I. Pobočíková, M. Michalková, Z. Sedliačková, and D. Jurášová, "Modelling the wind speed using exponentiated weibull distribution: case study of Poprad-Tatry, Slovakia," *Applied Sciences (Switzerland)*, vol. 13, no. 6, p. 4031, Mar. 2023, doi: 10.3390/app13064031.
[27] R. J. Rossi, *Mathematical statistics: An introduction to likelihood based inference*. 2018. doi: 10.1002/9781118771075.
[28] A. Alzaatreh, C. Lee, and F. Famoye, "A new method for generating families of continuous distributions," *Metron*, vol. 71, pp. 63–79, 2013, doi: 10.1007/s40300-013-0007-y.
[29] B. Anderson, "The failure of education in Papua's highlands," Inside Indonesia, 2013. [Online]. Available: https://www.insideindonesia.org/the-failure-of-education-in-papua-s-highlands.
[30] J. A. Mollet, "Educational investment in conflict areas of Indonesia: The case of West Papua Province," International Education Journal, vol. 8, no. 2, pp. 155–166, 2007..
[31] Education Sector Analytical and Capacity Development Partnership (ACDP), "Rural and remote area education strategic planning study for Tanah Papua," Jakarta, Indonesia, 2014. [Online]. Available: http://www.acdp-indonesia.org/wp-content/uploads/2014/08/ACDP039-RRA-Strategic-Planning-Study-for-Tanah-Papua-2014.pdf

Publisher's Note: Publisher stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.