BAREKENG: Journal of Mathematics and Its Applications
December 2026 Volume 20 Issue 4 Page 3355-3370
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss4pp3355-3370

REDUCING BIAS IN PREDICTING POVERTY PERCENTAGE AT THE VILLAGE LEVEL IN BENGKULU CITY USING SMALL AREA ESTIMATION

Etis Sunandi[1*], Pepi Novianti[2], Nurul Hidayati[3], Anggun Permatasari[4], Alya Saputri[5], Lutfiah Firlian[6], Indah Srilita Asmuyana[7]

[1,2,3,4,5,6,7]*Statistics Study Program, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Bengkulu*
Jln. WR. Supratman Kandang Limun, Kec. Muara Bangkahulu, Bengkulu, 38122, Indonesia

*Corresponding author's e-mail: * esunandi@unib.ac.id*

Article Info

Article History:
Received: 17th October 2025
Revised: 12th May 2026
Accepted: 07th June 2026
Available online: 24th August 2026

Keywords:
Bias;
MSE;
Poverty;
SAE-BB-APHL.

ABSTRACT

Analysis of poverty rates at the small area level, especially at the village level, often faces obstacles due to limited sample sizes. As a result, the use of direct estimates has large variances, making them less reliable. One solution to this problem is to use the Small Area Estimation (SAE) method. This study aims to apply the Beta-Binomial Small Area Estimation (SAE-BB-APHL) model by utilizing a second-order Laplace approach to estimate fixed effects that were fixed in the previous model, namely SAE-BB-HL. The SAE-BB-APHL model was applied to the case of poverty rates in each village in Bengkulu City. The results show a decrease in Mean Squared Error (MSE) in estimating the poverty rate when using SAE-BB-APHL compared to direct estimation and SAE-BB-HL. The MSE values produced by the direct estimation method in several villages were 0.20638 in village 1, 0.04865 in village 2, and 0.04865 in village 46. Meanwhile, the SAE-BB-APHL method successfully reduced the MSE to 0.00524 in village 1, 0.00493 in village 2, and 0.00630 village 46. In addition, the MSE in SAE-BB-APHL is also smaller than SAE-BB-HL as evidenced in Table 3. This decrease in MSE indicates that bias is reduced by using the SAE-BB-APHL method. Therefore, this approach can be a superior alternative in poverty analysis at the village level. More accurate estimation results are expected to assist the government in designing more targeted poverty alleviation policies.

How to cite this article:

E. Sunandi, P. Novianti, N. Hidayati, A. Permatasari, A. Saputri, L. Firlian and I. S. Asmuyana., "REDUCING BIAS IN PREDICTING POVERTY PERCENTAGE AT THE VILLAGE LEVEL IN BENGKULU CITY USING SMALL AREA ESTIMATION", *BAREKENG: J. Math. & App.*, vol. 20, no. 4, pp. 3355-3370, Dec, 2026.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

To estimate population parameters, such as districts, cities, subdistricts, or villages, various surveys usually apply a direct estimation approach. Accurate direct estimation requires a sufficiently large sample size. However, when the sample size is inadequate or even non-existent, this approach provides inaccurate estimates and produces high variance. Therefore, small area estimation (SAE) was developed to overcome this problem [1].

Small Area Estimation (SAE) is a statistical method for estimating parameters in a subpopulation with a small number of examples or even none at all, known as a “small area.” SAE works by increasing the accuracy of prediction results through indirect estimation by “borrowing strength” from explanatory information and using the relationship between explanatory variables and the variables being studied [2]. SAE model-based estimation has received much attention from statisticians. Based on the type of data available, small area models are classified into two types, namely area level models and unit level models.

The SAE model can be applied to data with continuous or discrete response variables, such as binary data. The SAE model for binary data is prone to overdispersion. According to [3], overdispersion can be overcome by using a mixed distribution, namely beta-binomial. Rao and Molina [2] have developed a beta-binomial SAE model. The proposed beta-binomial model is a covariate-free model with parameter estimation through a Bayesian approach. Meanwhile, according to [4], parameter estimation in the beta-binomial model can be done through Hierarchical Likelihood (HL). This method is claimed to be better than the Bayesian approach in terms of analytics and computation.

However, SAE-BB-HL still has bias in estimating fixed effects (β) and random effects (ν). This bias in estimation results in inaccurate estimation results, thereby reducing the accuracy and quality of the modeling results. Statistically, this condition is reflected in a high Mean Squared Error (MSE) value, indicating that the estimated results are no longer representative of the actual population parameter values. When these biased estimates are applied in real-world situations, the flawed data will mislead stakeholders and lead to critical errors in policy-making. Therefore, it is necessary to develop a method for reducing bias in the estimation of fixed effects (β) and random effects (ν) [5]. The development of a bias reduction method in SAE-BB-HL uses a second-order Laplace approximation to estimate fixed effects. This method is called SAE-BB-APHL [6]. The SAE-BB-APHL method can be implemented in various disciplines, one of which is in the socio-economic field, namely poverty.

Poverty is a condition in which an individual is unable to meet their own basic needs. The percentage of the poor population (Household Head Index) is an indicator that measures the proportion of the population with expenditures below the poverty line relative to the total population in a given region. The Central Statistics Agency (BPS) defines the poverty line as the minimum expenditure required to meet basic food and non-food needs [7]. In Indonesia, poverty remains one of the country’s most significant challenges, despite having experienced substantial economic growth over the past few decades. Although poverty rates have gradually decreased, millions of people still live below the poverty line, particularly in remote areas [8]. Globally, the urgency of the poverty issue is evident in the fact that poverty is one of the main targets in the Sustainable Development Goals (SDGs) agenda, namely the eradication of poverty [9].

Analysis of the poverty rate at the small regional level, particularly at the sub-district or village level, often faces obstacles due to limited sample sizes. As a result, the use of direct estimators produces large variances, making the estimation results less reliable. The solution to this is to use the Beta-Binomial SAE model with parameter estimation through Hierarchical Likelihood (HL). In previous studies [6] and [10], the Beta-Binomial SAE model with parameter estimation via Hierarchical Likelihood (HL) was applied; simulation results showed that the estimates of the fixed parameters, as well as the overdispersion parameter, were biased. Therefore, this study aims to develop a method to reduce bias in the SAE-BB-HL model by using a second-order Laplace approximation approach to estimate fixed effects, thereby improving the accuracy of the estimation results. The SAE-BB-APHL method was applied to the case of poverty rates in each subdistrict or village in Bengkulu City in 2022. More accurate estimation results are expected to assist the government in designing more targeted poverty alleviation policies.

2. RESEARCH METHODS

2.1 Direct Estimation

Direct estimation methods are statistical techniques used to estimate population parameters directly from survey data. These methods are particularly useful in survey sampling where the goal is to obtain reliable estimates for a specific region or subpopulation. This method is a classic approach to estimating a population based on a sample design. A direct estimator for an area uses only variables obtained from the sample units of that area. Direct estimation consists of several methods, one of which is simple random sampling (SRS), used in this study, which has equal weights because the probability of sample selection in each area unit is the same. The direct estimator for the SRS method is obtained using the following formula [2]:

$$\hat{Y}_i^{SRS} = \frac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij} \tag{1}$$

where $\hat{Y}_i^{SRS}$ is the i-th direct estimate, n_i is the i-th observation, and Y_{ij} are the i-th and j-th observations for $i = 1, 2, \dots$ and $j = 1, 2, \dots, n_i$. Because this study follows a binomial distribution and the predicted results are expressed as proportions, Eq. (1) is modified as follows:

$$p_i = \frac{y_i}{n_i} \tag{2}$$

where p_i is the proportion of the i-th area, y_i is the sum of the responses for area i-th, and n_i is the sum of the observations for area i-th. The SRS direct estimator for an infinite (unknown) population can be estimated using the formula:

$$MSE\left(\hat{Y}_i^{SRS}\right) = \frac{s_i^2}{n_i}, s_i^2 = \frac{\sum_{j=1}^{n_i}\left(Y_{ij} - \hat{Y}_i^{SRS}\right)^2}{n_i - 1} \tag{3}$$

where s_i^2 is the sample variance of the direct estimator SRS. Eq. (3) is then modified as follows:

$$MSE(p_i) = \frac{s_i^2(p_i)}{n_i}, s_i^2(p_i) = p_i * (1 - p_i) \tag{4}$$

where $s_i^2(\hat{p}_i)$ is the variance of the binomial distribution.

2.2 Small Area Estimation (SAE)

A small area is a subset of the population whose sample size is small with a variable of concern. Parameter estimation in a small area can be done by direct or indirect estimation. Direct estimation is an estimation based on sample data from the area. The result of direct estimation in a small area is an unbiased estimator but has a significant variance. This is because the estimated parameters are obtained from small samples. Therefore, the indirect estimation method is more appropriate for utilizing the strength of the surrounding area and data sources inside or outside the area whose statistics are to be obtained [2]. A small area estimation model has been developed by several researchers [10], [11], [12].

Based on the basic model, the small area model consists of two types: based on the area level and the unit level. The model used as the basis for developing the proposed model is the SAE Level Area Fay-Herriot model (1979), which is defined as follows: $y_i = \boldsymbol{x_i'\beta_x} + v_i + e_i$, for $i = 1, \dots, m$. The random variable area v_i is assumed to be normally distributed with zero expected value and variance σ_v^2. Meanwhile, the observation error is assumed to be normally distributed with zero expectation value and known variance σ_e^2. Then the variables v_i, e_i, and y_i are assumed to be independent. The next model is the Beta-Binomial Model [13]. The definition of this model is $y_i|p_i \sim Binomial\ (n_i, p_i)$ and $p_i|v_i \sim Beta(\alpha, \gamma)$ where $\alpha > 0, \gamma > 0$ are known.

The proposed model is a development of the compilation of the Fay-Herriot model and the Beta-Binomial Model, which is defined as follows [10]:

$$y_i|p_i \sim Binomial\ (n_i, p_i), i = 1, 2, \dots, m$$

$$p_i|v_i, \gamma \sim iid\ Beta(\alpha, \gamma), \alpha > 0, \gamma > 0$$

$$\log\left(\frac{p_i}{1 - p_i}\right) = \xi_i = \boldsymbol{x_i'\beta_x} + v_i \tag{5}$$

where y_i is the response variable of the i-th area, the parameter p_i is the proportion of the i-area. At the same time, v_i is an i-area random variable assuming $v_i \sim N(0, \sigma_v^2)$. The covariate $\boldsymbol{x}_i$ of size $p \times m$ is a constant covariate. The parameter $\boldsymbol{\beta}_x$ is a fixed effect parameter of size $p \times 1$. Then v_i and $y_i|p_i$ are assumed to be independent. Parameters γ are assumed to be fixed.

2.3 Hierarchical Likelihood (HL) Method

Estimating the parameters of the proposed model was done by maximizing the HL function and the Adjusted Profile Hierarchical Likelihood (APHL) function. GLMM is extended to HGLM by accommodating non-normal distributions for random effects [14]. They use HL to avoid complex integration. This method uses the logarithm of the probability density function with two variables (y, v) [10]. HL is a probability function with Henderson if both distributions are normal. If one or both distributions are not normal, the HL is a generalization of the joint probability. The function is determined by [6]. The HL formula can be written as follows:

$$h = h_0 + h_1 = log l_1(\beta, \phi; y|v) + log l_2(\lambda; v) \tag{6}$$

where $l_1(\beta, \phi; y|v)$ and $l_2(\lambda; v)$ are conditional probability functions $y|v$ and probability functions v. Vector $\boldsymbol{\beta}$ is a canonical parameter, (ϕ, λ) denotes the dispersion parameter, and λ is the distribution parameter v. Let $v(.)$ be the corresponding link function that defines HL such that $v = v(u)$. In the HL form, the selection ofthe random effect scale is essential in that the scale change requires Jacobian adjustment. Suppose that $v = v(u)$ is used to show the scale in which the random effect of v is assumed to be linear, and the predictor is linear. Using the score function of HL, the parameter v is estimated as follows [5]:

$$\frac{\partial h}{\partial \boldsymbol{\beta}} = \boldsymbol{0}, \frac{\partial h}{\partial \boldsymbol{v}} = \boldsymbol{0} \tag{7}$$

Therefore, the estimator $(\boldsymbol{\beta_x}, \boldsymbol{v})$ obtained through iterations of the Fisher Scoring algorithm as follows:

$$\begin{pmatrix} \boldsymbol{\beta_x} \\ \boldsymbol{v} \end{pmatrix}^{(r+1)} = \begin{pmatrix} \boldsymbol{\beta_z} \\ \boldsymbol{v} \end{pmatrix}^{(r)} + \begin{bmatrix} \boldsymbol{X'SUSX} & \boldsymbol{X'SUS} \\ \boldsymbol{SUSX} & \boldsymbol{WS + D^{-1}} \end{bmatrix}^{-1} \begin{pmatrix} \boldsymbol{u_{\beta_z}} \\ \boldsymbol{u_v} \end{pmatrix} \tag{8}$$

Where $\boldsymbol{S} = diag[p_i(1-p_i)]$, $\boldsymbol{WS} = diag(p_i(1-p_i)[u_i p_i(1-p_i) - \zeta_i(1-2p_i)])$, $\boldsymbol{D} = \sigma_v^2 \boldsymbol{I}$, and $\boldsymbol{U} = \begin{bmatrix} u_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & u_m \end{bmatrix}$.

Furthermore, the estimated dispersion parameter $\hat{\tau} = (\hat{\phi}, \hat{\sigma}_v^2)$ is the maximization solution of the APHL function APHL $p_{v,\boldsymbol{\beta}_x}(h)$ which is defined as follows:

$$\boldsymbol{p_{\beta,v}(h)} = \left(\boldsymbol{h} + \frac{1}{2}\log(2\pi \boldsymbol{H}^{-1})\right)\Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}, \boldsymbol{v}=\hat{\boldsymbol{v}}} \tag{9}$$

Where $\boldsymbol{H} = \begin{bmatrix} \boldsymbol{X'W_1X} & \boldsymbol{X'W_1Z} \\ \boldsymbol{Z'W_1X} & \boldsymbol{Z'W_1Z + W_2} \end{bmatrix}$, $\boldsymbol{W_1} = \left(\frac{\partial \mu}{\partial \eta}\right)^2 (\phi V(\mu))^{-1}$, and $\boldsymbol{W_2} = -\frac{\partial^2 h_1}{\partial v^2}$

In this study, predictions of the estimated proportion of the proposed model will also be carried out through the formula bellows [6]:

$$\hat{p}_i^{HL} = \frac{\exp(\boldsymbol{x}_i' \boldsymbol{\beta}_x + \boldsymbol{v}_i)}{1 + \exp(\boldsymbol{x}_i' \boldsymbol{\beta}_x + \boldsymbol{v}_i)} \tag{10}$$

for $i = 1, \dots, m$. the i-th MSE estimator of $\hat{p}_i^{HL}$ as follows:

$$mse_i(\hat{p}_i^{HL}) = \frac{1}{B}\sum_{b=1}^{B}(p_i - \hat{p}_i^{HL})^2 \tag{11}$$

2.4 Development of SAE-BB-APHL Model

The SAE-BB-APHL model is an improved version of the SAE-BB-HL model for estimating biased fixed effects parameters. The SAE-BB-HL model was developed in [9]. This is consistent with the opinions of previous researchers. For binary data analysis, joint maximization$(\boldsymbol{v}, \boldsymbol{\beta_x})$ of the biased $\boldsymbol{h}$ function is used to estimate $\boldsymbol{\beta_x}$. This problem can be overcome by using the APHL function $\boldsymbol{p_v(h)}$ to estimate $\boldsymbol{\beta_z}$ [5].

The function of $p_v(h)$ useful for obtaining APHL by eliminating random influences v. Scale $p_v(h) \approx \ell$ in the first order. Using the Laplace approach, we obtain [4]:

$$p_v(h) = \left(h - \frac{1}{2} logdet(\boldsymbol{W_1} + \boldsymbol{W_2})/(2\pi)\right)\Big|_{v=\hat{v}}, \tag{12}$$

For example,

$$h_b = h - \frac{1}{2} logdet(\boldsymbol{W_1} + \boldsymbol{W_2})/(2\pi)) \tag{13}$$

$$h_b = h - \frac{1}{2} logdet(\boldsymbol{W_1} + \boldsymbol{W_2})/(2\pi))$$
$$= \sum_{i=1}^{m} \left[\log\left\{\binom{n_i}{y_i}\right\} + \sum_{k=0}^{y_i-1} \log(p_i + k\phi) + \sum_{k=0}^{n_i-y_i-1} \log(1 - p_i + k\phi) - \sum_{k=0}^{n_i-1} \log(1 + k\phi)\right] + \log\left([2\pi\sigma_v^2]^{-\frac{1}{2}}\right) - \frac{\boldsymbol{v}'\boldsymbol{v}}{2\sigma_v^2} - \frac{1}{2}\log\left(\frac{|\boldsymbol{WS}+\boldsymbol{D}^{-1}|}{2\pi}\right) \tag{14}$$

Estimator $\boldsymbol{\beta}_x$ obtained by maximizing the function h_b. Note the following equation:

$$\frac{\partial h_b}{\boldsymbol{\beta}_x} = \frac{\partial}{\partial \boldsymbol{\beta}_x}\left(\sum_{i=1}^{m} \left[\log\left\{\binom{n_i}{y_i}\right\} + \sum_{k=0}^{y_i-1} \log(p_i + k\phi) + \sum_{k=0}^{n_i-y_i-1} \log(1 - p_i + k\phi) - \sum_{k=0}^{n_i-1} \log(1 + k\phi)\right]\right) - \frac{\partial}{\partial \boldsymbol{\beta}_x}\left(\frac{1}{2}\log\left(\frac{|\boldsymbol{WS}+\boldsymbol{D}^{-1}|}{2\pi}\right)\right) \tag{15}$$

$$\frac{\partial h_b}{\partial \boldsymbol{\beta}_x} = \boldsymbol{\zeta}'\boldsymbol{SX} - \frac{1}{2}\,\boldsymbol{e}' \tag{16}$$

Estimator $\boldsymbol{\beta}_x$ is obtained by maximizing the function $\frac{\partial h_b}{\boldsymbol{\beta}_x}$ which is not closed form. Meanwhile, the random effect estimation still uses $\boldsymbol{v}$ in HL. Therefore, iterative estimation using the Delta algorithm is necessary. Next, the second derivative of the function h_b with respect $\boldsymbol{\beta}_x$ is determined.

$$\frac{\partial^2 h_b}{\boldsymbol{\beta}'_x\boldsymbol{\beta}_x} = \frac{\partial}{\partial \boldsymbol{\beta}'_x}\left\{\boldsymbol{\zeta}'\boldsymbol{SX} - \frac{1}{2}\,\boldsymbol{e}'\right\} \tag{17}$$

$$\frac{\partial^2 h_b}{\boldsymbol{\beta}'_x\boldsymbol{\beta}_x} = \boldsymbol{X}'\boldsymbol{SUSX} - \frac{1}{2}\,\boldsymbol{Q} \tag{18}$$

where

$$\boldsymbol{Q} = \begin{bmatrix} tr(\boldsymbol{M}_1) & \cdots & \frac{\partial\left(tr\left[(\boldsymbol{WS}+\boldsymbol{D}^{-1})^{-1}\boldsymbol{J}_p\right]\right)}{\beta_{x1}} \\ \vdots & \ddots & \vdots \\ \frac{\partial\left(tr\left[(\boldsymbol{WS}+\boldsymbol{D}^{-1})^{-1}\boldsymbol{J}_1\right]\right)}{\beta_{xp}} & \cdots & tr(\boldsymbol{M}_p) \end{bmatrix} \tag{19}$$

Thus, the estimator $\boldsymbol{\beta}_x$ use the Fisher Scoring algorithm is calculated using the following equation:

$$\hat{\boldsymbol{\beta}}_x^{(k+1)} = \hat{\boldsymbol{\beta}}_x^{(k)} + \left[E\left\{-\frac{\partial^2 h_b}{x\boldsymbol{\beta}'_x}\right\}\right]^{-1}\left(\frac{\partial h_b}{\boldsymbol{\beta}_x}\right)' \tag{20}$$

MSE is used as a model evaluation metric in SAE-BB-APHL using the following formula:

$$mse_i(\hat{p}_i^{APHL}) = \frac{1}{B}\sum_{b=1}^{B}\left(p_i - \hat{p}_i^{APHL}\right)^2 \tag{21}$$

for $i = 1, \ldots, m, B = 500$, and $\hat{p}_i^{APHL}$ is the MSE SAE-BB-APHL estimator for the i-th region.

2.5 Poverty Rates

Poverty is a condition that occurs as a result of the inability to meet basic needs such as food, clothing, education, health, and housing or shelter. Poverty can also be caused by a scarcity of the means to meet basic needs and the complexity of accessing education and employment [15].

Ability to meet basic needs. With this approach, poverty is viewed as an inability from an economic perspective to meet basic food and non-food needs as measured from an expenditure perspective. The method used is to calculate the Poverty Line (GK), which consists of two components, namely the Food Poverty Line

(GKM) and the Non-Food Poverty Line (GKNM). The calculation of the Poverty Line is carried out separately for urban and rural areas [16], and [17].

2.6 Auxiliary Variables

SKTM (Certificate of Indigency) Recipients SKTM is a letter from the village/sub-district organization to the poor community, which is then taken to the sub-district office to be stamped so that it can be used. This SKTM allows the community to receive free medical treatment and care at local hospitals and SKTM services. In addition, SKTM can also be used to reduce education costs and other needs that require SKTM [18].

Health facilities are places or institutions that provide various health services, such as promotive, preventive, curative, and rehabilitative services, which can be in the form of medical centers, community health centers, clinics, hospitals, pharmacies, and laboratories.

Educational facilities are the resources and infrastructure that support teaching and learning activities to achieve educational goals. These facilities include classrooms, libraries, laboratories, and sports fields, as well as equipment such as computers, projectors, and textbooks [19].

A disaster is an event that cannot be predicted when it will occur and can cause injuries and fatalities, as well as damage and losses. A disaster is a series of events that threaten and disrupt the lives and livelihoods of communities, caused by natural and/or non-natural factors or human factors, resulting in fatalities, damage to the environment, property losses, and psychological impacts. Disasters can be grouped into two categories, namely natural and non-natural disasters. Natural disasters are caused by nature, such as earthquakes, tsunamis, volcanic eruptions, floods, droughts, tornadoes, landslides, and so on. Meanwhile, non-natural disasters are caused by epidemics, outbreaks, and so on [20].

The Equivalency Package Program consists of several levels, namely Package A, Package B, and Package C. Package A is equivalent to elementary school level, Package B is equivalent to junior high school level, and Package C is equivalent to senior high school level. With this program, it is hoped that the dropout rate can be reduced and those who have dropped out of school can obtain an equivalent education. According to one source, the government provides equivalency school programs, such as Kejar Package A, as part of its efforts to address the issue of school dropouts. This program provides opportunities for those who have dropped out of school to complete an education equivalent to the elementary school level [21].

The Community Reading Park (TBM) is an institution that was created by and for the community, which has the potential to empower citizens (the general public) to learn and obtain information/knowledge to improve their standard of living [19].

2.7 Data Sources

The empirical data comes from Statistics of Bengkulu Province Indonesia, Response Variable come from The National Socio-economic Survey (SUSENAS) and Auxiliary Variables comes from Village Potential (PODES) in 2022. This study made predictions the Proportion of poor people at the village level in Bengkulu City, Bengkulu Province in 2022. Then, this study using R Studio to visualization data, analysis data and prediction. The variables used can be seen in Table 1.

Table 1. The Research Variables

Variables		Code	Unit	Source
Response	The proportion of poor people	Y		SUSENAS 2022
Auxiliary Variables	Certificate of Indigency Recipients (Jumlah Penerima SKTM)	X_1	Household	PODES 2022
	Sum of Health Facilities	X_2	Unit	PODES 2022
	Sum of Education Facilities	X_3	Unit	PODES 2022
	Sum of Natural Disaster Posts	X_4	Unit	PODES 2022
	Sum of Package A/B/C Education Programs	X_5	Program	PODES 2022
	Sum of Community Reading Parks	X_6	Unit	PODES 2022

2.8 Analysis Stage

The step of analysis in this study are as follows:

1. Descriptive statistics of the research variables
2. Estimating model parameters

i. Determining the likelihood function: $h = \ell(\boldsymbol{\beta_z}, \phi, \sigma_v^2 | y_i, v_{i,}, \boldsymbol{z_i})$

ii. Determine the Adjusted Profile H-Likelihood (APHL) using the formula given in Eq. (14).

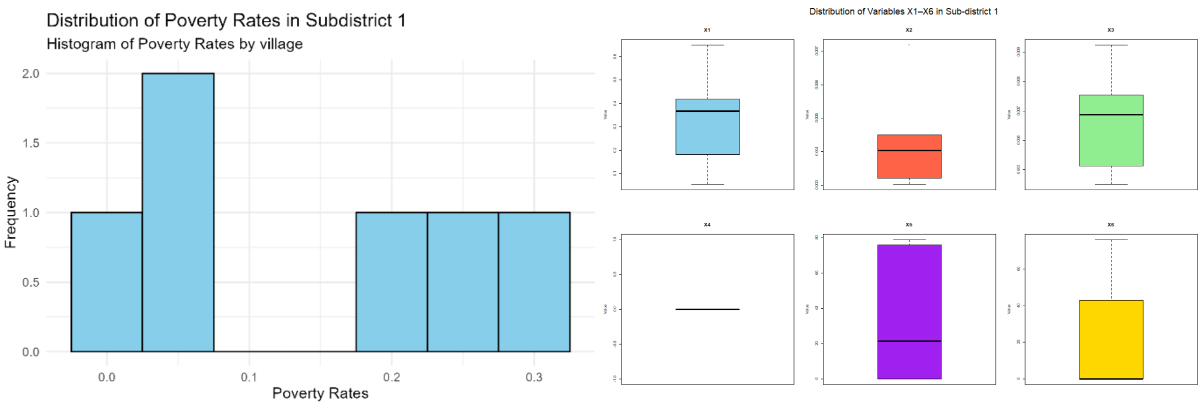
iii. Maximizing the function h and APHL for parameter estimation:

$$\hat{\varphi}^{HL} = (\widehat{\boldsymbol{\beta}}_{\boldsymbol{z}}, \hat{v}, \hat{\phi}, \hat{\sigma}_v^2)$$

3. Prediction of the model's proportion estimate using the formula $\hat{p}_i^{HL} = \frac{\exp(\boldsymbol{z}_i'\widehat{\boldsymbol{\beta}}_{\boldsymbol{z}}+\hat{v}_i)}{1+exp(\boldsymbol{z}_i'\widehat{\boldsymbol{\beta}}_{\boldsymbol{z}}+\hat{v}_i)}$.
4. Calculating the MSE using the bootstrap method: $MSE(\hat{p}_i^{HL}) = \frac{1}{B}\sum_{b=1}^{B}(p_i - \hat{p}_i^{HL})^2$
5. Summarizing the findings.

3. RESULTS AND DISCUSSION

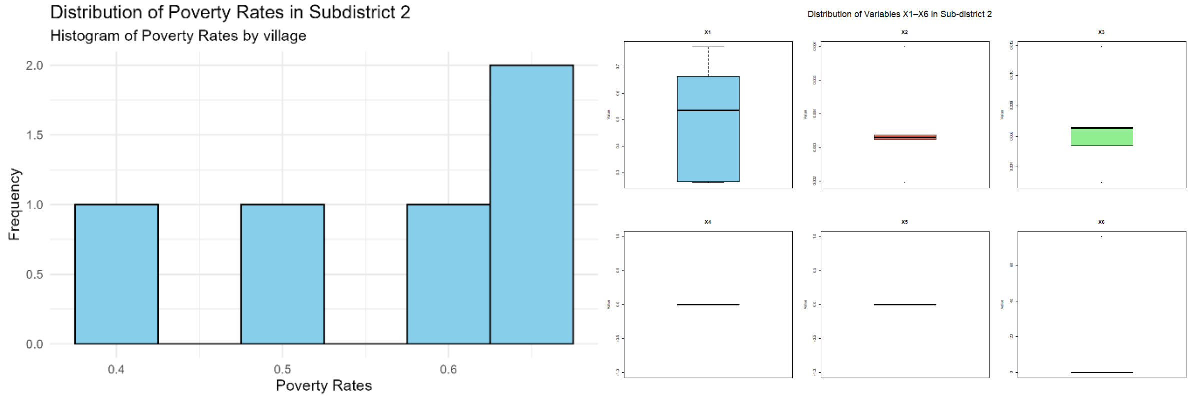
Bengkulu City is an urban area in Bengkulu Province, with an administrative government structure consisting of 9 sub-districts and 67 villages. Each village has distinct characteristics, both geographically, socially, and economically. Generally, villages in Bengkulu City can be grouped into two large areas: coastal villages and non-coastal villages [19]. In the 2022 SUSENAS data, 46 villages were selected as samples. Direct estimates of poverty rates per sub-district from the SUSENAS sample are presented below.



(a) (b)

Figure 1. (a) Distribution of Poverty Rates in subdistrict 1, (b) Distribution of Variable X in the sub-district 1

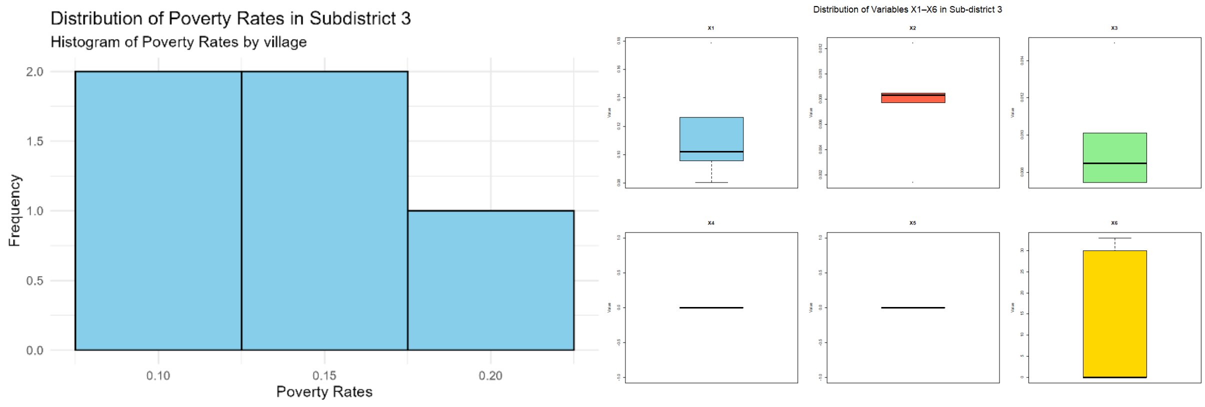
Based on Fig. 1, the distribution of Poverty Rates in Sub-district 1 shows that most villages are in the low to medium poverty rate range, which is around $0.05 - 0.10$ and the medium level at $0.20 - 0.30$. The frequency of villages with poverty rates below 0.05 and above 0.30 is relatively small. The box plot for variables X_1 to X_6 shows that variable X_1 to X_3 have a relatively stable distribution, although X_2 and X_3 have a narrow distribution. The variable X_4 is more constant with no variation. The variables X_5 and X_6 have a wider distribution with values that are much higher than the others.



(a) (b)

Figure 2. (a) Distribution of Poverty Rates in subdistrict 2, (b) Distribution of Variable X in the sub-district 2

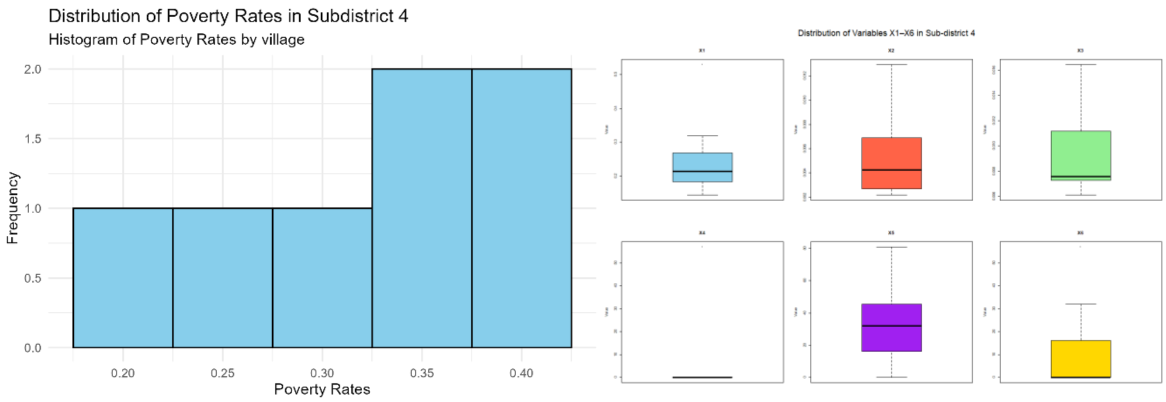
Based on Fig. 2, The distribution of Poverty Rates in Sub-district 2 shows that most villages are in the medium to high poverty rate range, around $0.40 - 0.70$. The highest frequency is seen in villages with poverty rates approaching 0.70, while the number of villages with poverty rates below 0.40 is relatively small. Meanwhile, the box plot results for variables X_1 to X_6 show that variable X_1 has a fairly wide spread with a relatively high median. Variables X_2, X_4, X_5, and X_6 appear constant with no significant variation, making them less informative for further analysis. Variable X_3 shows a slight variation although its spread is still narrow. Thus, in Sub-district 2, only X_1 and X_3 provide information on inter-village variation, while the other variables tend to be irrelevant due to minimal variation.



(a) (b)

Figure 3. (a) Distribution of Poverty Rates in subdistrict 3, (b) Distribution of Variable X in the sub-district 3

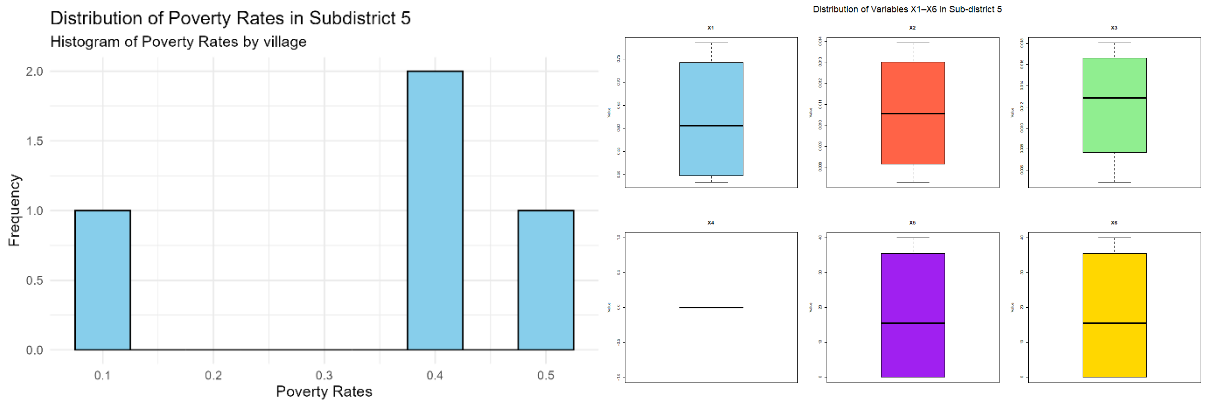
Based on Fig. 3, the distribution of poverty rates in Sub-district 3 shows variation among villages, with Poverty Rate values spread in the range of approximately $0.3 - 0.7$. Most villages are in the $0.6 - 0.7$ range, which means the majority of villages in this sub-district have a moderate to high poverty rate, although there are still villages with a lower poverty rate. Meanwhile, the distribution of variables $X_1 - X_7$ shows a fairly clear difference in scale. Variables X_1 and X_6 have relatively large values compared to the other variables, while X_2, X_3, and X_4 tend to be small with a narrower variation. Variables X_5 and X_7 are at a medium level with a moderate spread. This condition again highlights the need for data standardization so that further analysis can provide balanced results when assessing the contribution of each variable to the poverty rate.



(a) (b)

Figure 4. (a) Distribution of Poverty Rates in subdistrict 4, (b) Distribution of Variable X in the sub-district 4

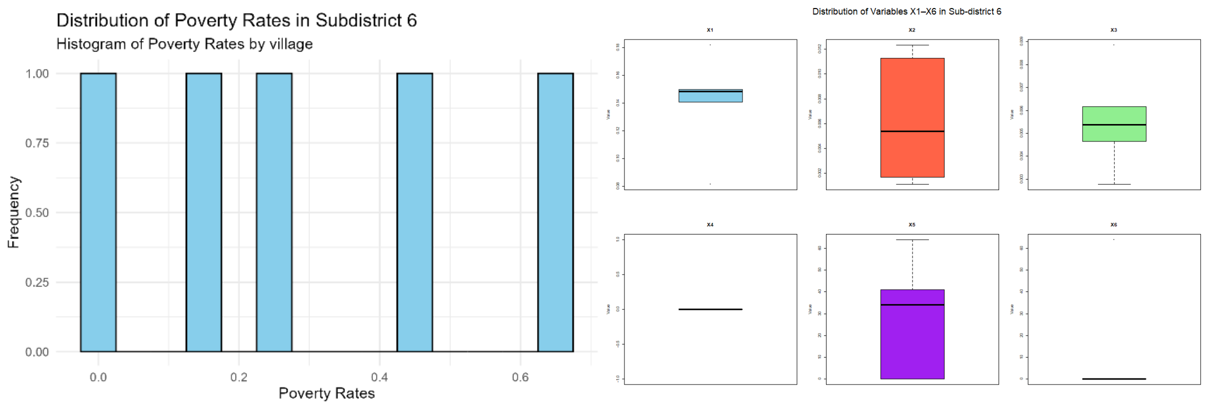
Based on Fig. 4, the distribution of Poverty Rates in Sub-district 4 is in the range of 0.18 to 0.40. The distribution shows that the villages in this sub-district have a medium to relatively high poverty rate. Many villages are concentrated in the $0.30 - 0.40$ range, while only a few villages are at a lower poverty rate. This indicates a tendency for inequality among villages, with most falling into the higher poverty group. From the box plot of variables X_1 to X_6, it is evident that variables X_1, X_2, X_3, X_5, and X_6 show a fairly clear variation among villages, with X_2 and X_5 having a wider range. Meanwhile, variable X_4 is almost constant, approaching zero, making it less informative. This condition illustrates that the explanatory factors in Sub-district 4 are relatively more varied compared to the previous sub-districts, particularly for variables X_2 and X_5.



(a) (b)

Figure 5. (a) Distribution of Poverty Rates in subdistrict 5, (b) Distribution of Variable X in the sub-district 5

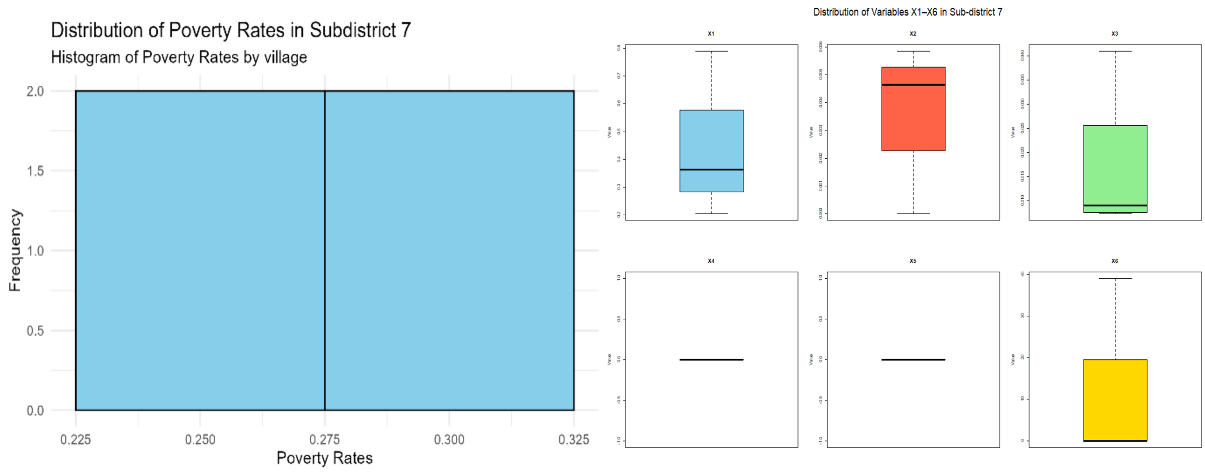
Based on Fig. 5, The distribution of Poverty Rates in Sub-district 5 shows a value range between 0.08 and 0.50. The distribution pattern appears quite varied, with one village at a low poverty rate, two villages at a high poverty rate around 0.40, and one village at a very high poverty rate around 0.50. This condition highlights the existence of a gap among villages within the sub-district, from those that are relatively prosperous to those facing a quite serious poverty level. From the box plot of variables $X_1 - X_6$, it's evident that X_1, X_2, and X_3 have a relatively widespread, indicating a variation in conditions among villages. Variable X_4 tends to be constant, while X_5 and X_6 show a wider distribution, especially X_6 which has extreme values in some villages. This indicates that explanatory factors, such as those represented by X_5 and X_6 , play an important role in explaining the variation in poverty in Sub-district 5.



(a) (b)

Figure 6. (a) Distribution of Poverty Rates in subdistrict 6, (b) Distribution of Variable X in the sub-district 6

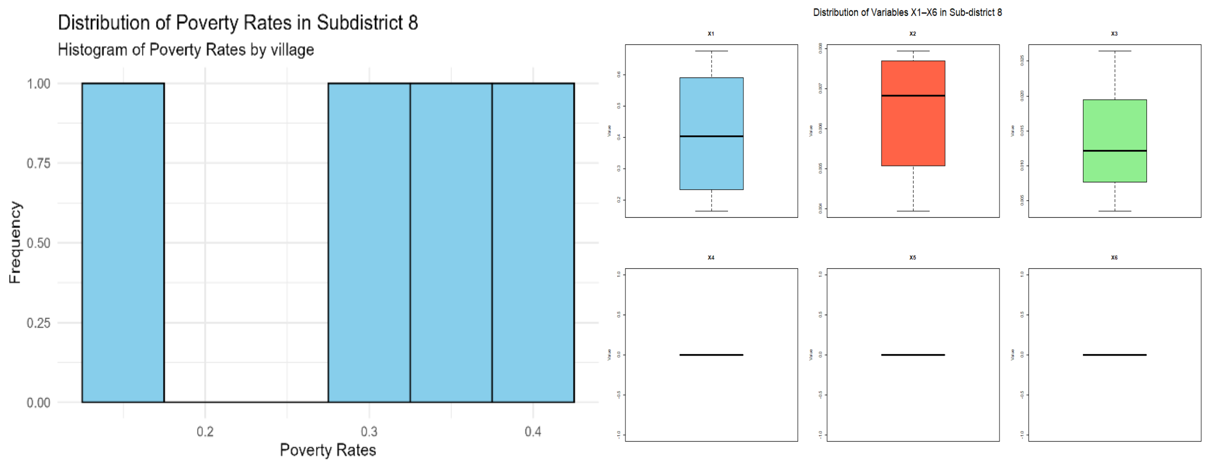
Based on Fig. 6 in Sub-district 6, the distribution of Poverty Rates shows a greater variation, with values spread evenly from 0.4 to 1.0. This pattern indicates a significant diversity in poverty conditions among villages, ranging from low to very high. Meanwhile, variables X_1 to X_7 show a similar pattern to Sub-district 5, where X_1 and X_6 have the highest values, while the other variables are relatively low. However, the variation in the intermediate variables, such as X_5 and X_7 is more visible, reflecting differences in characteristics among the villages.



(a) (b)

Figure 7. (a) Distribution of Poverty Rates in subdistrict 7, (b) Distribution of Variable X in the sub-district 7

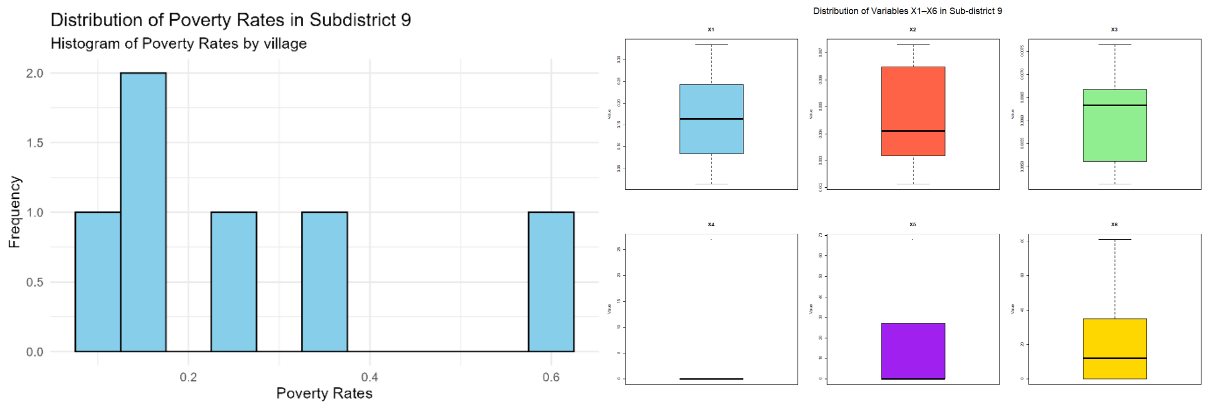
Based on Fig. 7, the distribution of Poverty Rates in Sub-district 7 is in the range of 0.23 to 0.32, with a relatively narrow pattern. This indicates that the poverty rate among villages in this sub-district tends to be uniform at a medium level, with no villages being significantly left behind or very advanced. On the box plot of variables $X_1 - X_6$, it is visible that X_1, X_2, and X_3 have a clearer variation compared to the other variables. X_2 in particular shows a wider spread, while X_4 and X_5 appear constant with no significant variation. Variable X_6 has a medium distribution with some differences among villages. Overall, Sub-district 7 shows a homogeneous poverty condition, with few internal factors driving differences among villages.



(a) (b)

Figure 8. (a) Distribution of Poverty Rates in subdistrict 8, (b) Distribution of Variable X in the sub-district 8

Based on Fig. 8, the distribution of Poverty Rates in Sub-district 8 is in the range of 0.15 to 0.40 with a relatively even spread. This shows that the poverty rate among villages varies from low to medium, without any villages experiencing extreme poverty. From the box plot of variables $X_1 - X_6$, it is visible that $X_1 - X_3$ have a clear spread, with X_2 showing higher variation than the others. In contrast, X_4 and X_5 appear constant with no variation, while X_6 has a more limited spread. This condition indicates that some socio-economic factors in Sub-district 8 are quite diverse, although there are variables that tend to be uniform among villages.



(a) (b)

Figure 9. (a) Distribution of Poverty Rates in subdistrict 9, (b) Distribution of Variable X in the sub-district 9

Based on Fig. 9, The distribution of Poverty Rates in Sub-district 9 ranges from 0.05 to 0.55 with a wider pattern compared to the previous sub-districts. Most villages are at a low to medium level, but there are also villages with a fairly high poverty rate. This reflects the existence of a gap between villages within this sub-district. On the box plot of variables $X_1 - X_6$, it's visible that $X_1 - X_3$ have a clear variation with a relatively balanced data distribution. Variable X_4 appears constant, while X_5 shows a widespread, and X_6 is at a medium level with limited variation. Overall, Sub-district 9 shows higher heterogeneity in terms of both poverty and the characteristics of its explanatory variables.

Table 2. Prediction of Poverty Rates Using the SAE-BB-APHL Model

Sub-District	Village	Sample Size	$\hat{p}_i^{APHL}$	$\hat{v}_i$	Sub-District	Village	Sample Size	$\hat{p}_i^{APHL}$	$\hat{v}_i$
1	1	20	0.4112	0.5672	5	24	10	0.2246	-2.2151
1	2	20	0.3328	-0.0957	5	25	8	0.2275	-1.9775
1	3	20	0.1341	-1.1956	5	26	9	0.0179	-5.2578
1	4	40	0.3539	0.2506	5	27	10	0.7462	-0.457
1	5	10	0.4409	-0.0473	6	28	19	0.6649	-0.0765
1	6	18	0.0177	-2.9943	6	29	10	0.0479	-3.402
2	7	40	0.4936	0.1684	6	30	10	0.8111	1.89512
2	8	19	0.5602	1.3327	6	31	10	0.3056	-2.1924
2	9	10	0.5261	0.9915	6	32	10	0.6345	-0.9406
2	10	19	0.4632	0.179	7	33	10	0.8335	0.57743
2	11	20	0.4387	-0.0942	7	34	10	0.5125	-0.9848
3	12	20	0.2557	0.4508	7	35	10	1	8.79186
3	13	10	0.2775	0.5303	7	36	9	0.9906	2.61029
3	14	8	0.2625	0.7781	8	37	19	0.7602	0.45898
3	15	10	0.3573	0.481	8	38	10	0.6108	-1.1686
3	16	9	0.1586	-0.3316	8	39	10	0.2409	-1.1344
4	17	10	0.3628	0.3077	8	40	10	0.5047	-1.1511
4	18	10	0.7065	1.2916	9	41	10	0.1857	-2.2328
4	19	20	0.5473	0.3073	9	42	10	0.4463	-0.7237
4	20	10	0.3681	0.4868	9	43	18	0.0411	-3.8121
4	21	20	0.4955	1.2369	9	44	10	0.9617	3.31105
4	22	10	0.3573	0.538	9	45	9	0.8586	1.35424
4	23	10	0.5074	0.8023	9	46	20	0.0242	-3.4477

Fig. 10 presented shows a comparison of poverty rate estimates in 46 villages using three different approaches: direct estimation, the SAE-BB-APHL and SAE-BB-HL method. The graph is divided into two panels; the first panel displays the results for villages 1–23, while the second panel shows the results for villages 24–46. The horizontal axis indicates the village number, while the vertical axis represents the proportion of poverty rates (Poverty Rates) on a scale ranging from 0 to approximately 0.7.

In general, the direct estimation results show considerable variation between villages. Extreme fluctuations in values in some regions reflect the weaknesses of the direct estimation method, which is highly dependent on sample size and data diversity at the local level. In areas with limited sample sizes, direct estimates tend to have high variance, making them less reliable in depicting true population conditions [2]. This is evident in several villages, such as Village 7 and Village 31, where the direct estimate values are much higher than those in other villages. This high variability has the potential to bias conclusions about socioeconomic conditions at the micro level.

In contrast, the SAE-BB-APHL method exhibits a more stable, consistent, and realistic estimation pattern across regions. The turquoise bars representing the SAE estimation results appear to have a more uniform distribution, indicating that this approach can reduce estimation uncertainty. One of the advantages of the SAE method is its ability to utilize additional information from other areas through the borrowing strength mechanism, thus improving estimates in areas with small sample sizes [6]. Thus, the SAE-BB-APHL method produces more efficient and reliable estimates than direct estimation.

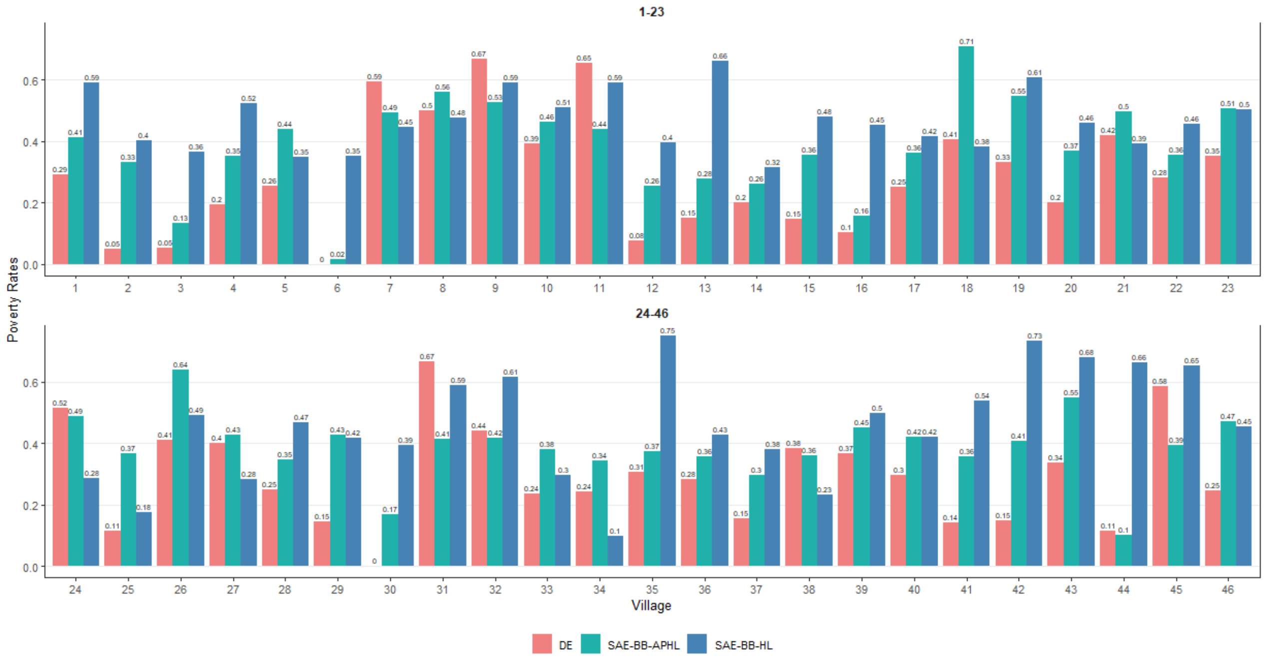


Figure 10. Comparison of Poverty Rates Prediction SAE-BB-APHL, SAE-BB-HL, and DE Estimation

Several interesting patterns can be observed from the graph. In Village 18, for example, the SAE estimate yields higher results than the direct estimate. This phenomenon can occur when the SAE model captures area effects not fully represented in the direct survey data, such as collective economic factors, spatial linkages, or social dynamics that indirectly influence poverty levels [2]. Conversely, the large differences between the two methods in certain villages, such as Village 7 and Village 31, may indicate sampling error or extreme local conditions that cause the direct estimate to deviate significantly from the true value.

Overall, these results reinforce the findings of various previous studies that the SAE method, particularly with the Bayesian approach or empirical best linear unbiased prediction (EBLUP), can produce more accurate and stable estimates than the direct method [2]. The application of the SAE-BB-APHL method to village poverty estimation has been shown to produce smoother and more consistent results across regions. With higher-quality estimates, this approach can be a crucial tool in supporting evidence-based policymaking, particularly in poverty alleviation programs at the local level.

Therefore, the use of the SAE-BB-APHL method not only contributes to improving the quality of poverty statistics but also strengthens the national statistical system by providing more accurate and reliable micro-level information. This approach can be a strategic solution for statistical institutions and local governments to overcome the limitations of macro-level surveys, which often fail to capture the diversity of poverty conditions at the village or sub-district level

The accuracy of the poverty rate estimates obtained using SAE-BB-APHL can be assessed using an evaluation matrix, namely Mean Squared Error (MSE). In Table 3 and Fig. 11 comparing the MSE of SAE-BB-APHL, SAE-BB-HL, and Direct Estimates.

Table 3. MSE Poverty Rates via SAE-BB-APHL, SAE-BB-HL, and Direct Estimation

Sub-District	Village	MSE APHL	MSE HL	MSE DE	Sub-District	Village	MSE APHL	MSE HL	MSE DE
1	1	0.00524	0.00602	0.20638	5	24	0.00524	0.00532	0.24974
1	2	0.00493	0.00565	0.04865	5	25	0.00526	0.00485	0.10122
1	3	0.00250	0.00540	0.04986	5	26	0.00512	0.00523	0.24221
1	4	0.00511	0.00603	0.15727	5	27	0.00510	0.00613	0.24000
1	5	0.00561	0.00537	0.19037	6	28	0.00498	0.00551	0.18750
1	6	0.00068	0.00457	0.00000	6	29	0.00549	0.00578	0.12493
2	7	0.00562	0.00585	0.24113	6	30	0.00268	0.00551	0.00000
2	8	0.00575	0.00551	0.25000	6	31	0.00615	0.00573	0.22222
2	9	0.00451	0.00518	0.22222	6	32	0.00543	0.00541	0.24654
2	10	0.00522	0.00578	0.23842	7	33	0.00491	0.00534	0.18075
2	11	0.00525	0.00567	0.22633	7	34	0.00493	0.00474	0.18408
3	12	0.00474	0.00534	0.07002	7	35	0.00565	0.00408	0.21302
3	13	0.00520	0.00476	0.12856	7	36	0.00521	0.00616	0.20250
3	14	0.00517	0.00577	0.16000	8	37	0.00505	0.00575	0.13081
3	15	0.00534	0.00520	0.12620	8	38	0.00557	0.00519	0.23669
3	16	0.00332	0.00570	0.09275	8	39	0.00555	0.00590	0.23222
4	17	0.00541	0.00597	0.18750	8	40	0.00590	0.00611	0.20816
4	18	0.00493	0.00524	0.24143	9	41	0.00592	0.00562	0.12245
4	19	0.00537	0.00586	0.22222	9	42	0.00562	0.00563	0.12620
4	20	0.00538	0.00593	0.16000	9	43	0.00553	0.00517	0.22383
4	21	0.00594	0.00574	0.24356	9	44	0.00236	0.00555	0.10122
4	22	0.00579	0.00555	0.20215	9	45	0.00583	0.00304	0.24306
4	23	0.00568	0.00594	0.22837	9	46	0.00630	0.00582	0.18595

Fig. 11 shows a comparison of the Mean Squared Error (MSE) values between the SAE BB APHL, SAE BB HL method, and Direct Estimation. The results indicate that the MSE of the SAE BB APHL and SAE BB HL method is much smaller than the DE MSE in almost all villages. The DE MSE values are relatively more varied compared to the APHL and HL MSE, which is more consistent and falls within the $0.00 - 0.01$ range. The SAE-BB-APHL method demonstrates stability and superiority by being able to reduce the variability of estimates. However, there are a few villages with a different pattern, where the MSE of both methods is almost the same or the difference is relatively small. This pattern confirms that the use of SAE APHL can significantly increase the accuracy and efficiency of poverty estimation at the village level compared to direct estimation, especially in areas with small sample sizes.

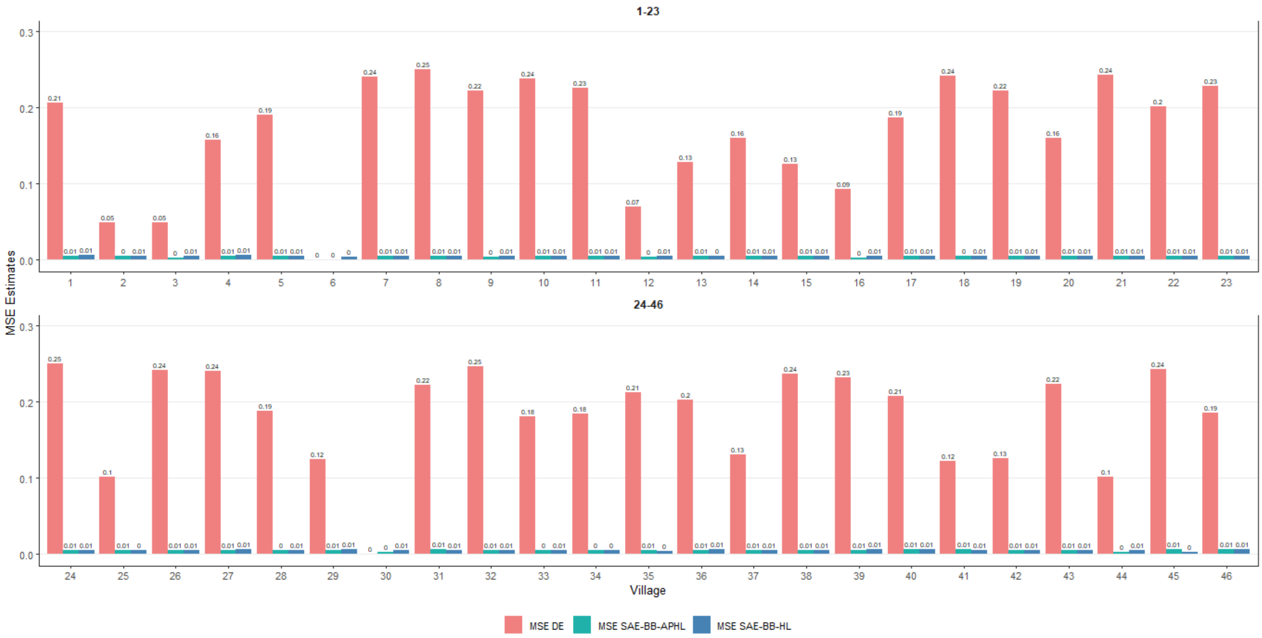


Figure 11. Comparison of MSE SAE BB APHL, SAE BB HL, and Direct Estimates

The overall visual pattern shows a striking difference between the three estimation techniques. The MSE for the Direct Estimation (DE) method (shown by the red bars) is significantly higher in all villages, mostly ranging between 0.10 and 0.25, and peaking around 0.25 in several villages, such as villages 8, 24, and 32. In contrast, the MSE of both Small Area Estimation (SAE) methods, namely as SAE-BB-APHL (green bars) and SAE-BB-HL (blue bars), remains consistently low, generally fluctuating between 0.00 and 0.01. This striking difference indicates that both SAE methods produce estimates with significantly lower error variance, suggesting that they provide more stable and accurate parameter estimates for small-area-level data.

From a methodological perspective, these results underscore the superiority of model-based estimation (SAE) over direct estimation (DE). The graph compares two variations of the SAE model: SAE-BB-HL (Hierarchical Likelihood) and SAE-BB-APHL (Adjusted Profile Hierarchical Likelihood). Both incorporate a Beta-Binomial distribution framework with a hierarchical likelihood structure, which effectively addresses the problem of overdispersion. This structure allows the model to incorporate random effects that capture region-specific variability. The key difference in the APHL approach is the "adjusted profile" function, which is designed to provide more robust parameter estimates and variance stabilization compared to the HL method. By leveraging additional information, both SAE methods employ a "borrowing-strength" mechanism that improves estimation reliability.

In contrast, the DE approach relies solely on observed sample data in each region. Consequently, its performance is highly sensitive to sample variability and data scarcity, resulting in much larger MSE values, as observed in the figure. The high degree of MSE fluctuation among direct estimates further highlights their instability and lack of precision. This finding reinforces a widely recognized limitation of direct estimation in small-area applications: its inability to produce statistically reliable results when local sample sizes are insufficient.

The observed superiority of the SAE-BB-APHL approach is consistent with the theoretical framework of Small Area Estimation (SAE) as articulated by [6]. According to their findings, model-based SAE methods particularly those incorporating a hierarchical likelihood component in the Beta-Binomial model offer substantial improvements in precision and efficiency compared to traditional design-based estimators. The integration of the Beta-Binomial distribution and adjusted profile likelihood in the SAE-BB-APHL model provides a flexible and computationally stable solution for addressing non-normality and heterogeneity in small area data. This makes it particularly suitable for poverty mapping and similar applications where micro-level precision is crucial for evidence-based policymaking.

In summary, the figure provides clear empirical evidence that the SAE-BB-APHL method outperforms the Direct Estimation approach in terms of accuracy and stability, as demonstrated by its consistently lower MSE values across villages. These results demonstrate that the SAE-BB-APHL framework not only improves estimation efficiency but also ensures greater robustness to data limitations common in small area analysis. Therefore, the application of the SAE-BB-APHL method represents a significant methodological improvement for producing accurate and reliable poverty estimates, offering valuable insights for policymakers and researchers involved in targeted regional development and poverty reduction initiatives.

4. CONCLUSION

The findings of this study clearly demonstrate the performance of the Small Area Beta-Binomial Estimation with Adjusted Profile Hierarchical Likelihood (SAE-BB-APHL) method compared to the Direct Estimation (DE) approach in producing reliable poverty estimates at the village level. Furthermore, when comparing SAE-BB-HL and SAE-BB-APHL, based on the estimation results, SAE-BB-APHL does indeed still produce overestimates when predicting the poverty rate. However, the MSE value for SAE-BB-APHL tends to be lower than that of SAE-BB-HL. This reduction in MSE is evidence of reduced bias and improved prediction accuracy. The SAE-BB-APHL method demonstrates its ability to produce estimates with higher precision and lower variability, even under conditions of limited or unbalanced sample data. These results reinforce the theoretical foundation of the model-based estimation framework, which leverages additional information and hierarchical structure to address the sampling limitations inherent in traditional design-based estimators. This indicates that SAE-BB-APHL is a suitable method for estimating poverty rates in Bengkulu City. These poverty estimation results are expected to assist policymakers in regional planning, particularly in remote areas within Bengkulu City. Future research plans include conducting a more thorough analysis of overestimation and bias reduction.

Author Contributions

Etis Sunandi: Formal Analysis, Funding Acquisition, Writing-Review and Editing. Pepi Novianti: Investigation, Resources, Validation, Writing-Review and Editing; Nurul Hidayati: Methodology, Visualization, Writing-Review and Editing. Anggun Permatasari: Software, Data Curation, Writing -Original Draft. Alya Saputri: Investigation, Project Administration, Writing-Original Draft. Lutfiah Firlian: Investigation, Resources, Validation. Indah Srilita Asmuyana: Resources, Writing-Original Draft. All authors discussed the results and contributed to the final manuscript.

Funding Statement

This research has been funded by PNBP University of Bengkulu for funding Fundamental Research with contract number 2094/UN30.15/PT/2025.

Acknowledgment

We would like to thank all parties who have provided valuable suggestions to improve the quality of this paper. We would like to thank Statistics Indonesia for providing data from SUSENAS and PODES of Bengkulu Province.

Declarations

The authors declare no competing interest.

Declaration of Generative AI and AI-assisted technologies

The authors declare that no generative AI or AI-assisted technologies were used in the preparation of this manuscript, including for writing, editing, data analysis, or the creation of tables and figures.

REFERENCES

[1] P. A. Aldini, 'PERFORMA PENGGUNAAN METODE SPATIAL EMPIRICAL BEST LINEAR UNBIASED PREDICTOR PADA SMALL AREA ESTIMATION', Thesis, Universitas Hasanuddin, 2024.

[2] J. N. K. Rao and I. Molina, *Small area estimation*, Second. Canada: John Wiley and Sons, Inc., Hoboken, New Jersey, 2015.

[3] A. J. Fernández, C. J. Pérez-González, A. Carrión-García, and V. B. Giner-, 'OVERDISPERSION EFFECTS ON RELIABILITY TEST PLANNING', *IEEE Trans*, vol. 72, no. No. 3, 2023, doi: https://doi.org/10.1109/TR.2022.3195055.

[4] Y. Lee and J. A. Nelder, 'HIERARCHICAL GENERALIZED LINEAR MODELS', *Journal of the Royal Statistical Society. Series B: Methodological*, vol. 58, no. 4, pp. 619–656, 1996. doi: https://doi.org/10.1111/j.2517-6161.1996.tb02105.x.

[5] J. Jiang, *Linear and Generalized Linear Mixed Models and Their Applications*. New York, 2007. [Online]. Available: http://www.springerlink.com/index/D7X7KX6772HQ2135.pdf. doi: https://doi.org/10.1007/978-0-387-47946-0.

[6] E. Sunandi, K. A. Notodiputro, I. Indahwati, and A. M. Soleh, 'BETA-BINOMIAL MODEL IN SMALL AREA ESTIMATION USING HIERARCHICAL LIKELIHOOD APPROACH', *AIP Conf. Proc. 3201, 060003*, 2024. doi: https://doi.org/10.14710/medstat.16.1.88-99.

[7] N. E. Khairiroh, 'KEMISKINAN DI INDONESIA: ANALISIS PENYEBAB, DAMPAK, DAN SOLUSI KEBIJAKAN', Maliki Interdisciplinary Journal (MIJ), vol. 3, no. Issue April, pp. 642–651, 2025, [Online]. Available: http://urj.uin-malang.ac.id/index.php/mij/index.

[8] BPS Provinsi Bengkulu, *Pembangunan Kependudukan Dalam Upaya Pengentasakan Kemiskinan Dan Pengembangan Wilayah*. Jakarta: Badan Pusat Statistik, 2024.

[9] Usman, Wartoyo, N. Haida, and N. Wahyuningsih, 'IMPLEMENTASI SUSTAINABLE DEVELOPMENT GOALS (SDGS) DI INDONESIA PERSPEKTIF EKONOMI ISLAM', *Al-Masharif : Jurnal Ilmu Ekonomi dan Keislaman*, vol. 11, no. 1, pp. 108–126, 2024. doi: https://doi.org/10.24952/masharif.v12i1.12446.

[10] E. Sunandi, K. A. Notodiputro, Indahwati, and A. M. Soleh, 'SMALL AREA ESTIMATION OF ILLITERACY RATES BASED ON BETA-BINOMIAL MODEL USING HIERARCHICAL LIKELIHOOD APPROACH', *Mathematics and Statistics*, vol. 11, no. 3, pp. 579–585, 2023. doi: https://doi.org/10.13189/ms.2023.110315.

[11] E. Tanur and A. Kurnia, 'SMALL AREA ESTIMATION FOR AUTOREGRESSIVE MODEL WITH MEASUREMENT kERROR IN THE AUXILIARY VARIABLE', *Communications in Mathematical Biology and Neuroscience*, vol. 83, pp. 1–29, 2022. doi: https://doi.org/10.28919/cmbn/7577.

[12] N. H. Pusponegoro, A. Kurnia, K. A. Notodiputro, A. M. Soleh, and E. T. Astuti, 'AREA SPECIFIC EFFECTS SELECTION OF SMALL AREA ESTIMATION FOR CONSTRUCTION OF REGIONAL CONSUMER PRICE INDICES IN INDONESIA', *Communications in Mathematical Biology and Neuroscience*, vol. 90, no. January 2018, pp. 1–16, 2021. doi: https://doi.org/10.28919/cmbn/6747.

[13] J. G. Skellam, 'A PROBABILITY DISTRIBUTION DERIVED FROM THE BINOMIAL DISTRIBUTION BY REGARDING THE PROBABILITY OF SUCCESS AS VARIABLE BETWEEN THE SETS OF TRIALS', *J R Stat Soc Series B Stat Methodol*, vol. 10, no. 2, pp. 257–261, 1948. doi: https://doi.org/10.1111/j.2517-6161.1948.tb00014.x.

[14] J. Wu and P. M. Bentler, 'APPLICATION OF H-LIKELIHOOD TO FACTOR ANALYSIS MODELS WITH BINARY RESPONSE DATA', *J Multivar Anal*, vol. 106, pp. 72–79, 2012. doi: https://doi.org/10.1016/j.jmva.2011.09.007.

[15] L. Nansadiqa, ‘ANALISIS PENGARUH PERTUMBUHAN EKONOMI TERHADAP KEMISKINAN DI INDONESIA’, *Jurnal HEI EMA*, vol. 3, pp. 46–60, 2024. doi: https://doi.org/10.61393/heiema.v3i2.233.
[16] BPS Provinsi Bengkulu, *Provinsi Bengkulu dalam Angka 2022*, vol. 4, no. 1. Bengkulu: BPS Provinsi Bengkulu, 2022.
[17] B. P. S. BPS Indonesia, *Pengeluaran Untuk Konsumsi Penduduk Indonesia 2016*, 1st ed. Jakarta: Badan Pusat Statistik Indonesia, 2016.
[18] I. Purnamasari, G. Marsha brata, F. T. Ramdani, and M. Y. G. Seran, ‘PELAYANAN PEMBUATAN SURAT KETERANGAN TIDAK MAMPU’, *Karimah Tauhid*, vol. 3, no. 2, pp. 1295–1302, 2024. doi: https://doi.org/10.30997/karimahtauhid.v3i2.11615.
[19] BPS Indonesia, *Statistik Potensi Desa Indonesia 2024*, vol. 15. Jakarta: Badan Pusat Statistik Indonesia, 2024.
[20] M. Danil, ‘MANAJEMEN BENCANA’, *Prosiding Mitigasi Bencana, Universitas Dharmawangsa*, pp. 7–14, 2021.
[21] N. L. Nurson and W. J. Sari, ‘PROGRAM PAKET KESETARAAN UPAYA MENGATASI ANGKA PUTUS SEKOLAH’, *Journal of Global and Multidisciplinary*, vol. 2, no. 1, pp. 838–844, 2024, [Online]. Available: https://journal.institercom-edu.org/index.php/multiple