

LOCALIZED DATA FOR EDUCATIONAL EQUITY: SMALL AREA ESTIMATION OF OUT-OF-SCHOOL CHILDREN IN BALI AND NUSA TENGGARA

Sherina Rafidah Khairunnisa¹, Azka Ubaidillah², Ahmad Sovi Hidayat³, Alya Nur Septiyana⁴, Shalihati Melani Putri⁵, Ahmad Tegar Prananggalih⁶, Arya Candra Kusuma^{7*}, Shafiyah Asy Syahidah⁸

¹Bappeda, Banyuwangi Regency Government

Jln. Ahmad Yani No. 100, Banyuwangi, Jawa Timur, Indonesia

²Department of Applied Statistics, Politeknik Statistika STIS

Jln. Otto Iskandardinata No 64C, Jakarta Timur, DKI Jakarta, 13330, Indonesia

³BPS Kaimana Regency, BPS – Statistic Indonesia

Kroy, Kaimana, Papua Barat, Indonesia

⁴BPS Sulawesi Barat Province, BPS – Statistic Indonesia

Jln, RE Martadinata No.10, Simboro Dan Kepulauan, Mamuju, Sulawesi Barat, Indonesia

⁵BPS Sleman Regency, BPS – Statistic Indonesia

Purbaya, Mlati, Sleman, Yogyakarta, Indonesia

⁶BPS Kepulauan Aru Regency, BPS – Statistic Indonesia

Jln. Pahlawan No. 11, Dobo, Kepulauan Aru, Maluku, Indonesia

⁷Directorate of Food Crops, Horticulture, and Plantation Statistics, BPS - Statistics Indonesia

Jln. Dr. Sutomo No. 6-8, 10710 Ps. Baru, Jakarta, Indonesia

⁸Directorate of Social Welfare Statistics, BPS - Statistics Indonesia

Jln. Dr. Sutomo No. 6-8, 10710 Ps. Baru, Jakarta, Indonesia

Corresponding author's e-mail: *aryacandra51@gmail.com

ABSTRACT

Article History:

Received: 27th September 2024

Revised: 3rd February 2025

Accepted: 4th March 2025

Published: 1st April 2025

Keywords:

Hierarchical Bayes;
Out-of-School Children;
Small Area Estimation.

This study aims to estimate the percentage of out-of-school children aged 7–17 years in Bali and Nusa Tenggara using the Small Area Estimation (SAE) method with a Hierarchical Bayes. One of the main challenges in education policy planning is the limited data available. National surveys, such as the National Socio-Economic Survey (Susenas), typically provide estimates only at the national and provincial levels, while more detailed data at the district level is often lacking. This limitation restricts the understanding of educational disparities at the local level and complicates the design of targeted policies. To address this issue, SAE Hierarchical Bayes provides a solution by producing more accurate district-level estimates, utilizing additional data without the need for new sampling. This method has proven to be cost-effective and efficient, particularly in regions with complex geography, such as Bali and Nusa Tenggara. The findings reveal that districts in East Nusa Tenggara generally exhibit a higher percentage of out-of-school children compared to the national average, indicating significant regional disparities that require attention. These findings highlight the urgency of improving educational infrastructure, particularly in underdeveloped areas of East Nusa Tenggara, to promote equitable access to education and reduce the number of children out of school



This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

How to cite this article:

S. R. Khairunnisa, A. Ubaidillah, A. S. Hidayat, A. N. Septiyana, S. M. Putri, A. T. Prananggalih, A. C. Kusuma and S. A. Syahidah., "LOCALIZED DATA FOR EDUCATIONAL EQUITY: SMALL AREA ESTIMATION OF OUT-OF-SCHOOL CHILDREN IN BALI AND NUSA TENGGARA," *BAREKENG: J. Math. & App.*, vol. 19, iss. 2, pp. 1179-1192, June, 2025.

Copyright © 2025 Author(s)

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

Education plays a critical role in shaping the future of nations by developing human capital, reducing inequality, and driving economic growth [1], [2]. While the Indonesian Constitution guarantees the right to education for every citizen, substantial disparities in access persist, particularly between urban and rural areas. Based on March 2023 Susenas, in rural areas, only 27.98% of the population aged 15 and above have completed secondary education, whereas, in urban areas, this figure increases to 49.16%, highlighting disparities in educational attainment that contribute to broader socio-economic inequalities [3]. These inequalities pose a significant obstacle to achieving “Indonesia Emas 2045”, which aspires to transform the nation into a prosperous and equitable society [4]. Addressing these challenges necessitates ensuring equitable access to education, supported by precise and comprehensive local data collection.

Across Indonesia, the phenomenon of out-of-school children remains a pressing challenge [5], [6], particularly in achieving equitable access to education. Even in the March 2023 Susenas, the percentage of out-of-school children in Indonesia reached 5.78%. This issue is particularly pronounced in regions like Bali and Nusa Tenggara, where a combination of geographical and socio-economic factors further exacerbate the problem. The archipelagic nature of these islands creates significant logistical challenges, with many remote areas lacking reliable transportation, infrastructure, and access to schools. Rural regions, particularly in Nusa Tenggara, also face severe poverty, which hinders families’ ability to prioritize education. Furthermore, in East Nusa Tenggara, the percentage of out-of-school children aged 7–17 years stands at 6.14%, exceeding the national average. Bali, with a 3.36% rate of children aged 7-17 years out-of-school, experiences a phenomenon where its dependence on tourism [7], [8] has led to a pattern in which some children forgo schooling to contribute to household income in the tourism sector [6]. These unique regional characteristics highlight the need for precise, localized data to better understand the distribution of out-of-school children and improve access to education in these underserved areas.

Data on the percentage of out-of-school children aged 7–17 years is available through the National Socio-Economic Survey (Susenas), but its estimates are limited to the national and provincial levels due to sample constraints. This makes it inadequate for effective policy-making at the district or city level. To overcome this, localized data and the application of Small Area Estimation (SAE) methods are essential for more precise estimates and targeted policies. Calculation of the percentage of out-of-school children aged 7–17 years in areas with minimal samples produces a high standard error, so the accuracy of the statistical data is questionable [9]. According to Statistics Indonesia, the relative standard error (RSE) value that is considered accurate is $\leq 25\%$ to produce good statistical results [10]. Bali-Nusa Tenggara is the 3rd island that has the highest percentage of districts with RSE of percentage of out-of-school children aged 7–17 years above 25% after Sumatra and Maluku. One solution that can be used to overcome this problem is the addition of samples in each district/city that has an RSE greater than 25%. On the other hand, additional samples will also increase the burden in terms of labor, time, and cost, especially in the Bali and Nusa Tenggara region, which consists of many small islands.

To address these challenges, Small Area Estimation (SAE) offers a powerful solution. SAE is a statistical technique that enhances estimation accuracy in regions with limited sample sizes by incorporating auxiliary data [11], such as economic and infrastructure variables, from related areas. This method effectively draws on information from surrounding regions to fill in gaps where data is scarce from surrounding areas to produce more reliable estimates at smaller levels [11]. In regions like Bali and Nusa Tenggara, where geographical barriers and limited access to data make it difficult to collect sufficient sample sizes, SAE becomes particularly valuable. It enables more accurate estimates despite the challenges posed by remote locations and fragmented infrastructure. Furthermore, the SAE method has been widely applied in various research fields [12], [13], [14]. By applying SAE, it becomes possible to generate precise and actionable estimates of the percentage of out-of-school children aged 7–17 years across different regions of Bali and Nusa Tenggara, providing policymakers with the detailed data needed to implement targeted educational interventions.

This study focuses on estimating the percentage of out-of-school children aged 7–17 years at the district level in Bali and Nusa Tenggara, which has not been previously conducted. Many studies, especially in education, have used SAE methods, such as the Gross Participation Rate of Primary and Secondary Education in Central Java [15], the Gross Participation Rate estimation of Higher Education in Indonesia [16], and the dropout rate in Semarang Regency [17]. Prior research has commonly utilized the total values of auxiliary variables derived from Podes (Potensi Desa/Village Potential) data to describe district/city conditions.

However, this approach may lead to inaccurate conclusions, as total values do not adequately reflect the distribution or characteristics of the district/city. Therefore, this study adopts a percentage- and ratio-based approach for auxiliary variables to provide a more representative depiction of district/city characteristics, ensuring a more accurate estimate. Based on the previous explanation, this study aims to provide an overview, estimate the percentage of out-of-school children aged 7–17 years at the district/city level in Bali and Nusa Tenggara using SAE, and evaluate the results.

2. RESEARCH METHODS

This research uses secondary data in the form of raw data from the March 2023 Susenas KOR Module and Podes 2021 obtained from Statistics Indonesia. Susenas KOR Module was selected to obtain direct estimates of the percentage of out-of-school children aged 7–17 years. This survey collects socio-economic data, including education indicators, and produces estimates down to the district/city level. Podes 2021, which provides a comprehensive overview of village/urban infrastructure, is used as the data source for the independent variables. The data used from Podes 2021 is aggregated from village-level records and combined at the district/city level. Besides that, the Ministry of Finance Data Portal is also used for the realization of social assistance data. The estimated/response variable (Y) in this study is the percentage of out-of-school children aged 7–17 years at the regency/city level in the provinces of Bali, Nusa Tenggara Barat (NTB), and Nusa Tenggara Timur (NTT). The auxiliary variables (X) used are 15 variables with the following details:

Table 1. List of Auxiliary Variables

Variable	Description	Category
X_1	District/city minimum wage	
X_2	Percentage of villages with workshops	
X_3	Percentage of villages with slum areas	
X_4	Percentage of villages with a salon	
X_5	Ratio of banks per village	
X_6	Ratio of village-owned enterprises (Bumdes) per village	Economy
X_7	Ratio of hotels per village	
X_8	Ratio of grocery stores/shops per village	
X_9	Ratio of minimarkets/supermarkets per village	
X_{10}	Ratio of semi-permanent markets per village	
X_{11}	Ratio of industrial centers per village	
X_{12}	Realization of social assistance	
X_{13}	Percentage of villages with internet access	
X_{14}	Ratio of early childhood education centers per village	Infrastructure
X_{15}	Ratio of high schools per village	

2.1 Out-of-school Children

According to the UNESCO Institute for Statistics, out-of-school children are the percentage of primary-school-age children who are not enrolled in primary or secondary school. Children in the official primary age group that are in preprimary education should be considered out of school. Based on Statistics Indonesia, out-of-school children are children aged SD/MI/equivalent, SMP/MTs/equivalent, and SMA/MA/equivalent (aged between 7–17 years old) who have never attended school, dropping out of school at a certain level, or completing certain levels but not continuing for other reasons such as geographical access difficulties, natural disasters, and conflict areas. In this study, the percentage of out-of-school children aged 7–17 years can be formulated in **Equation (1)**.

$$Y = \frac{\text{number of children aged 7 - 17 who are out - of - school}}{\text{number of children aged 7 - 17 years old}} \times 100\% \quad (1)$$

Various factors can influence the condition of out-of-school children, including economic factors [5], [18] and infrastructure factors [19]. One of the factors that influence children dropping out of school is

economic factors derived from parental income [20]. The minimum wage plays an important role in determining household economic conditions. An increase in the minimum wage can help increase family income, thus easing the burden of education costs and allowing children to continue attending school [21]. In addition, the existence of buildings such as workshops, salons, banks, BUMDes, hotels, shops/grocery stores, minimarkets/supermarkets, markets, and industrial centers can also be an indicator of economic activity taking place in the community. This is reinforced by Bali and Nusa Tenggara whose economies are supported by tourism, so the presence of hotels can have an impact on the economic activity of the activities of local communities [7], [8]. Slum dwellers are generally poor families [22]. Poor families prioritize basic needs such as food and shelter over education [23]. Therefore, the social assistance provided to poor families is expected to help them fulfill these basic needs so that they have better access to education [24]. In terms of infrastructure, access to the internet can improve access to information and knowledge [25]. Apart from that, educational progress is closely related to the availability of infrastructure, including school buildings. The existence of school buildings spread across various regions allows all school-aged children to gain easy access to education [26].

2.2 Direct Estimation

Direct estimation is an estimation method in which the calculation is only based on sample data from a domain or observed area. Direct estimation is defined as an estimation based only on the observed variable values from the survey sample [11]. The estimation number generated from the March 2023 Susenas data is estimated using the final sampling weight that has been subjected to various adjustments. For example, y_{ghij} represents the value of the characteristic Y for the j -th household, i -th census block, h -th stratum, g -th district/city, then the estimated proportion or percentage of the value of the characteristic Y is calculated based on Equation (2).

$$\hat{P}_Y = \frac{\sum_{ghij} W_{ghij}^{adj} y_{ghij}}{\sum_{ghij} W_{ghij}^{adj}} \quad (2)$$

Since Susenas applies a complex sampling design, the estimation of standard errors can be done with a Taylor linearization approach. This method treats the percentage or average as a ratio, $r = \frac{y}{x}$, where y is the total sample value, and x is the number of cases in the group or subgroup being considered. The calculation of standard error with Taylor linearization is expressed in Equation (3). However, direct estimates exhibit high variance in areas with small sample sizes, resulting in unreliable data. Increasing the sample size adds to the burden of labor, time, and cost, making the use of a small area estimation method essential.

$$SE = \sqrt{Var(\hat{R})} = \sqrt{\frac{1-f}{\hat{X}^2} \sum_{h=1}^2 \left[\frac{n'_{gh}}{n'_{gh} - 1} \sum_{i=1}^{n'_{gh}} \left(\hat{Z}_{ghi}^2 - \frac{\hat{Z}_{gh}^2}{n'_{gh}} \right) \right]} \quad (3)$$

where, $\hat{Z}_{ghi} = y_{ghi} - \hat{R}x_{ghi}$, $\hat{Z}_{gh} = \sum_i \hat{Z}_{ghi}$, $\hat{R} = \frac{\sum_{ghij} W_{ghij}^{adj} y_{ghij}}{\sum_{ghij} W_{ghij}^{adj} x_{ghij}}$.

2.3 Small Area Estimation Normal Hierarchical Bayes (SAE HB with Normal Distribution)

The Bayesian approach views parameters whose values are unknown as random variables. One of the methods included in the Bayesian approach is the Hierarchical Bayes (HB) method which utilizes a linking model by the Generalized Linear Mixed Model (GLMM) theory. The posterior distribution of the parameters of a variable under study is built using Bayes' theorem. As the name suggests, modeling using the HB method is carried out in stages by first deriving the posterior distribution of β and the variance component σ_u^2 and then deriving the posterior distribution of σ_u^2 [11]. The parameter estimation results from the HB method in terms of accuracy and reliability of the parameter estimates provided are better than those using multiple regression.

The HB model can maximize the results of auxiliary variable variance to explain the target variable. The ability to “borrow” information from other observations makes the HB model estimation results more stable. To obtain parameter estimates from the posterior distribution requires a very complex and difficult integral calculation, so the solution uses the Markov Chain Monte Carlo (MCMC) method [11]. The estimation result for inference with the Bayesian method is the formation of the posterior distribution obtained from a combination of prior (initial knowledge of the researcher) and likelihood (information on the data used).

The HB method is considered quite reliable in improving estimation results from direct estimation to make them more accurate. The HB method is commonly used because of its ease in handling various forms of data. This is in contrast to EBLUP, which is less appropriate to use when the data used is categorical [11]. One of the HB models proposed is SAE HB with normal distribution. The sampling model used for the HB with a normal distribution model is $\hat{\theta}_i | \theta_i, \boldsymbol{\beta}, \sigma_v^2 \sim N(\theta_i, \psi_i)$, $i = 1, \dots, m$, with the linking model as follows $\theta_i | \boldsymbol{\beta}, \sigma_v^2 \sim N(\mathbf{z}_i^T \boldsymbol{\beta}, b_i^2 \sigma_v^2)$, $i = 1, \dots, m$. Consider the case of known σ_v^2 and assume a “flat” prior on $\boldsymbol{\beta}$ given by $f(\boldsymbol{\beta}) \propto 1$ [11].

2.4 Research Steps

Data processing using R software version 4.2.2 (latest version used by the researcher to ensure compatibility with other packages) through the following stages [27]:

- (1) Conduct a literature review to determine auxiliary variables to be used. Literature review in research is generally based on economic factors [5] and infrastructure factors [19];
- (2) Calculate direct estimates of the percentage of out-of-school children aged 7–17 years at the regency/city level in the provinces of Bali, NTB, and NTT obtained from March 2023 Susenas data [11];
- (3) Explore the direct estimation data to find out the estimation method used [28];
- (4) Select the auxiliary variables based on the strength of their correlation with the direct estimates [11];
- (5) Determine MCMC parameters, such as the number of iterations, burn-in, thinning, and updates in the SAE HB modeling. These parameters are chosen based on initial exploratory analysis (trial-and-error approach) to ensure stable estimates. [11];
- (6) Formulate the SAE HB model with the determined auxiliary variables [11];
- (7) Evaluate the model by examining the RSE, ensuring it is below 25% for each district/city, and checking the validity of the SAE HB model by comparing the patterns and distributions of RSE and SAE HB estimates with direct estimates to ensure consistency [11];
- (8) Estimate the percentage of out-of-school children aged 7–17 years using the SAE HB model.

3. RESULTS AND DISCUSSION

3.1 Overview of Out Of School Children in Bali-Nusa Tenggara in 2023 from Direct Estimation

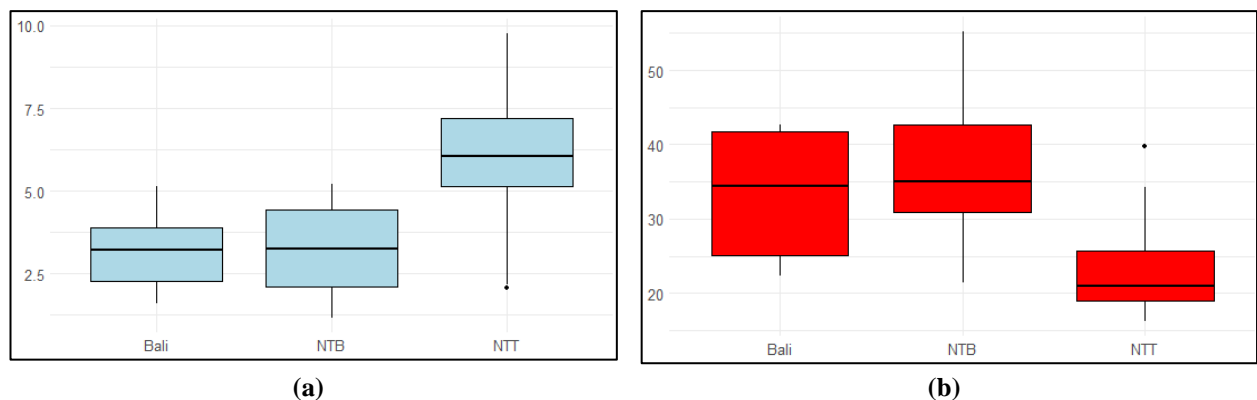
An overview of the percentage of out-of-school children 7–17 years in Bali and Nusa Tenggara in 2023 is presented in Table 2. The percentage of out-of-school children aged 7–17 years in NTT is generally higher than in Bali and NTB. Belu Regency in NTT has the highest percentage of out-of-school children aged 7–17 years among other districts/cities in Bali and Nusa Tenggara, at 9.7%. Meanwhile, the district/city with the lowest percentage of out-of-school children aged 7–17 years is Bima City in NTB, at 1.1%. Among the three provinces in Bali and Nusa Tenggara, it can be seen that NTT has a fairly long range compared to the others in the percentage of out-of-school children aged 7–17 years. This shows that there is a significant gap in access to education between the three provinces.

Table 2. Descriptive Statistics of The Percentage of Out-of-School Children Aged 7–17 Years in Bali and Nusa Tenggara in 2023

Descriptive Statistics	Bali	NTB	NTT
Minimum	1.60	1.17	2.08
Median	3.21	3.25	6.06
Mean	3.18	3.17	6.12
Maximum	5.14	5.21	9.77
Range	3.54	4.04	7.69

Note: Minimum, median, mean, maximum, and range represent the statistical summary of the percentage of out-of-school children aged 7–17 years in Bali, NTB, and NTT in 2023

Figure 1 shows the distribution of the percentage of out-of-school children aged 7–17 years in Bali and Nusa Tenggara Island in 2023. Based on the boxplot, the distribution of data in Bali and Nusa Tenggara Island shows a tendency to be normally distributed, as well as the distribution of data in each province. In addition, Sabu Raijua Regency is identified as a lower outlier in terms of the percentage of out-of-school children aged 7–17 years, indicating a significantly lower dropout rate compared to other areas. The low percentage of out-of-school children in Sabu Raijua may be attributed to teacher training and development, as improved teaching quality and teacher professionalism enhance access to education and reduce dropout rates [29], [30]. However, Sabu Raijua is also an upper outlier in the RSE for this same percentage, suggesting a high level of statistical variability or uncertainty in the data. This combination of lower dropout rates and higher RSE highlights the need for careful interpretation of the figures and may suggest underlying factors affecting educational outcomes in the region.

**Figure 1.** Boxplot of Percentage of Out-of-School Children 7–17 Years Old in Bali and Nusa Tenggara (a) Direct Estimation (%), (b) RSE (%)

To assess the distribution of the percentage of out-of-school children aged 7–17 years, we conducted a formal normality test using the Shapiro-Wilk test, as presented in **Table 3**. The Shapiro-Wilk test was chosen because it is particularly effective for evaluating normality in small sample sizes, typically less than 50 observations [31]. In this case, with 41 observations, the Shapiro-Wilk test is suitable for assessing normality. The normality test is assessed using the null hypothesis (H_0). If the p -value is below α (0.05), the data is not normally distributed. In this case, since p -values are above 0.05, the null hypothesis is not rejected, meaning the percentage of out-of-school children aged 7–17 years in each province within the Bali and Nusa Tenggara regions is normally distributed.

Table 3. Normality Test of Direct Estimation

Region	p -value	Decision	Conclusion
Bali	0.297	Fail to reject H_0	Normal
NTB	0.474	Fail to reject H_0	Normal
NTT	0.422	Fail to reject H_0	Normal

Note: p -values > 0.05 indicate that we failed to reject the null hypothesis (H_0).

After knowing the distribution of the direct estimation, an RSE check was conducted to see which regions still had RSE greater than 25%. Based on **Table 4**, it is known that there are 22 districts/cities out of 41 districts/cities in Bali-Nusa Tenggara Island that still have an RSE value of more than 25%. This shows that in Bali-Nusa Tenggara Island there are only about 46% or only 19 districts/cities that have a percentage of out-of-school children aged 7–17 years that is worth publishing because it meets the standard. Of the three

provinces, NTB is the province that has the highest number of districts/cities with RSE direct estimation greater than 25%. **Figure 1** shows that NTB has the most districts/cities with the highest RSE, namely Dompu, with an RSE of 55%. Besides that, Bali and NTB have similar patterns of direct estimate distribution and RSE. Therefore, the two provinces will be combined in the modeling process. The modeling will be divided into two, namely model 1 for Bali-NTB and model 2 for NTT.

Table 4. Number of Districts/Cities with RSE Direct Estimation Greater than 25%

Province	Number of Districts/Cities	Districts/Cities
Bali	7	Denpasar City, Bangli, Jembrana, Gianyar, Tabanan, Klungkung, Badung
NTB	9	Dompu, Sumbawa, Central Lombok, West Sumbawa, North Lombok, West Lombok, Mataram City, Bima City
NTT	6	Rote Ndao, Kupang, Lembata, Sabu Raijua, North Central Timor, Kupang City

3.2 Auxiliary Variable Selection

Before estimating the percentage of out-of-school children aged 7–17 years using the SAE method, it is necessary to select auxiliary variables. The auxiliary variables are selected based on the significance of the correlation between the auxiliary variables and the direct estimation (**Table 5**). The modeling is done by separating the Bali-NTB and NTT, so the auxiliary variables used in each province will be different. After selection, the auxiliary variables used in the SAE modeling of Bali-NTB include $X_1, X_5, X_6, X_7, X_8, X_9, X_{10}, X_{11}$, and X_{12} . Meanwhile, the auxiliary variables used in SAE modeling of NTT include $X_2, X_3, X_4, X_5, X_9, X_{13}, X_{14}$, and X_{15} .

Table 5. Correlation Between Auxiliary Variable and Direct Estimation

Model 1: Provinsi Bali-NTB		Model 2: Provinsi NTT	
Variable	Correlation	Variable	Correlation
X_1	-0.1684*	X_2	-0.4480*
X_5	-0.2082*	X_3	-0.4060*
X_6	0.4317*	X_4	-0.4576*
X_7	0.3882*	X_5	-0.5666*
X_8	0.3839*	X_9	-0.4648*
X_9	-0.1335*	X_{13}	-0.5409*
X_{10}	0.5339*	X_{14}	-0.3806*
X_{11}	-0.3168*	X_{15}	-0.4353*
X_{12}	0.3122*		

*) significant at $\alpha = 10\%$

To ensure the reliability of the selected variables, it is crucial to check for multicollinearity, which occurs when two or more independent variables in a model are highly correlated, leading to instability in the regression coefficients. This is commonly assessed using the Variance Inflation Factor (VIF). A VIF value greater than 10 is often considered indicative of problematic multicollinearity. As shown in **Table 6**, all auxiliary variables in the models for Bali-NTB and NTT have VIF values below 10, indicating that multicollinearity is not a concern and that the variables can be used confidently for further analysis.

Table 6. Multicollinearity Check

Model 1: Provinsi Bali-NTB		Model 2: Provinsi NTT	
Variabel	VIF	Variabel	VIF
X_1	5.3876	X_2	3.3772
X_5	3.0351	X_3	3.8892
X_6	1.6770	X_4	4.2213
X_7	1.9372	X_5	2.7137
X_8	3.7008	X_9	2.7478
X_9	6.4156	X_{13}	3.2166
X_{10}	1.8914	X_{14}	2.5883
X_{11}	1.1067	X_{15}	3.3568
X_{12}	1.0894		

3.3 SAE Hierarchical Bayes with Normal Distribution Model Analysis

The estimation of the percentage of out-of-school children aged 7–17 years is done using the SAE HB with a normal distribution approach because the distribution in the direct estimation has a normal distribution. SAE HB with normal distribution is considered more favorable and has a smaller Mean Square Error (MSE) compared to the EBLUP method and can accommodate EBLUP assumptions that are not met [32]. To obtain parameter estimates from complex posterior distributions, the integration process is difficult. Therefore, parameter estimation is often done numerically, one of which is the Markov Chain Monte Carlo (MCMC) technique. With this technique, there is a possibility of simulation in random sampling of complex stochastic models [33]. This is done by generating samples from the posterior distribution using Monte Carlo simulation iteratively through a Markov Chain process until it reaches a convergent or stationary condition [34]. Once the stationary condition is reached, the parameter samples are sampled from the posterior distribution. The Gibbs sampling algorithm is the simplest in MCMC. It can solve complex problems in multidimensional modeling into a series of small problems that are easy to solve.

In building the model, both for Bali-NTB and NTT, the total number of iterations, burn-in iterations, and thin need to be determined first by trial and error, starting from small values until the algorithm converges for all model parameters. After several simulation stages, the optimal model combination is achieved with 200 update iterations, 35,000 iterations per chain, a burn-in period of 15,000 iterations, and a thin of 10. The convergent condition can be seen in **Figure 2** and **Figure 3**, based on the trace plot, density plot, and autocorrelation plot for the model parameters.

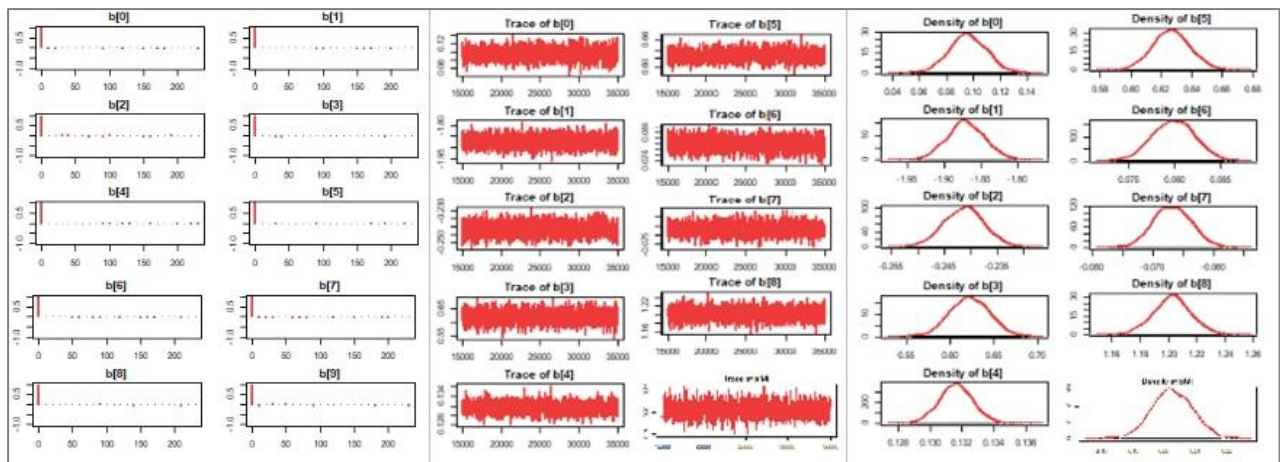


Figure 2. Visualization of Convergence with Autocorrelation Plot, Trace Plot, and Density Plot of the Bali-NTB Model

The autocorrelation plots in **Figure 2** and **Figure 3** show the correlation between samples for each iteration. Since the autocorrelation plot shows a cut-off after the first period, the parameter values of both models have converged. The trace plot shows the parameter values generated from some iteration values. Based on **Figure 2** and **Figure 3**, it can be seen that the trace plot of both models does not exhibit a periodic pattern so the MCMC technique has reached equilibrium so that the parameter values have converged. Meanwhile, the density plot shows the distribution of the generated parameter values. From **Figure 2** and **Figure 3**, it can be seen that the density plot has shown a smooth pattern so that the parameter values are proven to be convergent.

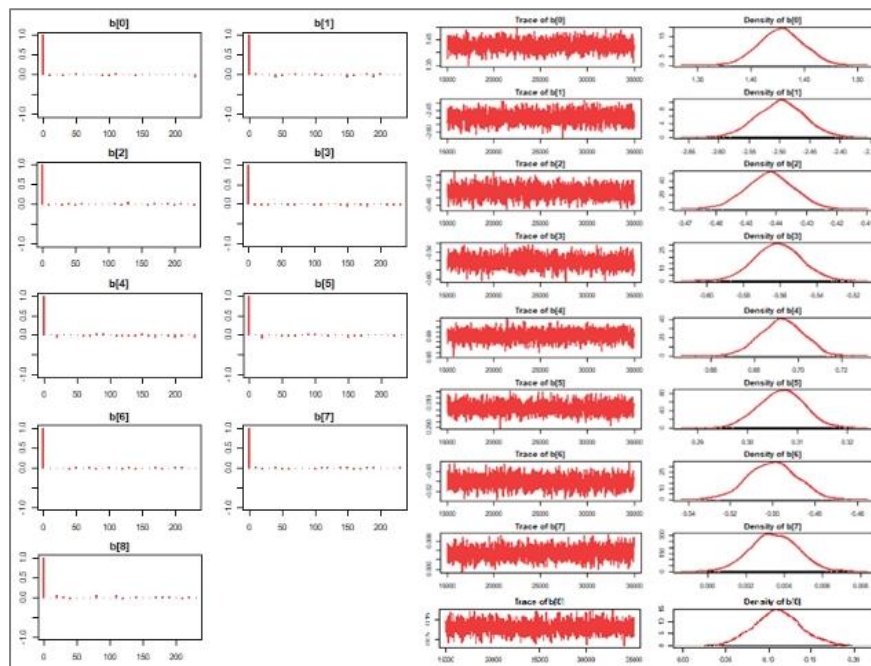


Figure 3. Visualization of Convergence with Autocorrelation Plot, Trace Plot, and Density Plot on NTT Model

After the MCMC simulation process with the Gibbs Sampling algorithm successfully converges the parameters, the next step is to see the significance of each variable. **Table 7** shows from 9 selected variables included in the Bali-NTB model, all auxiliary variables have a significant effect on the percentage of out-of-school children aged 7–17 years at the district/city level in Bali and NTB, which is marked by the credible interval value not passing zero. Meanwhile, of the 8 selected variables included in the NTT model, all auxiliary variables have a significant effect on the percentage of out-of-school children aged 7–17 years at the district or city level in NTT.

In the Bali-NTB model, the variables that have a positive and significant influence on the percentage of out-of-school children aged 7–17 years are the district/city minimum wage (X_1), the ratio of banks per village (X_5), ratio of Village-Owned Enterprises (BUMDes) per village (X_6), the ratio of minimarkets/supermarkets per village (X_9), the ratio of semi-permanent markets per village (X_{10}), and the realization of social assistance (X_{12}). Meanwhile, the ratio of hotels per village (X_7), the ratio of grocery stores/shops per village (X_8), and the ratio of industrial centers per village (X_{11}) have a negative influence on the percentage of out-of-school children aged 7–17 years at the district/city level. In the NTT model, the variables of the percentage of villages with workshops (X_2), the percentage of villages with slum areas (X_3), the percentage of villages with a salon (X_4), and the ratio of early childhood education centers per village (X_{14}) have a positive and significant influence on the percentage of out-of-school children aged 7–17 years at the district/city level in NTT. Meanwhile, variables that have a negative and significant influence include the ratio of banks per village (X_5), the ratio of minimarkets/supermarkets per village (X_9), the percentage of villages with internet access (X_{13}), and the ratio of high schools per village (X_{15}).

Table 7. SAE Estimation Coefficient Results from SAE HB with Normal Distribution

Variable	Mean	SD	Credible Interval				
			2.50%	25%	50%	75%	97.50%
Bali-NTB Model							
Intercept	-0.0792	0.0163	-0.1111	-0.0898	-0.0798	-0.0679	-0.0480
X_1	0.1402	0.0010	0.1383	0.1396	0.1402	0.1408	0.1422
X_5	0.5988	0.0243	0.5531	0.5818	0.5995	0.6149	0.6479
X_6	1.2128	0.0130	1.1874	1.2041	1.2130	1.2218	1.2375
X_7	-0.2384	0.0031	-0.2447	-0.2404	-0.2384	-0.2364	-0.2326
X_8	-0.0674	0.0030	-0.0732	-0.0694	-0.0674	-0.0654	-0.0614
X_9	0.6194	0.0108	0.5979	0.6122	0.6191	0.6266	0.6405
X_{10}	0.1912	0.0080	0.1755	0.1859	0.1915	0.1966	0.2070
X_{11}	-1.8532	0.0264	-1.9049	-1.8712	-1.8531	-1.8364	-1.7999
X_{12}	0.0819	0.0017	0.0785	0.0807	0.0819	0.0831	0.0852

Variable	Mean	SD	Credible Interval				
			2.50%	25%	50%	75%	97.50%
NTT Model							
Intercept	1.3514	0.0199	1.3118	1.3384	1.3515	1.3644	1.3908
X_2	0.2984	0.0057	0.2874	0.2945	0.2984	0.3022	0.3093
X_3	0.0043	0.0014	0.0016	0.0034	0.0043	0.0052	0.0069
X_4	0.7041	0.0104	0.6834	0.6974	0.7042	0.7110	0.7251
X_5	-2.5053	0.0451	-2.5932	-2.5380	-2.5038	-2.4742	-2.4201
X_9	-0.5502	0.0118	-0.5729	-0.5582	-0.5500	-0.5424	-0.5265
X_{13}	-0.4439	0.0074	-0.4583	-0.4488	-0.4440	-0.4386	-0.4297
X_{14}	0.1194	0.0243	0.0725	0.1033	0.1198	0.1361	0.1667
X_{15}	-0.5623	0.0119	-0.5863	-0.5701	-0.5620	-0.5545	-0.5394

3.4 Evaluation

Based on **Figure 4**, it is evident that the SAE HB with normal distribution model consistently produces a lower RSE than direct estimation across all districts and cities in the Bali-Nusa Tenggara region. This suggests that the SAE HB with normal distribution model offers more precise estimates. Moreover, the trend in RSE for the SAE HB with normal distribution closely mirrors the pattern observed in direct estimation, reinforcing the model's robustness. On average, the RSE for the SAE HB with normal distribution is 9.73%, reflecting a significant reduction of 65.87% compared to direct estimation, highlighting the shrinkage effect that improves the accuracy of the estimates.

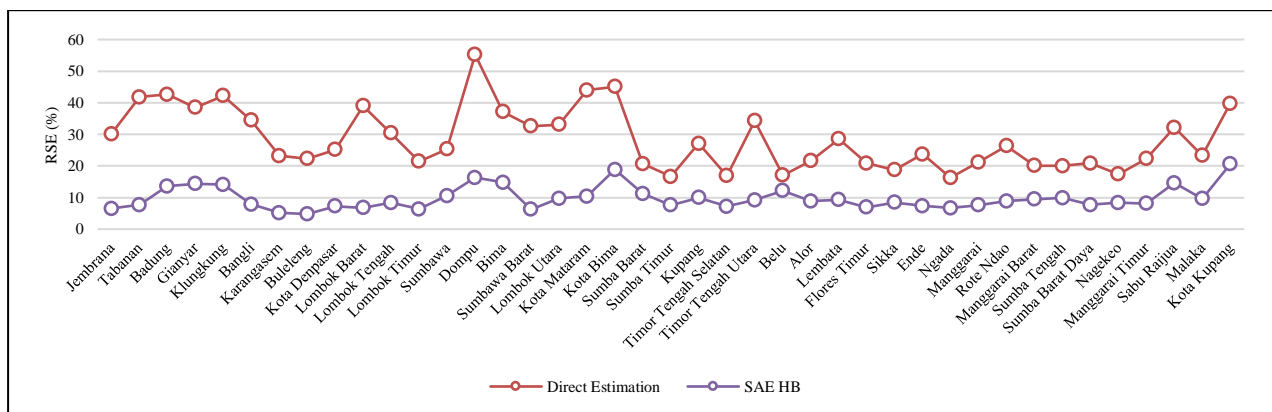


Figure 4. Comparison of RSE Estimation Results of the Percentage of Out-of-School Children Aged 7–17 Years in Bali and Nusa Tenggara

In addition, SAE HB with normal distribution modeling demonstrated notable success in reducing the Relative Standard Errors (RSEs) across all districts and cities in Bali and Nusa Tenggara to below 25% (as shown in **Table 8**). The significant reduction in RSE highlights the high accuracy and effectiveness of the modeling approach. This achievement improves the precision and reliability of data estimation, enabling regional governments to confidently design and implement evidence-based policies, such as targeted resource allocation or social program planning.

Table 8. Comparison of RSE Direct Estimate and SAE HB with Normal Distribution

RSE Level	Direct Estimate	SAE HB with normal distribution
(1)	(2)	(3)
RSE ≤ 25%	19	41
25% < RSE ≤ 50%	21	0
RSE > 50%	1	0

Based on **Figure 5**, the movement of the estimated percentage of out-of-school children aged 7–17 years in Bali and Nusa Tenggara Island in the SAE HB with normal distribution model also follows the movement of the direct estimation. This shows the consistency of SAE HB with normal distribution modeling to the estimated percentage of out-of-school children aged 7–17 years in Bali and Nusa Tenggara Island resulting from direct estimation.

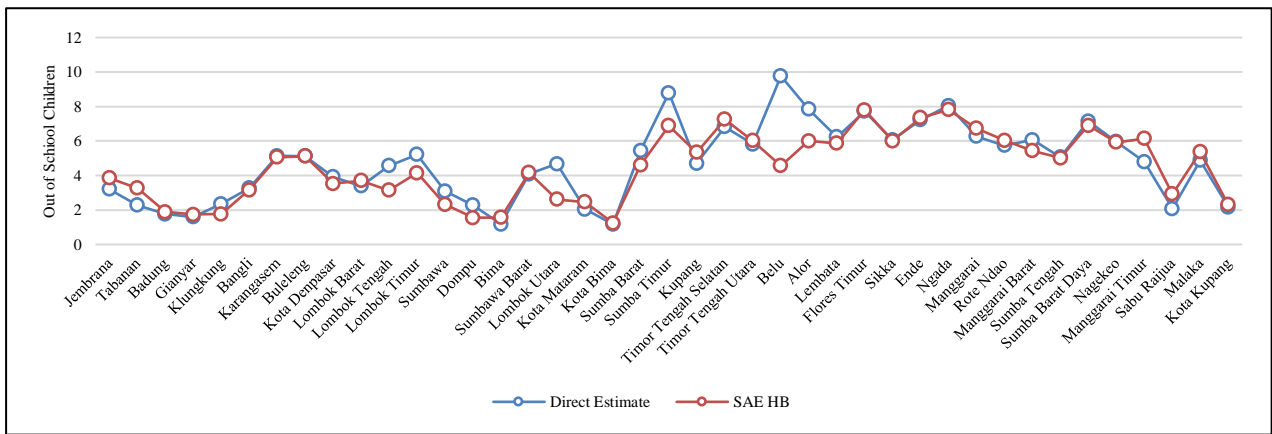


Figure 5. Comparison of Estimated Percentage of Out-of-School Children Aged 7–17 Years in Bali and Nusa Tenggara Islands

This result is also supported by **Figure 6**, which illustrates the average estimates of the percentage of out-of-school children aged 7–17 years in Bali and the Nusa Tenggara Islands at the provincial level. Notably, the averages fall well within the interquartile range depicted in the boxplot. This positioning within the interquartile range indicates a relatively consistent and stable percentage of out-of-school children aged 7–17 years across the provinces in these regions.

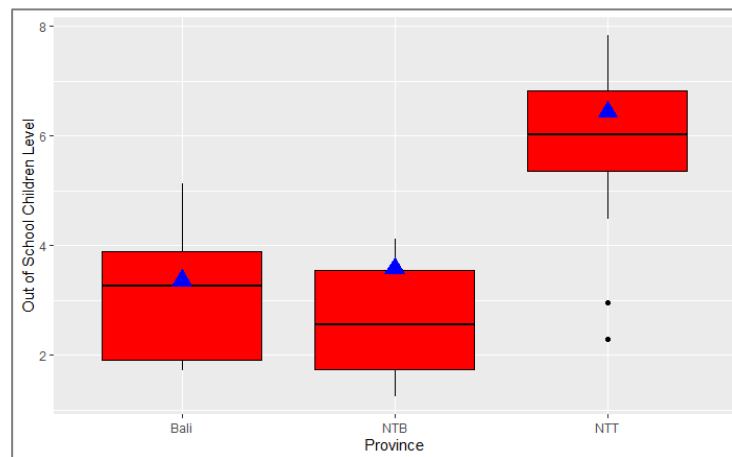


Figure 6. Boxplot of Estimated Percentage of Out-of-School Children Aged 7–17 Years at Regency/City Level in Bali and Nusatenggara

Based on **Table 9**, both the direct estimation method and the SAE HB with normal distribution method identify NTT as having the highest mean percentage of out-of-school children aged 7–17 years, with values of 6.12% and 5.83%, respectively, followed by Bali (3.18% for direct estimation, 3.27% for SAE HB with normal distribution) and NTB (3.17% for direct estimation, 2.69% for SAE HB with normal distribution). The range of minimum to maximum values for the direct estimation method is quite large, particularly in NTT (2.08%–9.77%), reflecting higher inequality compared to Bali and NTB. In contrast, the SAE HB with normal distribution method produces a narrower range of values, such as in NTT (2.32%–7.82%), providing more refined estimates by reducing the impact of extreme values or data imbalances.

Table 9. Comparison of Estimated Percentage of Out-of-School Children Aged 7–17 Years at Regency/City Level in Bali and Nusa Tenggara

Province	Direct Estimation				SAE HB with normal distribution			
	Mean	Median	Min	Max	Mean	Median	Min	Max
Bali	3.18	3.21	1.60	5.14	3.27	3.27	1.73	5.13
NTB	3.17	3.25	1.17	5.21	2.69	2.54	1.25	4.17
NTT	6.12	6.06	2.08	9.77	5.83	6.01	2.32	7.82

Figure 7 shows that the majority of districts/municipalities in Bali-Nusa Tenggara Island have an estimated percentage of children aged 7–17 who are not in school below the national rate (5.78%), especially in Bali and NTB. Meanwhile, the estimated percentage of out-of-school children aged 7–17 years above the

national rate is mainly owned by districts/municipalities in NTT. This is due to the limitations of these districts/municipalities, with 13 districts in NTT categorized as “underdeveloped, frontier, and outermost” regions according to Perpres No. 63/2020 on the establishment of underdeveloped regions for 2020-2024.

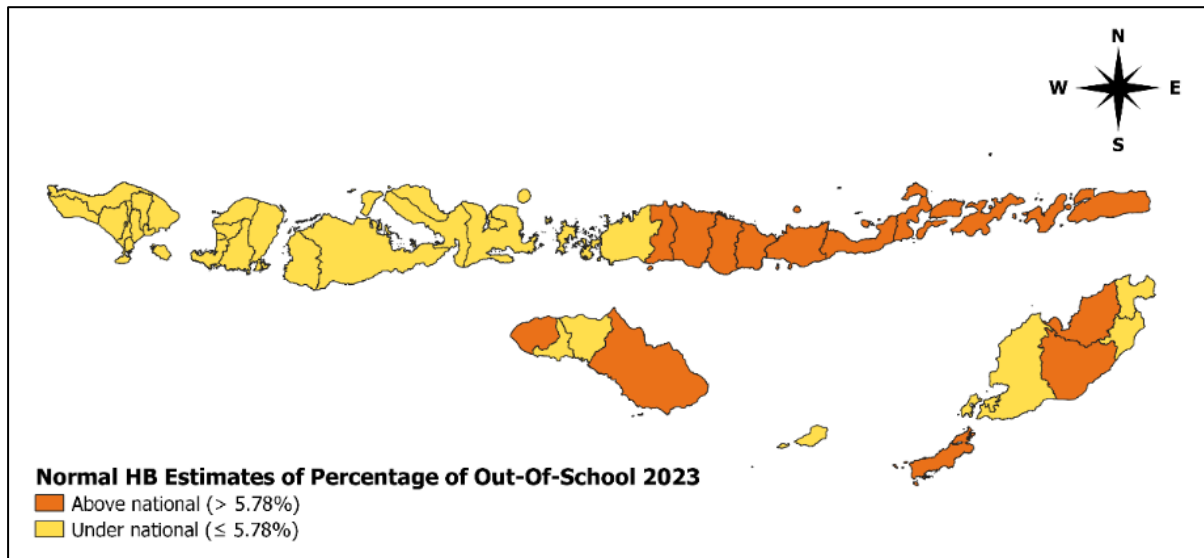


Figure 7. Map of Estimates of The Percentage Out-of-School Children Aged 7–17 Years on Bali-Nusa Tenggara Using SAE HB with Normal Distribution

4. CONCLUSIONS

Based on the results and previous discussion, the conclusions that can be drawn from this research include:

1. Referring to direct estimation, Belu Regency (NTT) has the highest percentage of out-of-school children aged 7–17 years among other districts/cities in Bali and Nusa Tenggara, while the lowest is Bima City (NTB). The distribution of direct estimates in each province, as indicated by both the graphical representation and formal testing, confirms that the data follows a normal distribution. NTB is the province that has the highest number of districts/cities with direct estimate RSEs greater than 25%.
2. Indirect estimation conducted using SAE HB with normal distribution shows a convergent plot, and all auxiliary variables significantly affect the response variable.
3. The use of SAE HB with normal distribution in estimating the percentage of out-of-school children aged 7–17 years at the district/city level in Bali and Nusa Tenggara can improve estimation accuracy compared to direct estimation with a significant reduction of 65.87% in RSE. Estimation using SAE HB with normal distribution shows that the majority of districts/municipalities with the percentage of out-of-school children aged 7–17 years above the national rate are located in NTT.

Suggestions that can be addressed to the government are the need for more attention to facilities and infrastructure that can support education in NTT. Meanwhile, a suggestion that can be considered for future research is to estimate the percentage of out-of-school children aged 7–17 years who are not in school at a level below the district/city level, such as subdistricts or villages. With this estimation, the results obtained can be a consideration for the formulation of targeted and specific policies for smaller areas.

ACKNOWLEDGMENT

The authors would like to thank the Directorate of Social Welfare Statistics Indonesia for providing the direct estimation data along with RSE in this study. The authors presented appreciation to the Department of Applied Statistics Politeknik Statistika STIS for the assistance in this research.

REFERENCES

- [1] F. A. Surbakti, J. E. Simanjuntak, T. Pascha S, and J. Ivanna, "POLICY EVALUATION OF ASSISTANCE PROGRAM OPERATIONAL SCHOOL (CASE STUDIES ON JUNIOR HIGH SCHOOL COUNTRY 17 MEDAN, BANDAR SELAMAT, KEC. MEDAN TEMBUNG, CITY MEDAN)," *Journal of General Education Science*, vol. 1, no. 2, pp. 104–108, 2023.
- [2] D. R. A. Sahek and N. B. M. Nasri, "A SYSTEMATIC REVIEW: ACCESS EQUALITY TOWARDS QUALITY EDUCATION," *Creat Educ*, vol. 10, no. 12, pp. 2947–2954, 2019, doi: 10.4236/ce.2019.1012219.
- [3] I. Widiyanti, A. V. Abdul, A. Nirwana, and A. Della Arlita, "ANCAMAN MELAWAN PUTUS SEKOLAH DENGAN DILEMA KUALITAS PENDIDIKAN INDONESIA," *JISIP: Jurnal Ilmu Sosial dan Pendidikan*, vol. 7, no. 3, pp. 2118–2126, 2023, doi: <http://dx.doi.org/10.58258/jisip.v7i3.5228>.
- [4] Bappenas, "VISI INDONESIA 2045," 2019. Accessed: Jul. 31, 2023. [Online]. Available: <https://perpustakaan.bappenas.go.id/e-library/dokumen/d547d8b7-1920-405f-8d11-36dcc350da45>
- [5] D. A. Maghfirah, "FAKTOR-FAKTOR PENYEBAB SISWA PUTUS SEKOLAH TINGKAT SMA/SMK NEGERI DI KOTA MATARAM," *Jurnal Spektrum Analisis Kebijakan Pendidikan*, vol. 8, no. 3, pp. 215–222, 2019.
- [6] F. Chandra, K. S. M. Parwati, and F. L. Amir, "ANALISIS FENOMENA SOSIAL PEKERJA ANAK DI DESTINASI WISATA PANTAI KUTA BALI," *MSJ: Majority Science Journal*, vol. 1, no. 3, pp. 64–68, 2023.
- [7] D. A. Hutama and M. S. W. Suliswanto, "THE INFLUENCE OF THE TOURISM SECTOR ON ECONOMIC GROWTH AND LABOR ABSORPTION IN THE PROVINCE OF BALI," *Journal of Indonesian Tourism, Hospitality, and Recreation*, vol. 6, no. 1, pp. 15–26, 2023.
- [8] T. Taupikurrahman and E. Suwandana, "ANALISIS SEKTOR PARIWISATA PROVINSI NUSA TENGGARA BARAT DAN DAMPAK MOTOGP MANDALIKA," *Jurnal Kepariwisata Indonesia: Jurnal Penelitian dan Pengembangan Kepariwisata Indonesia*, vol. 16, no. 2, pp. 78–100, Jun. 2024, doi: 10.47608/jki.v16i22022.78 - 100.
- [9] K. S. Button *et al.*, "POWER FAILURE: WHY SMALL SAMPLE SIZE UNDERMINES THE RELIABILITY OF NEUROSCIENCE," *Nat Rev Neurosci*, vol. 14, no. 5, pp. 365–376, May 2013, doi: 10.1038/nrn3475.
- [10] BPS-Statistics Indonesia, "SAMPLING ERROR ESTIMASI HASIL SURVEI PERTANIAN ANTAR SENSUS (SUTAS) 2018," 2019.
- [11] J. N. K. Rao and I. Molina, *SMALL AREA ESTIMATION*. Wiley, 2015. doi: 10.1002/9781118735855.
- [12] S. S. Khan, A. C. McLain, B. A. Olatosi, M. E. Torres, and J. M. Eberth, "SMALL AREA ESTIMATION OF COUNTY-LEVEL U.S. HIV-PREVALENT CASES," *Ann Epidemiol*, vol. 48, pp. 30-35.e9, Aug. 2020, doi: 10.1016/j.annepidem.2020.05.008.
- [13] M. Pratesi, L. Quattrociochi, G. Bertarelli, A. Gemignani, and C. Giusti, "SPATIAL DISTRIBUTION OF MULTIDIMENSIONAL EDUCATIONAL POVERTY IN ITALY USING SMALL AREA ESTIMATION," *Soc Indic Res*, vol. 156, no. 2–3, pp. 563–586, Aug. 2021, doi: 10.1007/s11205-020-02328-5.
- [14] A. Ubaidillah, K. A. Notodiputro, A. Kurnia, and I. W. Mangku, "MULTIVARIATE FAY-HERRIOT MODELS FOR SMALL AREA ESTIMATION WITH APPLICATION TO HOUSEHOLD CONSUMPTION PER CAPITA EXPENDITURE IN INDONESIA," *J Appl Stat*, vol. 46, no. 15, pp. 2845–2861, Nov. 2019, doi: 10.1080/02664763.2019.1615420.
- [15] G. Firmando and A. Ubaidillah, "PENDUGAAN AREA KECIL UNTUK ANGKA PARTISIPASI KASAR PENDIDIKAN DASAR DAN MENENGAH TINGKAT KABUPATEN/KOTA DI PROVINSI JAWA TENGAH TAHUN 2018," *Seminar Nasional Official Statistics*, vol. 2020, no. 1, pp. 651–661, Sep. 2021, doi: 10.34123/semnasoffstat.v2020i1.466.
- [16] R. Saadi and A. Ubaidillah, "PENDUGAAN AREA KECIL ANGKA PARTISIPASI KASAR PERGURUAN TINGGI DI INDONESIA MENGGUNAKAN MODEL SUBAREA TWOFOLD DENGAN PENDEKATAN HIERARCHICAL BAYES," *Seminar Nasional Official Statistics*, vol. 2022, no. 1, pp. 123–136, Nov. 2022, doi: 10.34123/semnasoffstat.v2022i1.1217.
- [17] N. F. J. Malik, A. Hoyyi, and D. Ispriyanti, "PENDUGAAN ANGKA PUTUS SEKOLAH DI KABUPATEN SEMARANG DENGAN METODE PREDIKSI TAK BIAS LINIER TERBAIK EMPIRIK PADA MODEL PENDUGAAN AREA KECIL," *Jurnal Gaussian*, vol. 3, no. 1, pp. 101–110, 2014, doi: <https://doi.org/10.14710/j.gauss.3.1.101%20-%20110>.
- [18] S. Mitra, S. K. Mishra, and R. K. Abhay, "OUT-OF-SCHOOL GIRLS IN INDIA: A STUDY OF SOCIOECONOMIC-SPATIAL DISPARITIES," *GeoJournal*, vol. 88, no. 1, pp. 341–357, Mar. 2022, doi: 10.1007/s10708-022-10579-7.
- [19] Parmo, M. H. Sucipto, and Sumarkan, "PENILAIAN KONDISI BANGUNAN GEDUNG SEKOLAH DASAR NEGERISTUDI KASUS DI SEKOLAH DASAR NEGERI SE-KABUPATEN MADIUN," *EMARA: Indonesian Journal of Architecture*, vol. 2, no. 1, pp. 42–52, 2016.
- [20] R. D. Guntur and M. Lobo, "STATISTICAL MODELLING FOR DROPPED OUT SCHOOL CHILDREN (DOS) IN EAST NUSA TENGGARA PROVINCE INDONESIA," *J Phys Conf Ser*, vol. 812, p. 012073, Feb. 2017, doi: 10.1088/1742-6596/812/1/012073.
- [21] D. W. Schanzenbach, J. Turner, and S. Turner, "RAISING STATE MINIMUM WAGES, LOWERING COMMUNITY COLLEGE ENROLLMENT," Cambridge, MA, Aug. 2023. doi: 10.3386/w31540.
- [22] A. Rigon, "DIVERSITY, JUSTICE AND SLUM UPGRADING: AN INTERSECTIONAL APPROACH TO URBAN DEVELOPMENT," *Habitat Int*, vol. 130, p. 102691, Dec. 2022, doi: 10.1016/j.habitatint.2022.102691.
- [23] M. Muhtar, "PEMENUHAN KEBUTUHAN DASAR KELUARGA MISKIN MELALUI KOMPLEMENTARITAS PROGAM BANTUAN SOSIAL (STUDI KASUS DI KABUPATEN BARITO KUALA)," *Sosio Konsepsia*, vol. 7, no. 3, pp. 119–129, Sep. 2018, doi: 10.33007/ska.v7i3.1443.

- [24] H. Rudji, "ANALISIS KONDISI DEMOGRAFI, EKONOMI DAN LINGKUNGAN PADA ANAK PUTUS SEKOLAH DI KECAMATAN BANAWA KABUPATEN DONGGALA (STUDI KASUS JENJANG PENDIDIKAN SEKOLAH MENENGAH PERTAMA)," *Jurnal Kolaboratif Sains*, vol. 6, no. 6, pp. 519–526, Jun. 2023, doi: 10.56338/jks.v6i6.3672.
- [25] D. A. Sprenger and A. Schwaninger, "TECHNOLOGY ACCEPTANCE OF FOUR DIGITAL LEARNING TECHNOLOGIES (CLASSROOM RESPONSE SYSTEM, CLASSROOM CHAT, E-LECTURES, AND MOBILE VIRTUAL REALITY) AFTER THREE MONTHS' USAGE," *International Journal of Educational Technology in Higher Education*, vol. 18, no. 1, p. 8, Dec. 2021, doi: 10.1186/s41239-021-00243-4.
- [26] A. A. Muhaimin, A. Gamal, M. A. S. Setianto, and W. L. Larasati, "THE SPATIAL JUSTICE OF SCHOOL DISTRIBUTION IN JAKARTA," *Heliyon*, vol. 8, no. 11, p. e11369, Nov. 2022, doi: 10.1016/j.heliyon.2022.e11369.
- [27] N. Tzavidis, L.-C. Zhang, A. Luna, T. Schmid, and N. Rojas-Perilla, "FROM START TO FINISH: A FRAMEWORK FOR THE PRODUCTION OF SMALL AREA OFFICIAL STATISTICS," *J R Stat Soc Ser A Stat Soc*, vol. 181, no. 4, pp. 927–979, Oct. 2018, doi: 10.1111/rssa.12364.
- [28] W. Kusuma and N. Iriawan, "SMALL AREA ESTIMATION OF EXPENDITURE PER-CAPITA IN BANYUWANGI WITH HIERARCHICAL BAYESIAN AND EMPIRICAL BAYES METHODS," *J Sci*, vol. 2, no. 3, pp. 43–48, 2017, doi: <http://dx.doi.org/10.12962/j23378530.v2i3.a3185>.
- [29] F. S. Tanggur et al., "MEMBINA KOMUNITAS BELAJAR GURU: BERBAGI PRAKTIK BAIK DAN PENGALAMAN UNTUK PENINGKATAN KUALITAS PEMBELAJARAN DI KABUPATEN SABU RAIJUA," *BERBAKTI: Jurnal Pengabdian Kepada Masyarakat*, vol. 2, pp. 323–334, Jan. 2025, doi: <https://doi.org/10.30822/berbakti.v2i3.3958>.
- [30] L. I. D. Koroh, K. E. Noya Nahak, F. S. Tanggur, and C. A. Naitili, "PELATIHAN PENGEMBANGAN KOMPETENSI GURU DI SABU RAIJUA," *Pengabdian Masyarakat Ilmu Pendidikan*, vol. 4, no. 1, pp. 73–79, 2024, doi: <https://doi.org/10.37792/pemimpin.v4i1.1231>.
- [31] Sugiyono, *METODE PENELITIAN KUANTITATIF, KUALITATIF, dan R&D*. Bandung: Alfabeta, 2019.
- [32] N. Permatasari and Widya Larasati, "PERBANDINGAN METODE SAE EBLUP DAN SAE HB PADA PENDUGAAN AREA KECIL (STUDI KASUS PENDUGAAN KEMISKINAN DI PROVINSI JAWA TIMUR)," *Jurnal Statistika dan Aplikasinya*, vol. 6, no. 1, pp. 96–108, Jun. 2022, doi: 10.21009/JSA.06109.
- [33] S. Brooks, A. Gelman, G. Jones, and X.-L. Meng, *HANDBOOK OF MARKOV CHAIN MONTE CARLO*. New York: Chapman and Hall/CRC, 2011. doi: 10.1201/b10905.
- [34] I. Ntzoufras, *BAYESIAN MODELING USING WINBUGS*. Wiley, 2009. doi: 10.1002/9780470434567.