Media Statistika

p-ISSN 1979 – 3693 e-ISSN 2477 – 0647
MEDIA STATISTIKA 16(1) 2023: 88-99
http://ejournal.undip.ac.id/index.php/media_statistika

BETA-BINOMIAL MODEL IN SMALL AREA ESTIMATION USING HIERARCHICAL LIKELIHOOD APPROACH

Etis Sunandi[1,2], Khairil Anwar Notodiputro[1], Indahwati[1], Agus M. Soleh[1]
[1] Department of Statistics, IPB University, Indonesia
[2] Statistics Program, Department of Mathematics, University of Bengkulu, Indonesia

e-mail: *khairil@apps.ipb.ac.id*

DOI: 10.14710/medstat.16.1.88-99

Article Info:
Received: 19 May 2023
Accepted: 1 December 2023
Available Online: 7 December 2023

Keywords:
Area Level; Binary Response; Illiteracy Rate; MSEP; Simulation; Small Sample

.

Abstract: Small Area Estimation is a statistical method used to estimate parameters in sub-populations with small or even no sample sizes. This research aims to evaluate the Beta-Binomial model's performance for estimating small areas at the area level. The estimation method used is Hierarchical Likelihood (HL). The data used are simulation data and empirical data. Simulation studies were used to investigate the proposed model. The estimator's Mean Squared Error of Prediction (MSEP) and Absolute Bias (AB) estimator values determine the best estimation criteria. An empirical study using data on the illiteracy rate at the sub-district level in Bengkulu Province. The results of the simulation study show that, in general, the parameter estimators are nearly unbiased. Proportion prediction has the same tendency as parameters. Finally, the HL estimator has a small MSEP estimator. The results of an empirical study show that the average illiteracy rate in Bengkulu province is quite diverse. Kepahiang District has the highest average illiteracy rate in Bengkulu Province in 2021.

1. INTRODUCTION

Small Area Estimation (SAE) is a statistical method for estimating parameters in a subpopulation with few or no samples. The development of SAE research is currently fairly rapid. Breidenbach et al. (2018) developed a small area estimation at the unit level and area level with the inhomogeneity constraint of the area of influence. The application of this study was carried out on remote sensing data. A small area quantitative estimator with a semiparametric density ratio model for the error distribution at the unit level nested error regression model has been developed by Chen & Liu (2019). Salvati et al. (2021) developed a small area estimation model on the data linkage. These data are mostly found in genetic data and informatics. Riley et al. (2021) developed a clinical prediction model and variable selection based on a small area model. This aimed to overcome overfitting because the technique reduces the predictor effect prediction towards zero and reduces the Mean Square Error Prediction (MSEP) on new subjects.

The SAE approach can be used with data that has either continuous or discrete response variables (Rao & Molina, 2015). One of the discrete data is binary data. Binary data is data that describes the number of "successful" events of an experiment. This data follows the Binomial Distribution, assuming that the random variables are independent and have the same (constant) "success" probability. However, it is not uncommon for assumptions to be violated in practice. One of the assumption violations is the problem of overdispersion in binary data observations. As stated by Morel & Nagaraj (1993) if it is assumed that the observations are independent with the same probability of success, the binomial model is suitable for analyzing the data. The grouping causes a correlation between the element units, and the probability of "success" in the group becomes unequal (varies). As a result, the data usually shows a more significant variance than the binomial model allows.

The problem of overdispersion that occurs in the observation of discrete data can be explained by two things: the correlation between random variables and the diversity of the probability of "success" on the random variables. Both are related events, so when there is a correlation between random variables, there will also be various "success" opportunities and vice versa. McCullagh & Nelder (1989) state that both events can occur because of groupings in the population. The emergence of overdispersion problems in discrete data will also impact the estimated value for variance. Where when the value of this variance estimator is used in calculating the confidence interval, it will produce a small average, causing a confidence interval that is too short. In addition, if the estimator value of this parameter is used for hypothesis testing, the results will tend to accept the null hypothesis. Overdispersion can be overcome and tested using a mixed distribution whose probability is not fixed but random and generated from other distributions. This mixed distribution has the potential as a modeling substance that uses data well (Wagner et al., 2015). One of the mixed distributions that can be used to overcome the problem of overdispersion in binomial data is the Beta-Binomial distribution (Harrison, 2015).

According to Lee & Nelder (2001), parameter estimation in a mixed model with a Beta-Binomial hierarchy can be done through Hierarchical Likelihood (HL). This method is claimed to be better than the Bayes approach analytically. The HL method can reduce the bias of estimating binary data parameters, which is a problem in the Generalized Linear Mixed Model (GLMM) (Yun & Lee, 2004).

Previous researchers, Rao and Molina (2015), developed the Beta-Binomial model in small area estimation through the Bayesian approach, both through Empirical Bayes (EB) and Hierarchical Bayes (HB). We have developed a Beta-Binomial model to estimate small areas (Sunandi et al., 2023). In this research, we are also still developing the Beta-Binomial model in small areas. The difference from previous research is that the estimation of the parameters of the fixed effect $\boldsymbol{\beta}_z$ and random variable $\mathbf{v}$ are done separately with the assumption of mutual independence. The results of this study are expected to provide solutions for scientific development in estimating binary data model parameters with small samples.

In accordance with the first Sustainable Development Goals point, we use socio-economic data as empirical data. The proposed model is applied to The National Socio-Economic Survey (SUSENAS), and Potential Village (PODES) data of Bengkulu Province. This province is one of the provinces on the island of Sumatra. according to the Statistics Indonesia-BPS report (2022), Bengkulu is the 2nd poorest province on the island of Sumatra (Statistics Indonesia, 2022). We aim to predict the illiteracy rate in Bengkulu province in 2021. The assumption in this study is that the illiteracy rate spreads following a Beta-Binomial distribution.

2. LITERATURE REVIEW

2.1 Small Area Estimation

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: $y_i = \boldsymbol{z}_{\boldsymbol{i}}'\boldsymbol{\beta}_{\boldsymbol{z}} + v_i + e_i,\ i = 1, \dots, m$. 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 (Skellam, 1948). 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:

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

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

$$\log\left(\frac{p_i}{1 - p_i}\right) = \xi_i = \boldsymbol{z}_{\boldsymbol{i}}'\boldsymbol{\beta}_{\boldsymbol{z}} + v_i$$

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 z_i of size $p \times 1$ is a constant covariate. The parameter $\boldsymbol{\beta}_z$ 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. Previous research using the Beta-Binomial model in SAE has been conducted by Yanuar et al., (2021). The estimation method used is the Bayes approach.

2.2 Hierarchical Likelihood

Estimating the parameters of the proposed model was done by maximizing the HL function and the Adjusted Profile Hierarchical Likelihood (APHL) function. Lee & Nelder (1996) extended GLMM to HGLM by accommodating a non-normal distribution for random effects. They use HL to avoid complex integration. This method uses the logarithm of the probability density function with two variables $(\boldsymbol{y}, \boldsymbol{v})$ (Wu & Bentler, 2012). HL is a probability function with Henderson (Lee & Nelder, 1996, 2001) 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 Lee et al. (2006). The HL formula can be written as follows:

$$h = h_o + h_1 = \log l_1(\boldsymbol{\beta}, \phi; \boldsymbol{y}|\boldsymbol{v}) + \log l_1(\lambda; \boldsymbol{v}) \tag{2}$$

where $l_1(\boldsymbol{\beta}, \phi; \boldsymbol{y}|\boldsymbol{v})$ and $l_2(\lambda; \boldsymbol{v})$ are conditional probability functions $\boldsymbol{y}|\boldsymbol{v}$ and probability functions $\boldsymbol{v}$. Vector $\boldsymbol{\beta}$ is a canonical parameter, (ϕ, λ) denotes the dispersion parameter, and λ is the distribution parameter $\boldsymbol{v}$. Let $\boldsymbol{v}(.)$ be the corresponding link function that defines HL such that $\boldsymbol{v} = \boldsymbol{v}(\boldsymbol{u})$. In the HL form, the selection of the random effect scale is essential in that the scale change requires Jacobian adjustment. Suppose that $\boldsymbol{v} = \boldsymbol{v}(\boldsymbol{u})$ is used to show the scale in which the random effect of $\boldsymbol{v}$ is assumed to be linear, and the predictor is linear. Using the score function of HL, the parameters $\boldsymbol{\beta}$ and $\boldsymbol{v}$ are estimated as follows (Jiang, 2007):

$$\frac{\partial h}{\partial \boldsymbol{\beta}} = \boldsymbol{0}, \frac{\partial h}{\partial \boldsymbol{v}} = \boldsymbol{0} \tag{3}$$

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

$$p_{\boldsymbol{\beta},\boldsymbol{v}}(h) \tag{3}$$

$$p_{\boldsymbol{\beta},\boldsymbol{v}}(h) = \left(h + \frac{1}{2}\log(|2\pi \boldsymbol{H}^{-1}|)\right)\Big|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}},\boldsymbol{v}=\widehat{\boldsymbol{v}}} \tag{4}$$

Where $\boldsymbol{H} = \begin{bmatrix} \boldsymbol{X}'\boldsymbol{W}_1\boldsymbol{X} & \boldsymbol{X}'\boldsymbol{W}_1\boldsymbol{Z} \\ \boldsymbol{Z}'\boldsymbol{W}_1\boldsymbol{X} & \boldsymbol{Z}'\boldsymbol{W}_1\boldsymbol{Z} + \boldsymbol{W}_2 \end{bmatrix}$, $\boldsymbol{W}_1 = \left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\eta}}\right)^2 \left(\phi V(\boldsymbol{\mu})\right)^{-1}$, and $\boldsymbol{W}_2 = -\frac{\partial^2 h_1}{\partial \boldsymbol{v}^2}$.

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

$$\hat{p}_i^{HL} = \frac{\exp\left(\boldsymbol{z}_i'\widehat{\boldsymbol{\beta}}_{\boldsymbol{z}} + \hat{v}_i\right)}{1 + \exp\left(\boldsymbol{z}_i'\widehat{\boldsymbol{\beta}}_{\boldsymbol{z}} + \hat{v}_i\right)} \tag{5}$$

3. MATERIAL AND METHOD

3.1. Data

This research uses simulation data and empirical data. The simulation data will be generated from distributions corresponding to a predetermined scenario. In the simulation study, the response variable data will be generated $y_i \sim BB(n_i, p_i, \phi)$, $i = 1,2, \dots, m = 50$. Initial values for the covariates and parameters are considered based on previous research (Bell et al., 2019; Burgard et al., 2020, 2021; Najera-Zuloaga et al., 2019).

The simulation study applied the SAE model with a variance for the random effect of $\sigma_v^2 = 2$. The value follows the research of Ybarra & Lohr (2008). Meanwhile, fixed effect parameters are given as three variables ($p = 3$) with initial values: $\boldsymbol{\beta}_Z \epsilon (1,1,1)$ and $\beta_0 = 0$. The effect of the three fixed effect parameters is assumed to be identical. The sample size of each area is specified as $n_i = 50$. This determination assumes that each area has the same sample size. the number of areas is specified as $m = 50$. The number of samples is determined as $n_i = 50$. The determination of this value is based on the minimum number of areas (sub-districts) in Indonesia. The initial value of the overdispersion parameter is $\phi = 1.5$. This is in accordance with the theory that overdispersion conditions are fulfilled if $\phi > 1$.

The empirical data comes from Statistics of Bengkulu Province, The National Socio-Economic Survey (SUSENAS), and Potential Village (PODES) 2021. The response variable comes from SUSENAS data. The auxiliary variables come from PODES data. This study made predictions on the proportion of illiteracy at the sub-district level in Bengkulu Province in 2021. The variables used can be seen in Table 1.

Table 1. The Research Variables

	Variable	Code
Response	The proportion of illiteracy	BBH
Auxiliary Variables	Average recipient letter description no able (SKTM)	Z1
	The number of malnutrition	Z2
	The number of facility health	Z3
	The number of activities eradicating illiteracy	Z4
	The number of facility education	Z5

3.2. Simulation Procedure

The summary of the simulation algorithm is shown in Figure 1. The details of the simulation algorithm are as follows:

1. Set the number of areas $m = 50$ and $n_i = 50$
2. Set the parameter value $\boldsymbol{\beta_z}$ size $p \times 1, \sigma_v^2 = 2,\ \alpha$ *and* γ according to the simulation scenario.
3. Generate covariates in the i-th area, $\boldsymbol{z_i}$ sized $p \times 1$ as covariates from a normal multivariate distribution with the expected value vector $-2\mathbf{1}_3$ and variance matrix $\boldsymbol{I_3}$.
4. Generate $v_i \sim N(0, \sigma_v^2)$
5. Set overdispersion parameter $\phi = 1.5$.
6. Calculate $\xi_i = logit\ (p_i) = \boldsymbol{z_i'\beta_z} + v_i$.
7. Calculate proportion parameters in the i-th area, k-th census block, and r-th household $p_i = \frac{\exp(\boldsymbol{z_i'\beta_z}+v_i)}{1+\exp(\boldsymbol{z_i'\beta_z}+v_i\ )}$
8. Generate the number of successful events in area-i, k-th census block, r-th household $y_i \sim BB(n_i, p_i, \phi)$
9. Estimating model parameters
   a. Determine the likelihood function $h\ =\ \ell(\boldsymbol{\beta_z}, \phi, \sigma_v^2 | y_i, v_{i,}, \boldsymbol{z_i})$
   b. Determine the Adjusted Profile H-Likelihood (APHL).
   c. Maximize h and APHL with Delta algorithm to estimate parameters
   d. $\hat{\varphi}^{HL}\ =\ (\boldsymbol{\widehat{\beta}_z}, \boldsymbol{\hat{v}}\ , \hat{\sigma}_v^2)$
10. Prediction of model proportion estimator using formula $\hat{p}_i^{HL} = \frac{\exp(\boldsymbol{z_i'\widehat{\beta}_z}+\hat{v}_i)}{1+exp(\boldsymbol{z_i'\widehat{\beta}_z}+\hat{v}_i)}$.

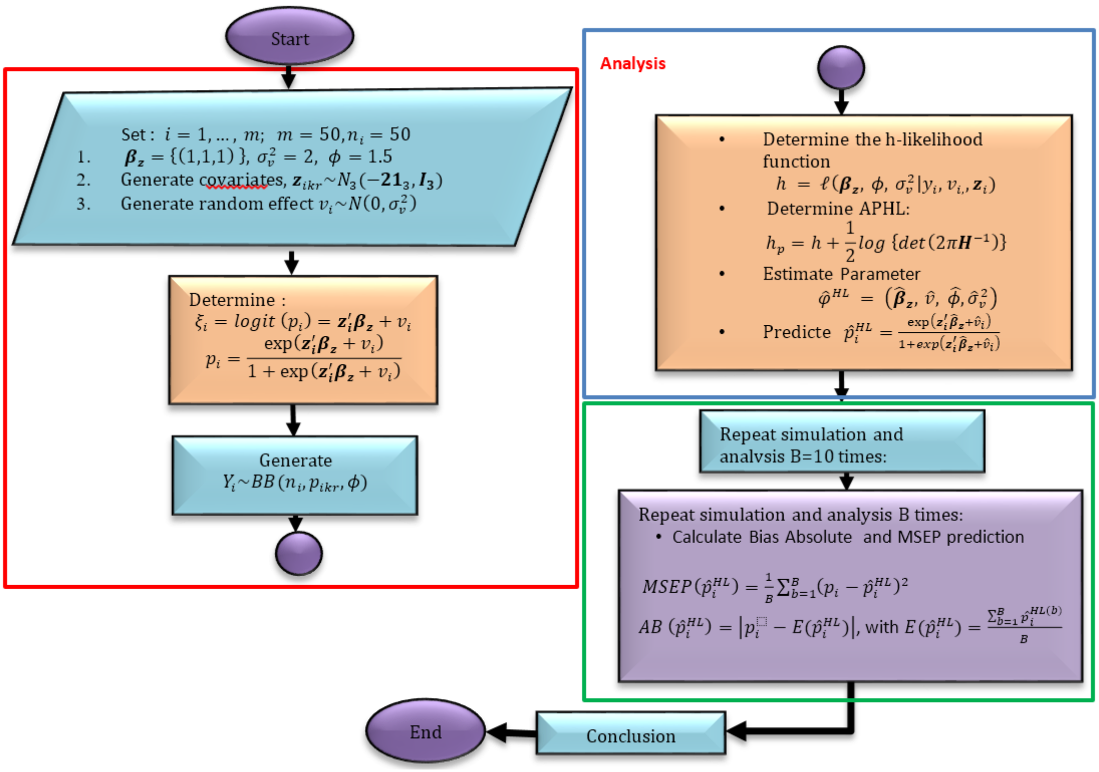

Figure 1. Simulation Algorithm

The measures of goodness of fit used in this research are the Mean Squared Error of Prediction (MSEP) and Absolute Bias (AB). The MSEP value is obtained from the average squared difference between the actual and predicted data. At the same time, the absolute bias parameter estimator is obtained from the difference between the parameter and the expected value. The MSEP measure is used to assess the accuracy of model predictions. Meanwhile, Absolute Bias is used to assess the goodness of precision of the built model parameter estimator. The simulation was repeated ten times. The formula of MSEP and AB is as follows

$$MSEP(\hat{p}_i^{HL}) = \frac{1}{B}\sum_{b=1}^{B}(p_i - \hat{p}_i^{HL})^2 \tag{6}$$

$$AB\ (\hat{p}_i^{HL}) = \left|p_i \ - E(\hat{p}_i^{HL})\right|, \text{ with } E(\hat{p}_i^{HL}) = \frac{\sum_{b=1}^{B}\hat{p}_i^{HL(b)}}{B} \tag{7}$$

4. RESULTS AND DISCUSSION

The simulation is designed to assess the performance of the proposed Model. In this study, a simulation was carried out with a scenario iteration of 13 which was repeated 10 times. Computing programming using the R program. The syntax is arranged from Najera's modified syntax Najera-Zuloaga et al., (2019).

The simulation results of the estimation of fixed effect parameters through HL are summarized in Table 2 and Figure 2. Based on the table, it can be seen that $\widehat{\boldsymbol{\beta}}_{\boldsymbol{Z}} = (0.237, 1.129, 1.383, 1.751)$. It is known that the initial value of the fixed effect parameter is $\boldsymbol{\beta}_{\boldsymbol{Z}} = (0,1,1,1)$. It appears that $\widehat{\boldsymbol{\beta}}_{\boldsymbol{Z}}$ is slightly biased. Based on the boxplot in Figure 2, the fixed effect estimators tend to underestimate. At the same time, the dispersion parameter on the area random variable is $\hat{\sigma}_v^2 = 1.3$. When compared with the initial value of the dispersion parameter set, $\sigma_v^2 = 2$, it can also be said that the assumption is biased. The overdispersion parameter estimator is $\hat{\phi}$ = 1.7. It shows that the model can detect that the data variance is more significant than the variance model. This condition follows the condition of the population.

Table 2. Fixed Effects Coefficients

	Estimate	Std. Error	t.value	p.value	Sig.
(Intercept)	0.237	0.400	0.594	0.553	
z1	1.129	0.388	2.909	0.004	**
z2	1.383	0.425	3.252	0.001	**
z3	1.751	0.630	2.780	0.005	**

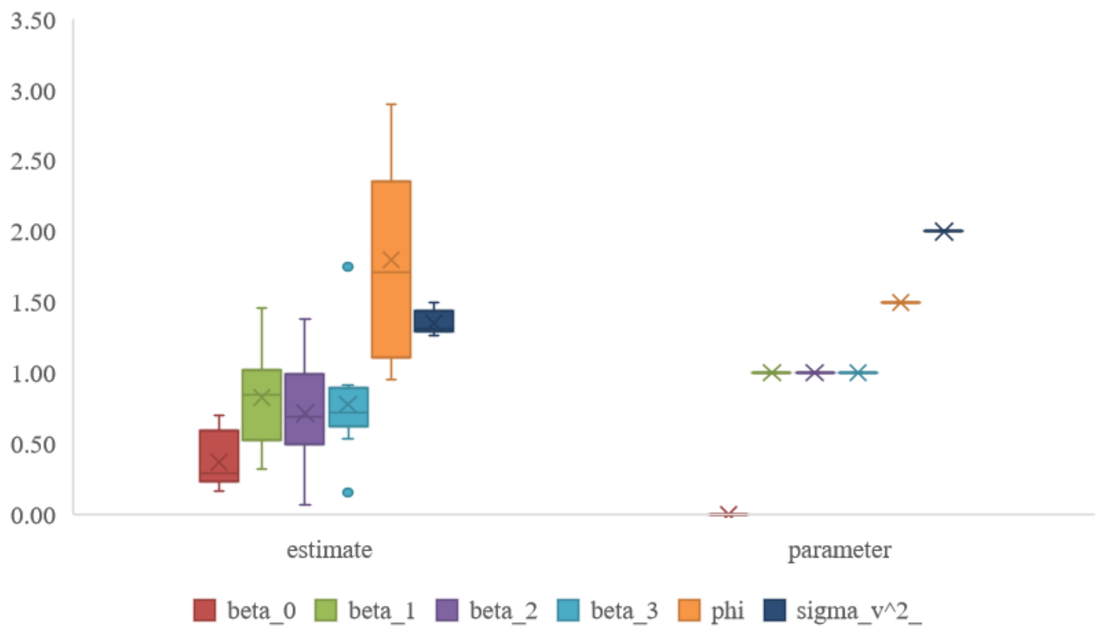

Figure 2. Comparison Between Estimates and Parameters

The results of this study follow what was stated in previous studies: in binary data analysis, the joint maximization $(\boldsymbol{v}, \boldsymbol{\beta})$ of the $\boldsymbol{h}$ function is biased in the estimation $\boldsymbol{\beta}$. The estimation can be overcome using $\boldsymbol{p}_{\boldsymbol{v}}(\boldsymbol{h})$ (Ha et al., 2017). Research using binary and HL models has been carried out by (Lokonon & Senou, 2020; Youngjo & Kim, 2020). The researchers showed that the second-order Laplace approach helps increase accuracy in estimating dispersion parameters. First-order corrections are often not frequent enough for non-Gaussian random effects, and second-order can reduce bias. Researchers (Lee & Nelder, 2001) proposed the use of the second-order Laplace Approach.

Table 3. Summary of Statistics and Measures of the Goodness of Fit the Model

	p_i	$\hat{p}_i^{HL}$	$AB(\hat{p}_i^{HL})$	$MSEP(\hat{p}_i^{HL})$
Minimum	0.002	0.003	0.001	0.043
Maximum	0.970	0.912	0.720	
Mean	0.328	0.274	0.137	
Median	0.212	0.195	0.071	

This study also carried out the proportion prediction in a small area. A summary of the comparison of the estimate and parameter of proportion is presented in Table 3. These results show that the proportion estimate of $(\hat{p}_i^{HL})$ is almost the same as the parameter. In addition, the table also shows that the absolute value of the bias $(\hat{p}_i^{HL})$ and MSEP $(\hat{p}_i^{HL})$ is relatively tiny. The mean of MSEP is 0.043 and the median of AB is 0.071. It can be said that this model works well. In Figure 3, the proportion (pi^{HL}) estimate from several areas has the same tendency. In other words, the proposed model has good flexibility, and it can be seen from the plot of the predictions that it can follow the distribution pattern of the observation data.

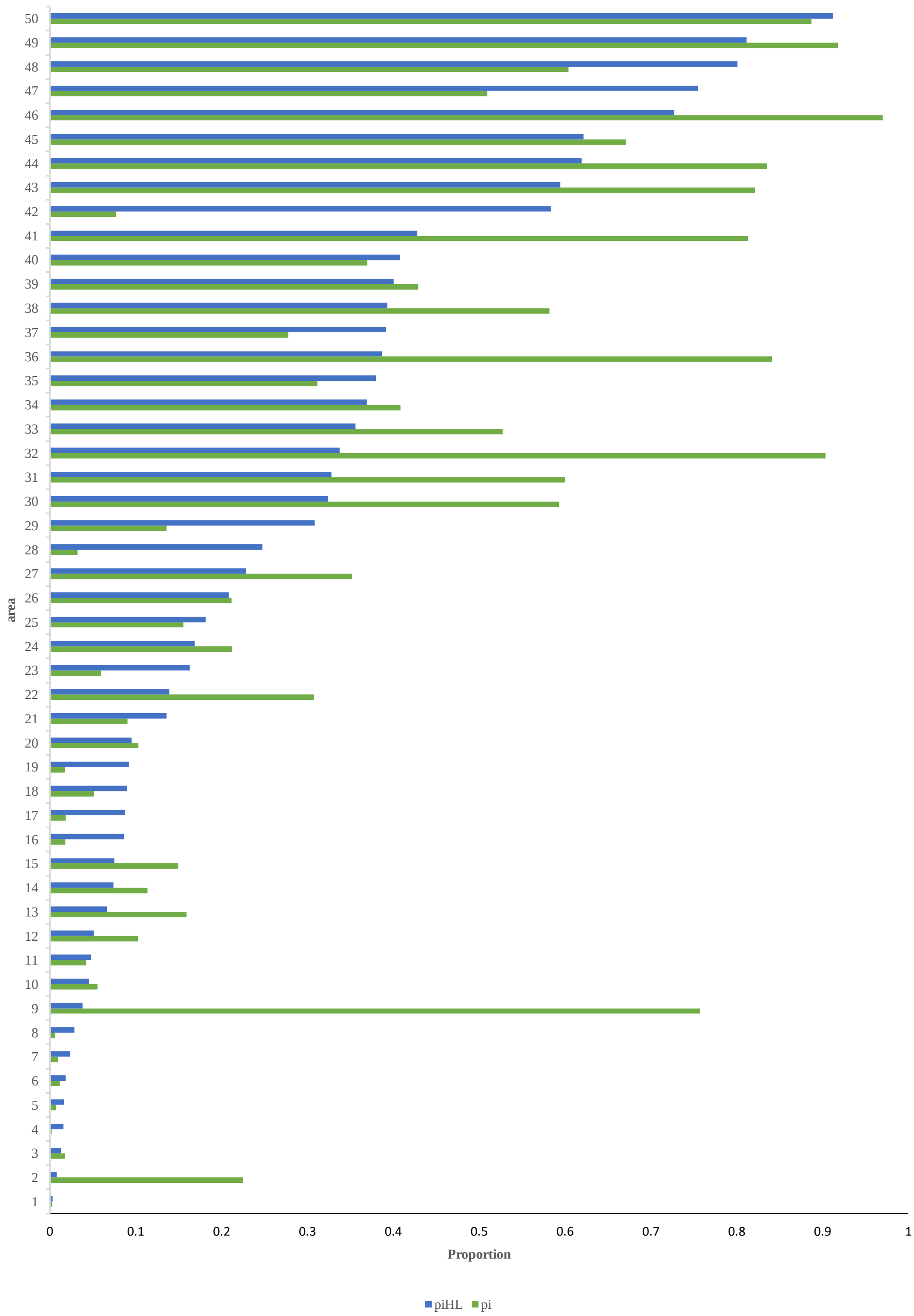

Figure 3. Comparison of Parameters and Estimators of HL Proportion

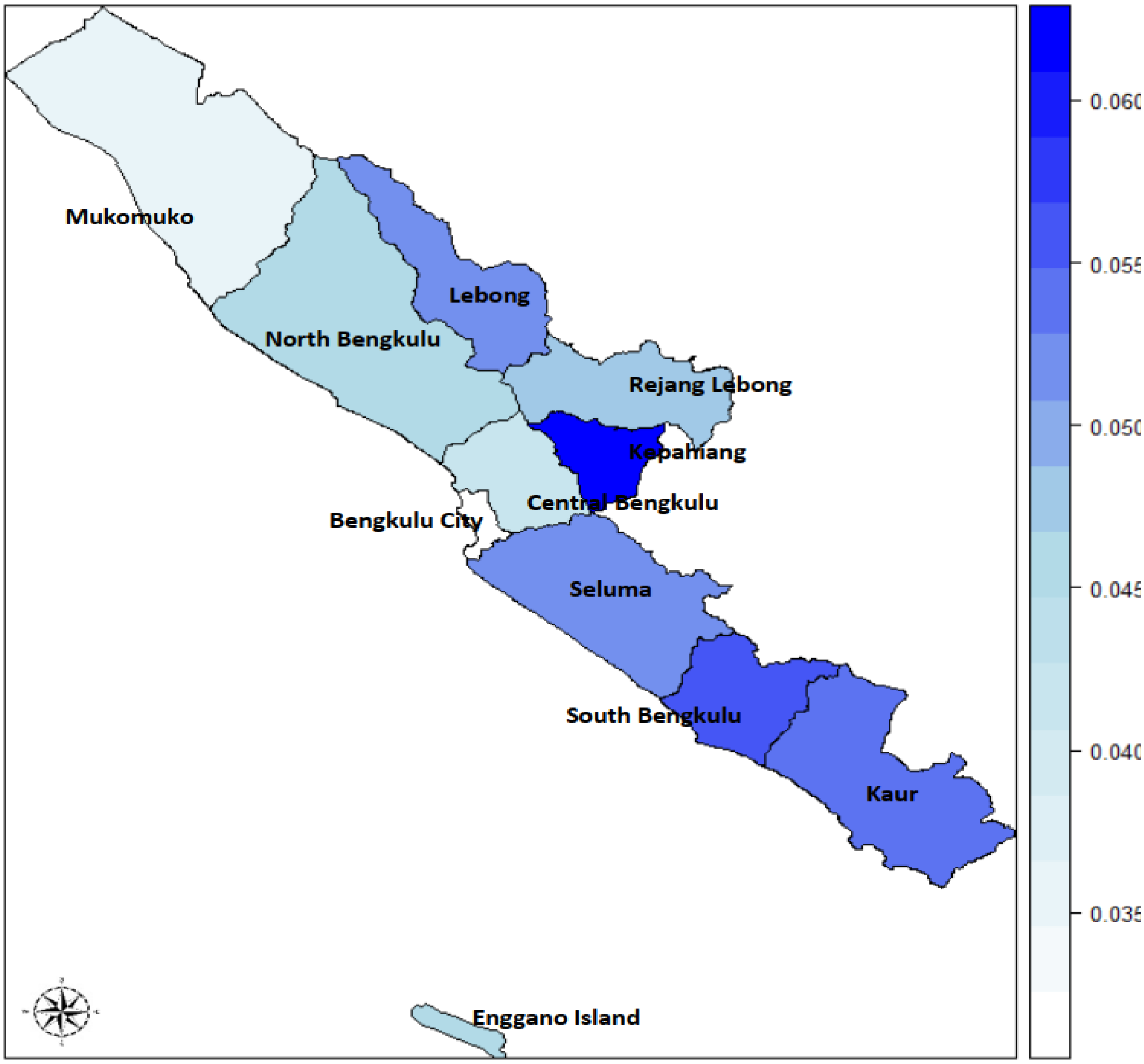

Figure 4. The Map Distribution of Illiteracy Rate Prediction

The application of the proposed model to the data of the illiteracy rate per sub-district in the province has been conducted. The distribution map of the predicted illiteracy rate per district through the proposed model can be seen in Figure 4. Kepahiang Regency has the largest mean illiteracy rate in Bengkulu Province, whereas the smallest mean illiteracy rate is located in Bengkulu City. On the other hand, some sub-districts are predicted to have no illiterate population by direct estimation. However, the proposed model predicts that the illiterate population is present in all sub-districts in Bengkulu Province. In general, the estimation results of the proposed model and the direct estimator have similar trends. For example, the direct estimator also predicts that Kota Bengkulu has fewer illiterate residents than the other regencies.

5. CONCLUSION

This study investigated the performance of Beta-Binomial Model. The results of the estimation of fixed effect parameters, $\widehat{\boldsymbol{\beta}}_Z$, tend to be almost unbiased. The dispersion parameter on the random effect variable is biased. The overdispersion parameter estimator is just about unbiased. It shows that the model can detect that the variance of data is more significant than the variance of the model. The proportion estimator ($\hat{p}_i^{HL}$) is near the parameter. The absolute bias ($\hat{p}_i^{HL}$) and MSEP ($\hat{p}_i^{HL}$) are both fairly small. The model works quite well. The proportion estimator ($\hat{p}_i^{HL}$) from various places has the same trend. In conclusion, the proposed model is flexible, and the prediction plot shows that it can adapt to the distribution shape of the observation data. Based on empirical studies, Kepahiang Regency has the highest average illiteracy rate. On the other hand, Bengkulu City is the region that has the lowest average illiteracy rate in Bengkulu province in 2021.

ACKNOWLEDGMENT

We would like to thank all those who have helped provide valuable suggestions for improving the quality of this paper. We also would like thank to Statistics Indonesia (BPS) for supporting data research. Our gratitude to the Ministry of Education and Culture of the Republic of Indonesia and The Directorate General of Higher Education, Research, and Technology for National Postgraduate Education Scholarships (BPP-DN) 2019 and funding the 2022 Doctoral Dissertation Research No. 3796/IT3.L1/PT.01.03/P/B/2022.

REFERENCES

Bell, W. R., Chung, H. C., Datta, G. S., & Franco, C. (2019). Measurement Error in Small Area Estimation: Functional Versus Structural Versus Naïve Models. *Survey Methodology*, *45*(1), 61–80.

Breidenbach, J., Magnussen, S., Rahlf, J., & Astrup, R. (2018). Unit-Level and Area-Level Small Area Estimation under Heteroscedasticity Using Digital Aerial Photogrammetry Data. *Remote Sensing of Environment*, *212*. https://doi.org/10.1016/j.rse.2018.04.028

Burgard, J. P., Esteban, M. D., Morales, D., & Pérez, A. (2020). A Fay–Herriot Model when Auxiliary Variables are Measured with Error. *Test*, *29*(1), 166–195. https://doi.org/10.1007/s11749-019-00649-3

Burgard, J. P., Esteban, M. D., Morales, D., & Pérez, A. (2021). Small Area Estimation Under A Measurement Error Bivariate Fay–Herriot Model. *Statistical Methods and Applications*, *30*(1), 79–108. https://doi.org/10.1007/s10260-020-00515-9

Chen, J. & Liu, Y. (2019). Small Area Quantile Estimation. *International Statistical Review*, *87*(S1). https://doi.org/10.1111/insr.12293

Ha, l Do, Jeong, J.-H., & Lee, Y. (2017). *Statistical Modelling of Survival Data with Random Effects H-Likelihood Approach* (M. Gail, J. M. Samet, B. Singer, & A. Tsiatis (eds.)). Springer.

Harrison, X. A. (2015). A Comparison of Observation-Level Randomeffect and Beta-Binomial Models for Modelling Overdispersion in Binomial Data in Ecology & Evolution. *PeerJ*, *2015*(7). https://doi.org/10.7717/peerj.1114

Jiang, J. (2007). Linear and Generalized Linear Mixed Models and Their Applications. In *Linear and Generalized Linear Mixed Models and Their Applications*. Springer. https://doi.org/10.1007/978-0-387-47946-0

Lee, Y., & Nelder, J. A. (1996). Hierarchical Generalized Linear Models. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(4). https://doi.org/10.1111/j.2517-6161.1996.tb02105.x

Lee, Y., & Nelder, J. A. (2001). Hierarchical Generalised Linear Models: A Synthesis Of

Generalised Linear Models, Random-Effect Models And Structured Dispersions. *Biometrika*, *88*(4). https://doi.org/10.1093/biomet/88.4.987

Lee, Y., Nelder, J. A., & Pawitan, Y. (2006). Generalized Linear Models with Random Effects: Unified Analysis via H-Likelihood. In *Generalized Linear Models with Random Effects: Unified Analysis via H-Likelihood*. https://doi.org/10.1111/j.1467-985x.2007.00485_4.x

Lokonon, B. E. & Senou, M. (2020). Effect of Misspecification of Random Effects Distribution on The Performance of Parameters Estimation Methods in Binary Logistic Mixed Models. *Afrika Statistika*, *15*(1). https://doi.org/10.16929/as/2020.2247.156

McCullagh, P. & Nelder, J. A. (1989). Generalized Linear Models. In *Chapman and Hall*. https://doi.org/10.1002/bimj.4710290217

Morel, J. G. & Nagaraj, N. K. (1993). A Finite Mixture Distribution for Modelling Multinomial Extra Variation. *Biometrika*, *80*(2). https://doi.org/10.1093/biomet/80.2.363

Najera-Zuloaga, J., Lee, D. J., & Arostegui, I. (2019). A Beta-Binomial Mixed-Effects Model Approach for Analysing Longitudinal Discrete and Bounded Outcomes. *Biometrical Journal*, *61*(3), 600–615. https://doi.org/10.1002/bimj.201700251

Rao, J. N. K. & Molina, I. (2015). Small Area Estimation: Second Edition. In *Small Area Estimation: Second Edition*. https://doi.org/10.1002/9781118735855

Riley, R. D., Snell, K. I. E., Martin, G. P., Whittle, R., Archer, L., Sperrin, M., & Collins, G. S. (2021). Penalization and Shrinkage Methods Produced Unreliable Clinical Prediction Models Especially when Sample Size was Small. *Journal of Clinical Epidemiology*, *132*. https://doi.org/10.1016/j.jclinepi.2020.12.005

Salvati, N., Fabrizi, E., Ranalli, M. G., & Chambers, R. L. (2021). Small Area Estimation with Linked Data. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, *83*(1). https://doi.org/10.1111/rssb.12401

Skellam, J. G. (1948). A Probability Distribution Derived from the Binomial Distribution by Regarding the Probability of Success as Variable between the Sets of Trials. *Journal of the Royal Statistical Society: Series B (Methodological)*, *10*(2). https://doi.org/10.1111/j.2517-6161.1948.tb00014.x

Statistics Indonesia, B. (2022). *Statistik Indonesia, Statistical Yearbook of Indonesia 2022*. Badan Pusat Statistik/BPS-Statistics Indonesia.

Sunandi, E., Notodiputro, K. A., Indahwati, & Soleh, A. M. (2023). Small Area Estimation of Illiteracy Rates based on Beta-Binomial Model using Hierarchical Likelihood Approach. *Mathematics and Statistics*, *11*(3), 579–585. https://doi.org/10.13189/ms.2023.110315

Wagner, B., Riggs, P., & Mikulich-Gilbertson, S. (2015). The Importance of Distribution-Choice In Modeling Substance Use Data: A Comparison of Negative Binomial, Beta Binomial, and Zero-Inflated Distributions. *American Journal of Drug and Alcohol Abuse*, *41*(6). https://doi.org/10.3109/00952990.2015.1056447

Wu, J. & Bentler, P. M. (2012). Application of H-likelihood to Factor Analysis Models with Binary Response Data. *Journal of Multivariate Analysis*, *106*. https://doi.org/10.1016/j.jmva.2011.09.007

Yanuar, F., Fajriyah, R., & Devianto, D. (2021). Small Area Estimation Method with Empirical Bayes Based on Beta Binomial Model in Generated Data. *Media Statistika*, *14*(1), 1–9. https://doi.org/https://doi.org/10.14710/medstat.14.1.1-9

Ybarra, L. M. R. & Lohr, S. L. (2008). Small Area Estimation when Auxiliary Information is Measured with Error. *Biometrika*, *95*(4), 919–931. https://doi.org/10.1093/biomet/asn048

Youngjo, L., & Kim, G. (2020). Properties of h-Likelihood Estimators in Clustered Data. *International Statistical Review*, *88*(2). https://doi.org/10.1111/insr.12354

Yun, S. & Lee, Y. (2004). Comparison of Hierarchical and Marginal Likelihood Estimators for Binary Outcomes. *Computational Statistics and Data Analysis*, *45*(3). https://doi.org/10.1016/S0167-9473(03)00033-1