Available online at http://scik.org
J. Math. Comput. Sci. 2022, 12:129
https://doi.org/10.28919/jmcs/7233
ISSN: 1927-5307

THE IMPLEMENTATION OF EMPIRICAL BEST LINEAR UNBIASED PREDICTION-FAY HERRIOT (EBLUP-FH) ON THE ESTIMATION OF AVERAGE PER CAPITA EXPENDITURE AT DISTRICT LEVEL IN WEST SUMATRA PROVINCE IN 2019

ARMALIA DESIYANTI[1,*], TONI TOHARUDIN[2], YUSEP SUPARMAN[2]

[1]Post-graduate Program in Applied Statistics, Faculty of Mathematics and Natural Sciences,
Universitas Padjadjaran, Indonesia

[2]Department of Statistics, Faculty of Mathematics and Natural Sciences,
Universitas Padjadjaran, Indonesia

Abstract: The problem of equitable distribution of welfare is still a sustainable development agenda that must be completed by all parties, including local governments. An indicator to describe the welfare of the population in an area is per capita income and per capita expenditure. The lack of data given by Statistics Indonesia has caused the government to not be optimal in policymaking and implementation because they require data presentation at a smaller regional level. This study aims to estimate per capita expenditure at the district level in West Sumatra Province in 2019 using the SAE EBLUP-FH method. Based on the results of the EBLUP-FH estimation, the distribution of per capita expenditure in West Sumatra Province is very varied among districts. Several districts that are geographically close to the provincial capital area tend to have a higher average per capita expenditure than other areas. Based on the

*Corresponding author
E-mail address: armalia21001@mail.unpad.ac.id
Received February 4, 2022

comparison between the direct estimator and the EBLUP-FH estimator, it was found that the RRMSE value of the EBLUP-FH estimator was smaller than the direct estimator. It can be said that the EBLUP-FH method can provide more precise estimation results.

Keywords: SAE; EBLUP; RRMSE; per capita expenditure.

2010 AMS Subject Classification: 62F10.

1. INTRODUCTION

The sustainable development agenda contained in the Sustainable Development Goals (SDGs) was created to fulfil the world's demands in addressing sustainable development problems, including issues of inequality, poverty, and the general welfare of the population. One indicator that can describe the population's welfare in an area is the population's per capita income [1][2]. As a provider of official statistics, the Indonesia Statistics (BPS) provides data of the population's per capita income in Indonesia using a per capita expenditure approach. The Indonesia Statistics (BPS) defines per capita expenditure as the cost incurred for the consumption of all household members for a month divided by the number of household members that has been adjusted for purchasing power parity. The composition of household expenditures for food and non-food in general can be used to assess the level of the population's economic welfare. The lower composition of household expenditure on food indicates an improving level of welfare.

The indicator of per capita expenditure is obtained by the Indonesia Statistics (BPS) through a survey, namely the National Socio-Economic Survey (SUSENAS). This survey is carried out twice a year, namely in March, which is to estimate up to the district/city level, and in September, which is to estimate the national and provincial levels. However, all of the data up to the district level is not yet available due to the lack of samples from each area. This condition causes the government to be not optimal in policymaking and implementation. In fact, the development planning and evaluation carried out by the government requires the presentation of data at a smaller regional level.

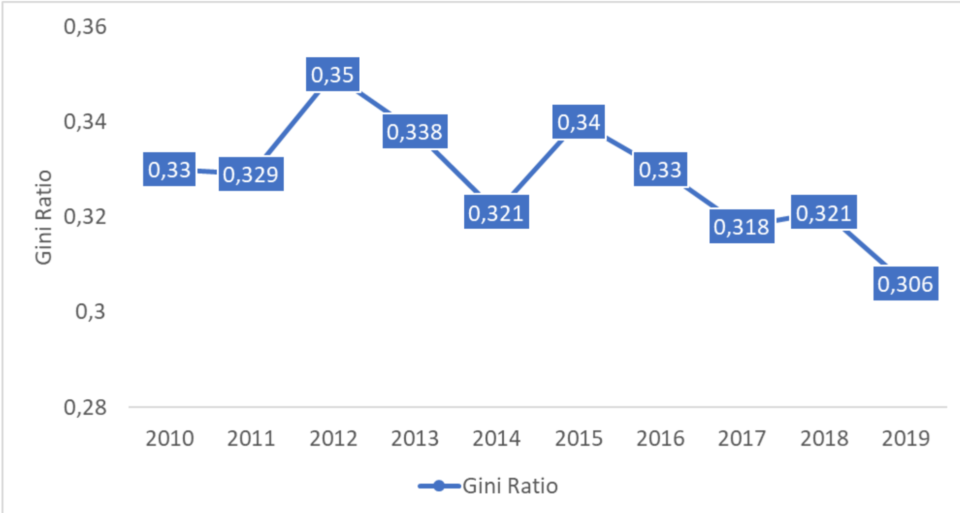
The presentation of data to the smallest level can be done by increasing the sample size, in which the statistics generated from this estimation are referred to as direct estimator. However, increasing the sample size will increase the cost of the survey as well. To overcome the problem of limited budget and resources, model-based estimator or indirect estimator can be used as an alternative solution. The Small Area Estimation (SAE) method is one of the most commonly used estimating methods[3][4].

To increase the effectiveness of the available sample size, the SAE estimator method was carried out by borrowing the power of the auxiliary variables associated with the predicted variables. The basic model of the Empirical Best Linear Unbiased Predictor (EBLUP) is one of the SAE models commonly used[5][6].

There have been many studies related to the implementation of the SAE method in estimating per capita expenditure. [7][8][9] apply SAE to estimate per capita expenditure using the Empirical Bayes method. [10][11][12] also estimates per capita expenditure using the Hierarchical Bayes method. The results of the studies show that the percentage of the population working in agriculture, the number of household members, the percentage of low-income families, the number of people currently attending school, the number of PLN (state electricity company) users, and population density have a significant effect on estimating per capita expenditure. In addition to the research of [8][11][13] also applies the Fay-Herriot method to SAE EBLUP by using a multivariate approach in estimating per capita expenditure per village in Bogor Regency. Based on the study results, it was found that the percentage of families using PLN and the percentage of agricultural families had a significant effect on the model.

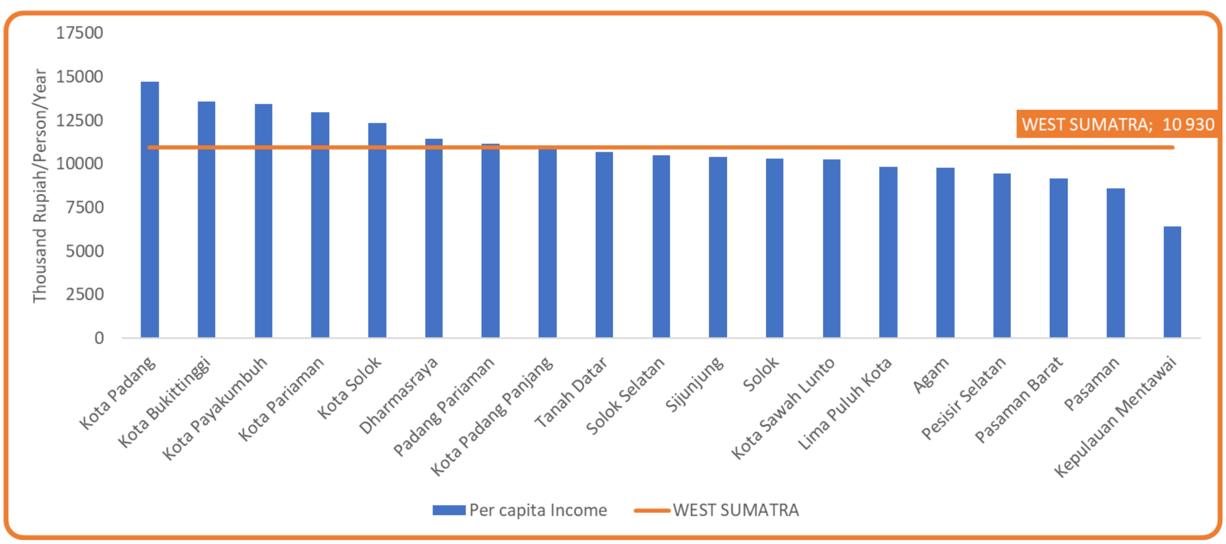
One of the problems of welfare and equity in Indonesia that needs to be evaluated is that the economic growth in West Sumatra Province over the last five years has shown a relatively stable value and is always above the national rate, which is always above 5 (five) percent. However, with the relatively stable economy, the Gini ratio coefficient in West Sumatra fluctuated every year during the period of 2010-2019 (Figure 1).

Figure 1: Gini Ratio of West Sumatra Province, 2010-2019



Viewed from the conditions of the last 20 years, the inequality occurring recently has actually increased from the Gini ratio coefficient value, which was only 0.28 in 1996 [14]. Per capita expenditure in West Sumatra in 2019 was ranked 17th in Indonesia, and its value has been relatively stable in the last five years. Based on Figure 2, it turns out that the adjusted per capita expenditure in West Sumatra Province has a relatively heterogeneous value between regencies/cities.

Figure 2: Adjusted per Capita Expenditure (Thousand Rupiah/Person/Year) in Regencies/City in West Sumatra Province in 2019



Based on the explanation of the background above, this study aims to analyze the comparison of the results of direct estimation and indirect estimation (EBLUP-FH) of per capita expenditure at the district level in West Sumatra Province in 2019 and analyze the distribution of per capita expenditure at the district level in West Sumatra Province in 2019 of the selected estimation results.

2. METHODOLOGY

2.1. Data Collection Method

The data used in this study was obtained from the March 2019 Susenas with a number of samples of 10,900 households spread over 175 districts out of a total of 179 districts in all regencies/cities in West Sumatra. Meanwhile, the auxiliary variables used in this study were obtained from PODES 2019. There were 50 proposed auxiliary variables used in this small area estimation (SAE). However, only 6 selected auxiliary variables will be used in this study based on backward elimination results.

Table 1: The Variables Used in The Study

No	Variable	Data	Source
1	Y	District level per capita expenditure	SUSENAS 2019
		HOUSING AND NEIGHBORHOODS	
2	Non Listrik	Number of families who are not electricity users	PODES 2019
3	Non PLN	Number of families using non- state electricity	PODES 2019
		EDUCATION AND HEALTH	
4	SD	Number of Elementary schools	PODES 2019
5	Praktik dokter	Number of places where doctors practice	PODES 2019
6	Poliklinik	Number of polyclinics/treatment centers	PODES 2019
		ECONOMY	
7	Minimarket	Number of minimarkets/supermarkets	PODES 2019

The kinds of software used in this research are R-Studio and Microsoft Excel.

2.2. Analysis Method

Direct Estimation

A direct estimator is an estimator generated from an approach based on the application of a design-based sampling model [15][16]. However, direct estimators obtained from survey results will produce large standard errors when used to estimate small areas [17].

Indirect Estimation

Indirect estimation is an estimation carried out in an area using additional information from auxiliary variables of the variables that will be estimated through the right model [5].

Small Area Estimation (SAE)

[17] explains that there are two basic models in SAE, namely the basic area level model and the basic unit level model:

a. Basic Area Level Model

According to [17], this basic area level model is a model based on the availability of supporting data in the form of auxiliary variables that only exist for certain area levels. The basic model for this area level estimator is known as the Fay-Herriot. The model built is as follows:

Theorem 2.1. *Fay-Herriot model*

$$\theta_i = \boldsymbol{x}_i^T \boldsymbol{\beta} + z_i \boldsymbol{v}_i \quad , \quad i = 1, 2, \dots, m \tag{2.1}$$

Where the number of areas is denoted by m, $\boldsymbol{x}_i^T$ as a vector of auxiliary variables, which are assumed to be related to the parameter to be estimated (θ); $\boldsymbol{\beta}$ is a vector of model parameters measuring p x 1; z_i is a known constant, and $\boldsymbol{v}_i$ is a random effect of small area. This random effect of a small area is assumed to be distributed $N\ (0, \sigma_v^2)$.

Theorem 2.2. *Assuming that the direct estimator model $\hat{\theta}_i$ is available, the estimator $\hat{\theta}_l$ can be calculated by the following equation:*

$$\hat{\theta}_l = \theta_i + e_i\,,\ i = 1, 2, \dots, m \tag{2.2}$$

Theorem 2.3. *The model based on equations (1) and (2) will produce a combined model (Mixed Model) as follows:*

$$\hat{\theta}_l = \boldsymbol{x}_i^T \boldsymbol{\beta} + b_i \boldsymbol{v}_i + e_i \quad , \quad i = 1, 2, \dots, m \tag{2.3}$$

b. Basic Unit Level Model

Theorem 2.4. *This basic area level model is a model based on the availability of supporting data in the form of auxiliary variables that only exist for certain unit levels (survey respondents). The model built is as follows:*

$$\theta_{ij} = \boldsymbol{x}_{ij}^T \boldsymbol{\beta} + \boldsymbol{v}_i + e_{ij} \quad , \quad i = 1, 2, \dots, m \ \text{ dan } \ j = 1, 2, \dots, n_i \tag{2.4}$$

Where the number of areas is denoted by m and the number of households in the ith area is denoted by *j*. The model used in this study is a basic area level model, namely at the district level.

Empirical Best Linear Unbiased Prediction-Fay Herriot (EBLUP-FH)

Theorem 2.5. *The basic area level model is a development of the GLMM model, which is known as the Fay-Herriot model* [18] *for a simple area level. The models used are:*

$$y_i = \theta_i + e_i = \boldsymbol{x}_i^T \boldsymbol{\beta} + \boldsymbol{u}_i + e_i \qquad , \; i = 1, 2, \ldots, m \tag{2.5}$$

Theorem 2.6. *The assumption of resolution of the above equation is that* $\boldsymbol{u}_i$ *and* e_i *are independent of each other and* σ_v^2 *is known. Based on the above equation, the BLUP estimator of* θ_i *will be:*

$$\hat{\theta}_i^{BLUP} = \boldsymbol{x}_i^T \tilde{\boldsymbol{\beta}} + \gamma_i \left(\hat{\theta}_i - \boldsymbol{x}_i^T \tilde{\boldsymbol{\beta}}\right) \tag{2.6}$$

Where $\gamma_i = \frac{\sigma_u^2}{\sigma_u^2 + \psi_i}$

The assumption that must be resolved by the calculation technique above is that the variance component of the random effect in the linear mixed model is known, so it is necessary to estimate the variance of the random effect first.

[19] state that to estimate the variance component of the random effect, several methods can be used, including the maximum likelihood method and the restricted maximum likelihood (REML) method. It was further explained that one of the drawbacks of using the maximum likelihood method in obtaining estimates for σ_u^2 is that it does not consider the loss of degrees of freedom due to estimating $\boldsymbol{\beta}$ with $\widehat{\boldsymbol{\beta}}$. This condition causes the estimation of variance in the model when replacing $\boldsymbol{\beta}$ with its estimator $\widehat{\boldsymbol{\beta}}$ to be biased.

Theorem 2.7. *According to [19], the REML method can solve the drawbacks. The use of data transformed in the REML method can consider the loss of degrees of freedom. Thus, the new estimator obtained by the REML method is as follows:*

$$\hat{\theta}_i^{EBLUP} = \hat{\gamma}_i + (1 - \hat{\gamma}_i)\, \boldsymbol{x}_i^T \tilde{\boldsymbol{\beta}} \tag{2.7}$$

Mean Square Error (MSE)

Theorem 2.8. *After obtaining the BLUP estimator value, which is denoted by* $\hat{\theta}_i^{BLUP}$, *then* *to see the goodness of the estimator value obtained, a measure called Mean Square Error (MSE) is used with the following equation:*

$$MSE\left(\hat{\theta}_i^{BLUP}\right) = g_{1i}(\sigma_v^2) + g_{2i}(\sigma_v^2) \tag{2.8}$$

with $g_{1i}(\sigma_v^2) = \frac{\sigma_v^2 \psi_i}{\sigma_v^2 + \psi_i} = \gamma_i \psi_i$ and $g_{2i}(\sigma_v^2) = (1-\psi_i)^2 \boldsymbol{x}_i^T \left[\sum_{i=1}^{m} \frac{x_i x_i^T}{\sigma_v^2 + \psi_i}\right]^{-1} \boldsymbol{x}_i$ and σ_v^2 value is not known.

Theorem 2.9. *Then, MSE for EBLUP is* [20]*:*

$$MSE\left(\hat{\theta}_i^{BLUP}\right) = g_{1i}(\sigma_v^2) + g_{2i}(\sigma_v^2) + 2g_{3i}(\sigma_v^2) \tag{2.9}$$

with $g_{3i}(\sigma_v^2) \approx \frac{\psi_i^2}{(\psi_i + \widehat{\sigma_v^2})}$ and $\widehat{\sigma_v^2}$ are obtained from REML estimation method.

According to [17], the comparison measure used to see the goodness of the estimator value obtained is the Relative Root Mean Square Error (RRMSE) value.

Theorem 2.10. *The estimator with a smaller RRMSE value is said to be better than the estimator with a larger RRMSE value. This RRMSE measure will be used to compare the direct estimation results and the EBLUP-FH estimator. The formula for calculating RRMSE is as follows:*

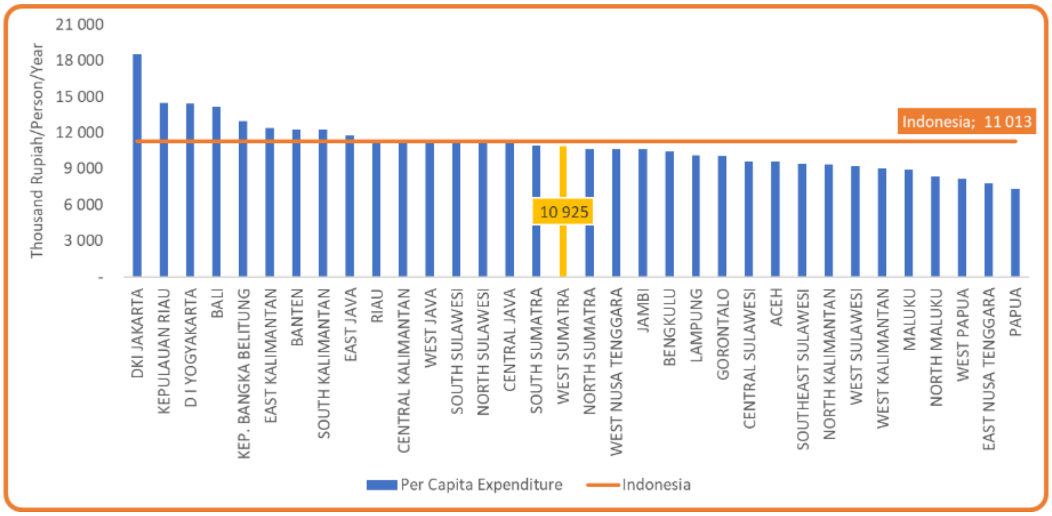
$$RRMSE\,(\theta) = \frac{\sqrt{MSE(\hat{\theta})}}{\hat{\theta}}\, x\ 100\ \% \tag{2.10}$$

3. RESULT AND DISCUSSION

Overview of Per Capita Expenditure of West Sumatra Province

In 2019, West Sumatra Province had an economic share of 1,53 percent, which was the 14th position of the Indonesian economy. Per capita expenditure in West Sumatra also shows a relatively stable and moderate value. In 2019, the per capita expenditure of West Sumatra Province was ranked 17th in Indonesia.

Figure 3: Adjusted per Capita Expenditure (Rupiah/Person/Year) per Province in Indonesia in 2019



The average per capita expenditure in West Sumatra is 10,9 million rupiah per capita per year. The highest per capita expenditure is in Padang city, reaching 14,7 million rupiah per capita per year. Meanwhile, the lowest per capita expenditure was found in the Mentawai Islands Regency, which reached 6,4 million rupiah per capita per year.

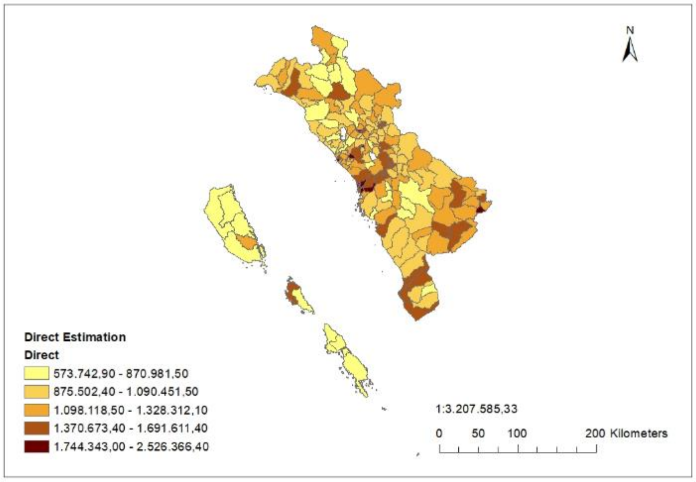
Direct Estimation

Direct estimation of expenditure per capita can only be carried out in areas sampled in the March 2019 Susenas. West Sumatra has a total of 179 districts, and only 175 districts were selected as sample. In this study, direct estimation was carried out at the district level. In direct estimation, in addition to getting an estimate of the studied variables, the estimation value for the sampling error variance is also obtained, which will be used in the EBLUP-FH estimation. In the March 2019 Susenas, the number of household samples selected for each district was different and relatively small, thus giving poor estimation results.

Based on the results of direct estimation, the districts with the highest average expenditure per capita per month are in the City of Padang, namely West Padang District (2,52 million rupiahs), East Padang District (2,06 million rupiahs), and North Padang District (1,99 million rupiah). On the other hand, the districts with the lowest average per capita expenditure per month are South

Pagai District, Mentawai Islands Regency (573 thousand rupiah), Tigo Lurah District, Solok Regency (580 thousand rupiah), and North Pagai District, Mentawai Islands Regency (635 thousand rupiah). The mapping of the direct estimation results can be seen in Figure 4. From Figure 4 we can see the distribution pattern of expenditure per capita at the district level as a result of direct estimation.

Figure 4: Mapping of Expenditure per Capita Direct Estimation Results



Selection of Auxiliary Variables and Testing the Normality of Direct Estimation

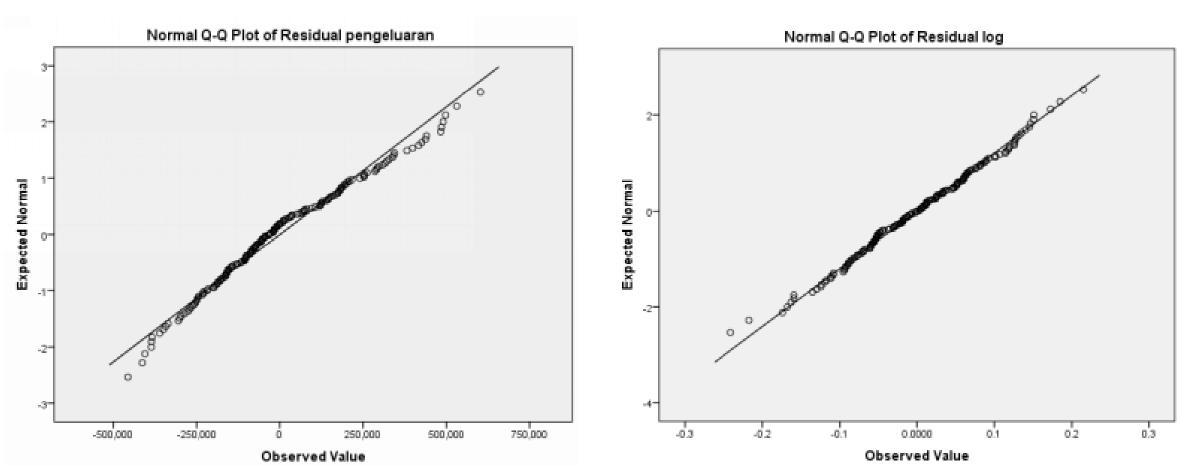
After obtaining the results of direct estimation of per capita expenditure at the district level, then the estimation is carried out using the EBLUP FH method. However, before estimating the EBLUP FH, the auxiliary variables must first be selected based on the correlation value and their significance to the direct estimation and testing the normality of the results of the direct per capita expenditure estimator.

Multicollinearity checking between auxiliary variables can be done by using Pearson Correlation between auxiliary variables. If the pairwise correlation coefficient of the two regressors/ auxiliary variables is very high, exceeding 0,8, then a serious multicollinearity problem occurs [21].

To ensure that multicollinearity violations do not occur between the auxiliary variables, another test is carried out using the Variance Inflation Factor (VIF). The higher the VIF value, the higher the collinearity of the X variable [21]. The absolute value of the Pearson correlation between the auxiliary variables is less than 0,8 and the VIF value is less than the number ten, which means that there is no multicollinearity between the auxiliary variables. The non-violation of the multicollinearity assumption indicates that the six auxiliary variables can be used at the modelling stage using EBLUP FH. After checking the multicollinearity between the auxiliary variables, the normality of the dependent variable was checked, namely expenditure per capita as a result of direct estimation using QQ-Plot.

Based on Figure 5, it can be seen that the distribution of the data resulting from the direct estimation of per capita expenditure has not been around the line, so it can be said that the per capita expenditure of the direct estimate has not met the assumption of normality. On the other hand, the distribution of data resulting from the logarithm transformation on per capita expenditure at the district level has coincided with the line so that it can be said that the per capita expenditure resulting from the logarithmic transformation is closer to the normal distribution. Because the logarithmic transformation data is used to overcome the data abnormality, then a reverse transformation will be carried out from the model to obtain the actual estimator for the mean value in each ith area, using the lognormal distribution.

Figure 5: QQ-Plot of Expenditure per Capita at District Level (a) Without Transformation (b) Logarithmic transformation



Indirect Estimation

The Fay-Herriot of EBLUP Estimator Model

The selection of auxiliary variables was carried out using the backward elimination method using a significance level of five percent, so that out of 50 proposed variables, six significant auxiliary variables were generated in the model. These variables include the number of non-electricity user families, the number of non-PLN electricity user families, the number of polyclinics/medical centers, the number of minimarkets/supermarkets, the number of SD/MI, and the number of doctor's practices. After obtaining the auxiliary variables, then the regression coefficient estimation is carried out on the six variables. The results of the regression coefficient estimation can be seen in Table 2.

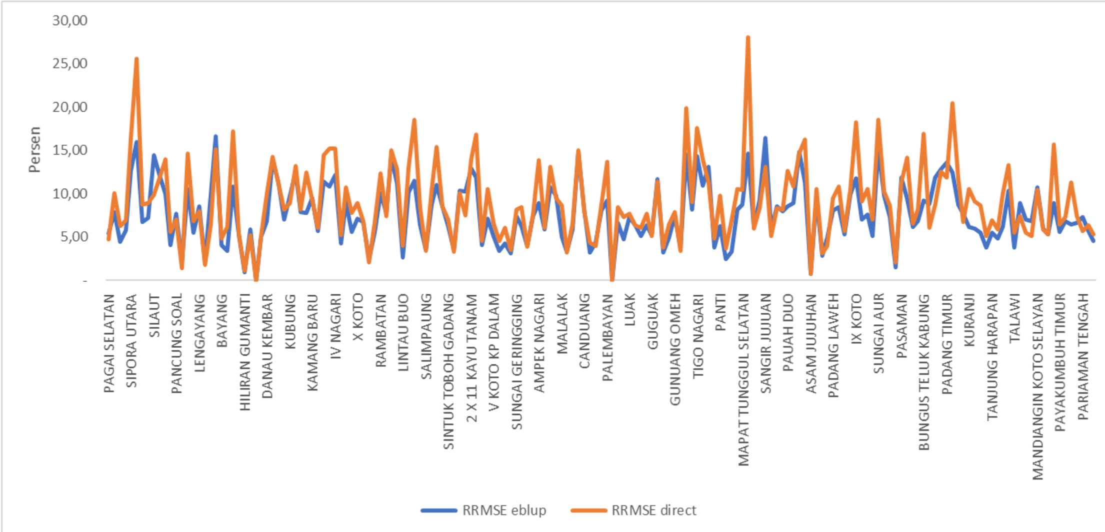
Table 2: Estimation of the Regression Coefficient with the EBLUP-FH Method

Variable	Beta	Std.error	p-value
Intercept	6,082	0,01739	< 2e-16
Non Listrik	-0,00008507	0,00002231	0,000192
Listrik Non PLN	-0,00006433	0,00002887	0,027153
Poliklinik/balai pengobatan	0,01112	0,005601	0,048646
Minimarket/swalayan	0,005014	0,001574	0,001721
SD	-0,003094	0,0007017	0,0000185
Tempat praktik dokter	0,004201	0,001556	0,007645

Before the estimated value of the regression coefficient in Table 2 is used to estimate per capita expenditure, it is necessary first to check the slope and slope of the direct estimate and small random effect area. Based on the processing results, it is shown that the random effect is close to a normal distribution. Thus, the estimated value of the regression coefficient of the auxiliary variables and the variance of random effects can be used to estimate per capita expenditure using the EBLUP FH method.

After estimating the regression coefficients, the next step is to compare the RRMSE values of the two estimates. The RRMSE values of the direct estimator and the EBLUP-FH estimator for per capita expenditure are shown in Figure 6 below.

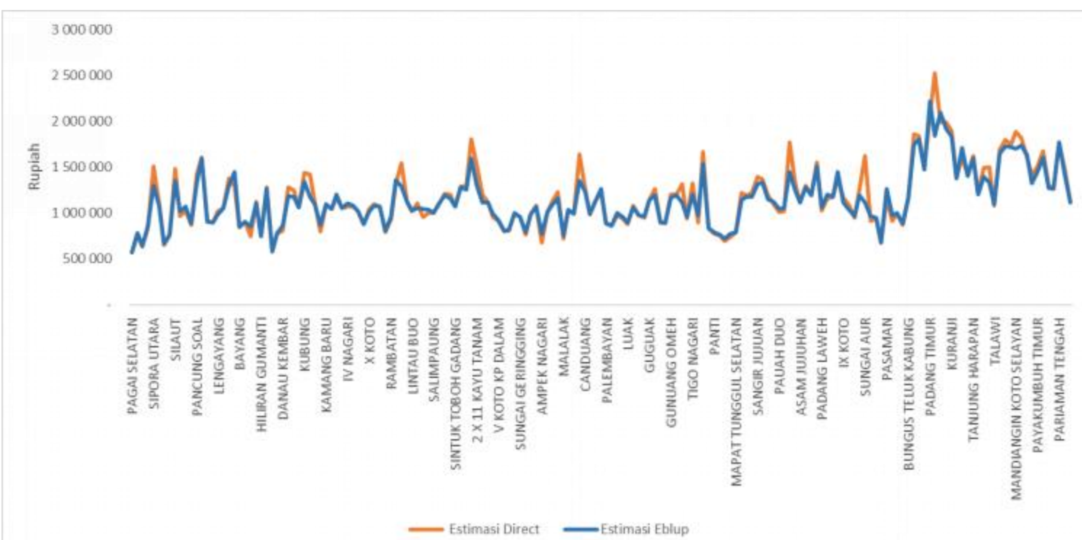
Figure 6: Comparison of RRMSE Value of Direct Estimator and Estimator of EBLUP-FH (%)



Based on Figure 6 above, it can be seen that the RRMSE value from the direct estimation has a very varied value when compared to the EBLUP-FH estimation. The RRMSE value of the EBLUP FH estimation that is relatively smaller than the RRMSE of the direct estimation indicates that the EBLUP FH estimation at the district level can improve the level of precision of the direct estimation results.

After comparing the RRMSE resulting from the direct estimator with the EBLUP-FH estimator, the next step is to compare the estimated value of the average per capita expenditure at the district level resulting from the direct estimator with the EBLUP-FH estimator.

Figure 7: Comparison of the Results of Direct Estimation and Estimation of EBLUP-FH (Rupiah)



In general, the EBLUP-FH estimation results in a per capita expenditure estimator, which is relatively not much different from the direct estimation results. However, the estimation value of EBLUP-FH tends to be lower than the direct estimation results due to the bias of the reverse log-normal transformation.

Based on the results of the EBLUP FH estimation, the districts with the highest average per capita expenditure per month are in the City of Padang, namely East Padang District (2,2 million rupiahs), North Padang District (2,09 million rupiahs), and Nanggalo District (1,92 million rupiah). On the other hand, the districts with the lowest average per capita expenditure per month are South Pagai District, Mentawai Islands Regency (572 thousand rupiah), Tigo Lurah District, Solok Regency (580 thousand rupiah), and North Pagai District, Mentawai Islands Regency (646 thousand rupiah).

Comparison of Results of Direct Estimation and Indirect Estimation

After making direct and indirect estimates using the EBLUP-FH method, the next step is to compare the goodness of fit of the two estimates. Goodness of fit criteria used in evaluating the results of direct and indirect estimation using the EBLUP FH method in this study is RRMSE.

Figure 8: Comparison of RRMSE Direct Estimation and EBLUP-FH (percent)

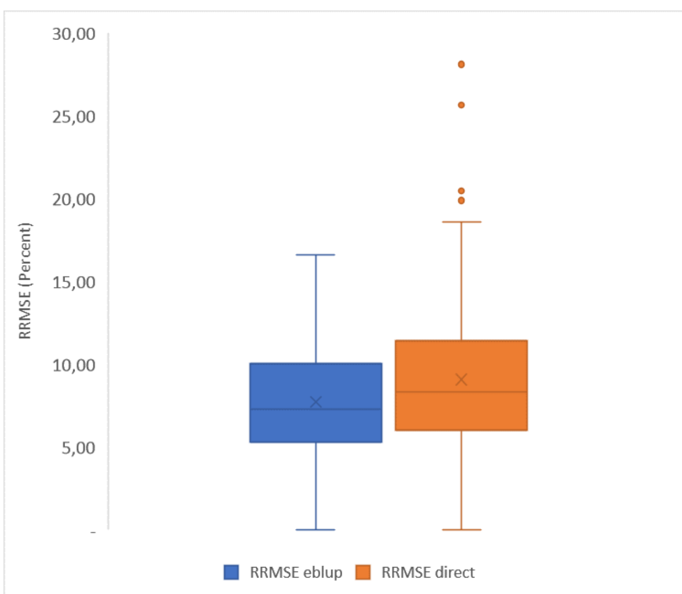


Table 3: Comparison of RRMSE of Direct Estimation Results and EBLUP-FH (percent)

Statistic	Min	Q1	Median	Mean	Q3	Max
$\hat{Y}$	0,76	6,03	8,33	9,05	11,37	28,08
$\widehat{Y^{EBLUP}}$	0,76	5,31	7,27	7,70	9,96	16,59

Based on Figure 8 and Table 3, it can be seen that the RRMSE value of the EBLUP-FH estimator is smaller than the RRMSE value of the direct estimator. In addition, the RRMSE boxplot is also denser than the direct estimator RRMSE boxplot, this indicates that the RRMSE value of the EBLUP FH estimator tends to be homogeneous.

Based on Table 4, it can be seen that the results of direct estimation and estimation of EBLUP FH do not appear to have significant differences.

Table 4: Comparison of Direct Estimation Results and EBLUP-FH (Rupiah)

Statistic	Min	Q1	Median	Mean	Q3	Max
$\hat{Y}$	573 742,90	939 254,20	1 107 408,40	1 170 242,66	1 326 664,60	2 526 366,40
$\widehat{Y^{EBLUP}}$	572 180,20	964 727,15	1 104 294,40	1 146 221,50	1 278 631,75	2 217 049,60

Estimation for Districts Not Affected by the Sample

After selecting the EBLUP-FH method as the best method for estimating per capita expenditure at the district level in West Sumatra, the next step is to estimate per capita expenditure for districts that are not included in the Susenas sample. Several districts not included in the Susenas sample have a direct estimate of per capita expenditure equal to zero. The estimation concept used for unsampled districts can use the synthetic estimator concept with the assumption that the district is homogeneous. The results of the synthetic estimator for per capita expenditure not covered by the sample in West Sumatra can be seen in Table 5.

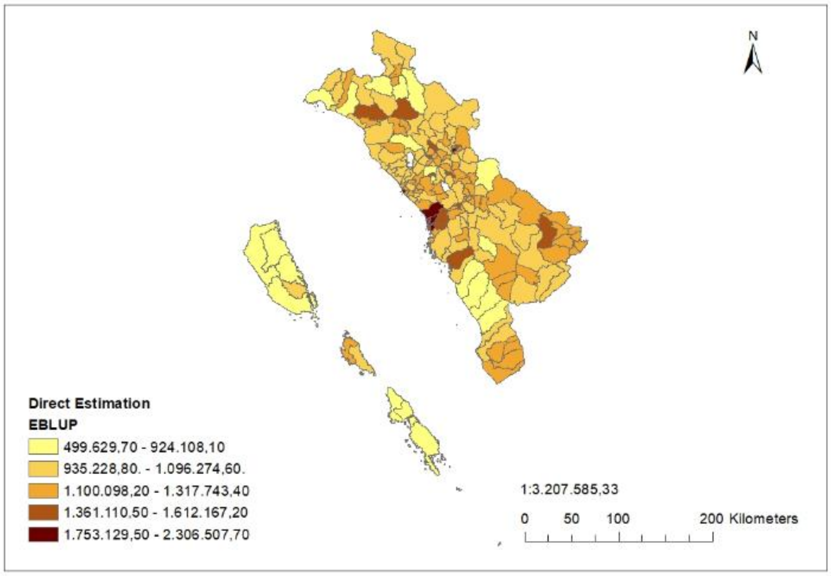
Table 5: Synthetic Estimator of Per Capita Expenditure for Districts that are Not Taken as Samples

Regency	**District**	**Direct Estimator (Rp)**	**RRMSE Direct Estimator (Percent)**	**Synthetic Estimator (Rp)**	**RRMSE Synthetic Estimator (Persen)**
Kep. Mentawai	Siberut Barat Daya	0	0	808 528,90	11,11
Kep. Mentawai	Siberut Barat	0	0	816 015,80	10,68
Pesisir Selatan	IV Nagari Bayang Utara	0	0	1 079 918,40	3,10
Padang Pariaman	Enam Lingkung	0	0	1 249 897,60	4,21

Mapping Average per Capita Expenditure District Level Selected Estimation Results

After obtaining the results of the per capita expenditure estimates for all districts, then the estimation results are presented through the district thematic maps so that further identification of the distribution pattern can be obtained. The mapping of indirect estimation results of per capita expenditure is shown in Figure 9.

Figure 9: Thematic Map of Indirect Estimation Results of Average Per Capita Expenditures at The District Level in West Sumatra, 2019

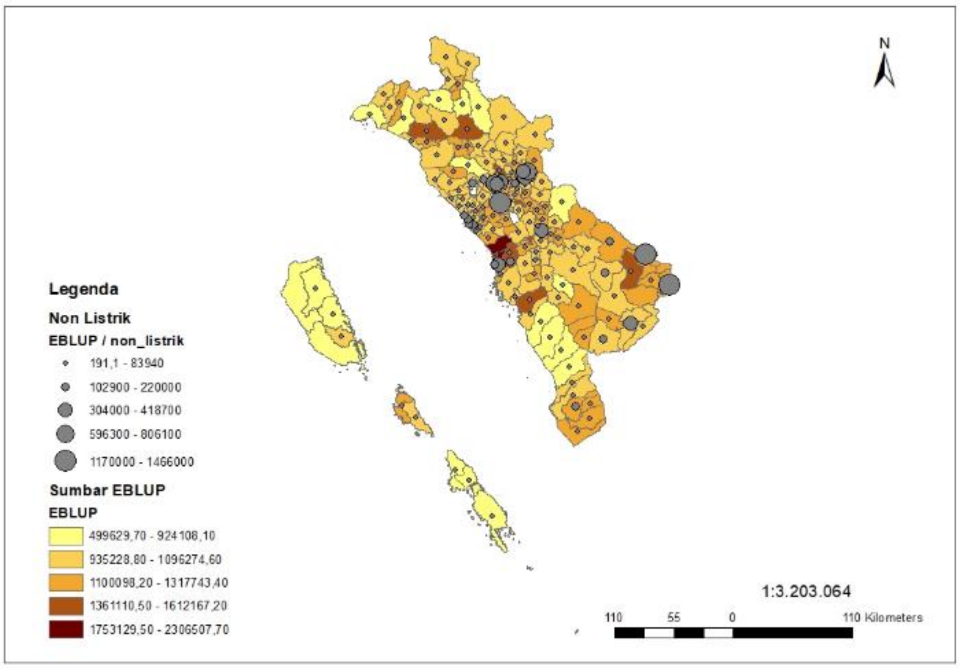


Evaluation of Auxiliary Variables

a. The Number of Families Who are Not Electricity Users

The variable of the number of families who do not use electricity has a negative coefficient, which means that the more households that do not use electricity, the average expenditure per capita at the district level in West Sumatra in 2019 tends to decrease.

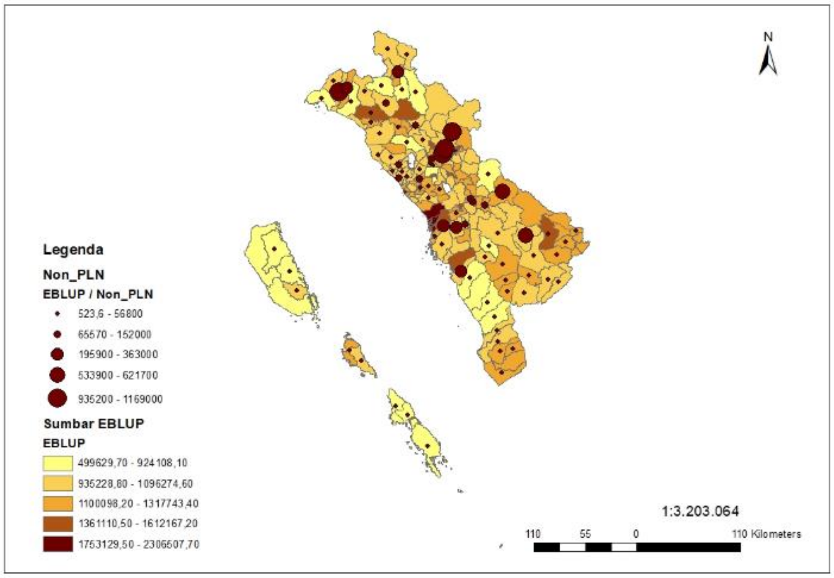
Figure 10: Thematic Map of Per Capita Expenditure and the Number of Families Who are Not Electricity Users in West Sumatra Province in 2019



b. The Number of Non-PLN Electricity Users

The variable of the number of families using non-PLN electricity has a negative coefficient, which means that the more households that use non-PLN electricity, the average per capita expenditure at the district level in West Sumatra in 2019 tends to decrease.

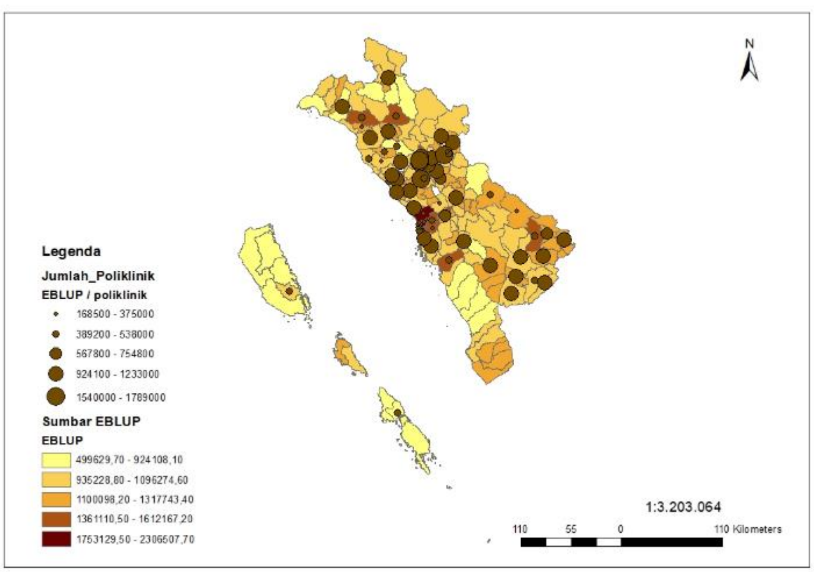
Figure 11: Thematic Map of Expenditures Per Capita and Number of Families of Non-PLN Electricity Users in West Sumatra Province in 2019



c. The Number of Polyclinics/Medical Centers

The variable number of polyclinics/medical centres has a positive coefficient, which means that the more the number of polyclinics/medicine centres, the average expenditure per capita at the district level in West Sumatra in 2019 tends to increase.

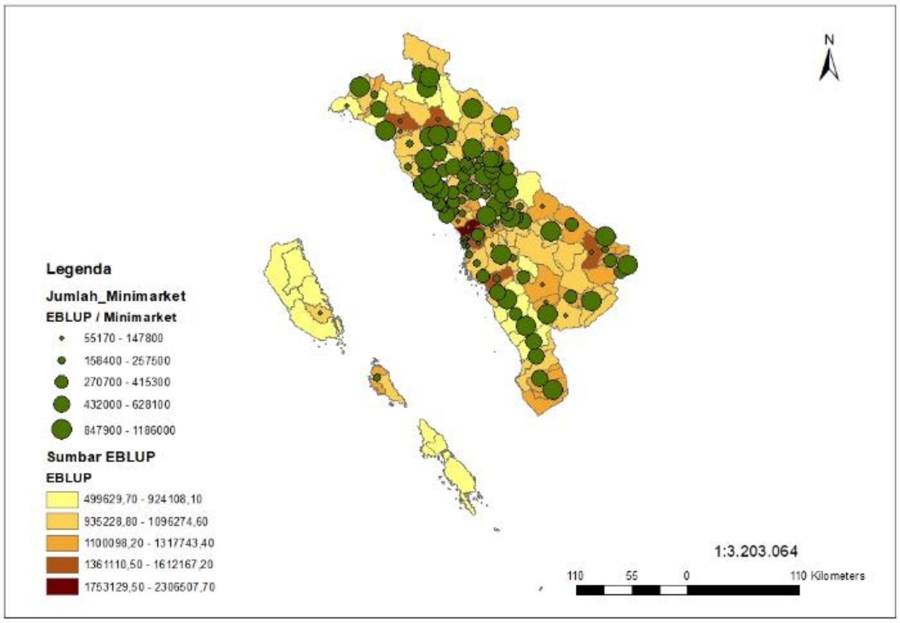
Figure 12: Thematic Map of Expenditure Per Capita and Number of Polyclinics/Medical Centers of West Sumatra Province in 2019



d. The Number of Supermarkets

The variable of the number of supermarkets has a positive coefficient, which means that the more minimarkets/supermarkets there are, the average expenditure per capita at the district level in West Sumatra in 2019 tends to increase.

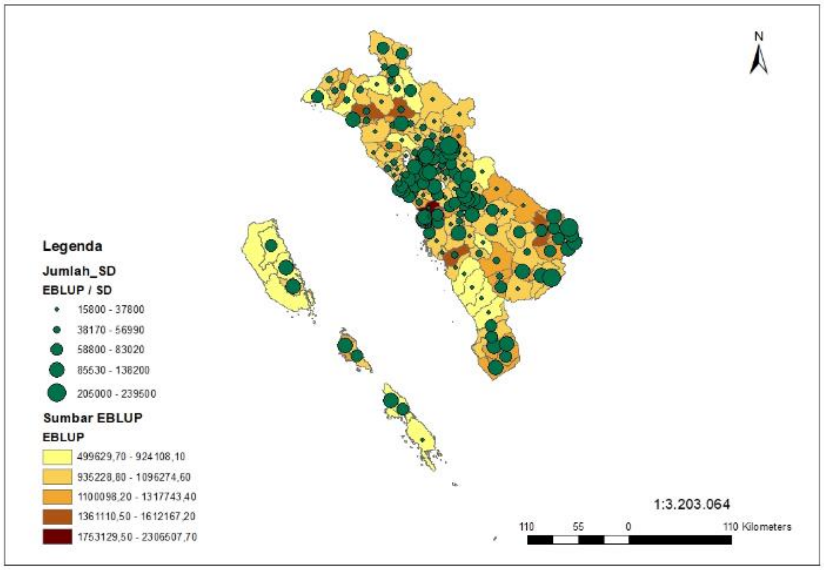
Figure 13: Thematic Map of Expenditures Per Capita and Number of Minimarkets/Supermarkets of West Sumatra Province in 2019



e. The Number of Elementary School

The variable of Elementary School has a negative coefficient, which means that the more elementary schools there are, the average per capita expenditure at the district level in West Sumatra in 2019 tends to decrease. When viewed from the thematic map in Figure 14, it appears that the number of SD/MI school buildings is spread unevenly and tends to cluster near the provincial capital, Padang; at the same time, other regions only have a small number of SD/MI buildings.

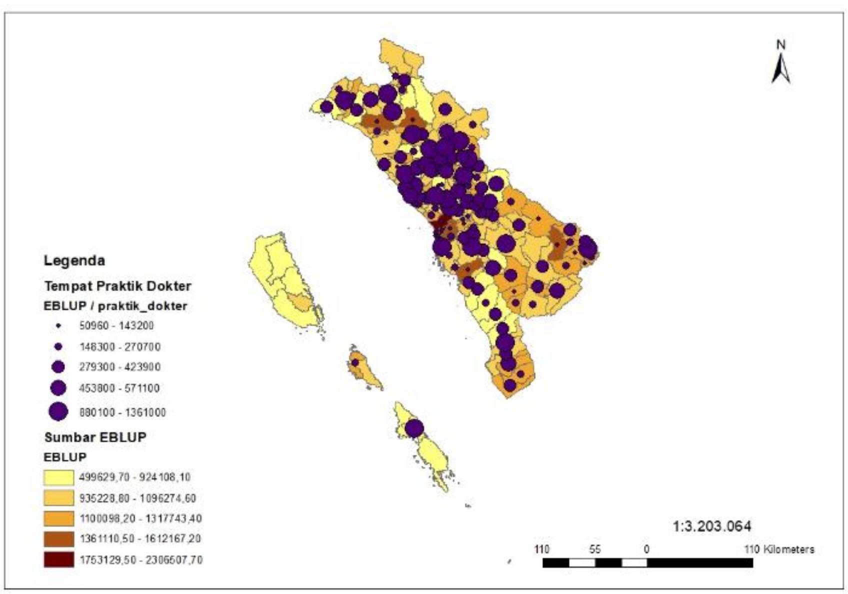
Figure 14: Thematic Map of Expenditures Per Capita and Number of Elementary Schools in West Sumatra Province in 2019



f. The Number of Doctor’s Practices

The variable of doctor’s practices has a positive coefficient, which means that the greater the number of doctor’s practices, the average expenditure per capita at the district level in West Sumatra in 2019 tends to increase.

Figure 15: Thematic Map of Per capita Expenditure and Number of Doctor’s Practice Places in West Sumatra Province in 2019



4. CONCLUSION

1. The comparison results show that the EBLUP-FH estimation has a smaller RRMSE value than the direct estimator. Thus, it can be said that the EBLUP FH method can provide more precise estimation results.
2. The per capita expenditure distribution tends to vary widely between districts. Districts that are geographically close to the provincial capital area tend to have a higher average per capita expenditure than other regions.

5. SUGGESTIONS

Based on the conclusions above, suggestions for this research are as follows:

1. The government can take equitable distribution of vital infrastructure policies, especially districts with lows per capita spending, such as equitable distribution of education and health infrastructure.
2. This research still uses a one-period and univariate model; therefore, future research is expected to include the effect of time or use multivariate analysis to obtain estimates with better precision.
3. To overcome the violation of the normality assumption, the next research can use other methods such as Box-Cox Transformation and Dual Power Transformation.

CONFLICT OF INTERESTS

The author(s) declare that there is no conflict of interests.

REFERENCES

[1] BPS, Pengeluaran per kapita, Sirusa, 2021. https://sirusa.bps.go.id/sirusa/index.php/indikator/197.

[2] M. H. Suryanarayana, Inclusive growth: A sustainable perspective, Indira Gandhi Inst. Dev. Res. Goregaon East, Mumbai, vol. 400, p. 65, 2013.

[3] Asian Development Bank, Introduction to small area estimation techniques: A practical guide for national statistics offices, no. May. 2020.

[4] B. Subandriyo, Estimasi angka partisipasi kasar perguruan tinggi provinsi papua melalui small area estimation, Semin. Nas. Off. Stat. 2019 (2020), 104–109.

[5] Y.D.I. Anggraeni, T. Nugrahadi, Penerapan Empirical best linear unbiased prediction fay herriot (Eblup Fh) dan spatial eblup fh pada data transformasi logaritma, Semin. Nas. Off. Stat. 2019 (2020), 93–103.

[6] S. Scholtus, R. Van De Laar, L. Willenborg, The memobust handbook on methodology for modern business statistics, Eurostat, 2014.

[7] N.A. Yensy, The small area estimation by using empirical bayes method, In: Proc. Int. Conf. Educ. Sci. Teach. Prof. (ICETeP 2020), 532 (2021), 343–347.

[8] H.F. dan Sutikno, Small area estimation terhadap pengeluaran per kapita di kabupaten sumenep dengan metode empirical bayes, Institut Teknologi Sepuluh November, 2011.

[9] A. Kurnia, K. A. Notodiputro, Eb-eblup mse estimator on small area estimation with application to bps data 1,2, no. June, pp. 1–6, 2006.

[10] K. Sadik, K. A. Notodiputro, Hierarchical Bayes estimation using time series and cross-sectional data: A case of per-capita expenditure in Indonesia, SAE2009 Conf. Small Area Estim., 2009.

[11] A.M.A. Satriya, Small area estimation pengeluaran per kapita di kabupaten bangkalan dengan metode hierarchical bayes, Institut Teknologi Sepuluh Nopember, 2016.

[12] T. Purwa, A.T. Rumiati, I. Zain, Small area estimation with bivariate hierarchical Bayes (HB) approach to estimate monthly average per capita expenditure of food and non-food commodities in province of Bali, IOP Conf. Ser. Mater. Sci. Eng. 546 (2019), 052054.

[13] H. Nuryadin, B. Susetyo, K. Sadik, Application of small area estimation of multivariate fay-herriot model (MFH) for the average of per capita expenditure in village level. Int. J. Sci. Eng. Res. 8 (2017), 1673-1676.

[14] M.K. Satrio, Determinan pertumbuhan inklusif di sumatera barat 2011-2016, Universitas Negeri Padang, 2016.

[15] P.K. Mukhopadhyay, A. McDowell, Small area estimation for survey data analysis using SAS software, SAS Glob. Forum 2011 (2011), 1–19.

[16] M. Hidiroglou, Small-area estimation: Theory and practice, JSM Proceedings, Surv. Res. methods Sect., 2007.

[17] J.N.K. Rao, Small area estimation. Wiley, 2003.

[18] Apriliansyah, I.Y. Wulansari, Penerapan Empirical Best Linear Unbiased Prediction (EBLUP) pada Pendugaan Tingkat Pengangguran Terbuka Level Kecamatan di Provinsi Banten, Semnasoffstat. 2021 (2021), 36–44.

[19] R. Chambers, Small area estimation: a review of methods based on the application of mixed models. Southampton Statistical Sciences Research Institute, WP M03/16, Southampton, 2003.

[20] L. Amaliana, I. Fithriani, T. Siswantining, Pendugaan mean squared error (MSE) pada model fay-herriot small area estimation (SAE), Pros. Semin. Nas. Mat. dan Pembelajarannya 2017, 9 (2017), 205–212.

[21] D.N. Gujarati, D.C. Porter, Basic Econometrics, Fifth Edit. New York: Douglas Reiner, 2009.