BAREKENG: Journal of Mathematics and Its Applications
June 2025 Volume 19 Issue 2 Page 1433-1444
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol19iss2pp1433-1444

ASSESSING UNEMPLOYMENT RATES IN TANAH DATAR REGENCY: INSIGHTS FROM SMALL AREA ESTIMATION

Rara Sandhy Winanda [1*], Yasyfin Ikrima Khairani[2], Fajar Wisga Permana[3]

[1,2]*Department of Mathematics, Faculty of Mathematis and Sciences, Universitas Negeri Padang*
Jln. Prof. Dr. Hamka, Air Tawar Barat, Padang, 25171, Indonesia

[3]*BPS-Statistics Payakumbuh Municipality*
Jln. Imam Bonjol No. 7, Payakumbuh, Sumatera Barat, 26225, Indonesia

Corresponding author's e-mail: * ***rarawinanda@fmipa.unp.ac.id***

ABSTRACT

Article History:

Received: 10th December 2024
Revised: 2nd February 2025
Accepted: 9th March 2025
Published: 1st April 2025

Keywords:

EBLUP Method;
RSE;
SAE;
Unemployment.

Unemployment is a significant issue in Indonesia's labor market. The unemployment rate is measured by the Open Unemployment Rate (OUR) through the National Labor Force Survey (SAKERNAS) conducted by BPS. In 2022, the OUR in Tanah Datar District reached its highest level in the past fifteen years. This rise in unemployment contrasts with the declining poverty rate, unlike other districts/cities in West Sumatra. To address the increasing unemployment, detailed information at the smallest administrative level is necessary. However, because the limited sample size in SAKERNAS does not allow for direct estimation of the OUR with sufficient accuracy, this study aims to overcome this limitation by estimating the OUR at the subdistrict level using indirect estimation through Small Area Estimation (SAE). The SAE method applied is Empirical Best Linear Unbiased Prediction (EBLUP), using the Restricted Maximum Likelihood (REML) estimation model. This research uses secondary data obtained from the National Labor Force Survey (SAKERNAS) of Tanah Datar Regency for the August 2022 period and the Village Potential data (PODES) of Tanah Datar Regency in 2021. The findings indicate that three subdistricts—Pariangan, Lintau Buo Utara, and Padang Ganting—have higher OUR values than Tanah Datar Regency in 2022, with rates of 6.00%, 6.01%, and 11.03%, respectively. The factor that influences the high OUR in these sub-districts is the variable percentage of the male population, which in this model has a large contribution to the calculation of OUR. The indirect estimations using EBLUP are deemed reliable, as the RSE value is below 25%. Therefore, the EBLUP indirect estimation results for OUR at the subdistrict level in Tanah Datar Regency can guide local government efforts to take targeted actions to reduce unemployment, especially in areas with high OUR.

How to cite this article:

R. S. Winanda, Y. I. Khairani and F. W. Permana., "ASSESING UNEMPLOYMENT RATES IN TANAH DATAR REGENCY: INSIGHT FROM SMALL AREA ESTIMATION," *BAREKENG: J. Math. & App.*, vol. 19, iss. 2, pp. 1433-1444, June, 2025.

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

Unemployment represents a significant challenge within Indonesia's labor sector. It typically occurs when the number of individuals seeking employment surpasses the availability of suitable job opportunities. This imbalance between labor supply and demand can lead to decreased productivity and reduced income for individuals, which in turn may contribute to poverty and other socio-economic issues. Productivity, in this case, can refer to the low productivity of labor that hinders overall economic growth, which is influenced by factors such as the amount of labor, the quality of labor, and the level of efficient use of resources in the economy [1]. In Indonesia, unemployment is measured using the Open Unemployment Rate (OUR), an indicator derived from the National Labor Force Survey (SAKERNAS). This survey is conducted biannually, with data collection in February to estimate unemployment at the provincial level and in August to provide estimates at the district and city levels [2].

According to data from the Central Statistics Agency (BPS), Indonesia's Open Unemployment Rate (OUR) stood at 5.32 percent in August 2023. In Tanah Datar Regency, the OUR fluctuated between 2018 and 2022, reaching 5.91 percent in 2022—an increase of 1.28 percent from 2021. The OUR in Tanah Datar Regency in 2022 was the highest recorded in the past fifteen years. Several factors contributed to this rise in unemployment, including growth in the labor force and the level of education [3]. Compared to other districts and cities in West Sumatra, the increase in the Open Unemployment Rate in Tanah Datar Regency is not in line with the decrease in the poverty rate in the area. In 2018, the poverty rate in Tanah Datar Regency reached 5.32% and continued to decline in subsequent years until, in 2022, the poverty rate reached 4.26%. A high unemployment rate can have a negative impact on individual welfare, as it results in a lack of income, thus reducing overall welfare and resulting in poverty. This phenomenon reflects the imbalance between the availability of quality employment and people's need for decent work [4]. In this context, the unemployment rate serves as the primary variable of interest for the study.

A crucial factor in facilitating strategies aimed at reducing unemployment is the availability of precise unemployment data. While sample surveys are commonly used to assess unemployment, they often fail to yield accurate measurements at more localized levels, such as subdistricts, villages, or neighborhoods. This data deficiency can pose significant challenges for policymakers in their efforts to mitigate unemployment [5]. To address this issue, an estimation method that enhances accuracy is essential. One viable approach is Small Area Estimation (SAE), which is a statistical technique designed to estimate parameters of subpopulations characterized by relatively small sample sizes [6].

In Small Area Estimation, two types of estimators can be employed: direct estimation and indirect estimation. Direct estimation relies on sample data collected from the specific area under study; however, its accuracy is often limited due to the small sample size, which results in substantial variation in the statistics obtained. In cases where the sample fails to adequately represent the population, indirect estimation becomes necessary. Indirect estimation in the context of Small Area Estimation (SAE) is very important as it provides a way to overcome the limitations that arise from direct estimation. Direct estimation, which relies on data collected directly from a specific area, tends to have low accuracy if the sample size is small or the available data is not representative enough [6]. This can happen in small areas or specific sub-groups where the number of respondents is limited, so direct estimation results can be highly variable or even unreliable. On the other hand, indirect estimation combines information from several areas or other data sources to improve the estimation accuracy.

Small Area Estimation encompasses three commonly utilized methods: Empirical Best Linear Unbiased Prediction (EBLUP), Empirical Bayes (EB), and Hierarchical Bayes (HB) [6]. This research employs the EBLUP approach to estimate the unemployment rate in Tanah Datar Regency. The use of the EBLUP method was chosen because Tanah Datar Regency is characterized by a small survey sample size at the sub-district level, making direct estimation less accurate. In addition, EBLUP offers flexibility in handling unknown variances through efficient estimation procedures, making it more reliable than other estimation methods in the context of small-scale data [7].

Previously, several studies have been conducted using the EBLUP method, including estimating the unemployment rate in Bogor City and estimating the open unemployment rate at the sub-district level in Banten Province. Based on these studies, direct estimation of the variables studied is inadequate due to the limited sample size. The research states that indirect estimators, specifically the EBLUP method, should be employed, as they yield more accurate predictions compared to direct estimators [2][8][9].

This research focuses on estimating the open unemployment rate in each sub-district in Tanah Datar Regency using the Small Area Estimation (SAE) with the EBLUP method. The main difference between the previous research and this research lies in the focus area and development approach. This research specifically develops a model to estimate the open unemployment rate in Tanah Datar Regency. Next, the model will be implemented to estimate the open unemployment rate in each subdistrict in the regency.

2. RESEARCH METHODS

The research focus area covers all sub-districts in Tanah Datar Regency, consisting of 14 sub-districts: X Koto, Batipuh, Batipuh Selatan, Lima Kaum, Lintau Buo, Lintau Buo Utara, Padang Ganting, Pariangan, Rambatan, Salimpaung, Sungai Tarab, Sungayang, Tanjung Baru, and Tanjung Emas. This research uses secondary data obtained from the National Labor Force Survey (SAKERNAS) of Tanah Datar Regency for the August 2022 period and the Village Potential data (PODES) of Tanah Datar Regency in 2021. This research uses the “SAE” package in R-Studio, SPSS, and Microsoft Excel.

2.1 Normality Test

The purpose of the normality test is to determine whether the data being used is normally distributed or not. The normality test used in this study is the Shapiro-Wilk test, as it is more effective for small sample sizes, specifically those with 50 or fewer samples. The test statistics and test criteria used in the Shapiro-Wilk test are as follows [10]:

$$W = \frac{1}{D}\left[\sum_{i=1}^{n} a_i (x_{n-i+1} - x_i)^2\right] \tag{1}$$

with

D	: $\sum_{i=1}^{n}(x_i - \bar{x})^2$
n	: the number of data (or observations)
a_i	: Coefficient test Shapiro-Wilk
x_{n-i+1}	: n-i+1th data
x_i	: i-th data
$\bar{x}$	: average of data

The Shapiro-Wilk test decision is based on the p-value and the W statistic. If the p-value is greater than $\alpha = 0.05$, the null hypothesis (H_0) is accepted, which means the data is normally distributed. If the p-value is smaller than 0.05, the null hypothesis is rejected, and the data is considered not normally distributed. In addition, if the value of the W statistic is close to 1, the data is normally distributed, while a smaller W value indicates data non-normality.

2.2 Pearson Correlation Test

Pearson correlation states whether or not there is a significant relationship between one variable and another. The assumptions or requirements that must be met in using Pearson correlation are that the variables connected are each normally distributed, the variables connected have a linear relationship, the data is randomly selected, the data connected have the same pair of subjects and the variables connected are interval data or ratio data. The test statistics and test criteria used in the Pearson correlation test are as follows [11]:

$$r_{xy} = \frac{m\sum_{i=1}^{m} X_iY_i - (\sum_{i=1}^{m} X_i)(\sum_{i=1}^{m} Y_i)}{\sqrt{\left\{m\sum_{i=1}^{m} X_i^2 - \left(\sum_{i=1}^{m} X_i\right)^2\right\}\left\{m\sum_{i=1}^{m} Y_i^2 - \left(\sum_{i=1}^{m} Y_i\right)^2\right\}}} \tag{2}$$

with

r_{xy} : Correlation coefficient

m : Sample size

$\sum_{i=1}^{m} X_i$: Number of observations of X

$\sum_{i=1}^{m} Y_i$: Number of observations of Y

The Pearson correlation test decision is based on the p-value and the correlation coefficient r. If the p-value is greater than $\alpha = 0.05$, the null hypothesis (H_0) is accepted, which means that there is no significant linear relationship between the two variables. If the p-value is smaller than 0.05, the null hypothesis is rejected, and it is concluded that there is a significant linear relationship. In addition, the r value indicates the strength and direction of the relationship, where a r value close to ± 1 indicates a strong linear relationship, while a r value close to 0 indicates a weak linear relationship or no linear relationship.

2.3 Small Area Estimation

Small area estimation (SAE) is a statistical technique employed to estimate parameters of subpopulations that typically have a relatively small sample size. The term "small area" generally refers to a confined geographical location, such as a sub-district, neighborhood, or village. An area is categorized as a small area when the sample collected from it is insufficient to produce direct estimates with a high degree of accuracy [6].

The development of small area parameter estimation models is primarily based on two main concepts: fixed effect models and small area random effects. Fixed effect models operate under the assumption that the variability within small areas can be explained by supplementary information. Conversely, small area random effects models assume that this variability is inherently random and can also be accounted for by additional information [7]. The integration of these two assumptions gives rise to a mixed-effects model.

In SAE, two commonly utilized models are the implicit and explicit models. The implicit model is applied based on the sampling design used during the direct estimation process. In contrast, the explicit model is formulated considering the influence of randomness in small areas, which is derived from the variability of auxiliary variables. Explicit model forms include Empirical Best Linear Unbiased Predictor (EBLUP) and Empirical Bayes [5][9]. Small area estimators can be categorized into two fundamental models: basic area-level models and basic unit-level models [6].

2.4 Empirical Best Linear Unbiased Prediction (EBLUP)

EBLUP is an estimator that stems from the inability of BLUP (Best Linear Unbiased Predictor) to estimate the unknown variance component. BLUP is a parameter estimator designed to minimize the mean squared error under the assumption that the variance component is known. However, in practice, accurately determining the variance component can be challenging, necessitating its estimation through sample data [6][11]. Consequently, the Empirical Best Linear Unbiased Predictor (EBLUP) model was developed as a technique for addressing mixed effects models, capable of minimizing the resulting mean squared error while assuming the variance component is known. EBLUP is a more straightforward method as it does not require the specification of prior or posterior distributions, unlike the Empirical Bayes (EB) and Hierarchical Bayes (HB) methods [11].

Generally, small area estimation (SAE) models can be categorized into two groups: area-level models, which rely on the availability of supporting data at specific area levels, and unit-level models, which depend on the availability of supporting data at specific unit levels [6]. In this study, an area-based model is utilized because the auxiliary variables are available at the area level, specifically at the sub-district level. The Fay-Herriot model serves as the foundational model for area-level analysis and is expressed as follows [6][11]:

$$\hat{\theta}_i = \boldsymbol{x}_i^T \boldsymbol{\beta} + v_i + e_i,\ \ i = 1, \dots, m \tag{3}$$

where x_i is a $p \times 1$ vector of area-level interacting variables, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ is a $p \times 1$ vector of regression coefficients for the interacting variable x_i, v_i is the small area random effect and e_i is the sampling error and m is the number of areas. $v_i \sim N(0, \sigma_v^2)$ where σ_v^2 is the variance of the random area effect, and $e_i \sim N(0, \psi_i)$ where ψ_i is known from the sample data, and v_i and e_i are mutually independent.

In the BLUP method, the variance of random effects (σ_v^2) is typically unknown and must be estimated prior to further analysis. To estimate the variance of random effects, the Residual Maximum Likelihood (REML) method can be employed. The EBLUP estimator of θ_i is expressed as follows [6]:

$$\hat{\theta}_i^{EBLUP} = \boldsymbol{x}_i^T \widehat{\boldsymbol{\beta}} + \gamma_i\left(\hat{\theta}_i - \boldsymbol{x}_i^T \widehat{\boldsymbol{\beta}}\right) \tag{4}$$

with $\gamma_i = \frac{\sigma_v^2}{(\psi_i + \sigma_v^2)}$.

2.5 Mean Square Error

The quality of the estimator can be assessed by examining the Mean Square Error (MSE) value derived from the estimator [6]. Generally, MSE is calculated as the average of the squares of the differences between the estimated values and the actual values. MSE serves as a metric for comparing two or more estimators. When generating estimators using the EBLUP method, the MSE can be calculated based on the specific estimation method employed. To obtain the MSE for the EBLUP method, it is essential to estimate the value of σ_v^2 using the Restricted Maximum Likelihood (REML) method [12].

$$MSE\left(\hat{\theta}_i^{EBLUP}\right) = g_{1i}(\sigma_v^2) + g_{2i}(\sigma_v^2) + 2g_{3i}(\sigma_v^2) \tag{5}$$

with $g_{1i}(\sigma_v^2) = \frac{\widehat{\sigma_v^2}\psi_i}{\left(\widehat{\sigma_v^2} + \psi_i\right)} = \widehat{\gamma_i}\psi_i$, $g_{2i}(\sigma_v^2) = (1 - \widehat{\gamma_i})^2 \boldsymbol{x}_i^T \left[\frac{x_i x_i^T}{\psi_i + \widehat{\sigma_v^2}}\right]^{-1} \boldsymbol{x}_i$ and $g_{3i}(\sigma_v^2) = \frac{\psi_i^2}{\left(\psi_i^2 + \widehat{\sigma_v^2}\right)^3}$

Once the Mean Squared Error (MSE) is determined, the precision of the EBLUP estimator can be assessed. The precision is calculated using the Relative Square Error (RSE), which measures how adequately the sample represents the population parameters being studied [13]. The quality of the estimation results from a survey can be evaluated based on the resulting RSE [14]. RSE is defined as the ratio of the square root of the MSE to the estimated response variable. A smaller RSE value indicates a better estimator. The RSE can be calculated using the following formula [15]:

$$RSE_i = \frac{\sqrt{MSE(\hat{\theta}_i)}}{\hat{\theta}_i} \times 100\% \tag{6}$$

The accuracy of an estimate can be evaluated based on the following criteria [14]:

a. RSE value $\leq$ 25%: The estimation result is considered accurate and can be utilized.

b. 25% $<$ RSE value $\leq$ 50: The accuracy of the estimation result is moderate, and caution is advised when using it.

c. RSE value $\geq$ 50%: The estimation result is deemed inaccurate and should not be used. In such cases, it is recommended to combine this estimate with other estimates to achieve an RSE value of $\leq$ 25%.

2.6 Unemployment

In labor indicators, unemployment refers to individuals who are not currently employed but are actively seeking work or preparing to start a new business. It also includes those who are not actively looking for work because they have secured employment but have yet to commence their job [8]. Unemployment can arise from an imbalance in the labor market, indicating that the supply of labor exceeds the demand for labor. Broadly speaking, the population of a country can be categorized into two groups: those who are part of the labor force and those who are not [16]. Individuals classified as part of the labor force are residents of working age, defined as those aged 15 years and older [8]. The unemployment rate can be calculated as the percentage ratio of the number of unemployed individuals to the total labor force [2]. The rate can be expressed mathematically as follows:

$$Open\ Unemployment\ Rate = \frac{Number\ of\ unemployed}{Total\ labor\ force} \times 100\%. \tag{7}$$

In straightforward terms, the analysis of unemployment is conducted using the Open Unemployment Rate (OUR), which is defined as the ratio of the number of unemployed individuals to the total labor force [17].

2.7 Auxiliary Variable

A suitable auxiliary variable is one that is closely linked to the variable of interest, or the response variable and is derived from reliable sources such as census data or administrative records. One commonly used source of census data is the Village Potential Data (PODES) [5][6]. The auxiliary variables that are thought to be correlated with the response variable Open unemployment rate at the sub-district level in Tanah Datar Regency are Number of elementary schools (X_1), Number of junior high schools (X_2), Number of high schools (X_3), Number of vocational schools (X_4), Number of universities (X_5), Number of PLN user households (X_6) [2], Number of markets and shops (X_7), Number of hotels and lodging (X_8), Number of micro and small industries (IMK) (X_9), Number of government-issued poor letters (X_{10}) and Percentage of male population (X_{11}) [16]. In small areas such as subdistricts, the sample size is often not large enough to produce accurate direct estimates. By using all relevant variables, indirect estimation methods utilize the correlation between variables to reduce variation and improve estimation precision. The variables used in this study were selected because of their direct relevance to the unemployment rate in Tanah Datar Regency.

2.8 Analysis Technique

The analysis technique involves a systematic approach to data processing and estimation to assess the open unemployment rate. The following steps outline the methodologies employed:

a. Data preparation

   Data preparation involves selecting the Tanah Datar District locus from both the SAKERNAS and PODES datasets.

b. Direct Estimation

   1) Calculate direct estimation

      Compute the direct estimation of the variable of interest, specifically the open unemployment rate, using **Equation (7)**.

   2) Test normality hypothesis

      Conduct a normality hypothesis test on the data obtained from the direct estimation of the unemployment rate using the Shapiro-Wilk test. If the data is found to be not normally distributed, apply the Arcsin transformation [6].

c. Indirect Estimation

   1) Prepare the auxiliary variable data

      Gather the auxiliary variable data obtained from the 2021 PODES (Potensi Desa) data.

   2) Conduct a correlation test

      Perform a correlation test on the auxiliary variables in relation to the response variable.

   3) Estimating the random effect, random effect variance and regression coefficient

      Utilize the Restricted Maximum Likelihood (REML) method in R software to estimate the random effect, random effect variance, and regression coefficient.

4) Estimating the unemployment rate of each sub-district
Apply the indirect estimator method with the EBLUP approach to estimate the unemployment rate for each sub-district using **Equation (4)**.

d. Measuring the precision of the EBLUP estimator

Calculate the Mean Squared Error (MSE) and Relative Square Error (RSE) values for the indirect estimator (EBLUP) using **Equation (5)** and **Equation (6)**.

e. Conclusion.
Summarize the findings and implications of the indirect estimation process and its effectiveness in assessing the unemployment rate.

3. RESULTS AND DISCUSSION

3.1 Direct Estimator of Open Unemployment Rate

Before estimating the variables using the EBLUP method, a direct estimation of the variables under study, specifically the open unemployment rate at the sub-district level was conducted. This direct estimation was carried out across 14 sub-districts in Tanah Datar Regency, utilizing SAKERNAS microdata from August 2022. The results of the direct estimation of the open unemployment rate in Tanah Datar regency can be seen in **Table 1**.

Tabel 1. Direct Estimator of OUR

Sub-district	OUR
X Koto	0.8772
Batipuh	5.2632
Batipuh Selatan	3.7736
Lima Kaum	5.4545
Lintau Buo	3.5714
Lintau Buo Utara	6.0606
Padang Ganting	11.1111
Pariangan	6.0000
Rambatan	5.7971
Salimpaung	5.4545
Sungai Tarab	1.4286
Sungayang	3.7500
Tanjung Baru	4.7059
Tanjung Emas	5.8824

Table 1 shows the results of the estimation of the OUR value at the sub-district level using the direct estimator based on **Equation (7)**. The following presents the descriptive statistics of the direct estimation results for the open unemployment rate for each sub-district in Tanah Datar Regency:

Table 2. Descriptive Statistics of OUR Direct Estimator

Statistics	Open Unemployment Rate
Average	4.9379
Variance	5.8661
Minimum	0.8772
Median	5.3589
Maximum	11.1111
Total Observations	14 Sub-district

Based on the results of the direct estimation, the average open unemployment rate is 4.94%, indicating that approximately 4 to 5 individuals out of every 100 in the labor force are unemployed. The lowest open unemployment rate obtained from the direct estimation is 0.87%. In contrast, the maximum recorded rate is 11.11%, which translates to 11 to 12 individuals out of every 100 residents within the labor force.

3.1.1 Normality Test

Following the direct estimation of the response variable, the open unemployment rate at the sub-district level in Tanah Datar Regency, it is essential to conduct a normality test on the data obtained from this estimation. The normality test serves to assess whether the data from the direct estimation is normally distributed. The results of the normality test for the direct estimation are presented in **Table 3**.

Table 3. Normality Test of Direct Estimation Results

Variable	Shapiro-Wilk		
	Statistics	**df**	**Sig**
Open Unemployment Rate	0.8822	14	0.0625

According to the test criteria outlined in the previous section, the data is considered normally distributed if the resulting significance value is greater than α. As shown in **Table 3**, the significance value in the column is 0.062, which is greater than 0.05. Therefore, the decision is to accept H_0, indicating that the data for the direct estimation of OUR follows a normal distribution.

3.1.2 Selection of Auxiliary Variables

Model-based indirect estimation is done by borrowing the strength of the area around the small area observed by using auxiliary variables. The auxiliary variables come from observations that do not contain errors, such as censuses or regional regressions. In this study, the auxiliary variables used came from the 2021 Tanah Datar District Village Potential (PODES) data collection [8]. The selection of auxiliary variables aims to determine whether a relationship exists with the observed variables. The greater the correlation between the open unemployment rate and the auxiliary variables, the better the estimation result.

The correlation test is used to measure the linear relationship between two variables with the condition that both variables are normally distributed. From the 11 auxiliary variables used, only 5 variables are normally distributed, which are then tested for correlation with the direct estimator OUR value variable. Using the Pearson Correlation test, the correlation coefficients and significance levels are determined, as shown in the following results:

Tabel 4. Pearson Correlation Value of Auxiliary Variable to OUR

Variable	$r_{x\hat{\theta}}$	P-Value
X_1	-0.2054	0.4813
X_2	-0.5266	0.0530
X_6	-0.2810	0.3304
X_7	-0.2935	0.3084
X_{11}	-0.7155	0.0040

Table 4 reveals that only one auxiliary variable is significantly correlated with the open unemployment rate at the 5% significance level: the percentage of the male population (X_{11}). The auxiliary variable chosen has a correlation with the OUR variable. So, only the percentage of the male population variable is used in the next step from the 5 normally distributed auxiliary variables.

3.2 SAE EBLUP Indirect Estimator

The auxiliary variables that are correlated with the OUR have been identified. The auxiliary variable used to build the EBLUP model is the percentage of the male population (X_{11}). Subsequently, indirect estimation can be carried out using the EBLUP method. The initial step involves determining the estimated values for the regression coefficient $(\hat{\beta})$, the random effect (v_i) and the variance of the random effect (σ_v^2) using the Restricted Maximum Likelihood (REML) method in R software. Based on the estimation results, the estimated variance of the random effect (σ_v^2) is 3.0480, while the estimated regression coefficients are provided in **Table 5**.

Tabel 5. **Regression Coefficient Estimation Results**

Variable	$\hat{\beta}_{EBLUPREML}$	P-Value
Intercept (X_0)	204.9672	0.0003
X_{11}	-4.0089	0.0004

Table 5 reveals that the intercept variable and variable X_{11} are significant at the 5% significance level. Before using these values to estimate OUR, a normality test was conducted on the random effects (v_i) using the Shapiro-Wilk test to ensure that the normality assumption was met. The test results show a significance value of 0.4897, which is greater than 0.05, so the null hypothesis that the random effects are normally distributed is not rejected. With the normality assumption fulfilled, the regression coefficient value $(\hat{\beta})$ and variance of the area random effects (σ_v^2) are considered valid and can be used to estimate the indirect open unemployment rate using the EBLUP method using **Equation (4)**. The following presents the results of indirect estimation using the EBLUP method in **Table 6**.

Tabel 6. **Indirect Estimator of OUR with EBLUP Method**

Sub-district	OUR of EBLUP
X Koto	0.8808
Batipuh	5.2725
Batipuh Selatan	3.7954
Lima Kaum	5.4398
Lintau Buo	3.5752
Lintau Buo Utara	6.0149
Padang Ganting	11.0331
Pariangan	6.0043
Rambatan	5.7499
Salimpaung	5.4670
Sungai Tarab	1.4374
Sungayang	3.7814
Tanjung Baru	4.6929
Tanjung Emas	5.8772

Table 6 shows the results of the estimation of the OUR value at the sub-district level using the EBLUP method. Based on Tanah Datar BPS data, the OUR value of Tanah Datar Regency in 2022 is 5.91%. There are three sub-districts that have an OUR value that exceeds the regency average of 5.91%, specifically Lintau Buo Utara at 6.01%, Padang Ganting at 11.03%, and Pariangan at 6.00%. This shows that these three sub-districts have relatively higher unemployment rates than other areas in Tanah Datar. In contrast, sub-districts such as X Koto and Sungai Tarab have much lower OUR values, which could reflect better local economic conditions or different labor characteristics. Below are the descriptive statistics of the indirect estimators calculated using the EBLUP method.

Tabel 7. **Descriptive Statistics of Indirect Estimates of OUR with The EBLUP Method**

Statistics	OUR of EBLUP – REML
Average	4.9301
Variance	5.7620
Minimum	0.8807
Median	5.3562
Maximum	11.0331
Total Observations	14 Sub-district

Table 7 reveals that the average OUR in the 14 sub-districts is 4.93%, which means four to five out of a hundred in the labor force are unemployed. The minimum OUR value was recorded at 0.88%, while the maximum value reached 11.03%. These results reflect significant differences in unemployment rates between regions, which could be caused by locally specific factors such as the availability of jobs or the economic conditions of each sub-district.

3.3 Determination of Precision Measure of EBLUP Estimator

To assess the precision of the EBLUP estimator, the Relative Square Error (RSE) can be calculated using equation 6 for each sub-district. The estimator is considered accurate and suitable for use if the RSE value meets certain criteria. Conversely, an estimator is deemed unsuitable if the RSE value exceeds these thresholds. Descriptive statistics for the RSE values of the indirect estimators using the EBLUP method are provided in **Table 8**.

Tabel 8. Descriptive Statistics RSE of EBLUP Estimator

Statistics	EBLUP
Average	5.0157
Minimum	2.8547
Median	4.2131
Maximum	10.6271
Total Observations	14 Sub-district

Table 8 indicates that the RSE values for the indirect estimation using the EBLUP method to estimate the open unemployment rate at the sub-district level in Tanah Datar Regency in 2022 range from 2.85% to 10.62%. This shows that the level of accuracy of the estimation model varies between sub-districts. Sub-districts with a lower RSE have a better level of estimation accuracy compared to sub-districts with a higher RSE. The small RSE value indicates that the EBLUP model has succeeded in minimizing prediction errors despite using limited data. Since all RSE values are below the 25% threshold, the EBLUP method is considered feasible and able to provide the estimation results of the open unemployment rate at the sub-district level in Tanah Datar regency in 2022.

4. CONCLUSIONS

The findings of this study reveal that one auxiliary variable is correlated with the OUR variable, specifically the percentage of the male population (x_{11}). The relationship between the percentage of male population and the unemployment rate reflects the fact that the demographic structure, specifically the number of males in the population, can affect the unemployment rate. Additionally, the RSE value of less than 25% indicates that the indirect estimator, based on the EBLUP method, is sufficiently accurate for estimating the OUR at the sub-district level in Tanah Datar Regency. Utilizing this indirect estimator through the EBLUP method, the average OUR at the sub-district level is calculated to be 4.93%. The highest OUR is recorded in Padang Ganting Sub-district at 11.03%, while the lowest is in X Koto Sub-district at 0.88%.

Three sub-districts have an OUR value exceeding the overall OUR for Tanah Datar Regency in 2022, which stands at 5.91%. These sub-districts are Pariangan, North Lintau Buo, and Padang Ganting. The results of the indirect estimation using the EBLUP method for the OUR value at the sub-district level in Tanah Datar can be utilized by the local government as a basis for planning and decision-making related to labor policies. The government can use this data to identify sub-districts with high OUR values, such as Padang Ganting sub-district, which has the highest OUR value of 11.03%. With this information, the government can prioritize interventions in these sub-districts, such as creating new jobs, providing local skills-based job training, or directing investment in certain economic sectors that are relevant to the potential of the region. Thus, the government can allocate resources more efficiently and reduce unemployment inequality between regions.

REFERENCES

[1] R. P. Wicaksono, I. K. G. Sukarsa, and I. P. E. N. Kencana, “MEMODELKAN TINGKAT PENGANGGURAN DI KOTA DENPASAR DENGAN PENDUGAAN AREA KECIL EMPIRICAL BAYES,” *E-Jurnal Matematika*, vol. 10, no. 2, p. 81, May 2021, doi: 10.24843/mtk.2021.v10.i02.p325.

[2] I. Yuni Wulansari, “PENERAPAN EMPIRICAL BEST LINEAR UNBIASED PREDICTION (EBLUP) PADA PENDUGAAN TINGKAT PENGANGGURAN TERBUKA LEVEL KECAMATAN DI PROVINSI BANTEN

(EMPIRICAL BEST LINEAR UNBIASED PREDICTION (EBLUP) FOR OPEN UNEMPLOYMENT RATE ESTIMATION AT SUB-DISTRICT LEVEL IN BANTEN PROVINCE),” 2021.

[3] I. K. G. Sukarsa, G. K. Gandhiadi, and I. P. E. N. Kencana, “UNEMPLOYMENT RATE ESTIMATION IN BALI PROVINCE: A SMALL AREA ESTIMATION APPROACH,” *BAREKENG: J. Il. Mat. & Ter*, vol. 16, no. 1, pp. 157–162, 2022, doi: 10.30598/barekengvol16iss1p157-162.

[4] R. C. Kurniawan, “ANALISIS PENGARUH PDRB, UMK, DAN INFLASI TERHADAP TINGKAT PENGANGGURAN TERBUKA DI KOTA MALANG TAHUN 1980-2011,” 2013.

[5] E. Ikhsan, C. A. Hidayat, and W. A. Nurizza, “EFISIENSI METODE EBLUP PADA SMALL AREA ESTIMATION STUDI KASUS: ESTIMASI PERSENTASE PENDUDUK MISKIN DI PROVINSI NUSA TENGGARA TIMUR TAHUN 2017 (THE EFFICIENCY OF EBLUP IN SMALL AREA ESTIMATION CASE STUDY: ESTIMATION OF POOR POGPULATION PERCENTAGE IN EAST NUSA TENGGARA PROVINCE IN 2017),” 2017.

[6] J. N. K. Rao and I. Molina, *SMALL AREA ESTIMATION*. John Wiley & Sons, Inc., 2015.

[7] A. Wijaya, Moh. Y. Darsyah, and I. J. Suprayitno, “SMALL AREA ESTIMATION PADA TINGKAT KEMISKINAN DI PROVINSI JAWA TENGAH DENGAN PENDEKATAN EMPIRICAL BEST LINIER UNBIASED PREDICTION,” 2017.

[8] S. Aprizkiyandari and A. Kurnia, “SMALL AREA ESTIMATION IN ESTIMATING UNEMPLOYMENT RATE IN BOGOR DISTRICT OF SAMPLED AND NON-SAMPLED AREAS USING A CALIBRATION MODELING APPROACH,” *IJCSN-International Journal of Computer Science and Network*, vol. 6, no. 6, 2017, [Online]. Available: www.IJCSN.orgImpactFactor:1.5760

[9] R. Amaliah, “PENDUGAAN AREA KECIL PADA TINGKAT PENGANGGURAN TERBUKA MENGGUNAKAN MODEL FAY-HERRIOT SPASIAL,” 2018.

[10] I. Sintia, M. Danil Pasarella, and D. Andi Nohe, “PERBANDINGAN TINGKAT KONSISTENSI UJI DISTRIBUSI NORMALITAS PADA KASUS TINGKAT PENGANGGURAN DI JAWA,” 2022.

[11] R. Ningtyas, R. Rahmawati, and Y. Wilandari, “PENERAPAN METODE EMPIRICAL BEST LINEAR UNBIASED PREDICTION (EBLUP) PADA MODEL PENDUGA AREA KECIL DALAM PENDUGAAN PENGELUARAN PER KAPITA DI KABUPATEN BREBES,” vol. 4, pp. 977–986, 2015, [Online]. Available: http://ejournal-s1.undip.ac.id/index.php/gaussian

[12] I. Molina, J. N. K. Rao, and G. S. Datta, “SMALL AREA ESTIMATION UNDER A FAY-HERRIOT MODEL WITH PRELIMINARY TESTING FOR THE PRESENCE OF RANDOM AREA EFFECTS,” 2015, [Online]. Available: www.statcan.gc.ca

[13] W. Larasati and N. Permatasari, “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, 2022.

[14] Kementerian Ketenagakerjaan Republik Indonesia, “KETENAGAKERJAAN DALAM DATA (EDISI 1),” 2024.

[15] E. Ikhsan, N. Yudelsa Ratu, and W. Avie Nurizza, “ESTIMASI SMALL AREA ESTIMATION ANGKA PARTISIPASI KASAR DI PERGURUAN TINGGI PULAU KALIMANTAN TAHUN 2018,” 2019.

[16] B. Hartono and R. Hapsari, “KAJIAN METODE SMALL AREA ESTIMATION UNTUK MENDUGA TINGKAT PENGANGGURAN TERBUKA,” vol. 1, pp. 95–106, 2018.

[17] BPS Tanah Datar, *INFORMASI KETENAGAKERJAAN KABUPATEN TANAH DATAR 2022*. 2022.