ComTech: Computer, Mathematics and Engineering Applications, 14(1), June 2023, 1−9
DOI: 10.21512/comtech.v14i1.8033
P-ISSN: 2087-1244
E-ISSN: 2476-907X

Spatial Modeling of Fixed Effect and Random Effect with Fast Double Bootstrap Approach

Wigbertus Ngabu[1]*, Henny Pramoedyo[2], Rahma Fitriani[3], and Ani Budi Astuti[4]

[1-4]Department of Statistics, Faculty of Mathematics and Natural Sciences, Brawijaya University
Jln. Veteran, Jawa Timur 65145, Indonesia
[1]bertongabu@gmail.com; [2]hennyp@ub.ac.id; [3]rahmafitriani@ub.ac.id; [4]ani_budi@ub.ac.id

Received: 30th December 2021/ **Revised:** 7th June 2022/ **Accepted:** 8th June 2022

How to Cite: Ngabu, W., Pramoedyo, H., Fitriani, R., & Astuti, A. B. (2023). Spatial Modeling of Fixed Effect and Random Effect with Fast Double Bootstrap Approach. *ComTech: Computer, Mathematics and Engineering Applications, 14*(1), 1−9. https://doi.org/10.21512/comtech.v14i1.8033

Abstract **-** The use of panel data on spatial regression has many advantages. However, testing the spatial dependency and parameter presumption generated in spatial regression of panel data becomes inaccurate when applied to regions with large numbers of small spatial units. One method of overcoming problems of small spatial unit sizes is the bootstrap method. The research aimed to combine cross-section and time-series panel data. The analysis was performed to extract information based on observations modified by the influences of space or location, known as spatial analysis of panels. The influence of location effects on spatial analysis was presented in the form of weighting. The research applied the Fast Double Bootstrap (FDB) method by modeling poverty rates on Flores Island. The results of the Hausman test show the right model, which is a random effect. Meanwhile, spatial dependency testing concludes spatial dependence and poverty modeling in Flores Island, which is more likely to be the Spatial Autoregressive Random (SAR) model. SAR random effect in modeling value has R^2 of 77,38% and does not meet the normality assumption. SAR effect in modeling the FDB approach can explain the diversity of poverty rate in the Flores Island with 88,64% and meets residual normality assumptions. The analysis with the FDB approach on spatial panels shows better results than the common spatial panels.

Keywords: spatial modeling; fixed effect; random effect; Fast Double Bootstrap (FDB)

I. INTRODUCTION

Panel data is made up of cross-section and time-series data. The consequence of observing one or more variables over a period is time-series data. Meanwhile, cross-section data is observed from one or more variables taken from multiple sample units or subjects over the same time. On panel data, the same individual units are observed over time (Schmidt, 2020).

Then, the spatial analysis of panels is influenced by the effects of space or location. The influence or effect of location on spatial analysis is presented in the form of weighting. The size of proximity depends on the knowledge of the size and shape of the observation units depicted on the map (Elhorst, 2017).

Spatial econometric models have been extensively studied in the past thirty years. They have fewer assumptions than classic econometric models, and to be precise, spatial econometrics can test spatial effects, including spatial dependence and heterogeneity. Therefore, spatial econometrics can obtain more reasonable and realistic conclusions than classic econometrics. Testing the existence of spatial dependence in a spatial econometric model has been a core issue. In addition, spatial econometric models include spatial lag and spatial error models (Liu & Yang, 2020). Previous research shows how dependent variables in the vicinity of the behavior affect other parts of the overall system behavior. Meanwhile, when spatial dependence exists in the error term, previous research is done on the influence of the error shock on neighboring region behavior (Wang & Lam, 2020). Currently, there are several methods to test the existence of spatial dependence in spatial cross-sectional data models. However, methods are rarely presented to test the existence of spatial dependence in spatial panel data models.

In spatial regression, the existence of spatial dependency is an absolute requirement for the use of this analysis. The common methods to test the existence of spatial dependence in spatial cross-sectional data models include Moran's I, Lagrange Multiplier (LM), Likelihood Ratio (LR), and Rao's score test. Moran's I test of spatial dependence is assumed to be a non-alternative hypothesis model. It can test spatial lag dependence and spatial error dependence. Therefore, Moran's I test has been the most commonly used (Wang, Yamamoto, & Liu, 2021).

Bootstrap is a method for estimating the distribution of an estimator or test statistic by resampling one's data or model estimated from the data (Galvao, Parker, & Xiao, 2021). So far, bootstrap is widely applied in cross-sectional data, time series, and panel data. Under serial and cross-sectional correlation, the first-order asymptotic null validity of the test is demonstrated. The consistency of the test under an alternate hypothesis is also demonstrated (Choi & Shin, 2020).

Different bootstrap methods have been developed for different types of regression, such as residuals bootstrap, block bootstrap (Djogbenou, MacKinnon, & Nielsen, 2019), wild bootstrap, and wild cluster bootstrap (Canay, Santos, & Shaikh, 2021). The block bootstrap is used in time series models and deals with panel data models and heteroscedasticity (Davidson & Trokić, 2020). Then, the pairs bootstrap is used in dynamic or heteroscedastic models in which the error term is unknowingly distributed (Lütkepohl & Schlaak, 2019). The sub-cluster wild bootstrap is a family of new methods that include the ordinary wild bootstrap as a limiting instance. The latter technique can perform very effectively in pure treatment models, where all observations within clusters are either treated or not. The most important criterion is that all cluster sizes should be comparable (MacKinnon & Webb, 2018).

There are several advantages of using the bootstrap approach. It does not require any assumptions about data distribution, or the term error is independent and normally distributed (Roodman, Nielsen, MacKinnon, & Webb, 2019). Using the bootstrap method in some conditions of regularity makes it possible to obtain a more accurate spread of the presumption than the usual statistical distribution. The bootstrap clusters method is often used in panel data, where the bootstrap clusters method works very well in practice (Du, Worthington, & Zelenyuk, 2018).

A number of bootstrap procedures are available in the literature. The bootstrap procedure considered in the research is the wild cluster bootstrap procedure, which is an extended version of the wild bootstrap proposed in a cluster setting. It is proven that this procedure is good to be performed in practice, despite the fact that the pairs cluster bootstrap works well in principle (MacKinnon, Nielsen, & Webb, 2021). In previous research, the comparison of the finite-sample size of the bootstrapped t-statistics resulting from the pairs cluster bootstrap and wild cluster bootstrap indicates that the wild cluster bootstrap performs better (MacKinnon & Webb, 2018).

Efron, in 1979, introduced the computational bootstrap method as an empirical problem-solving alternative. This method is proven more accurate than asymptotic methods under small sample conditions, and the distribution of parameters is unknown (LaFontaine, 2021). Then, Beran, in 1988, developed the double bootstrap method having better performance than the usual bootstrap method (Mameli, Musio, & Ventura, 2018). However, the double bootstrap method requires longer calculation time because it has to be calculated as many as $B1 + B1B2$ tests statistical values.

Meanwhile, Fast Double Bootstrap (FDB) assumes that the test statistics on the first stage bootstrap data set and the second stage bootstrap data set test statistics are mutually free. Thus, one replication on the bootstrap in the second stage is enough for each first-stage bootstrap data set. This method produces the same level of accuracy as the double bootstrap method but requires a much shorter processing time (Davidson & Trokić, 2020).

Spatial bootstrap tests based on residual Ordinary Least Square (OLS) are based on Moran's I statistics to test spatial correlations on models. The FDB method results in better Moran's I statistical and asymptotic assumption tests (Schuldt et al., 2019). The research uses the bootstrap method of the LM test for spatial dependency on panel data models with fixed effects. The version of the LM test consistency and bootstrap must be proven to obtain an asymptotic refinement of the LM bootstrap test (Ou, Long, & Li, 2019). The wild cluster bootstrap is used to investigate inference based on cluster residual for regression models with clustered errors. Asymptotic and bootstrap tests, as well as confidence intervals, are asymptotically valid under certain conditions. As the number of clusters approaches infinity, these restrictions limit the rate at which cluster sizes can grow. Edgeworth expansions for asymptotic and bootstrap test statistics are also derived (Djogbenou et al., 2019).

Poverty has been a significant issue in Indonesia for the past five years. As a result, for the next five years, it will become a strategic development concern. The research aims to discover poverty variables and measure their impacts to contribute to the eradication of poverty. Modeling poverty in Indonesia has two problems. To begin with, there is a spatial reliance in terms of poverty between regions. Second, certain key elements are left out of the empirical model due to data limitations. Both factors can cause estimation errors in regression parameters. Hence, a fixed effect panel spatial error model is used to model the poverty rate. It is found that poverty is influenced by the unemployment rate and economic well-being (Suparman & Ginanjar, 2021).

In addition, poverty that occurs over time is highlighted as part of an effort to develop a more accurate model. Previous research has aimed to use a geographic data panel analysis to determine the factors that influence the percentage of poor people in East Java province from 2012 to 2017. The Spatial Autoregressive (SAR) model with the concept of distance is the best model for this instance. It demonstrates that poverty in East Java has a spatial influence (Yolanda & Yunitaningtyas, 2019).

East Nusa Tenggara is one of the poorest provinces in Indonesia. Data from Statistics Indonesia showed the number of poor people in East Nusa Tenggara in March 2019 at 1,15 million (21,09%).

This number increased compared to March 2018, which was 1,14 million people. The problem of poverty in East Nusa Tenggara does not only lie in the high rates but also the high disparity between regions. Comparisons between villages/cities show a large disparity. This inequality occurs because of the high poverty rate in certain regions. Based on very diverse conditions, it causes differences in each region in East Nusa Tenggara, so spatial effect problems arise because of the effects of geographical factors. Flores Island is one of the islands in East Nusa Tenggara, consisting of eight districts. It is an archipelago with the highest poverty rate in East Nusa Tenggara. The problem of poverty has become one of the urgent problems in Indonesia. Poverty influences the well-being of the society on Flores Island (Badan Pusat Statistik Provinsi Nusa Tenggara Timur, 2020).

The research uses spatial autoregressive panel data models with the FDB approach to cluster poverty rate data in Flores Island. Poverty modeling in Flores Island has a small sample unit. In addition, the use of panel data on time series in 2018-2020 also has not been able to produce an adequate number of observations. This condition causes issues with testing spatial dependency due to small and non-normally distributed residual samples. Then, inference statistics are based on asymptotic normal distribution assumptions with reference to the law of large numbers and the central limit theorem in general. Inference statistics will cause gaps in confidence and improper statistical tests in a small sample. The Maximum Likelihood Estimator (MLE) and Ordinary Least Square (OLS) approach estimator accuracy are less accurate in small samples (Nieuwland et al., 2018). The asymptotic behavior of those statistics leads to poor estimation of the actual data if the sample size is not quite large. The FDB method can be utilized to fix this issue. It is possible to obtain a more accurate spread of the presumption FDB method than the usual statistical distribution under certain conditions of regularity (Du et al., 2018). The research aims to identify the procedures for using the FDB method on spatial regression with spatial fixed effect and random effect. The estimation model of the spatial fixed model and random effect is obtained through the FDB approach on the poverty rates in Flores Island.

II. METHODS

Data used are from Statistics Indonesia of East Nusa Tenggara for the 2018–2020 periods. The research is performed in the entire district of Flores Island. Dependent variables are poverty, while independent variables consist of previous Expected Years of Schooling (EYS), Gross Regional Domestic Product (GRDP), Average Life Expectancy (ALE), Regional minimum Wage (RMW), and Unemployment Rate (UR).

Poverty in Flores Island is characterized not only by a large number or percentage of people living in poverty but also by a wide inequality between regions. When comparing districts, there are significant differences. This disparity arises because of the high poverty rate in some places. Because of the wide range of variables in each East Nusa Tenggara location, the problem of spatial effects develops due to geographic considerations.

Data are tested using the SAR model with the FDB approach. Meanwhile, the analysis step is the first stage of determining the spatial weighting matrix based on contiguity and normalizing the line to obtain the matrix (W). The spatial weighting matrix (W) is a key component in defining the proximity of one place to another. It is calculated using data about the proximity between two areas (neighborhoods). The distance or contiguity between two regions can be used to generate a spatial weighting matrix (W). The spatial weighting matrix takes the following shape in Equation (1).

$$W = \begin{bmatrix} 0 & w_{12} & w_{12} & \cdots & w_{12} \\ w_{12} & 0 & w_{12} & \cdots & w_{12} \\ w_{12} & w_{12} & 0 & \cdots & w_{12} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{12} & w_{12} & w_{12} & \cdots & 0 \end{bmatrix} \quad (1)$$

The Rook Contiguity method is used in the study. It has $W_{ij} = 1$ for regions that overlap sides with areas of concern and $W_{ij} = 0$ for all other regions (Suryowati, Bekti, & Faradila, 2018). Weights line standardization of Rook Contiguity is shown in Equation (2).

$$\boldsymbol{w}^*_{ij} = \frac{\boldsymbol{w}_{ij}}{\sum_{j=1}^{N} \boldsymbol{w}_{ij}} \quad (2)$$

Then, choosing the right model between fixed and random effects with the Hausman test is performed. The modeling of the conformity test is conducted using the spatial Hausman test, which aims to compare fixed effects and random effects on SAR models. This test is done based on the assumption of whether the random effect matches the data used or not. If it does not match, the estimator of the random effect model is inconsistent and can decide to use the fixed-effect model (Lee & Yu, 2020). Hausman test statistics are shown in Equation (3). It shows $\hat{\beta}_{random}$ as the parameter estimator of the random effect model, $\hat{\beta}_{fixed}$ as the parameter estimator of the fixed effect model, and $\hat{\Sigma}_{random}$ and $\hat{\Sigma}_{fixed}$ as the matrix of the kovarian variant of each estimator.

$$H = NT\left(\hat{\beta}_{random} - \hat{\beta}_{fixed}\right)'\left(\hat{\Sigma}_{random} - \hat{\Sigma}_{fixed}\right)^{-1}\left(\hat{\beta}_{random} - \hat{\beta}_{fixed}\right) \quad (3)$$

The next step is performing spatial test dependencies for original data, namely Moran's I and LM tests. Spatial effects have been tested in spatial panel models before the spatial panel data models are established. At present, the most commonly used method to test spatial dependence is Moran's I test. The test statistics developed for the cross-section are extended to the panel data model. Moran's I test for residual approach is in Equation (4) (Li, Hong, & Peng, 2019). It has I as Moran's I test statistic of a spatial panel data model, $\boldsymbol{W}_{nT} = \boldsymbol{I}_T \otimes \boldsymbol{W}$ as the spatial weights matrix, $\otimes$ as the Kronecker product, and $\hat{\boldsymbol{\varepsilon}}$ as the residual. It shows that if the error term does not obey the classic distribution or heteroscedasticity, Moran's I test will lapse (Yang, 2018). The bootstrap method is an effective way to solve these problems.

$$\mathbf{I}_0=\frac{\hat{\varepsilon}' \mathbf{W}_{nT}\hat{\varepsilon}}{\hat{\varepsilon}'\hat{\varepsilon}} \tag{4}$$

The LM test is the most often utilized spatial autocorrelation specification test. The LM and robust LM tests are used, and statistical tests are performed to determine the presence of spatial dependency and the term autoregressive in the model (Elhorst, 2017). LM test in the lag model is in Equations (5) and (6) (Elhorst, 2017). Meanwhile, in Equations (7) and (8), LM tests error.

$$LM_\delta = \frac{\left[\varepsilon'(I_T \otimes W)y/\hat{\sigma}^2\right]^2}{J} \tag{5}$$

$$robustLM_\delta = \frac{\left(\hat{\varepsilon}'(I_T \otimes W)y/\hat{\sigma}^2 - \hat{\varepsilon}'(I_T \otimes W)\hat{\varepsilon}/\hat{\sigma}^2\right)^2}{J - TT_W} \tag{6}$$

$$LM_\delta = \frac{\left[\varepsilon'(I_T \otimes W)\hat{\varepsilon}/\hat{\sigma}^2\right]^2}{TT_W} \tag{7}$$

$$robustLM_\lambda = \frac{\left(\frac{\varepsilon'(I_T \otimes W)\hat{\varepsilon}}{\hat{\sigma}^2} - \left(\frac{TT_W}{J}\right)\left(\frac{\varepsilon'(I_T \otimes W)y}{\hat{\sigma}^2}\right)\right)^2}{TT_W\left(1-\frac{TT_W}{J}\right)^{-1}} \tag{8}$$

Next is the spatial autoregressive model of panel data (Elhorst, 2017). A spatial regression model using panel data with spatial effects is observed in the lag of dependent variables (λ=0). It is known as a SAR panel without spatial effects on the model error ($\rho \neq 0$). The spatial autoregressive fixed effect model is shown in Equation (9). Meanwhile, the spatial autoregressive random effect model is in Equations (10) and (11).

$$\boldsymbol{y} = [\boldsymbol{I}_{nT} - \hat{\rho}(\boldsymbol{I}_T \otimes \boldsymbol{W})]^{-1}(\boldsymbol{X}\hat{\boldsymbol{\beta}} + (\boldsymbol{I}_T \otimes \boldsymbol{\mu}) + \hat{\boldsymbol{\varepsilon}}) \tag{9}$$

$$\boldsymbol{y} = \rho(\boldsymbol{I}_T \otimes \boldsymbol{W}_{nT})\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \hat{\boldsymbol{\varepsilon}} \tag{10}$$

$$\boldsymbol{\varepsilon} = (\iota_T \otimes \boldsymbol{I}_N)\phi + v \tag{11}$$

The next stage was the spatial panel test using the FDB approach. The research develops bootstrap with the FDB method for calculating Moran's I value using panel data with small samples, as shown in Equations (12) and (13). Then, there is Moran's I FDB p-value in Equation (14). The $\hat{P}_I^*$ is Moran's I p-value at the bootstrap of the first stage, as shown in Equation (15).

$$I_b^{**} = \frac{\hat{\varepsilon}_b^{**\prime} W_{nT}\hat{\varepsilon}_b^{**}}{\hat{\varepsilon}_b^{**\prime}\hat{\varepsilon}_b^{**}} \tag{12}$$

$$I^{**} = \frac{1}{B}\sum_{b=1}^{B} I_b^{**} \tag{13}$$

$$\hat{P}_I^{**} = \frac{amount\left(I_b^{**} \geq Q^{**}(1-\hat{P}_I^*)\right)}{B} \tag{14}$$

$$\hat{P}_I^* = \frac{amount(I_b^* \geq I_0)}{B} \tag{15}$$

The presence of spatial dependencies between areas on dependent variables can be tested using LM lag tests. LM lags testing for the FDB approach is developed from the statistics of the LM lag test of Equation (5) on the bootstrap method (Ou et al., 2019). The value of FDB LM lags is stated in Equations (16) and (17).

$$LMlag_{fdb} = LM_\rho^{**} = \frac{1}{B}\sum_{b=1}^{B} LM_{\rho b}^{**} \tag{16}$$

$$LM_{\rho b}^{**} = \left(\frac{\hat{\varepsilon}_b^{**}(I_T \otimes W)y_b^{**}/\hat{\sigma}^2}{J_b^{**}}\right)^2 \tag{17}$$

Next, the research calculates the LM lag value from the first stage bootstrap data set (LM_ρ^*) and obtains the first stage bootstrap p-value $(\hat{P}_{LM\rho}^*)$. Then, the first stage bootstrap data set is resampled again in the second stage. The p-value of FDB LM lags can be obtained using Equations (18), (19), and (20).

$$\hat{P}_{LM\rho}^{**} = \frac{amount\left(LM_{\rho b}^{**} \geq Q^{**}(1-\hat{P}_{LM\rho}^*)\right)}{B} \tag{18}$$

$$\hat{P}_{LM\rho}^{*} = \frac{amount(LM_{\rho b}^{*} \geq LM_\rho)}{B} \tag{19}$$

$$LM_{\rho b}^{*} = \left(\frac{\hat{\varepsilon}_b^{*}(I_T \otimes W)y_b^{*}/\hat{\sigma}^{*2}}{J_b^{**}}\right)^2 \tag{20}$$

SAR modeling on fixed and random effect methods produces residual data sets $(\hat{\varepsilon})$. Bootstrap residual data sets are performed in as many as two stages to obtain fast double bootstrap replication $(\hat{\varepsilon}_b^{**})$. The model is estimated using the residual bootstrap data set from the second stage. As shown in Equation (21), the second stage residual bootstrap data set estimates y_b^{**} for the spatial fixed effect.

$$y_b^{**} = [I_{nT} - \hat{\rho}(I_T \otimes W)]^{-1}(X\hat{\beta} + (I_T \otimes \mu) + \hat{\varepsilon}_b^{**}) \tag{21}$$

The second stage of the residual bootstrap data set is utilized to estimate y_b^{**} for spatial random effect equations, as shown in Equation (22). The y_b^{**} that has been obtained on each replication data set is modeled with a SAR model to obtain an estimate of the parameters of each replication data set.

$$y_b^{**} = \rho(I_T \otimes W_{nT})y + X\beta + \varepsilon_b^{**} \tag{22}$$

$$\varepsilon_b^{**} = (\iota_T \otimes I_N)\phi + v \tag{23}$$

The spatial autoregressive model obtains the value of each parameter. The estimator value of the spatial lag parameter is obtained for each replication through the process of iterating the SAR equation. FDB parameter estimated value is shown in Equation (24).

$$\hat{\beta}^{**} = \frac{1}{B}\sum_{b=1}^{B}\hat{\beta}_b^{**} \tag{24}$$

The estimator value of the spatial lag autocorrelation coefficient for each replicated data set ($\hat{\rho}^{**}$) is obtained through the process of iterating the SAR model equations for each replication. Spatial lag autocorrelation coefficient presumption values with the FDB approach are shown in Equation (25).

$$\hat{\rho}^{**} = \frac{1}{B}\sum_{b=1}^{B}\hat{\rho}_b^{**} \tag{25}$$

There are steps in the FDB SAR fixed and random effects. A spatial regression data panel with spatial fixed effect and random effect is performed to obtain an estimate of parameters of $\hat{\rho}_0$ and $\hat{\beta}_0$ using the maximum likelihood method. Then, the residual is calculated as much as B replication, and residual resampling of the first stage is also conducted ($\hat{\varepsilon}_b^{*}$). Next, as many as B replication, residual resampling of the second stage ($\hat{\varepsilon}_b^{**}$) is performed. For each replication, calculation of y_b^{**}, $\hat{\beta}_b^{**}$. $\hat{\rho}_b^{**}$ is generated using y_b^{**} from fixed and random effects. The last step is calculating FDB p-values of $\hat{\beta}_b^{**}$ and $\hat{\rho}_b^{**}$.

III. RESULTS AND DISCUSSIONS

Spatial pattern analysis is used to see the distribution pattern and the relationship between variables between regions. In analyzing the spatial patterns of poverty levels and the factors that are thought to influence poverty levels in the Flores Island, the research uses thematic maps. Thematic maps are a visualization of data presented in map form. In this map, the areas on Flores Island are divided into three classes using the Equal Interval method: low, medium, and high. Data exploration shows that the useful images and information from the data should be considered without jumping to conclusions in general.

Figure 1 shows the pattern of poverty distribution of eight districts in the Flores Island used in the research. Figure 1 explains that during 2018–2020, the spatial pattern of poverty levels has been relatively the same. The territory with a percentage of poor people in the high category includes East Manggarai and Ende districts. As for areas with a low percentage of poor

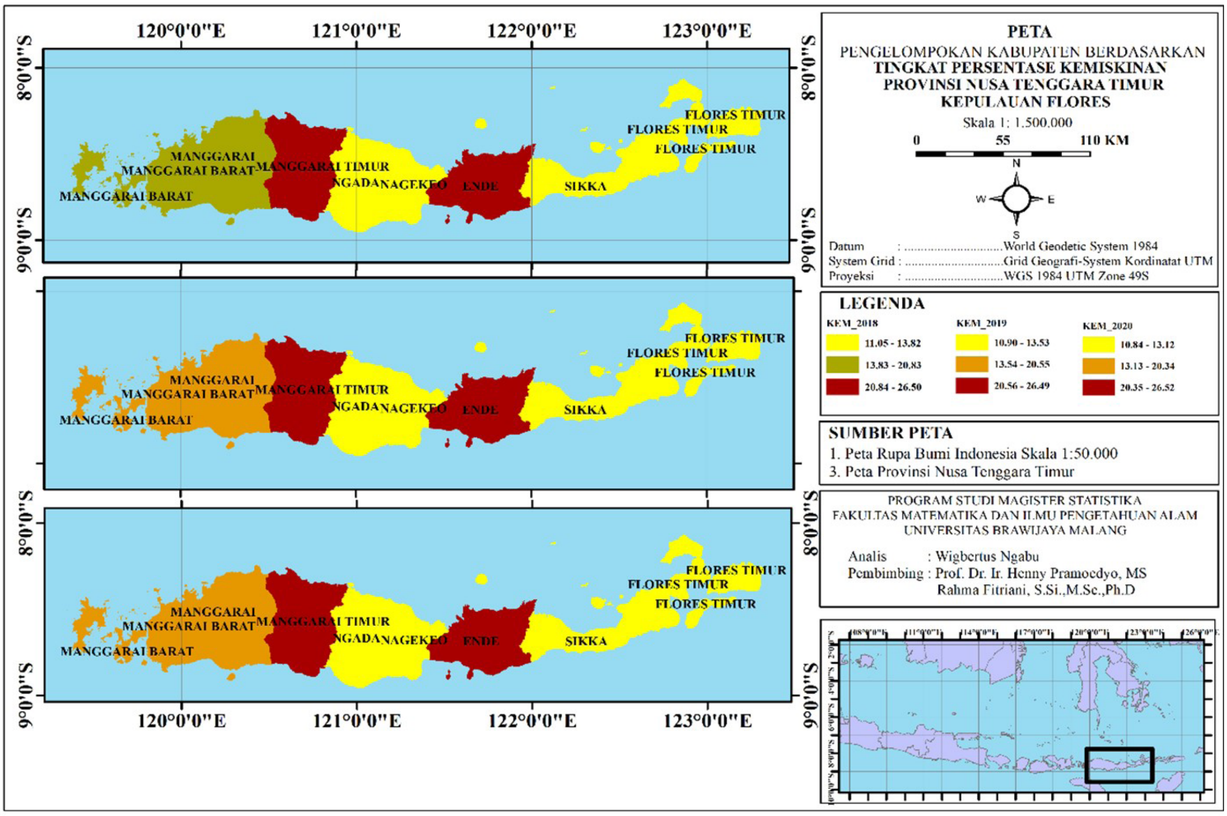

Figure 1 Thematic Map of Poverty Rates in the Flores Island in 2018-2020

people, there are East Flores, Sikka, Nagekeo, and Ngada districts. Then, the percentage of poor people in the moderate category includes West Manggarai and Manggarai districts. Among the regions with moderate and low percentages of poor population, the areas are close together. Hence, there is a spatial dependence between these regions.

Spatial modeling of poverty rates in the districts of Flores Island is conducted using the cluster of FDB approach utilizing residual values and dividing it into three clusters. The cluster method used in the panel data is based on location and time. Then, the cluster based on location is also used. SAR modeling of random effect with a cluster of FDB approach on the poverty rates in Flores Island results in an estimated parameter with 1.000 times repetition.

The first step is conducting the Hausman test to identify whether the fixed or random effects are more suitable to be used on the model. The test also aims to see whether there are random effects in the spatial panel data. Then, spatial dependency tests using Moran's I and LM are conducted to identify whether there is a dependency on spatial lag or error model.

Matrix weighting used is Rook Contiguity, followed by the Hausman test to produce the selected model. From Table 1, Hausman SAR and Spatial Error Model (SEM) tests obtain a bigger p-value of than α (0,05) or smaller t-value than $\chi^2_{(\alpha,N-K)}$= 7,815. Thus, it is concluded that the random effect model is better than the fixed model effect on poverty level modeling in Flores Island. Furthermore, testing is carried out using a random effect model.

Table 1 The Results of Hausman Test

Test Statistics	LM	P-Value
LM lag	15,902	$6{,}672 \times 10^{-5}$
Robust LM lag	10,036	0,001536
LM error	8,8527	0,003851
Robust LM error	2,4866	0,1148

Moran's I spatial dependency and LM tests are the next steps. Moran's I tests the spatial effects of poverty-level locations on Flores Island. Based on Table 2, the LM lag test results in a p-value of 6,672×10-5 and the LM error test with a p-value of 0,03851. These results indicate spatial dependencies are in the lag model and spatial error panel. The tested robust LM lag obtains a p-value of 0,00153 and a robust LM error with a p-value of 0,1148. It shows that the spatial model of the poverty rate panel in Flores Island is an autoregressive spatial model (SAR).

Then, a random effect in the spatial autoregressive test is conducted. It results in a coefficient of determination (R^2) value of 0,7738. Therefore, the residual normality assumption of the random effect spatial model is not met.

Table 2 LM Value Test in Spatial Regression Model

Test Statistics	LM	P-Value
LM lag	15,902	$6{,}672 \times 10^{-5}$
Robust LM lag	10,036	0,001536
LM error	8,8527	0,003851
Robust LM error	2,4866	0,1148

The next step is testing spatial panels with the FDB approach. The first step to model the poverty rate in districts of Flores Island with the FDB approach is to test spatial dependency using the FDB Moran's I and FDB Lagrange Multiplier (LM) tests. Then, the FDB robust LM is performed to identify whether spatial dependencies occur on the dependent variables in model errors or if there are autoregressive terms on SAR and SEM models using α=5%. The spatial dependency test for the projected FDB approach is shown in Table 3.

Table 3 Spatial Statistical Values Dependency on FDB Approaches

Test statistics	Values	P-Value
Moran's I	0,3616	0,0272
LM lag	1,3295	0,0101
Robust LM lag	1,0598	0,0085
LM error	1,2458	0,2791
Robust LM error	0,7555	0,7069

Table 3 shows the results of spatial dependency testing for the FDB approach. FDB Moran's I, LM lag test with 0,0272 and 0,0101 are smaller than α=0,05. The results indicate the spatial dependency on the percentage of poverty districts in Flores Island. Then, the FDB LM error with a p-value of 0,2791 shows no spatial dependency on the panel spatial error model. Meanwhile, FDB robust LM lag test obtains a p-value of 0,0085 and FDB robust LM error has a p-value of 0,7069. The results indicate that the spatial model of the poverty-level panel in Flores Island is SAR with the FDB approach.

The next step is to test the spatial autoregressive random effect model with the FDB approach. The spatial values of the autoregressive model with the FDB approach are shown in Table 4. From Table 4, the coefficient of determination (R^2) shows that the 88,64% of poverty rate in Flores Island can be explained by all five independent variables using spatial autoregressive random effects of the FDB approach. Then, variables that significantly affect the dependent variables in the SAR random effect model use α = 5%. It includes GRDP and RMW. It is indicated by a smaller p-value than α = 5%.

Table 4 SAR Random Effect Regression Coefficient Values with FDB Approach

Variable	Coefficient	P-value
EYS (X_1)	2,0579	0,3438
GRDP (X_2)	-0,3664	0,0133
ALE (X_3)	0,2795	0,7101
RMW (X_4)	-5,6983	0,0414
UR (X_5)	-1,0668	0,4217
R-Squared	0,8864	
ϕ	0,1224	
ρ	0,1600	

The significance test for each parameter produces two significant variables. So, a good model is formed using significant variables. The SAR random effect model with the FDB approach obtains the following calculation.

$$\hat{y}_{it} = -0.160 \sum_{i=1}^{n} w_{ij} y_{jt} - 0.3664\ X_{2t} - 5.5683\ X_{4t} + 0.1224$$

Then a Spatial Autoregressive random effect model is formed for each location. For example, the random effect Spatial Autoregressive (SAR) model for the Ende district is as follows.

$$\hat{y}_{Sikka\ ke-t} = -0.0800\ y_{Flores\ timur\ ke-t} - 0.0800\ y_{Ende\ ke-t} - 0.3664\ X_{2\ Sikka\ ke-t} - 5.5683\ X_{4\ Sikka\ ke-t} + 0.1224$$

From the mentioned model, the poverty rate in East Flores and Ende districts has a role of 0,0800% in the poverty in Sikka district. Each reduction in the percentage of poverty in the two adjacent areas have an impact on reducing the percentage of poverty in Sikka district by 0,0800%. Then, the residual normality assumption is met at a fairly small observation measure using the FDB SAR random effect approach. The results from the FDB approach lead to better results. Moreover, there are improvements in testing assumptions of normality. The plot normality of SAR random effect parameters with the FDB approach is presented in Figure 2. Based on Figure 2, the estimated value of parameters obtained by the FDB SAR random effect approach with a looping of 1.000 times meets the normal distribution (limiting normal distribution).

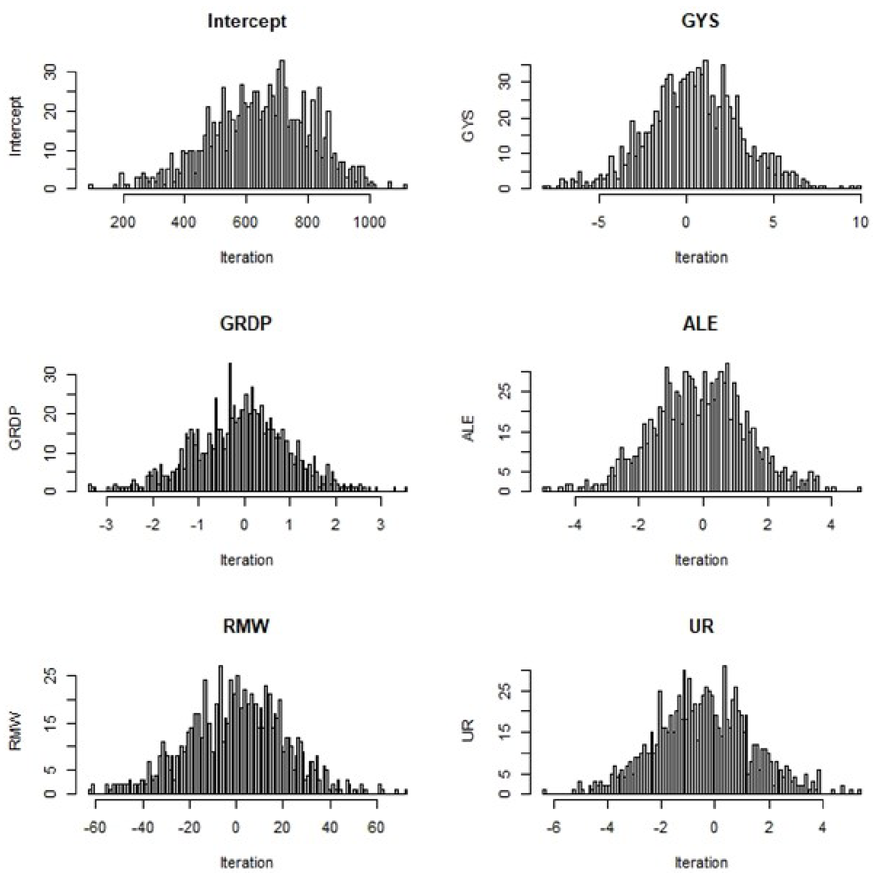

Figure 2 Normality Plot with Fast Double Bootstrap Spatial Autoregressive Random Effect

IV. CONCLUSIONS

Analysis of SAR random effect modeling of poverty rates in Flores Island with the FDB approach produces a more accurate distribution of presumption and effective value than the usual SAR random effect method. The random effect in spatial modeling autoregressive model with the FDB approach results in a higher R-squared value of 88,64% of the original spatial autoregressive random effect model, namely R-squared 77,38%. It shows that cluster in FDB of Spatial Autoregressive (SAR) random effect can explain the diversity of poverty rates in Flores Island by 88,64% using five variables under the research. In addition, the normality assumption is met at a fairly small observation size through the FDB SAR random effect approach. The results reveal that independent variables providing significant effects on poverty rates in Flores Island are gross regional domestic product (PDRB) and Regional Minimum Wage (RMW). Spatial testing for the FDB approach leads to better results. In addition, there are improvements in assumptions of small samples.

The limitations in the research are more focused on examining the use of FDB in spatial regression of panel data with spatial random and fixed effects. It obtains a spatial random effect estimation model with an FDP approach to the problem of poverty levels in Flores Island, East Nusa Tenggara Province. Then, it compares the more effective results of using the FDP method. In future research, spatial dependency statistical tests must be developed with an FDB approach that considers the existence of outlier data. Then, it needs to be developed with a Bayesian approach. The Bayesian approach will accommodate the problem of not fulfilling the normal distribution and small sample size.

REFERENCES

Badan Pusat Statistik Provinsi Nusa Tenggara Timur. (2020). *Ringkasan data dan informasi kemiskinan Provinsi Nusa Tenggara Timur 2019*. Retrieved from https://ntt.bps.go.id/publication/2020/03/06/e311fb1c494cced6e63de93a/ringkasan-data-dan-informasi-kemiskinan-provinsi-nusa-tenggara-timur-2019.html

Canay, I. A., Santos, A., & Shaikh, A. M. (2021). The wild bootstrap with a “small” number of “large” clusters. *The Review of Economics and Statistics, 103*(2), 346–363.

Choi, J. E., & Shin, D. W. (2020). Block bootstrapping for a panel mean break test. *Journal of the Korean Statistical Society, 49*(3), 802–821.

Davidson, R., & Trokić, M. (2020). The fast iterated bootstrap. *Journal of Econometrics, 218*(2), 451–475.

Djogbenou, A. A., MacKinnon, J. G., & Nielsen, M. Ø. (2019). Asymptotic theory and wild bootstrap inference with clustered errors. *Journal of Econometrics, 212*(2), 393–412.

Du, K., Worthington, A. C., & Zelenyuk, V. (2018). Data envelopment analysis, truncated regression and double-bootstrap for panel data with application to Chinese banking. *European Journal of Operational Research, 265*(2), 748–764.

Elhorst, J. P. (2017). Spatial panel data analysis. *Encyclopedia of GIS, 2*, 2050–2058.

Galvao, A. F., Parker, T., & Xiao, Z. (2021). *Bootstrap inference for panel data quantile regression*. Retrieved from https://arxiv.org/abs/2111.03626

LaFontaine, D. (2021). The history of bootstrapping: Tracing the development of resampling with replacement. *The Mathematics Enthusiast, 18*(1), 78–99.

Lee, L. F., & Yu, J. (2020). Initial conditions of dynamic panel data models: On within and between equations. *The Econometrics Journal, 23*(1), 115–136.

Li, L., Hong, X., & Peng, K. (2019). A spatial panel analysis of carbon emissions, economic growth and high-technology industry in China. *Structural Change and Economic Dynamics, 49*(June), 83–92.

Liu, S. F., & Yang, Z. (2020). Robust estimation and inference of spatial panel data models with fixed effects. *Japanese Journal of Statistics and Data Science, 3*, 257–311.

Lütkepohl, H., & Schlaak, T. (2019). Bootstrapping impulse responses of structural vector autoregressive models identified through GARCH. *Journal of Economic Dynamics and Control, 101*(April), 41–61.

MacKinnon, J. G., Nielsen, M. Ø., & Webb, M. D. (2021). Wild bootstrap and asymptotic inference with multiway clustering. *Journal of Business & Economic Statistics, 39*(2), 505–519.

MacKinnon, J. G., & Webb, M. D. (2018). The wild bootstrap for few (treated) clusters. *The Econometrics Journal, 21*(2), 114–135.

Mameli, V., Musio, M., & Ventura, L. (2018). Bootstrap adjustments of signed scoring rule root statistics. *Communications in Statistics-Simulation and Computation, 47*(4), 1204–1215.

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., ... & Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *eLIFE, 7*, 1–24.

Ou, B., Long, Z., & Li, W. (2019). Bootstrap LM tests for spatial dependence in panel data models with fixed effects. *Journal of Systems Science and Information, 7*(4), 330–343.

Roodman, D., Nielsen, M. Ø., MacKinnon, J. G., & Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal, 19*(1), 4–60.

Schmidt, P. (2020). *Econometrics*. CRC Press.

Schuldt, A., Ebeling, A., Kunz, M., Staab, M., Guimarães-Steinicke, C., Bachmann, D., ... & Eisenhauer, N. (2019). Multiple plant diversity components drive consumer communities across ecosystems. *Nature Communications, 10*, 1–11.

Suparman, Y., & Ginanjar, I. (2021). A fixed effect panel spatial error model in identifying factors of poverty in West Java Province. *Journal of Physics: Conference Series, 1776*, 1–10.

Suryowati, K., Bekti, R. D., & Faradila, A. (2018). A comparison of weights matrices on computation of dengue spatial autocorrelation. *IOP Conference Series: Materials Science and Engineering* (Vol. 335, pp. 1–9). IOP Publishing.

Wang, Z., & Lam, N. S. N. (2020). Extending Getis–Ord statistics to account for local space–time autocorrelation in spatial panel data. *The Professional Geographer, 72*(3), 411–420.

Wang, J., Yamamoto, T., & Liu, K. (2021). Spatial dependence and spillover effects in customized bus demand: Empirical evidence using spatial dynamic panel models. *Transport Policy, 105*(May), 166–180.

Yang, Z. (2018). Bootstrap LM tests for higher-order spatial effects in spatial linear regression models. *Empirical Economics, 55*, 35–68.

Yolanda, A. M., & Yunitaningtyas, K. (2019). Spatial data panel analysis for poverty in East Java Province 2012-2017. *Journal of Physics: Conference Series, 1265*, 1–9.