

# Geo-additive Models in Small Area Estimation of Poverty

Novi Hidayat Pusponegoro<sup>1</sup>, Anik Djuraidah<sup>2</sup>, Anwar Fitrianto<sup>3</sup>, I Made Sumertajaya<sup>4</sup>

<sup>1</sup> Department of Statistics, STIS Polytechnic of Statistics

Jalan Otto Iskandardinata No.64C, Jakarta Timur, Indonesia

<sup>1</sup> novie@stis.ac.id

<sup>1,2,3,4</sup> Department of Statistics, Bogor Agriculture University

Jalan Raya Dramaga, Bogor, Indonesia

<sup>2</sup> anikdjuraidah@gmail.com

<sup>3</sup> anwarstat@gmail.com

<sup>4</sup> imsjaya@gmail.com

Received on 08-12-2018, revised on 19-02-2019, accepted on 09-04-2019

## Abstract

The main flaw in direct estimation especially in poverty researches is the sample adequacy fulfilment, otherwise it will produce large estimate parameter variance. The Small Area Estimation (SAE) developed to handle that flaw. Since, the small area estimation techniques require “borrow strength” across the neighbor areas thus SAE was developed by integrating spatial information into the model, known as Spatial SAE. SAE and spatial SAE model require the fulfilment of covariate linearity assumption as well as the normality of the response distribution that is sometimes violated, and the geo-additive model offers to handle that violation using the smoothing function. The purpose of this paper is to compare the SAE, Spatial SAE and Geo-additive model in estimating a sub-district level mean of per capita income using the poverty survey data in Bangka Belitung province in 2017. The findings of the paper are the Geo-additive is the best fit model based on AIC, so form of relationship between response and covariate is the most important part of modelling.

**Keywords:** EBLUP, spatial EBLUP, geo-additive, income per capita

## I. INTRODUCTION

Spatial data is data with territorial reference containing two information that are observation information and region information. For example: the number of people with certain diseases in some areas of the country [1]. The regional information can present spatial patterns for data presented in geographical units such as municipalities, districts or sub-districts. In social studies, geographic information is also able to provide the pattern of poverty [2, 3, 4, 5, 6, 7 and 8].

Policy makers and researchers take into consideration the information of poverty such as the well-being or the living condition of people in a certain area, since poverty is the first Sustainable development goals that carried by the United Nations. Nevertheless, the poverty estimation only allows for larger regions or larger population subgroups since the limitations in its data collection methodologies. Less of sample adequacy to deliver direct estimation is one of those limitation, so the Small Area Estimation (SAE) developed by sample survey statisticians. The existing auxiliary information from the neighborhood areas are useful in small area estimation techniques [9].

The SAE model is a mixed model with variance within the sub-populations can be explained by fixed and random effect. Respectively, those effects correspond with the variance in auxiliary and specific variance of sub-populations. Fay and Herriot were the first researchers that developed a small area statistics based on linear mixed model with empirical Best Linear Unbiased Prediction (BLUP) as a solution method [10]. Borrowing strength across neighbor

areas is the determination of small area estimation techniques and EBLUP approach was developed by integrating spatial information into the model. Spatial Empirical Best Linear Unbiased Prediction (SEBLUP) is a parameter estimation method that take into account the random effects of spatial correlation in EBLUP. The SEBLUP method is developed from area-level SAE and it escalated the variance covariance structure of small area models with spatial correlations between areas. This method was proposed by Salvati [11] and has been applied by Petrucci et.al [12], Singh et al. [13, 14].

The normality assumptions on the response variable are required in the estimation of the linear mixed model. However, many data found had not normal distributions, so logarithm transformation is needed in order to make the response variable distribution fit to normal distribution [15]. In fact, the issue of spatial modelling is not only about the distribution but also form of correlation between covariate spatial data and response variables are non-linear relationships [16]. Thus, appropriate estimation methods are needed to address the problem.

An alternative method that can be used in the estimation of small area statistics is the Geo-additive Model. Geo-additive model was proposed by Kammann & Wand [17], is a method of combining kriging and additive models presented in the form of linear mixed model. This model has the flexibility in specifying the form of relationship between responses with spatial information, because it forms a smoothing function. Since, the Fay-Herriot model and the spatial Fay-Herriot model require the fulfillment of covariate linearity assumption as well as the normality of the response distribution that is sometimes violated and the geo-additive model offers those handling using the smoothing function. Therefore, the purpose of this paper is to verify the spatial effect by comparing the SAE, Spatial SAE and Geo-additive model in order to estimate the mean of per capita income of each sub-district using the poverty survey data in Bangka Belitung province at 2017 derived from the Institute of statistics.

This paper is organized into five sections with the background, motivation and purpose of this research are presented at first section. Second section is a review of the latest literature that is related to this research topic. This section also clarifies the position of this research among similar researches that has been done. The methodology and empirical results are presented in section 3 and 4. The final section summarizes the main findings of the analysis and discusses the further possible researches.

## II. SPATIAL INFORMATION IN SMALL AREA ESTIMATION

Survey is an effective way of collecting data, but the necessity of estimating population characteristics for insufficient sample adequacy leads to the usage of small area estimation methods as one of the indirect estimation methods. Based on the availability of supporting data, there are two basic model types of SAE, i.e. area level model and unit level model [9]. Area level model is a special form of mixed linear model where the auxiliary information are supporting data that exist only for a given area level. While the unit level model is a model with auxiliary information are individually data that compatible to the response and it well-known as nested error regression model.

Area level SAE models can be presented in the form of a linear model with spatial patterns that accounted as random effects and it can be used to improve the estimation [10]. The formula of the model can be expressed as:

$$\hat{\boldsymbol{\vartheta}} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{u} + \mathbf{e}, \quad (1)$$

where  $\boldsymbol{\vartheta}$  be the  $m \times 1$  vector of the parameters of inferential interest and assume that the direct estimator  $\hat{\boldsymbol{\vartheta}}$  is available,  $\mathbf{e}$  is a vector of independent sampling errors with mean vector  $\mathbf{0}$  and known diagonal variance matrix  $\mathbf{R} = \text{diag}(\varphi_i)$ ,  $\varphi_i$  representing the sampling variance of the direct estimators of the area parameters of interest. Thus,  $\boldsymbol{\alpha}$  is the  $p \times 1$  vector of regression parameters,  $\mathbf{u}$  is the  $m \times 1$  vector of independent random area specific effects with zero mean and  $m \times m$  covariance matrix  $\Sigma_u = \sigma_u^2 I_m$ . The Empirical Best Linear Unbiased Predictor (EBLUP) is extensively used to obtain model based indirect estimators of small area parameters  $\boldsymbol{\vartheta}$  and associated measures of variability, this is of the form:

$$\hat{\boldsymbol{\vartheta}}_i^{FH} = \hat{\gamma}_i \hat{\boldsymbol{\vartheta}}_i + (1 - \hat{\gamma}_i) \mathbf{x}_i^T \hat{\boldsymbol{\alpha}} \quad (2)$$

where  $\hat{\gamma}_i = \hat{\sigma}_u^2 / (\hat{\sigma}_u^2 + \varphi_i)$  and the weighted least squares estimate of  $\boldsymbol{\alpha}$  is  $\hat{\boldsymbol{\alpha}}$  with weights  $(\hat{\sigma}_u^2 + \varphi_i)^{-1}$ .

One of the Fay-Herriot model assumption is the effect between areas is independent, whereas in fact it will be violated. Based on Tobler's first law of geography which is "everything is related to everything else but the related things are more related than distant things". Kriging was the very first model that took account spatial correlation that assumed spatial dependency follows the Conditional Autoregressive process [18], and it has been applied mainly in mining and geology [19, 20]. Thus, Simultaneous Autoregressive (SAR) process developed with the spatial dependence is incorporated into the error component of a random factor [21].

Based on, the SAR model which its random effect is a function of the spatial weighted matrix and spatial autoregressive coefficients, subsequently the spatial SAE model developed. The formula of spatial SAE model can be written as:

$$\hat{\boldsymbol{\vartheta}} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{D}\boldsymbol{\nu} + \boldsymbol{\epsilon}, \quad (3)$$

where  $m \times 1$  vector of spatially correlated random area effects  $\boldsymbol{\nu}$  is given by the following autoregressive process with spatial autoregressive coefficient  $\rho$ :

$$\boldsymbol{\nu} = \rho \mathbf{W} \boldsymbol{\nu} + \boldsymbol{u} \Rightarrow \boldsymbol{\nu} = (\mathbf{I}_m - \rho \mathbf{W})^{-1} \boldsymbol{u},$$

and  $D$  is a  $m \times m$  matrix of known positive constants, and  $W$  is  $m \times m$  spatial interaction matrix. The covariance matrix of  $\boldsymbol{\nu}$  is the  $m \times m$  Simultaneously Autoregressive (SAR)  $\mathbf{G}(\delta) = \hat{\sigma}_u^2 [(\mathbf{I} - \rho \mathbf{W}^T)(\mathbf{I} - \rho \mathbf{W})]^{-1}$  and the covariance matrix of  $\hat{\boldsymbol{\vartheta}}$  is given by  $\mathbf{V}(\delta) = \mathbf{R} + \mathbf{D}\mathbf{G}\mathbf{D}^T$  where  $\delta = (\hat{\sigma}_u^2, \rho)$ . So that, the estimator of  $\boldsymbol{\vartheta}$  named as Spatial Empirical Best Linear Unbiased Predictor (SEBLUP) is:

$$\hat{\boldsymbol{\vartheta}}_i^{SFH} = \mathbf{x}_i \hat{\boldsymbol{\alpha}} + \mathbf{b}_i^T \hat{\mathbf{G}} \mathbf{D}^T (\mathbf{R} + \mathbf{D} \hat{\mathbf{G}} \mathbf{D}^T)^{-1} + (\hat{\boldsymbol{\theta}} - \mathbf{x} \hat{\boldsymbol{\alpha}}) \quad (4)$$

where  $\hat{\boldsymbol{\alpha}} = (\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{X})^{-1} \mathbf{X} \hat{\mathbf{V}}^{-1} \hat{\boldsymbol{\vartheta}}$  and  $\mathbf{b}_i^T$  is a  $1 \times m$  vector  $(0, 0, \dots, 0, 1, \dots, 0)$  with value 1 for the  $i$ -th area. Geographic information and spatial models can be valuable in small area parameters estimation if those are handled appropriately [11, 12]. Geographic boundaries of small areas are generally defined according to administrative criteria is an example of the geographic information [13]. Thus, the random effects between neighboring regions can also be replaced by the criteria of contiguity.

Lately, the fields of spatial statistics is widely used in order to describe the spatial structure or analyze the geographical pattern of the interesting phenomena by spatial location data. Considering the usage of spatial location data, it needs bivariate smoothing techniques to obtain a surface estimate that exploit knowledge of the spatial coordinates (latitude and longitude), such as kernel estimate or kriging. Not only produce a map, the usage of bivariate smoothing also can be applied to handle the non-linear relation between any two continuous predictors and a response variable [22, 23]. Moreover, the spatial information alone does not properly explain the pattern of the response variable, even some covariates are in more complex model.

In order to handle the spatial distribution of analysis while to take account the possible non-linear covariate effects, Kammann and Wand developed the Geo-additive model [17]. Kammann and Wand merged kriging with the additive model to obtain a single mixed model. Thus, the first half of the model formulation involves a low rank mixed model representation of additive models. Then, incorporation of the geographical component is achieved by expressing kriging as a linear mixed model. The Geo-additive model can be written as:

$$y_i = \beta_0 + f(s_i) + g(t_i) + \beta_1^T \mathbf{x}_i + S(\mathbf{x}_i) + \varepsilon_i \quad (5)$$

where in the first part of the models,  $t_i$  and  $y_i$  represents measurements on two predictors  $s$  and  $t$  and a response variable  $y$  for  $i$ -th unit,  $f$  and  $g$  are smooth, functions of  $s$  and  $t$  respectively. The second part of the model is the simple universal kriging model with  $S$  is the geographical component, with  $\mathbf{x}_i \in R^2$  represents the geographical location and  $\{S(\mathbf{x}): \mathbf{x}_i \in R^2\}$  is a stationary zero-mean stochastic process. The error  $\varepsilon_i$  assume to be independent zero-mean random variables with common variance  $\sigma^2$  and independently distributed [22]. Using the thin plate spline, Kammann and Wand [17] chose the penalized spline to construct the  $f$  and  $g$  function which are written as follows:

$$f(s_i) = \beta_{0s} + \beta_s s_i + \sum_{k=1}^{K_s} u_k^s z_k^s(s_i)$$

$$g(t_i) = \beta_{0t} + \beta_t t_i + \sum_{k=1}^{K_t} u_k^t z_k^t(t_i)$$

where  $u_k^s \sim N(0, \hat{\sigma}_s^2)$ ,  $u_k^t \sim N(0, \hat{\sigma}_t^2)$  and  $(z_1^t, K, z_K^t) = (t_i - K_k^t)(z_1^s, K, z_K^s) = (s_i - K_k^s)$  for  $1 \leq i \leq 1 \leq k \leq K$ . The geo-additive model can be expressed as a linear mixed model which allows for estimation and inference using standard methodology then the model representation in the matrix form can be written as follows:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon} \quad (6)$$

with  $\mathbf{X} = [1 \ s_i \ t_i \ x_i]_{1 \leq i \leq n} = [\beta_0 \ \beta_s \ \beta_t \ \beta_x^T]^T$  and  $\mathbf{Z} = (\mathbf{Z}_s \ | \ \mathbf{Z}_t \ | \ \mathbf{Z}_x)$ . Furthermore  $\mathbf{u}' = (\mathbf{u}_s \ \mathbf{u}_t \ \bar{\mathbf{u}})$  and  $Cov(\boldsymbol{\varepsilon}) = \begin{pmatrix} \sigma_s^2 I & 0 & 0 & 0 \\ 0 & \sigma_t^2 I & 0 & 0 \\ 0 & 0 & \sigma_x^2 I & 0 \\ 0 & 0 & 0 & \sigma_\varepsilon^2 I \end{pmatrix}$ .

Utilizing the general linear mixed model framework of small area estimation models and additive models, it can be defined SAE geo-additive model. Thus, we can say that in a geo-additive model the linear mixed model structure allows to include the area-specific effect as an additional random components. In particular, a geo-additive SAE model has two random effect components: the area-specific effects and the spatial effects [24].

Nowadays, many studies have developed or applied geo-additive models [see 24, 25] such as in small area approach, in long-term data and so on. Another extension to the work of Kammann and Wand is the geo-additive model proposed to analyze the spatial distribution of grain weight, a commonly used indicator of wheat production, taking into account its nonlinear relations with other crop features [26]. Spatial models also useful in the small areas estimation of official statistics [27, 28]. And lately, there is an improvement approach in smoothing function determination with K-Spline [16]. Based on the above literature, this paper conducted Fay-Herriot model and compared the spatial model such as Spatial SAE and geo-additive models in the poverty analysis, with Euclidean distance as spatial weighted.

### III. DATA AND METHODOLOGY

#### A. Data

Bangka Belitung is a province with the fewest sub-district in Sumatra. It has 387 sub-districts that spread in seven municipalities i.e. Bangka, West Bangka, Central Bangka, South Bangka, Pangkalpinang, Belitung and East Belitung. Bangka Belitung also has the smallest head count index among provinces in Sumatra that is 5.3% in 2017 with poverty line is Rp.441579.91. However, its poverty line value is the highest value among provinces in Sumatra. The poverty line represents the sum of the minimum food and non-food needs expenditure values, the higher the value reflects the high prices of the goods in the market or can also indicate the high income of the population in a region. Since 2006, East Belitung has the highest poverty line among municipalities in Bangka Belitung Province.

One of substantial poverty indicator is regional mean of per capita income. In this paper, mean of per capita income is estimated from some auxiliary variables such as infrastructure, electricity and regional accessibility. Many researchers had studied that infrastructure and electricity as a proxy of technological development and its existence can elevate access to basic services such as education and health [30, 31, and 32]. Moreover, the level of regional accessibility also has significant relationship with poverty [33, 34]. Regional accessibility that defined by the distance of region to city center represents access to social economic facilities. Thus, this study use 2 auxiliary variable which are household total in each sub-district that have access to electricity from Indonesia Electricity Company and distance the sub-district center to the municipality office. Data set of this study derived from the poverty survey data in Bangka Belitung province at 2017 by the STIS Polytechnic of Statistics that consist of 139 sub-districts data.

### B. Methodology

Fay-Herriot (FH) model, Spatial Fay-Herriot (SFH) and the Geo-additive model (GAM) were conducted and compared in order to fulfill the purpose this study. As response variable, sub-district mean per capita income distribution is skewed [15] thus it was transformed in log value. Using FH model, the spatial information accounted as random effect and formulation of the model can be expressed as follows:

$$\log \text{income}_i = \alpha_1 \text{electricity} + \alpha_2 \text{distance} + \boldsymbol{u} \quad (7)$$

where:

- $\log \text{income}_i$  : the  $i$ -th sub-district log mean per capita income,  
 $\text{electricity}_i$  : household total in the  $i$ -th sub-district that have access to electricity from Indonesia Electricity Company,  
 $\text{distance}_i$  : distance the  $i$ -th sub-district center from the municipality office,  
 $\boldsymbol{u}$  : Independent random area specific effects.

On the other hand SFH model replaces independent random area specific effects by function of matrix spatial correlation that is invers Euclidean distance matrix. In order to handle the spatial distribution analysis while to take account the possible non-linear covariate effects, the GAM formulation can be written as:

$$\log \text{income}_i = \beta_0 + s(\text{electricity}) + s(\text{distance}) + s(\text{spatial location data}) \quad (8)$$

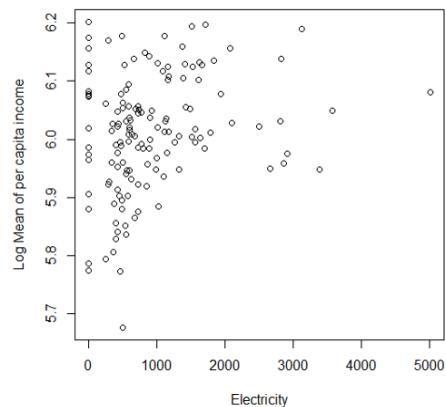
where  $s(\text{electricity})$  and  $s(\text{distance})$  are smoothing function of household total in each sub-district that have access to electricity from Indonesia Electricity Company and distance the sub-district center to the municipality office. While,  $s(\text{spatial location data})$  is a smoothing function of 2 variables that is longitude and latitude.

Akaike's Information Criteria (AIC) is used as information criteria [36], in order to select the best fit model for the poverty data in Bangka Belitung province despite of the parsimonious model.

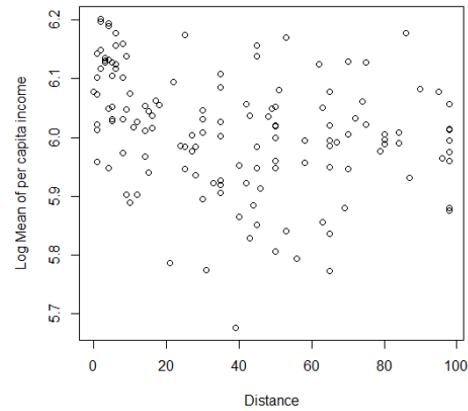
## IV. RESULTS AND ANALYSIS

First of all, data exploration is a crucial step in data analysis. It is used to examine the distribution of the variables and the correlation patterns among them. Based on [37], it is known that relationship between the auxiliary variables and 139 sub-district log mean per capita income in Bangka Belitung province are not sufficiently linier but also conceive of spatial correlations.

This paper gives an explanation about the correlation among the variables even in non linear form. The preliminary analysis are conducted in order to knots setting of GAM. Based on the correlation among auxiliary variables and log of mean per capita income that presented in Fig. 1 and Fig. 2, there are three possible points that are used as the basis at the smoothing function specification. The three points are  $K = 6, 10, 20$  and  $K = 16, 25, 50$  respectively for electricity access and distance variables. Furthermore, the selection evaluation results are presented in Table 1.



*Fig. 1. Scatter plot of household total that have access to electricity from Indonesia Electricity Company and the log sub-district mean per capita income*



*Fig. 2. Scatter plot of distance and the log sub-district mean per capita income*

Table 1 presents that AIC for the knots setting in electricity are stable when the knots setting in distance are K= 16 and 50. So does, the p-values of the parameters estimate in Table 1. Thus, this paper sets K= 6 and 16 for electricity access and distance since they have minimum AIC and p-values among others. In order to obtain the fit model, AIC of the FH model, SFH model and GAM are -250.725, -251,472 and -285.714, respectively.

TABLE 1. AIC AND P-VALUES OF SIGNIFICANT PARAMETERS BY KNOTS SETTING IN ELECTRICITY AND DISTANCE

Knot in electricity	Knot in distance								
	16			25			50		
	AIC	P-values of distance	P-values of spatial location	AIC	P-values of distance	P-values of spatial location	AIC	P-values of distance	P-values of spatial location
6	-285.71	0.052	0.059	-285.62	0.055	0.059	-285.58	0.059	0.059
10	-285.71	0.052	0.059	-285.92	0.070	0.051	-285.58	0.051	0.059
20	-285.71	0.052	0.059	-283.93	0.034	0.089	-285.58	0.089	0.059

Based on Table 2, parameter estimate of FH and SFH model have similar values and levels of significant. The same results have been shown by [38]. EBLUP as parameter estimate method of FH, can not be used to estimate the effect of spatial data since it does take into account spatial information in its model. On the other hand SEBLUP uses spatial data however the result is not significant to estimate the effect of spatial information. Thus, it can be concluded that spatial information in SFH model does not provide the actual effect because it is not well formulated. However, spatial information have significant effect to estimate mean per capita income in GAM.

TABLE 2. PARAMETER ESTIMATE RESULTS

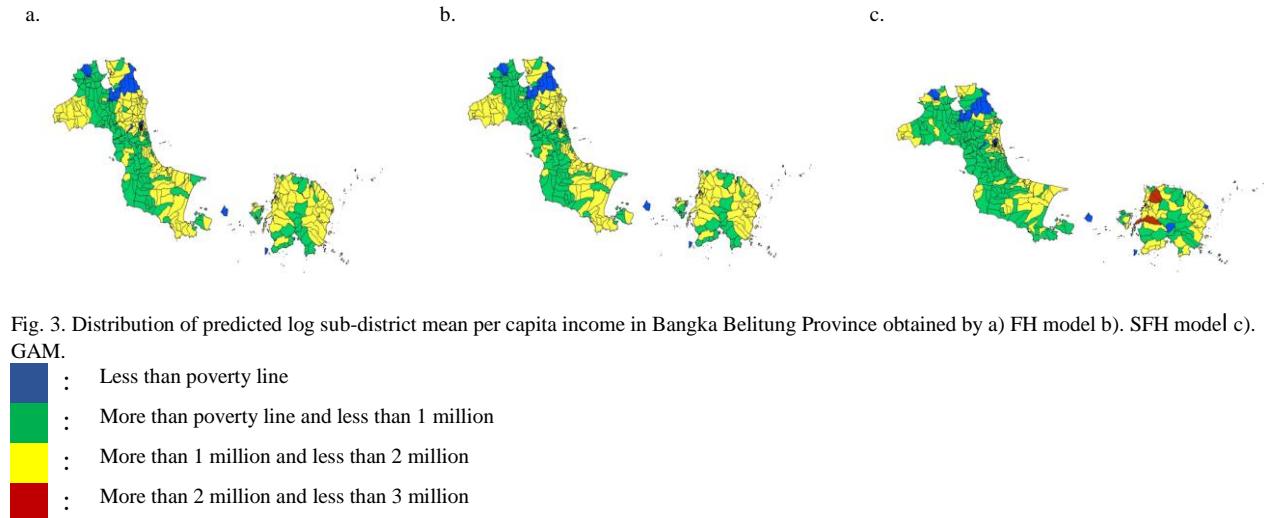
Parameter	FH	SFH	GAM
Intercept	6.02333100	6.02541700	6.011944
Electricity	0.00002396**	0.00002318**	Not sig
Distance	-0.00086498**	-0.00089630**	Sig*
Spatial Data	-	Not Sig	Sig*

\* = Significant at level 5%

\*\* = Significant at level 10%

The distributions of predicted log sub-district mean per capita income in Bangka Belitung province are obtained by those three models and shown in Fig. 3. Fig. 3 illustrates that FH Model and SFH model have the same results in sub-district grouping. Sum of sub-district in Bangka Belitung province which their mean per capita income less than poverty line is 30. And mostly, 198 sub districts agglomerate in third group where the range of mean per capita income

between 1 and 2 millions rupiah. The rest of them agglomerate in second group. However, the Geo-additive model defines different sub district grouping than the other two models. There are 32, 166 and 164 sub-districts where mean per capita income less than poverty line, between poverty line and 1 million rupiah and between 1 and 2 millions rupiah. Even, the model determine the fourth group where its mean per capita income range between 2 and 3 millions rupiah. This group consists of 4 sub-district in East Belitung.



## V. CONCLUSIONS

This paper shows that Geo-additive is the fit model to poverty data in Bangka Belitung province since it have lowest AIC value and it is one and only model that has significant spatial effect . In sum, distance to social economic facilities and spatial location data are the significant auxiliary variables effected the sub-district mean per capita income. The predicted values of mean per capita income which discussed in the previous section bring up to another discussion such as bias of the direct estimate values. Thus, circumspection is a must especially in small area estimation discussion since the responses are the direct estimate values that arose from many different sampling techniques. In order to get deeper spatial effect it is necessary to develop a geo-additive model involving time (temporal) or spatial and time (spatial temporal) effects so that the dynamics of the effect can be examined.

## REFERENCES

- [1] A. S. Fotheringham, “Geographically weighted regression: The analysis of spatially varying relationships”, London: Wiley; (2002).
- [2] S. Friedman and D. T. Lichten, 1998. Spatial inequality and poverty among American children, *Population Research and Policy Review*, 17(2), pp.91–109.
- [3] D. A. Cotter, 2002. Poor people in poor places: Local opportunity structures and household poverty, *Rural Sociology*, 67(4), pp.534–555.
- [4] L. Lobao and R. Saenz, 2002. Spatial inequality and diversity as an emerging research area, *Rural Sociology*, 67(4), pp.497–511.
- [5] G. Betti, F. Ballini, and L. Neri, “Poverty and Inequality mapping in Albania”, Conference of World Bank and Institute of Statistics, Tirana, Albania, (2003).
- [6] L. Neri, F. Ballini, and G. Betti, 2005. Poverty and inequality mapping in transition countries, *Statistics in Transition*, 7, pp.135-157
- [7] I. Molina and J. N. K. Rao, 2010. Small area estimation of Poverty Indicators, *The Canadian Journal of Statistics*, 38(3), pp.369-385.
- [8] A. M. Zewdye, N. M. Aidi, and B. Sartono. 2015. Spatial Econometric Model of Poverty in Java Island, *American Journal of Theoretical and Applied Statistics*, 4(6), pp.420-425.
- [9] J. N. K. Rao, “Small Area Estimation”, London: Wiley; (2003).
- [10] R. E Fay and R. A. Herriot, 1979. Estimation of income from small places: an application of James –stein procedures to census data, *Journal of American Statistical Association*, 74(366), pp.269-277.
- [11] N. Salvati, 2004. Small area estimation by spatial models: the spatial empirical best linear unbiased prediction (Spatial EBLUP), Working paper 2004/03, University of Florence: Department of Statistics “G Parenti”, Florence, Italy.
- [12] A. Petrucci, M. Pratesi, and N. Salvati, 2005. Geographic information in small area estimation: small area models and spatially correlated random area effects, *Statistics in Transition*, 7(3), pp.609–623.
- [13] B. B. Singh, G. K. Shukla, and D. Kundu, 2005. Spatio-temporal models in small area estimation, *Statistics Canada*, 31(2), pp.183-195.
- [14] M. Pratesi and N. Salvati N, 2008. Small area estimation: the EBLUP estimator based on spatially correlated random area effects, *Statistical Methods & Applications*, 17(1), pp.113–141.
- [15] A. Kurnia, “The Best Predictions for Empirical Logarithm Transformation Models in Small Area Estimation with Application on SUSENAS Data”, Ph.D dissertation, Department of Statistics. Bogor Agricultural University, Bogor, Indonesia, (2009).

- [16] Y. Vandendijck, C. Faes and N. Hens, "Geostatistical Analysis using K-splines in the Geoadditive Model", The 30th International Workshop on Statistical Modelling, Johannes Kepler Universitat Linz, (2015).
- [17] E. E. Kammann and M. P. Wand, 2003. Geoadditive models, *Applied Statistics*, 52(1), pp.1–18.
- [18] N. Cressie, 1990. The origins of kriging, *Mathematical Geology*, 22(3), pp.239–252.
- [19] B. B. Trangmar, R. S. Yost, and G. Uehara, 1986. Application of geostatistics to spatial studies of soil properties, *Advances in Agronomy*, 38, pp.45–94.
- [20] P. Goovaerts, 1999. Geostatistics in soil science: State-of-the-art and perspectives, *Geoderma*, 89 (1-2), pp.1-45.
- [21] L. Anselin, "Spatial Econometrics: Method and Models". Boston: Kluwer Academic Publishers, (1992).
- [22] N. A. C. Cressie, "Statistics for Spatial Data". New York: Wiley-Interscience, (1993).
- [23] D. Ruppert, M. P. Wand, and R. J. Carroll, "Semiparametric Regression". New York: Cambridge University Press, (2003).
- [24] C. Bocci, "Geoadditive models for data with spatial information", Ph. D dissertation, Department of Statistics, University of Florence, Italy, (2009).
- [25] D. Ruppert, M. P. Wand, and R. J. Carroll, 2009. Semiparametric regression during 2003–2007, *Electronic Journal of Statistics*, 1(3), pp.1193–1256.
- [26] B. Cafarelli and A. Castrignano, 2011. The use of geoadditive models to estimate the spatial distribution of grain weight in an agronomic field: a comparison with kriging with external drift, *Environmetrics*, 22(6), pp.769–780.
- [27] A. Petrucci and M. Pratesi, 2013. Spatial models in small area estimation in the Context of official statistics, *Statistica Applicata - Italian Journal of Applied Statistics*, 24 (1), pp.9-27
- [28] M. Pratesi, "Analysis of Poverty Data by Small Area Estimation". London: Wiley, (2016).
- [29] F. J. Paraguanas and A. A. Kamil, "Spatial Econometrics Modeling of Poverty", The-8th WSEAS International Conference on Applied Mathematics, Tenerife, Spain, (2005).
- [30] A. M. Balisacan, E. M. Pernia and A. Asra, 2003. Revisiting Growth and Poverty Reduction in Indonesia, *Bulletin of Indonesian Economic Studies*, 39(3), pp.329-353.
- [31] J. W. Fedderke and Z. Bogetic, "Infrastructure and Growth in South Africa: Direct and Indirect Productivity Impacts of Nineteen Infrastructure Measure", Workshop on Infrastructure and Growth of Economic Research Southern Africa, Cape Town, South Africa, (2005).
- [32] S. S. Nugroho, 2015. The Roles of Basic Infrastructure on Poverty Alleviation in Indonesia, *Kajian Ekonomi dan Keuangan*, 19(1), pp.1-96.
- [33] A. W. Astuti and M. Musiyam, 2009. Kemiskinan dan Perkembangan Wilayah di Kabupaten Boyolali, *Forum Geografi*, 23(1), pp.71 – 85.
- [34] R. Rahmawati R and Anik Djuraidah. Regresi Terboboti Geografis dengan Pembobot Kernel Kuadrat Ganda untuk Data Kemiskinan di Kabupaten Jember. *Forum Statistika dan Komputasi*. (Oct 2010), 15(2): 32-37. ISSN : 0853-8115
- [35] T. Piketty and E. Saez, 2003. Income Inequality in the United States 1913-1998, *Quarterly Journal of Economics*, 68(1), pp.1-39.
- [36] Dzaij JJ, Lanza ST, Xu S, 2012. Sensitivity and Specificity of Information Criteria, the Methodology Center, the Pennsylvania State University. Retrieved from <http://methodology.psu.edu>.
- [37] N.H Pusponegoro and R.N Rachmawati. 2018. Spatial Empirical Best Linear Unbiased Prediction in Small Area Estimation of Poverty, *Procedia Computer Science*, 135, pp. 712–718.
- [38] H.A. Zainuddin, "Kajian Transformasi Logaritma untuk Penduga Spatial Empirical Best Linear Unbiased Prediction pada Penduga Area Kecil", Thesis, Department of Statistics. Bogor Agricultural University, Bogor, Indonesia, (2016).