Spatial modeling of poverty level in Kepulauan Bangka Belitung province

Andre Ridho Perdana Hutapea\textsuperscript{1} and Agung Priyo Utomo\textsuperscript{2}

\textsuperscript{1} BPS – Statistics Indonesia, Jakarta, Indonesia
\textsuperscript{2} Politeknik Statistika STIS, Jakarta, Indonesia

Abstract. Poverty is still the biggest problem in the world, including in Indonesia. Based on BPS-Statistics data, in 2017, there are 10.64 percent or about 27.77 millions of Indonesian people are poor. One of the regions in Indonesia that experienced a significant slowdown in poverty reduction is Kepulauan Bangka Belitung Province. Differences in characteristics between regions in Kepulauan Bangka Belitung allow poverty levels to have different causal factors in each region. This study aims to identify and analyze variables that affect poverty levels in Kepulauan Bangka Belitung Province by considering differences in characteristics between regions. The unit of analysis in this study is the sub-district and analyzed using the Geographically Weighted Regression (GWR) model. This study uses data of 6,523 households from all regencies/cities in Kepulauan Bangka Belitung Province, which collected from the field study by students of Politeknik Statistika STIS in 2017. The results of the analysis show that the effect of education levels, health conditions, unemployment rates, the agricultural sector, and labor productivity on poverty levels varies spatially. These results indicate that handling of the poverty problem should pay attention to differences in characteristics between regions (sub-districts).

1. Introduction
End poverty in all its forms everywhere is the first goal of the Sustainable Development Goals (SDGs). However, poverty is still one of the biggest problems in the world, including in Indonesia. Based on data from Badan Pusat Statistik (BPS – Statistics Indonesia), in 2017, there are 10.64 percent or about 27.77 millions of Indonesia people are poor \cite{1,3}.

One of the regions in Indonesia that experienced a significant slowdown in poverty reduction is Kepulauan Bangka Belitung Province. Based on BPS data, the poverty rate in Bangka Belitung in the period 2002-2017 fluctuates from year to year, although it shows a declining trend. In 2002 it was 11.62 percent and in September 2017 it was 5.30 percent \cite{3}, but the decline slowed significantly starting in 2012. In September 2012 it was 5.37 percent while in September 2017 it was 5.30 percent. This shows that the reduction in poverty levels in Kepulauan Bangka Belitung over the past five years was only 0.07 percent, which was the smallest decrease compared to the other provinces in Indonesia. Although the poverty rate has decreased, the number of poor people in Bangka Belitung actually increased by 5.99 thousand people, from 70.21 thousand people in September 2012 to 76.20 thousand people in September 2017 \cite{1}.

Poverty is a complex problem so it needs to be reviewed from various aspects. Identification and analysis of variables that cause poverty need to be done well so that poverty alleviation policies can run effectively. Geographical targeting can be a significant alternative for allocating resources for...
poverty alleviation in developing countries [5]. The lack of a spatial perspective on poverty reduction is reflected in the uniformity of poverty reduction programs implemented in almost all places without regard to the specific characteristics of the region or location where the poor live in [11].

Kepulauan Bangka Belitung as an archipelago has various regional characteristics. Geographical conditions that are divided into Bangka and Belitung islands allow for differences in natural resources, access to infrastructure and economic activities, and human resources which will cause differences in characteristics between regions. Different regional characteristics will affect the level of poverty. This can be seen from BPS data which shows that poverty rates in Bangka Belitung have variations between regions [2]. The percentage of poor people in the regency/city on Belitung island is higher than the percentage of poor people in the regency/city on Bangka island.

The problem that might arise in analyzing the factors that cause poverty in Kepulauan Bangka Belitung is the spatial non-stationarity due to differences in characteristics between regions in Bangka Belitung. Relationships that vary between regions will not be well represented by the global regression model and these global values may be very misleading locally [7].

A variable may be very relevant in one region but not relevant in another region, parameters describing the relationship between certain independent variables and poverty levels may be negative in some regions but positive in other regions, and the same model might accurately explain the causes of poverty in one region but not in other regions. When there is a spatial non-stationarity, the measurement of a relationship depends on the location where the measurement was made [7]. Therefore, we need a method that can produce different local models in each region according to their respective characteristics. This study aims to identify and analyze variables that affect poverty levels in Kepulauan Bangka Belitung Province by considering differences in characteristics between regions.

The method used is Geographically Weighted Regression (GWR). The GWR model can produce different local models in each region. In addition, the results of the GWR model can find out what factors vary spatially.

Research on the factors that cause poverty with the GWR model is not much done. Previous research has been carried in Malawi [4], Bangladesh [8], and Papua [13].

2. Methodology
The global regression model, e.g. OLS, will produce the same poverty level model in each region. The different regional characteristics in Kepulauan Bangka Belitung Province allow these relationships to differ according to regions that can indicate spatial non-stationarity. The GWR model can be used when there is spatial non-stationarity. In the GWR model, the magnitude of the effect of each variable on poverty levels in Kepulauan Bangka Belitung Province, \( \beta(u_i, v_i) \), is different because the measurement depends on the location where the measurement was made.

In the GWR model, each region has its own equation. After the estimation of \( \beta(u_i, v_i) \) is obtained, a t-test is performed to determine whether each independent variable has a significant effect on poverty levels in an area. Furthermore, a goodness of fit test is performed to determine whether the GWR model is better than the OLS model. The final step, comparing the GWR model with the global regression model. This comparison was done to see the ability of the GWR model to predict poverty levels in Bangka Belitung compared to the global regression model.

The data used in this study is the data from the field study which were collected by students of Politeknik Statistika STIS in 2017. Then the data was processed and estimated to obtain aggregate data at the sub-district level. In accordance with the availability of data, this study uses the results of mapping in 2010 as a reference so that the number of observation units is 43 sub-districts. The following table and figure show the names and location map of 43 sub-districts in the Kepulauan Bangka Belitung.
Table 1. Names of the sub-district in Kepulauan Bangka Belitung Province.

| i | Name of sub-district | i | Name of sub-district | i | Name of sub-district | i | Name of sub-district |
|---|----------------------|---|----------------------|---|----------------------|---|----------------------|
| 1 | Mentok              | 12| Merawang            | 23| Koba                | 34| Badau               |
| 2 | Simpang Teritip     | 13| Mendo Barat         | 24| Lubuk Besar         | 35| Tanjung Pandan      |
| 3 | Jebus               | 14| Gerunggang          | 25| Simpang Rimba       | 36| Sijuk               |
| 4 | Kelapa              | 15| Pangkal Balam       | 26| Payung              | 37| Kelapa Kampit       |
| 5 | Tempilang           | 16| Rangkui             | 27| Pulau Besar         | 38| Simpang Renggian    |
| 6 | Belinyu             | 17| Taman Sari          | 28| Air Gegas           | 39| Dendang             |
| 7 | Riau Silip          | 18| Bukit Intan         | 29| Toboali            | 40| Simpang Pesak       |
| 8 | Bakam               | 19| Sungai Selan        | 30| Tukak Sadai        | 41| Gantung             |
| 9 | Pemali              | 20| Simpang Katis       | 31| Lepar Pongok       | 42| Manggar             |
|10 | Sungai Liat         | 21| Pangkalalan Baru    | 32| Selat Nasik        | 43| Damar               |
|11 | Puding Besar        | 22| Namang              | 33| Membalong          |      |                     |

Figure 1. Map of 43 sub-districts in the Kepulauan Bangka Belitung Province.

The independent variables used in this study are based on the sustainable livelihood approach [9]. One outcome of sustainable livelihood is poverty reduction. To obtain this outcome, the population must have a livelihood resource. Livelihood resources are capital that can be used by residents to create livelihoods. The capital consists of human capital (e.g., education and health level), economic/financial capital (e.g., labor conditions), natural resource capital, and social capital. The models used in this study are:

\[ y_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_{1i} + \beta_2(u_i, v_i)x_{2i} + \beta_3(u_i, v_i)x_{3i} + \beta_4(u_i, v_i)x_{4i} + \beta_5(u_i, v_i)x_{5i} + \epsilon_i \text{ where } i = 1, 2, \ldots, 43 \] (1)

with description of the variables,

- \( y_i \) : poverty level in the i-th sub-district;
- \( x_{1i} \) : percentage of the labor force with less than senior high school education in the i-th sub-district;
- \( x_{2i} \) : percentage of the labor force experiencing malnutrition in the i-th sub-district;
- \( x_{3i} \) : percentage of the unemployed labor force in the i-th sub-district;
- \( x_{4i} \) : percentage of the agricultural sector workforce in the i-th sub-district;
- \( x_{5i} \) : labor productivity in the i-th sub-district.
the coordinate points data (longitude, latitude) denoted by \((u_i, v_i)\) are centroids in each sub-district obtained from the mapping.

3. Results

3.1. Descriptive Analysis
The distribution pattern of the percentage of poor people in 43 sub-districts in Kepulauan Bangka Belitung Province in 2017 is shown in Figure 2. It can be seen that there are differences in poverty levels between sub-districts on Bangka and Belitung Islands. Sub-districts on Bangka Island (except in the Pangkalpinang City) have a relatively higher poverty rate compared to sub-districts on Belitung Island.

Figure 2. Percentage of poor population in Kepulauan Bangka Belitung Province according to sub-district.

In addition, it can be seen the distribution pattern of poverty levels of sub-districts in Bangka Belitung cluster in certain areas. This indicates a spatial effect in the data, so further tests need to be carried out to find this out.
The distribution pattern of the percentage of the labor force educated below high school, the percentage of the labor force experiencing malnutrition, the percentage of the labor force unemployed, the percentage of the agricultural sector labor force, and labor productivity have different relationship patterns to the distribution pattern of poverty levels in Kepulauan Bangka Belitung Province (Figure 3). The difference in the relationship pattern indicates the existence of positive or negative relationships of each variable to the poverty level. This indicates that the relationship between the five independent variables and the poverty level in Kepulauan Bangka Belitung Province varies spatially (there may be spatial non-stationarity).

3.2. Models Fitting

The independent variables were used to calibrate Ordinary Least Square (OLS) regression and Geographically Weighted Regression (GWR), which were later evaluated to see their performance in modeling poverty in Kepulauan Bangka Belitung Province.

The first step is to estimate an OLS model to examine the model assumptions and identify the spatial effect that occurs. Table 2 shows there is no multicollinearity, while Table 3 shows the existence of spatial autocorrelation and spatial heterogeneity, which means that the OLS assumptions are violated. So, modeling the relationship between independent variables and dependent variable needs to consider where the measurements were made and consider the effects of spatial dependence.

![Figure 3. Distribution pattern of (a) percentage of labor force with less than senior high school education; (b) percentage of labor force experiencing malnutrition; (c) percentage of unemployed labor force; (d) percentage of the agricultural sector workforce; (e) labor productivity; in Kepulauan Bangka Belitung Province according to sub-district.](image)

| Variables                                          | $\hat{\beta}$ | p-value | VIF |
|----------------------------------------------------|---------------|---------|-----|
| Intercept                                          | 38.8905       | 0.0126  | -   |
| Labor force with less than senior high school education (%) | -0.3702       | 0.0356  | 6.30 |
| Labor force experiencing malnutrition (%)          | -0.2370       | 0.5801  | 1.30 |
| The unemployed labor force (%)                     | 0.9036        | 0.0950  | 1.83 |
| Agricultural sector workforce (%)                  | 0.4881        | 0.0002  | 6.59 |
| Labor productivity (rupiah)                        | -0.0019       | 0.0343  | 1.76 |

The second step is the GWR model fitting. The type of kernel used is adaptive spatial kernel due to the distribution of data points or the centroids in sub-districts in Kepulauan Bangka Belitung Province are not evenly distributed (mainly due to separate between the Bangka Island and Belitung Island). Therefore, there are two alternative weighting functions with bandwidth optimum, namely adaptive Gaussian and adaptive bisquare.
Table 4. Comparison of Gaussian and Bisquare weighting functions.

| Indicators   | Adaptive Gaussian | Adaptive Bisquare |
|--------------|-------------------|-------------------|
| Optimum Bandwidth | 9                 | 21                |
| $R^2$        | 0.7987            | 0.8528            |
| Adjusted $R^2$ | 0.5551           | 0.6342            |
| RSS          | 847.93            | 619.91            |
| AIC          | 290.82            | 282.28            |

Four indicators ($R^2$, adjusted $R^2$, RSS, and AIC) show that the Adaptive Bisquare is a better weighting function for GWR modeling (Table 4). An optimum bandwidth of 21 means there are 21 closest neighboring sub-district that have a significant effect on the estimated value generated by GWR’s equation in a sub-district.

The goodness of fit test result shows that the GWR model is more appropriate than the OLS model in explaining the relationship of poverty level with independent variables at 5 percent level of significance (Table 5). This result provides further evidence that the magnitude of the effect of independent variables on the poverty level in the Bangka Belitung is different in each sub-district.

Table 5. Goodness of Fit Test of the GWR model.

| Source          | Sum of square | Degree of Freedom | Mean Square | F-statistics | p-value |
|-----------------|---------------|-------------------|-------------|--------------|---------|
| OLS Residual    | 2160.96       | 37.000            |             |              |         |
| GWR Improvement | 1541.05       | 19.103            | 80.671      |              |         |
| GWR Residual    | 619.91        | 17.897            | 34.637      | 2.3290       | 0.0428* |

* Significant at $\alpha = 5\%$

The spatial variations test result of the estimated local parameters can be seen in Table 6. The value of the difference of criterion on all parameters of regression worth negative, this means that the magnitude of the effect of all independent variables used on the poverty level in Bangka Belitung varies spatially (different by region). Therefore, policies related to these variables are adjusted to the needs of each sub-district.

Table 6. Variation test of local parameter estimates.

| Variables                                      | Difference of Criterion |
|------------------------------------------------|-------------------------|
| Intercept                                      | -380.8267               |
| Labor force with less than senior high school  | -9.4902                 |
| education (%)                                 |                         |
| Labor force experiencing malnutrition (%)     | -1.9111                 |
| The unemployed labor force (%)                | -1.8948                 |
| Agricultural sector workforce (%)             | -186.5140               |
| Labor productivity (rupiah)                   | -93.0504                |

Table 7 provides a summary of the parameter estimates of the GWR model. The local estimated parameters in the GWR model to all independent variables varied in each sub-district, in contrast to global estimated parameters (OLS model), which only has a single constant value for each variable. On the agriculture sector workforce variable, for example, the magnitude of the effect of this variable on the poverty level varies from -0.0284 to 0.9665.
Table 7. Estimated parameters of the GWR model.

| Variables                                      | $\hat{\beta}$ local | $\hat{\beta}$ maximum local | Sign $\hat{\beta}$ local = $\hat{\beta}$ global (%) | Amount of $\hat{\beta}$ Significant* |
|-----------------------------------------------|----------------------|-------------------------------|------------------------------------------------------|--------------------------------------|
| Intercept                                     | -18.1971             | 85.4962                       | 60.47                                               | 3                                    |
| Labor force with less than senior high school | -1.1207              | 0.1702                        | 55.81                                               | 2                                    |
| education (%)                                 |                      |                               |                                                     |                                      |
| Labor force experiencing malnutrition (%)     | -1.7106              | 0.6388                        | 79.07                                               | 0                                    |
| The unemployed labor force (%)                | -0.6408              | 3.2529                        | 62.79                                               | 2                                    |
| Agricultural sector workforce (%)             | -0.0284              | 0.9665                        | 97.67                                               | 21                                   |
| Labor productivity (rupiah)                   | -0.0043              | 0.0008                        | 55.81                                               | 0                                    |

* Significant at $\alpha = 5$

3.3. Analysis of GWR Model Parameter Estimation Results

Labor force with less than senior high school education has a varied effect from negative to positive on the poverty level in Bangka Belitung. Based on Figure 4, it can be seen that the positive effect of education level on poverty level occurs in the sub-districts in Pangkalpinang City and surrounding areas, while the negative effect of education level on poverty level occurs in the sub-districts in South Bangka Regency and sub-districts in Belitung Island.

Differences directions of these effects can occur due to the ability differences of the workforce [6]. Workers with the same level of education may have different income levels because they have different levels of ability. Workers that educated less than senior high school in the regions which have negative effect relatively have a higher abilities, seen from their average labor productivity (11.071 thousand rupiah) and average income (1.924 million rupiah) is higher than average labor productivity (10.748 thousand rupiah) and average income (1.851 million rupiah) of workers that educated less than senior high school in the regions which have positive effect. It is supported by the fact that the majority (50.91 percent) of the workforce that educated less than senior high school in Kepulauan Bangka Belitung Province work in the agricultural sector which requires more ability and experience in the field than a high level of education. Therefore, increasing the level of education of the labor force needs to be done by following the development of the agricultural sector from traditional agriculture to modern agriculture in order to increase labor productivity and income in Kepulauan Bangka Belitung Province.

Figure 5 shows the distribution of the estimated local parameters of the labor force experiencing malnutrition variable. The positive effect occurs in the sub-districts in West Bangka and South Bangka Regency, while negative effect occurs in the sub-districts in Pangkalpinang City, Bangka and Central Bangka Regency, and Belitung Island. The negative effect can be caused by the small number of workers who experience malnutrition in these regions which is poor. Improving the health condition of the labor force needs to be done because with better health, higher labor productivity and income will be more easily achieved [12].
Figure 5. $\hat{\beta}(u_i, v_i)$ of labor force experiencing malnutrition variable.

Figure 6 shows the distribution of the estimated local parameters of the unemployed labor force variable. Positive effect occurs in the sub-districts in West Bangka, Bangka, South Bangka Regency, and in the Belitung Island, while the negative effect occurs in the sub-districts in Pangkalpinang City and surrounding areas. The negative effect can be caused due to a lot of unemployed (86.45 percent) were living in non-poor households in these regions.

Figure 6. $\hat{\beta}(u_i, v_i)$ of the unemployed labor force variable.

Figure 7 shows the distribution of the estimated local parameters of the agricultural sector workforce variable. Almost all estimated local parameters value (97.67 percent) have effect that is consistent with the estimated value of the global parameter, which has a positive local parameter estimation value in each sub-district so that it can be said, in general, the agricultural sector workforce has a positive effect on the poverty level in the Kepulauan Bangka Belitung Province. The positive effect is due to the large number of agricultural sector workforce who is poor (21.06 percent), in contrast to number of non-agricultural sector workforce who is poor only 8.53 percent. This is supported by the fact that the average productivity (10.949 thousand rupiah) and average income (1.747 million rupiahs) of agricultural sector workers are lower than the average productivity (12.660 thousand rupiahs) and average income (2.399 million rupiahs) of non-agricultural sector workers.

Figure 7. $\hat{\beta}(u_i, v_i)$ of agricultural sector workforce variable.

Figure 8 shows the distribution of the estimated local parameters of the labor productivity variable. Positive effect occurs in the sub-districts in West Bangka and Bangka Regency, and Pangkalpinang
City, while the negative effect occurs in the sub-districts in Bangka Tengah and South Bangka Regency, and in the Belitung Island.

Differences directions of these effects can occur due to differences in the number of working hours. The high labor productivity in regions that have a positive effect is more due to the lower number of working hours.

3.4. Comparison of GWR and Global Regression Model

The ability of the GWR model on modeling the poverty level in the Bangka Belitung can be evaluated by comparing the model diagnostics through several indicators such as the value of $R^2$ and AIC to the OLS model. In addition, it can be evaluated by comparing the residual of the model [10] and the statistical value of Moran’s I from the residuals generated by the model [7]. Table 8 summarizes these indicators.

| Indicators                        | OLS  | GWR  |
|-----------------------------------|------|------|
| Coefficient of Determination ($R^2$) | 0.4870 | 0.8528 |
| Akaike Information Criterion (AIC) | 302.65 | 282.28 |
| Residual Sum of Square (RSS)      | 2160.96 | 619.91 |
| Moran’s I                         | 0.2065 | -0.1038 |

Based on Table 8, it can be concluded that by taking into account the spatial location (spatial non-stationarity) in the modeling process, GWR is able to produce a better model of poverty level than OLS model. This result is supported visually through a scatter plot between observed value and predicted value of the poverty level (Figure 9), which shows a tighter cluster where predicted values are closer to the regression line for GWR model compared to the OLS model [14].
poverty level by (a) OLS; (b) GWR.

Another way to evaluate model performance is to compare the model residual [10]. GWR model had a lower RSS than OLS model. Lower residuals show that the model GWR can predict the poverty level in Bangka Belitung with more accurate than the OLS model.

4. Conclusion
The distribution pattern of the poverty level in Kepulauan Bangka Belitung province shows that there are differences in poverty level among sub-districts in Bangka Island and Belitung Island. Most of the sub-districts on Bangka Island have relatively higher poverty level compared to sub-districts on Belitung Island. Simpang Teritip in West Bangka, Bakam in Bangka Regency, and Tukak Sadai in South Bangka Regency are sub-districts with the highest poverty level, while the sub-districts in the Belitung Island, both in Belitung Regency and East Belitung Regency, as well as the sub-districts in Pangkalpinang City, are sub-districts with relatively low poverty level.

The GWR model showed that the magnitude of the effect of education levels, health conditions, unemployment rates, the agricultural sector, and labor productivity to poverty level varies between sub-districts in Kepulauan Bangka Belitung Province. Therefore, the handling of the poverty problem should pay attention to differences in characteristics between regions (sub-districts). The result of the goodness of fit test showed that the GWR model is better than the OLS model in modeling the poverty level in Kepulauan Bangka Belitung Province.

References
[1] Badan Pusat Statistik 2018 *Jumlah Penduduk Miskin Menurut Provinsi 2007-2018* (Accessed on March 26, 2018 through [https://www.bps.go.id/dynamictable/2016/01/18/20119/jumlah-penduduk-miskin-menurut-provinsi-2007-2018.html](https://www.bps.go.id/dynamictable/2016/01/18/20119/jumlah-penduduk-miskin-menurut-provinsi-2007-2018.html))

[2] Badan Pusat Statistik 2018 *Persentase Penduduk Miskin Menurut Kabupaten/Kota 2002-2018* (Accessed on March 26, 2018 through [https://babel.bps.go.id/dynamictable/2017/06/05/358/persentase-penduduk-miskin-menurut-kabupaten-kota-2002-2018.html](https://babel.bps.go.id/dynamictable/2017/06/05/358/persentase-penduduk-miskin-menurut-kabupaten-kota-2002-2018.html))

[3] Badan Pusat Statistik 2018 *Persentase Penduduk Miskin Menurut Provinsi 2007-2018* (Accessed on March 26, 2018 through [https://www.bps.go.id/dynamictable/2016/08/18/1219/persentase-penduduk-miskin-menurut-provinsi-2007---2018.html](https://www.bps.go.id/dynamictable/2016/08/18/1219/persentase-penduduk-miskin-menurut-provinsi-2007---2018.html))

[4] Benson T, Chamberlin J and Rhinehart I 2005 An investigation of the spatial determinants of the local prevalence of poverty in rural malawi *Food Policy* 30 532-50

[5] Bigman D and Fofack H 2000 Geographical targeting for poverty alleviation: an introduction to the special issue *The World Bank Economic Review* 14 129-45

[6] Borjas G J 2013 *Labour Economics 6th Edition* (New York: McGraw-Hill/Irwin)

[7] Fotheringham A S, Brunsdon C and Charlton M E 2002 *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships* (Chicester: John Wiley & Sons)

[8] Kam S P, Hossain M, Bose M L and Villano L S 2005 Spatial patterns of rural poverty and their relationship with welfare-influencing factors in bangladesh *Food Policy* 30 551-67

[9] Scoones I 1998 Sustainable rural livelihoods: a framework for analysis *IDS Working Paper* 72

[10] Shrestha P M 2006 *Comparison of Ordinary Least Square Regression, Spatial Autoregression, and Geographically Weighted Regression For Modeling Forest Structural Attributes Using A Geographical Information System (GIS)/Remote Sensing (RS) Approach Thesis* (Canada: University of Calgary)

[11] SMERU 2008 *The Poverty Map of Indonesia: Genesis and Significance* (Jakarta: The SMERU Research Institute)

[12] Todaro M P and Smith S C 2004 *Pembangunan Ekonomi di Dunia Ketiga Edisi kedelapan* (Jakarta: Erlangga)
[13] Wahyuni R N T and Damayanti A 2014 Poverty-causing factors in papua province: spatial heterogeneity analysis J. Ekonomi dan Pembangunan Indonesia 14 128-44

[14] Wang Q, Ni J and Tenhunen J 2005 Application of a geographically weighted regression analysis to estimate net primary production of chinese forest ecosystems Global Ecology and Biogeography 14 379-93