A Space-Time Scan Statistic for Detecting Poverty Hotspots in Java Island

Mochamad Fahim, Erfiani, Bagus Sartono

Abstract— Poverty rate reduction is one of eight goals in Millenium Development Goals (MDGs). Most of Indonesian population lived in Java Island, it reached 56.82 percents in 2015. It is also followed by high poverty rate, where poverty rate of three provinces (Jawa Tengah, DIY, and Jawa Timur) is greater than Indonesia. This condition made Java Island as a great contributor to Indonesia's poverty rate. Therefore, it is needed to know when and where poverty hotspots were detected. This research focused on poverty hotspot detection in Java Island based on data from BPS-Central Bureau of Statistic for the situation of 2011-2015 using space-time scan statistic. As a result, we managed to identify eight hotspots that were statistically significant. Among those eight hotspots, seven were detected in the period of 2011-2013 while the other one hotspot was in 2014-2015. According to the target that had MDGs set in 2015, it meant that the effort of poverty rate reduction in Java Island were quite succesfull in the end of MDGs program, it was shown from only one hotspot detected, which only one member namely Malang Regency.

Index Terms— hotspot, poverty, space-time scan statistic

1 Introduction

Poverty is one of undamental problems concerned by government in many countries. Thus, nation leaders and the representatives of 189 countries as the members of United Nation have approved Millennium Development Goals (MDGs) whose main goal is to overcome poverty and hunger [1].

The attempt to overcome poverty which has been done by government encounters some problems. According to Harmadi (2015), one of the problems in resolving poverty in Indonesia is due to unequal population distribution and also geographical problem. Geographical problem here is related to territorial location [2].

Based on data released by Central Bureau of Statistics, more than half of Indonesian population resides in Java Island. The percentage reaches 56.82 percent in 2015 [3]. Consistent with the number of population in Java Island, the number of poor people in Java is also relatively big compared to other province. The concern is high poverty rate in Java. Even Central Java Province, Yogyakarta Special Region, and East Java have higher poverty rate than it nationally. Poverty rate in Central Jave Province, Yogyakarta Special Region, and East Java is 13.32 percent, 13.16 percent, and 12.28 percent, respectively. Meanwhile, national poverty rate is counted 11.13 percent [3].

The high poverty rate in Java Island accompanied by great number of population will significantly affect to national poverty rate. Therefore, it needs an attempt to know when and where poverty hotspots occur in Java Island.

- *Mochamad Fahim is currently pursuing masters degree program in Department of Statistics in Bogor Agricultural University, Indonesia, PH +62-81242913264. E-mail: fahim.mochamad@gmail.com*
- *Erfiani is a Lecturer, Department of Statistics, Bogor Agricultural University, Bogor, Indonesia. E-mail: erfiani_ipb@yahoo.com*
- *Bagus Sartono is a Lecturer, Department of Statistics, Bogor Agricultural University, Bogor, Indonesia. E-mail: bagusco@gmail.com*

Millennium Development Goals (MDGs) program ended in 2015. Nevertheless, the target to reduce percentage of poor people has not been reached. Recorded in September 2015, the percentage of poor people in Indonesia still showed 11,13 percent [3]. Meanwhile, the target to reach in MDGs in 2015 was 7,55 percent. Therefore, an analysis of poverty in several consecutive periods is need until the end of MDGs period in 2015.

According to Rahmawati & Djuraidah [4], poverty problem is highly influenced by space and neighboring. The influence of location in poverty problem enables region with high poverty rate to affect surrounding regions. This condition allows groups of region with similar poverty rate to form.

Previous studies related to how poverty hotspots established in Indonesia have been widely conducted. Nurcahyani and Purhadi conducted a study to detect poverty hotspot using Flexibility Shaped Spatial Scan Statistic method [5]. Hasibuan studied poverty spatial pattern in Indramayu using Flexibility Shaped Spatial Scan Statistics to detect poverty hotspots [6]. Kusumaningrum investigated poverty hotspots, unemployment, and food insecurity in Java Island using Spatial Scan Statistic and Upper Level Set (ULS) Scan Statistics methods [7]. Siswantining conducted a study to detect poverty hotspots using Spatial Scan Statistic method in several villages in Jember Regency. The study was conducted by estimating small areas to village level on calories consumption and household expenses variables [8]. Septiani conducted a study to detect poverty hotspots in village iin Bogor Regency and Bogor Municipality using Flexibility Shaped Spatial Scan Statistics method [9].

The previous studies mentioned, employed Spatial Scan Statistic, Flexibility Shaped Spatial Scan Statistic, and Upper Level Set (ULS) Scan Statistic methods to detect poverty hotspots. The three methods are the development of Scan Statistic method which has considered spatial element in detecting poverty hotspots. However, those methods can only be implemented in data of 1 (one) point of time or cross section.

A method to detect hotspots of a subject using time series by involving the location influence (spatial) has been developed. Space-Time Scan Statistic has statistical method which has an ability to detect hotspots of a subject in spatial data and time series [10].

Referring to previous description, the problem discussed in this study is how to detect poverty hotspots in Java Island by including spatial and time elements. The goal of the study is to detect poverty hotspots in Java Island by including spatial and time influence. Information about poverty hotspots is expected to be reference for government in the attempt to overcome poverty.

2 Research Methode

2.1 Data

Data employed in this study was secondary data which was the number of poor people, population number, and latitude and longitude coordinates. The data scope of this study was regencies/municipalities in Java Island in 2011-2015. The type of spatial data used in this study was point spatial data. The point employed as the main location referred to the concept used by Central Bureau of Statistics of the Republic of Indonesia in recording *Potensi Desa*, which was coordinate of regional office, here is coordinate of Regent/Major Office.

2.2 Data Analysis Procedures

Analysis method used was descriptive analysis and space-time scan statistic. Descriptive analysis was used to provide a general overview about variables being used. Meanwhile, space-time scan statistic was used to detect poverty hotspots.

Analysis procedures in this study were as follows:

1. Data exploration
   Exploring data to observe descriptively the development of poverty rate in regency/municipality in Java Island during 2011-2015.
2. Detecting global spatial correlation
   Detecting global spatial correlation was conducted using spatial autocorrelation. Statistical measure employed was Moran's Global I Index [12]. This value is correlation coefficient in limits -1 and 1. Hypothesis testing was employed to ensure this index, where hyphotesis null was there is no spatial correlation.
3. Implementing zone formation algorithm.
   Zone is a potential area to form hotspot (hotspot candidate). Kulldroff *et al.* [11] developed spatial scan statistic [10] by including data time element, called space-time scan statistic. Spatial scan statistic enables to detect two-dimension hotspots in point process. This method employs circular window to identify a case in the zone to detect high potential location of hotspot compared to other region.
   The steps of zone formation algorithm are as follows:
   a) Choosing 1 regency/municipality randomly, where the selected location of regency is represented by latitude and longitude coordinates.
   b) Estimating Euclidean distance of selected regency/municipality to other regency/municipality.
   c) Arranging the distances from the nearest to the farthest one, so that an array is established.
   d) Repeating step a-c for each regency/municipality, provided that the previously selected regency/municipality is not re-selected.
   e) Choosing 1 regency/municipality randomly, where the location of selected regency is represented by latitude and longitude coordinates.
   f) Making scanning window in form of a circle whose central point is latitude-longitude coordinates and magnifies circular radius based on array order in step c.
   g) Repeating steps e and f for each regency/municipality, provided that the previously selected regency/municipality is not re-selected.
   h) Estimating relative risk of each established zone. After that, discarding scanning window which has relative risk value less than1. Until this step, hotspot candidate is obtained.
   i) Estimating log likelihood ratio value of hotspot candidate obtained in step h, with formulation as in [11]. Then, likelihood ratio value used to established hypothesis with null hypothesis is proportion of case inside cylinder is the same with it outside cylinder.
4. Estimating p-value of each hotspot candidate established in step 3 using Monte Carlo Simulation, where p-value shows whether a null hypothesis is accepted or rejected. Monte carlo simulation was done with 999 replication.
5. Presenting hotspots using thematic map.
6. Interpret detected hotspots.

3 Result and Discussion

3.1 Data Exploration

Data exploration was done by presenting the distribution of poverty rate in form of thematic map. Thematic map was made for each research year. Regency/municipality was categorized into 2, regency/municipality with poverty rate greater than national poverty rate and regency/municipality with poverty rate lower than national poverty rate. Regency/municipality with poverty rate greater than national poverty rate is presented in Figure 1 in grey color.

Figure 1 appears that the distribution pattern of regencies/municipalities with poverty rate greater than national poverty rate tends to be the same from 2011 to 2015. The change of distribution pattern over years was not really significant.

In 2011, there were 57 regencies/municipalities with poverty rate greater than national poverty rate. In 2012, it increased to 62 regencies/municipalities with addition of Kepulauan Seribu Regency (Jakarta), Tasikmalaya Regency (West Java), Temanggung Regency (Central Java), Surakarta Municipality (Central Java), and Ponorogo Regency (East Java).

Composition change from 2012 to 2013 was not really significant. Pasuruan Regency and Magetan Regency whose poverty rate in 2012 was still low than national poverty rate, in 2013 it increased more than national poverty rate. Meanwhile, Probolinggo Municipality at the end of 2012 had poverty rate

greater than national poverty rate became lower than it in 2013.

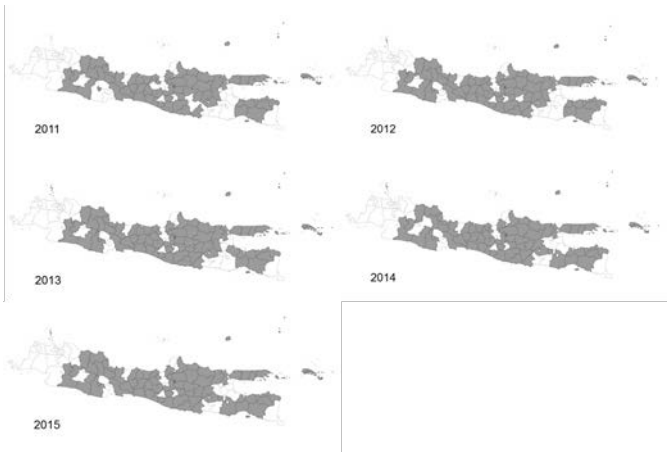

Fig. 1. Map of regencies/municipalities in Java Island with poverty rate greater than national poverty rate in 2011-2015

Meanwhile, composition change from 2013 to 2014 and from 2014 to 2015 was not really significant. Composition change from 2013 to 2014 was that Sumedang Regency, Surakarta Municipality, Pasuruan Regency, and Jombang Regency in 2014 had poverty rate lower than national poverty rate. Meanwhile, Malang Regency in 2014 had poverty rate greater than national poverty rate. In 2015, there was an addition of 1 regency from the condition in 2014, which was Sumedang Regency whose lower poverty rate in 2014, it turned out to be greater in 2015 than national poverty rate.

3.2 Spatial Autocorrelation Detection

Spatial effect testing needs to be done to make sure the presence or the absence of spatial dependency of poverty rate between regencies/municipalities in Java Island. Based on the estimation of spatial influence in poverty rate during 2011 to 2015, it was obtained Global Moran's I Index as presented in Table I respectively.

TABLE 1
GLOBAL MORAN'S I INDEX AUTOCORRELATION

Year	Global Moran's I		
	Observed	Expected	p-value
2011	0.2283319	-0.0085470	0.00*
2012	0.2270212	-0.0085470	0.00*
2013	0.2151253	-0.0085470	0.00*
2014	0.1967752	-0.0085470	0.00*
2015	0.1879452	-0.0085470	0.00*

* significant at 0.05 significant level

There was no scale of Global Moran's I Index during the time which had null value, this shows that there was spatial autocorrelation of poverty rate in regencies/municipalities in Java Island, which is poverty rate in adjacent regencies/municipalities tend to be similar. Positive correlation is also strengthened with the result of hypothesis testing where for each year from 2011 to 2015 had p-value smaller than 0.05 so that null hypothesis was rejected. The rejection of null hypothesis represented that there was positive spatial autocorrelation of poverty rate in regencies/municipalities in Java Islands.

3.3 Hotspots

Hotspots are established through 2 steps. The first is implementing zone formation algorithm, as potential are to form hotspots (hotspot candidates). The second one is determining poverty hotspots from the established hotspots.

In zone formation step, it was done an estimation of relative risk value of each established zone. Based on Table 2, it appeared that the nine established zones have relative risk value greater than 1. This shows that the nine established zones have high risk to be hotspot candidate. It can be seen that the highest relative risk is in zone 2 and the lowest one is in zones 7 and 8.

In zone establishment step was also done an estimation of log likelihood ratio value. This value will later be the reference in determining hotspots from the existing hotspot candidates. Log likelihood ratio value is sorted from the biggest to the smallest.

TABLE 2
POVERTY HOTSPOTS

i-th hotspot candidates	Time frame	Relative risk	Log likelihood ratio	p-value
1	2011-2013	1.72	803454.09	0.001*
2	2011-2013	1.99	762173.30	0.001*
3	2011-2013	1.41	300105.67	0.001*
4	2011-2013	1.24	108778.78	0.001*
5	2011-2013	1.18	19740.78	0.001*
6	2011-2013	1.04	971.34	0.001*
7	2011-2013	1.01	61.22	0.001*
8	2014-2015	1.01	45.03	0.001*
9	2011-2013	1.03	4.44	0.627

* significant at 0.05 significant level

The determination of hotspot was conducted to make sure statistically which hotspot candidates to be the hotspots. The determination of hotspots was done using *Monte Carlo* simulation with replication of 999 times. *Monte Carlo* simulation was done to each hotspot candidate to obtain p-value from each hotspot candidate. That p-value was then used as the reference for decision making of null hypothesis testing. Null hypothesis established was proportion of hotspots inside scanning cylinder was the same with it outside the scanning cylinder, or no hotspot establishment.

Table 2 also presented the result of *Monte Carlo* simulation to 9 hotspot candidates which had been previously formed. It appeared that hotspot candidate 1 to 8 had p-value of 0,001, where p-value is smaller than significant level of 0.05. Hence, it was determined that null hypothesis was rejected and it was

concluded that candidates 1 to 8 were poverty hotspots. Meanwhile, the 9th hotspot candidate had p-value of 0.627 where p-value greater than significant level of 0.05, so that null hypothesis was not rejected and it was concluded that the 9th hotspot candidate was not a hotspot.

To make it easy to see the position of poverty hotspots, visualization of poverty hotspot establishment was presented using satellite map as seen in Figure 2.

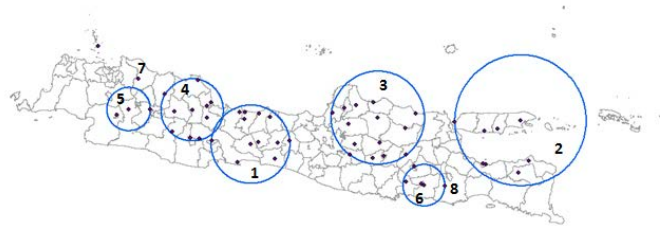

Fig. 2. Visualization of poverty hotspots in Java Island.

In Figure 2, it was seen circles which were the representatives of scanning cylinder in space time scan statistic. It was seen 8 circles with different sizes/radius. The eight circles were scanning windows of scanning process result. Areas inside scanning window were poverty hotspots in Java Island. It was seen that in Figure 2, it was difficult to know how far the radius of each hotspot and which regencies/municipalities included. Therefore, information about radius and regencies/municipalities included in hotspot was presented in Table 3.

TABLE 3
POVERTY HOTSPOTS BY ITS RADIUS AND COVERAGE

i-th hotspots	Radius (km)	Regencies/municipalities
1	74.36	Banyumas, Purbalingga, Cilacap, Tegal, Banjarnegara, Kebumen, Pemalang, Tegal*, Pekalongan, Brebes, Banjar*, Wonosobo.
2	124.29	Sumenep, Pamekasan, Sampang, Situbondo, Bondowoso, Probolinggo, Probolinggo*, Bangkalan.
3	88.42	Blora, Rembang, Pati, Ngawi, Bojonegoro, Grobogan, Kudus, Sragen, Tuban, Kota Madiun, Madiun, Magetan, Demak, Karanganyar , Nganjuk.
4	59.14	Majalengka, Cirebon, Kuningan, Sumedang, Cirebon*, Tasikmalaya*, Garut, Ciamis, Indramayu, Subang.
5	41.15	Cianjur, Sukabumi*, Bandung Barat.
6	39.45	Blitar, Blitar*, Tulungagung, Kediri.
7	0.00	Karawang
8	0.00	Malang

* Municipality

3.3 Characteristics of Poverty Hotspots

Presidential Regulation Number 5 of 2010 stated that poverty is one of 3 developmental strategies stated in triple track strategy which is pro growth, pro job, and pro poor. The analysis of pro growth strategy in poverty hotspots was done by observing dominant economical sector which contributed to the economy of each hotspot and the source of economical growth to be connected to pro job strategy by observing the characteristics of poor population labors in each hotspot.

The first and the seventh hotspots had the same characterwastic which was their regional economy was dominated by manufactur sector with source of economical growth derived from manufactur sector and its domination of job field for poor population was not in agricultural sector. The fourth and the fifth hotspots also had the same characterwastic which was the absence of dominant economical sector with source of economical sector derives from wholesale & retail trade sector and domination of job field for poor population was not in agricultural sector. The same characterwastic was also seen in the third and the eight hotspots where the economy was dominated by manufactur sector with source of economical growth was obtained from agricultural sector and domination of job field for poor population was in agricultural sector.

An interesting fact was seen in the second and the sixth hotspots. In the second hotspot, the economy and source of economical growth derives from agricultural sector, but almost half of poor population work in agricultural sector, the percentage reaches 45.66 percent. It meant that agriculture as the mainstay of the economic sector has not been able to reliably for poverty reduction.

Meanwhile, in the sixth hotspot, the economy was dominated by agricultural sector with domination reaches 26.39 percent, much more than wholesale & retail trade sector which occupy the second rank with contribution of 19.20 percent. However, the source of economical growth comes from wholesale & retail trade sector. The domination of job field for poor population in the sixth hotspot does not come from agricultural sector with domination reaches 35.11 percent, whereas in agricultural sector the percentage only reaches 26.59 percent. It described to us that agriculture as the mainstay of the economic sector has not been fullest role in poverty reduction.

4 Conclusion

This study concludes that during 2011-2015 in Java Island had been established 8 poverty hotspots. As many as 7 poverty hotspots were established during 2011 to 2013 and 1 hotspot was formed during 2014 to 2015. The attempt to overcome poverty in Java Island was a success regarding MDGs program in 2015 resulted in only 1 poverty hotspot with only 1 member which is Malang Regency.

The establishment of poverty hotspots in regencies/municipalities in Java Island can be used as evaluation for government in relevant province/regency/municipality to

overcome poverty fit to the characteristics of each poverty hotspot. Recommendation for future research is that to conduct estimating missing data. This is because in this study Pangandaran Regency is divided from Ciamis Regency so that the poverty data has just been available for 2015.

ACKNOWLEDGMENT

The authors wish to thank BPS-Central Bureau of Statistic for providing data in this research. This work was supported in part by a grant from BPS.

REFERENCES

[1] United Nation, "*Millenium Development Goals*", http://www.undp.org/content/undp/en/home/sdgoverview/mdg_goals.html. 2000.

[2] SHB. Harmadi, "Response for Technocratic Plans of RPJMN 2015-2019 Field of Social Culture, Infrastructure, and Spatial Area, with an Emphasis on Substance of Poverty" unpublished.

[3] Central Bureau of Statistic, *Statistik Indonesia 2016*. Jakarta, pp. 186, 2016.

[4] A. Djuraedah, R. Rahmawati, "*Analysis of Geographically Weighted Reggression (GWR) with Gaussian Kernel Weighted for Poverty Data*," *Proc. National Conference on Statistics*, pp. 322-326, 2011.

[5] H. Nurcahayani, Purhadi, "*Modeling Spatial Poverty with Mixed Geographically Weighted Regression and Poisson Flexibility Shaped Spatial Scan Statistic*," *Proc. National Conference on Mathematics and Mathematics Education*, pp. 97-106, 2013.

[6] LS. Hasibuan, "Spatial Patterns of Poverty in Indramayu,"Master Thesis, Dept. of Statistics, Bogor Agricultural Univ., Bogor, 2007.

[7] D. Kusumaningrum, "Hotspot Analysis on Poverty, Unemployment and Food Security in Java, Indonesia", Master Thesis, Dept. of Statistics, Bogor Agricultural Univ., Bogor, 2010.

[8] T. Siswantining, "Geoinformatic on Small Area and Its Application for Detecting Poverty Hotspots in Jember, East Java," PhD Dissertation, Dept. of Statistics, Bogor Agricultural Univ., Bogor, 2013.

[9] D. Septiani, "A Flexibility Shape Spatial Scan Statistics for Detecting Poverty Hotspots," Master Thesis, Dept. of Statistics, Bogor Agricultural Univ., Bogor, 2006.

[10] M. Kulldroff, "A Spatial Scan Statistics," *Communication in Statistics: Theory and Methods*, vol. 26, no. 6, pp. 1481-1496, 1997.

[11] M. Kulldroff, WF. Athas, EJ. Feuer, BA. Miller, CR. Key, "Evaluating Clusters Alarm: A Space-Time Scan Statistics and Brain Cancer in Los Alamos,"*American Journal of Public Health*, vol. 88, pp. 1377-1380, 1998.

[12] J. Lee, DWS Wong, *Statistical Analysis with ARCview GIS*. New York: John Wiley & Son, Inc., 2001.