A FLEXIBLY SHAPED SPATIAL SCAN STATISTIC
FOR DETECTING POVERTY HOTSPOTS

DINA SEPTIANI

DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
BOGOR AGRICULTURAL UNIVERSITY
BOGOR
2006
ABSTRACT

DNA SEPTIANI. A Flexibly Shaped Spatial Scan Statistic for Detecting Poverty Hotspots. Under the direction of ASEP SAEFUDDIN and BAGUS SARTONO.

Poverty always becomes the crucial and complex problem in developing countries. Government conducts anti-poverty policies for reducing poverty. If poverty-reduction is expected to succeed, there must be well-founded approach and widespread support. Any approach and active participation of numerous parties will be no use if it is not supported by programs that are well-planned, executed in stages, and sustainable. Poverty hotspot detection is very useful to provide early stage to conduct planning strategy for poverty reduction.

In this research, a flexibly shaped spatial scan statistic, proposed by Tango and Takahashi (2005) is an approach for detecting geographical hotspots that allow flexibility in shape. This scan statistic imposes an arbitrary shaped window with maximum length cluster \( K \). The test statistic is based on likelihood ratio test under Poisson assumption and evaluated using Monte Carlo hypothesis testing.

The object of this research is Bogor, which is divided into 493 villages. As the results, 34 significant hotspots are detected among 91 hotspot candidates at significance level \( \alpha = 0.05 \) and \( K = 15 \). These significant poverty hotspots consist of 199 villages in Bogor. There is no geographical overlap among all reported hotspots. The MLC that gains the highest log likelihood ratio, in poverty hotspots mapping, is concentrated in the western of Bogor.

Keyword: Hotspot, Poverty, Flexibly shaped spatial scan statistic.
A FLEXIBLY SHAPED SPATIAL SCAN STATISTIC
FOR DETECTING POVERTY HOTSPOTS

By:
DINA SEPTIANI
G14102041

Thesis

For the partial fulfillment for the degree of Bachelor of Sciences
Faculty of Mathematics and Natural Sciences
Bogor Agricultural University

DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
BOGOR AGRICULTURAL UNIVERSITY
BOGOR
2006
Title: A FLEXIBLY SHAPE SPATIAL SCAN STATISTIC FOR DETECTING POVERTY HOTSPOTS
Name: Dina Septiani
NRP: G14102041

Approved by:

Advisor I, Advisor II,

Dr. Ir. Asep Saeufuddin, M.Sc
NIP. 130938261

Bagus Sartono, M.Si
NIP. 132311923

Acknowledged by:
Dean of Faculty of Mathematics and Natural Sciences
Bogor Agricultural University

Prof. Dr. Ir. Yonny Koesmaryono, MS.
NIP. 131473999

Passed examination date:
FOREWORD

Alhamdulillahi rabbi ‘alamin, praise and grateful to Allah SWT for all blessing and favour that made this thesis with the title “A Flexibly Shaped Spatial Scan Statistic for Detecting Poverty Hotspots” completed. The idea for making this thesis came from recommendation of my advisor, Dr. Ir. Asep Saefuddin, M.Sc.

This thesis could not been written without the support from many people, therefore, in this opportunity, the author would like to say thank you to:

1. Dr. Ir. Asep Saefuddin, M.Sc. and Bagus Sartono, S.Si, M.Si for the valuable advice, theoretical guidance, supports, and critics during this research.
2. All lecturers at Department of Statistics for sharing their knowledge and also staffs of Department of Statistics for helping my study: Bu dede, Bu Markonah, Bu Sulis, Bang Sudin, Pak Ian, Mang Herman, Bang Dur.
3. My beloved family for all prayer, supports, love, and advices.
4. Indra for his loves, caring, and supports.
5. For all STK’39: Rani, Rizka, and Ekek (thanks for being my panellists and your supports). Karin, Ochi, Neng Fitri (thanks for all your support), Cey (thanks for your help). Echi (thanks for the format of thesis), Rian ‘Jack’ my partner at PT. Fortune Indonesia, Uput and others who can not be mentioned here, I am very thankful for the friendship and the times that we had together.
6. All STK’38, STK’40, and STK’41.
7. Mba Ita, Mba Ila, Mba Marta, Nono, Hilme, Uci, Mire, Tria, Mba Gadi, Mba Wulan and all alumni Asrama Putri IPB Baranangsiang.
8. Mba Nur and Sari for support.

The author realized that perfection only belongs to Allah SWT and that there is still a lot of work in this thesis. Hopefully, this thesis can be useful for anybody who reads it.

Bogor, August 2006

Author
CONTENTS

LIST OF TABLES .......................................................................................................................... vi
LIST OF FIGURES .......................................................................................................................... vi
LIST OF FIGURES .......................................................................................................................... vi
INTRODUCTION .................................................................................................................................
  Background ................................................................................................................................... 1
  Objective ...................................................................................................................................... 1

LITERATURE REVIEW ......................................................................................................................
  Poverty ......................................................................................................................................... 1
  Hotspot ........................................................................................................................................ 2
  Flexibly Spatial Scan Statistic ....................................................................................................... 2
  Likelihood Ratio Test .................................................................................................................. 3
  Relative Risk ............................................................................................................................... 3
  Monte Carlo Hypothesis Testing .................................................................................................. 3

MATERIAL AND METHODS ...........................................................................................................
  Source of Data ............................................................................................................................. 4
  Methods ..................................................................................................................................... 4

RESULTS AND DISCUSSION ...........................................................................................................
  Preliminary Description .............................................................................................................. 4
  Hotspot Detection ...................................................................................................................... 5

CONCLUSION .................................................................................................................................
COMMENDATION ............................................................................................................................
REFERENCES .................................................................................................................................
PENDICES ......................................................................................................................................

Page
vi
vi
vi
1
1
2
2
3
3
4
4
4
5
7
7
7
9
LIST OF TABLES

1. Villages with the 20th highest poverty rate in Bogor .................................................. 4
2. The five most significant poverty hotspots ................................................................. 5
3. Villages that comprised Hotspot 1 .............................................................................. 5
4. Villages that comprised Hotspot 2 .............................................................................. 6
5. Villages that comprised Hotspot 3 .............................................................................. 6
6. Villages that comprised Hotspot 4 .............................................................................. 6
7. Villages that comprised Hotspot 5 .............................................................................. 6

LIST OF FIGURES

Limitation of circular window ......................................................................................... 1
An irregularly shaped window and connected regions .................................................... 2
The disconnected regions ................................................................................................. 3

LIST OF APPENDICES

1. The significant poverty hotspots ................................................................................ 10
2. Scanning result of five most significant poverty hotspots ....................................... 11
3. Mapping of five most significant poverty hotspots in Bogor .................................. 12
4. Mapping of all significant poverty hotspots in Bogor .............................................. 13
Hotspot

Hotspot is defined as something unusual, abnormal, aberration, outbreak, elevated cluster, or critical area (Patil and Taillie, 2004). Hotspots are locations or regions that have consistently high levels of disease and may have characteristics unlike those of surrounding areas (Haran, Molineros, and Patil, 2006).

Hotspot clusters were generated by setting relative risk in some counties to be larger than one (Song and Kulldorff, 2003). A poverty hotspot represents an area characterized by certain local characteristics which could also expand and affect other neighboring areas (Betti, Ballini, and Neri, 2006).

**Flexibly Spatial Scan Statistic**

Flexibly spatial scan statistic is an approach of scan statistic that can detect regular shaped clusters. A flexibly spatial scan statistic imposes an *irregularly shaped window* (zone) on each region by connecting adjacent regions (Tango and Takahashi, 2005). An example of an irregular shaped window is depicted in Figure 2.

![Figure 2. An irregularly shaped window and connected regions.](image)

Consider the situation where an entire study area is divided into $m$ regions. The number of cases (events) in the region $i$ is noted by the random variable $N_i$ with observed value $n_i$, $i = 1, ..., m$. $N_i$ are independent Poisson variables such that

$$N_i \sim \text{Poisson}(\lambda_i), i = 1, ..., m \quad \cdots \cdots \cdots \cdots (1)$$

where $\text{Poisson}(\cdot)$ denotes Poisson distribution with mean $\lambda$ and the $\lambda_i$ are the expected number of cases in the region $i$.

A Poisson model is used when the number of cases is compared to an underlying population at risk derived from the census (Huang, Kulldorff, and Gregorio, 2005). For any given region $i$, the set of irregularly shaped windows with length $k$ consisting of $k$ connected regions including $i$ and let $k$ moves from 1 to the pre-set maximum $K$.

To avoid detecting a cluster of *unlikely peculiar shape*, the connected regions are restricted as the subsets of the set of regions $i$ and $(K - 1)$-nearest neighbors to the region $i$ where $K$ is a pre-specified maximum length of cluster.

Let $z_{a(j)} = j, ..., j_k$ denote the $j$-th window which is a set of $k$ regions connected starting from the region $i$, where $j_k$ is the number of $j$ satisfying $a(j) \subseteq a(k)$ for $k = 1, ..., K$. Then, all the windows (zones) to be scanned are included in the set

$$Z = \{z_{a(j)} | 1 \leq i \leq m, 1 \leq k \leq K, 1 \leq j \leq j_k\}. \quad (2)$$

The window of a scan statistic is often thought of as an interval, area, or volume of fixed size and shape, which then moves across the study area (Kulldorff, 1999).

In flexibly spatial scan statistic, a window refers to a collection of connected regions that is flexible in size and shape.

These are algorithms to find the set $Z$ that is defined as the set of arbitrarily shaped windows (zones) $z$ within a pre-specified maximum length $K$:

1. First, set an $m \times m$ matrix adjacent $A = (a_{ij})$ such as

   $$a_{ij} = \begin{cases} 1 & \text{(region } i \text{ and } j \text{ connected)} \\ 0 & \text{(otherwise)} \end{cases}$$

   and set $Z = \emptyset$ and $i_0 = 0$

2. Let $i_1 \leftarrow i_0 + 1$ and $i_k (k = 1, 2, ..., m)$ be the starting region. Then, create the set consisting of $(K - 1)$-nearest neighbors to the starting region $i_0$ and $i_k$ itself:

   $$W_k = \{i_0, i_1, i_2, ..., i_k \}$$

   where $i_k$ is the $k$-th nearest to $i_0$.

3. Consider all the set $z \subset W_k$, which includes the starting region $i_0$. For any given such set $z$, repeat the following steps 4–7.

4. Divide the set $z$ into two disjoint sets: $z_0 = \{i_0\}$ and $z_1$ which contains the other regions of $z$.

5. Make two new sets $z_0$ and $z_1$. $z_0$ consists of the regions of $z_1$ that are connected to some regions of $z_0$. On the other hand, $z_1$ consists of the regions of $z_1$ that are not connected to any regions of $z_0$. Then,
replace $z_0$ and $z_1$ by $z'_0$ and $z'_1$, respectively.

6. Repeat the step 5 recursively until either $z_0$ or $z_1$ becomes null first.

7. Make a decision as follows: $z$ is said to be "connected" when $z_1$ becomes null first and "disconnected" when $z_0$ becomes null first. If $z$ is "connected", $z$ is added to the set $Z$. If $z$ is "disconnected", $z$ is discarded.

Figure 2 depicts that the collection of shaded regions is connected and, therefore, constitute a window in $Z$. Figure 3 depicts that the collection of shaded regions is disconnected.

Repeat the steps 2–7 until the set $Z$ consisting of arbitrarily shaped windows $z$ whose maximum length is $K$ is finally obtained.

Likelihood Ratio Test

For each region and length of the scanning window, the alternative hypothesis is that there is at least one window $z$ for which the underlying risk is higher (elevated risk) inside the window when compared with outside. In other words, considering the following hypothesis:

$H_0 : E[N(z)] = \xi(z)$, for all $z$, $H_1 : E[N(z)] > \xi(z)$, for some $z$

where $N(\cdot)$ and $\xi(\cdot)$ denote the random number of cases and the null expected number of cases within the specified window, respectively. For each window, it is possible to compute the likelihood to observe the observed number of cases within and outside the window, respectively. Under the Poisson assumption, the test statistic, which was constructed with the likelihood ratio test (Kulldorff, 1997), is given by

$$
\sup_{z \in Z} \left( \frac{n(z)}{\xi(z)} \right)^{n(z)} \left( \frac{n(z)}{\xi(z)} \right)^{n(z)} \left( \frac{n(z)}{\xi(z)} \right) \left( \frac{n(z)}{\xi(z)} \right) \left( \frac{n(z)}{\xi(z)} \right)
$$

where $Z$ indicates all the regions outside the window $z$, and $n(\cdot)$ denotes the observed number of cases within the specified window and $\xi(\cdot)$ is the indicator function.

$\xi(z)$ is equal to 1 when the window has more cases than expected under the null-hypothesis and it constitutes window $z$ as hotspot candidate, and 0 otherwise.

The likelihood function is maximized over all window locations and sizes, and the one with the maximum likelihood constitutes the most likely cluster (MLC).

In addition to the most likely cluster, there are secondary clusters with high values of the likelihood ratio. Secondary clusters are reported when those clusters have no geographical overlap with other reported clusters that have higher likelihood ratio. The statistical significance of a secondary cluster is evaluated irrespectively of any other cluster, by comparing its likelihood ratio with maximum likelihood ratio from the generated random data sets (Jung, Kulldorff, and Klassen, 2005).

Relative Risk

The relative risk, any non-negative number, representing how much more common event (case) is in this location is compared to the baseline. Setting a value of one is equivalent to not doing any adjustment. A value of greater than one is used to adjust for an increased risk (it constitute a window as hotspot candidate) and a value of less than one to adjust for lower risk. A relative risk of zero is used to adjust for missing data for that particular location.

Relative risk is calculated by:

$$
RR = \frac{n(z)}{\xi(z)}
$$

(6)

Monte Carlo Hypothesis Testing

Monte Carlo simulation (Dwass, 1957 in Kulldorff, 1999) is required to find the distribution of the test statistic under the null hypothesis. This simulation is used because the calculation for obtaining the distribution of the scan statistic is quite complex.

The p-value is determined through Monte Carlo hypothesis testing, by comparing the rank of the maximum likelihood from the real data set with the maximum likelihoods from the random data sets.

If $R$ is the rank of the test statistic from the real data set and $M$ is number of simulation among all data sets, then
Monte Carlo hypothesis testing for a scan statistic is a four-step procedure:
1. Calculate the value of test statistic for real data.
2. Create a large number of random data sets generated under null hypothesis for each window (M data sets are generated under \( H_0 \)).
3. Calculate the value of the test statistic for each of the random replications (\( \lambda \) is computed for each simulated data set).
4. Sort the value of the test statistic, from the real and random data sets, and note the rank of the one calculated from the real data set. At the \( \alpha \) significance level, \( H_0 \) is rejected if the rank of the \( \lambda \) obtained from real data is among the \( \alpha(M+1) \) largest \( \lambda \)'s.

**MATERIAL AND METHODS**

**Source of Data**
The data used in this study are secondary data from *Poliensi Desa* (Podes) survey conducted in 2006 by Statistics Indonesia. The target of this research is Bogor consisting of Bogor Regency, which comprises 31 districts, and Bogor Municipality, which comprises 6 districts. Bogor Regency consists of 425 villages while Bogor Municipality consists of 68 villages.

To detect poverty hotspots, this scanning procedure was required:
1. Number of pre-prosperous family or prosperous family level 1 (poor family).
2. Population of family,
3. Central coordinate for each region (in this study, region refers to village), and
4. Matrix of adjacent area (contiguity matrix).

**Methods**
The methods of this research were:
Conduct preliminary data description
Due to mathematical computation limitation, data were transformed by dividing them by 10.
Construct spatial poverty hotspot candidates with maximum length clusters \( \lambda = 15 \), as suggested by Tango and Takahashi (2005).
Calculate the value of test statistic for each hotspot.
Detect significant poverty hotspots by using Monte Carlo-based hypothesis testing with number of simulation \( M = 999 \) replications and significance level \( \alpha = 0.05 \).

6. Define the relative risk of significant hotspots.
7. Present map of poverty hotspots. This map is drawn by using MapInfo Professional 7.5 SCP.

Step 1 and 2 are performed by using Microsoft Excel Professional 2003. Meanwhile, step 3 to 6 are performed by FlexScan Version 1.1.2

**RESULTS AND DISCUSSION**

Preliminary Description
Overall population of family in Bogor is 1,094,480 families and 29.44% of them are categorized as poor family.
Population of family in Bogor Regency is 901,819 families and 32.93% of them are categorized as poor family. Meanwhile, there are 13.06% of 192,661 families in Bogor Municipality are categorized as poor family.
Average of poor family at district level of Bogor Regency is 9,573 families per district with coefficient of variation 42.24%. Average of poor family at district level of Bogor Municipality is 4,193 families per district with coefficient of variation 33.40%.
Average of poor family at village level of Bogor Regency is 700 families per village with coefficient of variation 61.23%. Average of poor family at village level of Bogor Municipality is 370 families per village with coefficient of variation 81.36%.

Based on the previous explanation, there is a wide spatial heterogeneity of poor people at village levels.

Table 1 shows twenty villages that hold the highest poverty rate in Bogor. Singabrja in Tenjo District is the poorest village (100.00%) that all of families at this village are living under poverty. Purwassari and Sukadamai in Dramaga District, respectively, hold the second and third highest poverty rate in Bogor.

<table>
<thead>
<tr>
<th>Area Name</th>
<th>District</th>
<th>Poverty Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singabrja</td>
<td>Tenjo</td>
<td>100.00%</td>
</tr>
<tr>
<td>Purwassari</td>
<td>Dramaga</td>
<td>99.29%</td>
</tr>
<tr>
<td>Sukadamai</td>
<td>Dramaga</td>
<td>94.52%</td>
</tr>
<tr>
<td>Harkatjaya</td>
<td>Sukajaya</td>
<td>90.88%</td>
</tr>
<tr>
<td>Claruten Udik</td>
<td>Cibungbulang</td>
<td>90.58%</td>
</tr>
<tr>
<td>Citaruna</td>
<td>Sukajaya</td>
<td>88.97%</td>
</tr>
<tr>
<td>Area Name</td>
<td>District</td>
<td>Poverty rate</td>
</tr>
<tr>
<td>------------</td>
<td>----------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Cisaruan</td>
<td>Sukajaya</td>
<td>88.79%</td>
</tr>
<tr>
<td>Cisau</td>
<td>Sukajaya</td>
<td>88.51%</td>
</tr>
<tr>
<td>Karasari</td>
<td>Sukajaya</td>
<td>87.06%</td>
</tr>
<tr>
<td>Pasir Madang</td>
<td>Sukajaya</td>
<td>87.06%</td>
</tr>
<tr>
<td>Sukaraja</td>
<td>Sukajaya</td>
<td>86.22%</td>
</tr>
<tr>
<td>Jatibon</td>
<td>Sukajaya</td>
<td>81.14%</td>
</tr>
<tr>
<td>Karasari</td>
<td>Sukajaya</td>
<td>81.02%</td>
</tr>
<tr>
<td>Cibuburung</td>
<td>Sukajaya</td>
<td>77.66%</td>
</tr>
<tr>
<td>Tapai</td>
<td>Sukajaya</td>
<td>77.38%</td>
</tr>
<tr>
<td>Cibuburung</td>
<td>Sukajaya</td>
<td>77.29%</td>
</tr>
<tr>
<td>Cibuburung</td>
<td>Sukajaya</td>
<td>77.20%</td>
</tr>
<tr>
<td>Cibuburung</td>
<td>Sukajaya</td>
<td>77.04%</td>
</tr>
<tr>
<td>Sukamulya</td>
<td>Sukajaya</td>
<td>76.66%</td>
</tr>
<tr>
<td>Sukamulya</td>
<td>Sukajaya</td>
<td>76.28%</td>
</tr>
<tr>
<td>Sukamulya</td>
<td>Sukajaya</td>
<td>76.17%</td>
</tr>
</tbody>
</table>

**Hotspot Detection**

Consider that the expected relative risk is to one. It means that there is the equally likely within any region in Bogor. The estimated probability of poverty within any region in Bogor is 0.294.

This scanning yields 91 hotspot candidates. These hotspot candidates have active risk value larger than one and log likelihood ratio maximized over all window rotations and sizes.

At significance level $\alpha = 0.05$ and $K = 15$, 34 significant poverty hotspots are detected (Appendix 1). There are 199 villages that are included in these significant poverty hotspots. There are no geographical overlap among all reported hotspots.

Table 2 shows the scanning result that presents the five most significant hotspots. At significance level $\alpha = 0.05$, null-hypothesis is rejected, meaning that there is an elevated risk inside the window (hotspot) when compared to the outside.

Hotspot 1 that attains the highest log likelihood ratio (LLR) can also named as the Most Likely Cluster (MLC). Relative risk of poverty in Hotspot 1 is equal to 2.31, which means that the probability of poverty in Hotspot 1 is 2.31 times higher than the expected.

### Table 2. The five most significant poverty hotspots

<table>
<thead>
<tr>
<th>Scan Area</th>
<th># units</th>
<th>LLR</th>
<th>P-value</th>
<th>RR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hotspot 1</td>
<td>13</td>
<td>3347.59</td>
<td>0.001</td>
<td>2.31</td>
</tr>
</tbody>
</table>

The average distance from villages within Hotspot 1 to the nearest city is 71.85 km. It indicates that the accessibility of these villages is quite far from the city center. The main economic potency of Hotspot 1 is farming with the proportion of farming family is 65.69%. This hotspot is not an industrial area. There is neither bank nor cooperation that exists in this hotspot.

The second rank of log likelihood ratio is gained by Hotspot 2 (Table 2). There are 63.62% of 21,168 families in this hotspot that are categorized as poor family (Appendix 2). Relative risk of poverty in Hotspot 2 is equal to 2.16, which means that the probability of poverty in Hotspot 2 is 2.16 times higher than the expected.

Villages that comprised Hotspot 2, mentioned in Table 4, are concentrated in two districts, Pamekasan and Cibuburung.

### Table 4. Villages that comprised Hotspot 2

<table>
<thead>
<tr>
<th>No.</th>
<th>Village</th>
<th>District</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cisauh</td>
<td>Pamekasan</td>
</tr>
<tr>
<td>2</td>
<td>Gunung Picung</td>
<td>Pamekasan</td>
</tr>
<tr>
<td>3</td>
<td>Cibening</td>
<td>Pamekasan</td>
</tr>
</tbody>
</table>
### Table 5. Villages that comprised Hotspot 3

<table>
<thead>
<tr>
<th>No.</th>
<th>Village</th>
<th>District</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Purwasari</td>
<td>Dramaga</td>
</tr>
<tr>
<td>2</td>
<td>Petir</td>
<td>Dramaga</td>
</tr>
<tr>
<td>3</td>
<td>Suka Damai</td>
<td>Dramaga</td>
</tr>
<tr>
<td>4</td>
<td>Sukawening</td>
<td>Dramaga</td>
</tr>
<tr>
<td>5</td>
<td>Neglasari</td>
<td>Dramaga</td>
</tr>
<tr>
<td>6</td>
<td>Sinar Sari</td>
<td>Dramaga</td>
</tr>
<tr>
<td>7</td>
<td>Ciherang</td>
<td>Dramaga</td>
</tr>
<tr>
<td>8</td>
<td>Dramaga</td>
<td>Dramaga</td>
</tr>
<tr>
<td>9</td>
<td>Sukaharja</td>
<td>Ciomas</td>
</tr>
</tbody>
</table>

The average distance from villages within Hotspot 3 to the nearest city is 24.78 km. The mean economic potency of Hotspot 3 is farming, mining, manufacturing, and commerce. The proportion of farming family is 22.78%. There are 4 middle industries and 2 small industries in this hotspot. This hotspot has six units of cooperation but bank does not exist.

### Table 6. Villages that comprised Hotspot 4

<table>
<thead>
<tr>
<th>No.</th>
<th>Village</th>
<th>District</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kopo</td>
<td>Cisarua</td>
</tr>
<tr>
<td>2</td>
<td>Leuwimalang</td>
<td>Cisarua</td>
</tr>
<tr>
<td>3</td>
<td>Cipayung Girang</td>
<td>Mega Mendung</td>
</tr>
<tr>
<td>4</td>
<td>Gunung Geulis</td>
<td>Sukaraja</td>
</tr>
<tr>
<td>5</td>
<td>Cijayanti</td>
<td>Babakan Madang</td>
</tr>
<tr>
<td>6</td>
<td>Bojong Koneng</td>
<td>Babakan Madang</td>
</tr>
<tr>
<td>7</td>
<td>Sumur Batu</td>
<td>Babakan Madang</td>
</tr>
<tr>
<td>8</td>
<td>Hambalang</td>
<td>Citeureup</td>
</tr>
<tr>
<td>9</td>
<td>Pasir Muki</td>
<td>Citeureup</td>
</tr>
</tbody>
</table>

Hotspot 5 attains the fifth rank of log likelihood ratio (Table 2). The population of family in Hotspot 5 is 21,568 families and 54.51% of them are living as poor family (Appendix 2).

Relative risk of poverty in Hotspot 5 is equal to 1.85, which means that the probability of poverty in Hotspot 5 is 1.85 times higher than the expected.
REFERENCES


CONCLUSION

At significance level $\alpha = 0.05$ and $K = 15$, significant hotspots are detected among 91 hotspot candidates. These significant poverty hotspots consist of 199 villages. There is no graphical overlap among all reported hotspots.

The Most Likely Cluster (MLC) for the region of Bogor that consists of 13 villages attains the highest log likelihood ratio 7.59 and relative risk 2.13. The MLC comprises two districts that are Sukajaya and Jasinga. The MLC in poverty hotspots mapping is concentrated in the western of Bogor.

Poverty hotspot detection is very useful to provide early stage to conduct planning strategy for poverty reduction. The characteristics underlying the hotspots have important policy implications for poverty reduction programs.

RECOMMENDATION

There are several limitations of flexibility in spatial scan statistic. To detect noncircular clusters (hotspots), this spatial scan statistic needs more efficient algorithm for larger cluster sizes and for reducing parameter estimation.
<table>
<thead>
<tr>
<th>Hotspot</th>
<th>Number of units</th>
<th>LLR</th>
<th>P-value</th>
<th>RR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13</td>
<td>3347.59</td>
<td>0.001</td>
<td>2.309</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>3216.96</td>
<td>0.001</td>
<td>2.158</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>2831.73</td>
<td>0.001</td>
<td>2.287</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>1923.06</td>
<td>0.001</td>
<td>1.915</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>1883.59</td>
<td>0.001</td>
<td>1.852</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>1649.61</td>
<td>0.001</td>
<td>1.845</td>
</tr>
<tr>
<td>7</td>
<td>11</td>
<td>1350.37</td>
<td>0.001</td>
<td>1.775</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>1195.76</td>
<td>0.001</td>
<td>1.822</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>1111.23</td>
<td>0.001</td>
<td>2.056</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>1006.53</td>
<td>0.001</td>
<td>1.731</td>
</tr>
<tr>
<td>11</td>
<td>9</td>
<td>972.001</td>
<td>0.001</td>
<td>1.851</td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>885.936</td>
<td>0.001</td>
<td>1.646</td>
</tr>
<tr>
<td>13</td>
<td>12</td>
<td>773.654</td>
<td>0.001</td>
<td>1.537</td>
</tr>
<tr>
<td>14</td>
<td>7</td>
<td>749.812</td>
<td>0.001</td>
<td>1.911</td>
</tr>
<tr>
<td>15</td>
<td>8</td>
<td>634.517</td>
<td>0.001</td>
<td>1.575</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>619.564</td>
<td>0.001</td>
<td>2.757</td>
</tr>
<tr>
<td>17</td>
<td>8</td>
<td>611.784</td>
<td>0.001</td>
<td>1.535</td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>383.505</td>
<td>0.001</td>
<td>2.307</td>
</tr>
<tr>
<td>19</td>
<td>8</td>
<td>367.174</td>
<td>0.001</td>
<td>1.509</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>326.038</td>
<td>0.001</td>
<td>1.710</td>
</tr>
<tr>
<td>21</td>
<td>4</td>
<td>263.476</td>
<td>0.001</td>
<td>1.636</td>
</tr>
<tr>
<td>22</td>
<td>8</td>
<td>245.19</td>
<td>0.001</td>
<td>1.330</td>
</tr>
<tr>
<td>23</td>
<td>1</td>
<td>231.785</td>
<td>0.001</td>
<td>2.281</td>
</tr>
<tr>
<td>24</td>
<td>3</td>
<td>208.611</td>
<td>0.001</td>
<td>1.512</td>
</tr>
<tr>
<td>25</td>
<td>1</td>
<td>199.932</td>
<td>0.001</td>
<td>2.226</td>
</tr>
<tr>
<td>26</td>
<td>1</td>
<td>195.077</td>
<td>0.001</td>
<td>2.395</td>
</tr>
<tr>
<td>27</td>
<td>2</td>
<td>185.747</td>
<td>0.001</td>
<td>1.587</td>
</tr>
<tr>
<td>28</td>
<td>1</td>
<td>178.089</td>
<td>0.002</td>
<td>1.960</td>
</tr>
<tr>
<td>29</td>
<td>1</td>
<td>176.831</td>
<td>0.002</td>
<td>2.099</td>
</tr>
<tr>
<td>30</td>
<td>3</td>
<td>171.433</td>
<td>0.002</td>
<td>1.547</td>
</tr>
<tr>
<td>31</td>
<td>1</td>
<td>161.97</td>
<td>0.002</td>
<td>1.849</td>
</tr>
<tr>
<td>32</td>
<td>1</td>
<td>154.779</td>
<td>0.002</td>
<td>1.527</td>
</tr>
<tr>
<td>33</td>
<td>10</td>
<td>151.092</td>
<td>0.003</td>
<td>1.194</td>
</tr>
<tr>
<td>34</td>
<td>2</td>
<td>138.276</td>
<td>0.004</td>
<td>1.456</td>
</tr>
</tbody>
</table>

Overall Number of Poor family: 322,185 families
Overall Population: 1,094,480 families

RR = 1 → q = 0.294
## Appendix 2. Scanning result of five most significant poverty hotspots

<table>
<thead>
<tr>
<th>Scan Area</th>
<th>Number of units</th>
<th>( q )</th>
<th>( N(z) )</th>
<th>Population</th>
<th>( E(z) )</th>
<th>%poverty</th>
<th>LLR</th>
<th>P-value</th>
<th>RR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hotspot 1</td>
<td>13</td>
<td>0.294</td>
<td>12127</td>
<td>17842</td>
<td>5252.1972</td>
<td>67.97%</td>
<td>3347.59</td>
<td>0.001</td>
<td>2.31</td>
</tr>
<tr>
<td>Hotspot 2</td>
<td>13</td>
<td>0.294</td>
<td>13468</td>
<td>21168</td>
<td>6231.2807</td>
<td>63.62%</td>
<td>3216.96</td>
<td>0.001</td>
<td>2.16</td>
</tr>
<tr>
<td>Hotspot 3</td>
<td>9</td>
<td>0.294</td>
<td>10495</td>
<td>15582</td>
<td>4586.9149</td>
<td>67.35%</td>
<td>2831.73</td>
<td>0.001</td>
<td>2.29</td>
</tr>
<tr>
<td>Hotspot 4</td>
<td>9</td>
<td>0.294</td>
<td>10934</td>
<td>19397</td>
<td>5709.9467</td>
<td>56.37%</td>
<td>1923.06</td>
<td>0.001</td>
<td>1.91</td>
</tr>
<tr>
<td>Hotspot 5</td>
<td>11</td>
<td>0.294</td>
<td>11756</td>
<td>21568</td>
<td>6349.0297</td>
<td>54.51%</td>
<td>1883.59</td>
<td>0.001</td>
<td>1.85</td>
</tr>
</tbody>
</table>
Appendix 2: Mapping of five most significant poverty hotspots in Bogor