Spatial Scan Statistic for AIDS Hotspots Detection at Regencies and Municipalities in Java

Heri Japar Sodik, Asep Saeufuddin, Dian Kusumaningrum

Department of Statistics, Institut Pertanian Bogor

Abstract
Scan statistic is a statistical method with many potential applications. It is designed to detect a local excess of events and to test if such an excess can reasonably have occurred by chance. It can be applied on wide area of interest, such as on disease clustering. In this research spatial scan statistic was applied important to detect the locations of the significant hotspot clusters of severe and dangerous disease, AIDS. This research applied the probability distribution generating events of poison under the null hypothesis, supported by likelihood ratio test analysis to get the statistic test value and a Monte Carlo simulation to obtain the significant hotspot cluster from calculation of those statistic test values in computational process algorithm. Analysis with 19 percents of the population at risk of maximum spatial cluster size and the circular spatial window shape, spatial scan statistic presented the significant hotspot clusters of AIDS cases in Java. The places considered as hotspot are the municipalities of Center Jakarta and North Jakarta, Bandung, Malang, Surabaya, Semarang, Yogyakarta, and Salatiga. Based on the fact, government has to provide AIDS eradication and surveillance program in the hotspot area while the program of prevention must be implemented in the surrounding area (non hotspot area).

Keywords: scan statistic, spatial scan statistic, AIDS hotspot, maximum spatial cluster size

Introduction
Background
AIDS is the most severe infection of HIV. HIV has killed so many CD4+ T cells (vital organs of the human immune system). Once a person is infected by HIV to fewer than 200 of these cells per microliter of blood causing cellular immunity is lost. Global HIV prevalence (the proportion of people living with the virus) appears to have leveled off. However, the number of people living with HIV has increased to 33.2 million in 2007 from 29.0 million in 2001. There were 2.5 million new AIDS cases and 2.1 million people have died from AIDS-related illnesses. Meanwhile in Asia, an estimate of 4.9 million people are living with HIV (2007), including the 440,000 people who have just recently got infected in the past year. Approximately 300,000 people have died from AIDS-related illnesses. There were approximately 20% more new HIV cases in East Asia compared to 2001. According to United Nations Programme on HIV/AIDS (UNAIDS) (2008), HIV prevalence is highest in South-East Asia, with wide variation of epidemic trends between different countries. Myanmar, Thailand and Cambodia show declines in prevalence, but the epidemic is growing at a particularly high rate in Indonesia and Viet Nam.

The HIV epidemic in Indonesia is among the fastest-growing in Asia. Although adult national prevalence is still low, almost 200,000 people are living with HIV across this heavily populated archipelago (Centers for Disease Control and Environmental Health and National AIDS Commission, 2006 in UNAIDS, 2008). HIV epidemic is initially growing rapidly among injecting drug users in Bali, Jakarta and West Java. Now encompasses many of their non-injecting sex partners, as well as prisoners, sex workers and their clients, and has spread to 32 provinces and 169 districts. In 2000, half of these provinces had reported HIV or AIDS case and in the next 7 years in 2007, all provinces had it (Aidsindonesia, 2008). Every year people who suffer from AIDS in Indonesia increases. Most of them are caused by injecting drug. The spreading of AIDS is also caused by blood transfusion, homosexual, transsexual, sex worker and their clients, etc. AIDS is one of the killing and spreading disease. In fact, new born babies can also from suffer AIDS because they are infected by their parents.

HIV and AIDS can affect economic growth by reducing the availability of human capital. Without proper nutrition, health care and medicine that is available in developed countries, large numbers of people are becoming victim to AIDS. They will not only be unable to work, but will also require significant medical care (Wikipedia, 2008). It is forecasted that AIDS will likely cause a collapse of economies and societies in countries with a significant AIDS population. Base on wiktionary (2008), in some heavily infected areas such as some african countries, the epidemic has left behind many orphans cared for by elderly grandparents. Mortality increases in this region, causing a decrease of skilled population and labor force. Less labor force will be predominantly young people, with reduced knowledge and work experience leading to a decline in productivity. An increase in workers’ time off to look after sick family members or for sick leave will also cause lower productivity. High mortality will also weaken the mechanisms that generate human
capital and investment in people, through loss of income and the death of parents. The death of mainly young adults caused by AIDS seriously weakens the tax paying population. This will reduce the resources available for public expenditures such as education and health services resulting an increasing pressure for the state's finances and causes economical slower growth (Wiktionary, 2008).

AIDS is evidently serious and dangerous problem for Indonesia. So, the research of AIDS in Indonesia as one of the AIDS epidemic in South-East Asia (UNAIDS, 2008) is very important. Java as the capital state of Indonesia and also has most of the big cities in Indonesia, have been estimated to have the significant cases of AIDS. On a smaller level such as regency and municipality in Java’s provinces, it is more accurate to see the AIDS outbreaks. Java has many regencies and municipalities causing barriers for the government to prevent AIDS outbreaks. It will be very useful if the government knows where the statistically significant high rate AIDS cases are. The spatial scan statistic has the ability to identify the significant hotspot of AIDS cases in a cluster of regencies and municipalities in Java. This is supported by mapping the hotspots to see the distribution of significant hotspot clusters visually. So, if the hotspots of AIDS are known, then government and people can easily take action, control and prevent AIDS outbreaks.

Objective
The objectives of this study are the following:
1. To detect the statistically significant hotspots cluster of AIDS case in Java.
2. To perform geographical surveillance of those hotspots area.
3. To be an early warning to the government and the people of those hotspots area and their surroundings.

Literature Review
AIDS
HIV (Human Immunodeficiency Virus) is either of two related viruses (HIV-1 and HIV-2) that progressively destroy the body's immune system and can lead to AIDS. AIDS (Acquired Immune Deficiency Syndrome) is a set of symptoms and infections resulting from the damage to the human immune system caused by HIV (Wiktionary, 2008). This condition progressively reduces the effectiveness of the immune system and leaves individuals susceptible to opportunistic infections and tumors. HIV is transmitted through direct contact of a mucous membrane or the bloodstream with a bodily fluid containing HIV, such as blood, semen, vaginal fluid, preseminal fluid, and breast milk. The transmission can involve anal, vaginal or oral sex, blood transfusion, contaminated hypodermic needles, exchange between mother and baby during pregnancy, childbirth, or breastfeeding, or other exposure to one of the above bodily fluids.

The symptoms of AIDS are primarily the result of conditions that do not normally develop in individuals with healthy immune systems. Most of these conditions are infections caused by bacteria, viruses, fungi and parasites that are normally controlled by the elements of the immune system that HIV damages. Opportunistic infections are common in people with AIDS. HIV affects nearly every organ system. People with AIDS also have an increased risk of developing various cancers such as Kaposi's sarcoma, cervical cancer and cancers of the immune system known as lymphomas. Additionally, people with AIDS often have systemic symptoms of infection like fevers, sweats (particularly at night), swollen glands, chills, weakness, and weight loss. The specific opportunistic infections that AIDS patients develop depend in part on the prevalence of these infections in the geographic area in which the patient lives (Wiktionary, 2008).

Hotspot
Hotspot is something unusual, anomaly, aberration, outbreak, elevated cluster, critical resource area, etc (Pauli and Taille, 2004). The declared need for monitoring, etiology, management, or early warning. The responsible factors may be natural, accidental, or intentional. Hotspot were generated by setting the relative risk in some counties to be larger than one (Song and Kulldorff, 2003).

Spatial Scan Statistic
Scan statistic is statistical method to detect clusters in a point process. Spatial scan statistic is used to determine whether a spatial point process contains a localized cluster of points somewhere in a region of interest. In the original problem, studied by Naus (1965), there is a homogeneous Poisson process on a rectangular region.

The spatial scan statistic deals with the following situation. A region R of Euclidian space is subdivided into cells defined (denote by A). Data are available in the form of a count on each cell A. In addition, A size value P(A) is associated with each cell. The cell sizes P(A) are assumed known and fixed, while the cell counts N(A) are independent random variables. Poisson distributional settings are commonly studied that P(A) is a positive real number and N(A) ~ Poisson (μ(A)P(A)), where μ(A) > 0 is an unknown parameter attached to cell A.

The spatial scan statistic seeks to identify clusters of cells that have an elevated response compared with the rest of the region. Elevated response means large values for the rates, μ(A) = N(A)/P(A), instead of for the raw counts N(A). Cell counts are thus adjusted for cell sizes before comparing cell responses.

Anderson and Titterington (1997) presented the following algorithm for a circular window of fixed
diameter $d$ on a homogeneous Poisson (assuming homogenous variations) process:
1. Identify the locations $(x,y)$ of two events no more than distance $d$ apart.
2. Construct the two circles of diameter $d$ for which $x$ and $y$ lie on the circumference.
3. Identify the number of events that lie on or inside each of the two circles and let $n$ be the larger of those two numbers.
4. Repeat steps 1 to 3 for all relevant pairs of location and report the largest of the resulting $n$ values as being the scan statistics.

The relative risk is a non-negative number, representing how much more common disease is in the location and time period compared to the baseline. Setting a value of one is equivalent to not doing any adjustments and a value of less than one to adjust for lower risk A value of greater than one is used to adjust for an increased risk. A cluster with a relative risk (RR) value greater than one is defined as a candidate of hotspot. A relative risk of zero is used to adjust for missing data for that particular time and location (Kulldorff, 2006). The relative risk is calculated by (Kulldorff, 2006):

$$RR = \frac{n(Z)}{E(c)}$$

where $n(Z)$ is the number of observed cases, and $E(c)$ is the expected number of cases in a location which is calculated by:

$$E(c) = p(C)$$

Where $p$ is the number of population in the cluster of interest, while $C$ and $P$ are the total number of cases and total number of population.

Currently, it is known that spatial scan statistic window has some limitation as the following:

1. Circles have been used for the scanning window, resulting in rather low power for detection of irregularly shaped cluster. Alternatively, an irregularly shaped cluster may reported as a series of circular clusters.
2. The response variable has been defined on the cells of a tessellated geographic region, preventing application to responses defined on a network.
3. Reflecting the epidemiological origins of the spatial scan statistics, response distributions have been taken as discrete specifically binomial or poisson.

**Poisson Models**

Let $N$ denote a spatial point process where $N(A)$ is the random number of points in the set $A \subset \mathbb{R}$. As the window moves over the study area it defines a collection $\Omega$ of cluster $Z \subset \mathbb{R}$. Interchangeably, $Z$ will be used to denote both a subset of $\mathbb{R}$ and a set of parameters defining the cluster.

Under the Poisson model, points are generated by an inhomogeneous Poisson (assuming inhomogeneous variance) process. In the model, the cluster has probability $p$ of being a point, while the probability for individuals outside the cluster is $q$. There is exactly one cluster $Z \subset \mathbb{R}$ such that $N(A) \sim \text{Pois}(p(\mu(A \cap Z) + q\mu(A \cap Z^c)) \forall A$. The hypothesis is $H_0 : p = q$ and $H_1 : p > q$, $Z \in \Omega$. Under $H_0$, $N(A) \sim \text{Pois}(p\mu(A)) \forall A$. Note that one of the parameters, $Z$, disappears under the null hypothesis.

The likelihood ratio test is needed to get a value of the statistical test. The likelihood function for the Poisson model is a little more complex (Kulldorff, 1997). For the numerator we first take the supremum over all $p$ and $q$ for fixed $Z$. The likelihood function is maximum when $p = n_x / \mu(Z)$ and $q = (n_x - n_z) / (\mu(R) - \mu(Z))$, so

$$L(Z) = \begin{cases} e^{-n_x} \frac{n_x^{n_z}}{\mu(Z)^{n_z}} \frac{n_z}{\mu(R)^{n_z}} & \text{if } \frac{n_x}{\mu(Z)} > \frac{n_z}{\mu(R) - \mu(Z)} \\ e^{-n_x} \left( \frac{n_x}{\mu(Z)} \right)^n \Pi_{x \in R} \mu(x) \end{cases}$$

The test statistic $\lambda$ of the likelihood ratio test can now be written as:

$$\lambda = \sup_{n_x \geq 0} L(Z)$$

$$= \sup_{n_x \geq 0} \left( \frac{\mu(Z)}{\mu(R)} \right)^{n_z} \left( \frac{\mu(Z)}{\mu(R) - \mu(Z)} \right)^{n_x - n_z} \Pi_{x \in R} \mu(x)$$

$$= \left( \frac{n_x}{\mu(Z)} \right) \left( \frac{\mu(Z)}{\mu(R) - \mu(Z)} \right)$$

$$\frac{n_x}{\mu(Z)} \geq \left( \frac{n_x - n_z}{\mu(Z)} \right)$$

If there is at least one cluster $Z$ such that

$$\frac{n_x}{\mu(Z)} \geq \left( \frac{n_x - n_z}{\mu(Z)} \right)$$

and $\lambda = 1$ otherwise. $l(\cdot)$ is the indicator function.

In order to find the value of the statistic test, we need a way to calculate the likelihood ratio as it is maximized over the collection of cluster in the alternative hypothesis. This might seem like a daunting task since the number of cluster could easily be infinite. Two properties allows us to reduce it to a
finite problem. The number of observed points is always finite and for a fixed number of points the likelihood decreases as the measure of the moving window increases.

Monte Carlo-Based Hypothesis Testing

Simulation is a method used to imitate a real-life system, especially when other analyzes are too mathematically complex or too difficult to reproduce (crystalball on adithm02, 2008). Monte Carlo simulation can be defined as a method to generate random sample data based on some known distribution for numerical experiments (Teknomo, 2008). Once the value of the test statistic has been calculated, it is easy to do the inference. We can not expect to find the distribution of the test statistic in closed analytical form. Instead we rely on Monte Carlo-Based hypothesis testing.

With a Monte Carlo test, the significance of an observed test statistic calculated from a set data is assessed by comparing it with a distribution obtained by generating alternative sets of data from some assumed model. If the assumed model implies that all data orderings are equally likely then this amounts to a randomization distribution.

Monte Carlo-based hypothesis testing was proposed by Dwass (1957), who pointed out that probability of falsely rejecting the null hypothesis is exactly according to the significance level, in spite of the simulation involved. Mantel (1967) proposed its use in terms of spatial points processes, while Turnbull et al. (1990) was the first to use in the context of a multidimensional scan statistic. Monte Carlo hypothesis testing for a scan statistic is a four-step procedure (Kulldorff, 1997):

1. Calculate the value of the test statistic for the real data.
2. Create a large number of random data sets generated under the null hypothesis.
3. Calculate the value of the test statistic for each of the random replications.
4. Sort the values of the test statistic from the real and random data sets, and note the rank of the one calculated from the real data sets. If it is ranked in the highest a percent, then reject the null hypothesis at a percent significance level.

For example, when we condition on the total number of points n, with 9999 such replications, the test is significant at the 5 percent level of a if the value of the test statistic for the real data sets is among the 500 highest values of the test statistic coming from the replications.

The p-value is obtained through Monte Carlo hypothesis testing (Dwass 1957 in Kulldorff 1997), by comparing the rank of the maximum likelihood from the real data sets with the maximum likelihood from the random data sets. If this rank is R, then p-value = R / 1 + #simulation.

Material and Method

Data Sources

This study is based on cumulative data of AIDS cases until December 31st 2007 quoted from Ministry of Health of Indonesia and people of year 2007 from independent project surveyor, at all regencies and municipalities in Java. Additionally, the data of latitude and longitude coordinate of those locations from Java digital polygon map also are used.

Methods

The methods used in this study are:

1. Provide and input the data that are needed for the analysis of this study. The required data are population, number of cases, and latitude and longitude coordinates.
2. Create the hypothesis testing model under the poison probability model assumption using the most compatible maximum spatial cluster size of the population. The cases evidently needed 19% of maximum spatial cluster size.
3. Scan the data with purely spatial type analysis on high rates scan for the areas to run the hotspot clusters’ process.
4. Calculate the relative risk of each regions scan window. Discard scan window with relative risk less than 1.
5. Estimate the value of the statistical test using the likelihood ratio test.
6. Perform the 999 Monte Carlo replication for statistical significant value of each hotspot candidate. The null hypothesis of no cluster was rejected when the simulated p-value was less than or equal to 0.05 for the primary and secondary clusters.
7. Interpret the results and present the significant hotspot clusters in a thematic java map.

Softwares which were used in this study are: Microsoft Excel 2007, MapInfo Professional 7.8 and SaTScan V6.1.

Results and Discussions

Java’s AIDS Cases

Nowadays Java island consists of six provinces: DKI Jakarta, West Java, Center Java, East Java, DI Yogyakarta, and a new province, Banten. There are 32 municipalities and 83 regencies in Java, so the total is 115 areas. Based on aidssindonesia (2008) the observed cumulative AIDS cases until December 2007 were 6332 cases from the total of population 135.760.230 in 28 municipalities and 82 regencies. Five regions were not included due to the Java digital polygon map limitations. Hence this study only researched cluster within municipalities and regencies registered on an existing digital map. So, in the thematic map there are not these five regions too because they did not register in the Java digital polygon map. But it was not a problem because the five regions do not have any reported cases of AIDS.
All of them have represented. Thus, those regions was not an influence for the analysis. The top ten AIDS cases evidently occurred in some municipalities only. There were not any regencies which include these top ten AIDS cases. It was not something strange, because most of AIDS cases grown vastly in the city. AIDS cases usually related with the culture of the people in a region. The people in the city is very susceptible for AIDS. But, based on the research, regencies' AIDS cases are still an influence the analysis. So, spatial scan statistic needs the surrounding area of those municipalities. The top ten AIDS cases based on descriptive statistics can be seen on the table below:

Table 1. Top Ten Municipalities on AIDS Case.

<table>
<thead>
<tr>
<th>Municipality</th>
<th>Total case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center Jakarta</td>
<td>1523 *</td>
</tr>
<tr>
<td>Bandung</td>
<td>1128 *</td>
</tr>
<tr>
<td>North Jakarta</td>
<td>710 *</td>
</tr>
<tr>
<td>Surabaya</td>
<td>493 *</td>
</tr>
<tr>
<td>Malang</td>
<td>394 *</td>
</tr>
<tr>
<td>West Jakarta</td>
<td>348</td>
</tr>
<tr>
<td>South Jakarta</td>
<td>329</td>
</tr>
<tr>
<td>Bogor</td>
<td>227</td>
</tr>
<tr>
<td>Semarang</td>
<td>167 *</td>
</tr>
<tr>
<td>Bekasi</td>
<td>155</td>
</tr>
</tbody>
</table>

These municipalities were not guaranteed to be hotspots. The municipalities which have sign * are to be hotspot. Many factors can cause an area to become hotspot such as surrounding area cases, population, likelihood ratio test value, relative risk value, p-value, etc. Spatial scan statistic considers many things to get compatible results of hotspot detection. Using the SaTScan software, the calculations for 6332 AIDS cases in 110 area on Intel Celeron processor of 1.86 GHz and 80 GB HDD notebook took 4 seconds to get an output of hotspot clusters.

Detected Significant Hotspots

There were seven significant AIDS hotspots cluster detected. They were the municipalities of Center Jakarta and North Jakarta, Bandung, Malang, Surabaya, Semarang, Yogyakarta, and Salatiga. All of them have the p-value less than 0.05 of significance level and the relative risk values greater than one. Based on kulidoff (2006), the program of SaTScan will scan for clusters of geographic size between zero and some upper limit defined by the user. By applying purely spatial scan statistic with 19% of maximum spatial cluster size, circular spatial window shape, and 999 Monte Carlo replications under the poisson model assumption, those seven significant AIDS hotspots cluster are obtained. Among the maximum spatial cluster size that has been analyzed in this research, 19% of maximum spatial cluster size is most reasonable and accurate. For further detail, see appendix 1. Center Jakarta and North Jakarta were to be the most likely cluster for AIDS cases with a radius of 4.45 kilometers from its centroid. AIDS cases in those area were unusual or very high compared with the others. Center Jakarta has the highest rate case of AIDS because it is known that more than 70% were infected by injection drug (Sihaloha, 2007). North Jakarta is known to be an AIDS spread area due to geographic position which is near coastal area with high mobility society. North Jakarta is also known for its biggest prostitution place in Indonesia called Kramat Tunggak (jakartautara, 2008). If they were under the poisson probability model assumption, the total cases should be only ±109 cases. But, the observed cases was 2233. Number of observed case in this cluster was 20.641 times higher than expected case. Hence, the relative risk of the people which suffer AIDS in Center Jakarta and North Jakarta is 31.341times than in the surrounding areas. With the highest relative risk and statistic value of log likelihood ratio (5047.995520) calculated by Monte Carlo simulation procedures, this cluster was the most significant hotspot. Such condition is known as the most likely cluster. Appendix 2 shows the output of SaTScan analysis completely.

There were six secondary clusters besides a most likely cluster. They were municipalities of Bandung, Malang, Surabaya, Semarang, Yogyakarta, and Salatiga. This result is reasonable because base on AIDS Solve Comission (Komisi Penanggulangan AIDS/KPA) (2008), the highest rate of AIDS case in West Java is in Bandung municipality. KPA regional of Bandung is planning to build an AIDS Center there (Aidindonesia, 2008). Based on the data of Health Department of Indonesia Republic on Surabaya-wealth (2008), East Java has the fourth level of highest AIDS case in Indonesia. According to the KPA chairman regional of Malang on ANTARA (2008), every month an average of ten people are vonised HIV/AIDS positive in Malang and most of them live in Malang municipality. Additionally, Surabaya is a high rate risk municipality of AIDS in East Java because actually it has one of the biggest prostitution place in Indonesia called Doli. In Yogyakarta, based on a survey result in 2004, Yogyakarta municipality is on a concentrated epidemic level. On this level, 5% or more of the population have high rate risk to be infected by HIV(Aidindonesia, 2008). Based on the head of Center Java health official sited on menkokesra (2008), Center Java is the ninth level of high rate case of AIDS in Indonesia. Semarang and Salatiga municipalities are two of ten area with AIDS high risk in Center Java. The thematic map of the significant hotspot clusters is depicted on Appendix 3.

The six secondary clusters were also significant hotspots. It has a p-value of 0.001, similar with the
most likely cluster except for Salatiga. This cluster had a p-value of 0.024. The municipality of Salatiga was unpredictable because the other significant hotspot clusters of AIDS’ case were big, busy, or crowded cities. Salatiga is a hotspot because of the most surrounding areas case of Salatiga have few cases with a rather large population compared to Salatiga. Such as regencies of Semarang, Boyolali, Magelang, grobogan, and Klaten. So, it might be an influence to Salatiga becoming a hotspot cluster. In fact, based on infosalatiga.wordpress (2008), Salatiga is the fifth high risk of HIV/AIDS in Center Java. Salatiga was the opposite of West Jakarta. Although West Jakarta has more cases than Salatiga, but most of the surroundings area of West Jakarta have high cases. Salatiga is an example that even though the case number was not too high compared to the other hotspot areas, but the location is geography strategic causing the relative risk and log likelihood ratio value to be a significant hotspot. This can be a weakness of spatial scan statistic analysis if the number of cases of Salatiga is very small compared with the surrounding areas. In addition, the maximum spatial cluster size is not an influence to the analysis. It means that even we change the value of maximum spatial cluster size many times, but the result of SaTScan shows that significant hotspot is location with very small case. The following shows the table of AIDS significant hotspot clusters including the statistic value of Relative Risk (RR), Log Likelihood Ratio (LLR), and P-value.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>RR</th>
<th>LLR</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center Jakarta and North Jakarta</td>
<td>31.341</td>
<td>5047.995</td>
<td>0.001</td>
</tr>
<tr>
<td>Bandung</td>
<td>13.181</td>
<td>1769.864</td>
<td>0.001</td>
</tr>
<tr>
<td>Malang</td>
<td>10.459</td>
<td>558.158</td>
<td>0.001</td>
</tr>
<tr>
<td>Surabaya</td>
<td>3.880</td>
<td>291.505</td>
<td>0.001</td>
</tr>
<tr>
<td>Semarang</td>
<td>3.847</td>
<td>100.172</td>
<td>0.001</td>
</tr>
<tr>
<td>Yogyakarta</td>
<td>2.313</td>
<td>17.762</td>
<td>0.001</td>
</tr>
<tr>
<td>Salatiga</td>
<td>2.744</td>
<td>7.834</td>
<td>0.024</td>
</tr>
</tbody>
</table>

in Java. Government and all Indonesians should focus and be more serious to control and prevent AIDS spreading from those hotspot cluster. Government has to provide AIDS eradication and surveillance program in the hotspot area while the program of prevention must be implemented in the surrounding area (non hotspot area). Hence, hopefully AIDS cases do not spread out to other people in surrounding areas. Avoid and prohibit drugs, avoid free sex or having sex with one’s legitimate partner, enforce religious education at all levels, and spread out knowledge about the danger of AIDS in advance are some of solution to solve that problem. It will minimize the people which suffer AIDS especially in Java.

Conclusion

There were seven AIDS’ significant hotspot clusters at 5% significance level which are located in the municipalities of Center Jakarta and North Jakarta, Bandung, Malang, Surabaya, Semarang, Yogyakarta, and Salatiga. Center Jakarta and North Jakarta are to be the most likely cluster. More care, control, priority, and full concentration in significant hotspot area by government and people are very useful to solve AIDS cases outbreak.

Spatial scan statistic has a weakness if a hotspot cluster has very small number of cases, although if compared with the surrounding areas, it is higher. Additionally, the maximum spatial cluster size is not influence the analysis and the result is not compatible with the fact.

Recommendations

The next research should try the other methods for AIDS’ hotspot detection such as temporal scan statistic or space time scan statistic. It will be interesting to compare when and where the hotspot cluster occurs. The result will be more useful for the people, government, and institution that have a concern on AIDS cases. Besides that, it is interesting to find the data which are compatible with the distribution probability model of Bernoulli, ordinal, and exponential such which are provided in the SaTScan software. It will be also interesting to is make a research to find how small of cases and what other factors in an area that will become a hotspot cluster.

References

Adithme02. 2008. Simulasi Monte Carlo. [June 25th, 2008]

Aidsindonesia. 2008. AIDS di Kota Bandung. [August 18th, 2008]

Aidsindonesia. 2008. AIDS di Kota Yogyakarta. [August 18th, 2008]


[May 7th, 2008]

[August 18th, 2008]

Bungsu, Anita Putri. 2006. The Spatial Scan Statistic For Detecting Dengue Fever Hotspots in Bogor Municipality, Departement of Statistics, Mathematics and Natural Science Faculty, Bogor Agricultural University. [Thesis]

[August 18th, 2008]

[September 14th, 2008]


[May 7th, 2008]

[May 7th, 2008]

[September 15th, 2008]

[August 18th, 2008]


[May 7th, 2008]

[May 7th, 2008]

[August 21st, 2008]


[July 21st, 2008]