
Article

Distribution and Spatial Pattern Analysis on Dengue Cases in Seremban District, Negeri Sembilan, Malaysia

Nuriah Abd Majid *, Nurafiqah Muhamad Nazi and Ahmad Fariz Mohamed
Institute for Environment and Development (LESTARI), University Kebangsaan Malaysia, Bangi 43600, Selangor Darul Ehsan, Malaysia
* Correspondence: nuriah@ukm.edu.my; Tel.: +603-89217639

Received: 24 April 2019; Accepted: 28 May 2019; Published: 28 June 2019

Abstract: Dengue fever disease increases alongside urbanization rate in tropical countries. Hence, the need to visualize the distribution pattern of increases is vital for the management of dengue cases, especially in Malaysia. Thus, the dengue surveillance system is proposed for the monitoring of dengue cases using computer-generated modeling for spatial distribution patterns, which is important for management and control. The present study performed distribution and spatial pattern analysis of dengue cases reported in the growing Seremban district in Negeri Sembilan, Malaysia in 2008 and 2009. The purpose of the study is to evaluate the pattern of distribution and determine whether it is clustered or dispersed. A total of 1401 and 1056 cases for dengue-related diseases were reported by the Ministry of Health Malaysia in Seremban district in the years 2008 and 2009, respectively. Three spatial statistical analysis were conducted: Spatial mean center, directional distribution, and standard distant on distribution of dengue cases reported. This study found that the distribution pattern for dengue cases is clustered. Spatial mean center and directional distribution for both sets of years have slight differences. Meanwhile, standard distance for dengue cases reported in the year 2008 is 22,085.82 m, which is bigger than dengue cases reported in 2009, showing a standard distance of 20,318.35 m. More sets of cases throughout years are required in further studies to identify factors that contribute to dengue epidemiology in the Seremban district undergoing urbanization.

Keywords: dengue; spatial mean center; standard distance; directional distribution

1. Introduction

Dengue is a mosquito-borne disease spreading globally and increasing gradually every year. There was about 2.2 million dengue reported globally in 2010, which rose up to 3.2 million cases in 2015 [1]. Intensified dengue cases among the human population in 128 countries has become a major threat, as dengue transmission shows negative signs of decline even though many researches and mitigation programs have been worked on. Dengue fever is one of the major vector-borne viral diseases in the world, mainly in the tropical and subtropical regions [2,3]. It is a most significant public health issue globally, and approximately one-third of the population of the tropical and subtropical regions are exposed to the risk of dengue [4]. About 40% of the world’s population is now at dengue risk, as estimated by the World Health Organization (WHO) [5]. Dengue cases have been reported in tropical areas, especially in the Asia Pacific countries such as Cambodia, China, Laos, Malaysia, Philippine, Singapore, Vietnam, and Australia, as well as countries in the Pacific Islands [6]. It is found that the main vectors of dengue are active through Aedes mosquitoes when the surrounding temperature is less than 28 °C. Aedes mosquitoes also require clean stagnant water to lay eggs and reproduce [7]. Thus, the tropical and subtropical countries are the most suitable areas for the spread of dengue cases. Dengue fever was first reported in Malaysia in the early 1900s. Later on, in 1962, dengue hemorrhagic fever emerged in Penang [8–10]. Malaysia is a developing country with climate category
of hot and humid throughout the year. Malaysia receives on average 250 centimeters of rainfall in a year and the average temperature is 27 °C [11]. This condition is favored by domestic *Aedes* mosquitoes that live with the human population and reside under the roof, balcony, and porch [12]. It is known that the dengue virus spreads briskly due to the high adaptive tendency of the *Aedes* species with climate change, common areas suitable as breeding sites, strong population mobility, and international trade [13–15]. There are two female *Aedes* mosquito species, namely *Aedes aegypti* as the main dengue vector and *Aedes albopictus* as the secondary vector, which prefer to breed outdoors. Female *Aedes* mosquitoes bite and suck on human blood to obtain protein and iron mineral, thus increasing the chances to transmit dengue viruses [12,16,17]. The bite of *Aedes* mosquitoes transmits dengue viruses from four flavivirus (DEN-1, DEN-2, DEN-3, and DEN-4) and causes dengue fever and dengue hemorrhagic fever that can lead to death [18]. Early symptoms of dengue fever include high body temperature up to 40 °C, headache, eye pain, muscle or joint pain, nausea, swollen gland, and petechial rash. Abdominal symptoms, bleeding tendencies, and respiratory or circulatory disturbances are typical symptoms of dengue hemorrhagic fever [5]. Multiple dengue prevention programs conducted from different perspective, from the field of clinical, engineering, and vector control biology have been done each year.

As part of rigorous efforts in dengue epidemiology control and assessment in Malaysia, the Ministry of Health (MOH) has collaborated with the Ministry of Science, Technology, and Innovation (MOSTI) to implement dengue transmission monitoring using remote sensing (RS), namely iDengue (idengue.remotesensing.gov.my), under supervision of Remote Sensing Agency Malaysia [19]. This surveillance system monitors dengue transmission across the country by updating dengue fevers reported in every hospital and medical institution on daily basis. This initiative was established to deliver information on dengue cases statistics and dengue hotspots to the public for further preventive measure to combat *Aedes* mosquitoes and dengue fever.

Prior to that, researchers have worked together in the academic analysis on dengue surveillance with aims to estimate dengue transmission using predictive mathematical-based modeling and spatial-temporal analysis of mosquito distribution and localities. The outcomes of these researches proposed great potential for cost-efficiency on dengue surveillance and control in regional and national scales [20].

Therefore, the epidemic transmission monitoring using GIS-based mapping is gradually become important for disease control and management [21–24]. GIS and spatial statistical analysis are two important tools used for monitoring and mapping the vector-borne study and its geographical pattern [25,26]. Over the years, geoprocessing and digital mapping techniques were applied for public health issues analysis, as well as the use of spatial analysis programs to visualize the spatial distribution patterns of epidemiological data [27,28]. These techniques allow the development of models to predict the possibilities of risk of disease and territorial infestation, mapping environmental and social conditions associated with such patterns [29]. Most of the studies analyzed spatial pattern of dengue cases reported using GIS software to determine the correlation between diseases statistical analysis with factors contributing to dengue outbreak [30–33]. Moreover, spatial statistics is a useful technique for the disease mapping, intensity, and environmental factor analysis. For example, cluster analysis, hotspot analysis, and spatial autocorrelation are commonly used by the researchers [34–38].

The purpose of this study is to evaluate the pattern of distribution of dengue cases and determine whether it is clustered or dispersed in the Seremban district in order to contribute to the effective decision-making process to control dengue outbreak in the district. Geographic information system (GIS) software ArcMap, version 10.5 was used to analyze and visualize the cluster pattern in geographical context [39]. Three approaches were performed to test the null hypothesis that there were no differences between dengue cases reported in the years 2008 and 2009. Spatial statistical analysis was conducted for both sets of year, focusing on spatial mean center, directional distribution, and standard distant analysis. Application of this knowledge would strengthen disease surveillance system in the field of health and GIS technology.
2. Materials and Methods

2.1. Study Area

This study is focused in the Seremban district, which is the capital of Negeri Sembilan, Malaysia and located 60 km south of Kuala Lumpur with coordinates of 02°43′20″N and 101°56′30″E that covers about 959 km² of land area (Figure 1). The average temperature in Seremban is 31 °C, with rainfall precipitation ranging from 250 mm to 320 mm [40]. Seremban, which is located east of Negeri Sembilan, has less slope than the middle part of the state. It is the significant district where extensive land use changes and human activities increase every year, which in turn promote population growth [41]. With a population of 620,100 in 2015, Seremban currently has the highest dengue hotspot localities within the district, mainly in the residential areas [19]. It is reported that a cumulative 1864 patients were admitted due to dengue-related diseases with death toll of six at the end of 2018 [42]. Along with the rise of development growth in the Seremban district that has been identified as part of Kuala Lumpur Extended Mega Urban Region (KLEMUR), which attracts population concentration to reside within its urbanized town, dengue epidemiology needs to be controlled for its urban dwellers to be free from fear of health problems in this district [43].

Figure 1. Study site.
2.2. Data Collection

Demographic data of dengue cases reported from hospitalized patients in every government and private hospital in Seremban in the years 2008 and 2009 obtained from Ministry of Health Malaysia contains: (1) Location; (2) number of cases; and (3) date start and last case recorded. During the years surveyed, a total of 2457 dengue cases were reported throughout the city of Seremban, with positive dengue fever selected for this study. Latitude and longitude coordinates of patients' locality were extracted from the data collection to perform distribution and spatial analysis.

2.3. Statistical Analysis

Statistical analysis is used to identify and confirm spatial patterns, such as the center of a group of features, the directional trend, and whether features form clusters. It also classifies and symbolizes the data can obscure or overemphasize patterns. Statistical functions analyze the underlying data and measures that can be used to confirm the existence and strength of the pattern [39]. Thus, the statistical analyses in this study will help to provide answers for linking distribution pattern of dengue cases with cause of the dengue.

Geographical distribution is to measure a set of features for calculation of value which represents a characteristic of the distribution, such as the center, orientation, or compactness [39]. This study focused on two methods: Spatial distribution and spatial distribution pattern analysis. Spatial distribution includes spatial mean center, standard distance, and directional distribution analysis. Spatial distribution pattern analysis is the average nearest neighbor method. Mapping of dengue epidemiology distribution trend identifies relationship with specific physical characteristics. Spatial pattern analysis was performed using the average nearest neighbor method. Distribution and pattern analysis were conducted to understand whether there were any dominant distributions and patterns for the dengue outbreak [44,45]. Pattern distribution details were minimized using pattern analysis. Clustered pattern showed there was a factor that causes dengue epidemiology in that area.

2.4. Spatial Mean Center

Spatial mean center is the average coordinates $x$ and $y$ of all features in the study area. It is useful to detect any changes in distribution or to compare features of distribution. This analysis shows leaning-centered phenomena, especially, to visualize the means of dengue cases reported in Seremban. Spatial mean center is to study changes in distribution detection to compare type and feature of distributions. Spatial mean center is able to create a new feature point classification where every feature represents mean center. The mean of $x$ and $y$ show the value of mean center, it means that the dimensional field was included as the product of features.

Mean center calculated as below:

$$
\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{1}
$$

$$
\bar{Y} = \frac{\sum_{i=1}^{n} y_i}{n} \tag{2}
$$

where $x_i$ and $y_i$ are coordinates for $i$, and $n$ is the total number of the features.

Weighted mean is as below:

$$
\bar{X}_w = \frac{\sum_{i=1}^{n} W_i x_i}{\sum_{i=1}^{n} W_i} \tag{3}
$$

$$
\bar{Y}_w = \frac{\sum_{i=1}^{n} W_i y_i}{\sum_{i=1}^{n} W_i} \tag{4}
$$

where $W_i$ is the weigh on characteristic $i$. 
Method to calculate center for three dimension is the $z$ attribute for every feature:

$$\bar{Z} = \frac{\sum_{i=1}^{n} Z_i}{n}$$

(5)

$$\bar{Z}_w = \frac{\sum_{i=1}^{n} W_i z_i}{\sum_{i=1}^{n} W_i}$$

(6)

2.5. Standard Distant

Standard distant measure distribution density, which provides a single value, represents the distribution around the mean center. The scatter of points around the center needs measurement. The value is the distance represented by deviations. Therefore, measures of spatial dispersion give information about the dengue spread around a center. The expression used is standard distance, which is equivalent to the statistician’s standard deviation. Standard distant is a tool used to create a round polygon of points dispersed around a center. In this study, a round of polygon represent dispersed dengue cases reported in each year from the mean center [39].

Standard distant shows a new feature class which includes a round, centered polygon on mean center (a single mean and a circle for every case). Every round polygon is created with radii that have the same value as the standard distant value. The attribute value for every round polygon is the $x$-coordinate mean centered polygon, $y$-coordinate mean centered, and standard distant (radius circle).

Calculation of standard distance is as follows:

$$SD = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{X})^2}{n} + \frac{\sum_{i=1}^{n} (y_i - \bar{Y})^2}{n}}$$

(7)

where $x_i$ and $y_i$ is the coordinates for feature $i$, $\{\bar{X}, \bar{Y}\}$ show mean center features, and $n$ is equal to the total of feature number.

Weighted standard distance is as follows:

$$SD_w = \sqrt{\frac{\sum_{i=1}^{n} W_i (x_i - \bar{X})^2}{\sum_{i=1}^{n} W_i} + \frac{\sum_{i=1}^{n} W_i (y_i - \bar{Y}_w)^2}{\sum_{i=1}^{n} W_i}}$$

(8)

where $w_i$ is the weigh on feature $i$ and $\{\bar{X}_w, \bar{Y}_w\}$ indicates mean center.

2.6. Standard Deviational Ellipse

As a GIS tool for delineating spatial point data, standard deviational Ellipse (SDE) was used to summarize spatial characteristics of geographical features, which are the central tendency, dispersion, and directional trends [39]. The common method to measure trend of one set location is to calculate standard distance separately on axis $x$ and $y$. Both steps determine ellipse axis while including distribution characteristics. Ellipse is also known as standard deviational ellipse as it is calculated by standard deviation from the $x$-coordinate and $y$-coordinate from mean center to determine ellipse axis. Ellipse allows distribution characteristics to go longitudinally and have certain orientation. Graphing ellipses was used in disease surveillance studies to predict spatial spread trends, since the central tendency and dispersion are two principal aspects concerned by epidemiologists [46,47]. The calculations are as in Equations 9a and 9b below:

$$SDE_x = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{X})^2}{n}}$$

(9a)
where \( x_i \) and \( y_i \) are the coordinates for feature \( i \), \([\bar{X}, \bar{Y}]\) shows mean center, and \( n \) is the total number of features.

Angle of rotation calculated as Equation (2) below:

\[
\tan \theta = \frac{A + B}{C}
\]

\[
A = \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} y_i^2
\]

\[
B = \sqrt{(\sum_{i=1}^{n} x_i^2)^2 - (\sum_{i=1}^{n} y_i^2)^2}
\]

\[
C = \sum_{i=1}^{n} x_i y_i
\]

where \( x_i \) and \( y_i \) deviation derived from \( xy \) coordinates from mean center.

Standard deviation for \( x \) and \( y \)-axis are as follows:

\[
\sigma_x = \sqrt{\frac{\sum_{i=1}^{n} (x_i \cos \theta - (\bar{y} \sin \theta)^2}{n}}
\]

\[
\sigma_y = \sqrt{\frac{\sum_{i=1}^{n} (x_i \sin \theta - (\bar{y} \cos \theta)^2}{n}}
\]

Standard deviational ellipse creates new feature class that includes centered ellipse polygon on mean center for all characteristics (or for cases when values were defined) [39]. Attribute values from ellipse polygon included standard distance (long and short axis) and ellipse orientation. Orientation represents long axis rotation measured clockwise from noon. In this study, SDE was conducted to analyze quantitatively on the orientation of dengue fever cases for identifying to what extent that those dengue outbreak might have spread. Other than that, it also identified the number of standard deviation (1, 2, or 3). When features have normal spatial distribution (means density in center and becoming less dense towards the edge), one standard deviation makes up to 68% from all centroid input features. Two standard deviations will make up about 95% from all features, and three standard deviations will cover about 99% of all centroid features.

Spatial distribution pattern mapping was conducted by pattern analysis using average nearest neighbor (ANN) analysis that measures the average distance from each point in the study area to its nearest point. The average distance is compared to the expected average distance. In doing so, an ANN ratio is created, which, in simple terms, is the ratio of observed/expected. If the ratio is less than 1, we can say that the data exhibits a clustered pattern, whereas a value greater than 1 indicates a dispersed pattern in our data [39]. In this study, this distribution pattern mapping of dengue fever cases was observed, and the shape of the ANN curve as a function of neighbor order can provide insight into the spatial arrangement of points relative to one another whether it is clustered, random, or dispersed. ArcMap 10.5 software was used to perform these statistical analyses. The application formula used in ArcMap 10.5 uses mathematical equations in the software to develop every analysis with different functions.
3. Results

Statistical analysis is important to evaluate the spatial distribution of dengue and to understand the factors that may be related to the occurrence of the cases. Spatial distribution pattern is used to identify geographic patterns. It is important to understand how geographic phenomena behave using formulation spatial mean center, directional distribution, and standard distance. It was found in the study area that 1401 dengue cases were reported in 2008. Dengue cases decline about 28% in 2009 from previous year, when 1056 patients were hospitalized with dengue symptoms.

Spatial mean center is the analysis where phenomena leans to be centered. It is the average of x and y that sync all features in the study area. It is useful to detect changes in distribution or to compare types of feature distribution. Mean center point create new point feature class where every feature represent mean center. X and Y means show the value of center, case and dimension field means included as the product of features. Figure 2 shows the location of every dengue cases reported in the years 2008 and 2009 with their mean centers, respectively. Spatial mean center for both dengue cases overlap with each other and are located on the west of this study site. Mean center for dengue cases in the year 2008 is located at x-coordinate 441919.907209 and y-coordinate 299749.74661. Meanwhile, the mean center for dengue cases for the year 2009 is located at x-coordinate 441789.096591 and y-coordinate 299165.549242. This is parallel to the distribution of the primary dengue cases reported in both years.

Figure 2. Spatial mean centers for dengue cases in the years 2008 and 2009.
Figure 3 shows the directional distribution for dengue cases reported. The directions for distribution of dengue cases in both years shows slight differences. Ellipse polygon of the year 2008 is bigger than the year 2009 derived from analysis, that covers 68% of the features with rotation 93.05° and 94.34° on the long axis, respectively. Both ellipses situated at the east of the study site. Hence, this shows that dengue fever outbreak distribution in both years have the same directions, which are centered at the west of the Seremban district.

![Legend](image)

**Legend**
- †: Dengue cases 2008
- *: Dengue cases 2009
- Blue: Directional distribution 2008
- Orange: Directional distribution 2009
- White: Study site

Figure 4 shows standard distant of dengue cases in the Seremban district. Standard distance for dengue cases reported in the year 2008 is 22,085.82 m. Meanwhile, dengue cases reported in 2009 show standard distance of 20,318.35 m. This analysis shows that both years develop a minor difference of dengue transmission with smaller dispersal of standard distance in the year 2009 compared to 2008.
4. Discussion

Dengue Cases Distribution Pattern Analysis

Figure 5 shows the spatial distribution pattern map. This pattern mapping was conducted by pattern analysis using average nearest neighbor method. Figure 5 represent nearest distance analysis method. The z-score is -55.4447 and there is a less than 1% likelihood. Nearest neighbor ratio is 0.22 (Figure 6). According to theory, if the index (nearest neighbor ratio) is less than 1, the pattern displayed is clustered. Meanwhile, if the index is bigger than 1, trend direction is toward the disperse distribution [39]. In this case, the pattern of this analysis is clustered. This phenomenon indicates that the hotspot for dengue cases to occur in both the years 2008 and 2009 is located at the west of this study area.
Pattern distribution analysis was performed to determine and understand whether there are any dominant patterns in dengue case distribution. Details of pattern distribution were filtered by conducting pattern analysis. Clustered pattern indicates there were a factor causing dengue epidemiology in the study area (Figure 7). High density of population located at the west of Seremban district could be the reason of high dengue incidence between these two years [46]. Thus, this would help the authority to enhance prevention and control over dengue spread at the specific location by planning of care provisions in areas of greater vulnerability in order to reduce the severity of the disease spread.
Table 1 summarized nearest distance method analysis. Z-score reading shows that pattern less than 1% likelihood are significant. According to theory, when the index (average nearest neighbor ratio) is less than 1, the pattern shown is clustered. Meanwhile, if the index is bigger than 1, pattern direction is dispersed. Thus, distribution pattern of dengue cases reported in this study is clustered. Z-score reading is less than 1% indicates distribution pattern of dengue cases is significant.

Table 1. Spatial pattern for dengue cases in the Seremban district.

<table>
<thead>
<tr>
<th>Year</th>
<th>Dengue Cases Reported</th>
<th>Nearest Neighbor Ratio</th>
<th>z-Score</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>0.225698</td>
<td>−55.444729</td>
<td>0.000000</td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>0.264112</td>
<td>−45.748278</td>
<td>0.000000</td>
<td></td>
</tr>
</tbody>
</table>

Every dengue case reported displayed clustered pattern. However, dengue cases in the year 2008 were the more clustered compared to the year 2009. This shows that there is a factor affecting the dengue transmission during that year that causes higher dengue outbreak compared to the next year. Rapid urban growth with extensive land use change processes through development of new township and infrastructure, along with the needs of population growth, caused breeding ecosystem for the dengue vector, especially Aedes mosquitoes [48]. The urban growth provides good spaces and places for breeding, while higher population of humans and animals provides good food sources for the Aedes mosquitoes [7,16]. Thus, this could be the key factor, as the cluster areas are located within the city center of the Seremban district.

5. Conclusions

Spatial statistical distribution and pattern analyses play important roles in understanding dengue cases epidemiology in the study area. This is because distribution and spatial pattern are derived from certain processes. These three spatial statistical analyses (spatial mean center, standard distance, and directional distribution) show that the dengue cases in the years 2008 and 2009 are both clustered at the west of the study site. This analysis found that there are differences between the direction of dengue cases distribution, even it is minimal. Pattern distribution analysis was conducted to enhance the finding results, and it was found that the clustered pattern is dominant in this study area with an average nearest neighbor ratio of less than 1. This proved that there is a factor affecting the rise in dengue cases reported in the year 2008 compared to 2009. Above all, this shows that spatial modeling should take into account of dengue cases reported in other years for further studies.

We observed that the population in this region frequently migrates and rapidly urbanizes, which may be contributing factors in the emergence of centered dengue epidemic at the west of the Seremban district [48]. The future urbanization, globalization, environment, human behavior, and lack of control
of vectors have contributed to the global distribution of dengue. Furthermore, if unprecedented population growth and continued urbanization continue as projected, there will be gradual increases in the severity, frequency, geographical distribution, and magnitude of dengue epidemics in the future [13].

Dengue fever outbreaks in Malaysia reached 80,615 cases reported in the year 2018, with a death toll of 147. The Seremban district recorded a total of 1864 dengue fever cases with six deaths caused by dengue virus, which is 2.31% from cumulative of cases reported across the country [42]. Although this number contributed to the rise of dengue transmission in Malaysia on a small scale, it still requires monitoring to ensure the quality of public health in the urbanized Seremban district and preventive measure should be applied to reduce the number.

The lack of numbers on data sets per year for dengue fever cases reported was been found to create a gap of knowledge in this study. Further studies conducted should include a range of at least five years of data sets for dengue outbreak in the country to analyze the distribution pattern and direction of the transmission. Studies on factors affecting dengue outbreak in the Seremban district would also help as a tool in mitigating dengue fever.

**Author Contributions:** N.A.M. conceived and designed the research. N.M.N. did most of the analysis, contributed extensively to data processing and preparation. A.F.M. has given many suggestions for improving the article for language editing. All authors read and approved the final manuscript.

**Funding:** This project was funded by the project under the University Kebangsaan Malaysia research funds with grant no. (GGPM-2018-030).

**Acknowledgments:** This study honor the late researcher Shaharudin Idrus from Institute for Environment and Development (LESTARI), Universiti Kebangsaan Malaysia for his contribution towards the knowledge of this study.

**Conflicts of Interest:** The authors declare no conflicts of interest.

**References**


34. Oliveira, F.L.P.; Cançado, A.L.F.; de Souza, G.; Moreira, G.J.P.; Kulldorff, M. Border analysis for spatial clusters. *Int. J. Health Geogr.* **2018**, *17*, 5. [CrossRef]
