

Free and open GIS source software for spatial epidemiology and geospatial health in Malaysia: a comparative analysis of the software usability

Wan Hairi Wan Mahsin¹, Abdul Rauf Abdul Rasam^{1, 2*}, Noraain Mohamed Saraf¹ and Nafisah Khalid¹

Centre of Studies for Surveying Science and Geomatics, Faculty of Architecture, Planning and Surveying Universiti Teknologi MARA, Shah AlamSelangor, Malaysia¹

Environmental and Social Health (ESH) Group, Health and Wellbeing Excellence Entities, Universiti Teknologi MARA, Shah AlamSelangor, Malaysia²

Received: 10-November-2020; Revised: 20-April-2021; Accepted: 25-April-2021

©2021 Wan Hairi Wan Mahsin et al. This is an open access article distributed under the Creative Commons Attribution (CC BY) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

The use of free, open Geographical Information System (GIS) source software in Malaysia has evolved rapidly in the recent years, from a generating map digitally into a set of methods with various applications. Free open GIS source software has been utilised in health management fields due to its cost factor and user-friendly feature. Therefore, a comparison of selected free, open GIS source software focusing on the distribution of tuberculosis (TB) cases is presented in this paper to examine their functions and technical usability for local spatial epidemiological applications of Sarawak, Malaysia. Three health-based free open GIS source software are selected in this comparative study i.e., gvSIG, GeoDa and MapWindows GIS. These software's are possible alternatives to the licenced software (ArcGIS). The functions and output of each software are analysed using mapping, spatial analysis, and online system visualization. The free GIS open-source software might not be able to compete with licenced software in terms of advanced analysis. However, for basic mapping and spatial analysis functions such as heatmap mapping and hotspot detection, free GIS open-source software is a sufficient choice, especially for non-genetic or non-computer-based users.

Keywords

GIS free open source, GvSIG, GeoDa, MapWindows GIS, Geospatial health, Tuberculosis disease.

1. Introduction

Tuberculosis (TB) disease is one of the oldest diseases that exist on earth and remain one of the top 10 infectious killers in the world. Although TB can be cured and prevented, it is very important for the public to know and take precautions to stop the spreading of the disease since TB is ranked as the number two leading cause of death among the global population [1]. TB is caused by bacteria known as Mycobacterium tuberculosis that usually affect the lung. TB mortality is related to factors such as being infected with human immunodeficiency virus (HIV), multi vaccine resistance and limited access to health service [1]. In 2018, World Health Organisation (WHO) estimated almost 10 million TB cases around the world and 1.3 million death cases caused by TB [2].

Malaysia reported a total of 25,837 TB cases in 2018 which showed a 1.3% decline compared to 26,168 TB cases reported in 2017. In 2018, the range of death caused by TB disease worldwide was 6.6 cases per every 100,000 people, which showed an increase of 0.1 cases compared to 6.5 cases per every 100,000 people reported in 2017 [3].

Due to the rising need in geographical information system (GIS) application and the advancement in computing software technology over the last decade. Several scientists have been working hard to develop free software's and tools which are known as open-source GIS applications (Open GISc). Many of these open-source software's are helpful, especially for university students, researcher and those who cannot afford to buy commercial and expensive software such as ArcGIS software. The licence of commercial software is expensive and there are also added to tool functions which occurs more cost [4]. Said Open GISc could enhance the transparency of research and

*Author for correspondence

scientific practice utilizing GIS. Map overlay and statistical system (MOSS) and geographic resources, analysis support system (GRASS) are among first open-source created by the U.S. Department to help them in the environmental planning and land management [5].

GIS has been utilised in TB datasets management and analysis [6–8]. Nowadays, there are numerous free open GIS source software that can be freely downloaded from the Internet including the software used in this study i.e., gvSIG, GeoDa and MapWindows GIS. Although these free platforms provide functions essential to deliver health and disease data or information to related users [9–11], an empirical comparative analysis need to conduct to quantify the software usability for the local health applications mainly for disease mapping, spatial analysis, and online system visualization. The free platform can help the local health department to monitor and identify hot spot areas that usually record high rates of TB cases.

This study attempts to explore the potentials of free and open GIS platform for local health organizations such as the Ministry of Health Malaysia and Sarawak Health Department to monitor and manage the disease datasets because of limitation of funding to purchase costly commercial GIS software. Moreover, to operate such software's, local health department officers may require special technical skills and training and this may incur additional cost. It is important to note that developing an interoperable GIS-based health information system of public health information technology is essential for a health surveillance programme [10–13]. In particular, the objectives of the study are; i) to examine the capabilities of the selected Free open GIS source software, ii) to analyse the spatial pattern of TB distribution in Sarawak, and iii) to evaluate the capabilities or usability of the selected open GIS source software. Considering the time and materials available, this investigative study is only carried out with the three selected software and basic GIS operations, including gvSIG, GeoDa and MapWindows GIS.

2.Literature review

2.1GIS and open-source software

GIS is a software that is used to create, manage, analyze and visualize geographic data, i.e., data with a reference to a place on earth. A GIS helps users to display, grasp, query, analyse, and visualise data in a number of ways, including maps, globes, papers, and charts, to expose relationships, patterns, and trends. By

looking at data in a manner that is readily interpreted and communicated, a GIS may also support users in answering questions and solving problems. GIS software is widely used in the assessment of locations for new shops, the management of power and gas lines, the development of charts, and the interpretation of results. For this multitude of applications different types of GIS functions are required and different categories of GIS software exist, which provide a particular set of functions needed to fulfil certain data management tasks.

Open-source software (OSS) has widely been proposed as the right solution to the unsustainability of most IT systems in developing countries. OSS is thought to support the production of local information technology that can gain more competitive and efficient development outcomes in poor contexts such as the public sector in Africa [14] due to its non-proprietary existence. In general, OSS will help the growth of local IT skills, allowing for greater local localization diversity while also maintaining low costs and interoperability. Open-source GIS commonly used for mapping and analysis the data [15,16]. It is also a platform that provides a freely available software and source code which can be downloaded by any user through online network or browser.

The benefit of using open-source GIS is it has no licensing fees with a provided highly customized applications and tools based on client's needs. This an open-source application can freely access and modifies the source code for, such as system for automated geoscientific analyses (SAGA) GIS, geographic resources, analysis support system (GRASS) GIS and Quantum GIS (QGIS) [11] that are usually used by a group of volunteer programmer. In the field of applied GIS, three parallel and mutually reinforcing forces of progress are open access data, open-source applications, and cloud computing [17]. Governments and organisations are gradually supporting open geospatial research, which includes open geospatial data and open-source geospatial software [18]. In terms of potential innovations, the global geospatial community predicts that open-source geospatial applications will continue to rise in popularity [19, 20].

ArcGIS, GRASS and QGIS are examples of well-known GIS software used in Malaysia that are very familiar among the researcher and other user that involved in GIS projects. In the context of health and disease application, gvSIG, GeoDa and MapWindows GIS has also potential to be implemented in the

country. MapWindows GIS is statistical software for epidemiology developed by the Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia. The software allows for the development of automated surveys, data entry, and analysis [14, 15].

GeoDa is a free software package that can conduct spatial data analysis, geovisualization, spatial autocorrelation and spatial modeling. Open GeoDa is one of the open-source version of legacy GeoDa. It has sophisticated spatial analysis, multivariate exploratory data analysis, and global and local spatial autocorrelation capabilities. It also has simple linear regression capabilities [1]. gvSIG is a Java-based open-source GIS, that is suitable for data visualizations and simple analysis. It is a free extension from the WHO and used for service availability mapping [16]. In Malaysia, [6] explored the capabilities of the Free and OSS for Geospatial (FOSS4G) disease mapping and analysing of TB in Petaling, specifically QGIS, SAGA GIS. It reveals that this GIS software have basic functions of disease mapping and analysis.

2.2 GIS free and open-source software (FOSS) for health application

OSS has widely been applied as the right solution to the unsustainability of most IT systems in developing countries. The social conditions need to be shaped in the countries in related to technical facilities, organisational planning and development process. Due to its non-proprietary nature, OSS is believed to provide the production of the local information technology that can benefit more sustainable and effective development results in the framework of poor contexts such as the public sector in Africa [14]. GIS is becoming a vital tool in healthcare applications covering database management, planning, risk assessment, service area mapping, location identifications and others [12, 13].

GIS OSS can raise the functions of disease mapping and healthcare analysis towards more advanced operations, especially for spatial analysing and mapping distribution of disease cases [6,10,11,16]. [11, 16] utilized QGIS to process and produce the output map of health datasets. Also included the analysis of the distribution pattern of drug addiction by creating a hotspot map of drug addiction in the study area [11]. In Kenya, EPI Info used to explore how the implementation and usage of technologies has

influenced the frameworks of the health management information systems, as well as to create an OSS business model research process based on all of the results [14].

3. Methods

3.1 Planning and Study area

Figure 1 shows the overall of research methodology of the study, start of project planning to result and conclusion. The state of Sarawak in Malaysia is selected as the study area. It is one of the largest states, among 13 other states with the total area 124,451 km² which is almost the same size as West Malaysia. Sarawak is located in the northwest of the Borneo Island and it consists of 4 capital cities. The capital city is Kuching, which is located at the west of the state, and it is the economic centre of the state.

3.2 Software

The digitized base map of the study area is created in ArcMap shapefile 10.6 by ArcGIS desktop application. Then, the map is saved as a shapefile so that it can be used in the other software's. Next, the table of TB data cases is inserted in the attribute table in the ArcMap 10.6 where the point is produced based on the coordinate in the table of data. The shapefile created using ArcMap which contain Sarawak base map and TB data are inserted and processed in the EPI gvSIG software, GeoDa software and MapWindows. The data are then analysed using a tool available in the GeoDa software, MapWindows GIS, JumpGIS and other softwares.

3.3 Data acquisition

Two base maps are digitized in this study and then converted into the shapefile format. The Malaysian base map is used as an indicator or keymap for the main map, while the Sarawak state base map is used as the platform for mapping the TB distribution in the state. The main data set used in this research is the TB data which were originally officially obtained from the Ministry of Health, Malaysia. The disease data are presented in a table and then they are converted into a spatial point shapefile. Each point data consists of coordinate values and the number of cases in the nearby areas as shown in *Figure 2*.

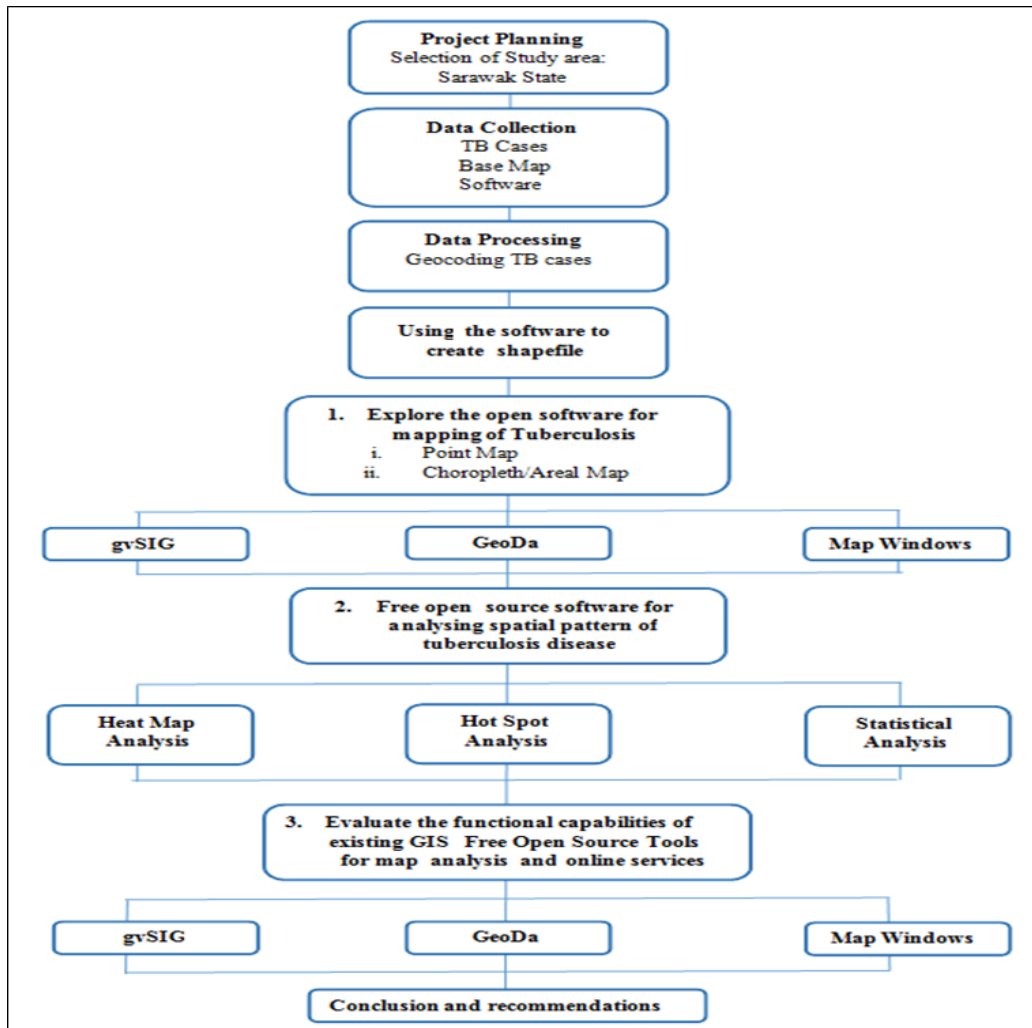


Figure 1 Research methodology

3.4 Data processing

The geocoding process is performed to convert the addresses (MS Excel format) of the patients into coordinate through ArcGIS. ArcGIS can store the address data in the attribute table, but cannot display the data without spatial reference. Geocoding process uses Google Drive software to determine the coordinate of TB patient. After the geocoding process, the disease mapping is conducted in shapefile format using WGS 1984 and then overlaid with Sarawak boundary shapefile and the disease cases.

For the data analysis, the first technique used is proximity analysis. Proximity analysis is an analytical technique used to determine the relationship between a selected point and its neighbours. Two types of proximity analysis used in this study are buffer map and choropleth map. These two maps have its function and display different result analyses. The second technique used is spatial analysis. Spatial analysis is a type of geographical analysis which seeks to explain patterns of human behaviour and its spatial expression in terms of mathematics and geometry, i.e., locational analysis. Three types of spatial analyses are employed in this task, including hot spot analysis, kernel density and K analysis.

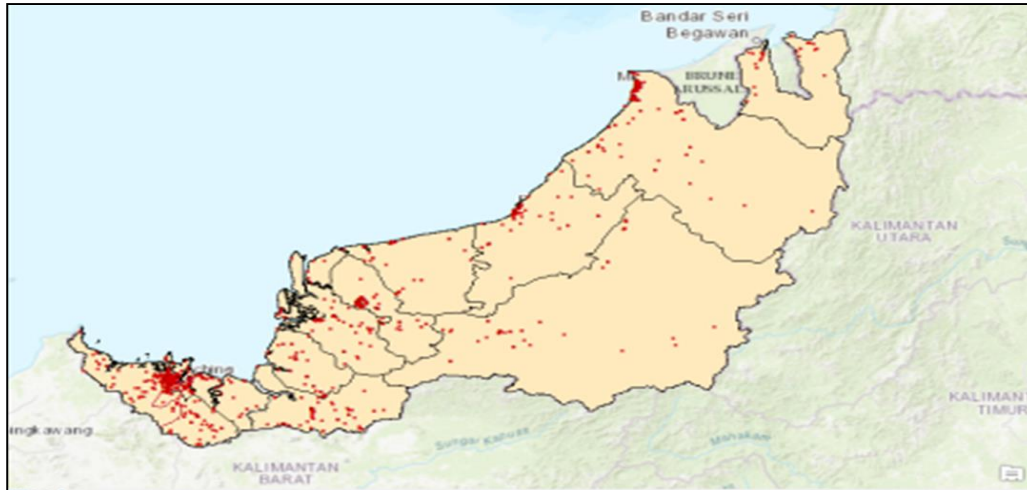


Figure 2 Sarawak base map and tuberculosis cases in ArcGIS platform





4. Results and discussion






4.1 The selection of existing free open-source software

There are 15 official GIS open-source software introduced to the public and these software's are explored in this study to analyse the GIS data and basic operations. Out of the 15 GIS open-source software, 9

software are further selected for this study namely ArcGIS, Epi Info, GeoDa, OpenJump, Saga GIS, MapWindows, GeoFIS, Whitebox GIS and gvSIG. Each selected software has its advantages and disadvantages depending on the data and analysis used. *Table 1* shows a list of the selected software examined in this study.

Table 1 List of the selected open-source software

No.	Software	Comment(s)	Download Site
1.	 ArcMap 10.6	Generally, ArcGIS is a very powerful tool with a stipulated set of tools but it is costly.	Not provided (Proprietary software)
2.		It can be run on a smaller amount of computer RAM since it is a small-sized software, but there are only limited features or functions available in the software.	https://www.cdc.gov/epiinfo/support/downloads.html
3.		It can be used to analyse spatial data, but it does not provide layout functionality or features.	https://geodacenter.github.io/download.html
4.		It can be used to customize the appearance of your information, but sometimes the changes made are often lost and	http://www.download82.com/download/windows/openjump/

No.	Software	Comment(s)	Download Site
5.	 Saga GIS	<p>there is a need for backups.</p> <p>It can accommodate large volumes of vectors and information, but every element added, appear separately. Thus, making it troublesome at times.</p>	https://sourceforge.net/projects/saga-gis/
6.	 MapWindows	<p>It can support data created from other software and provides a layout function similar to ArcMap but the layout is automatic customised and cannot be set manually.</p>	https://mapwindow-gis.soft112.com/
7.	 GeoFIS	<p>It can be used to display the data on screen by using coordinates, but unfortunately it cannot be used to display in layout form.</p>	https://www.geofis.org/en/install/install-on-windows/
8.	 Whitebox GIS	<p>It has vector and raster manipulation tools, but the data management is complicated.</p>	https://jblindsay.github.io/ghrg/Whitebox/download.shtml
9.	 gvSIG	<p>It works with different vector formats, but limited options are available to change the layout of the fields in the table.</p>	http://www.gvsig.com/en/products/gvsig-desktop/downloads

4.1.1 Comparison of the functions and usability of the selected open GIS source software

The functions and usability of the selected GIS OSS are determined in terms of basic functions, disease mapping, analysis and information system design. The functions and usability of these software are explored as well as downloaded software evaluation. There are a few basic criteria or functions that have been examined for a fair comparison between the selected GIS open-source software, including shapefile, digitizing shapefile insert, layout display, legend

feature and analysis function. Some of the selected software is slightly more complicated compared to the others. For example, for the gvSIG software, the user needs to change the setting of the language before the software can be used. This is because the original version of gvSIG is either set in German or Spanish language depending on the site the user uses to download it. *Table 2* shows the comparison in terms of functions and usability of the nine out of 15- software selected in this study.

Table 2 Qualitative checklist features and functions of the selected OSS for health application

Software	Digitize spatial	Spatial data insert	Layout	Legend features	Analysis	Information
ArcGIS	/	/	/	/	/	Since ArcGIS is a licence software, then it has all the basic functions for GIS study needs.
Epi Info	-	/	/	-	/	Epi Info software only has a function to import data but do not have a function to create data like shapefile.
GeoDa	/	/	-	-	/	GeoDa software can be used to create a shapefile, but it does not have a layout and layout feature functions such as legend and scale.
OpenJump	-	/	/	/	/	OpenJump has all basic functions, but it does not have a create data function.
Saga GIS	/	/	-	-	/The The	Saga GIS software have all needed for data, but for display it in layout, it need another software to do it.
Map Windows	/	/	/	/	/	MapWindows has all basic functions including layout and layout features.
GeoFIS	-	/	-	-	/	In GeoFIS software, it is more likely used as a display data on screen.
Whitebox GIS	-	/	-	-	/	Whitebox GIS is the same as SAGA GIS, and does not provide a layout function.
gvSIG	/	/	-	-	/	For gvSIG, it can create a shapefile data by using a basemap it provides but it does not have a layout function.

4.2 General distribution of tuberculosis cases (point map)

Sarawak is divided into 12 districts, namely Kuching, Samarahan, Serian, Sri Aman, Betong, Kapit, Sibul, Mukah, Bintulu, Miri and Limbang. From the distribution of TB cases, the number of cases is divided based on the divisions as shown in *Table 3*. The TB distribution in Sarawak shows that higher cases are concentrated in Kuching, Miri and Sibul than other divisions.

4.2.1 Disease mapping of tuberculosis pattern

- Mapping of TB cases in ArcGIS and open-source software

In ArcGIS, the map is generated by using layout view and inserting the element of maps such as north arrow, legend, and scale bar (*Figure 3*). The map of the TB pattern distribution is also created using the selected

GIS open sources (*Figure 4*, *Figure 5* and *Figure 6*). This map point is created after the processing is complete to show the distribution pattern of TB cases in 2016 in the state. In MapWindows, the map is generated by using the new print setup and inserting the elements of maps such as the north arrow, legend, and scale bar (*Figure 4*).

In gvSIG, TB map is generated by using project manager (MAP) and inserting the element of maps such as the north arrow, legend, and scale bar as shown in *Figure 5*. A similar map of TB cases is also created by using GeoDa (*Figure 6*). The GeoDa software can only import one data or shapefile per processing. The data is displayed without a legend, map scale north arrow and others as provided by ArcGIS software.

Table 3 Number of tuberculosis cases in Sarawak in 2016

No.	Division	Area perimeter (Km ²)	Number of cases (Person)
1.	Kuching	4,560	367
2.	Samarahan	4,967	74
3.	Serian	2039.9	86
4.	Sri Aman	5,467	84
5.	Betong	4,181	52
6.	Sarikei	4,332	76

7.	Kapit	38,934	55
8.	Sibu	4129.5	151
9.	Mukah	6,998	55
10.	Bintulu	12,166	92
11.	Miri	26,777	203
12.	Limbang	7,788	29
Total		124,451	1325

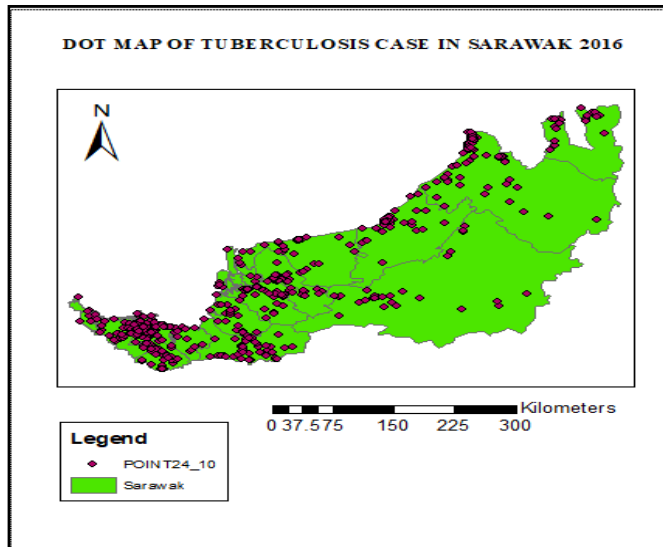


Figure 3 TB points map using ArcGIS

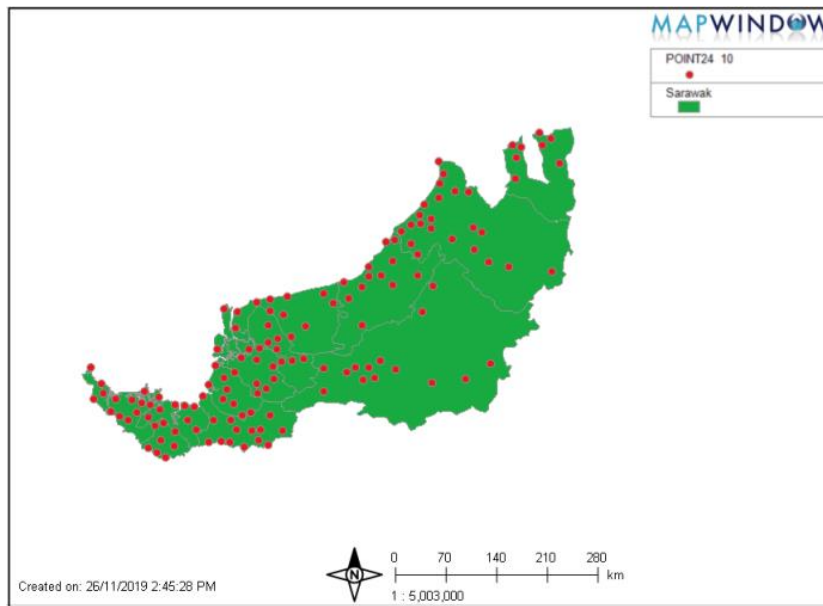


Figure 4 TB points map using Map Window GIS

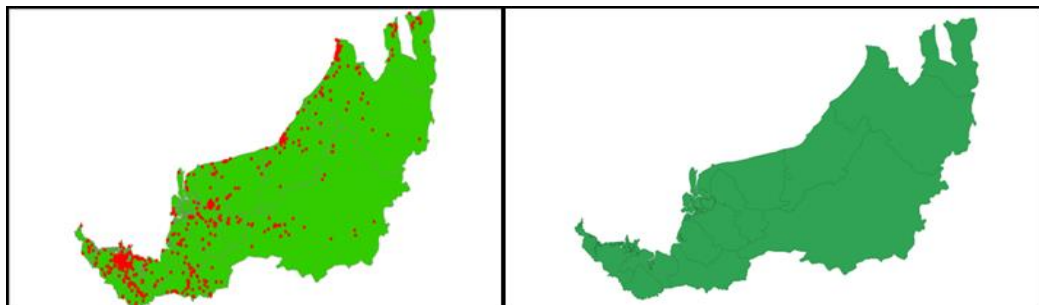


Figure 5 TB points map using gvSIG

Figure 6 TB points map using GeoDa

Table 4 The density of tuberculosis cases by division in Sarawak

No.	Division	Area perimeter km ² (v)	Number of cases/patients (m)	Density $P = \frac{m}{v}$
1.	Kuching	4,560	367	0.0805
2.	Samarahan	4,967	74	0.0149
3.	Serian	2039.9	86	0.0422
4.	Sri Aman	5,467	84	0.0154
5.	Betong	4,181	52	0.0124
6.	Sarikei	4,332	76	0.0175
7.	Kapit	38,934	55	0.0014
8.	Sibu	4129.5	151	0.0366
9.	Mukah	6,998	55	0.0079
10.	Bintulu	12,166	92	0.0076
11.	Miri	26,777	203	0.0008
12.	Limbang	7,788	29	0.0037
Total		124,451	1325	0.2409

Proximity analysis, mapping of tuberculosis pattern

TB distribution pattern (choropleth mapped) used to display the TB pattern distribution in the divisions of Sarawak recorded in a year. This mapping shows two types of choropleth products i.e., choropleth map based on the divisions in the state and a choropleth map based on a density of TB in each division. Table 4 indicates the data of area perimeter, the number of cases and the density of the disease for each division.

Comparative results of choropleth map in ArcGIS, gvSIG, GeoDa and MapWindow GIS

Table 5 illustrates the comparison of the results of four different GIS software, namely ArcGIS, gvSIG, GeoDa and MapWindows GIS using TB cases and TB density in Sarawak. Most of the open sources have similar functions with ArcGIS. For example, MapWindows GIS has the same additional features such as legend, north arrow, scale bar and most importantly, the map layout. This is then followed by GeoDa which only provides the legend feature. Finally, the gvSIG software used a density factor, but it has a limited minimum value to display the final output of the division colour map.

Comparative results of buffering map in ArcGIS, gvSIG, and MapWindow GIS

Table 6 shows the comparison of the results of buffer point map using 3 different GIS software i.e., ArcGIS, gvSIG, and MapWindows GIS. The distance buffer used in this proximity analysis is 5 km. The result reveals that MapWindows GIS has the most familiar function of the ArcGIS software and followed by gvSIG. However, both software (MapWindows GIS and gvSIG) cannot display, or create multiple buffers unlike ArcGIS.

4.3 Spatial pattern analysis of tuberculosis cases

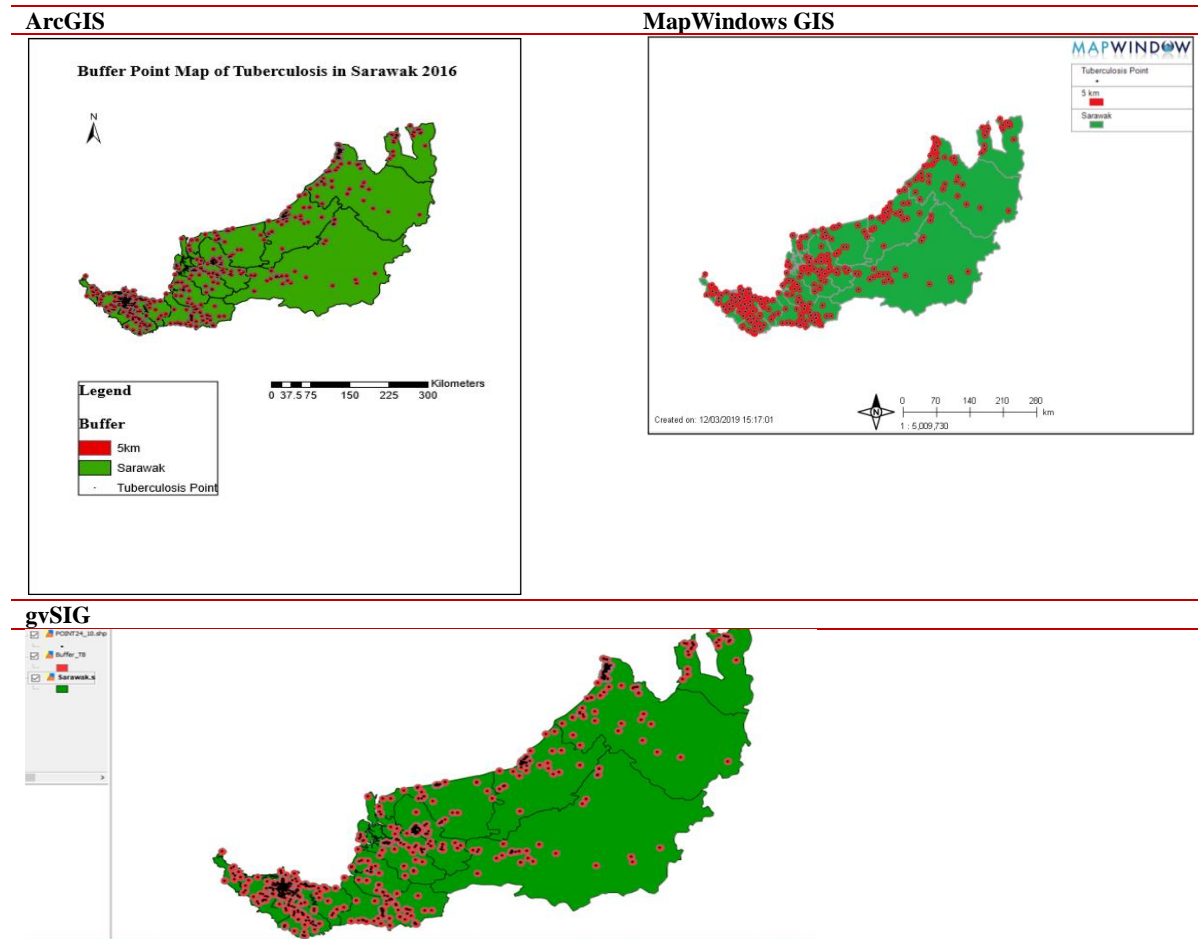
4.3.1 ArcGIS Software

In ArcGIS software, Spatial Statistic toolbox used for the analysis of spatial distribution, pattern, process and relationship with statistical tools. These toolboxes used as a clustering method and technique to examine the cases based on the case locations and the spread distribution.

Table 5 Comparison of the results of choropleth map based on the TB cases and density

Factor	Tuberculosis case	Tuberculosis density
Software		
ArcGIS		
gvSIG		
GeoDa		
MapWindow GIS		

Table 6 Comparison of the results for buffering point map of TB



Hot map analysis, kernel density

Hot map analysis identifies statistically significant spatial clusters of high values hot spots and low values cold spots. It creates a new output feature class with a z-score, p-value, and confidence level bins for each feature in the Input Feature Class. The hot spots of TB point data are also displayed and most of the higher number of hot spot area is identified in Kuching with 99% confidence level as stated in the legend (Figure 7). Other areas seem to be normal and not significant distribution. In ArcGIS software, the Kernel Density used to calculate the density of linear features in the neighbourhood of each output raster cell. Its smooth curved surface is fitted over each line. Its value is the greatest on the line and diminishes as you move away from the line, reaching zero at the search radius for the line. Figure 8 displays the Kernel Density of TB in Sarawak. It can see that the highest reading can be seen in a few areas in Sarawak whereby the most notable reading is in Kuching, followed by Miri and a few red spots in Sibu with the mean reading between 4 000 –

4 500m, while the mean for other areas is normal which is between 0 – 500m.

GeoDa Software

In GeoDa software, Spatial Map Explore and Cluster is used for analysing spatial distribution, pattern, process, and relationship with statistical tools. These toolboxes are used as clustering method and technique to examine the cases based on the locations of TB cases, density and its distribution and spread.

K-Means Analysis, K-median Analysis and K-Medoids Analysis

In GeoDa software, K-means clustering is one of the simplest and popular unsupervised machine learning algorithms. The K-means algorithm identifies k number of centres, and then allocates every data point to the nearest cluster, while keeping the centroids as small as possible. K-means analysis map is produced by using GeoDa software by clicking the cluster button in the menu bar and select K-means, then select

the interval. For this analysis 5 intervals are used and the result is as shown in *Figure 9*. From the result, the pattern of TB cases is categorized as clustered with K-means, the total within-cluster sum of the square is 0.0634681, and between-cluster sum of the square is 10.9365. The ratio of the between-cluster sum of the

square is 0.99423. The ratio of the between-cluster sum of the square is the standard deviation that shows the probability that the pattern would be an outcome of random chance.

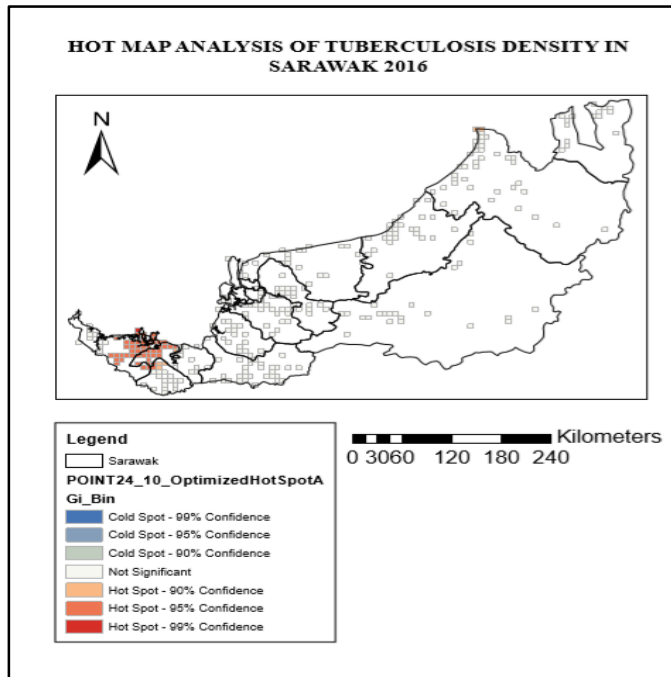


Figure 7 Hot map analysis of TB density in ArcGIS

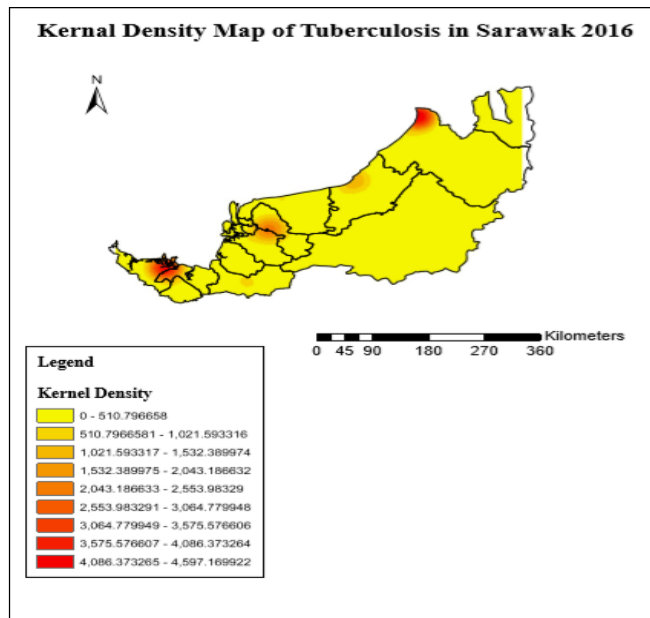


Figure 8 Kernel density map of TB density in ArcGIS

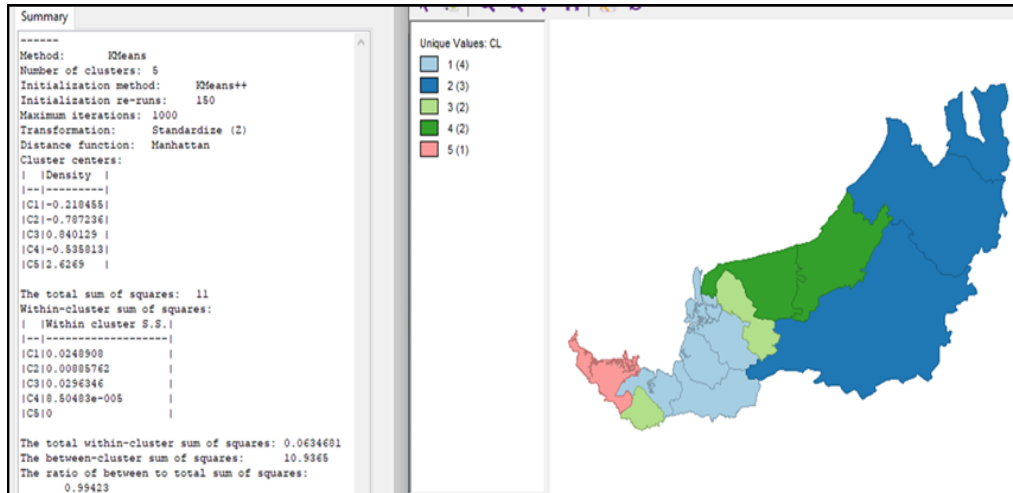


Figure 9 K Mean analysis of tuberculosis density in GeoDa

K-medians clustering is a cluster analysis algorithm. It is a variation of K-means clustering where instead of calculating the mean for each cluster to determine its centroid, one instead calculates the median. In this step, all objects are assigned to their nearest median. K-median analysis map can be produced by using GeoDa software by clicking the cluster button in the menu bar and then selecting the K-median. After that, select the interval, 5 intervals in this case and the result is as shown in *Figure 10*. The pattern of TB cases is categorized as clustered with a K-median; a total within-cluster sum of a square is 0.00013903, and between-cluster sum of a square is 0.00017557. The ratio of the between-cluster sum of a square is -0.992143. The ratio of the between-cluster sum of the square is the standard deviation that shows the probability that the pattern would be an outcome of random chance.

In the K-medoids or partitioning around medoids algorithm is a clustering algorithm reminiscent of the k-means algorithm. It is more robust to noise and outliers as compared to k-means because it minimizes a sum of pairwise dissimilarities instead of a sum of squared Euclidean distances. K-median analysis map is produced by using GeoDa software by clicking the cluster button in the menu bar and selecting K-median. After that, choose the interval, in this case, 5 intervals and the result is as shown in *Figure 11*. The *Table 7* shows the result of spatial autocorrelation. The total within-cluster sum of the square is 5.13866, and between-cluster sum of the square is 5.06134. The ratio of the between-cluster sum of the square is 0.532094. The ratio of the between-cluster sum of the square is the standard deviation that shows the probability that the pattern would be an outcome of random chance.

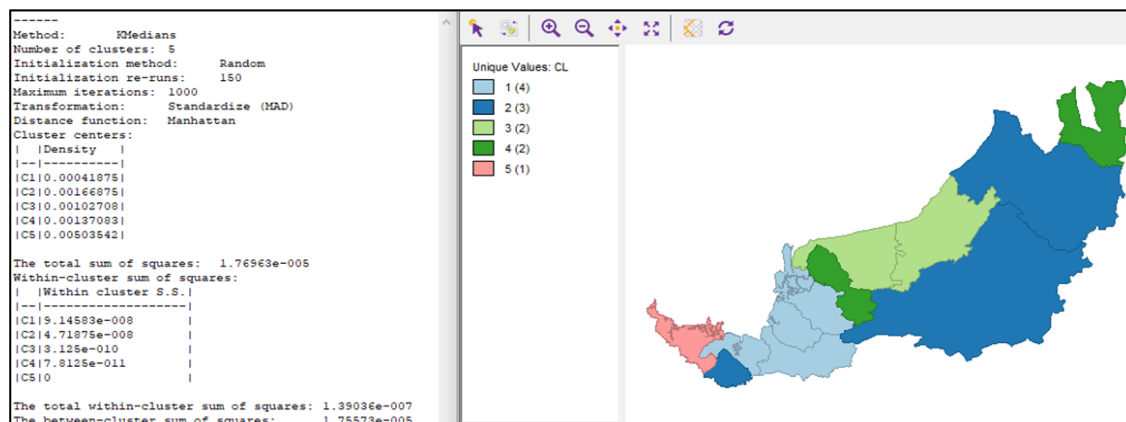


Figure 10 K-median analysis of tuberculosis density in GeoDa

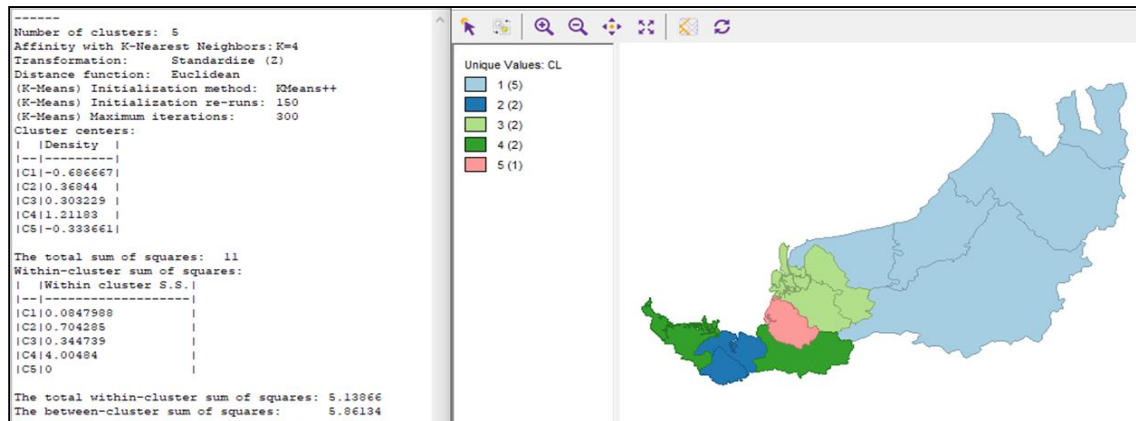


Figure 11 K-Medoid analysis of tuberculosis density in GeoDa

4.4 Comparative results of online system geovisualization

Table 7 shows the comparison of the result geovisualization using two different platforms i.e., ArcGIS Online, and MangoMap. There are three (3) techniques of visualization used, namely buffer point map, a heat map and dasymmetric map. From the table, it can see that the buffer point map technique can be created using MangoMap with the distance of 5 km within the point. For the second technique which is a heat map visualization, it shows that only ArcGIS online is capable of creating a heat map by using density data, while MangoMap is only capable of creating an interpolation map by using point. Finally, for dasymmetric map, it shows that ArcGIS online has no problem in visualizing the dasymmetric map while ManggoMap can also display the boundary without data extraction capability.

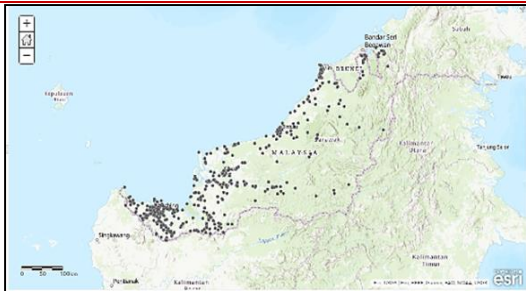
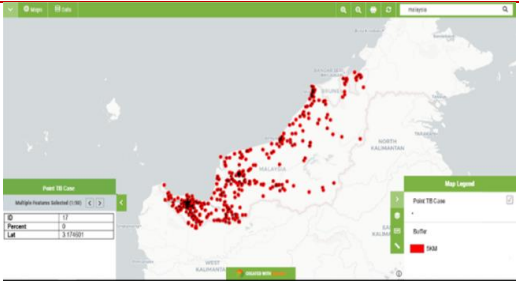
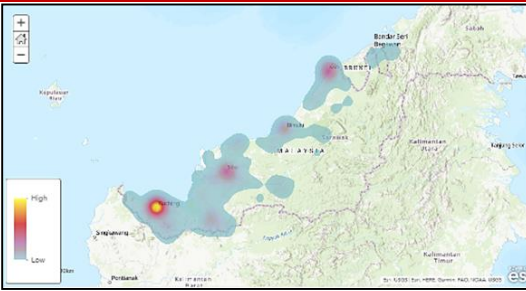
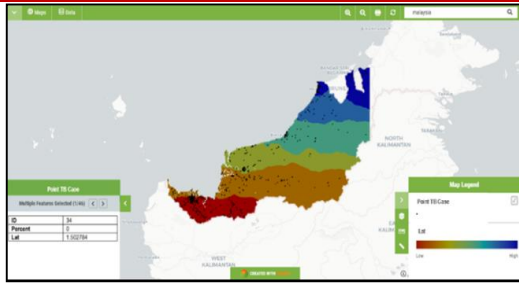
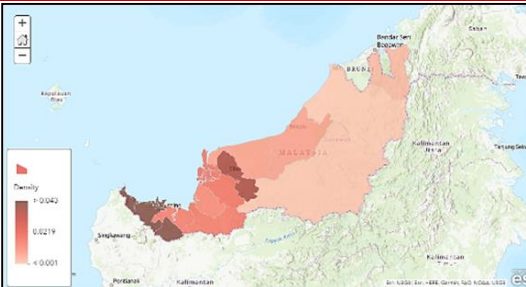
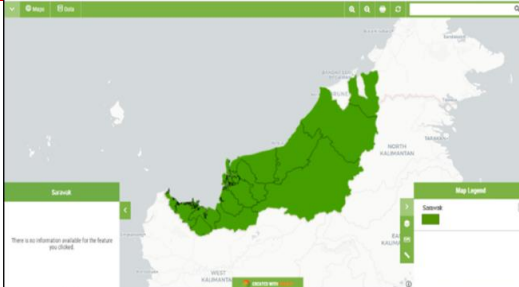
Overall, the capabilities of the selected open-source platform compared to the ArcGIS have been described in this study. Although the three free OSS can map and show the disease distribution, MapWindows GIS has potential and an application to be used in the disease management in the state. The MapWindows GIS can create an analysis system in displaying the highest heat spot of TB in the state, especially in Kuching, Miri and Sibiu while other districts are slightly normal situation. The free GIS open-source software might not be able to compete with the licenced software such as ArcGIS in term of advanced analysis, but for a basic analysis function such as to detect the hotspot area, free GIS OSS such as MapWindows might be the best choice for the non-geomatic users. It is also free and easy to be used for a mapping and basic analysis function as provided in ArcGIS. Combination of GIS open and free source software can manage and analyse the disease or human-based datasets effectively [10–15].

5. Conclusion and future work

This study has evaluated the functions and usability of the selected free GIS OSS in health and disease applications in Malaysia, in particular spatial TB data mapping and analysis in Sarawak. Various techniques and tools have been used in the software to analyse and map the pattern of TB cases. There are three main software usability applied in this study such as disease mapping, proximity analysis, spatial pattern analysis and online visualization system. From the hot spot map analysis of TB, it shows that the Kuching division recorded the highest concentration of cases since it has a small area size but has a high incidence of the disease. The comparison of three selected free GIS open-source software and ArcGIS reveals that the most applicable software for analysing and mapping disease in Malaysia is MapWindows GIS, and followed by GeoDA and gvSIG. Similarly, for a better layout and layout features, MapWindows GIS also shows the best result since it has the same function as ArcGIS in layout view output. However, GeoDa has a few advantages in terms of data analysis compared to other software because it can calculate and interpret the K mean calculation and interpret it as a map. In terms of the online geovisualization, MapWindows has an advantage because it has MangoMap online platform which can perform buffer point map and interpolation mapping. The free GIS open-source software explored in this study can be used for standard disease pattern analysis and display system online. These comparative results will help the health departments in the country to further explore the open-source platforms and utilise the tools for disease surveillance and create public awareness in the community. This study has been limited by the software selection and the testing time frame, and it is

therefore the further study might include exploring other software with advanced spatial functions.

Table 7 Comparative results of geovisualization by using ArcGIS Online and MangoMap

Platform	ArcGIS Online	MangoMap
Map		
Buffer Point Map		
Heat Map		
Dasymetric Map		

Acknowledgment

The authors gratefully acknowledge the assistance of the Ministry of Higher Education (MOHE) and Universiti Teknologi MARA Selangor for providing Fundamental Research Grant Scheme (FRGS) 600-IRMI/FRGS 5/3 (093/2019). The research is registered in the National Medical Research Register, Malaysia (ID: NMR R-15-2499-24207). The authors are also thankful to the Ministry of Health Malaysia for providing TB datasets used in this study.

Conflicts of interest

The authors have no conflicts of interest to declare.

References

[1] Berra TZ, Queiroz AA, Yamamura M, Arroyo LH, Garcia MC, Popolin MP, et al. Spatial risk of tuberculosis mortality and social vulnerability in Northeast Brazil. *Journal of the Brazilian Society of Tropical Medicine*. 2017; 50(5):693-7.

[2] <https://apps.who.int/iris/bitstream/handle/10665/336069/9789240013131-eng.pdf>. Accessed 10 October 2020.

[3] <https://kpkesehatan.com/2019/03/25/kenyataan-akhbar-kpk-25-mac-2019-sempena-sambutan-hari-tibi-sedunia-2019/?fbclid=IwAR0iZcSRCWJw63q2avREbinrTOjb2sg6BiHNqa4FepqHjhTsOBJXJUFBY>. Accessed 10 October 2020.

[4] Singleton AD, Spielman S, Brunson C. Establishing a framework for open geographic information science. *International Journal of Geographical Information Science*. 2016; 30(8):1507-21.

[5] <https://thenewstack.io/humble-beginnings-world-changing>. Accessed 10 October 2020.

[6] Azewan MD, Rasam AR. Disease mapping and health analysis using free and open source software for geospatial (FOSS4G): an exploratory qualitative study of tuberculosis. In *charting the sustainable future of ASEAN in science and technology 2020* (pp. 495-506). Springer, Singapore.

- [7] Abdul Rasam AR, Shariff NM, Dony JF. Identifying high-risk populations of tuberculosis using environmental factors and GIS based multi-criteria decision making method. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*. 2016; 42(4):9-13.
- [8] Rasam AA, Shariff NM, Dony JF, Misni A. Socio-environmental factors and tuberculosis: an exploratory spatial analysis in peninsular Malaysia. *International Journal of Engineering and Technology*. 2018; 7(3):187-92.
- [9] Evans B, Sabel CE. Open-source web-based geographical information system for health exposure assessment. *International Journal of Health Geographics*. 2012; 11(1):1-11.
- [10] Bui TQ, Pham HM. Web-based GIS for spatial pattern detection: application to malaria incidence in Vietnam. *SpringerPlus*. 2016; 5:1-14.
- [11] Jamaluddin AS, Rasam AR, Khalid N, Halim MA. Exploring open source GIS for drug mapping and analysis in Malaysia. In *international conference on system engineering and technology 2018* (pp. 91-6). IEEE.
- [12] Abdul Rasam AR, Shariff NM, Dony JF, Maheswaran P. Mapping risk areas of tuberculosis using knowledge-driven GIS model in shah alam, Malaysia. *Pertanika Journal of Social Sciences & Humanities*. 2017; 25:135-44.
- [13] Rasam AA, Shariff NM, Dony JF, Othman F. Spatial and statistics for profiling risk factors of diseases: a case study of tuberculosis in Malaysia. In *IOP conference series: earth and environmental science 2019* (pp. 1-10). IOP Publishing.
- [14] Bernardi R. IT innovation in a health information system in Kenya: Implications for a sustainable open-source software model in developing countries. In *proceedings of the 10th international conference on social implications of computers in developing countries*, Dubai 2009.
- [15] Sadoun B, Al-Bayari O, Al-Azizi J, Al Rawashdeh S. The BAU GIS system using open source mapwindow. *Human-Centric Computing and Information Sciences*. 2015; 5(1):1-12.
- [16] Duarte L, Teodoro AC, Freitas A. GIS open source application as a support to a hospital morbidity database-hospital GIS. In *international conference on geographical information systems theory, applications and management 2017* (pp. 169-76). SCITEPRESS.
- [17] Pettit C, Stimson B, Barton J, Goldie X, Greenwood P, Lovelace R, et al. Open access, open source and cloud computing: a glimpse into the future of GIS. In *Handbook of Planning Support Science*. 2020. Edward Elgar Publishing.
- [18] Mobasheri A, Mitasova H, Neteler M, Singleton A, Ledoux H, Brovelli MA. Highlighting recent trends in open source geospatial science and software. *Transation GIS*. 2020; 24(5):1141-6.
- [19] Coetzee S, Ivánová I, Mitasova H, Brovelli MA. Open geospatial software and data: a review of the current state and a perspective into the future. *ISPRS*

International Journal of Geo-Information. 2020; 9(2):1-30.

- [20] Minghini M, Mobasheri A, Rautenbach V, Brovelli MA. Geospatial openness: from software to standards & data. *Open Geospatial Data, Software and Standards*. 2020; 5:1-5.



Wan Hairi Wan Mahsin graduated from Universiti Teknologi MARA (UiTM) in Bachelor of Surveying Sciences and Geomatics (Honours). He is currently working in land surveying field. His team was awarded a Silver Medal in the 5th International Built Environment Undergraduate Research Competition (BEURC) 2020 in University of Malaya (UM), Malaysia. His research interest is GIS and Health Information Management System.
Email: w.hairi93@gmail.com



Gs Ts Dr Abdul Rauf Abdul Rasam is a senior lecturer and researcher in the Faculty of Architecture, Planning and Surveying, Universiti Teknologi MARA (UiTM) Shah Alam, Selangor, Malaysia. He obtained his PhD from Universiti Sains Malaysia (USM) in GIS (Geospatial Modelling and Geovisualisation). He has published a number of indexed papers on various topics on the Geospatial and Geomatics such as GIS, Cartography, Mapping, Remote Sensing and related fields of Computer Science, Social, Health and Environmental Studies. In addition, he has received several innovation awards from national and international organizations.
Email: rauf@uitm.edu.my



Sr Gs Noraain Mohamed Saraf is a lecturer at Faculty of Surveying Science and Geomatics, Universiti Teknologi MARA (UiTM) and has served for 3 years. She obtained a Bachelor Degree and Master of Science from UiTM. Her research interests focus on geospatial technology includes 3D LiDAR, Remote Sensing, and GIS. She is registered as a member of the Royal Institution of Surveyors Malaysia (RISM) and Institution of Geospatial and Remote Sensing Malaysia (IGRSM).
Email: noraainms@uitm.edu.my



Dr. Nafisah Khalid is a senior lecturer at Faculty of Architecture, Planning and Surveying, Universiti Teknologi MARA (UiTM). She has a PhD in Built Environment from Universiti Teknologi MARA. Her research interests focus on GIS, Remote Sensing and Environment.
Email: nafisahkhalid@uitm.edu.my