

GIS-Based Platform for Disease Mapping and Hotspot Detection in Southern Malawi

Chitani Jarves Bob Tobias

Department of Humanities (Geography), University of Livingstonia, Malawi

*Corresponding author: Chitani Jarves Bob Tobias, Department of Humanities (Geography) University of Livingstonia, Malawi.

Submitted: 25 January 2025 Accepted: 31 January 2025 Published: 05 February 2025

doi <https://doi.org/10.63620/MKWJCD.2025.1004>

Citation: Chitani Jarves, B. T. (2025). GIS-Based Platform for Disease Mapping and Hotspot Detection in Southern Malawi. *Wor Jour of Clin Der*, 2(1), 01-08.

Abstract

Geographic information systems (GIS) are critical for mapping the geographical distribution of illness prevalence, disease transmission patterns, and spatially modeling environmental elements of disease occurrence. This study was purposively to map and cluster five diseases such as cholera, diabetes, malaria, asthma and hypertension using GIS techniques in Southern Malawi. Getis-Ord G_i^* statistic was used to determine spatial distribution of diseases. Data were acquired from Malawi Demographic Health Survey (DHS) and assessed geographically using spatial-statistical analytical techniques, including the kernel and hotspot models. These models were developed to investigate and exhibit the dissimilar patterns of the selected diseases, as well as to depict places of high concentration, and they may be useful in understanding local patterns of diseases and their geographical correlations. The study found that asthma is high in Mangochi, Balaka, Blantyre, Thyolo, Mulanje, Chikwawa and Nsanje districts. It was also revealed that hypertension patients were highly concentrated in Blantyre, Thyolo, Chikwawa, Zomba, Mangochi and Nsanje Districts. Mangochi, Phalombe, Nsanje and Chikwawa are high risk areas of Malaria and diabetes is highest in Blantyre. Cholera is high in Mangochi, Balaka, Blantyre, Thyolo, Mulanje, Chikwawa and Nsanje districts.

Keywords: GIS, Disease Mapping, Disease Clustering, Spatial, Hotspots

Introduction

A considerable incentive for taking into account disease mapping as an integral part of the process of development, emanates from the WHO (2006) strategic and operational framework (SOF). Disease mapping has long been regarded as one of the most essential public health challenges, stemming from a knowledge of the link between health and geographical location. Understanding this association has long piqued the interest of scientists and academics. The disease mapping approach was first used in 1854, when Dr. Snow mapped a cholera epidemic throughout the city of London, England [1-3]. However, mapping results from GIS are one of the most useful uses of GIS in public health (Wilkinson, Grundy, Landon, Stevenson and 1998). One of the most essential GIS technologies, disease mapping technology, has been rapidly evolving [2]. GIS is increasingly using this technology as a useful tool in disease surveillance [4].

A basic overview of the illness status in a certain geographic area is provided by disease mapping, which uses disease maps as visual representations of complex geographic data [5]. Investigating the causes of diseases is the focus of epidemiology, and these causes frequently vary in frequency and spatial distribution. Because of this fluctuation in space, remote sensing was able to find a niche in the field of public health as disease researchers looked into the climatic and environmental elements that might be causing an illness. Studies have shown that the application of remote sensing to disease mapping and epidemiology, as well as to enhance surveillance and control efforts, has great potential (ibd).

The majority of epidemiological studies that had mapped vector-borne diseases in the context of environmental factors associated with those diseases had made assumptions based on established scientific evidence that those environmental factors

were associated with the disease outcome of interest [6]. To date, a growing number of disease mapping and epidemiology studies have used environmental and climatic data to map and forecast disease spread in specific geographic areas. Such studies were frequently used to assist direct and focus the deployment of health treatments to regions identified as having a high burden of the mapped disease.

Disease mapping, disease clustering, and ecological analysis are the three primary types of studies on the geographical distribution of diseases. GIS-based disease mapping relies on identifying a number of factors, the most important of which are disease occurrence locations, patterns of disease spread, environmental risk factors that lead to disease spread, and socioeconomic data, in order to analyze spatial relationships within the affected area [7].

GIS-based disease clustering studies try to determine if a disease is spatially grouped and, if so, where the clusters are located. Furthermore, these investigations are based on an analysis of prospective environmental risk factors that may exist at a given area and constitute a potential hazard [8]. Ecological analysis studies are significant in epidemiological research because they focus on examining illness distribution in connection to explanatory factors, often at an aggregated spatial level [9]. Healthcare research depends largely on GIS-based mapping and clustering technologies to provide a broad visualization aimed at controlling disease transmission and identifying disadvantaged people on a regional level (Koch, 2005) [10].

Several studies have proven that GIS is an effective tool for disease mapping and clustering. For example, people in the 1980s in the United States plotted the sites of disease presence and possible dissemination pathways. They also attempted to manage the illness by clustering. To really appreciate the role of GIS, Braga et al used Kernel and Bayesian approaches to map lung cancer illness in two Italian cities and establish disease rates at the level of urban clusters. Eisen and Eisen also examined the use of geographic information systems (GIS) to detect and control illnesses such as malaria and West Nile virus. In the same vein, Rasam, Ghazali, Noor, Mohd, Hamid, Bazlan and Ahmad, employed GIS-based cholera disease mapping to determine disease distribution patterns in Malaysia's Sabah region. To assess the pattern of illness spread, the disease mapping relied on the cohort approach. As a result, it was evident that the disease was being passed from person to person via polluted water.

Furthermore, a GIS-based measles disease mapping approach was employed in the United States to locate disease incidence areas and establish distribution patterns. In a more generic sense, it became evident that illness rates were high in places where children had not received disease immunization. The disease mapping approach aided in identifying regions in need of health services and determining the optimal sites for these services, particularly once the maps displayed the locations and quantity of illness cases, as well as the time period associated with disease transmission [11].

To obtain precise disease mapping outputs, the quality of the spatial data used in the analysis should be evaluated at the analysis level based on a number of elements, including (1) positional accuracy, (2) thematic accuracy, (3) temporal accuracy, (4) com-

pleteness, (5) logical consistency, and (6) usability (Santos, Medeiros, dos Santos, and Filho, 2017) When mapping illness cases or counts inside tracts, disease incidence data may be portrayed differently than when mapping disease structures based mostly on estimates of sophisticated models [9]. The geographical incidence of illness employs the locations of disease cases as the primary unit of observation.

Residential addresses can provide important statistics and information about environmental dangers and possible exposure [12]. The portrayal of illness rates at specified sets of places is the most basic feasible mapping form. The locations of case events should be mapped for events. To count inside tracts, a representation of the number of events within tracts of arbitrary regions with the locations and quantities of case events should be provided [8]. GIS software can conduct a variety of analytical and statistical procedures for disease mapping and grouping. By integrating and modeling geographical data, these approaches can assist with pinpoint cases and exposures, define spatial patterns, identify illness clusters, connect multiple sets of spatial data, and test statistical hypotheses [13].

The most important analytical and statistical methods of disease mapping and clustering within GIS software are: (1) a kernel density estimation method, which is used to produce spatial distribution maps of epidemic diseases by modeling disease risk prediction, (2) a weighted standard deviational ellipses method, which can compare the spatial distributions of the diseases and identify their possible spatial directions, and (3) a hotspot analysis method, which is used to identify hotspots of disease transmission [14].

In many studies, several analytical and statistical tools were used to analyze health services using GIS. Murad for example, explored the use of GIS to establish health center catchment areas using the straight-line allocation (SLA) strategy, which determines health center catchment areas based on a closest proximity approach. Murad used the Euclidean distance and drive-time methodologies to assess the accessibility of public health clinics in Jeddah. A cumulative model was developed to measure the amount of accessibility in each municipal area based on estimated distances from a road and health services. In this vein, Euclidean distance was also used [15].

The purpose of this paper is to answer the questions of how GIS can be used to map disease locations in the southern Malawi and how this technology can be used to analyze spatial clustering and determine whether any unusual clusters of health conditions exist, as well as which places have high or low disease prevalence.

Methods And Materials

The study used Getis-Ord G_i^* statistic to determine spatial distribution of diseases. This is based on hotspot analysis, and it is one of the analytical methods used to determine spatial diffusion of diseases [15]. For example, Saxena, Nagpal, Das, Srivastava, Gupta, Kumar, Jeyaseelan, and Baraik, defined a geographical distribution pattern, identified hotspots, and mapped a directional distribution trend of *Plasmodium vivax* (Pv) and *Plasmodium falciparum* (Pf) occurrences in Ranchi, India, using the Getis-Ord G_i^* and Standard Deviational Ellipse techniques. The study depended on malaria epidemiology data from 328 subcenters of the district's 14 primary health centers (PHCs)

from 2007 to 2009. In this study, a GIS was utilized to estimate the prevalence of five chosen illnesses using data from Demographic Health Survey (DHS). To predict the spatial spread of health problems in the Malawi's southern region, the Getis-Ord Gi* statistic model was applied. The Getis-Ord Gi* statistic was calculated for each feature in each dataset using ArcGIS' hotspot analysis tool.

Data analysis was done using a spatial statistic function known as standard deviational ellipses, which quantifies the orientation and direction of features and provides a tool for abstracting spatial patterns. This form of study is useful for comparing the distributions of health condition groups. Although placing elements on a map in GIS analysis might convey a feeling of direction, calculating the standard deviational ellipse explains the trend. This tool calculated the standard deviational ellipse using either the feature locations or the locations modified by an attribute value linked with the features. This is known as a "weighted standard deviational ellipse" [15].

Many satellite studies have emphasized the necessity of assessing the geographical distribution of health condition categories using the function of standard deviational ellipses [15]. Eryando, Susanna, Pratiwi, and Nugraha, for example, employed the standard deviational ellipse function, statistical analysis, overlap analysis, and environmental factors to map the geographical distribution of malaria in Sukabumi, Indonesia, where a malaria outbreak occurred from 2004 to 2012. The study focused on data gathering by Global Positioning System (GPS) plotting and field surveys based on data from Sukabumi district healthcare clinics on positive malaria cases. Nevertheless, the distribution of standard deviation ellipses was biased toward the northwest and southeast. Malaria was spread by environmental conditions such as an abnormality in rainfall and temperatures, particularly in warm and high places [15].

The physical environment of the area under study, aided vector growth and metabolism. Mapping the geographical distribution of malaria gives an initial visualization that might aid in the development of potential intervention priorities. As a result, the application of the standard deviational ellipse function aided in identifying the geographical elements that contributed to the incidence of malaria, as well as establishing prevalence trends based on specific geographic patterns. Similarly, Dong, Yang, Xu, Liu, and Chen, employed the standard deviational ellipse function to evaluate the directional trend and existence of spatial-temporal clustering of influenza A(H7N9) in China between March 2013 and December 2014. In this study, standard deviational ellipse function was used to evaluate the directional trend and existence of spatial-temporal clustering of hypertension, malaria, diabetes, cholera and asthma in the southern region of Malawi.

Study Site

Malawi's southern area, which covers around one-third of the country, is the most developed and has the biggest population of 7,750,000 of the country's 19 million inhabitants. Thirteen of Malawi's 28 districts are situated in the Southern Region. These include Balaka, Blantyre, Chikhwawa, Chiradzulu, Machinga, Mangochi, Mulanje, Mwanza, Neno, Nsanje, Phalombe, Thyolo, and Zomba. Malawi's southern area is home to some of the greatest wildlife parks in the country, as well as the second-highest mountain peak in the southern hemisphere. These are Majete game reserve, which is located southwest of Blantyre and is home to an estimated 4,000 animals, Lengwe National Park, which is located 80 kilometers south of Blantyre, and Liwonde National Park, which is located 120 kilometers north of Blantyre on the Shire River. Malawi's elephant population is concentrated in Liwonde. The Mulanje Massif peaks are quite stunning. They are Malawi's highest peak, rising to 3,0002 m (9,824 ft). They are 65 kilometers east of Blantyre. Yao, Sena, Nyanja and Lhomwe are the native languages spoken in the southern part of the country [16].

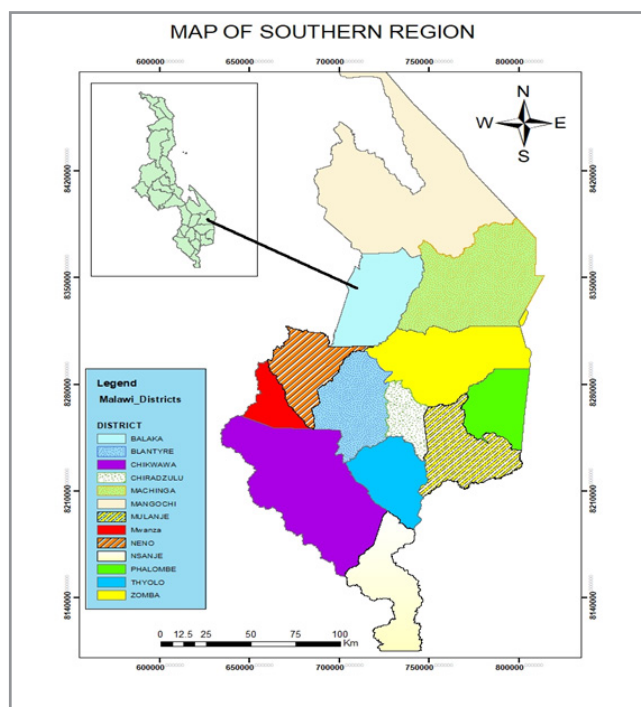


Figure 1: Map of Southern Region of Malawi

Results

Spatial Distribution of Hypertension Disease

The study found that hypertension patients were highly concentrated in Blantyre, Thyolo, Chikwawa, Zomba, Mangochi and Nsanje Districts (Figure 1). Through in-depth interviews, the study established that in urban areas many people have hypertension due to lifestyles choices of foods for example. Some districts such as Chikwawa, Nsanje, Mangochi have high temperatures. These areas are also densely populated. This demonstrates that hypertension patients follow population density patterns.

The study also used (SDE) to detect stages of hypertension. The standard deviational ellipse was employed at each phase to study the directional trend of hypertension using statistical scans to discover its spatial-temporal clusters. This is also in line with study conducted by Dong et al in China. The study discovered that A(H7N9) was in China particularly in central and western regions. As a result, identifying the spatial-temporal patterns of the epidemic revealed broad insights into the mechanics of the virus's propagation throughout China. In the southern region of Malawi, a weighted standard deviational ellipse was employed based on the number of patients with hypertension.

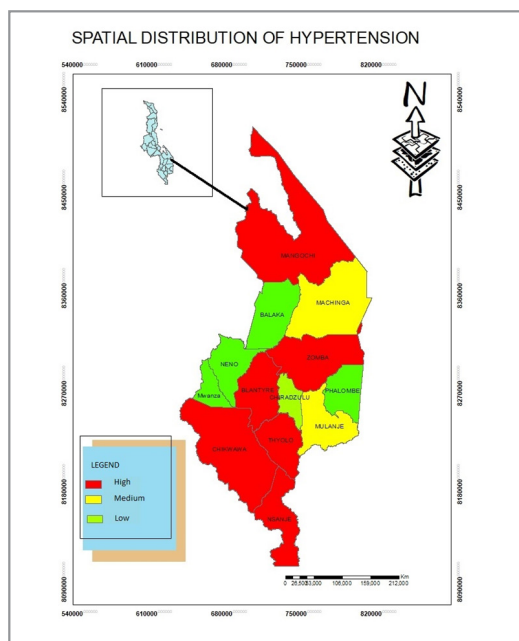


Figure 2: Spatial Distribution of Hypertension Disease

Spatial Distribution of Malaria Disease

The study found that Mangochi, Phalombe, Nsanje and Chikwawa are high risk areas of Malaria while Balaka, Machinga, Zomba and Blantyre have moderate incidents of Malaria. Through Focus group discussion, the study found that malaria occurrence is so alarming due to water bodies that host mosquitoes. The

standard deviational ellipse was employed to evaluate the directional trend of Malaria in the southern region districts of Malawi. Dong et al in China also used SDE to detect the spread of A(H7N9) where this disease spread rapidly in the central and western regions of China posing substantive threat to the inhabitants.

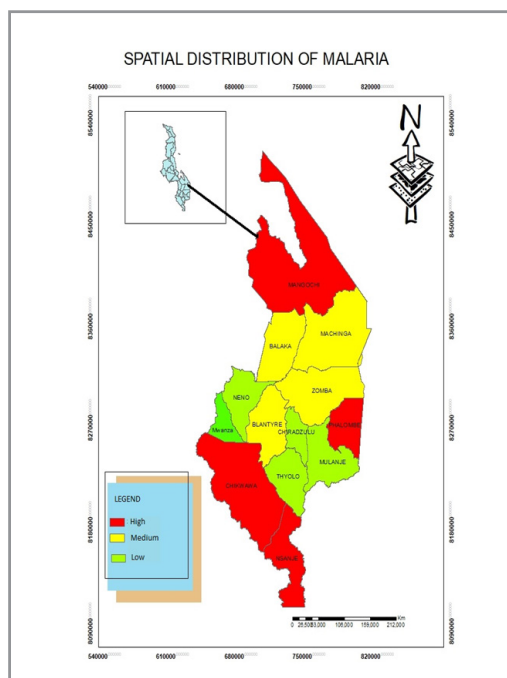


Figure 3: Spatial Distribution of Malaria

Spatial Distribution of Diabetes in the Southern Region of Malawi

The study found that diabetes is highest in Blantyre as compared to Mangochi, Neno and Mulanje and other districts in the southern region. The study found that lifestyles led to the occurrence of diabetes. The study also used SDE to detect the distribution pattern of diabetes clusters.

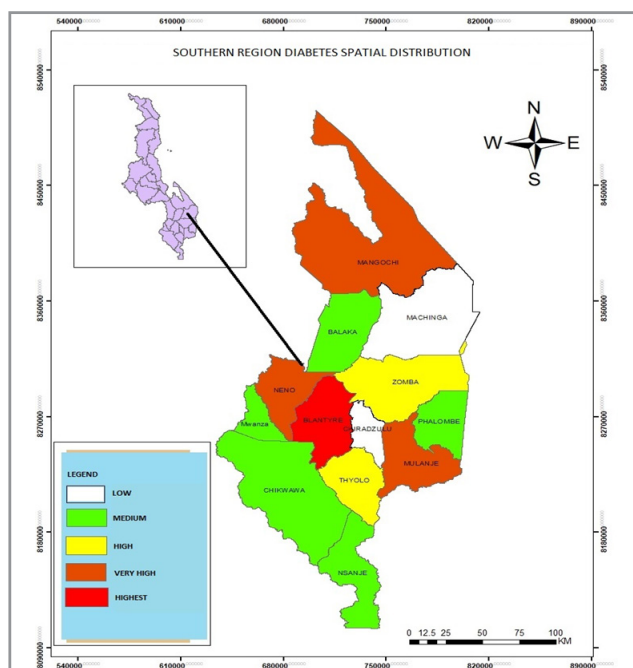


Figure 4: Spatial Distribution of Diabetes

Spatial Distribution of Asthma Disease

Based on findings, asthma is high in Mangochi, Balaka, Blantyre, Thyolo, Mulanje, Chikwawa and Nsanje districts. The study found that climatic variables led to the incidents of asthma. For example, districts such as Mangochi, Chikwawa, Balaka,

and Nsanje are dusty due to arid conditions while Thyolo, Mulanje and Blantyre are cool areas. In addition, Blantyre is highly industrialized where a lot more carbons are produced, affecting the inhabitants in these areas.

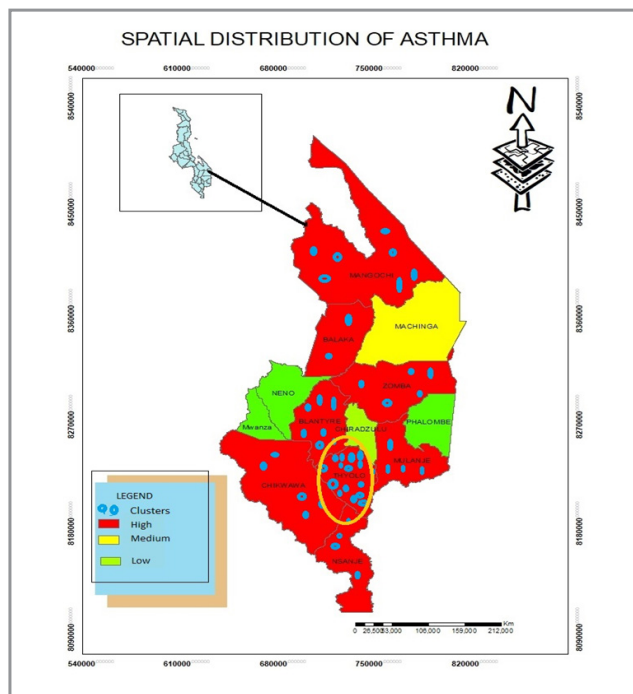
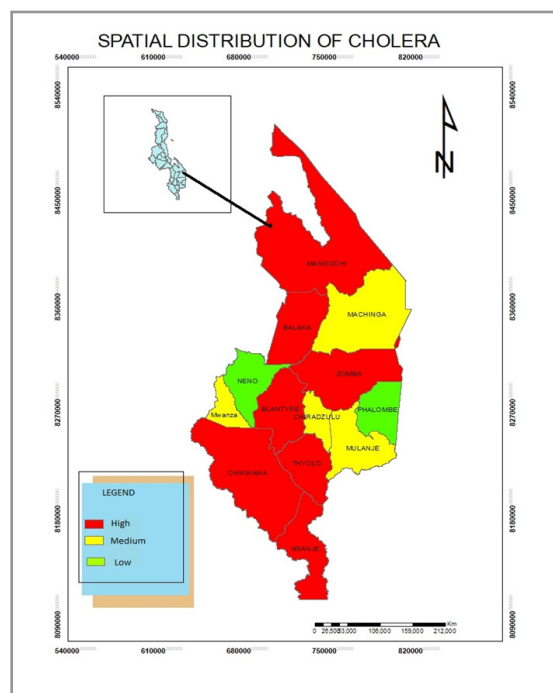


Figure 5: Spatial Distribution of Asthma Disease

Spatial Distribution of Cholera Disease

The study results show that cholera, a waterborne disease is also high in Mangochi, Balaka, Blantyre, Thyolo, Mulanje, Chikwawa and Nsanje districts. The study found that due to poor sanitation, cholera spreads rapidly in these areas.



density services for the specified kinds of diseases such as asthma and diabetes in the southern Malawi are shown in figures 4 and 5. The use of Geographic Information Systems (GIS) in health care research has grown, and its applications have gotten increasingly complex. For example, GIS may be used to examine cancer incidence in a community exposed to airborne dioxins by exposure modeling. In order to examine the link between the environment and health, statistical tools and epidemiological approaches capable of extensive analysis and visualization must be developed. Nonetheless, current GIS technologies have aided in disease control and decision-making by effectively updating, mapping, and monitoring illness.

Furthermore, GIS can map illness transmission patterns and offer a representation of how diseases develop by modeling the environmental risk variables that cause the disease.

Based on practical evidences, the conventional technique of disease control is arduous, costly, and time-consuming. The fundamental causes of illnesses and sources of infection, on the other hand, can be established by using GIS techniques to map disease prevalence, density, and geographic distribution. For example, Srinath, Szonyi, Esteve-Gassent, Lupiani, Gautam, used GIS to evaluate and map illness distribution in the US state of Texas, as well as explore environmental variables that contribute to disease spread. They stressed in their study that there is a strong association between disease distribution and environmental variables, necessitating the creation of new spatial methodologies to improve epidemiological research [15].

According to Geanuracos, public health experts are progressively changing their focus away from disease etiology models that focus solely on individual risk factors and toward models that take into account the complex and strong influences of the socio-physical environment. Many diseases, such as sexually transmitted infections (STIs) and HIV/AIDS, have been seen to spread excessively inside endemic regions or cores. These neighborhoods are frequently marked by high degrees of racial segregation, poor socioeconomic position, and high rates of homicide and other criminal activity.

Epidemiology is founded on the investigation of disease clusters to determine whether such clusters require further investigation, whether they are likely to be chance occurrences, or whether they reflect a rational interpretation of the spatial distribution of the at-risk population. Spatial clustering algorithms are exploratory tools that aid academics and policymakers in comprehending complicated geographical patterns. Knowing whether or not clusters exist and where they are situated is critical for health research and policy development. Responding to community concerns, on the other hand, targets just a subset of possible clusters and is likely to overlook clusters in groups with less political and economic clout [15].

Spatial clustering analysis is an excellent tool in public health monitoring, especially when spatial clustering approaches are combined; this will certainly contribute to uncovering the sources and causes of health concerns through more exploratory and investigative analyses [15].

In the southern Malawi, GIS was used to describe and simulate the geographical clusters of five diseases namely diabetes, hypertension, cholera, malaria and asthma. The kernel model and

the Getis-Ord Gi* statistic model were used to accomplish this [16-20].

Conclusions

GIS technology has been used to map many diseases in the public health sector. GIS technology advancements have aided in the improvement of healthcare by allowing for disease mapping and clustering, the detection of disease transmission patterns, the identification of the geographical and temporal distribution of infection vectors, and the management and monitoring of infectious diseases. Diabetes, hypertension, malaria, cholera and asthma were chosen for examination in this study in the southern region districts of Malawi [21-25].

Data about these disorders was gathered from health centers data and DHS. The locations of clusters of the selected diseases were defined using GIS-based clustering methods, including kernel and Getis-Ord Gi* statistic models. The findings demonstrate separate clusters for each disease, which will help Malawi's southern region health practitioners optimize the availability of health services in the resultant cluster locations [26-35].

Abbreviations and Acronyms

GIS: Geographical Information Systems.

Acknowledgements

I would like to thank the University of Livingstonia for financial support.

Conflict of Interest

The authors have no conflict of interest.

References

1. Geraghty, E. (2016). Why health is so spatial.
2. Kirby, R., Delmelle, E., Eberth, J. (2016). Advances in spatial epidemiology and geographic Cartographies of Disease: Maps, Mapping, and Medicine; ESRI Press: Redlands, CA, Place, space, and health: GIS and epidemiology, 14, 384.
3. GIS Geography. (2015). The remarkable history of GIS.
4. Fletcher-Lartey, S. M., Caprarelli, G. (2011). Application of GIS technology in public health: For the prediction, prevention, and control of vector-borne diseases. *Annual Review of Entomology*, 56(1), 41-61.
5. Eryando, T., Susanna, D., Pratiwi, D., Nugraha, F. (2020). Standard deviational ellipse (SDE) ESRI. In GIS for public health today and tomorrow.
6. Tran, A., Ippoliti, C., Balenghien, T., Conte, A., Gely, M., Calistri, P. (2013). *A Geographical, USA*, Wiley & Sons: Chichester, UK, 32-34.
7. Murad, A. (2025). Using geographical information systems for defining the accessibility to health.
8. Lawson, A., Browne, W., Vidal-Rodeiro, C. (2000). Disease Mapping with Win bugs and Mlwin; John Lawson, A., Cressie, N. Spatial statistical methods for environmental epidemiology.
9. Maheswaran, R., Craglia, M. (2004). GIS in Public Health Practice; CRC Press: New York, NY, models for malaria surveillance, case study: Sukabumi district-Indonesia, 11, 130.
10. American Hospital Association. (2005). Mapping Medicare Disparities. Analytics tools for infectious disease epidemiology: A systematic review. *J. Biomed. Inform*, 51, 287-298.
11. Sones, M. (2019). Reveal: Mapping and Tracking the Spread of Deadly Diseases.
12. Lawson, A.; Kleinman, K. (2005). Spatial and Syndromic

13. Carroll, L. N., Au, A. P., Detwiler, L. T., Fu, T., Painter, I. S., Abernethy, N. F., & Souris, M. (2014). Exploring spatial patterns and hotspots of diarrhea cholera mapping and analysis towards a local scale predictive modelling.
14. Gatrell, A., Senior, M. (1999). Health and health care application. Handbook of Statistics, 18, 357-396.
15. Murad, A., Khashoggi. (2020). Using GIS for planning public general hospitals at Jeddah city. Environ. Des.
16. Malawi Project. (2022). <https://www.malawiproject.org/malawis-southern-region/>
17. Braga, M., Cislighi, C., Luppi, G., Tasco, C. A. (2014). Multipurpose, interactive mortality atlas care facilities in Jeddah City, Saudi Arabia. Geospatial Health, 8(3), 661-669.
18. Carnes, A., Ogneva-Himmelberger, Y. (2018). Temporal variations in the distribution of west and geographical information systems. In Proceedings of the 2nd International Conference on Asthma and Associated Factors in Saudi Arabia: A Systematic Review and Meta-Analysis. Biomedical Research International, 9.
19. Dong, W., Yang, K., Xu, Q., Liu, L., Chen, J. (1998). Spatio-temporal pattern analysis, 179-189.
20. Esri. (2020) How Directional Distribution Standard Deviation Ellipse Works.
21. Esri. (2020) How Hot Spot Analysis (Getis-Ord Gi*) Evaluation of the spread of human infections with avian influenza A(H7N9) virus in China, 2013–2014. BMC Infect. Dis. 2017, 17, 704.
22. Geanuracos, C. (2024). Use of geographic information systems for planning HIV prevention.
23. Hassan, A. N., Beck, L. R., Dister, S. (2013). Prediction of villages at risk for filariasis transmission. Health Informatics, 5.
24. Hussain, S., Farhan, S., Alnasser, S. (2009). Time trends and regional variation in prevalence in Chiang Mai, Thailand. International Journal of Health Geographics, 8(36).
25. Indian J. Med Res. (2012). Information system-based multicriteria evaluation to map areas at risk for Rift Valley fever vector-borne transmission in Italy. Indian Journal of Medical Research, 136(6), 776-782.
26. Transboundary and Emerging Diseases. (2013). Information system-based multicriteria evaluation to map areas at risk for Rift Valley fever vector-borne transmission in Italy. Transboundary and Emerging Diseases, 60(Suppl 2), 14-23.
27. Indumathi Srinath., Barbara Szonyi., Maria Esteve-Gassent., Blanca Lupiani., Raju Gautam., Alfonso Clavijo., & Renata Ivanek-Miojevic. (2011). Information systems: A tool to map and analyze disease spread. Online Journal of Public Health Interventions for High-Risk Youths, 97(12), 1974-1981.
28. Latif, Z.A., Mohamad, M. H. (2023). Mapping of dengue outbreak distribution using spatial statistics.
29. Murad, A. (2024). Defining health catchment areas in Jeddah city, Saudi Arabia: An example
30. Photis, Y. N. (2017). Disease and health care geographies: Mapping trends and patterns in a GIS. positional accuracy assessment of geospatial data. Bull. Geod. Sci., 23, 405-413.
31. Rasam, A., Ghazali, R., Noor, A., Mohd, W., Hamid, J., Bazlan, M., & Filho, J. L. (2023). Use of geostatistics on absolute.
32. Saxena, R., Nagpal, B., Das, M., Srivastava, A., Gupta, S., Kumar, A., & Baraik, V. (2025). A spatial statistical approach to analyze malaria situation at micro level for Sci., 3, 3-22.
33. Srinath, I., Szonyi, B., Esteve-Gassent, M., Lupiani, B., Gautam, R. (2016). Geographical Successes and challenges. Parasitology, 143, 1-15.
34. Longley, P., Goodchild, M., Maguire, D., Rhind, D. Eds., John Wiley & Sons. (1999). System Principles and Application Chichester, UK, 925-938.
35. Ali Hassan. N., Beck L. R., Dister, S. (1998). Prediction of villages at risk for filariasis transmission in the Nile Delta using remote sensing and geographic information system technologies. Journal of the Egyptian Society of Parasitology, 28(1), 75-87.