GIS IN EPIDEMIOLOGIC SPATIAL DYNAMICS

Arpita Chakraborty[1]
[1] M.Tech GIS, NIIT University, Neemrana, Rajasthan

Abstract:

With the growth in science and technology, GIS has emerged as a solution to many problems. GIS along with other powerful tools such as remote sensing and GPS can be effectively and efficiently utilized for epidemiologic surveillance, monitoring and management. Health professionals are slowly realizing the interlinkage that exists between heath information and geography, and thus the emergence of GIS, which is a multidisciplinary subject, combining the concepts of information technology and geography in one. GIS allows its users to incorporate different layers of health and health related information stored in various layers into one large database file, and allows them to organize, visualize and analyze the data layers more effectively. With this sort of powerful tools it can be said that somewhere down the road, an individual's history of where he or she has lived will become an important component of medical record. To understand the environment to which they have been exposed to will become important information for clinical practices, just like an individual's genetic makeup is. Thus combining GIS with the existing medical model will surely uplift the efficiency of public health services.

About the Author:

Ms. Arpita Chakraborty (M.Tech GIS)

Currently pursuing Masters in GIS, with a high motivation to learn more.
She has worked on subject based projects such as Cartographic and Statistics project on World disaster scenario with perspective of human development index (HDI), GIS Programming project on creating ArcGIS Plugins, GIS Project Management project on public health GIS, Advanced GIS Programming project on MapQuest (Tourist Web Application) using ArcGIS Resources.

E mail ID: arpita.chakraborty@st.niituniversity.in
Contact No: +91 – 7737386481

Introduction

Epidemiology is the study of the patterns, causes, and effects of health and disease conditions in defined populations. It is the cornerstone of public health, and informs policy decisions and evidence-based medicine by identifying risk factors for disease and targets for preventive medicine. Considering this as the basic definition, data collection of health and disease condition is an important task. Along with that, if we combine the geographical location of the incidences, we get a more wholesome picture of the scenario.

The term geo-medicine has been coined recently which integrates Geographic Information Software with clinical databases. Various maps can be produced as an information product from the system, which could be used to provide significant insights that might improve individual and population health by revealing hidden spatial patterns. In healthcare, this approach is useful in finding correlations between health conditions and the geographic areas where patients live. Another area in which GIS can make use of its strength is in understanding the accessibility of health care facilities to the population. This approach will helps in planning and modeling of systems especially in the rural areas. GIS and epidemic study, also called spatial epidemiology or health geography considers the spatial and temporal characteristics of epidemiology, which can help identify causes and effects of the disease. Various spatial attribute can be used for statistical modeling of epidemics using GIS technology. These statistical models can then be used to communicate any correlation between epidemic incidents and their physical, social and environmental surroundings. This enables one to get a bigger picture of the whole scenario, and thus implement respective solution to overcome the crisis. The science of epidemiology serves as the foundation and logic of interventions made in the interest of public health. Results and facts, shown in a GIS, can be used by the planning of health services by policy makers. The ability to visualize a situation allows for policy makers to make well founded decisions about where to allocate different resources. For all this possible analysis and representations **ESRI's ArcGIS tools** are most suitable.

Epidemic and Spatial Dynamics

Spatial dynamics uses mathematical models to model real world complex problems such as that of an epidemic situation. A model is basically a simplified description of a complex entity or process. A model is comprised of entities, attributes, relations and rules. GIS architecture can be utilized to model the complex phenomenon. GIS stores the entities in a database as records/ tables and attributes as the properties of these tables. The relationship between the entities in GIS database is shown through linked tables, which share a common attribute. Finally the rule in the GIS holds all of it together and allows for accurate results while querying the created database. In other words, a dynamical spatial model is a computational representation of a real-world process where a location on the earth's surface changes in response to variations on external and internal dynamics on the landscape. Spatial dynamics basically ponders on the spatial data which is dynamic. In case of epidemics, a model would consider the infectives (people having the disease) to be dynamic (mobile) and the susceptibles to be static. A disease occurs, turns into a contagious one, and finally results in an epidemic, hence we can say that epidemic has a lot to do with neighbouring location, i.e. space. It is due to the dynamic characteristics of epidemics and the spatial aspect which contributes to it becoming and epidemic that epidemic and spatial dynamics are closely linked.

Howeever, in this paper we will be more focused on the GIS aspects of epidemics- its advantages and functions in the field of epidemic monitoring and managemnent.

Study Area

Delhi (also known as the National Capital Territory of India) is a metropolitan region in India that includes the national capital city, New Delhi. With a population of 22 million in 2011, it is the world's second most populous city and the largest city in India in terms of area. The corresponding population density was 11,297 persons per km^2.Out of total no. of 807 Medical Institutions in Delhi in 2010, 676 are in private Sector and 131 institutions are in Govt. /Public Sector. Total no. of hospital beds in these institutions are 40342 which includes 23120 beds in Govt. /Public Sector Institutions and 17222 beds in Private Sector Institutions. With estimated population of Delhi in March 2010 on the basis of Census 2011, bed population ratio was 2.45 beds per thousand populations.

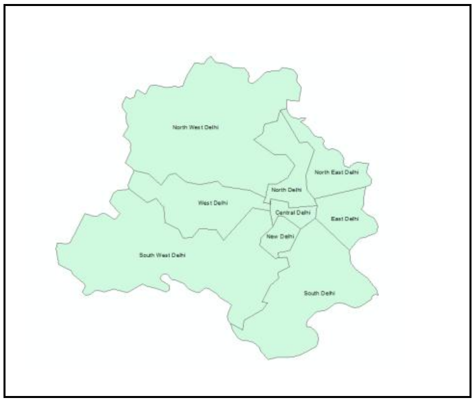

Fig: 1 – Delhi State (district/zone wise)

GIS and its functionality

Geographic Information System (GIS), has many advantages over the traditional technology in the epidemic surveillance field. GIS has its advantages in geospatial data management, visualisation and mapping and spatial analysis. Its strength can be effecitively used in epidemic mapping to reveal hidden spatial trends and distributions which may otherwise be difficult to detect. Since epidemic has a lot to do with it's surroundings and neighbouring locations, incorporating GIS for monitoring and management of epidemic can be very useful. Foremost we must realize that infections are dynamic, hence the concept of spatial dynamics. With the emerging technology, we have started to relaize a more global persepective, both conceptually and geographically, where we consider social, climatic and environmental conditions driving disease emergence. The main functionalities of Geographic Information System in epidemic management is explained as stated under:

a) **Spatial Data Management**

A spatial database is a database that is optimized to store and query data related to objects in space, including points, line, and polygons. It is a collection of spatially referenced data that acts as a model of reality. Spatial data can be classified into 2 types, i.e. raster data and vector data.

An example of raster data is satellite imagery, where each pixel stores a measured value. Vector data includes points, line and polygons. GIS technology can work with both these types of data. For epidemic management, we require an extensive database comprising of hospital locations, health data, social and environmental data. Once data has been accumulated, various spatial queries can be carried out i.e. spatial range queries, nearest neighbor queries, spatial join queries etc. Foe epidemic surveillance, monitoring and management, we can make use of all the above mentioned queries. Example, once the database of hospital locations is ready, we can plot the point features (hospital locations) on the base map of the study location, we can run spatial range query. A spatial range query will highlight objects of a particular type within a particular distance from a given location. If we have identified the incident location of an epidemic, we can use this toll to find all hospitals in the range of 25 Km from that particular location. Nearest neighbor query, locates objects (in this case hospital) that are closest to a given location, allowing patients to locate the nearest possible hospital.

b) Visualization and Mapping

In a GIS system, spatio-temporal data can be retrieved and processed to provide users with necessary information for their decision. The system is friendlier with users if the information is visualized. Epidemic arises at a place and spread too many other locations. It is an event changing in space and time, moving from place to others and appearing from this time moment to other. It is really difficult for people to recognize its evolution. The visualization techniques are applied here to provide users with changes of its location and time on screen of computer. We can map the hospital locations, and have a different layer corresponding to population density of the study location. On superimposing the two layers we can derive some useful patters, which can help in the mitigation process of the epidemic. “A picture speaks a thousand word”, and what is a map, but a useful, information deriving picture. It conveys much more than simple texts and tabular information. It will reveal trends and patterns which can be used in decision making process. Visualization of the epidemic information can help us locate the hot spots, which are critically impacted and help can be routed to that particular location. It offers powerful tools to present spatial information to the level of individual occurrence, and conduct predictive modeling. It determines geographical distribution and variation of diseases, and their prevalence and incidence. Visualization and mapping can explore the spatial patterns and correlations of diseases and many factors such as census and environment. With the growing technology, dynamic web maps have come into existence. These web maps operate in online mode, and permits a dynamic link between database and maps so that any updates in the database is automatically reflected in the maps. These maps can be easily accessed from the web, and thus prove to be useful even for the general public.

Moreover, GIS can help in generating thematic maps - ranged color maps or proportional symbol maps to denote the intensity of a disease or a vector.

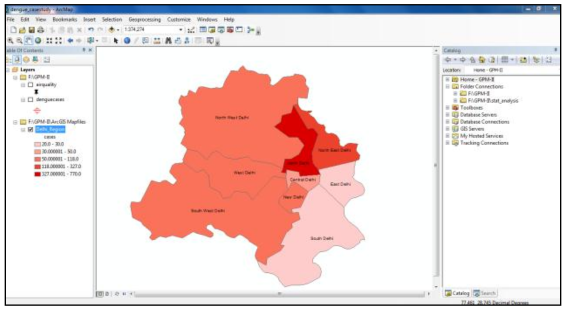

Fig: 2- Thematic Map showing vulnerable zones in Delhi State using **ArcMap**

In comparison with tables and charts, maps developed using GIS can be extremely effective means for communicating messages clearly even to those who are not familiar with technology. It allows policy makers to easily understand and visualize the problems in relation to the resources, and effectively target resources to those communities in need.

c) Spatial Analysis

Data analysis transforms data into useful information, and spatial analysis is the type of analysis which is dependent on space, i.e. location of the entity being analysed. Since in spatial analysis, location plays a major role, we can that in spatial analysis if we move the feature(s) the result changes or the analysis modifies the geometry or creates new geometries. Basically, spatial analysis examines relationships between geographic features collectively and to use the relationships to describe the real-world phenomena that map features represent. In order to perform spatial analysis, one must have sufficient knowledge of the problem and/or question to be answered, knowledge of the data, knowledge concerning the GIS capabilities and statistical analysis. To comprehend spatial information we need to work with three basic elements namely, location (the x, y co-ordinate), attribute data (describing the non-spatial characteristics of the locations) and topology (describes the spatial relationship between spatial features). For epidemic study, spatial analysis is very important as it shows any spatial correlation that may exist. The various spatial analyses done for epidemic surveillance includes the followings:

- Buffer analysis

Buffer analysis is used for identifying areas surrounding geographic features. The process involves generating a buffer around existing geographic features and then identifying or selecting features based on whether they fall inside or outside the boundary of the buffer. In case of an epidemic, for example, a radius of 10 km around a hospital to depict its catchment area. Buffer can be created around a point, line or polygon feature. In this study, we will be concentrating only on point feature buffering. Let us say for example we want to create a buffer around hospital locations (which are stored as points in a separate layer).A point is the basic unit of resolution in any GIS system. Buffering point data involves the creation of a circular polygon about the point of interest. The radius of this circular polygon is called the buffer distance. In this scheme the buffer distance or the radius of the circle could be fixed for all points in a layer or the user could specify it.

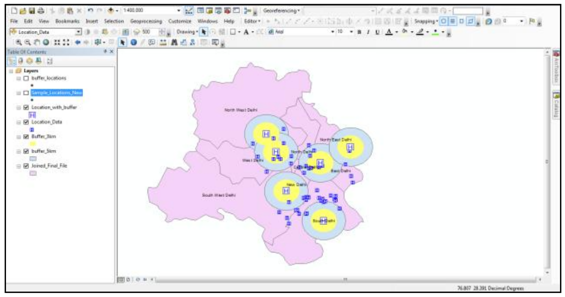

Fig: 3- Buffer Analysis using **ArcToolbox (ArcGIS)**

If we superimpose the buffered data with census information, we can see hospital to patient ratio of each buffer zone. Where the ratio is extremely low, new hospitals or clinics can be set up in those regions.

- Statistical analysis

Statistical analysis measures two categories of spatial category:

a. Identifying characteristics of a distribution
b. Quantifying geographic patterns

Spatial statistics are important for better understanding of the geographic phenomena, pinpointing the causes of specific geographic patterns, summarize the distribution in a single number and make decisions with a higher level of confidence.

Spatial autocorrelation is a statistical tool that can be used in epidemic study. ESRI states it's a measure of the degree to which a set of spatial features and their associated data values tend to be clustered together in space (positive spatial autocorrelation) or dispersed (negative spatial autocorrelation).Global Moran's I is used to establish the autocorrelation parameter on the basis of the numerical value of Moran I index .This can help identify the vulnerable clusters affected by the disease.

Hot Spot analysis is another statistical tool used to identify significant clusters. Density mapping can locate clusters in the data set, but cannot guarantee its significance.

Here, we have done Hot Spot considering vulnerable locations facing the epidemic condition.

For Hot Spot Analysis, we work with input feature class (not point data). For each feature the Gi^* value is computed internally using the following formula:

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j} x_j - \bar{X} \sum_{j=1}^{n} w_{i,j}}{S \sqrt{\frac{\left[n \sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j} \right)^2 \right]}{n-1}}}$$

Where x_j is the attribute value for feature j and $w_{i,j}$ is the spatial weight between features i and j, n is the no. of features considered.

The output is a Z score (GiZScore) and P value (GiPValue) for each feature .A high Z score and small P value for a feature indicates a significant hot spot. The higher (or lower) the Z score, the more intense the clustering. A Z score near zero means no spatial clustering.

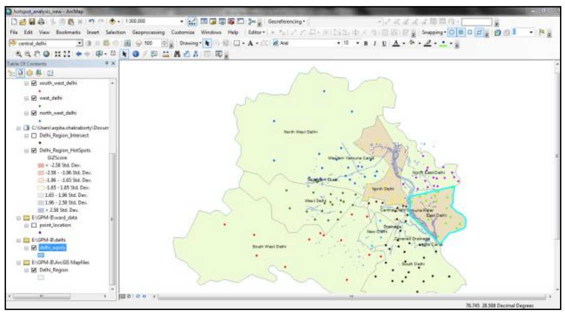

Fig: 4 - Showing Hot spots near Yamuna River, Delhi using **ArcGIS Toolkit**

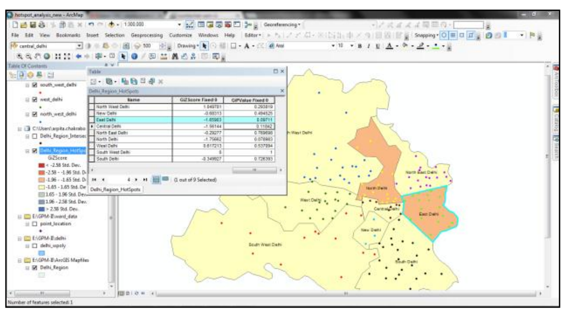

Fig: 5 - East and North Delhi Hot spots for Dengue cases using **ArcGIS Toolkit**

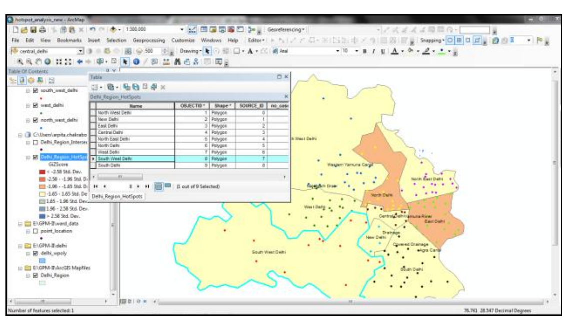

Fig: 6 - South West Delhi showing no spatial clustering using **ArcGIS Toolkit** (GiZScore=0, PValue=1)

Hot Spot analysis show dengue cases maximum near the stagnant portion of Yamuna River in Delhi state. North Delhi and East Delhi segment are relative hot spots. Low P value of 0.078983 and 0.09711 respectively suggest they are significant clusters. Similarly, South Delhi region shows a P value of 1 and Z score of 0, thus representing no spatial clustering.

Kernal Density Mapping calculates the density of features in a neighborhood around those features. It can be calculated for both point and line features. For this epidemic study of dengue cases in Delhi, we did Kernel Density Mapping for the dengue incidence locations. It was used for the purpose of estimation and identification of pattern of spatial diffusion. The Red region shows spatial clustering of dengue cases whereas, blue region show spatial dispersion.

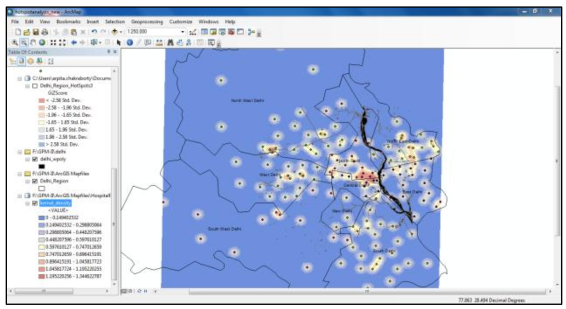

Fig: 7- Kernel Density Mapping using **ArcGIS Toolkit**

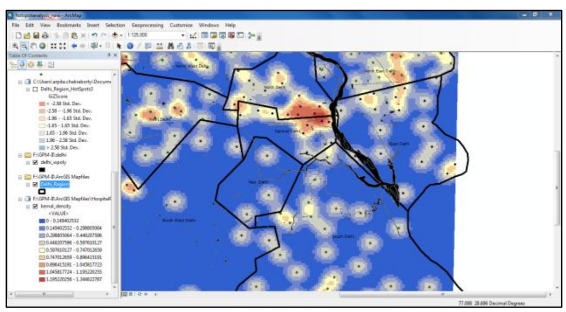

Fig: 8- Zoom in view of Kernel Density Map using **ArcGIS Toolkit**

This Kernel Density Map shows spatial clustering in the Central and North Delhi district, marked by the red region. We can also see that the spatial clustering is observed near the stagnant portions of Yamuna River. This visual interpretation can help in the decision making process for the mitigation of the disease.

Conclusion:

The use of GIS within epidemiology offers many advantages. Visualization of data is one of the greatest strengths, which enables us to identify patterns and relationships. By discerning patterns of disease greater insight is gained in the underlying factors affecting health of people and populations.

GIS analysis shows that the most vulnerable zones in Delhi are North and East Delhi zones with maximum number of dengue cases recorded. Moreover, it also can be interpreted that the vulnerable region is around the stagnant portions of the Yamuna River flowing through Delhi. The safest/least vulnerable zone is the South West Delhi zone, showing no spatial clustering of disease in the hot spot analysis. These resulting maps can be used by policymakers for health service planning.
All analysis in this paper was done using ArcGIS Software, both for spatial analysis and visualization and mapping purposes.

References:

1. Rosan van Wilgenburg, SPINlab, VU Amsterdam, **GIS and epidemiology, a SWOT analysis.**

2. Pierre Chaplet, Bertrand Lefebvre, University of Rouen(France), **Conceptualizing the Urban Healthcare System**

3. Krishna Prasad Bhandari[1], PLN Raju[2] and B. S. Sokhi[3] , **Application of GIS Modelling for Dengue Fever prone area based on socio-cultural and environmental factors- A case study of Delhi city**

   [1]Head of department science and humanities, Western Region campus, Institute of Engineering, Tribhuvan University, Pokhara, Nepal
   [2] In- charge, Geoinformatics Division, Indian Institute of Remote Sensing, NRSA DehraDun, India
   [3]Head HUSA Division, Indian Institute of Remote Sensing, NRSA, DehraDun, India,

4. Ranvir Singh, Jawaharlal Nehru University, New Delhi, **GIS in Health Management Information System: A case study of Delhi**
5. *Dr. C.P. Johnson[1], Dr. Jasmin Johnson[2],* **GIS: A Tool for Monitoring and Management of Epidemics**
   [1] Geomatics Group, C-DAC, Pune
   [2] B.J. Medical College, Pune