# Spatial and Temporal Variation of Japanese Encephalitis Disease and Detection of Disease Hotspots: A Case Study of Gorakhpur District, Uttar Pradesh, India

Shipra Verma [a], Rajan Dev Gupta [b]

[a] GIS Cell, Motilal Nehru National Institute of Technology Allahabad- 211004, U.P.  (India),
shipraenvsc@gmail.com,
[b] Civil Engineering Department, Motilal Nehru National Institute of Technology, Allahabad-211004, U.P. (India)
gupta.rdg@gmail.com, rdg@mnnit.ac.in

**Commission VIII, WG VIII/2**

**KEY WORDS:** Japanese Encephalitis, GIS, Spatial analysis, Clustering, Hotspot

**ABSTRACT:**

In recent times, Japanese Encephalitis (JE) has emerged as a serious public health problem. In India, JE outbreaks were recently reported in Uttar Pradesh, Gorakhpur. The present study presents an approach to use GIS for analyzing the reported cases of JE in the Gorakhpur district based on spatial analysis to bring out the spatial and temporal dynamics of the JE epidemic. The study investigates spatio-temporal pattern of the occurrence of disease and detection of the JE hotspot. Spatial patterns of the JE disease can provide an understanding of geographical changes. Geospatial distribution of the JE disease outbreak is being investigated since 2005 in this study. The JE incidence data for the years 2005 to 2010 is used. The data is then geo-coded at block level. Spatial analysis is used to evaluate autocorrelation in JE distribution and to test the cases that are clustered or dispersed in space. The Inverse Distance Weighting interpolation technique is used to predict the pattern of JE incidence distribution prevalent across the study area. Moran's I Index (Moran's I) statistics is used to evaluate autocorrelation in spatial distribution. The Getis-Ord Gi*(d) is used to identify the disease areas. The results represent spatial disease patterns from 2005 to 2010, depicting spatially clustered patterns with significant differences between the blocks. It is observed that the blocks on the built up areas reported higher incidences.

## 1. INTRODUCTION

The environment is highly conducive to the propagation of Vector Borne Disease (VBD) all over the world. JE is a VBD that occurs in most of Asia. It is a major public health problem in India, and mainly affects children and young adults (Mutheneni *et al.*, 2014). It is estimated that three billion people live in Japanese Encephalitis Virus (JEV) prone endemic regions, and the disease incidence is about 50,000 cases and 10,000 deaths annually (Fischer *et al.*, 2008). As per World Health Organization estimate, JE claims 10,000 to 15,000 lives a year (WHO, 2011). Epidemics of JE were recognized as early as 1871 in Japan and were common in Japan, Korea, and China in the first half of 1900s. The occurrence of this disease has been reported in several states of India like Bihar, Andhra Pradesh, and Orissa. In recent times, it has emerged as a serious health problem in rural areas of Eastern Uttar Pradesh (UP). Among 175 districts, 80 districts are classified as endemic, of which 54 (68%) were from UP alone (Sabenson, 2008, Saxena *et al.*, 2009). The scourge of the disease is most severe in the Gorakhpur district (Singh. 2007). It is difficult to eradicate JE because it is transmitted from natural reservoirs, like pigs, waddling birds, which are important amplifying hosts and man is involved as an accidental host (Khinchi *et al.*, 2010, Fischer *et al.*, 2010). The epidemiology of JE virus is not well understood and less research has been done (Upadhyayula , 2012). It is important to recognize the spatial and temporal characteristics of the JE disease transmission. Geographic Information System (GIS) plays a vital role in the surveillance and control of the vector-borne diseases. as it has the capability to scrutinize various factors associated with the disease through the spatial analysis (Moore *et al.*, 1999) and statistical techniques (Gupta *et al.*, 2000). The present study evaluates the JE diffusion pattern from starting outbreaks of year 2005 till 2010. The aim is to demonstrate the potential use of GIS based methodology to explore and present the dynamic epidemic records geographically.

## 2. STUDY AREA

Gorakhpur occupies the North Eastern corner of the State of Uttar Pradesh, and is located on the bank of river Rapti and Rohni, a tributary of Ganges originating in Nepal, that sometimes causes severe floods. Based on the geographical distribution, the district occupies the 15th position in the State in terms of size, covering an area of 3,483.8 km$^2$. The study area lies between latitudes 26°13′N and 27°29′N and longitudes 83°05′E and 83°56′E (Figure 1). In census 2001, Gorakhpur had a population of 37,69,456. The population of Gorakhpur district contributes to 2.22 percent of the total population of Uttar Pradesh. The majority of the cases of JE came from

Eastern Uttar Pradesh, in particular, Gorakhpur, which is the paddy growing Terai area.

The habitats have been modified for the development of agriculture in the region, which has resulted in the vast expansion of water bodies which support mosquito breeding. These environmental changes probably have triggered JE in the region. Information available on the mosquito species, particularly JE vectors prevailing in Gorakhpur district are very little. Every year a number of cases and deaths occur in this division due to this notorious disease. Year 2005 was more dreadful, between29 July to 31 December when 4114 JE cases and 1003 deaths were reported from Gorakhpur division (Srivastava *et al*. 2008).
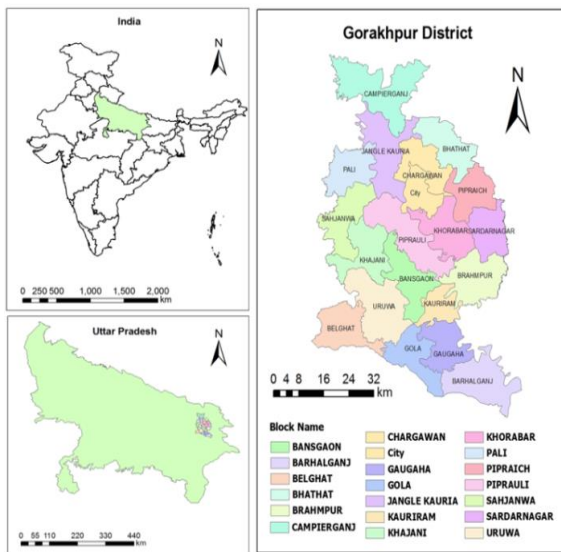


Figure 1: Location Map of Study Area

The Gorakhpur district correspond seven Tehsils namely Bansgaon, Campierganj, Chauri chaura, Gola, Khajani, Sadar Sahjanwa and nineteen blocks.

The Tehsil are divided into nineteen blocks. Sadar tehsil corresponds to five blocks, namely Jungle Kauria, Chargawa, Bhatahat, Khorabar and Piparaich. Chauri Chaura is into two blocks, i.e. Sardar Nagar and Brahmpur. Sahjanwa, Khanjani and Campierganj are divided into two blocks, namely Pali, Piparauli, Belghat, Khanjani and other Campierganj and Jungle Kauria. Tehsil Bansgaon is divided into three blocks Bansgaon, Kauriram and Gagaha and Gola tehsil in three blocks namely Gola, Uruwa and Barhalganj.

## 3. DATA COLLECTION

The data of JE reported cases from 2005 to 2010 is used in this study. The data was obtained from District Hospital of Gorakhpur. Population data of 19 blocks of district was collected from Census data 2001. After the outbreak of JE cases in 2005, all the patients with symptoms of brain fever and headache were considered JE infectious, most of them being children.

## 4. METHODOLOGY

### 4.1 Spatial analysis

Gorakhpur district comprises 7 tehsils namely Chauri Chura, Gola, Khajani, Bansgaon, Campierganj, Sadar and Sahjanwa. The tehsils are subdivided into 19 blocks in the district. The collected for each year from 2005 to 2010, JE incidence per year at every block level is analyzed. For conducting a GIS-based analysis of the spatial distribution of JE, a block-wise polygon layer was generated based on the administrative boundary map at a scale of 1:50,000 obtained from district planning map, Gorakhpur. All JE cases were geo-coded and mapped to the block layers of polygon by the administrative code using the GIS software Arc GIS 9.3 (ESRI Inc., Redlands, CA, USA).

### 4.2 Spatial mapping

Data from all the JE cases was collected and geo-coded at block level. Mapping incidence of JE disease is the first step in spatial analysis and needs to be done carefully. Blocks in high number of population have more incidences, which closely reflect in the statistical variability. A spatial filtering was performed to reduce the problem in differentiation of range in the data that comes from the high variance characteristic populations. The disease incidence rate is computed from the total number of cases per 10,000 at each block, divided by the total number of people at risk within the blocks.

### 4.3 Space-time analysis

The understating on the spatio-temporal dynamics of the JE outbreaks at spatial distribution, ranging from local to global areas is a critically observed. GIS software and improved analysis techniques provide opportunities to study and model spatio-temporal dynamics of the disease outbreak (Jeeffo *et al.*, 2011). GIS software contains space-time data analysis tools, which are quite useful in the study of spatial epidemiology.

### 4.4 Inverse Distance Weighted (IDW) analysis

IDW is a method to interpolate the surface based on the assumption that the surface is more influenced by nearby similarities and distant by more dissimilarity (Leung *et al.*, 1998). The IDW interpolation technique is used to visualize the surface of disease incidence rates of JE cases across the whole district. The spatial distribution mapping of the JE disease requires a potential risk on the district. The IDW interpolation technique is commonly used in the GIS tool for producing surfaces using interpolation of scatter point (Hu *et al.*, 2007). GIS tools and spatial statistics enable the spatial distribution and predict the disease outbreaks (Pathirana *et al.*, 2009). Since the IDW method is an exact method and is more accurate, it sustains the entire probability distribution of incidence values, which serves as a reliable data to analyze the disease spread (Kheirandish *et al.*, 2012).

### 4.5 Spatial autocorrelation

The word autocorrelation refers to the correlation of a variable across geographical locations (Hayran, 2004). The Moran's I Index (Moran's I) statistics is used to evaluate autocorrelation

in JE spatial distribution and tests with high (or low) incidence rates. Moran's I is a spatial autocorrelation analysis tool used to measure the presence of clustering of similar values (Tsai *et al.*, 2009). Spatial autocorrelation of the ArcGIS 10 spatial analyst tool is used to detect the values approaching units, which are situated in similar geographic units or near dissimilar units, as indicated by numbers between -1 to 1. The value of 1 is for the presence of disease cases that are geographically similar and -1 when dissimilar. If the value indicates 0, it means it is an independent value and there is an absence of autocorrelation among geographical units. The indices are evaluated, considering the original location of the blocks. The spatial patterns of JE usually present a strong clustered autocorrelation due to the spatial relationships between cases and mechanism of the disease, involving distance and neighborhood.

### 4.6 Cluster-outlier detection

The cluster-outlier (CO) type field distinguishes between a statistically significant (P<0.01) cluster of high values (High-High), cluster of low values (Low-Low), outliers in which a high value is surround primarily by low values (High-Low), and outliers in which a low value is surrounded primarily by high values (Low-High). A positive value of 'I' indicates that the feature is surrounded by features with similar values and this feature is the part of a cluster. A negative value for 'I' indicates that the feature is surrounded by features with dissimilar values. Moran's I index are interpreted within the context of the computed Z-score or p-value. The 99% significance level (P<0.01) is used to indicate significant clusters of spatial autocorrelation.

### 4.7 Hotspot detection

Hotspot is defined as a condition indicating some form of clustering in a spatial autocorrelation (Osei *et al.*, 2008). This led to use of the Getis-OrdGI* (d), which can separate clusters of high values from clusters of low values (Getis *et al.*, 1992). Moreover, clusters of cases that occur randomly can have an influence on the spread of infectious disease. The Gi* (d) statistics is useful for determining the spatial dependence of neighboring observations (Scott *et al.*, 2010). The result expresses the Z-score and p-value of the calculated value of Gi*(d) in comparison with the normal distribution of statistics calculated by simulation (Bhunia *et al.*, 2013). These values represent the statistical significance of the spatial clustering of values, given the conceptualization of spatial relationships and the scale of analysis in distance parameter.

The cluster and hotspot detection are computed on a yearly basis by global clustering analysis for the cases of each block. Mapping location and movement for each year computed hotspots, which highlighted the movement pattern of the disease. Next, locations with significant value of clustering were plotted on a map from 2005 to 2010 to display the specific locations of the JE hotspot.

## 5. RESULTS

JE occurred in most part of central blocks, causing severe health problems on the affected population.

The After combining the JE cases for all the blocks, in 2005, the total number of JE cases was 879. These total cases were 458 in 2006, 647 in 2007, 645 in 2008 and 668 in 2009. But in 2010, the total number of cases was 923 which is the highest recorded incidence for the decade.

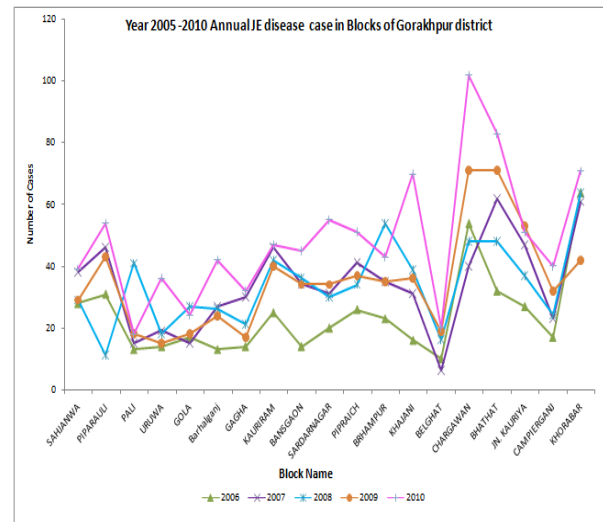The detailed block wise epidemiology data from 2005 to2010 is shown in Figure 2.



Figure 2: Block wise number of JE cases in the district from 2005 to 2010

### 5.1 Mapping of the JE disease

Disease incidence mapping at block level needs to be analyzed to avoid misinterpretation. The statistical method is helpful to understand the incidence rate but the variability on the absolute position is not clear. The results of smooth maps allow the analysis of the pattern and occurrence of cases per 10,000 populations in the blocks. The high incidence per 10,000 populations was reported in 2005, in 2006 it was lower but gradually increased and was highest in 2010. Map in Figure 3 shows that the disease occurred everywhere in the blocks and in central areas of the district along with eastern part of block, which is in main city and spreads to the rural land.

The results of analysis show that JE cases are spread all around the district. The Figure 3 shows 'very high endemic blocks' in 'blue' color for the blocks namely Chargawan, Khorabar, Bhatahat in the urban region of the district, 'high endemic blocks' in 'Pink' color which are Piparaich and Jungle Kauria, 'medium endemic blocks' with 'red' color including Piparauli, Sardar nagar, Khajani, Brahmpur and Kauriram, 'low endemic blocks' with 'yellow' color that are Bansgaon, Sahjanwa, Campierganj, the 'light yellow blocks' with very low number of diseases includes Uruwa, Pali, Belghat, Gola, and Gaugha blocks of the district.
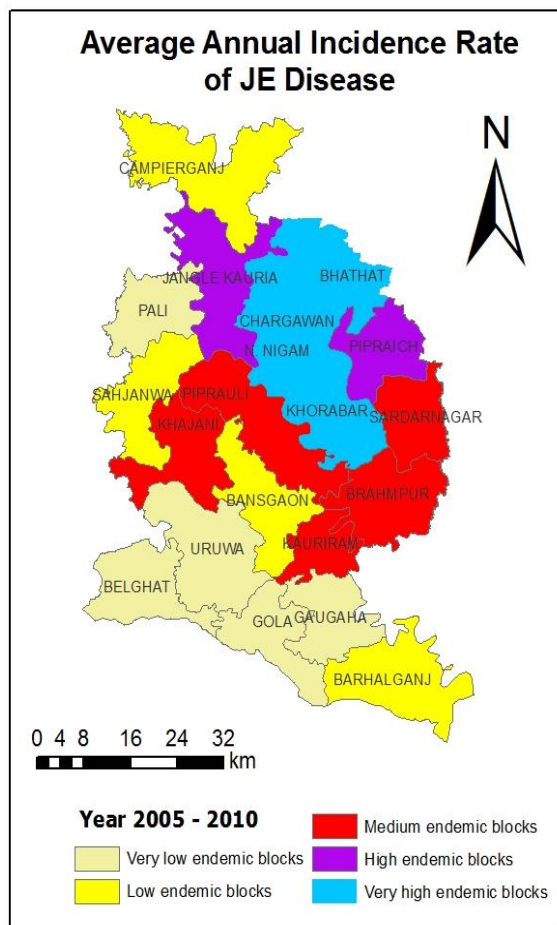
Figure 3: Classification of blocks based on severity of the JE disease from 2005 to 2010

The results of the analysis, as shown in Figure 3, conclude that most of the highest endemic blocks are located in the central part of the district including the city portion of the district.

The interpolated IDW map of incidence rates of JE disease from 2005 to 2010 are used to observe the high risk areas during the successive period. This is shown in Figure 4.
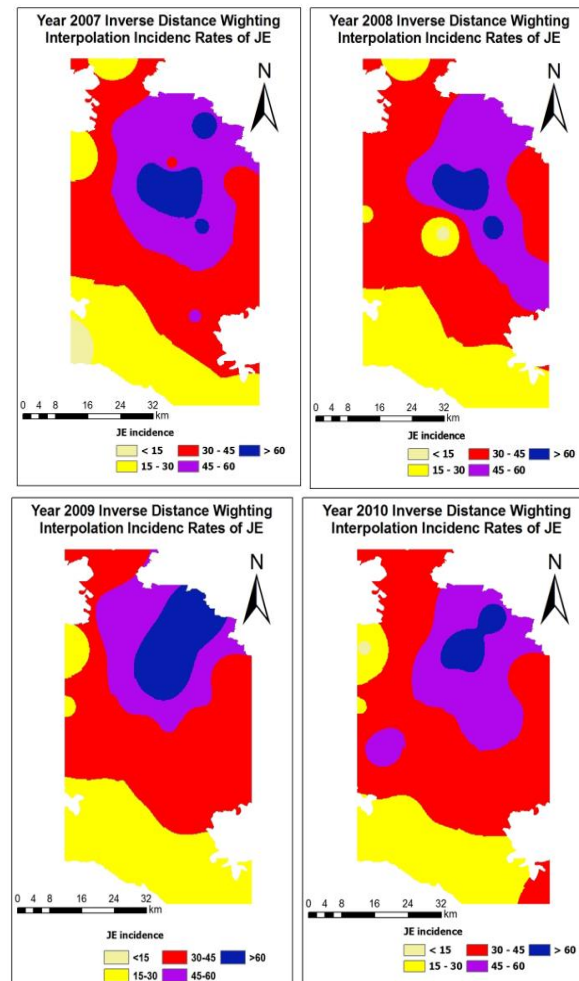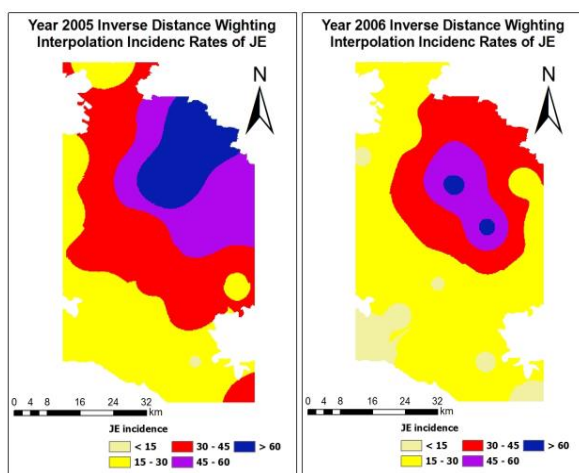




Figure 4: Maps showing the IDW interpolation rates of JE disease from 2005 to 2010.

From the IDW maps shown above, it can be observed that the pattern of disease varied geographically across the district. The high level of intensity grows in the center, which is shown in dark blue color and visualizes that the darker area has the higher risk of JE. In 2005, high-risk JE incidence foci are found in the center to southern part, in 2006, the risk decreased and is visualized in the central part. There is a concern that the disease gradually increased from 2008 to 2010 in middle to south-eastern part.

## 5.2 Spatial-temporal distribution of JE disease

This analysis is performed on the incidence rates of the JE disease to test whether the cases are distributed randomly over space and, whether the autocorrelation technique can identify spatial disease clusters for statistical significance.

The spatial autocorrelation analysis with Moran's I index is explained in Table 1. The cluster results of 2005 to 2010 show that the spatial distribution of the JE disease are not uniform in all blocks of the district in each year. It presents the expected clustered pattern for an infectious disease. Moran's I Index shows that cases distribution in the year 2005 produced a Moran's I value of 0.07. This suggests that the disease pattern

is a result of random chance. Further, higher Moran's I (value being 0.09) is observed in the year 2010.

Table 1: Spatial autocorrelation analysis of JE

| Year | Z-score | Moran's I | Pattern |
| --- | --- | --- | --- |
| 2005 | 5.85 | 0.07 | Clustering |
| 2006 | 5.7 | 0.02 | Clustering |
| 2007 | 2.47 | 0.03 | Clustering |
| 2008 | 1.55 | 0.02 | Clustering |
| 2009 | 2.85 | 0.05 | Clustering |
| 2010 | 4.93 | 0.09 | Clustering |

The analysis of results shows that the larger the Z-score value, the more intense is the clustering of high values as hotspot. Similarly, the smaller the Z-score is, the more intense clustering of low values as coldspot. In the Figure 5, darker areas indicate statistically significant hotspots, while light areas represent significant cold spot areas. These locations are considered as hotspots for making future strategy to control the outbreak of JE disease.

Using spatial temporal analysis in the GIS environment, the JE cases in Gorakhpur district are mapped and analyzed. The maps in Figure 5 show spatial patterns of JE cases. It can be observed that most of the cases are spread in the central and north-western part of the district in year 2010. It also illustrates that the hotspot is found in the centre of the district covering urban and around semi-urban areas.

The coldspots are found in the southwest and south-north areas in the district. The study significantly extracted the hotspot of JE cases over a six year period. During the collection of data from the block Chargawn, the disease cases were 81 in 2005 and predominantly increased to 102 cases in year 2010. Piparauli, Bhatahat, Khajani and Jungle Kauria blocks are also an area of concern due to the high incidence rate.
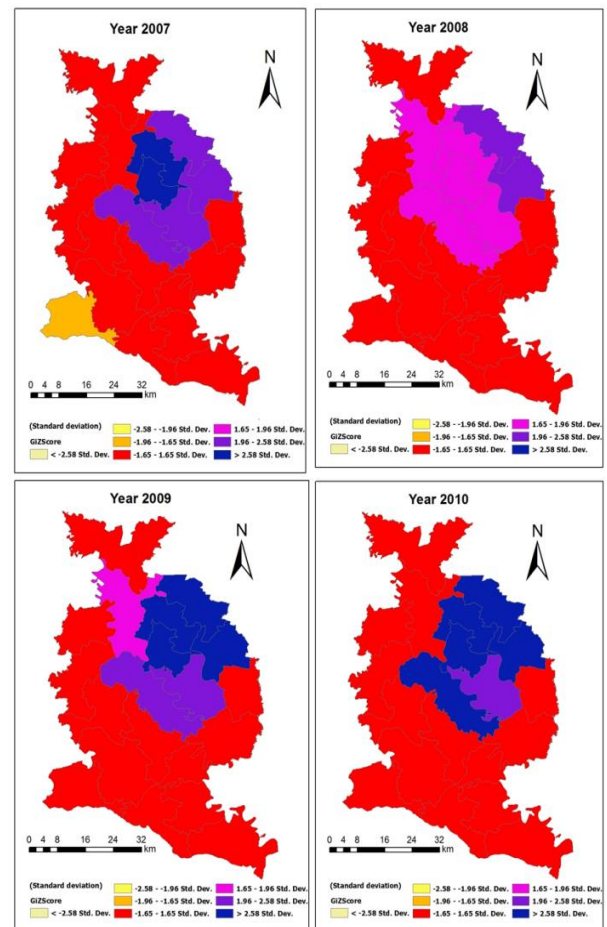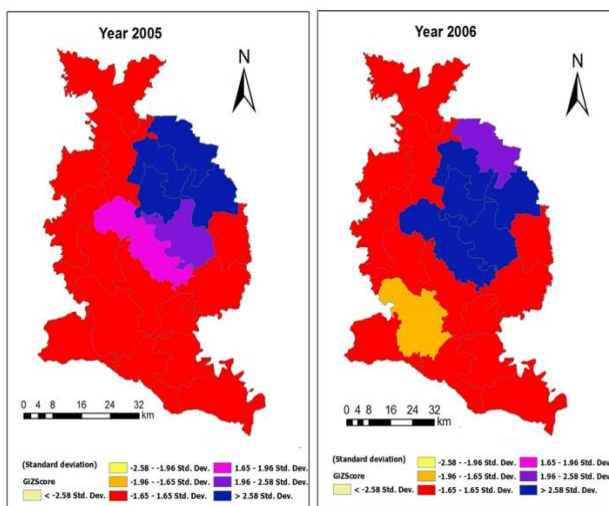




Figure 5: Hotspots of JE cases of years 2005-2010

JE cases map is made from the collective number of cases for each year in block during the study period. It is confirmed from this study that JE cases are spread all around the study area, showing a hotspot in the middle part, the buildup zone of the study area.

## 6. DISCUSSION

The present results indicate that spatial mapping and temporal analysis can play an important role in the distribution of JE disease in epidemiology and control the disease geographically. The spatial association between cases is subject to the measure of spatial proximity to visualize the closeness of disease as compared to a conventional method. Based on these analyses, a health planner will be able to perceive the clustering of areas with high detection rate of cases. Temporal analysis of the JE disease suggested that the annual incidence rates is fluctuating considerably, with the peak incidence rate in 2010.

Quantitative spatial analysis technique using Moran's I and Getis-Ord Gi*(d) statistics demonstrated the spatial distribution pattern clusters and identified hotspots of JE cases in Gorakhpur district. The positive Moran's I values of the spatial autocorrelation in disease distribution indicates that the disease is accumulated in some particular part of the area. However, in our analysis, 0.09 is the highest value of Moran's I statistics. The result of analysis of the JE disease shows a high cluster

pattern in the urban areas of the study area. The high intensity of JE disease occurrence may be attributed to the conducive surrounding factor like high agriculture practices, presence of wetland areas such as Ramgarh, Jhacia, and Kwada tal and high pig population density which in turn increase the susceptibility to JE disease.

## 7. CONCLUSION

The result of the GIS based study makes it possible to measure high and low risk areas of JE incidence in a particular blocks of the Gorakhpur district. The visualized JE maps showed that the proposed methods and tools namely Moran's I Index (Moran's I) and Getis-Ord Gi*(d) can be beneficial for public health officers to understand the patterns of diseases.

The present analysis is performed based on the block-wise data of the area. It will be better if the proposed methodology is adopted to perform the analysis at village level in the district.

The present study shows that GIS based spatial statistical techniques may provide an opportunity to clarify and quantify the outbreak of JE disease high endemic areas.

## REFERENCES

1. WHO. (2011). Annual Report of Child health Division, Department of Health Services IPD Section.

2. Bhunia, Shankar. Gauri., Shreekant, Kesari., Nandini, Chatterjee., Kumar, Vijay. and Pradeep, Das. (2013) Spatial and temporal variation and hotspot detection of kala-azar disease in Vaishali district (Bihar), India. BMC Infectious Diseases, 13:64.

3. Fischer, M., Hills, S., Staples, E., Johnson, B., Yaich, M. and Solomon, T. (2008). Japanese encephalitis prevention and control: advances, challenges, and new initiatives. In: Scheld WM, Hammer SM, Hughes JM. Emerging infections, ASM Press., Washington DC, pp 93–124.

4. Fischer, M., Lindsey, N., Staples, JE. and Hills, S. (2010). Centre for Disease Control (CDC) and Prevention. Japanese encephalitis vaccines: recommendations of the Advisory Committee on Immunization Practices (ACIP). Morbidity and mortality weekly report. Recommendations and reports 59 (RR-1): 1–27.

5. Getis, A. and Ord, JK. (1992). The analysis of spatial association by use of distance statistics. Geographical Analysis, 24: 189–206.

6. Gupta, R.D., Garg, P.K. and Arora, M. (2000). Use of statistical techniques for inter-block disparities analysis under GIS environment, Journal GIS India, Vol. 9, No.6.

7. Hu, W., Tong, S., Mengersen, K. and Oldenburg, B. (2007). Exploratory analysis of social and environmental factors associated with the incidence of Ross River viruses in Brisben, Australia.Am J Trop Med Hyg, 76:814–819.

8. Hayran, Mutlu. (2004). Analyzing factors associated with cancer occurrence: A geographical systems approach. Turkish Journal of Cancer Volume 34, No. 2, 2004.

9. Jeefoo, Phaisarn., Tripathi, Kumar. Nitin. and Marc, Souris. (2011). Spatio-Temporal Diffusion Pattern and Hotspot Detection of Dengue in Chachoengsao Province, Thailand. Int. J. Environ. Res. Public Health, 8, 51-74.

10. Kheirandish, S., Liaghat, M., Azahar, TM., and Gohari, A. (2012) Comparison of interpolation methods in prediction the pattern of basal stem rot disease in palm oil plantation. Geoinformatica - An International Journal (GIIJ), 2(1):12–16.

11. Khinchi, R. Y., Kumar, A., and Yadav, S. (2010). Study of acute encephalitis syndrome in children. Journal of College of Medical Sciences-Nepal, Vol. 6, No. 1, 7-13.

12. Leung, Y., and Leung, K. S. (1993). An intelligent expert-system shell for knowledge-based geographical information systems. 1. The tools. International Journal of Geographical Information Systems, Vol. 7, pp. 189-199.

13. Moore, A. Dale., and Carpenter, E. Tim., (1999). Spatial Analytical Methods and Geographic Information Systems: Use in Health Research and Epidemiology. Epidemiologic. Reviews Vol. 21, No. 2.

14. Muheneni, Rao. Srinivasa., Upadhyayula, Murty. Suryanarayana., and Natarajan, Arunachalam. (2014). Prevalence of japanese encephalitis and its modulation by weather variables. Academic Journals, Vol. 6(1), pp. 52-59, January.

15. Osei, FB., and Duker, AA., (2008) Spatial and demographic patterns of Cholera in Ashanti region-Ghana. Int J Health Geogr, 7:44.

16. Pathirana, Sumith., Kawabata, Masato. and Goonatilake Rohitha. (2009). Study Of potential risk of dengue disease outbreak in sri lanka using GIS and statistical modelling. Journal of Rural and Tropical Public Health, Vol 8, p. 8-17.

17. Sabesan, Shanmugavelu., Konuganti, Raju. Hari, Kishan. and Perumal, Vanamail., (2008). Spatial Delimitation, Forecasting and Control of Japanese Encephalitis: India – A Case Study. The Open Parasitology Journal, 2008, 2, 59-63.

18. Saxena, K. Shailendra., Mishra, Niraj., Saxena, Rakhi., Singh, Maneesh. and Mathur, Asha. (2009). Trend of Japanese encephalitis in North India: evidence from thirty-eight acute encephalitis cases and appraisal of niceties. J Infect Dev Ctries 2009; 3(7):517-530.

19. Scott, M. Lauren. and Janikas V. Mark. (2010). Spatial Statistics in ArcGIS. Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications.

20. Singh, Arvind. (2007). Menace of Japanese encephalitis inrural areas of eastern Uttar Pradesh. Current Science, Vol. 93, No. 12, 25 December.

21. Srivastava, V.K., Ajay, Singh. and Thapar B.R., (2008). Field evaluation of malathion fogging against Japanese encephalitis vector, Culex tritaeniorhynchus. J Vector Borne Dis 45, September 2008, pp. 249–250.

22. Upadhyayula, Murty. Suryanarayana., Mutheneni, Rao. Srinivasa., Nayanoori, Hari. Krishna., Natarajan, Arunachalam. and Prashant, Goswami. (2012). Impact of weather variables on mosquitoes infected with Japanese encephalitis virus in Kurnool district, Andhra Pradesh. Asian Pacific Journal of Tropical Medicine 337-341.