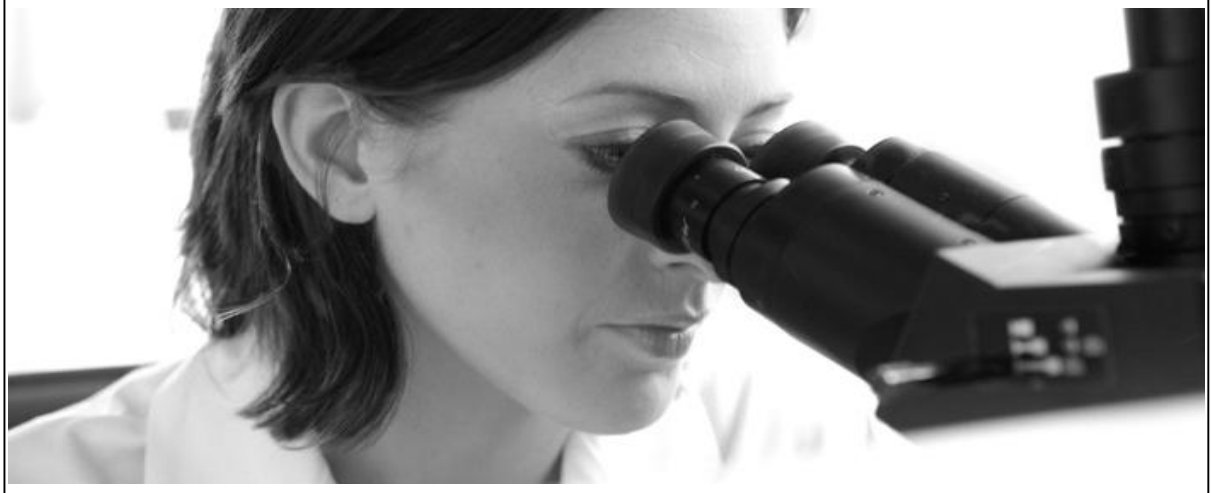
ISSN: 0976-3031

International Journal of Recent Scientific Research

Impact factor: 5.114

GEOSPATIAL ANALYSIS FOR VILLAGE LEVEL SOCIAL PROFILING: A CASE STUDY OF SOLAPUR DISTRICT



Yedage A S and Baviskar S.P

Volume: 6 Issue: 10

THE PUBLICATION OF
INTERNATIONAL JOURNAL OF RECENT SCIENTIFIC RESEARCH
http://www.recentscientific.com
E-mail: recentscientific@gmail.com

Available Online at http://www.recentscientific.com

ISSN: 0976-3031

International Journal of Recent Scientific Research
Vol. 6, Issue, 10, pp. 6815-6820, October, 2015

International Journal of Recent Scientific Research

RESEARCH ARTICLE

GEOSPATIAL ANALYSIS FOR VILLAGE LEVEL SOCIAL PROFILING: A CASE STUDY OF SOLAPUR DISTRICT

Yedage A S and Baviskar S.P

Department Of Environmental Science, School of Earth Sciences, University of Solapur, Kegaon, Solapur, Maharashtra, India

ARTICLE INFO

Article History:

Received 15th July, 2015
Received in revised form
21st August, 2015
Accepted 06th September, 2015
Published online 28st
October, 2015

Key words:

GIS, Spatial Autocorrelation, Getis-Ord Gi*, Hot Spot

ABSTRACT

Geographic Information Systems (GIS) and location-based services represent the next golden age of maps and cartography. In this paper, we consider some of the technological changes like use of GIS, which could be useful for generating a system for village level information in India. For the sustainable development of the district as a whole, the primary requirement is to collect the information at village level for spatial statistical analysis related to demographic parameters like male/female ratio and Schedule caste population etc. Study region include 11 talukas and 1151 villages and 11 urban areas in Solapur district. For Spatial Statistics, both the techniques viz; Spatial Autocorrelation and Getis-Ord Gi*a facility to identify trends, patterns and clustering of the data was used. Hot Spot analysis of Scheduled Caste population showed a major Hot Spot in the west parts of Solapur district and strong Cold Spot in the North East parts of the district. The male/female ratio indicates the highway influence and migration of male population towards urban centres as there is a Cold spot around the Solapur city. A flexible and user friendly information system was developed to assist planners for village level planning with reference to managing the resources.

INTRODUCTION

Geographic Information Systems (GIS) and location-based services represent the next golden age of maps and cartography. The tools fueling this age are computers, the Internet, and the Global Positioning System (GPS). Present era is known as Information Technology (IT) era. Now a day, the information about any phenomena related to human beings and their activities is getting importance. This information is essential for different administrative departments for planning and development process. Developments in information and communication technologies (ICTs) during the last quarter of the 20th century heralded an information age in which economic and social activity has been widened, deepened and transformed. The more optimistic projections suggest that a computerized and networked world would not only ensure a more widespread and rapid growth of employment, productivity and output, but would also improve access to facilities that enhance the quality of life. In this project we consider some of the technological changes like use of GIS, which could be useful for generating a system for village level information in developing countries like India. Geographical Information Systems and the underlying Spatial Data Infrastructures have potential to assist in planning, (Adinarayana *et al* 2004)

The potential contribution of GIS to Village Level Information System

The expectations that GIS generate village improvement in developing countries stem from three sources. The first is their role as an instrument for continuing education and lifelong learning in developing countries, to be informed about and trained in the use of advances in knowledge. The second is their use as a delivery mechanism to poor and remote locations of a wide variety of services varying from improved public information education to emergency advice, including advice on dealing with and mitigating the consequences of natural disasters. The third source is their potential use as a mechanism to increase the transparency and efficiency of governance which would, in turn, improve the available and delivery of publicly provided information services. Geographical Information System is a computer aided tool for generating, storing, manipulating, retrieving at will and displaying the spatial and non spatial data. It has benefited different location based problems, like disaster management, crime analysis, etc. It has been useful for analyzing the trends and patterns of various spatially distributed phenomena. A well-organized information system through the DISNIC programme of the National Informatics Centre (NIC), Ministry of Communication and Information Technology, Government of India (http://www.nic.in), has emerged in the country at

Corresponding author:* **Yedage A S
Department Of Environmental Science, School of Earth Sciences, University of Solapur, Kegaon, Solapur, Maharashtra, India

district/sub-district level to facilitate development planning, responsive administration and promote informatics culture at the district/sub-district level.

Need of Study

For the sustainable development of the district as a whole, the primary requirement is to collect the information at village level. The administrative officers and planners are in a need of an information system for understanding the current situation in the district. The facilities, market places, health centers, education and economic activities at the village as a unit will give them an idea of the requirements of villages (Adinarayana *et al* 2002).

Now a day Geoinformatics is effectively applied in various fields like tourism, health facilities, utilities, navigation etc. Information facility is one of prime and important sector. It is related to everyone in the society i.e. student, parents, researchers, teaching community and government officials. If one resourceful information system is developed by considering all essential and important aspects of information facilities, that will be very useful for all. Planners need to have at their hands a sophisticated data Management system to handle huge spatially correlated data.

The emergence of Remote Sensing and Geographic Information System as a powerful tool for spatial analysis and storage has in effect alleviated the problem by computerization of the spatial data. This new technology can reduce the time and cost to the planners in organizing the data in arriving at precise conclusion and decisions (Ravindra *et al* 2007).While generating the information system at village as the smallest unit, all basic needs in village are taken into consideration; such as health and educational facilities, and location based facilities, and the information about to population like total population, male population, female population, male female ratio, etc.

Study Area and Data Sources

Study Area: The study area of the project is six districts of Maharashtra State which are coinciding with the *Hot dry Semi-arid Agro ecological zone* of the State. This is the most vibrant area as per the climatic conditions. In the western zone i.e. Pune, Satara, Ahmednagar cities with their surroundings, are the only industrial tracts in the study area. Other region comprises of agriculture and related activities as a major economic activity. The Solapur district was assigned to our group (Figure No.2.1).

Location: Solapur district is situated in the south part of the state. Geographically this lies between 75°56' 45.44'' East Longitude and 17° 41' 13.86'' North Latitude. Solapur is surrounded by pune Ahmadnager and Osmanabad in the north, Osmanabad in the east, Satara and Sangli in west and Sangli with the Karnataka State in the South. There are 11 talukas in Solapur district. This district is having different physical and climatic condition across the talukas

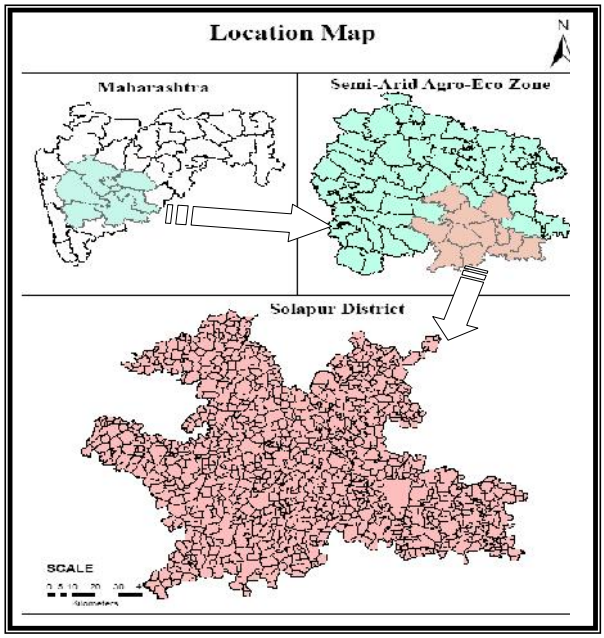


Figure 1 Location Map of the Study Area

Physiographi*c*- The whole part of Solapur district is plain area. These all part comes under the rain shadow area. The district is drained by three main rivers namely Bhima, Man and Sina. Bhima is the largest river which is flowing through central parts of the district hence has fertile soils and therefore is highly populated.

Climatic condition - The climate of the Solapur district is hot and dry, on whole extremely amiable and is characterized by a hot summer and general dryness during major part of the year except during south-west monsoon season. In the cold season which lasts from November to February, the air is dry and invigorating. The period from March to the first week of June is the hot season. It is followed by the south-west monsoon season which lasts till the end of September. October and November constitute the post-monsoon or the retreating south-west monsoon season. The western hilly region receives more rainfall. The average annual rainfall in the district is 501.800 mm. (19.76"). May is the hottest month & December is the coldest month. On few occasions thin film of fog is being observed in the early mornings of December & January.

Demography - As per census 2001 Solapur is having population 2624259. Males constitute 51.96% of the population and females 48.05%.

Data Used

The data is considered as a heart of the project. Data and Methodology are the important components of GIS. The selected topic for the study is village level Information system; hence it contains variety of data regarding villages. Name of the village, its population, facilities of village etc all information regarding of village come under village level information. The important and basic information about educational center means colleges and schools also come in this study. The data can be easily categorized in spatial and non- spatial or attribute data. As the study area of the project is

vast and spreads across 6 districts. Hence all possible data is collected through available sources; hence only secondary data is used in the project.

Spatial data –Vector map of Maharashtra at taluka level Maharashtra State Library at 1:100000 scale. Raster maps of Talukas having village as the smallest unit from Census 2001.

Non Spatial data **–**District census data, Census 2001 was collected from census office, Mumbai. The census data is categorized in the following groups; Population, education, finance, health, agriculture, etc.

Software Used

All soft wares are having their own positive and negative points but all are having near about same functionality as well as capability**.** The GIS software has all functions of digitization, georeferencing, attribution and analysis etc. In this project commercial as well as open source soft wares are used as per the need of project. The list of soft wares and some tasks are mentioned below.

GIS Software

1. ***ArcGIS 9.2*** - Georeferencing, Registration of maps, Projection transformation Data attachment, Query Processing and analysis using Spatial Statistics tools
2. ***Kosmo GIS 1.2*** - Digitizing and creation of Layers
3. ***MySQL*** - database server for generating, editing and storing village boundary database using Kosmo GIS in multi-user concurrent edition mode.

Other Software

1. ***Microsoft Excel*****:** Data Tabulation, Calculations, Graphical presentation etc
2. ***Microsoft Word*****:** Write-up and Formation of Project.

METHODOLOGY

Mapping work

1. Georeferencing the Maps – All the scanned raster maps of talukas were georeferenced with the help of Maharashtra vector map using ArcGIS 9.2 soft ware. The rasters were reprojected in UTM projection using WGS 84 datum.
2. Creation of Layers – The vector map of 6 districts were clipped and put in the MySQL server for editing. It was being accessed using Kosmo GIS 1.2 software by various users at the same time. Using this server application, the village level map of all six districts was generated.

After the generation of village level vector map, the village codes from Census village directory data were assigned to each polygon for their unique identification i.e. primary key.

Then, it was converted to a shape file using Kosmo 1.2 soft ware. Using Arc GIS software, the final vector was made error free by generating topology (generation of mathematical relationship between geographical features).

The vector map was divided into 6 parts as per the districts. The district vector was then assigned to the groups for further analysis (figure 1).

Analysis

Some of the Census parameters were analyzed using some spatial statistical tools (Spatial autocorrelation, Hot spot analysis) in Arc GIS 9.2 Soft ware.

The parameters are given below

1. Scheduled Caste Population
2. Scheduled Tribes population
3. Male / Female Ratio
4. Population Density

Spatial Analysis

Spatial Statistics has a facility to identify trends, patterns and clustering of the data. In this study, there were basically 2 techniques used –

1. Spatial Autocorrelation.
2. Hot Spot Analysis.

Spatial Autocorrelation

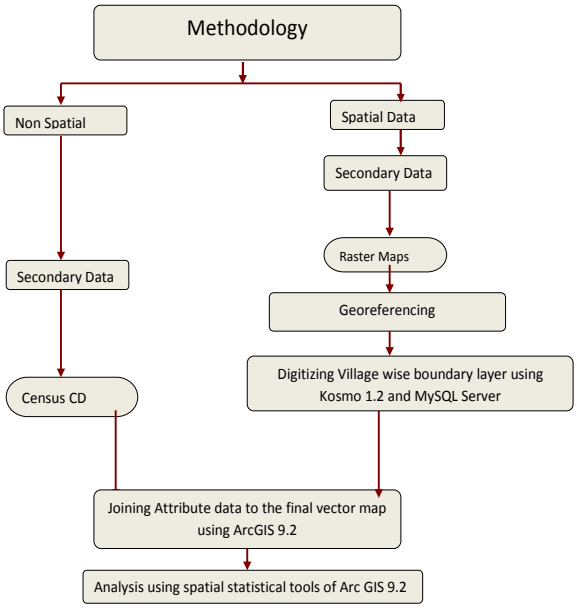
Given a set of features and an associated attribute, Spatial Autocorrelation evaluates whether the pattern expressed is clustered, dispersed, or random.

The tool calculates the *Moran's I Index* value and a Z score evaluating the significance of the index value. In general, a Moran's Index value near +1.0 indicates clustering while an index value near -1.0 indicates dispersion. However, without looking at statistical significance we have no basis for knowing if the Index value is anything more than random chance (Ebdon 1985) (Goodchild 1986) Z scores are measures of standard deviation.

For example, if a tool returns a Z score of +2.5 it is interpreted as "+2.5 standard deviations away from the mean". Z score values are associated with a standard normal distribution. This distribution relates standard deviations with probabilities and allows significance and confidence to be attached to Z scores.

Very high or a very low Z scores are found in the tails of the normal distribution. From the graph above, it is evident that the probabilities in the tails of the distribution are very low. When you perform a feature pattern analysis and it yields either a very high or a very low Z score, this indicates it is very unlikely that the observed pattern is some version of the theoretical spatial pattern represented by your null hypothesis.

Flowchart



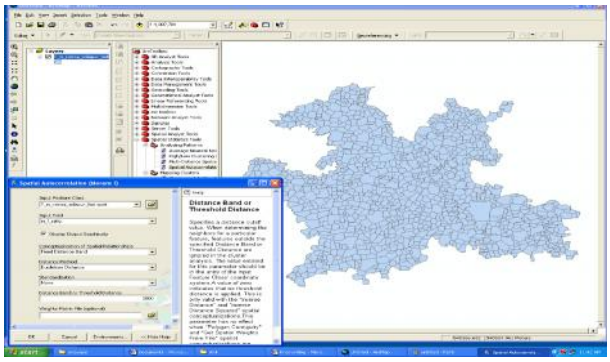


Figure no 1 Performing Spatial Autocorrelation tool in Arc GIS soft ware

In figure 1, the process and output of the spatial autocorrelation is displayed. The input for this analysis are; shape file, parameter which has to be analyzed, distance method (Euclidean distance), and threshold distance. In the output of Spatial Autocorrelation, (Figure 2), it gives Moran's I Index, Z score value and its significance.

Hot Spot Analysis - The Hot Spot Analysis tool calculates the Getis-Ord Gi* statistic for each feature in a weighted set of features. (Goodchild 1986) The G-statistic tells you whether features with high values or features with low values tend to cluster in a study area. This tool works by looking at each feature within the context of neighboring features (Scott 2005). If a feature's value is high, and the values for all of its neighboring features are also high, it is a part of a hot spot. The local sum for a feature and its neighbours is compared proportionally to the sum of all features; when the local sum is much different than the expected local sum, and that difference is too large to be the result of random chance, a statistically significant Z score is the result. The Gi* statistic is actually a Z score. For statistically significant positive Z Scores, the larger the Z score is, the more intense the clustering of high values (Scott 2005)

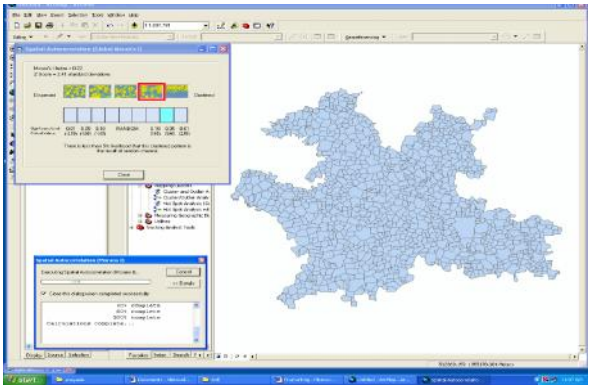


Figure no. 2 Result of Spatial Autocorrelation Tool in Arc GIS software

RESULTS AND DISCUSSION

The district vector maps were analyzed using some statistical tools available in Arc GIS 9.2 soft ware. For this purpose, some of the census parameters like population of scheduled caste and tribes. Even some population analysis like density and male/female ratio were also used for finding their spatial distribution patterns. The process of spatial autocorrelation is repetitively done to identify the largest value of Z score for that specific parameter. It will be further used as an input in the Hot Spot analysis. The distance of 44600 meters (Table 1) gave highest Z score value and is taken as an input for further analysis.

Table 1 Moran's I Index for SC Population (Solapur District)

Distance	Z-Score	Moran's I	Remarks
1100	0.55	0.38	random
1500	1.73	0.31	cluster
2000	2.29	0.2	cluster
3000	5.35	0.99	cluster
5000	8.19	0.14	cluster
10000	12.5	0.1	cluster
15000	14.81	0.09	cluster
20000	14.46	0.06	cluster
30000	17.51	0.05	cluster
35000	18.73	0.05	cluster
40000	20.23	0.04	cluster
42500	*20.56*	*0.04*	*Cluster*
45000	19.07	0.03	cluster
50000	19.74	0.03	cluster
55000	18.13	0.03	cluster
60000	17.85	0.02	cluster
75000	10.76	0.01	cluster
100000	3.27	0	cluster

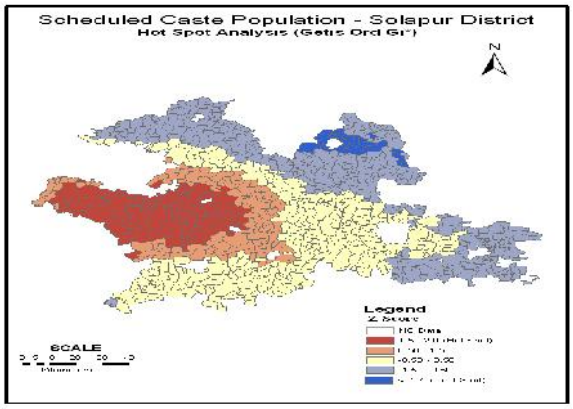


Figure No. 3 Hot Spot analysis for Scheduled Caste Population

In the map (Figure 4.3), the dark red color area shows villages with high Z score values of Scheduled Caste Population is called as Hot Spot .The dark blue color area shows less Z score values of Scheduled Caste Population is called as Cold Spot. The north-East part of Solapur district (most of Barshi taluka) shows low concentration of Scheduled Caste Population and West part shows high concentration of Scheduled Cast Population in the major parts of Malshiras and North part of Sangola talukas. The white color shows not available data for that village.

Male/ Female Ratio

The proportion of females to males is an important issue for the development of the society. In the recent years, the female population is decreasing in comparison with the males very drastically. For this purpose this analysis is taken in this study. Here new attribute has been generated for the district vector. The formula for the male female ratio is (Female population in a specified region/ Male population of that specified region) The Spatial autocorrelation analysis highlighted the distance of 4000 as the threshold distance (table 2) as it holds the highest Z score value of 6.32.

*1000 = female population per 1000 males

Table 2 Moran's I Index for Male / Female Ratio (Solapur District)

Distance	Z Score	Moran's I	Remark	
1100	23.92	16.71	cluster	
1200	13.15	5.3	cluster	
1400	6.24	1.49	cluster	
1500	4.92	0.89	cluster	
2000	2.41	0.22	cluster	
2200	1.71	0.12	cluster	
2300	1.58	0.1	random	
2400	5.28	0.31	cluster	
3000	2.68	0.1	cluster	
4000	6.32	0.14	cluster	Hotspot
5000	5.11	0.09	cluster	
8000	5.81	0.06	cluster	
10000	5.03	0.04	cluster	
12000	5.96	0.04	cluster	
15000	4.17	0.02	cluster	
18000	3.36	0.02	cluster	
20000	3.06	0.01	cluster	
25000	2.39	0.01	cluster	
30000	2.68	0.1	cluster	
40000	2.77	0.01	cluster	
50000	2.62	0	cluster	
60000	2.14	0	cluster	
75000	3.35	0	cluster	

The Male / Female Ratio attribute of Solapur district has been taken for the calculation of Hot Spot analysis, and the Z score output of the same then is classified into 6 classes using standard deviation and mean and mapped (Figure (4). In above map the dark red color area shows villages with high Z score value called as Hot Spot. The Hot spot area is covering most of the south part of Sangola and some part of the North Solapur talukas. This indicates that the female population of those talukas is comparable to the male population. The dark blue color area shows less value of Z score, called as cold area. The central part of Solapur district shows less number of female and therefore it is an alarming area as per the male female proportion is concerned.

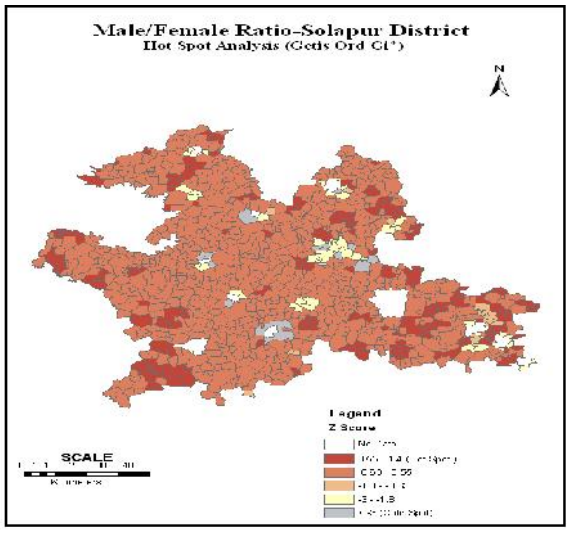


Figure 4 Hot Spot analysis for Male/Female Ratio

CONCLUSIONS

The study in this project is focused on generation of Village information system, and analyzing the data for clustering and pattern recognition. 6 districts in the semi arid agro ecological zone have been studied taking some demographical data. In Solapur district has in total 11 talukas and 1151 villages and 11 urban areas. The technology of Geographical Information System (GIS) has immense capability of handling, analyzing spatial and non spatial data. The whole study had 2 distinct parts i.e. generation of village information system and spatial statistical analysis for some demographic parameters. It gave the following conclusions. Concurrent editing of the vector gave better results than doing the same job separately. Hot Spot analysis of Scheduled Caste population showed a major Hot Spot in the west parts of Solapur district and strong Cold Spot in the North East parts of the district. Hot Spot analysis of male/female ratio indicates the highway influence and migration of male population towards urban centres as there is a Cold spot around the Solapur city i.e. less number of females to 1000 male population and in contrast to those two hot spots in the west depicting the comparable population of females to that of males.The fuzzy approach can be used for site specific analysis and the membership grades can be used for queries for successive suitability rating for the area. The query shell is flexible for an expert user to query for his preset conditions and check for the spatial extent and accordingly suggest plans. Furthermore, the spatial analysis tool of the system is a handy option for the planners to correlate spatial information and hence to zero in areas of the planners interest, such site suitability assessment for and developmental activities, in this project report used different analysis method and done this analysis part in ArcGIS Software.

Limitations

- Time did not permit to add other additional layers like Road, railways, etc.
- Analysis of demographic parameters could not be validated using other sources of data.

- Due to lack of time the coverage of area was limited to only the semi arid agro ecological zone of the Maharashtra State.

References

1. Adinarayana, J., Raj, F. J., and Sharma, V. (2004). Village level information system – a tool for decentralized planning at district level in India, *International Journal of Environmental Informatics*, 4(2): 56-64
2. Adinarayana, J., Maitra, S. and Dent, D. (2000). A spatial decision support system for land use planning at district level in India. The Land, 4(2), 111-130.
3. Census of India, Solapur District Census Handbook, Village & Town Directory (2005)., Village & Town wise Primary Census Abstract, Series 14, Maharashtra, Director of Census Operations, Mumbai, Maharashtra
4. Burrough, P.A. (1986). Principles of Geographical Information System for Land Resources Assessment, Clarendon Press - Oxford, England
5. Campbell, H. And Masser, I. (1995). *GIS and Organizations, Taylor and Francis*, London.Ebdon, David. (1985).Statistics in Geography, Blackwell.
6. Goodchild, M.F. (1986). *Spatial Autocorrelation*, Catmog 47, Geo Books.Mitchell, Andy. (2005). The ESRI Guide to GIS Analysis, Volume 2, ESRI Press.
7. Maji A., Krishna, N.D.R., Sehgal, J and Velayutham, M. (1998). "Dissemination of Soil information for district level planning using GIS", GIS India (21-25)
8. Ravidra, A. and Jaishankar, J (2007).GIS based information system for village level planning, access on 22 December 2012 NIC (National Informatics Centre). Ministry of Communication and Information Technology, Government of India, New Delhi.
9. Scott, L. and Warmerdam, N. (2005). Extend Crime Analysis with ArcGIS Spatial Statistics Tools in Arc User Online

How to cite this article:

Yedage A S and Baviskar S.P.2015, Geospatial Analysis for Village Level Social Profiling: A Case Study of Solapur District. *Int J Recent Sci Res.* 6(10), pp. 6815-6820.

****

International Journal of Recent Scientific Research

ISSN 0976-3031