Spatial Data Mining and Modeling for Visualisation of Rapid Urbanisation

*Uttam Kumar[1], Mukhopadhyay C[2]., and Ramachandra T. V[3]**

[1] Department of Management Studies and Centre for Sustainable Technologies, Indian Institute of Science, Bangalore
[2] Department of Management Studies, Indian Institute of Science, Bangalore
[3] Centre for Ecological Sciences, Centre for Sustainable Technologies, Centre for Infrastructure, Sustainable Transportation and Urban Planning Indian Institute of Science, Bangalore, India
Email: [1]uttam@ces.iisc.ernet.in, [2]cm@mgmt.iisc.ernet.in, [3]cestvr@ces.iisc.ernet.in

Abstract

Rapid urbanisation in India has posed serious challenges to the decision makers in regional planning involving plethora of issues including provision of basic amenities (like electricity, water, sanitation, transport, etc.). Urban planning entails an understanding of landscape and urban dynamics with causal factors. Identifying, delineating and mapping landscapes on temporal scale provide an opportunity to monitor the changes, which is important for natural resource management and sustainable planning activities. Multi-source, multi-sensor, multi-temporal, multi-frequency or multi-polarization remote sensing data with efficient classification algorithms and pattern recognition techniques aid in capturing these dynamics. This paper analyses the landscape dynamics of Greater Bangalore by: (i) characterisation of direct impervious surface, (ii) computation of forest fragmentation indices and (iii) modeling to quantify and categorise urban changes. Linear unmixing is used for solving the mixed pixel problem of coarse resolution super spectral MODIS data for impervious surface characterisation. Fragmentation indices were used to classify forests – interior, perforated, edge, transitional, patch and undetermined. Based on this, urban growth model was developed to determine the type of urban growth – Infill, Expansion and Outlying growth. This helped in visualising urban growth poles and consequence of earlier policy decisions that can help in evolving strategies for effective land use policies.

Keywords: *Landscape, Orthogonal subspace projection, Forest fragmentation, Urban growth model*

1. Introduction

Urbanisation is a form of metropolitan growth in response to incomprehensible combination of economic, social, and political forces and to the physical geography of an area. This could be planned (in the form of townships) or unplanned (organic) Many organic cities are now undergoing redevelopment for economic purposes with new roads, infrastructure improvements, etc. It results in the increase in population, in proportion to the region's rural population. This phenomenona is very rapid in India with urban population growing at around 2.3 percent per annum The 21st century is witnessing "rapid urbanisation of the world's population", as the global proportion of urban population rose dramatically from 13% (220 million in 1900) to 29% (732 million, in 1950) to 49% (3.2 billion, in 2005) and is projected to rise to 60% (4 9 billion) by 2030 [1] An increased urban population and growth in urban areas is inadvertent with dramatic increase in population mainly due to migration. There are 35 urban agglomerations/cities having a population of more than one million in India (in 2001). Of the 4000 plus urban agglomerations, about 38 percent reside in just 35 urban areas, thus indicating the magnitude of urbanisation. Overall rise in population of urban poor or increase in travel times owing to congestion in road networks are indicators of urban governance and planning. As Indian cities continue to urbanise, the land use (LU) plays a determining role in the quality of land, water and air environment This pose serious challenges to the decision makers in the city planning and management process involving plethora of issues like infrastructure development, traffic congestion, and basic amenities (such as electricity, water, and sanitation), floods, heat island, loss of aquatic eco-system, carbon footprint, enhanced level of pollution, higher instances of diseases, etc. [2] Thus, the administration at all levels: local bodies, federal and national governments are facing the brunt of rapid urban growth. It is imperative for planning and governance to facilitate, augment and service the requisite infrastructure over time systematically. This requires an understanding of landscape characterisation focusing on agents of urban dynamics.

Problem Statement – Identifying, delineating and mapping landscapes on temporal scale provide an opportunity to monitor the changes, which is important for natural resource management and sustainable planning activities. Multi-source, multi-sensor, multi-temporal, multi-frequency or multi-polarization remote sensing (RS) data [3] with pattern recognition techniques [4] aid in capturing the dynamics. Pattern classification and modeling are very useful in landscape characterisation for addressing the problems of sprawling cities and urban management. The objectives of this work are to understand the landscape dynamics in Greater Bangalore through visualization by i) modeling impervious surface, ii) computation of forest fragmentation indices, iii) quantification and categorisation of urban changes and (iv) modeling of urban dynamics.

Visuaisation - Current Trend and Techniques: Geovisualisation helps to explore real-world environments and model 'what if' scenarios based on spatio-temporal data. This aids as a tool for modeling the environmental interests in policy planning. Some of the well known applications are 3D photorealistic representations to simulate urban redevelopment and possible pollution diffusion over the next few years. Techniques used in geovisualisation are agent based cellular automata models, genetic algorithms, etc.

2. Methods

2.1 Orthogonal Subspace Projection (OSP) – Pixel usually contain contributions from more than one class, when object size is smaller than the pixel resolution (mixed pixels) except in higher spatial resolution data. Standard classification techniques, which attribute a single class to the entire pixel, are therefore inappropriate to low spatial resolution satellite images from space-borne sensors This motivates the development of algorithms which unmix the coarse spatial data or, in other words, perform classification at a sub-pixel level. The main objective is to find out the proportion of each category in a given pixel, or in other words, unmix the pixel to identify the categories present within the pixel. In order to address this problem, unmixing techniques have been developed to exploit sub-pixel level information for image analysis [5], where sub-pixel class composition is estimated through the use of techniques, such as linear mixture modeling [6], supervised fuzzy-c means classification and artificial neural networks, etc

Linear unmixing is based on an assumption that the spectral radiance measured by the sensor consists of the radiances reflected collectively in proportion to the sub-pixel area covered by each material. Let K be the number of spectral bands in the multispectral / superspectral data set, and P, the number of distinct classes of objects in the physical scene. Associated with each pixel is a K-dimensional vector **y** whose components are the gray values corresponding to the K bands. Let $\mathbf{E} = [e_1, e_2, \ldots, e_P]$, where $\{e_i\}$ is a column vector representing the spectral signature of the i^{th} target material or category. The column vectors of the K x P matrix **E** are called end-members. For a given pixel, the abundance fraction of the p^{th} target material present in a pixel is denoted by α_p, and these values are the components of the P- dimensional abundance vector α. Assuming the linear mixture model, the observation vector **y** is related to **E** by

$$y = E\alpha + \bar{\eta} \quad (1)$$

where $\bar{\eta}$ accounts for the measurement noise. We further assume that the components of the noise vector $\bar{\eta}$ are zero-mean random variables that are independent and identically distributed. Therefore, the covariance matrix of the noise vector is $\sigma^2\mathbf{I}$ where σ^2 is the variance, and **I** is K x K identity matrix. The conventional approach [7] to extract the abundance values is to minimise $\| y - E\alpha \|$, and the estimate for the abundance is

$$\alpha = (E^T E)^{-1} E^T y \quad (3)$$

where,

$$\lambda = \frac{2(1^T(E^TE)^{-1}E^Ty\text{-}1)}{1^T(E^TE)^{-1}1} \quad (4)$$

Chang (2005) [8] came up with a technique called Orthogonal Subspace Projection (OSP) in the signal processing domain, which is used here as a classification technique based on two aspects: 1) how to best utilise the target knowledge provided *a priori* and 2) how to effectively make use of numerous spectral bands. Briefly, the technique involves (i) finding an operator which eliminates undesired spectral signatures, and then (ii) choosing a vector operator which maximises the signal to noise ratio (SNR) of the residual spectral signature. In order to find the abundance of the p^{th} target material (α_p), let the corresponding spectral signature of the desired target material be denoted as **d**. The term **Eα** in equation (1) can be rewritten to separate the desired spectral signature **d** from the rest as:

$$E\alpha = \mathbf{d}\alpha_p + \mathbf{R}\mathbf{r} \quad (5)$$ where

r contains the abundance of the rest of the end-members, and **R** is a K x P - 1 matrix containing the columns of **E** except for the column vector **d**. We rewrite (1) as

$$\mathbf{y} = \mathbf{d}\alpha_p + \mathbf{R}\mathbf{r} + \bar{\eta} \quad (6)$$

The interfering signatures present in **R** can be removed from (6) by the operator,

$$P = (\mathbf{I} - \mathbf{R}(\mathbf{R}^T\mathbf{R})^{-1}\mathbf{R}^T) \quad (7)$$

which is used to project the vector **y** into a space orthogonal to the space spanned by the interfering spectral signatures. Therefore, operating on **y** with **P**, and noting that **PR** = 0, we get

$$Py = P\mathbf{d}\alpha_p + P\bar{\eta} \quad (8)$$

The next step is to find an operator w^T which maximizes the SNR given by

$$SNR = \frac{(w^TP\mathbf{d}\alpha_p)^2}{\varepsilon\{(w^TP\bar{\eta})^2\}} = \frac{w^TP\mathbf{d}\alpha_p^2\mathbf{d}^TP^Tw)^2}{w^TP\varepsilon\{(\bar{\eta}\bar{\eta}^T)P^Tw} \quad (9)$$

$$= \left(\frac{\alpha_p^2}{\sigma}\right)^2 \frac{w^TP\mathbf{d}\mathbf{d}^TP^Tw}{w^TPP^Tw}$$

Maximizing the SNR leads to the generalized eigenvalue problem: $P\mathbf{d}\mathbf{d}^TP^Tw = \lambda PP^Tw$

The eigenvector corresponding to the maximum eigenvalue is the vector '**w**'. It can be shown that the **w** which maximises the SNR is given by

w = k**d**. (10)

Therefore, an optimal estimate of α_p is given by

$$\hat{\alpha}_p = \frac{y^TP^TPy}{\mathbf{d}^TP^TP\mathbf{d}} \quad (11)$$

In the absence of noise, the estimate matches with the exact value as in (6). The value of α is the abundance of the p^{th} class (in an abundance map) ranging from 0 to 1 in any given pixel. 0 indicates absence of a particular class and 1 indicates full presence of that class in that particular pixel. There are as many abundance maps as the number of classes. Intermediate values between 0 and 1 may represent a fraction of that class. If the objective is to map the impervious surface, then OSP renders an impervious surface abundance map, where each abundance pixel shows the proportion of impervious material (buildings / houses / roads / paved surfaces, etc.) in that pixel.

2.2 Forest fragmentation – Forest fragmentation is the process whereby a large, continuous area of forest is both reduced in area and divided into two or more fragments. The decline in the size of the forest and the increasing isolation between the two remnant patches of the forest has been the major cause of declining biodiversity [9]. The primary concern is direct loss of forest area, and all disturbed forests are subject to "edge effects" of one kind or another. Forest fragmentation is of additional concern, insofar as the edge effect is mitigated by the residual spatial pattern [10].

Land cover (LC) map indicate only the location and type of forest, and further analysis is needed to quantify the forest fragmentation. Total extent of forest and its occurrence as adjacent pixels, fixed-area windows surrounding each forest pixel is used for calculating type of fragmentation. The result is stored at the location of the centre pixel. Thus, a pixel value in the derived map refers to between-pixel fragmentation around the corresponding forest location. As an example [11] if *Pf* is the proportion of pixels in the window that are forested and *Pff* is the proportion of all adjacent (cardinal directions only) pixel pairs that include at least one forest pixel, for which both pixels are forested then, Pff estimates the conditional probability that, given a pixel of forest, its neighbour is also forest. The six fragmentation model that identifies six fragmentation categories are: (1) interior, for which Pf = 1.0; (2), patch, Pf < 0.4; (3) transtitional, 0.4 < Pf < 0.6; (4) edge, Pf > 0.6 and Pf-Pff > 0; (5) perforated, Pf > 0.6 and Pf-Pff < 0, and (6) undetermined, Pf > 0.6 and Pf = Pff.

When Pff is larger than Pf, the implication is that forest is clumped; the probability that an immediate neighbour is also forest is greater than the average probability of forest within the window. Conversely, when Pff is smaller than Pf, the implication is that whatever is nonforest, is clumped. The difference (Pf - Pff) characterizes a gradient from forest clumping (edge) to nonforest clumping (perforated). When Pff = Pf, the model cannot distinguish forest or nonforest clumping. The case of Pf = 1 (interior) represents a completely forested window for which Pff must be 1.

Forest fragmentation indices [12] have two parts: (1) Total forest proportion (TFP) and Forest continuity (FC). TFP provides a basic assessment of forest cover in a region ranging from 0 to 1. FC value examines only the forested areas within the analysis region. Its measure specifically utilises the results from the forest fragmentation model.

$$TFP = \frac{total\ forest\ area}{total\ non\text{-}water\ area} \quad (12)$$

$$FC = \frac{\text{weighted area}}{\text{total forest area}} - \frac{\text{area of largest interior forest patch}}{\text{total forest area}} \quad (13)$$

Weighted values for the weighted forest area (WFA) are derived from the median Pf value for each fragmentation class (equation 14):

$$WFA = (1\ 0 * \text{interior}) + (0\ 8 * (\text{perforated} + \text{edge} + \text{undetermined}) + (0\ 5 + \text{transitional}) + (0\ 2 * \text{patch}) \quad (14)$$

The rationale is that, given two regions of equal forest cover, the one with more interior forest would have a higher weighted area, and thus be less fragmented. To separate further regions based on the level of fragmentation, the weight area ratio is multiplied by the ratio of the largest interior forest patch to total forest area for the region. FC ranges from 0 to 1.

2.3 Urban Dynamics Characterisation – An urban dynamics model has spatially detailed data with fine spatial grain, examines the whole landscape and assesses urban growth in all areas, avoids spatial averaging, maintains spatial pattern and configuration, and is consistent over time. The urban growth model used here is based on the modified forest fragmentation model developed by Ritters *et al.*, (2000) [13]. The urban dynamics model uses urban and non-urban classes to create fragmentation image instead of forest and non-forest in forest fragmentation model. Input data consist of multi-dates LC with a minimum of three classes; urban (developed), non-urban (non-developed), and water. Instead of using a forest versus non-forest binary image to create a forest fragmentation map, the first step of the urban growth model uses a non-developed versus developed image to create a "non-developed" fragmentation image. The single date fragmentation map consist of three "fragmented" classes that are combination of the five classes previously described (interior, patch, transitional, edge and perforated). The first is interior, which occurs when all pixels in a *3 x 3* window are non-developed. The second class is perforated, which occurs when between 60 % and less than 100 % of pixels in a *3 x 3* window are non-developed. The final class is patch, which occurs when fewer than 60 % of pixels in a *3 x 3* window are non-developed. Two dates of fragmentation maps are used to create change map. There are three types of change classes – the first type consists of no change classes (including developed, water and interior), the second type includes improbable changes, likely due to classification error, and the third type consists of classes that represent urban growth. The classes that indicate urban growth are outlined in table 1 along with their corresponding urban growth classes. The change classes determine the type of urban growth. *Infill* is the development of a small area surrounded by existing developed land. *Expansion* is the spreading out of urban LC from existing developed land. *Outlying growth* is an interior pixel that changes to developed, and is further classified as either *isolated*, *linear branching*, or *clustered branching* (see table 1).

Table 1: Change classes that represent urban growth and the corresponding growth types [13]

Significant Change Classes	Type of Growth
Patch to Developed	Infill Growth
Perforated to Developed	Expansion Growth
Interior to Developed	Outlying Growth: *1. Isolated Growth,* *2. Linear Branching Growth,* *3. Clustered Branching Growth*

Their distinction is made using a set of rules. An isolated growth is defined as a new, small area of construction surrounded by non-urban land and some distance from other developed areas. A linear branching growth is a road, corridor, or linear development surrounded by non-urban land and some distance from other urban areas. A clustered branching growth is indicative of a new, large and dense development in a previously undeveloped area. The urban growth model produces an urban growth map, which consists of five types of growth as well as developed land, water, and non-developed land.

3. Study area and Data

Greater Bangalore is the principal administrative, cultural, commercial, industrial, and knowledge capital of the state of Karnataka with an area of 741 square kilometers (sq. km) and lies between the latitudes 12°39'00" to 131°3'00"N and longitude 77°22'00" to 77°52'00"E (figure 1). Bangalore has grown spatially more than ten times (741 sq. km) since 1949 (69 sq. km.) and is the fifth largest metropolis in India currently with a population of about 7 million. The mean annual total rainfall is about 880 mm with about 60 rainy days a year over the last ten years. The summer temperature ranges from 18° C – 38° C, while the winter temperature ranges from 12° C – 25° C. Thus, Bangalore enjoys a salubrious climate all round the year. Survey of India (SOI) toposheets of 1:50000 and 1:250000 scales were used to generate base layers of the city boundary and water bodies. Field data for ground control points (GCPs) and training sites for supervised classification of RS data were collected with a handheld GPS. RS data used are

- Landsat MSS of 1973, TM of 1992, ETM+ of 2000 [downloaded from http://glcf.umiacs.umd.edu/data/]

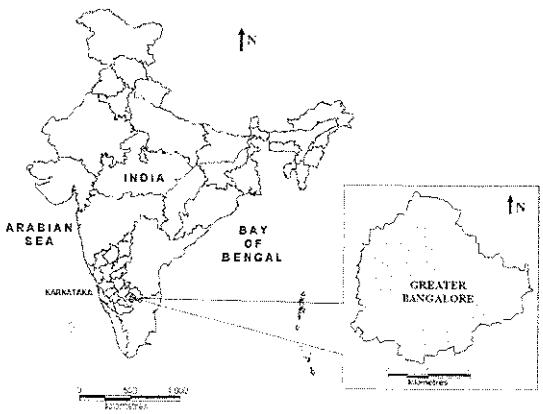

Figure 1: Study area – Greater Bangalore, India. (Self Compiled)

- IRS (Indian Remote Sensing) LISS (Linear Imaging Self Scanner)-III of 1999 and 2006 [procured from NRSA, Hyderabad, India]
- MODIS of 2006, 2007 (Moderate Resolution Imaging Spectroradiometer) Surface Reflectance 7 bands product [downloaded from http://edcdaac.usgs.gov/main.asp]
- Google Earth data (http://earth.google.com) served in pre and post classification process and validation of the results.

4. Results and discussions

4.1 Classification of IRS LISS-III (2006) data: IRS LISS-III data acquired on 28th March, 2006 having 3 spectral bands (Green, Red and NIR) with 23.5 m spatial resolution were georegistered with known GCPs and geometrically corrected and resampled to 25 m (1360 x 1400 pixels) which helped in pixel level comparison of abundance maps obtained by unmixing MODIS data (of 2006) with LISS-III classified map. The class spectral characteristics for four LC categories (buildings/urban, vegetation, water bodies and others) using LISS-III MSS bands 2, 3 and 4 were obtained from the training pixels spectra to assess their inter-class separability and the images were classified using Gaussian Maximum Likelihood classifier with training data uniformly distributed over the study area collected with pre calibrated GPS (figure 2).

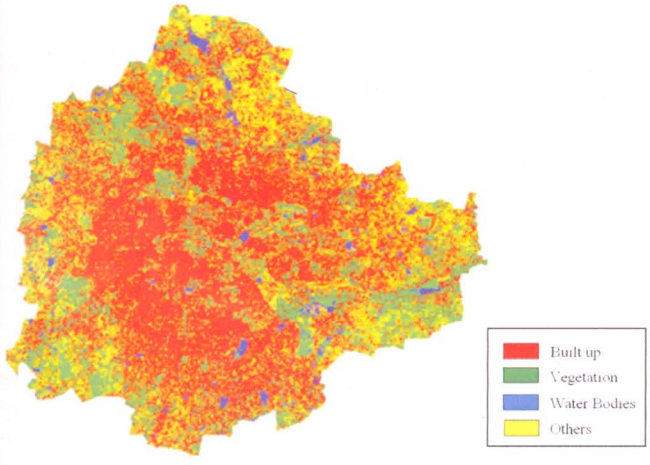

Figure 2: LC classification using LISS-III MSS (2006) (Self Compiled)

Table 2: LC details using LISS -III MSS (Self Compiled)

Class	Area (Ha)	Area (%)	P* (%)	U* (%)	O* (%)
Builtup	29535	43.2	85.2	84.5	84.6
Vegetation	19696	28.8	83.5	88.1	
Water bodies	1073	01.6	90.7	80.2	
Others	18017	26.4	78.3	77.7	

P* - Producer's Accuracy, U* - User's Accuracy, O* - Overall Accuracy

This was validated with the representative field data (training sets collected covering the entire city and validation covering ~ 10% of the study area) and also using Google Earth image (http://www.earth.google.com). LC statistics, producer's accuracy, user's accuracy and overall accuracy computed are listed in table 2. The urban pixels were extracted from the classified image and were used as a benchmark to validate the output obtained from OSP.

4.2 Unmixing of MODIS data using OSP: MODIS data (of 2006) having 7 bands with band 1 and 2 at 250 m and band 3 to 7 at 500 m spatial resolution were georegistered with GCPs, geometrically corrected and bands 3 to 7 were resampled to 250 m (136 x 140 pixels) so that all 7 bands are at same spatial resolution. Unmixing of MODIS data (250 m) were done to obtain abundance map of the impervious surface. Visual inspection as well as accuracy assessment of abundance image showed that OSP algorithm maps impervious surface spatially and proportionally similar to supervised classified image of LISS-III (of 2006). The proportions of urban pixels in the image (figure 3) range from 0 to 1, with 0 indicating absence and increasing value showing higher abundance.

Accuracy assessment at pixel level: A pixel of MODIS (250 m) corresponds to a kernel of 10 x 10 pixels of LISS-III spatially (resampled to 25 m). The builtup abundance map was validated pixel by pixel with the LISS-III builtup map. The proportion of builtup percentage obtained in these 136 x 140 cells were compared using three measures of performance – Root Mean Square (RMS) error (0.160), Bivariate Distribution Functions (BDFs) between LISS-III

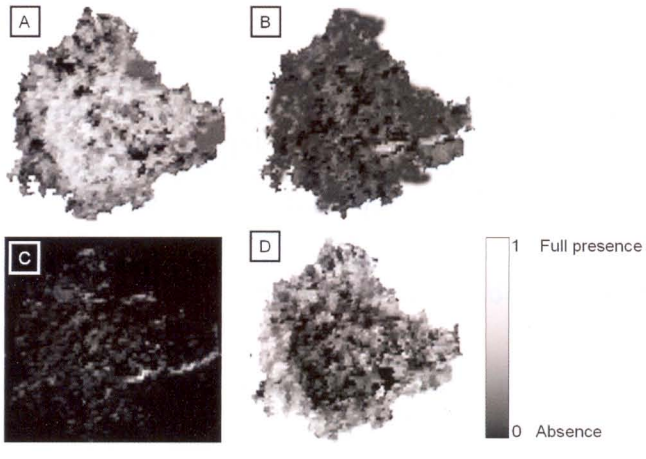

Figure 3: LC abundance maps (of 2006) for builtup, vegetation, water bodies and others. (Self Compiled)

classified (real proportion) and MODIS based abundance map (predicted proportions), and Wilcoxon signed rank test. The bivariate plot (figure 4) was helpful in visualising the accuracy of prediction by mixture models. While most of the points fall between the two 20 % difference lines, some points are outside the lines indicating that there

could be higher or lower estimation than the actual values. The Pearson's product-moment correlation was 0.84 (p-value < $2.2e^{-16}$). 10 % of the difference pixels between the LISS-III proportion and the MODIS abundance were randomly selected at the same spatial locations in the two images by random number generation. These difference data values failed when subjected to normality test (table 3).

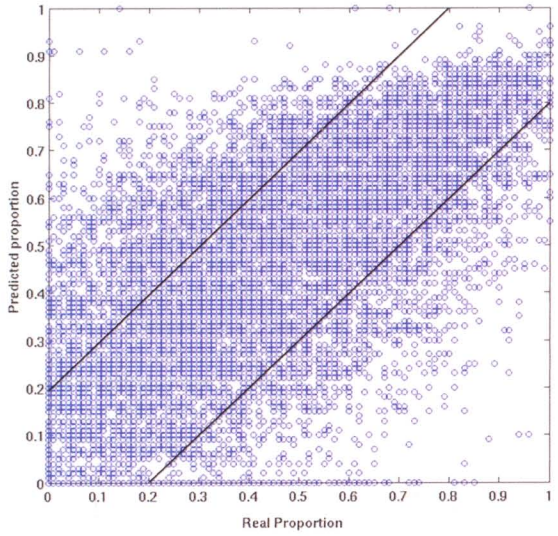

Figure 4: BDF of LISS-III classified and MODIS based abundance map. (Self Compiled)

Table 3: Normality tests for the difference in LISS-III and MODIS proportions (Self Compiled)

Normality test	p-value
Shapiro-Wilk	$2.2e^{-16}$
Anderson-Darling	$2.2e^{-16}$
Kolmogorov-Smirnov	$2.2e^{-16}$
Cramer-von Mises	Infinity
Lilliefors	$2.2e^{-16}$

The data points corresponding to builtup proportion in a MODIS pixel and corresponding LISS-III pixels were then subjected to box-plot and histogram analysis. The distribution was symmetrical, hence Wilcoxon signed rank test with continuity correction was performed (r = 0.66, p-value < $2.2e^{-16}$), which showed less evidence against null hypothesis – *"true location of median of the difference between two data sets is equal to 0"*.

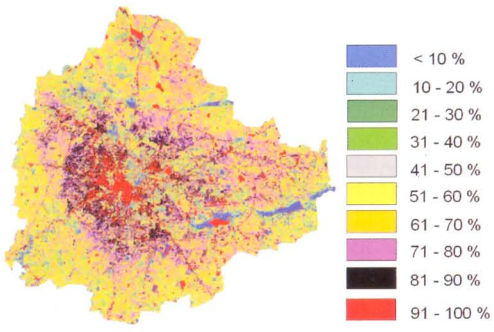

Figure 5: Percent impervious surface layer for the Bangalore city. (Self Compiled)

The abundance image was further analysed with a minimum detectable threshold of 10 % and in increments of 10 % (i.e., 20 to 30 %, 30-40 %, ..., 90 to 100 %) as shown in figure 5. Objects representing different brightness classes of impervious surface are typically selected to be mapped based on the tonal variations in the builtup landscape, which constitute spectral classes. The results hold great promise for improved accessibility and accuracy of this important urban indictor at the local and city level. Table 4 depicts the percentage impervious surface from abundance map.

4.3 Forest fragmentation indices: Supervised classification was performed on temporal datasets for Bangalore city with the Bayesian classifier using the open source programs (i.gensig, i.class and i.maxlik) of Geographic Resources Analysis Support System (http://wgbis.ces.iisc.ernet.in/grass). Accuracy assessment was done using field knowledge, and visual interpretation which showed an overall accuracy of 72 %, 75 %, 71 %, 77 % 85 % and 55 % for the year 1973, 1992, 1999, 2000, 2006 and 2007. The class statistics is given in table 5.

Table 4: Percent Impervious surface from abundance map (Self Compiled)

Percent Impervious surface	Area	
	Ha	%
0 - 10 %	7019.57	9.97
11 - 20 %	5853.89	8.31
21 - 30 %	7410.17	10.52
31 - 40 %	10126.27	14.38
41 - 50 %	10199.57	14.48
51 - 60 %	9149.63	12.99
61 - 70 %	9863.94	14.01
71 - 80 %	7050.32	10.01
81 - 90 %	3485.47	4.95
91 - 100 %	274.69	0.39

From the classified raster data, urban class was extracted and converted to vector representation for computation of precise area in hectares indicating 466 % increase in built up area from 1973 to 2007 as evident from temporal analysis leading to a sharp decline of 61% area in water bodies mostly attributing to intense urbanisation process. Rapid urbanisation and sprawl has many detrimental effects including the loss of valuable agricultural and eco-sensitive (e.g. wetlands, forests) lands, enhanced energy consumption and greenhouse gas emissions from increasing private vehicle use and also from modern glazed architectures. Vegetation has decreased by 63 % from 1973 to 2007. Using the results from classification, forest fragmentation indices were computed as shown in figure 6 and the statistics are presented in table 6.

Applying forest fragmentation indices to a time series LC data provided a quantitative assessment of the pattern of forest fragmentation at each date and a means for tracking trends in forest fragmentation. Patch forest which

was only 2.45 % in 1973 has increased up to 9.6 % in 2006 and interior forest has decreased from 54 % in 1972 to 27.7 % in 2006.

Table 5: Greater Bangalore land cover statistics (Self Compiled)

Class → Year ↓		Built up	Vegetation	Water Bodies	Others
1973	Ha	5448	46639	2324	13903
	%	7.97	68.27	3.40	20.35
1992	Ha	18650	31579	1790	16303
	%	27.30	46.22	2.60	23.86
1999	Ha	23532	31421	1574	11794
	%	34.44	45.99	2.30	17.26
2000	Ha	24163	31272	1542	11346
	%	35.37	45.77	2.26	16.61
2006	Ha	29535	19696	1073	18017
	%	43.23	28.83	1.57	26.37
2007	Ha	30876	17298	1005	19143
	%	45.19	25.32	1.47	28.01

Using forest fragmentation indices, the state of the fragmentation for the city was developed. While the forest fragmentation map produced valuable information, it was difficult to visualise the state of forest fragmentation for tracking the trends and to identify the areas where forest restoration might prove appropriate to reduce the impact of deforestation. The values of TFP and FC calculated (table 7) for the temporal data were plotted on a graph that specifies six conditions of forest fragmentation in figure 7. The TFP designations revealed that forest fragmentation become more severe as forest cover decreases from 100 percent cover towards 80 percent. Between 60 and 80 percent forest cover, the opportunity for re-introduction of forest to connect forest patches is greatest, and below 60 percent, forest patches become small and more fragmented. The FC regions were evenly split and designated high forest continuity (above 0.5) or low forest continuity (below 0.5). The amount of forest and continuity were high in 1973 and decreased substantially in 2006 (figure 6). Forest edges and perforations have less meaning where the amount of forest is low and percolation theory [14] identifies two critical values of Pf. In a completely forested landscape, for a completely random forest conversion on an infinite grid (and evaluating adjacency in cardinal directions), the residual forest is guaranteed to occur in identifiable patches when Pf falls below a critical value of about 0.4. Below that value, the nonforest pixels form a continuous path across the window. Conversely, as long as the residual forest is above a critical value of about 0.6, the forest pixels form such a path, and forest edges and perforations have more meaning. These values are used to define *patch* and *transitional* categories recognising that they are approximate because actual LC pattern is not random.

Table 6: Composition of forest fragmentation (Self Compiled)

Fragmentation Categories	1973		1992		1999		2000		2006	
	Ha	%	Ha	%	Ha	%	Ha	%	Ha	%
Patch	1120	2.45	1950	6.34	1788.6	5.91	1944	6.47	1810	9.57
Transitional	3789	8.28	4259	13.84	3816	12.62	3999	13.31	3956	20.92
Edge	14717	32.16	10709	34.81	9516	31.46	10593	35.25	6050	31.99
Perforated	1453	3.18	1447	4.70	1410	4.66	1219	4.06	1859	9.83
Interior	24675	53.93	12401	40.31	13715	45.35	12291	40.91	5236	27.69
Total	45756	100	30766	100	30244	100	30047	100	18909	100

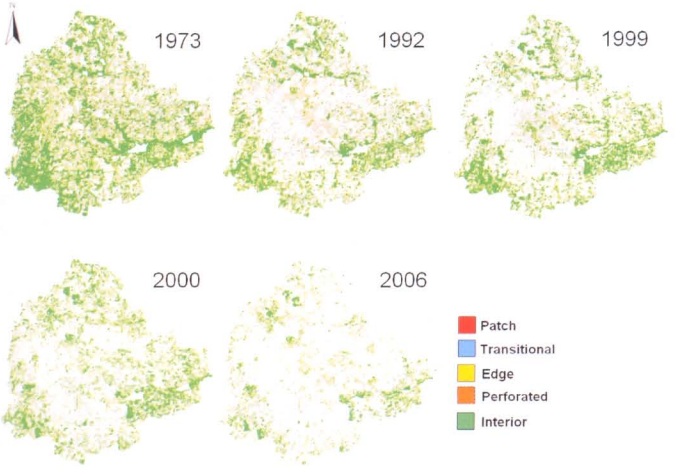

Figure 6: Forest fragmentation map (1972, 1992, 1999, 2000 and 2006). (Self Compiled)

Table 7: Total forest proportion and forest continuity (Self Compiled)

State of forest fragmentation index	1973	1992	1999	2000	2006
TFP	0.71	0.48	0.47	0.46	0.29
FC	0.45	0.30	0.34	0.30	0.19

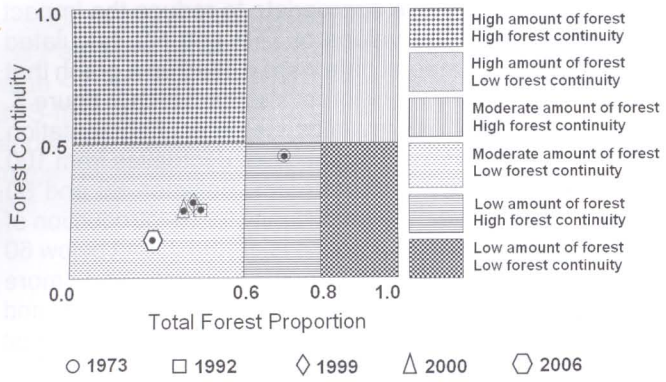

Figure 7: Six forest fragmentation conditions based on the values for Total Forest Proportion and Forest Continuity. (Self Compiled)

Forest fragmentation also depends on the scale of analysis (window size) and various consequences of increasing the window size are reported in [11]. The measurements are also sensitive to pixel size. Higher fragmentation occurs when finer grain maps are used over a fixed extent (window size). Finer grain maps identify more nonforest area where forest cover is dominant but not exclusive. The strict criterion for interior forest is more difficult to satisfy over larger areas. Although knowledge of the feasible parameter space is not critical, there are geometric constraints. For example, it is not possible to obtain a low value of Pff when Pf is large. Percolation theory applies strictly to maps resulting from random processes; hence, the critical values of Pf (0.4 and 0.6) are only approximate and may vary with actual pattern. As a practical matter, when Pf > 0.6, nonforest types generally appeared as "islands" on a forest background, and when Pf < 0.4, forests appeared as "islands" on a nonforest background.

4.4 Urban Dynamics Characterisation: Input data consisted of multi date images having three classes: urban, non-urban, and water bodies obtained from the supervised classification algorithm. Each single date fragmentation map consisted of three fragmentation classes – interior, perforated and patch. The final step in the urban growth model was to use two date fragmented image to create change map (1973 to 1992, 1992 to 1999, 1999 to 2000 and 2000 to 2006). The model produced urban growth maps, which consisted of developed land, non-developed land, water and the three types of growth (Infill growth, Expansion and Outlying growth) as shown in figure 8 and statistics are given in table 8.

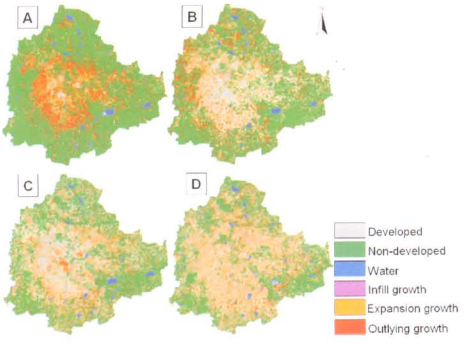

Figure 8: Urban growth map (A) 1973 to 1992, (B) 1992 to 1999, (C) 1999 to 2000, (D) 2000 to 2006. (Self Compiled)

Outlying growth have been further classified as – isolated, linear branching and clustered growth (figure 9). There were no perfect linear growth (linear branching surrounded by non- developed land). In all the cases, linear branching were surrounded by one of the other growth classes such as expansion growth, infill growth etc. Combining several temporal urban growth maps created an informative picture of the urban dynamics. Ward wise, impermeable areas were plotted against the urban population density to understand the increase in population density with the increase in the impermeable area. Increased paved surface and concentrated human activities has led to urban heat island effect, as there is increase in temperature between urban pixels compared to its surroundings. Visualisation helps the Local decision makers to see the implications of earlier decisions and policies and evolve appropriate LU policies for sustainable planning and management.

5. Conclusions

Urbanisation and urban sprawl seems inevitable in Indian cities as long as the growth is coupled with lack of holistic approaches in governance. Sections of the society will continue to cope up with skewed economic growth while attempting to protect natural resources. Visualisation of urban dynamics based on data mining and modeling aid in capturing the information in an effective way. This would help in better planning for provision of basic amenities and also to adopt sustainable urban strategies. Orthogonal subspace projection mapped the impervious surface accurately evident from accuracy assessment. Forest fragmentation indices and the urban dynamics model show the change in type of land cover classes from vegetation to urban. Interior forest which was present in the city up to 54 % (in 1973) has come down to 28 % in 2006. Urban growth characterised by developed, infill, expansion and outlying types evidently illustrate the phenomenon of urbanising Greater Bangalore.

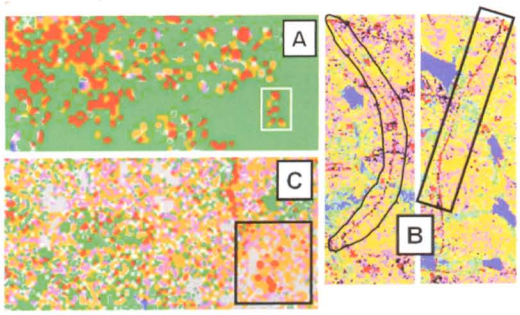

Figure 9: Types of urban outlying growth highlighted in boxes – (A) isolated growth, (B) linear branching (road/corridor), (C) clustered growth. (Self Compiled)

Acknowledgement

We thank Indian Institute of Science for financial and infrastructure support. NRSA, Hyderabad is acknowledged for providing LISS-III data. MODIS data were downloaded from http://glcf.umiacs.umd.edu/index.shtml. Google earth data (http://earth.google.com) was used as reference data for image analysis. We are grateful to Prof. N. V. Joshi for valuable suggestions during the spatial data analysis and model development.

Table 8: Changes in urban growth types from 1973 to 2006 (Self Compiled)

Urban Growth Type	1973 - 1992		1992 - 1999		1999 - 2000		2000 - 2006	
	Ha	%	Ha	%	Ha	%	Ha	%
Developed	6889	1.0	3818	5.6	2876	4.2	32380	4.74
Non-developed	48529	71	44034	64.6	42610	63	37922	55.54
Water	905	1.3	692	1.0	1434	2.1	952	1.39
Infill	720	1.1	2146	3.2	3088	4.5	4232	6.20
Expansion	5720	8.4	10078	14.8	13110	19	16412	24.04
Outlying	11520	17	7453	10.9	5054	7.4	5526	8.09
Total	68082	100	68220	100	68172	100	68282	100

References

[1] "*World Urbanization Prospects*," Revision, Population Division, Department of Economic and Social Affaris, UN, 2005

[2] T. V Ramachandra, and U. Kumar, "*Geoinformatics for Urbanisation and Urban Sprawl pattern analysis. In: Geoinformatics for Natural Resource Management*," (Eds. Joshi et al.). Nova Science Publishers, NY (In press), 2008.

[3] M. K. Arora, and S. Mathur, "*Multi-source classification using artificial neural network in rugged terrain*," GeoCarto International, vol. 16, no. 3, pp. 37-44, 2001.

[4] R. O. Duda, P. E. Hart, and D. G. Stork, "*Pattern classification*," A Wiley-Interscience Publication, Second Edition, ISBN 9814-12-602-0, 2000.

[5] Y. E. Shimabukuro, and A. J. Smith, "*The least-squares mixing models to generate fraction images derived from remote sensing multispectral data*," IEEE Transactions on Geoscience and Remote Sensing, vol. 29, no. 1, pp. 16–20, 1991.

[6] U. Kumar, N. Kerle, and T. V. Ramachandra, "Contrained linear spectral unmixing technique for regional land cover mapping using MODIS data," In: Innovations and advanced techniques in systems, computing sciences and software engineering / ed by Elleithy, K., Berlin: Springer, 2008. ISBN: 978-1-4020-8734-9. paper 87. 9 p, 2008.

[7] A. A. Nielsen, "*Spectral mixture analysis: Linear and semi-parametric full and iterated partial unmixing in multi-and hyperspectral image data*," International Journal of Computer Vision, vol. 42, no. 1, pp. 17-37, 2001.

[8] C-I Chang, "*Orthogonal Subspace Projection (OSP) Revisited: A Comprehensive Study and Analysis*," IEEE Transactions on Geoscience and Remote Sensing, vol 43, no. 3, pp. 502–518, 2005.

[9] J. D. Hurd, E. H. Wilson, S. G. Lammey, and D. L. Civco, "*Characterization of Forest Fragmentation and Urban Sprawl using time sequential Landsat Imagery*," In *ASPRS 2001 Annual Convention, St. Louis, MO, 2001.*

[10] S. A. Levin, "*The problem of pattern and scale in ecology*," *Ecology*, vol. 73, pp. 1943-1967, 1992.

[11] K., J. Riitters, R. O'Neill, Wickham, B. Jones, and E. Smith, "*Global-scale patterns of forest fragmentation*," Conservation Ecology, vol. 4, no. 2-3, 2000. http://www.consecol.org/vol4/iss2/art3/

[12] J. D. Hurd, E. H. Wilson, and D. L. Civco, "*Development of a Forest Fragmentation Index to Quantify the rate of forest change*," In ASPRS-ACSM Annual Conference and FIG XXII Congress, April 22-26, 2002.

[13] D. L. Civco, J. D. Hurd, E. H. Wilson, C. L. Arnold, and M. P. Prisioe, "*Quantifying and Describing Urbanizing landscape in the Northeast United States*," Photogrammetry Engineering and Remote Sensing, vol. 68, no. 10, pp. 1083 - 1090, 2002.

[14] D. Stauffer, "*Introduction to percolation theory*," Taylor and Francis, Philadelphia, Pennsylvania, USA, 1985.

------ ☺ ☺ ☺ ☺ ☺ ☺ ☺ ☺ ------