Urban growth dynamics and modelling using remote sensing data and multivariate statistical techniques

Manish Kumar[1,*], R. B. Singh[2], Ram Pravesh[3], Pankaj Kumar[2], Dinesh Kumar Tripathi[4] and Netrananda Sahu[2]

[1]Department of Geography, Kalindi College, University of Delhi, Delhi 110 008, India
[2]Delhi School of Economics, Department of Geography, University of Delhi, Delhi 110 008, India
[3]Department of Geography, Kumaun University, SSJ Campus, Almora 263 601, India
[4]Department of Geography, Kamla Nehru Institute of Physical and Social Sciences, Sultanpur 228 118, India

In this article, sprawl area of impervious surfaces and their spatial and temporal variability have been studied for Pune city over a period of 19 years, i.e. 1992–2011. Statistical techniques and image classification approach have been adopted to quantify the urban sprawl and its spatial and temporal characteristics. For this purpose, satellite images were obtained from various sensors, viz. Landsat Thematic Mapper and Landsat Enhanced Thematic Mapper Plus. To establish the relationship between urban sprawl and its causative factors, multivariate statistical technique has been used. The determinants of causal factors of urban sprawl such as population, α-population density, β-population density, workforce engaged in secondary and tertiary sectors, road density, and gender gap in literacy collectively explain the 93.09% variation in urban growth. The result also depicts that incessant growth in the built-up area in Pune city has surpassed the rate of population growth. From 1992 to 2011, population in the region grew by 75.40% while the amount of built-up land grew by 227.3%, i.e. more than three times the rate of population growth. To understand the future urban growth of Pune city, a foresight approach is being developed that allows long-term projections. This depicts that by the year 2051, the built-up area in the municipal limits would rise to 212.27 sq. km, which may be nearly 50.0% more than that in 2011 (141.50 sq. km). The vegetative areas, open spaces and areas around the highways are expected to become major targets for urban sprawl due to further increase in the pressure on land.

Keywords: Remote sensing, statistical techniques, spatial and temporal variability, urban sprawl.

THE urban areas of developing countries are experiencing a rapid state of change. The urban centres of the world are growing fast in terms of geographical area and population. According to a report by the United Nations[1], the trend of global urbanization shows that 66% of the world's population is projected to be urban by 2050, with 90% of this expansion being anticipated in the developing countries. The concentration of the world population will be in Asia (52%) and Africa (21%) for most of the urban area by the year 2050. Since independence, the percentage of urban population to the total population in India has been steadily increasing. With more than 30% of urban population, India is the second largest urban system in the world. It is estimated that by 2025, half of India's population will become urban. The urban population of Indian cities is highly concentrated in larger cities which are growing rapidly at the cost of small urban centres.

Usually urban growth is determined by the population concentration in an area. Urbanization drives the change in land use/cover pattern. Various activities such as urban planning and management, land and water resources management, service and marketing analysis, etc. need precise information on the extent of urban growth. In order to accommodate the increasing population or other urban land uses, urban local bodies are required to dedicate more effort, attention and time to manage the use of land resources[2]. Estimation of urban sprawl by traditional surveying and mapping techniques is expensive and time-consuming. Hence, nowadays remote sensing and GIS techniques are used extensively for mapping and monitoring of urban sprawl[3]. Several studies have been made on urban sprawl[4–8].

Geospatial techniques play a vital role in quantifying and modelling urban landscape, which is not possible through traditional mapping techniques. For the analysis and modelling of urban sprawl and land-use change, remote sensing techniques have already established their importance in mapping urban areas and also as a data source[9–13]. Spatially consistent datasets that cover large areas with both high spatial detail and high temporal frequency are provided by remote sensing. Also, remote sensing is both cost- and time-effective, and is therefore a popular technique for the analysis of urban sprawl[14,15].

*For correspondence. (e-mail: manish.ks1@gmail.com)

Remote sensing, GIS and database management systems have helped in quantifying, monitoring and modelling the urban sprawl phenomenon. For nearly three decades, extensive research efforts have been made for urban change detection using remotely sensed images[16–19]. A post-classification or an image-to-image comparison has supported these studies.

Remote sensing and GIS along with statistical techniques have been used to estimate, quantify and monitor the urban growth pattern[20–24]. To establish the relationship between built-up area and various urban development indicators, multivariate statistical techniques have been used.

Sudhira *et al.*[14] conducted a study to model the future urban growth and estimate the dynamics of urban sprawl pattern using remote sensing and GIS techniques and other datasets. Kumar *et al.*[25] considered Indore city, Madhya Pradesh for a similar study. Three temporal satellite remote sensing data were used for the period 1990–2000 to analyse Indore city. Jat *et al.*[26] studied Ajmer city, Rajasthan for the period 1977–2005 using satellite remote sensing data. Results of their study showed that the built-up growth in Ajmer city is three times higher than the population growth. Punia and Singh[27] conducted a study on Jaipur city, Rajasthan using entropy approach to quantify urban sprawl. The results of their study revealed that the built-up growth rate of Jaipur city has surpassed the growth rate of population. Rawat and Kumar[28] conducted a study on land use/cover pattern and dynamics of Hawalbagh block, Almora district, Uttarakhand. A study of satellite data from 1990 to 2010, i.e. 20 years showed that the built-up area has sharply increased on agricultural and vegetation lands around Almora town area due to construction of new buildings. Roshan *et al.*[29] have established the relationship between climate variables and components of urban sprawl using regression and correlation methods in Tehran, and found significant relationship between climate change and urban sprawl. Polyzos *et al.*[30] analysed empirically the urban sprawl-driven land-use changes and their regional and economic development implications in Greece using ordinal regression model and accessibility, housing, population and natural resources as major driving forces of urban sprawl. Majid and Mohammad[31] studied the dynamics and prediction of land use/cover changes, population growth and urban expansion in Srinagar city, Jammu and Kashmir using integrated approaches of remote sensing, GIS, multivariate statistical techniques and regression models. Andrew *et al.*[32] developed a database on urban expansion and predicted future urban growth pattern for the city of Niamey, Niger using geospatial and statistical modelling approaches. This study suggests the adverse environmental and social consequences of unrestrained urban growth in the study area. Goswami and Khire[33] studied urban sprawl analysis of Ahmedabad city, Gujarat, analysing multi-temporal Landsat Data (TM and ETM+) using remote sensing techniques. They highlighted the usefulness of a database on land use/cover for urban planners and decision-makers.

Here we examine the usefulness of remote sensing data and statistical techniques for urban growth dynamics and modelling of spatial and temporal variability. Urban growth of Pune city, Maharashtra in the last 19 years (1992–2011) has been estimated of four different years. Multivariate regression analysis was used to establish a relationship between urban growth and causal factors such as population, α-population density, β-population density, workforce engaged in secondary and tertiary sectors, road density and gender gap in literacy.

Study area

Pune city is located between 18°25′46″N–18°37′17″N lat. and 73°44′58″E–73°57′46″E long. In 1950, under the Bombay Provisional Municipal Corporation (BMPC) Act, 1949, the Pune Municipal Corporation (PMC) was established. PMC includes four main zones; it is further sub-divided into 14 administrative wards (Figure 1). The built-up growth of PMC is significantly interconnected and reinforced by its nearby clusters. Since 1951, Pune city has developed spatially more than six times. Presently, it is the ninth largest city in India with a population of about three million. In 2001, the population of Pune city was about 2,538,473, which increased to 3,115,431 in 2011, accounting for 22.73% growth in a decade. Density of population has also increased from 10,410 persons/sq. km in 2001 to 12,777 persons/sq. km in 2011.

Pune city is growing in the pattern of concentric rings. Since the establishment of PMC, small village areas have been added which has increased the area of jurisdiction of PMC. The factors aiding the growth of Pune city are mainly the flourishing of the IT sector as well as economic development in the automobile industry in and around the city. The outskirt growth has resulted into increased residential areas, transportation nodes and facilities. Industrial growth is mostly found in the northwest and southeast corridor along major roads entering the city. The population of Pune city has increased manifold in the last 60 years and after 1981, the increase has been very high. The continually growing population has put pressure on the adjoining areas and the city is extending outwards. The built-up expansion has taken place in all directions, but more extensively in the southern, southwestern and eastern directions. In the eastern part of the city, important changes in land use/cover have been noticed. The major land use in the heart of Pune city or the Central Business District (CBD) is under commercial and residential activities. The older part of the city is congested and overcrowded with little open space and narrow roads. Large sections of high- and middle-income groups have mainly settled down in the fringe areas.

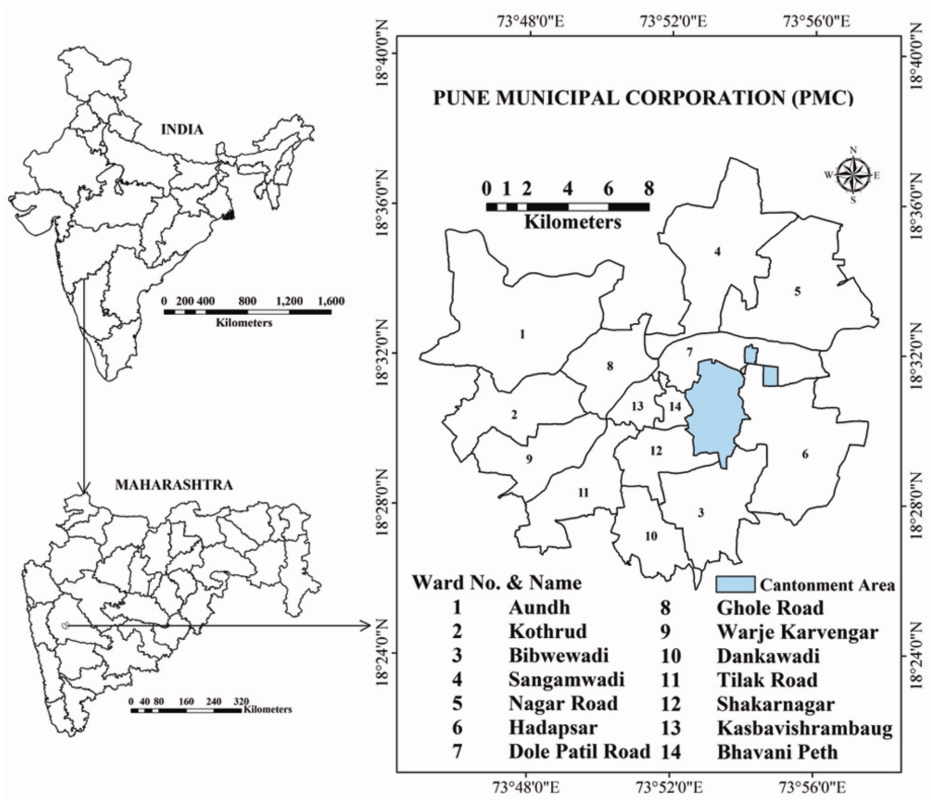

Figure 1. Study area.

Table 1. Data used

Data	Path and row	Year	Spatial resolution/scale (m)	Source
Satellite images				
Landsat TM	147, 047	February 1992	30	GLCF
Landsat ETM+	147, 047	November 1999	28.5	EarthExplorer
Landsat ETM+	147, 047	October 2006	28.5	GLCF
Landsat ETM+	147, 047	January 2011	28.5	EarthExplorer
Additional data				
Toposheet	–	–	1 : 50,000	Survey of India
PMC map	–	2011	–	PMC
Demographic data	–	1991, 2001, 2011	–	Census of India
Road layer	–	2011	–	Google Earth

Methodology

Multi-temporal images for four years (1992, 1999, 2006 and 2011) were selected for analysis. The freely available Landsat images used have been acquired from the Global Land Cover Facility (GLCF) and EarthExplorer website of the United States Geological Survey, because it is the only source that has an enough temporal data. Another reason for selecting these images was their availability at similar resolution. Few additional data were also acquired for the study. The Survey of India topographical sheet on 1 : 50,000 scale, PMC map and demographic data from the Census of India were also used for the study (Table 1).

The ERDAS Imagine 13.0 software was used for image processing at various stages of image analysis. ArcGIS 10.2 software was used for spatial analysis and generating thematic layers.

At the classification level, maximum likelihood algorithm of supervised classification was applied to the images. This algorithm is based on the probability that each pixel is classified to a specific class. By creating training areas for each class, all the images were classified. Only four main classes were considered at the time of classification: impervious built-up area, vegetation land, bare land or open space and water body[34]. Table 2 presents the selected classes and their description.

Table 2. Description of land use/cover classes

Land use/cover	Description
Urban or built-up area	Residential, commercial, and services and industrial, road, other means of transportation and urban features
Vegetation land	Forest, scattered trees, parks
Open space/vacant land	Exposed soil, landfill sites, area of active excavation, open space in built-up land
Water body	River, lakes, ponds, reservoirs, etc.

Table 3. Mapping accuracy of land use/cover

Land use/cover	Overall accuracy (%)	Kappa coefficient
1992	85.59	0.7396
1999	87.39	0.8023
2006	92.79	0.8841
2011	87.27	0.7747

A post-classification technique was used for land use/cover change detection. In order to produce the change detection map and quantify the changes more efficiently, a pixel-based comparison was done. To quantify the changes for the 19-year period (1992–2011), classified images of two different decades' data were compared using cross-tabulation matrix. A change matrix and land encroachment map were generated using ERDAS software. The matrix showed the gains and losses in each land use/cover classes during 1992–2011.

Field data were collected owing to two reasons. First, to obtain GPS data for ground verification of doubtful areas. The ground control points were used to correct the misclassified areas using recode option in ERDAS Imagine software. Secondly, to estimate the mapping accuracy of classified images. Table 3 shows the results of the classification accuracy.

In general, urban development is controlled by some locational factors such as distance from heart of the urban centres and major roads. As the distance from urban centres and roads increases, the density of built-up area decreases rapidly. The density gradient relationships between average densities of built-up area growth and distances from city centre and major roads were calculated using regression techniques.

In order to develop the probable relationship between percentage built-up area (dependent variable) and causal factors (independent variables) of growth, regression analysis was carried out. In this study the selected independent variables are total population, α-population density, β-population density, workforce engaged in secondary and tertiary sectors, road density and gender gap in literacy.

Results and discussion

Population growth and built-up area

According to the Census of India, 2011, Pune city has a population of more than three million. During the last 60 years, the population of Pune city has increased more than six times, from 0.48 million in 1951 to 3.11 million in 2011. The average decadal growth rate from 1951 to 2011 is 36.54%. This unprecedented growth has been due to industrialization and expansion of IT sector around PMC/Pimpri–Chinchwad Municipal Corporation (PCMC) area. The maximum distribution of population was found mostly in the area of Hadapsar and Bibvewadi ward, mainly due to the establishment of IT companies. The decadal growth pattern of Pune city has shown a sharp decline from 50.08% decadal growth rate in 1991–2001 to 22.73% decadal growth rate in 2001–2011. This is mainly due to the development of industrial centre of PCMC. It may be acted as counter magnet for the development of Pune city. The higher growth rate and economic activities in the Pune region attract several migrants not only from the adjoining regions, but also from different parts of the country. The population density has increased from 3907 persons/sq. km in 1951 to 12,777 persons/sq. km in 2011. The population density gradually increased from 1951 to 2011, except in 2001 when it was 10,410 persons/sq. km. This increase is primarily due to the addition of 23 villages within the limit of PMC area.

Urban growth for the years 1992, 1999, 2006 and 2011 was estimated as growth in the built-up area, which was obtained from multi-spectral and multi-temporal satellite images. Built-up area increased from 43.22 sq. km in year 1992 to 141.50 sq. km in 2011. Results show that the built-up area development rate of Pune city has surpassed the rate of population growth (Table 4). From 1992 to 2011, the population grew by 75.40% while the built-up area grew by 227.3%, which is higher than three times the rate of population growth (Table 4 and Figure 2). This indicates that the land was utilized at a faster rate due to higher rate of urbanization. It further depicts that over the last two decades, the per capita land consumption has increased exceptionally.

Figure 3 shows the spatial distributional pattern of ward-wise built-up area development during 1992–2011. Built-up area growth is higher in the fringe areas (ward nos 1–6) along the major roads, industrial and commercial hubs compared to core of the city. When the city expanded, the settlements with low population also grew; the development of these settlements took place outside the core of the city due to the search for better employment opportunities and good accessibility of vacant land.

Table 4. Urban growth statistics for Pune city during 1992–2011

Year	Built-up area (sq. km)	Increase in built-up area (sq. km)	Percentage increase in built-up area	Projected population	Absolute increase in population	Percentage growth in population
1992	43.22	–	–	1,776,136	–	–
1999	80.19	36.97	85.5391	2,369,084	592,948	33.38
2006	118.59	38.4	47.88627	2,826,970	457,886	19.33
2011	141.5	22.91	19.31866	3,115,431	288,461	10.2
1992–2011	–	98.28	227.3947		1,339,295	75.4

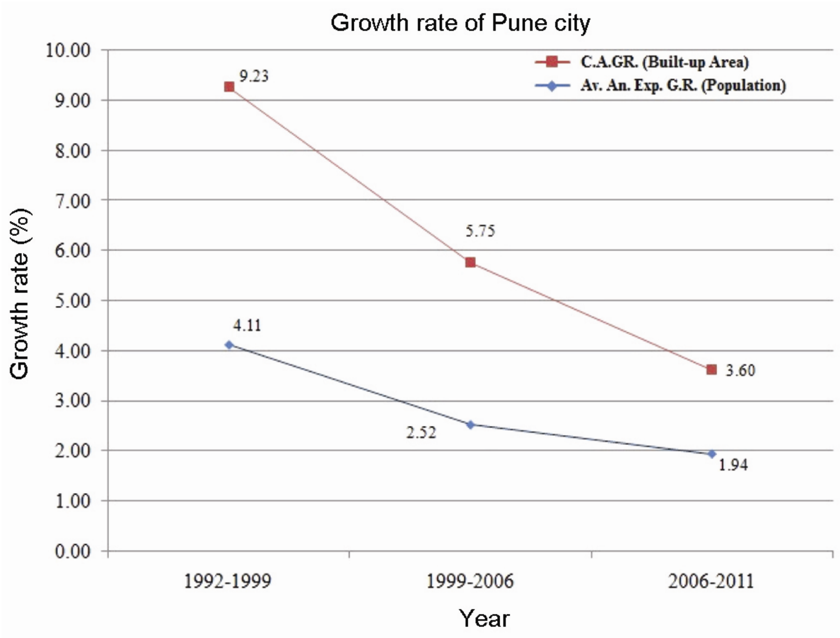

Figure 2. Comparison of the built-up area and population growth of Pune city (1992–2011). C.A.GR., Compound annual growth rate; Av.An.Exp.G.R., Means average annual exponential growth rate.

Land use/cover pattern and change

Table 3 presents the overall accuracy and Kappa coefficient for all the classified images. For better classification results, random sets of 300 samples were produced. With the help of reference images, the classification results were compared with the true information classes. Due to coarse classification, only four classes were used in this study; therefore, higher accuracy was obtained.

The results reveal that both gain and loss occurred in the land use/cover pattern of Pune city (Table 5 and Figure 4). The classified land use/cover maps depict that the total built-up area for 1992 was 43.22 sq. km. It increased to 80.19 sq. km by 1999, 118.59 sq. km by 2006 and finally reached 141.50 sq. km in 2011. This shows about 227.34% growth in built-up area over a period of two decades (Figure 5). With respect to vegetation cover, in 1992 it was 43.43 sq. km and decreased to 12.17 sq. km by 2011. Likewise, open space was about 160.05 sq. km in 1992 and decreased to 93.39 sq. km in 2011. The total decrease in vegetation cover and open space during the study period was 72% and 41% respectively. The change in land use/cover was mainly due to overexploitation of land for built-up purpose.

A change detection matrix (Table 6) and map (Figure 6) were prepared for a better understanding of land conversion for different land categories during the last two decades. They reveal that about 0.43 sq. km area of water body is converted into vegetation, 0.6 sq. km area under open space and 0.68 sq. km area under built-up land; about 0.77 sq. km area of vegetation land has been transformed into water body, 17.11 sq. km into open space and 18.54 sq. km into built-up land; about 0.53 sq. km area of open space is changed into water body, 4.73 sq. km area under vegetation and 79.06 sq. km area under built-up land.

Density gradient analysis

It is observed that urban development usually occurs around city centres and along major roads (Figures 7 and 8). To determine the impact of these locational functions on the spatial pattern of land development, proximity

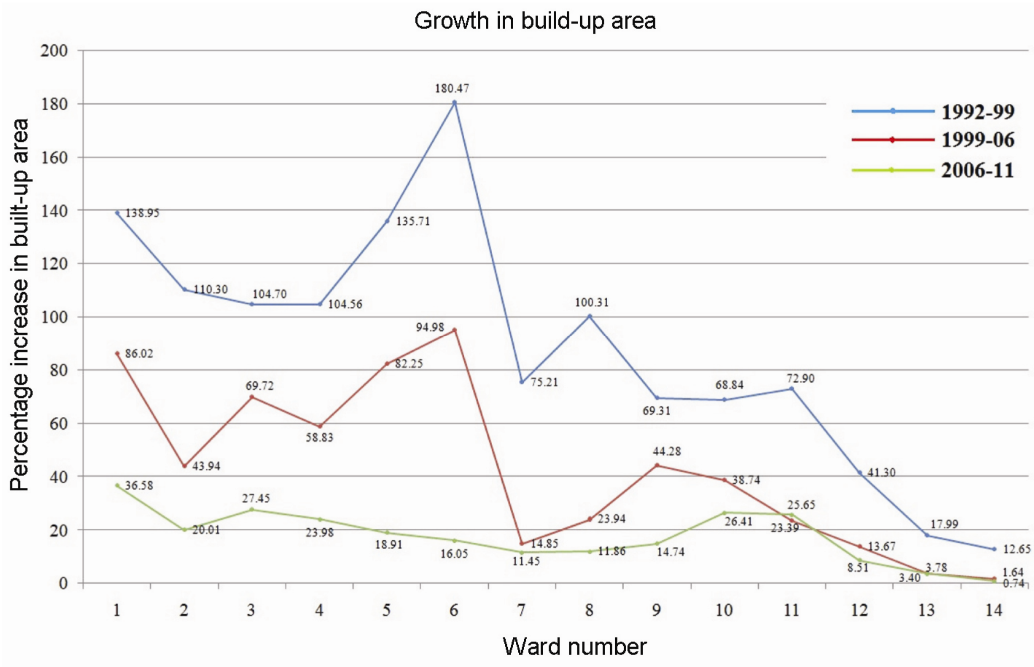

Figure 3. Ward-wise percentage growth in built-up area of Pune city from 1992 to 2011.

Table 5. Status of land use/cover in all categories (sq. m)

Year	1992	1999	2006	2011
Water body	3.0312	2.7414	3.1833	2.6676
Vegetation	43.434	57.7305	49.3857	12.1779
Open space	160.0524	109.0755	78.5817	93.3921
Built-up area	43.2288	80.199	118.5957	141.5088

Source: Landsat TM and ETM+ satellite data for 1992, 1999, 2006 and 2011.

(buffer) analysis of GIS was done. In order to calculate the density of land development in each buffer zone, buffer zones were created around the city centre and major roads. The distance decay function of built-up area development can be seen in Figure 9 as scatter plots. The relationship between average density of built-up development from the city centres and major roads can be summarized using regression analysis. Exponential model has been found to be the highest correlation coefficient [eq. (1) ($R^2 = 0.955$)] to show the relationship between average density of built-up development and distance from major roads compared to linear, quadratic, logarithmic and power distributions.

$$y = 0.745e^{-0.0002962553785x}, \tag{1}$$

where y is the average density of built-up development and x is the distance from major roads. On the other hand, logarithmic model has been found to be the best-fitted correlation coefficient [eq. (2) ($R^2 = 0.919$)] to show the relationship between average density of built-up development and distance from city centres compared to linear, exponential, quadratic and power distributions.

$$y = -0.186 \ln(x) + 2.199, \tag{2}$$

where y is the average density of built-up growth and x is the distance from the city centre. The decline in the density of built-up area growth was much faster from roads than from city centre (Figure 8).

Modelling of urban sprawl

The determinants of causal factors of urban sprawl modelling are population (P), α-population density (α-PD), β-population density (β-PD), workforce engaged in secondary and tertiary sectors (WST), road density (RD) and gender gap in literacy (GGP). The proportion of built-up area to total area of a ward is the percentage of built-up area. The proportion of population in every ward to the built-up area of that ward is the α-population density. The proportion of population in every ward to the total area of that ward is the β-population density. Sometimes, β-population density is also known as population density. In the present study, built-up area plays an important role in the analysis. The percentage built-up area, α- and β-population densities are calculated and examined ward-wise and categorized as sub-zones. Population data (ward-wise) for the year 2011 were downloaded from the Census of India website, which includes some other key factors of urban sprawl like workforce engaged in secondary and tertiary sectors and gender gap in literacy. The

Table 6. Land use/cover change matrix of Pune city (sq. km)

Land use/cover category		Year 1992 Water body	Vegetation land	Open space	Built-up
Year 2011	Water body	1.32	0.77	0.53	0
	Vegetation land	0.43	7.01	4.73	0
	Open space	0.6	17.11	75.73	0
	Built-up area	0.68	18.54	79.06	43.22

Source: Landsat TM and ETM+ satellite data for 1992 and 2011.

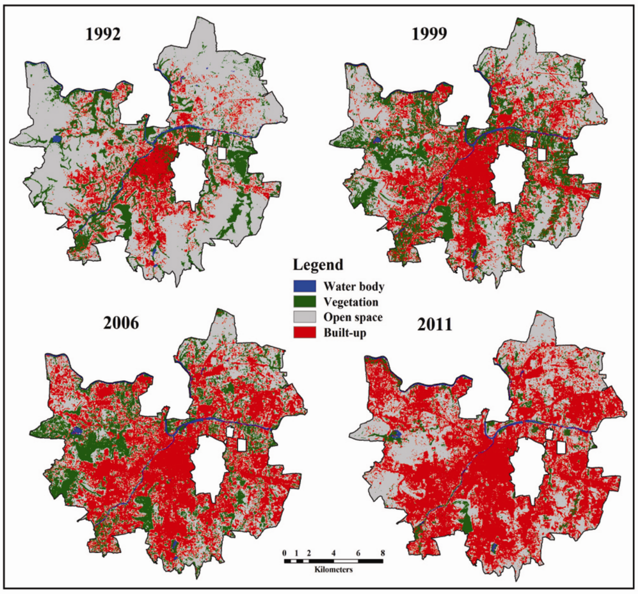

Figure 4. Land use/cover status during 1992–2011.

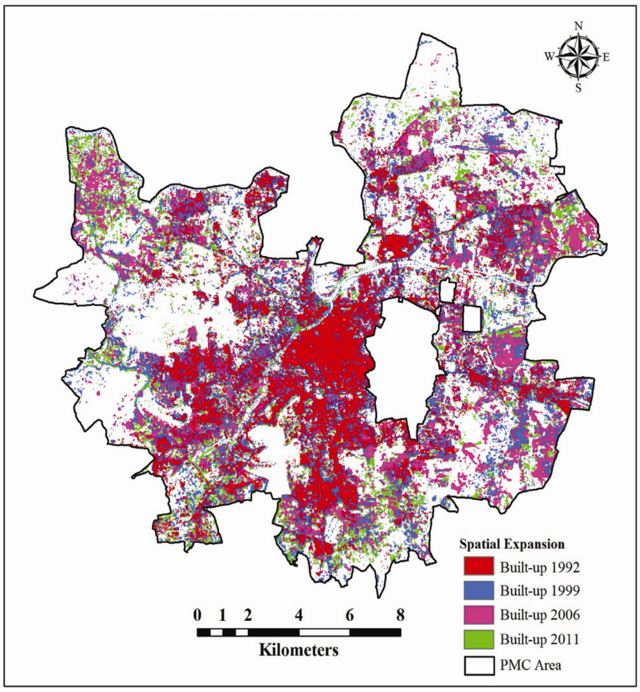

Figure 5. Spatial expansion of built-up area during 1992–2011.

physical growth of a city is largely driven by its tertiary and secondary sectors. Hence, ward-wise workforce engaged in secondary and tertiary sectors was calculated. Literacy is the key indicator of urbanization. When the population of any region grows and become urbanized, the gap between male and female literacy decreases. Urban growth follows the major transportation nodes. In order to calculate road density, ward-wise road network has been digitized. With the help of these causal factors, urban sprawl modelling was carried out.

In order to investigate the possible relationship between dependent (percentage built-up area) and independent variables (causal factors of sprawl), regression analysis was carried out. To determine the nature of significance of independent variables (causal factors of sprawl), several regression models such as linear, quadratic, exponential and logarithmic were considered for the study. The regression models represent individual contribution of independent variables on urban growth dynamics. A variety of associations and their statistical factors are shown in Appendices 1–6. Some of the significant relationships are as follows.

Road density:

$$y = 4.11x + 22.28, \tag{3}$$

where $R^2 = 0.867$, y is the percentage built-up area and x is the road density.

β-Population density:

$$y = -(2.15 \times 10^{-8})x^2 + (2.43 \times 10^{-3})x + 31.002, \tag{4}$$

where $R^2 = 0.827$, y is the percentage built-up area and x is the β-population density.

α-Population density:

$$y = -(8.67 \times 10^{-13})x^3 + (9.33 \times 10^{-8})x^2 - (1.83 \times 10^{-3})x + 88.25, \tag{5}$$

where $R^2 = 0.570$, y is the percentage built-up area and x is the α-population density.

Table 7. Model summary of stepwise regression

Model	R	R^2	Increase in R^2	Adjusted R^2	Increase in adjusted R^2	Standard error of the estimate
1	0.9310[a]	0.8668	–	0.8557	–	6.5469
2	0.9359[b]	0.8760	0.0091	0.8534	–0.0023	6.5990
3	0.9600[c]	0.9216	0.0456	0.8980	0.0446	5.5044
4	0.9642[d]	0.9296	0.0081	0.8983	0.0003	5.4961
5	0.9645[e]	0.9303	0.0007	0.8867	–0.0116	5.8014
6	0.9649[f]	0.9309	0.0007	0.8717	–0.015	6.1730

[a]Predictors: (constant), α-PD; [b]Predictors: (constant), α-PD, P; [c]Predictors: (constant), α-PD, P, β-PD; [d]Predictors: (constant), α-PD, P, β-PD, RD; [e]Predictors: (constant), α-PD, P, β-PD, RD, GGL; [f]Predictors: (constant), α-PD, P, β-PD, RD, GGL, WST.

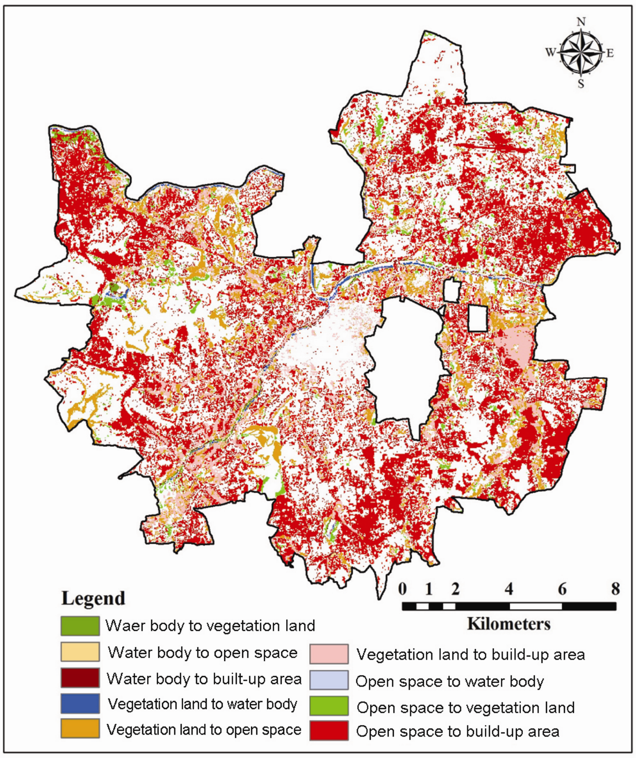

Figure 6. Land use/cover change during 1992–2011.

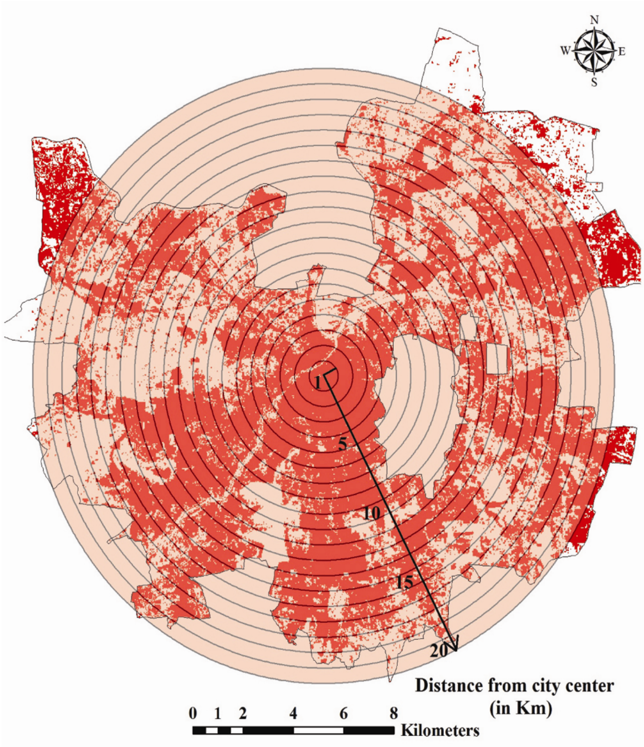

Figure 7. Built-up area growth from city centre.

Gender gap in literacy:

$$y = (2.15 \times 10^{-7})x^2 - (6.46 \times 10^{-3})x + 102.06, \quad (6)$$

where $R^2 = 0.426$, y is the percentage built-up area and x is the gender gap in literacy.

Workforce engaged in secondary and tertiary sectors:

$$y = (2.14 \times 10^{-14})x^3 - (8.09 \times 10^{-4})x + 117.98, \quad (7)$$

where $R^2 = 0.106$, y is the percentage built-up area and x is the workforce engaged in secondary and tertiary sectors.

Population:

$$y = -(1.85 \times 10^{-16})x^3 - (1.19 \times 10^{-4})x + 88.25, \quad (8)$$

where $R^2 = 0.059$, y is the percentage built-up area and x is the population.

Linear regression model has been found to be the higher correlation coefficient [eq. (3) ($R^2 = 0.867$)] to show the relationship between percentage built-up area and road density compared to quadratic, logarithmic, cubic, exponential and power distributions. Relationship between percentage built-up area and β-population density is quadratic. Quadratic regression outcomes depict higher correlation coefficient ($R^2 = 0.827$). Relationships between α-population density and percentage built-up area have been found to be cubic with higher correlation coefficient ($R^2 = 0.867$). Quadratic regression model has been found to be the higher correlation coefficient ($R^2 = 0.426$) to explain the relationship between percentage built-up area and gender gap in literacy. Relationships between percentage built-up area and workforce engaged in secondary and tertiary sectors have been found to be cubic regression with higher correlation coefficient ($R^2 = 0.106$). Relationships between percentage built-up area and population have been found to be cubic with highest correlation coefficient ($R^2 = 0.059$). The results of quadratic and linear regression depict that population density and road density have substantial impact on percentage built-up area.

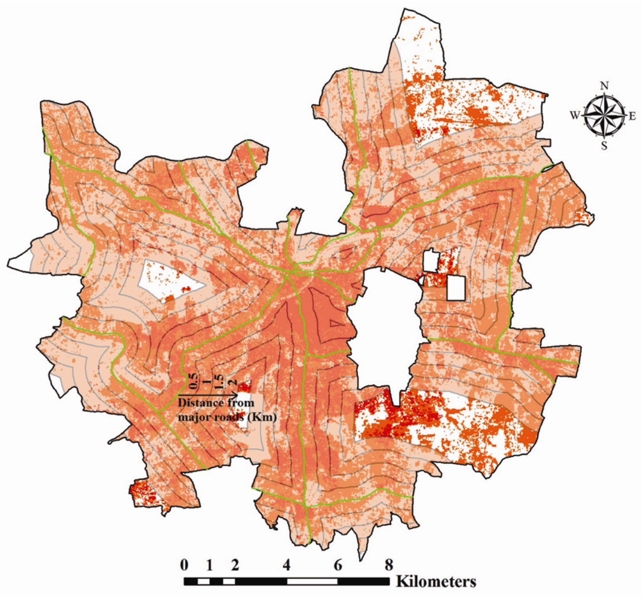

Figure 8. Built-up area growth from major roads.

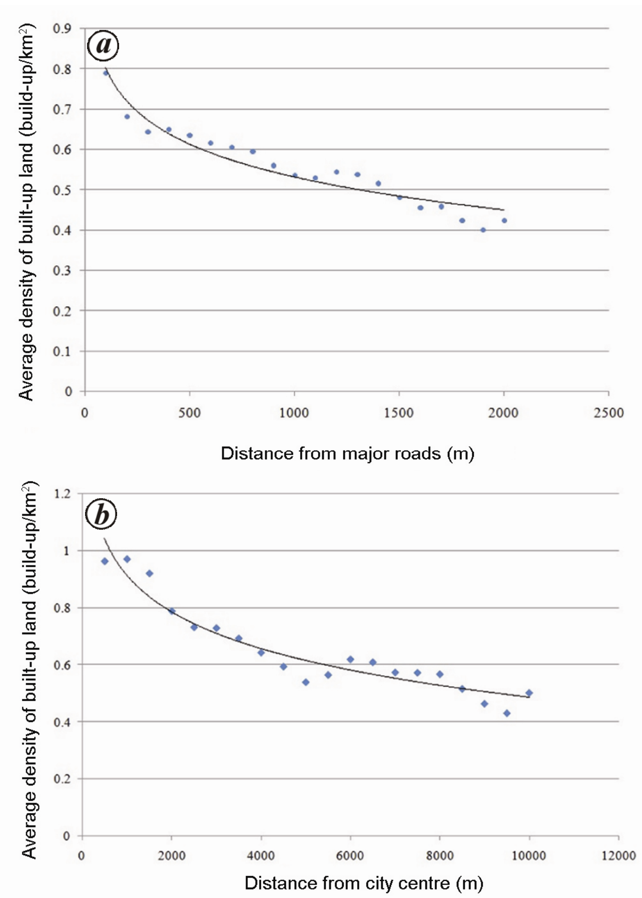

Figure 9. The relationship between average densities of land development and distances from (***a***) major roads and (***b***) city centre.

Stepwise multivariate regression analysis was also carried out to estimate the collective impact of independent variables (Table 7). In this regression model, it is presumed that the relationships between the factors are linear. Using each of the independent variables (causal factors) in the stepwise regression, eq. (9) shows the highest correlation coefficient ($R^2 = 0.9309$) which collectively explains the 93.09% variation in the urban growth. The α-population density explains the highest proportion (86.68%) as it represents the distinctive characteristic of city development (Appendix 7).

$$\begin{aligned}\mathrm{PB} = {} & 34.86 + 3.06\mathrm{RD} + (1.50 \times 10^{-3})\beta \\ & - \mathrm{PD} - (1.32 \times 10^{-3})\alpha - \mathrm{PD} \\ & + (3.15 \times 10^{-4})\mathrm{WST} - (1.19 \times 10^{-4})\mathrm{P} \\ & + (5.05 \times 10^{-4})\mathrm{GGL}. \qquad (9)\end{aligned}$$

Predicting scenarios and future growth pattern of urban sprawl

To deal with the future urban growth, a particular foresight approach is being developed that allows long-term projections. Equation (10) predicts the urban sprawl of Pune city considering that road density, population and β-population density are available from the database. The relationships between these variables are linear. Considering these independent variables (causal factors) in the stepwise regression shows the highest correlation coefficient ($R^2 = 0.8793$; eq. (10))

$$\begin{aligned}y = {} & 19.67 + 3.51\mathrm{RD} + (2.01 \times 10^{-4})\beta \\ & - \mathrm{PD} + (2.22 \times 10^{-5})\mathrm{P}. \qquad (10)\end{aligned}$$

Table 8. Prediction of built-up area growth for Pune city

Year	Built-up (sq, km)
2021	150.03
2031	161.25
2041	180.07
2051	212.27

Appendix 1. Coefficients of causal factors and percentage built-up area by linear regression analyses

Dependent variable (y)	Independent variable (x)	Model equation ($y = ax + b$)	R	R^2	Adjusted R^2	Standard error of estimate
Percentage built-up area	Population	$y = -(8.68 \times 10^{-5})x + 83.46$	0.241	0.058	–0.020	17.410
Percentage built-up area	α-Population density	$y = (9.29 \times 10^{-4})x + 39.28$	0.738	0.545	0.507	12.103
Percentage built-up area	β-Population density	$y = (9.36 \times 10^{-4})x + 46.57$	0.848	0.719	0.696	9.504
Percentage built-up area	Road density	$y = 4.11x + 22.28$	0.931	0.867	0.856	6.547
Percentage built-up area	Workforce in secondary and tertiary sectors	$y = -(2.90 \times 10^{-4})x + 88.71$	0.301	0.090	0.015	17.110
Percentage built-up area	Gender gap in literacy	$y = -(1.92 \times 10^{-3})x + 83.04$	0.551	0.304	0.246	14.968

Appendix 2. Coefficients of causal factors and percentage built-up area by quadratic regression analyses

Dependent variable (y)	Independent variable (x)	Model equation ($y = ax^2 + bx + c$)	R	R^2	Adjusted R^2	Standard error of estimate
Percentage built-up area	Population	$y = -(3.09 \times 10^{-11})x^2 - (7.22 \times 10^{-5})x + 81.81$	0.241	0.058	–0.113	18.184
Percentage built-up area	α-Population density	$y = -(1.06 \times 10^{-9})x^2 + (1.01 \times 10^{-3})x + 38.04$	0.738	0.545	0.462	12.638
Percentage built-up area	β-Population density	$y = -(2.15 \times 10^{-8})x^2 + (2.43 \times 10^{-3})x + 31.002$	0.909	0.827	0.796	7.794
Percentage built-up area	Road density	$y = -(3.65 \times 10^{-2})x^2 + 4.99x + 17.61$	0.931	0.867	0.843	6.822
Percentage built-up area	Workforce engaged in secondary and tertiary sectors	$y = (4.93 \times 10^{-9})x^2 - (1.16 \times 10^{-3})x + 125.76$	0.318	0.101	–0.062	17.764
Percentage built-up area	Gender gap in literacy	$y = (2.15 \times 10^{-7})x^2 - (6.46 \times 10^{-3})x + 102.06$	0.653	0.426	0.322	14.195

Appendix 3. Coefficients of causal factors and percentage built-up area by cubic regression analyses

Dependent variable (y)	Independent variable (x)	Model equation ($y = ax^3 + bx^2 + cx + d$)	R	R^2	Adjusted R^2	Standard error of estimate
Percentage built-up area	Population	$y = -(1.85 \times 10^{-16})x^3 - (1.19 \times 10^{-4})x + 88.25$	0.242	0.059	–0.113	18.181
Percentage built-up area	α-Population density	$y = -(8.67 \times 10^{-13})x^3 + (9.33 \times 10^{-8})x^2 - (1.83 \times 10^{-3})x + 88.25$	0.755	0.570	0.441	12.884
Percentage built-up area	β-Population density	$y = -(7.84 \times 10^{-14})x^3 - (1.37 \times 10^{-8})x^2 + (2.23 \times 10^{-3})x + 32.28$	0.910	0.827	0.775	8.168
Percentage built-up area	Road density	$y = -(3.65 \times 10^{-2})x^2 + 4.99x + 17.61$	0.931	0.867	0.843	6.822
Percentage built-up area	Workforce engaged in secondary and tertiary sectors	$y = (2.14 \times 10^{-14})x^3 - (8.09 \times 10^{-4})x + 117.98$	0.325	0.106	–0.057	17.719
Percentage built-up area	Gender gap in literacy	$y = -(5.81 \times 10^{-13})x^3 + (2.33 \times 10^{-7})x^2 - (6.63 \times 10^{-3})x + 102.43$	0.653	0.426	0.254	14.888

Appendix 4. Coefficients of causal factors and percentage built-up area by logarithmic regression analyses

Dependent variable (y)	Independent variable (x)	Model equation ($y = a \log x + b$)	R	R^2	Adjusted R^2	Standard error of estimate
Percentage built-up area	Population	$y = -19.14 \log x + 299.37$	0.233	0.054	–0.024	17.445
Percentage built-up area	α-Population density	$y = 26.60 \log x - 204.34$	0.695	0.483	0.440	12.903
Percentage built-up area	β-Population density	$y = 23.33 \log x - 160.17$	0.885	0.784	0.766	8.340
Percentage built-up area	Road density	$y = 45.85 \log x - 39.53$	0.927	0.859	0.848	6.726
Percentage built-up area	Workforce engaged in secondary and tertiary sectors	$y = -25.05 \log x + 347.892$	0.305	0.093	0.017	17.087
Percentage built-up area	Gender gap in literacy	$y = -16.27 \log x + 211.15$	0.626	0.392	0.341	13.993

Using eq. (10), the percentage built-up area for 2021 and 2051 is found to be 56.66% and 84.99% respectively (Table 8). This depicts that by 2051, built-up area in the municipal limits would rise to 212.27 sq. km, which may be nearly 50.0% more than that (141.50 sq. km) in 2011. The vegetation land, open space and areas around the highways are expected to become major targets for urban sprawl due to further increase in the pressure on land.

Appendix 5. Coefficients of causal factors and percentage built-up area by exponential regression analyses

Dependent variable (y)	Independent variable (x)	Model equation ($y = be^{ax}$)	R	R^2	Adjusted R^2	Standard error of estimate
Percentage built-up area	Population	$y = 80.48e^{-(1.17 \times 10^{-6})x}$	0.208	0.043	–0.037	0.275
Percentage built-up area	α-Population density	$y = 43.94e^{(1.29 \times 10^{-5})x}$	0.654	0.428	0.380	0.212
Percentage built-up area	β-Population density	$y = 48.25e^{(1.34 \times 10^{-5})x}$	0.775	0.600	0.567	0.178
Percentage built-up area	Road density	$y = 33.44e^{(6.06 \times 10^{-2})x}$	0.878	0.771	0.751	0.135
Percentage built-up area	Workforce engaged in secondary and tertiary sectors	$y = 86.20e^{-(3.88 \times 10^{-6})x}$	0.257	0.066	–0.012	0.271
Percentage built-up area	Gender gap in literacy	$y = 81.88e^{-(2.82 \times 10^{-5})x}$	0.517	0.268	0.207	0.240

Appendix 6. Coefficients of causal factors and percentage built-up area by power law regression analyses

Dependent variable (y)	Independent variable (x)	Model equation ($y = bx^a$)	R	R^2	Adjusted R^2	Standard error of estimate
Percentage built-up area	Population	$y = 1534.66x^{-0.26}$	0.203	0.041	–0.039	0.275
Percentage built-up area	α-Population density	$y = 1.39x^{0.376}$	0.628	0.394	0.343	0.219
Percentage built-up area	β-Population density	$y = 2.18x^{0.348}$	0.844	0.713	0.689	0.150
Percentage built-up area	Road density	$y = 13.11x^{0.687}$	0.888	0.788	0.770	0.129
Percentage built-up area	Workforce engaged in secondary and tertiary sectors	$y = 2968.49x^{-0.342}$	0.265	0.070	–0.007	0.271
Percentage built-up area	Gender gap in literacy	$y = 524.38x^{-0.236}$	0.580	0.337	0.282	0.229

Appendix 7. Stepwise regression equations

Stepwise regression equations	R^2
$PB = 22.28 + 4.11RD$	0.8668
$PB = 25.61 + 3.40RD + (2.06 \times 10^{-4})\beta - PD$	0.8760
$PB = 44.54 + 2.53RD + (1.47 \times 10^{-3})\beta - PD - (1.26 \times 10^{-3})\alpha - PD$	0.9216
$PB = 35.65 + 2.65RD + (1.52 \times 10^{-3})\beta - PD - (1.31 \times 10^{-3})\alpha - PD + (9.44 \times 10^{-6})WST$	0.9296
$PB = 35.16 + 2.65RD + (1.53 \times 10^{-3})\beta - PD - (1.31 \times 10^{-3})\alpha - PD + (2.17 \times 10^{-4})WST - (4.58 \times 10^{-5})P$	0.9303
$PB = 34.86 + 3.06RD + (1.50 \times 10^{-3})\beta - PD - (1.32 \times 10^{-3})\alpha - PD + (3.15 \times 10^{-4})WST - (1.19 \times 10^{-4})P + (5.05 \times 10^{-4})GGL$	0.9309

Future growth of Pune city is directed by key projects such as Pune International Airport (northeast Pune), development of townships like Megapolis and Blue Ridge in Hingewadi, Mumbai–Pune Expressway, Magarpatta city Information Technology township in Hadapsar (west Pune), Rajiv Gandhi Infotech Park in Hingewadi (east Pune), Bus Rapid Transit System (BRTS) in the city, and Delhi–Mumbai Infrastructure Corridor (northwest Pune). The boundaries of the municipal limit are also expected to change in the coming years in the city and Pune Metropolitan Region. These future spatial growth centres are mainly around employment and major transportation nodes. The future growth of PMC area will be mainly towards east, west and north directions and not in the south because of hilly outcrops. The growth will also be restricted in the northeast direction due to the airport funnel area, and in the northeast direction because of PCMC.

Conclusion

In this article, urban growth pattern of Pune city over a period of two decades has been presented. The result shows that built-up area has increased from 43.22 sq. km in 1992 to 141.50 sq. km in 2011. This continuous increase in built-up area has surpassed the rate of population growth. A relationship between urban sprawl and some of its causative factors have been established using multivariate regression analysis. Analysis of the causal factors of urban growth collectively explains the 93.09% variation in it. The results of regression analysis depict that α-population density is the most significant variable in the urban growth pattern. It has been found that the amount of built-up land grew by 227.3% over the period of nearly 19 years and by 2051, the built-up area in the region would rise to 212.27 sq. km, which would be nearly 50.0% more than the sprawl of 141.50 sq. km in 2011. In addition, future research and development of Pune city would also be influenced by other causal factors such as socio-economic change, future government investments corridors, development of small and medium towns around hinterland, industrialization, tourism initiatives, constrains of physical features, distances from major sites, etc.

1. UN, World Urbanization Prospects: The 2014 Revision, Department of Economic and Social Affairs, United Nations, New York, 2014, p. 2.

2. Jat, M. K., Garg, P. K. and Khare, D., Modeling of urban growth using spatial analysis techniques: a case study of Ajmer city (India). *Int. J. Remote Sensing*, 2008, **29**(2), 543–567.
3. Epstein, J., Payne, K. and Kramer, E., Techniques for mapping suburban sprawl. *Photogramm. Eng. Remote Sensing*, 2002, **63**(9), 913–918.
4. Batty, M., Xie, Y. and Sun, Z., The dynamics of urban sprawl. Working Paper Series, No. 15, Centre for Advanced Spatial Analysis, University College, London, 1999.
5. Torrens, P. M. and Alberti, M., Measuring sprawl. Working Paper No. 27, Centre for Advanced Spatial Analysis, University College, London, 2000; http://www.casa.ac.uk/working papers/.
6. Hurd, J. D., Wilson, E. H., Lammey, S. G. and Civco, D. L., Characterization of forest fragmentation and urban sprawl using time sequential Landsat Imagery. In Proceedings of the ASPRS Annual Convention, St. Louis, MO, USA, 23–27 April 2007.
7. Jantz, C. A., Goetz, S. J. and Scott, J., Analysis of scale dependencies in an urban land-use-change model. *Int. J. Geogr. Inf. Sci.*, 2005, **19**(2), 217–241.
8. Yang, X. and Liu, Z., Use of satellite derived landscape imperviousness index to characterize urban spatial growth. *Comput., Environ. Urban Syst.*, 2005, **29**, 524–540.
9. Batty, M. and Howes, D., Predicting temporal patterns in urban development from remote imagery. In *Remote Sensing and Urban Analysis* (eds Donnay, J. P., Barnsley, M. J. and Longley, P. A.), Taylor and Francis, London, pp. 185–204.
10. Clarke, K. C., Parks, B. O. and Crane, M. P., *Geographic Information Systems and Environmental Modeling*, Prentice Hall, New Jersey, 2002.
11. Donnay, J. P., Barnsley, M. J. and Longley, P. A. (eds), In *Remote Sensing and Urban Analysis*, Taylor and Francis, London, 2001, pp. 3–18.
12. Herold, M., Menz, G. and Clarke, K. C., Remote sensing and urban growth models – demands and perspectives. In Symposium on Remote Sensing of Urban Areas, Regensburg, Germany, 2001, vol. 35.
13. Jensen, J. R. and Cowen, D. C., Remote sensing of urban/suburban infrastructure and socio-economic attributes. *Photogramm. Eng. Remote Sensing*, 1999, **65**(5), 611–622.
14. Sudhira, H. S., Ramachandra, T. V. and Jagadish, K. S., Urban sprawl: metrics, dynamics and modelling using GIS. *Int. J. Appl. Earth Obs.*, 2004, **5**, 29–39.
15. Haack, B. N. and Rafter, A., Urban growth analysis and modelling in the Kathmandu valley, Nepal. *Habitat Int*., 2006, **30**(4), 1056–1065.
16. Gomarasca, M. A., Brivio, P. A., Pagnoni, F. and Galli, A., One century of land use changes in the metropolitan area of Milan (Italy). *Int. J. Remote Sensing*, 1993, **14**(2), 211–223.
17. Green, K., Kempka, D. and Lackey, L., Using remote sensing to detect and monitor land-cover and land-use change. *Photogramm. Eng. Remote Sensin*g, 1994, **60**, 331–337.
18. Yeh, A. G. O. and Li, X., Measurement and monitoring of urban sprawl in a rapidly growing region using entropy. *Photogramm. Eng. Remote Sensing*, 2001, **67**(1), 83–90.
19. Yang, X. and Lo, C. P., Modelling urban growth and landscape changes in the Atlanta metropolitan area. *Int. J. Geogr. Inf. Sci*., 2003, **17**(5), 463–488.
20. Lo, C. P., Modeling the population of China using DMSP operational Linescan system nighttime data. *Photogramm. Eng. Remote Sensing*, 2001, **67**, 1037–1047.
21. Lo, C. P. and Yang, X., Drivers of land-use/land-cover changes and dynamic modelling for the Atlanta, Georgia metropolitan area. *Photogramm. Eng. Remote Sensing,* 2002, **68**(10), 1062–1073.
22. Weng, Q., Modeling urban growth effects on surface runoff with the integration of remote sensing and GIS. *Environ. Manage.*, 2001, **28**(6), 737–748.
23. Cheng, J. and Masser, I., Urban growth pattern modelling: a case study of Wuhan City, PR China. *Landsc. Urban Plann.*, 2003, **62**, 199–217.
24. Chabaeva, A. A., Civco, D. L. and Prisloe, S., Development of a population density regression model to calculate imperviousness. In ASPRS Annual Conference Proceedings, Denver, CO, USA, 2004.
25. Kumar, J. A. V., Pathan, S. K. and Bhanderi, R. J., Spatio-temporal analysis for monitoring urban growth – a case study of Indore city. *J. Indian Soc. Remote Sensing*, 2007, **35**(1), 11–20.
26. Jat, M. K., Garg, P. K. and Khare, D., Monitoring and modelling of urban sprawl using remote sensing and GIS techniques. *Int. J. Appl. Earth Obs. Geoinf.*, 2008, **10**, 26–43.
27. Punia, M. and Singh, L., Entropy approach for assessment of urban growth: a case study of Jaipur, India. *J. Indian Soc. Remote Sensing*, 2012, **40**(2), 231–244.
28. Rawat, J. S. and Kumar, M., Monitoring land use/cover change using remote sensing and GIS techniques: a case study of Hawalbagh block, district Almora, Uttarakhand, India. *Egypt. J. Remote Sensing Space Sci.*, 2015, **18**, 77–84.
29. Roshan, R., Shahraki, S. Z., Sauri, D. and Borna, R., Urban sprawl and climatic changes in Tehran, Iran. *J. Environ. Health. Sci. Eng*., 2010, **7**(1), 43–52.
30. Polyzos, S., Minetos, D. and Niavis, S., Driving factors and empirical analysis of urban sprawl in Greece, *Theor. Empirical Res. Urban Manage.*, 2013, **8**(1), 5–29.
31. Majid, F. and Mohammad, M., Dynamics and forecasting of population growth and urban expansion in Srinagar city – a geospatial approach. *Int. Arch. Photogramm., Remote Sensing Spatial Inf. Sci.*, 2014, **11**(8), 709–716.
32. Andrew, M., Twumasi, Y. A., Shou, L. K. and Coleman, T. L., Predicting urban growth of a developing country city using a statistical modeling approach. *Int. J. Geomat. Geosci*., 2015, **5**(4), 603–613.
33. Goswami, M. and Khire, M. V., Land use and land cover change detection for urban sprawl analysis of Ahmadabad city using multitemporal landsat data. *Int. J. Adv. Remote Sensing GIS*, 2016, **5**(4), 1670–1677.
34. Anderson, J. R., Hardy, E. E., Roach, J. T. and Witmer, R. E., A land use and land cover classification system for use with remote sensor data. US Geological Survey Professional Paper, No. 964, USGS, Washington, DC, 1976, USA, p. 28.

Received 4 May 2016; revised accepted 6 July 2017

doi: 10.18520/cs/v114/i10/2080-2091