AI-based Built-up Area Extraction

Introduction

India, with an urban population of 329.8 million (Census of India 2011), has World's second largest urban population, accounting for over 10.0 % of the global urban population (United Nations 2015). The urban population, which accounts for 31.2% of total population of India, has witnessed over 31.8% growth rate during 2001-2011 decade, far outpacing the rural growth rate of 12.2% (Census of India 2011). This has also contributed towards an increase in the total number of urban settlements in the country from 5,161 in year 2001 to 7,935 in year 2011 (Census of India 2011). While cities are growing at such a rapid pace, the corresponding efforts to plan and mitigate the effects of increasing urbanisation are few and largely confined to major metropolitan areas. Much of the urban growth in the country is unplanned and haphazard, leading to loss of agricultural land, costly and inefficient infrastructure, high cost of transportation, and deteriorating quality of life. The artificial surfaces are very dynamic and therefore require regular update. Built-up area serves as an indicator of socio-economic growth of a country, besides being an alternate for population growth of a region. Furthermore, several studies require built-up area for quantifying anthropogenic impacts on environment and climate change.

Remote sensing has been widely used for mapping urban sprawl and analysing the growth pattern of urban areas. Landsat data with 30 m spatial resolution has extensively been used in several urban sprawl studies all over the World. Landsat data is being further supported by the Sentinel-2 data with 12 m spatial resolution, Indian Remote Sensing (IRS) satellite data with 23.5 m resolution data acquired by the Linear Imaging Self Scanner (LISS-3) sensor and 5.8 m resolution data acquired by LISS-4 sensor, and the 20 m spatial resolution SPOT data acquired by the French satellite. Banzhaf et al. (2009) recommended 10 m SPOT 5 data as more suitable alternative for monitoring urban growth in European cities. Several studies have also attempted to understand the growth dynamics and spatial structure of cities using various sprawl metrics derived from the coarse resolution remote sensing data. Zhu et al. (2006), however, observed that 7.5 m or better spatial resolution is optimal for urban fragmentation analysis. The studies on urban sprawl have extensively relied upon the multi-date coarse resolution data acquired by Landsat series satellites, particularly owing to its consistent data quality and the continuity of data availability since 1978. The Landsat data, however, is sub-optimal for urban analysis (Zhu et al. 2006, Banzhaf et al. 2009). The primary aim of this study is to map urban sprawl using 5.8 m or better spatial resolution satellite images acquired by Indian Remote Sensing satellites. This spatial resolution is not only suitable for urban sprawl studies as indicated by previous researchers, but is also effective in capturing small scattered growth in urban periphery. The comparative analysis of urban sprawl in multiple cities, however, requires application of uniform methodology and common datasets to ensure consistent interpretation of results, as intended in this study.

The information on urban growth in India is currently available from several independent isolated studies that are difficult to compare and fail to provide holistic picture of urbanisation in the country. The understanding of growth processes in Indian cities, both in space and time domain, is constrained by the non-availability of reliable baseline data. Mumbai (Pathan et al. 1993), Indore (Kumar et al. 2007), Mangalore (Sudhira et al. 2004), Ajmer (Jat et al. 2008), Kolkata (Bhatta 2009), Bangalore (Ramachandra et al. 2012), Jamnagar (Jain et al. 2014) and Pune (Kantakumar et al. 2016) are few of remote sensing data based studies on Indian cities. In addition to these single-city independent studies, many national and global level land-use land-cover (LULC) and settlement datasets are available that provide built-up area. These studies however provide widely varying estimate of built-up surfaces, which can be very misleading as shown in Figure 1. These datasets used varying satellite data and methodologies. The ArcGIS LULC uses Sentinel-2 data (10 m GSD), Global Human Settlement Layer (GHSL) uses Landsat data (30 m GSD), NRSC LULC uses LISS-3 data (23 m GSD) and NRSC National Resource Census uses AWiFS data (56 m GSD). These studies have relied on coarse resolution data acquired by that has either resulted in omission of small settlements or commission of fragmented open spaces.

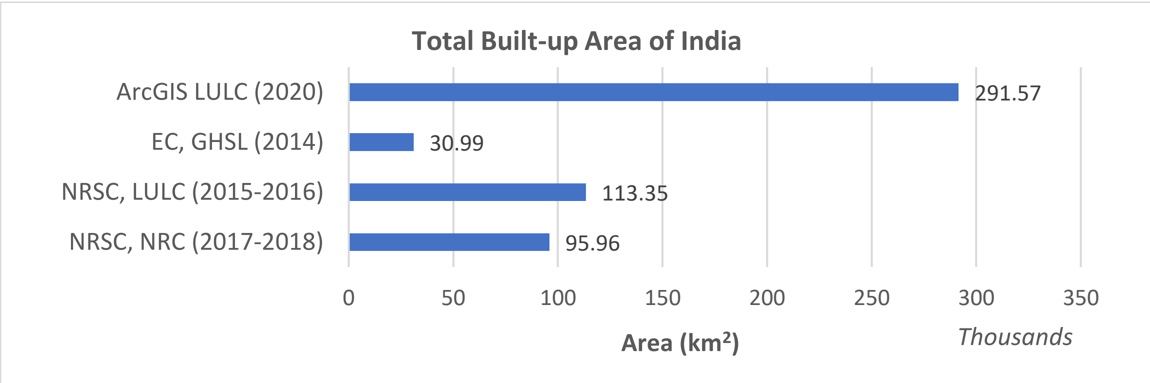


Figure 1: Built-up Area in Multiple Cities

This study intends to estimate built-up area of all class-1 cities of India, including cities identified under the Government of India's flagship AMRUT programme, using DL/ML technique from LISS-4 data. Indian Space Research Organisation (ISRO) has undertaken several projects on mapping Indian cities and creating systematic database of urban areas. Several past and ongoing projects such as Large-Scale Mapping (35 cities mapped at 1:10,000 scale for year 2005), National Urban Information System (138 cities mapped 1:10,000 scale for 2010), and the more recently AMRUT (500 cities at 1:4000 scale) are intended to create large scale base maps, appropriate for preparation of Development Plans. Updating of these maps, however, will be time consuming process considering the level of details incorporated in these datasets. The acquisition of historical imageries will be equally challenging in case of very large-scale mapping projects. The aim of this study is to design and develop an urban sprawl information system for monitoring urbanisation of 500 Indian cities proposed under AMRUT mission to help in understanding impact of urbanisation on urban environment and quantify landscape fragmentation using space-based inputs and GIS. The study will utilize multi-date optical remote sensing data acquire by Indian Remote Sensing Satellites with spatial resolution of 5.8 m or better to ascertain the spatial and temporal variations in the evolution of 500 Indian cities and Class-1 cities. It will further attempt to characterize urban morphology by considering different aspects of urban sprawl in next phase, quantifying sub-urban areas, fringe open spaces, open space fragmentation, etc.

Objectives

Following are the major project objectives:

1) Extraction of built-up area of Class-1 cities of India from Resourcesat-2 LISS-IV data for 2023-2024 timeframe using Artificial Intelligence (AI).
2) Development of web-based information system for urban sprawl and enabling comparison across multiple cities.

Methodology

Pre-processing of Satellite Images

The Resourcesat-2/2A LISS-4 images were pre-processed before extracting AI-based built-up area. In total, 445 satellite images were downloaded from NRSC Bhuvan website and processed to extract built-up area for 484 cities. For each city, bands were initially georeferenced using Cartosat-2 series data with approx. 1-pixel accuracy. These bands were converted from DN to TOA Reflectance using the band meta data information available with the individual satellite scene. It involves DN value at each pixel, sun's position, solar irradiance, cosine of the solar zenith angle for calculating spectral radiance values (Sharma et. al, 2008) using below equation (1).

Spectral Radiance ($W/m^2/sr/\mu m$)= (DN- Bias)* (Gain)/(Solar Irradiance)*(cos(solar zenith angle))---(1)

Where, DN is the digital number of the LISS IV pixel, Bias is the offset value for the sensor, Gain is the scaling factor for the sensor and cos(solar zenith angle) is the cosine of the solar zenith angle. The TOA reflectance bands were stacked in standard False color composition (FCC) and clipped to the city's ROI. Cities FCC image in geotiff file format with 8-bit unsigned integer is used for training and extracting built-up area from AI-model.

AI-based Built-up Area Extraction

The spatial database of urban sprawl of all Class- 1 cities of India (including all cities identified under the AMRUT) has been created using IRS Resourcesat-2/2A LISS-4 images acquired during 2023-2024 timeframe. The methodology for extraction of built-up area from LISS-4 images uses Deep Learning-based image segmentation model. Convolutional Neural Networks (CNNs) have been successfully used for extracting information from satellite imagery. The CNN architecture based on UNet and Atrous Spatial Pyramid Pooling (ASPP) concept, termed as UNet-AP (Rastogi et.al, 2020) has been utilized. The model uses supervised image segmentation techniques, where each pixel is classified in pre-defined class (urban or non-urban). Satellite based impervious surface extraction is done from the proposed UNet-AP model as shown in Figure 2.

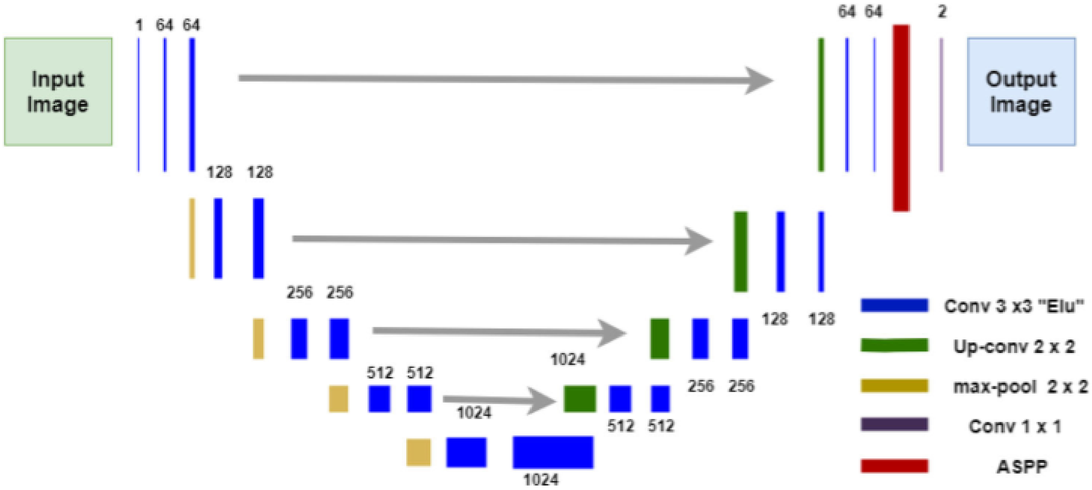


Figure 2: UNet-AP Deep Learning Hybrid Model for Delineating Built-up Areas from LISS-4 Image

The model uses hybrid model UNet and Deeplab V3+ Atrous Spatial Pyramid Pooling network. The model has the capability to extract built-up surface from small, medium and large cities built upon varied terrain conditions. The multi-level window of ASPP in the model helps in delineating the exact boundaries of manmade structures of varied shape size from complex urban environment. The built-up area thus extracted was checked for its quality.

Quality Checking (QC) of AI based built-up area extraction

Quality checking is required to ensure correctness of output result before its utilization for further analysis. AI based extracted built-up area is in the form of raster data which is checked with respect to original satellite image. Area having built-up probability greater than 25% was used to check whether it is matching with respective image showing built-up area by visual interpretation technique. For some of the cases built-up area probability needs to change (increase or decrease) for better built-up area shape matching. If shape of the city and built-up area extracted is almost matching with image data, then it passes the QC, else it fails. Figure 3 and Figure 4 shows examples of cities where QC passed and failed respectively. The cities where quality of built-up area extraction was acceptable, were subjected to post-processing.

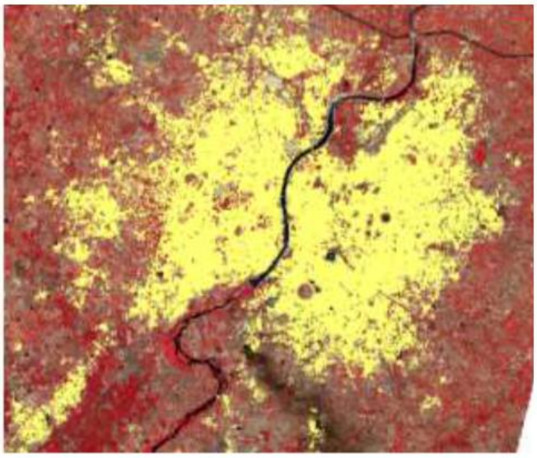


Figure 3: QC Passed AI-extracted Built-up Area of Ahmedabad City, Gujarat

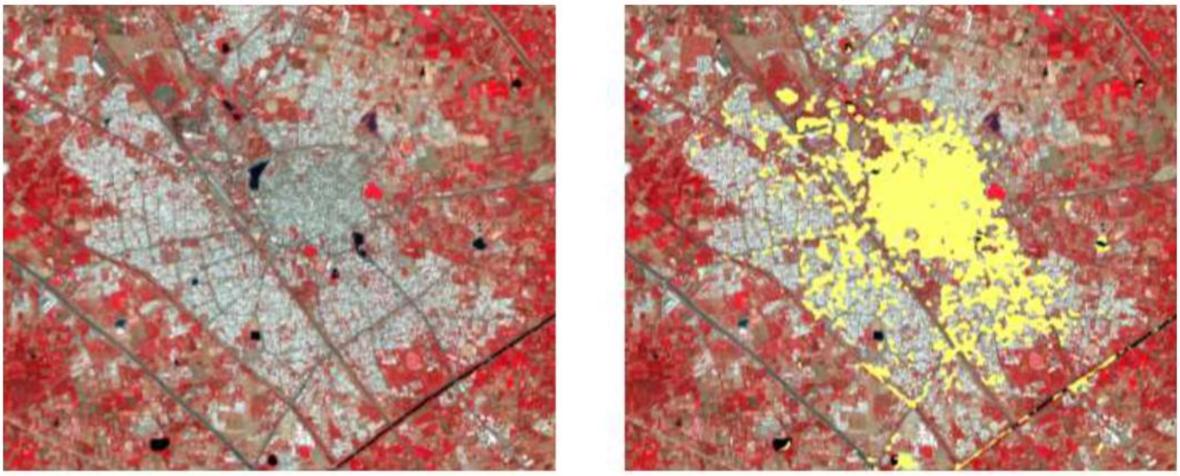


Figure 4: QC Failed AI-extracted Built-up Area of Nadiad City, Gujarat

Table 1 and Table 2 shows different types of omission and commission-errors detected in the AI based extracted output. The commission errors included (i) Salt pan extracted as built-up area; (ii) Mining area extracted as built-up area; (iii) River part extracted as built-up area; (iv) Sea-beach extracted as built-up area; and (v) Scrub land extracted as built-up. The omission errors included (i) Sparse built-up; (ii) Industrial built-up area; (iii) Built-up with vegetation.

Table 1: Examples of Commission Errors

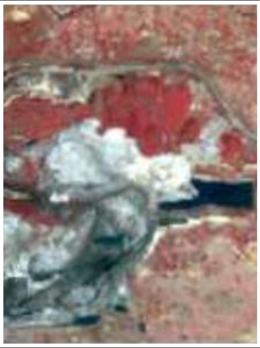
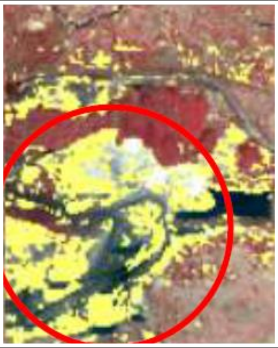
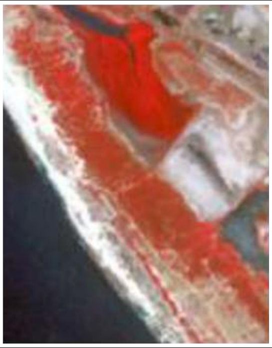
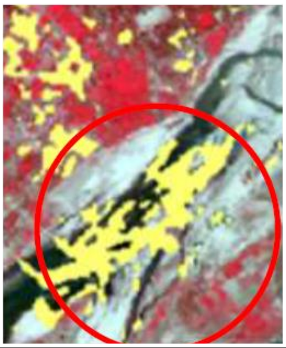
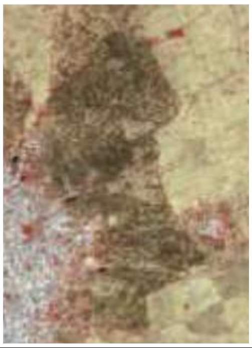
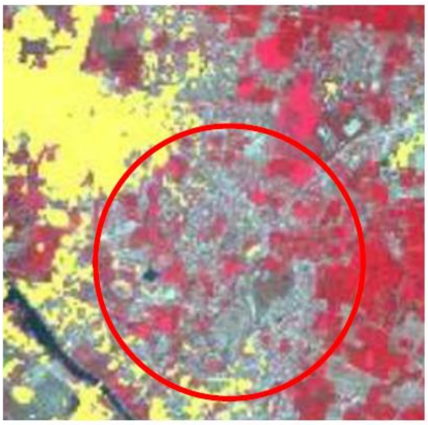
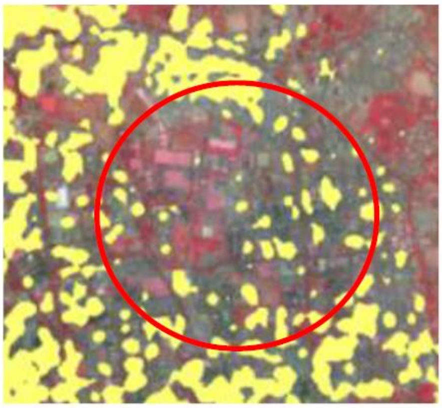
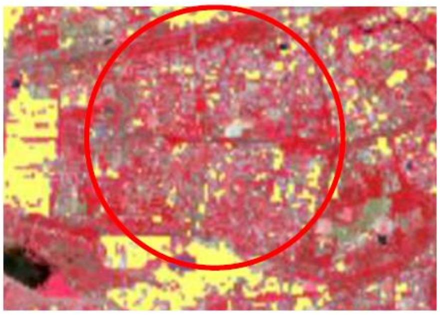
Sr. No.	Error Description	Image	Classification
1	Salt pan extracted as built-up area		
2	Mining area extracted as built-up area		
3	Sea-beach extracted as built-up area		
4	River bed extracted as built-up area		
5	Scrub land extracted as built-up		

Table 2: Examples of Omission Errors

Sr. No.	Description	Image with Classification
1	Sparse built-up not detected	
2	Industrial built-up not detected	
3	Built-up with vegetation not detected	

Post processing of AI based built-up output

The post processing involved vectoring the data, applying simplification and eliminating polygons smaller than 5625 sq. meter area. Following procedure was adopted for post-processing of AI based output for finalization of built-up area layer and ready to publish on VEDAS portal.

1. Select built-up probability output image and generate a binary raster with pixels greater than 25% probability of being built-up pixels, assigned value of one and others assigned value of zero. This will generate a binary built-up raster. A lower value of probability may be defined in case the city has large areas with sparse built-up.
2. Convert built-up raster to vector layer.

3. Add Area field and calculate area in square meters.
4. Select polygons with area less than 5,625 sq. m, corresponding to Minimum Mappable Unit (MMU) for 3 mm x 3 mm at 1:25,000 scale. Eliminate selected polygons in GIS.
5. Reselect all built-up area polygons and remove all non-built-up polygons.
6. Edit built-up area layer to remove commission errors by selecting such polygons and delete them manually after visual interpretation.
7. Add omission areas by visual interpretation techniques wherever required.
8. For smoothening of boundaries, use Simplify function in GIS (Tolerance = 5 m)
9. Convert projection parameters to geographic coordinate system.
10. The layers thus accepted were published on the VEDAS portal (https://vedas.sac.gov.in/urban).

WebGIS based Urban Sprawl Information System (USIS)

A WebGIS-based Urban Sprawl Information System (USIS) is designed to monitor, analyze, and visualize urban sprawl. The current system architecture is shown in Figure 5. Developed system uses AI extracted builtup area output to track the expansion of urban areas, providing tools for decision-makers, urban planners, and policymakers to better understand and manage urban growth enabling the assessment of temporal changes leveraging advanced GIS technologies. It also integrates other geospatial datasets including high resolution satellite imagery from BHUVAN (Bhuvan), LISSIV datasets of all cities, Sentinel 2 and AWiFS FCC, administrative boundaries, SISDP LULC and transportation infrastructure for value addition. Builtup area of 500 cities are available in shape files format having polygon features containing area and city code as attribute. These spatial datasets are merged statewise and organised into spatial database. This database is designed using PostgreSQL for storing and managing spatial datasets including builtup area of all cities. PostgreSQL is an open-source, spatial relational database management system which is designed for handling large amounts of spatial and non-spatial data. It is known for its robustness, scalability, and compliance with SQL standards. It supports spatial (geometry), relational (tables, rows, columns) and non-relational (JSON, key-value) data models (Postgres).

To publish spatial datasets in database as Open Geospatial Consortium (OGC) compliant Web Map Service (WMS), WebGIS package GeoServer (version 2.26) is used. GeoServer (GeoServer) is a Free and Open Source Software (FOSS) for publishing and sharing geospatial data using open standards such as Web Map Service (WMS), Web Feature Service (WFS), and Web Coverage Service (WCS) (GeoServer). Developed application aims to provide analytical tool to track the expansion of urban areas in selected City / State. To achieve this, REST API is developed in Python Flask to provide builtup area statistics for all cities for selected State or India. API is also developed to provide builtup area statistics for specific city also. Application also provides other GIS functionalities i.e. feature information, measure distance / area, upload KML, GeoJSON or Shapefile. User Interface is developed using HTML, JavaScript libraries i.e. VueJS (VueJS), OpenLayers etc. OpenLayers (OpenLayers) is an open-source JavaScript library used for displaying maps and geographic data in web applications, provide GIS tools i.e. mouse position, zooming, panning, calculating distance/area etc on the web.

When User select State, City and dataperiod for accessing a builtup area map or online data analysis, it will reach the web server. Here, Apache Tomcat (version 9) web server is used. Tomcat parses the request and forwards it to desired server. To access online map, request is forwarded to GeoServer. For spatial data and its analysis REST based Application Programming Interface (API)s are developed. For analysing built area statistics using graphs and tables, APIs are called. These flasks based APIs fetch data from spatial databases and perform required calculations i.e. citywise builtup area for selected state, builtup area for selected city, top ten statistics etc. These results are disseminated to users using developed Interface. The GeoServer serves WMS services while data request for graphs and tables are served through APIs from PostgreSQL database according to user request. Subsequently the Web server sends the results in the form of GIS enabled images or APIs output in JSON format to the user Interface.

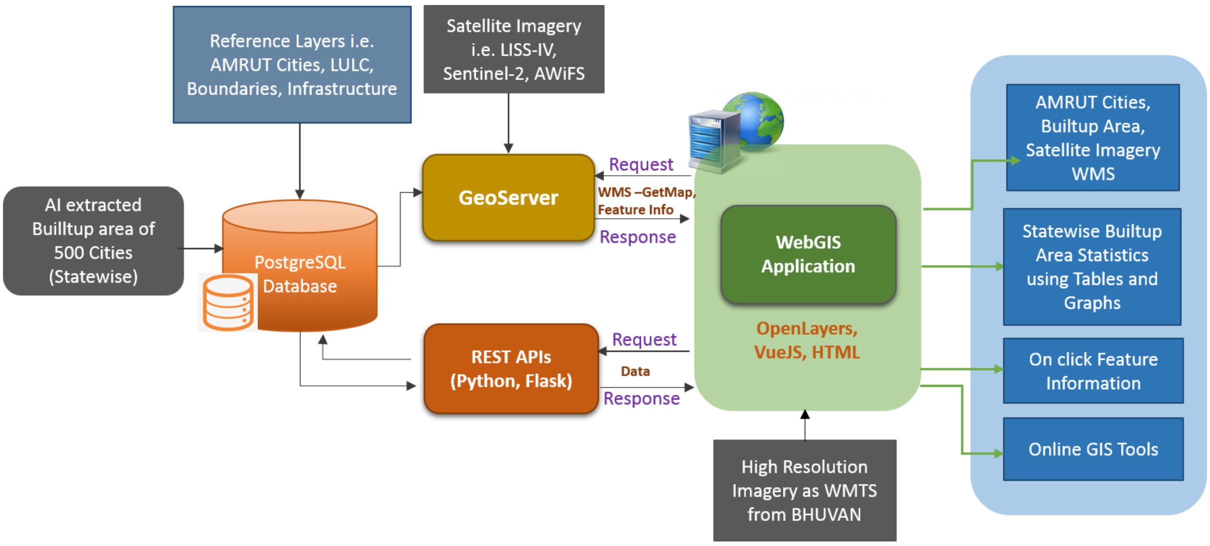


Figure 5: WebGIS Architecture of Urban Sprawl Information System

Results

AI-derived Built-up Area

The study extracted built-up area of all 500 cities of India from LISS-4 data using AI-based approach. Figure 6 shows the cities for which the built-up area for 2023-2024 timeframe has been extracted as green dot (484 cities), whereas 16 cities marked with red dot, are remaining due to quality issues. Figure 7 shows the built-up area of all cities of Punjab state.

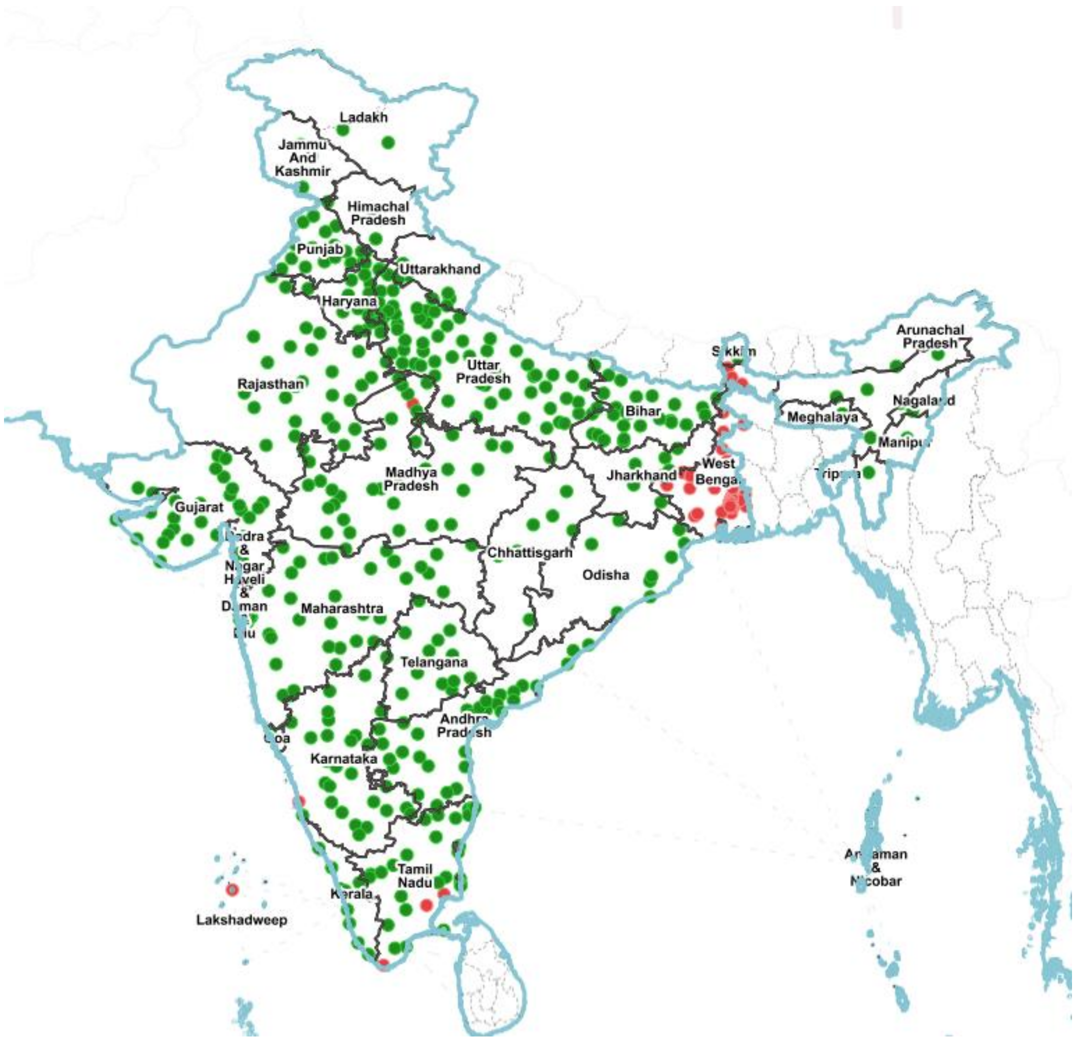


Figure 6: Cities with Built-up Area Extraction Completed

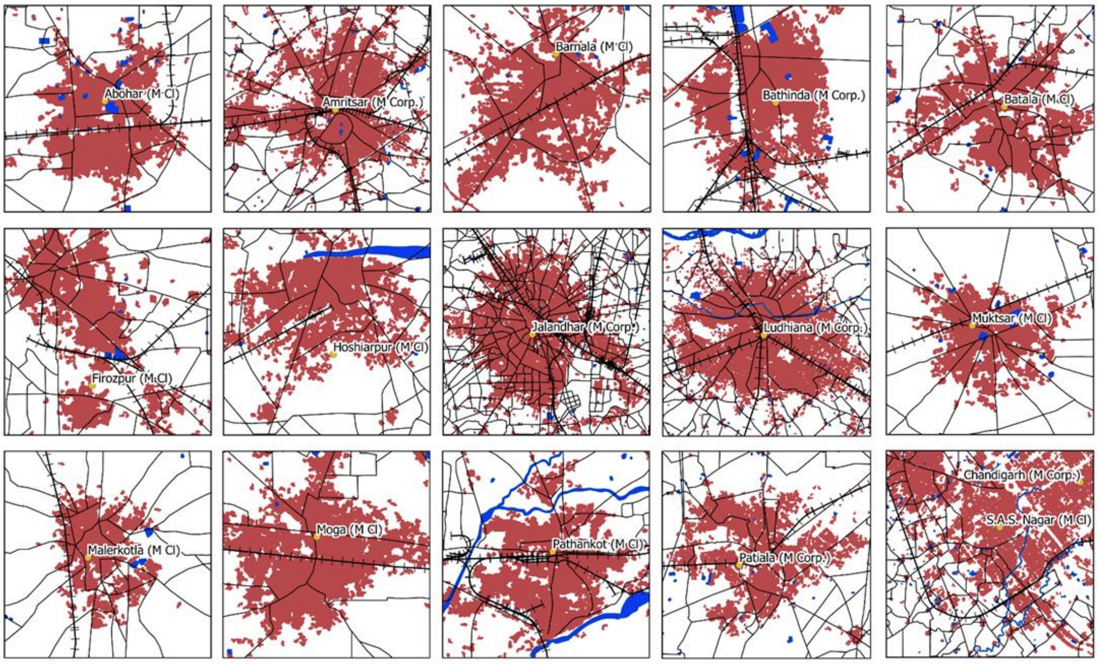


Figure 7: AI-derived Built-up Area of AMRUT Cities of Punjab State

Urban Sprawl Information System

A web-based application on VEDAS portal was created to enable visualisation and analysis of built-up area of 484 cities. It enables state and national level comparison of built-up area of cities. Figure 8 shows built-up area of cities of Gujarat on the web application. It shows the area of top ten cities in the state in term of built-up area as bar chart and a table containing built-up area of all cities in Gujarat for comparison.

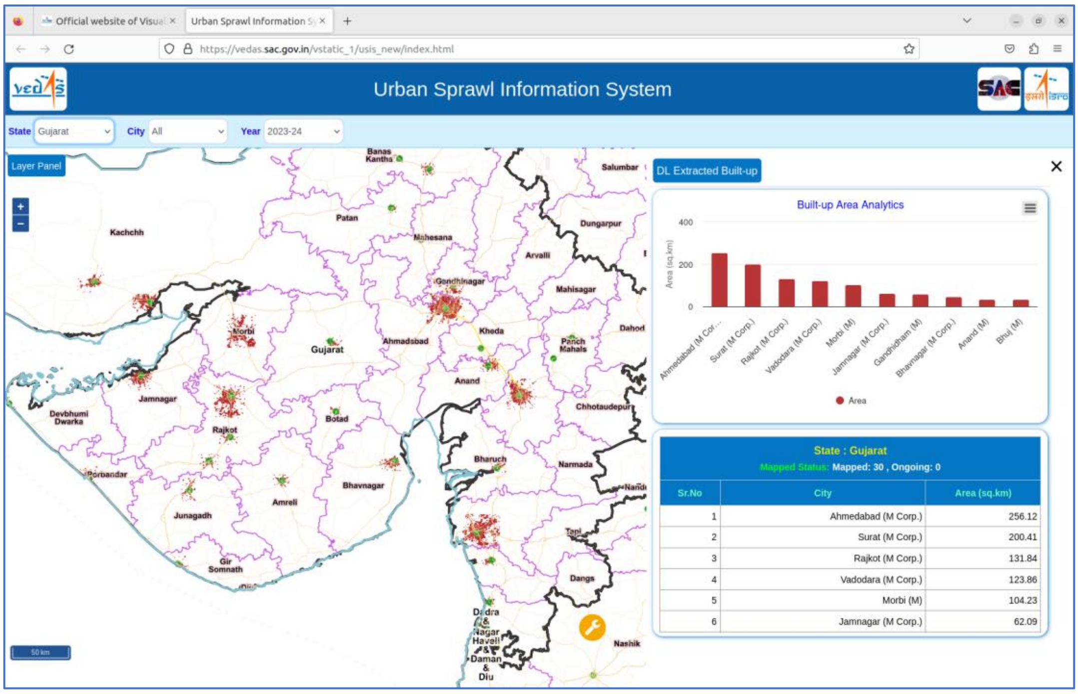


Figure 8: Cities of Gujarat as viewed on urban sprawl information system

Figure9 & Figure10 shows built-up area of Ahmedabad and Bengaluru city. Figure9 shows built-up area superimposed on the Very High Resolution Base Map fetched from Bhuvan portal of ISRO. Several indicators of urban form and related spatial metrics can be computed using this built-up data.

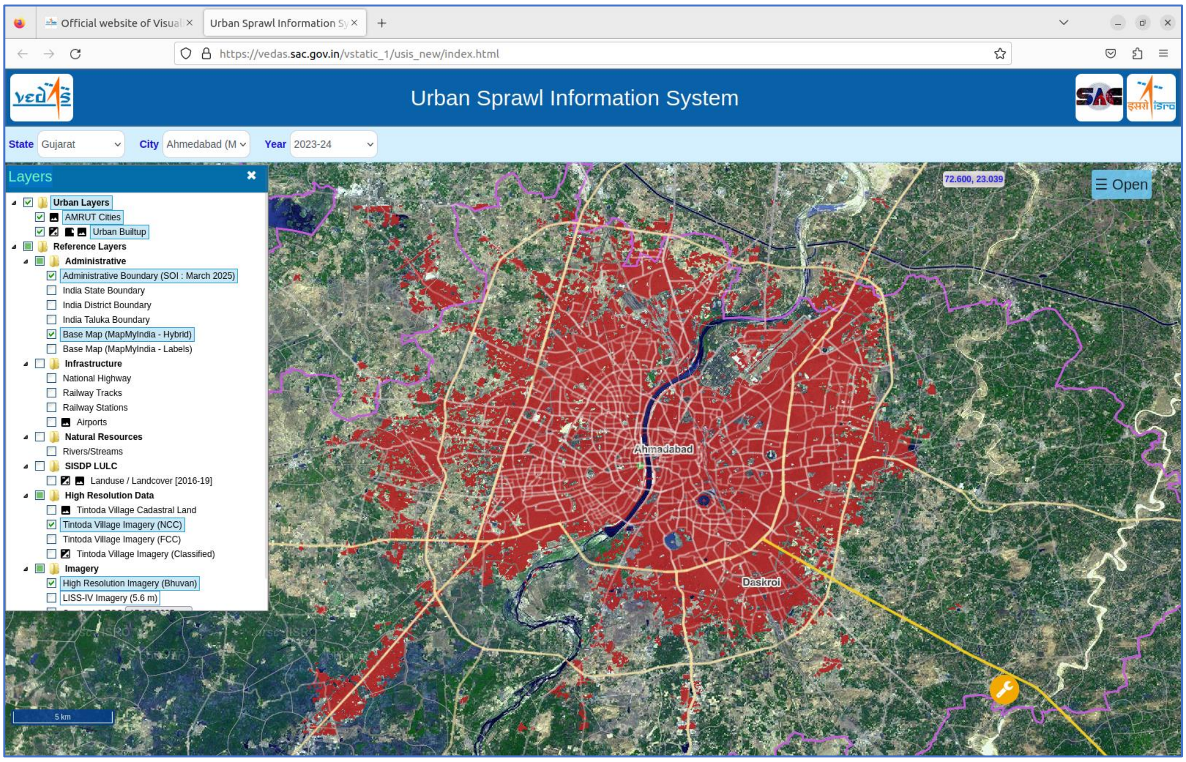


Figure 9: AI-extracted Built-up Area of Ahmedabad city overlaid on Bhuvan Base Map

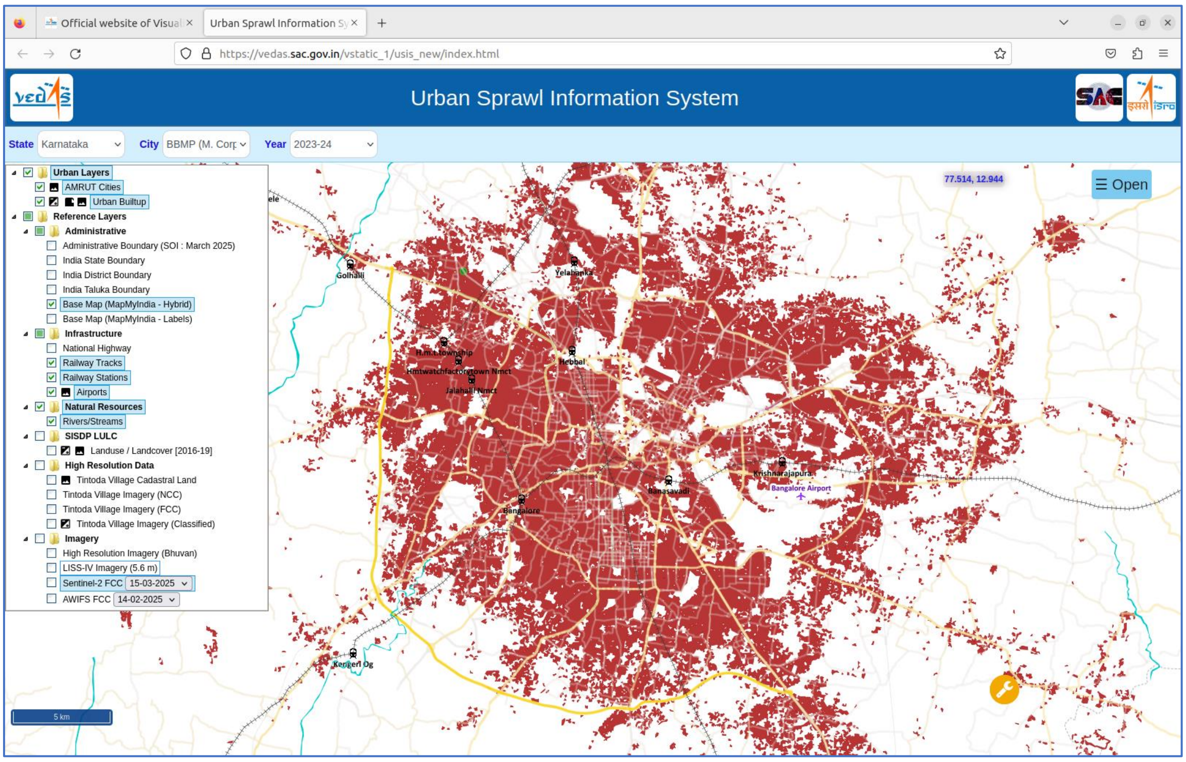


Figure 10: Bengaluru City AI-extracted Built-up Area

Conclusion and Way Forward

The AI-based model performed well in almost all cities of India. However, in few cities where the built-up area was covered with vegetation and water bodies, such as cities in West Bengal, or where the proportion of such sparse built-up area was high, the quality checking of AI-model results were not satisfactory. In such cases, alternate techniques and datasets are being considered. Figure 1 shows application of Support Vector Machine (SVM) for built-up area extraction in Asansol city of West Bengal. Similarly, attempts are being made to use alternate data such as fusion of microwave data with optical data for identifying built-up mixed with vegetation.

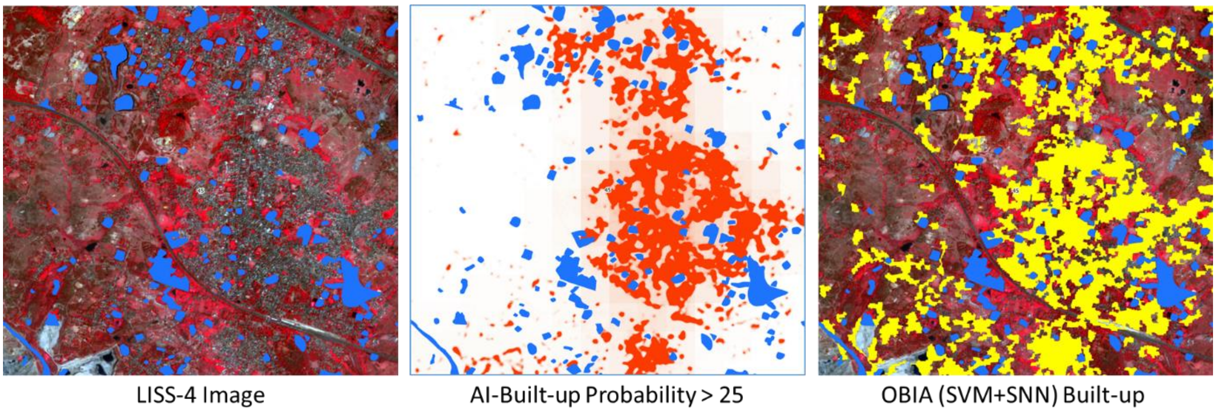


Figure 11: Alternate Approaches for Built-up Area Extraction to improve Accuracy

The built-up area thus extracted will be used as input for urban landscape analysis tool (ULAT) to estimate urban areas and urban footprints, further classifying built-up area as core urban built-up, sub-urban built-up and rural built-up area, and the surrounding open spaces as fringe open space, captured open space and rural open space. This will be used to compute percent of sub-urban area and open space fragmentation indices of all cities using ULAT Tool as shown in 12 for Amreli, Anand and Bhuj cities of Gujarat.

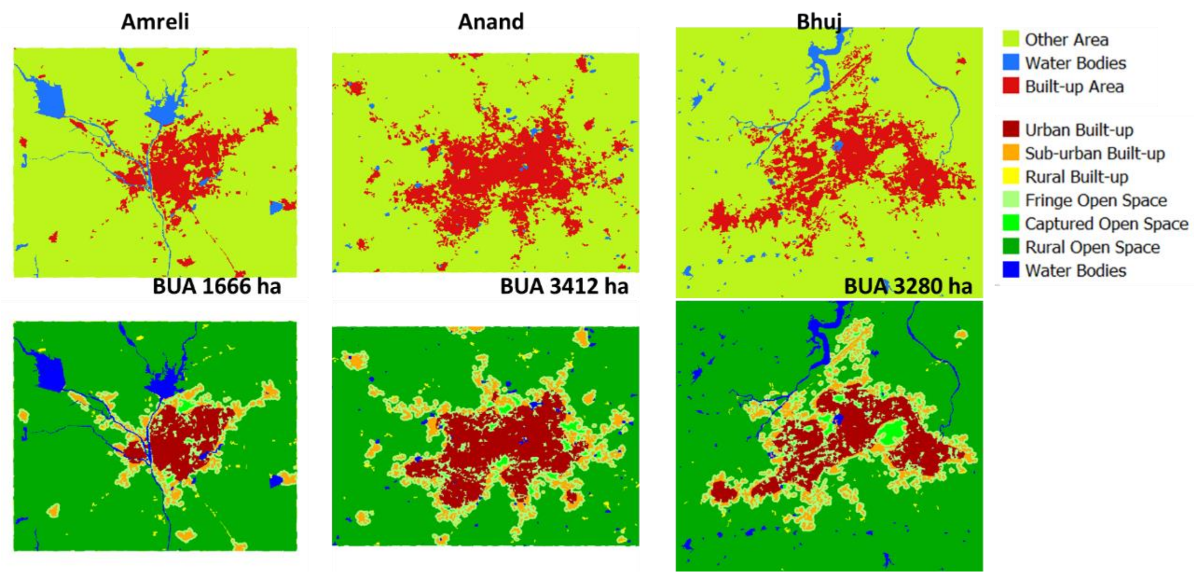


Figure 52: Urban Landscape Analysis

The project will subsequently also utilize the LISS-4 data of 2018-2019 timeframe to extract built-up area using AI-model and assess urban sprawl in all these cities in future. A web-based application on VEDAS portal was created to enable visualisation and analysis of built-up area. It enables state and national level comparison of built-up area of cities. The Urban Sprawl Information System web-application provides interactive map of built-up area of all 484 cities completed so far, a bar chart showing top ten cities in terms of built-up area (within a state or in India), tabular list of all cities containing built-up area for comparison (within a state or in India), overlaying of multiple base maps and spatial information layers such as Bhuvan and MapMyIndia, recent satellite imageries from Sentinel-2 and AWiFS, etc. The application will be useful in managing urbanisation in the country. It will not only help urban planners in Central Government, State Governments and Urban Local Government Bodies in understanding urban morphology and identifying growth drivers, but will also assist researchers, academia and citizens in working towards sustainable urbanisation.

References

1. Banzhaf, E., Grescho, V. and Kindler, A. 2009. Monitoring urban to peri urban development -urban development with integrated remote sensing and GIS information: a Leipzig, Germany case study, International Journal of Remote Sensing, 30:7, pp. 1675-1696, DOI: 10.1080/01431160802642297.
2. Bhatta, B., 2009, Analysis of urban growth pattern using remote sensing and GIS: a case study of Kolkata, India. International Journal of Remote Sensing, 30:18, pp. 4733-4746, DOI: 10.1080/01431160802651967.
3. Census of India, 2011, Primary Census Abstracts (Census of India 2011), Office of Registrar General of India, Government of India, New Delhi.
4. Jain, G., Sharma, S., Vyas, A., and Rajawat, A. S. 2014. Characterizing multi-dimensionality of urban sprawl in Jamnagar, India using multi-date remote sensing data, in The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-8, pp. 761-766. DOI: 10.5194/isprsarchives-XL-8-761-2014.
5. Jat, M. K., Garg, P. K. and Khare, D. 2008. Modelling of urban growth using spatial analysis techniques: a case study of Ajmer city (India). International Journal of Remote Sensing, 29:2, pp. 543-567, DOI: 10.1080/01431160701280983.
6. Kantakumar, L. N., Kumar, S., Schneider, K. 2016. Spatiotemporal urban expansion in Pune metropolis, India using remote sensing, Habitat International, 51, pp. 11-22, DOI: 10.1016/j.habitatint.2015.10.007.
7. Kumar, J. A. V., Pathan, S., and Bhanderi, R. 2007. Spatio-temporal analysis for monitoring urban growth - A case study of Indore city, Journal of the Indian Society of Remote Sensing, 35 (1), pp. 11–20.
8. Pathan, S. K., Sastry, S. V. C., Dhinwa, P. S., Rao, M., Majumdar, K., L., Kumar, S., D., Patkar, V. N., Phatak, V. N. 1993. Urban growth trend analysis using GIS techniques—a case study of the Bombay metropolitan region, International Journal of Remote Sensing, 14(17), pp. 3169- 3179.

9. Ramachandra, T. V., Aithal, B. H. and Sanna, D. D. 2012. Insights to urban dynamics through landscape spatial pattern analysis, International Journal of Applied Earth Observation and Geoinformation, 18, pp. 329–343, DOI:10.1016/j.jag.2012.03.005.
10. Rastogi, K., Bodani, P., Sharma, S. 2020. Automatic building footprint extraction from very high-resolution imagery using deep learning techniques, Geocarto International, DOI: 10.1080/10106049.2020.1778100
11. Sharma, Anu, Badarinath, K. and Roy, Parth. (2008). Corrections for atmospheric and adjacency effects on high resolution sensor data a case study using IRS-P6 LISS IV data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B8. Beijing 2008.
12. Sudhira, H.S., Ramachandra, T.V. and Jagadish, K.S. 2004. Urban sprawl: metrics, dynamics and modelling using GIS. International Journal of Applied Earth Observation and Geoinformation, 5, pp. 29-39, DOI: 10.1016/j.jag.2003.08.002.
13. Zhu, M., Jiang, N., Li, J., Xu, J. and Fan, Y. 2006. The effects of sensor spatial resolution and changing grain size on fragmentation indices in urban landscape, International Journal of Remote Sensing, 27:21, pp. 4791-4805, DOI: 10.1080/01431160600702442.