P4 - Spatiotemporal Analysis of Urban Expansion Using NDBI from Satellite Imagery

by

Pranav Guruprasad Rao
B.Tech. (Hons.) in CSE (2025)
IIT Palakkad

INDIA SPACE ACADEMY

SUMMER TRAINING PROGRAM 2025

Contents

Chapter 1

Introduction

This project aims to study the urban expansion of five cities, namely Bengaluru, Coimbatore, Hyderabad, Indore and Mumbai, in the time span from 2019 to 2024. This shall be done via Satellite Imagery using the NDBI metric. We used Google Earth Engine for this project.

1.1 Objective

To analyze and quantify spatiotemporal patterns of urban expansion over a defined period by employing remote sensing-based spectral indices such as the Normalized Difference Built-up Index (NDBI).

1.2 Remote Sensing

We obtain the data for our study via Satellite Imagery, which is a subcategory of Remote Sensing.

Remove Sensing refers to the non-physical acquisition of information of an object or a phenomenon. Some examples include SONAR, Satellite Imagery and RADAR. It forms a fundamental aspect of our daily life aiding in Weather Forecasting, GPS navigation, Soil Health Assessment, etc.

In Satellite Imaging, we record the transmitted and reflected light from the object on the sensor. Different sensors are built for different resolutions and EM bands. Due to the inherent property of the object, it may selectively choose to reflect certain frequencies of light (this property being known as its *spectral signature*) which can be useful in identifying it.

1.3 Remote Sensing Indices

Remote Sensing indices are formulas that combine reflectance data from two or more spectral bands to highlight specific features or conditions on the Earth's surface. The specific index we shall look into is the **Normalised Difference Built-up indices (NDBI)** which uses the NIR and SWIR bands to emphasize manufactured built-up areas

Formula -

$$NDBI = \frac{(SWIR - NIR)}{(SWIR + NIR)} \tag{1.1}$$

NDBI ranges from -1 to 1; higher the index value, more the urbanisation. Usually we associate negative NDBI with water bodies and positive value with human settlement.

1.4 Geographic Information System (GIS)

The data collected from Remote Sensing is usually processed in a Geographic Information System (GIS).

Geographic Information System is the combination of Computer Hardware, Software, Personnel, Data and its processing methods which is used to manipulate, analyse and present information tied to a spatial location. It provides the location, condition, trends & patterns and modelling associated with a spatial location. Some examples of GIS Software include QGIS and Google Earth Engine.

1.5 Urban Expansion

An Urban area is a densely populated built environment (i.e. human made settlements), such as a city or a town, with a high concentration of infrastructure as well as people. These areas evolve from rural areas with a gradual increase in influx of people seeking to profit from its commercial and employment advantage.

This process of Urban Expansion is also known as *Urbanization*, with the cities growing bigger over time. The rate of growth varies between city to city, as depended on different factors such as changes in industry, connectivity, living space, government policies, etc. Although this is a natural growth of an economy, this expansion could prove to be hazardous to both the citizens and the environment if it were unplanned and unchecked. This is one aspect that we shall be observing in this project.

1.6 Organization of The Report

This chapter provided you with the basic introduction to the topics discussed in this report as well as the motivation behind them. In the upcoming chapters, we shall see the detailed explanation of the methodology, data and results.

In Chapter 2, we take a look into our Study Areas. In Chapter 3, we discuss the data collected from satellite used for the analysis - Sentinel 2A. In Chapter 4, we delve into the methodology used. In Chapter 5 & 6, we finally look into the Results and Conclusion.

We have also created a Github Repository for the code and data.

Chapter 2

Study Areas

We shall look at 5 different cities in our project - Bengaluru, Coimbatore, Hyderabad, Indore and Mumbai. These cities were chosen as a varying scale of urbanisation.

2.1 Bengaluru

Bengaluru (2.1(a)) is the state capital of Karnataka and the leading IT hub of India. Known as the *Silicon Valley of India*, it is home to countless startups, R&D centers, and multinational tech corporations. Beyond its tech-centric identity, Bengaluru is also known for its beautiful parks and green spaces, earning its name as the *Garden City*.

2.2 Coimbatore

Often called the *Manchester of South India*, Coimbatore (2.2(a)) is a major industrial city in Tamil Nadu. Its economy is deeply rooted in the textile industry, but it's also a significant hub for manufacturing, particularly auto parts and engineering goods. Nestled at the foothills of the Western Ghats, the city serves as a gateway to several popular hill stations, including Ooty.

2.3 Hyderabad

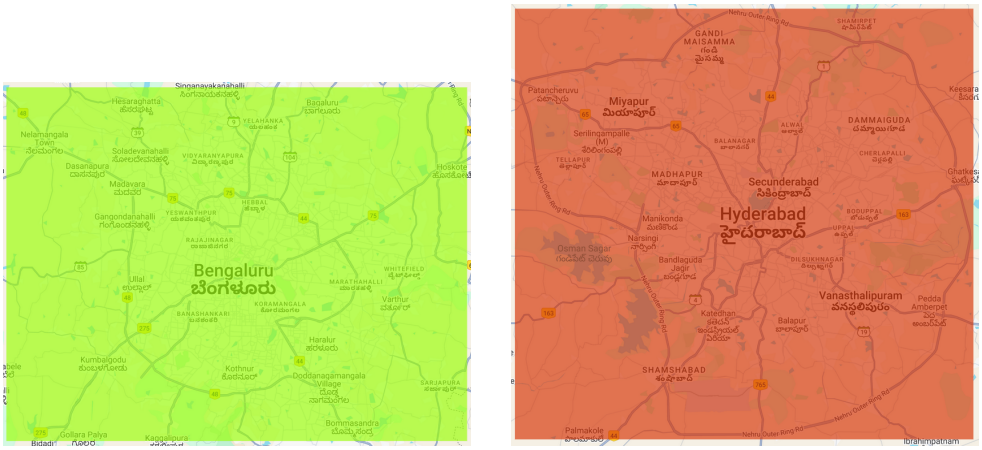
Hyderabad (2.1(b)), the *City of Pearls*, is a bustling metropolis that blends history with modernity. It has emerged as a major IT and pharmaceutical hub, hosting the offices of many global tech giants. The city's rich history is evident in landmarks like the Charminar and Golconda Fort.

2.4 Indore

Recognized as India's cleanest city for several consecutive years, Indore (2.2(b)) is the commercial capital of Madhya Pradesh. It's a key center for trade, education, and finance in Central India.

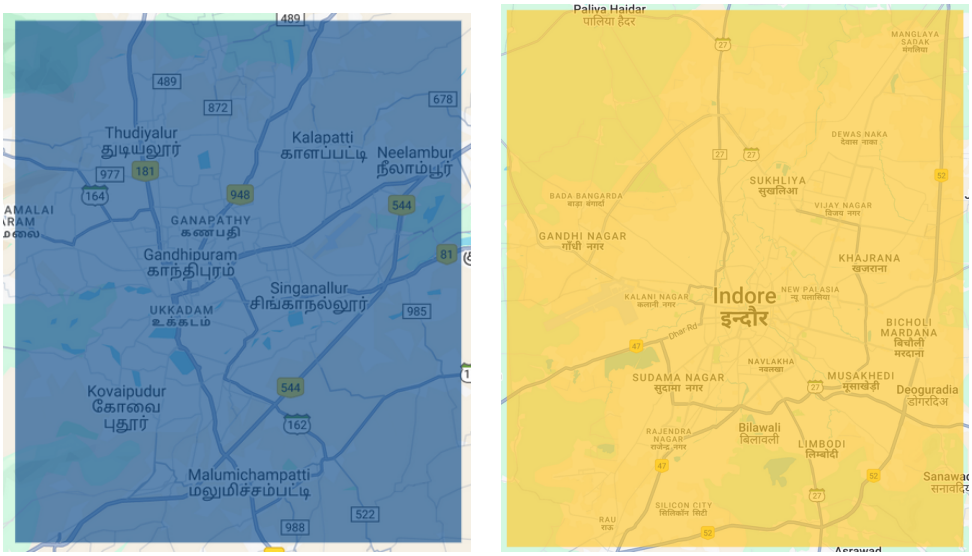
2.5 Mumbai

Mumbai (2.3) is the financial, and commercial capital of India and the state capital of Maharashtra. Located on the Salsette island off the Konkan coast, It is also the most populous city of the nation. As the heart of the nation's economy, it hosts the headquarters of major corporations and financial institutions like the Bombay Stock Exchange.



(a) Bangalore (b) Hyderabad

Fig. 2.1: Bangalore & Hyderabad AOI



(a) Coimbatore (b) Indore

Fig. 2.2: Coimbatore & Indore AOI

Fig. 2.3: Mumbai AOI

Chapter 3

Data Used

We shall be utilising the **Sentinel 2A** satellite data from 1 Jan 2019 to 31 Dec 2024. The images obtained from Google Earth Data Catalog are corrected for maximizing the *Surface Reflectance*. We apply further corrections as follows:

3.0.1 Cloud Masking

We first extract the Quality Assessment band which contains the pixel-level quality flags stored as bit information. We then create bit masks for cloud cover and Cirrus clouds and check for the condition of cloud masking. Else, we apply the mask and convert the digital numbers to reflectance values scaled by 1e-4.

3.1 Sentinel 2A

It is a sun-synchronous satellite operated by the **European Space Agency (ESA)** as part of the Copernicus Programme. It has high temporal (10 days) and spatial resolution (10 m) with a swath width of 290 km. It has a multi-spectral sensor the following bands given in 3.1. For this project, we utilise the bands B8 and B11 for the study.

Band	Resolution	Central Wavelength	Description
B1	60 m	443 nm	Ultra Blue (Coastal and Aerosol)
B2	10 m	490 nm	Blue
B3	10 m	560 nm	Green
B4	10 m	665 nm	Red
B5	20 m	705 nm	Visible and Near Infrared (VNIR)
B6	20 m	740 nm	Visible and Near Infrared (VNIR)
B7	20 m	783 nm	Visible and Near Infrared (VNIR)
B8	10 m	842 nm	Visible and Near Infrared (VNIR)
B8a	20 m	865 nm	Visible and Near Infrared (VNIR)
B9	60 m	940 nm	Short Wave Infrared (SWIR)
B10	60 m	1375 nm	Short Wave Infrared (SWIR)
B11	20 m	1610 nm	Short Wave Infrared (SWIR)
B12	20 m	2190 nm	Short Wave Infrared (SWIR)

Fig. 3.1: Sentinel 2A bands

Chapter 4

Methodology

We shall now pour over the complete methodology of the project. There are three stages - Dataset Creation, Classification and Analysis

4.0.1 Dataset Creation

We must first decide on the study area(s) for our project. We decide on the five cities based on their different urbanisation levels. We must then decide on the time span to be used - 2019 to 2024 in our case.

Next, we look into the required satellite for the data (we chose Sentinel 2A). Within this satellite, we choose the appropriate bands required - the True RGB bands, the NIR and the SWIR band required for the NDBI calculation.

4.0.2 Classification

For each study area, we retrieve the images per year (after being properly preprocessed). So, in total we shall have 25 raw images that display the regions in the NDBI band.

We classify the complete image based on the per pixel NDBI value:

$$class = \begin{cases} 1, & \text{if } NDBI \leq 0 \\ 2, & \text{otherwise} \end{cases}$$

We obtain the total pixels classified as urban area. We calculate their total area and note the total urban area in the image (in units of sqkm).

4.0.3 Analysis

We export and combine the areas and the classified images for each region and perform analysis on them.

4.0.4 Overall Algorithm

Algorithm 1 Urban Expansion Analysis

yearList $\leftarrow$ $range(2019, 2024)$
regions $\leftarrow$ [Bengaluru, Coimbatore, Hyderabad, Indore, Mumbai]
areas $\leftarrow$ []
for $(i \leftarrow 0; i <$ regions.$length; i++)$: **do**
    region $\leftarrow$ regions$[i]$
    for $(j \leftarrow 0; j <$ yearList.$length; j++$ **do**
        year $\leftarrow$ yearList$[j]$
        start $\leftarrow$ (1/1/year)
        end $\leftarrow$ (31/12/year)
        image $\leftarrow$ dataset.filterDate(start, end).filterBounds(region).maskClouds()
        ndbi $\leftarrow$ image.normalizedDifference([SWIR, NIR])
        class $\leftarrow \begin{cases} 1, & \text{if NDBI} \leq 0 \\ 2, & \text{otherwise} \end{cases}$
        urbanClass $\leftarrow$ class $== 0$
        totalArea $\leftarrow$ $(\sum_{pixel}$ (urbanClass$[pixel] *$ pixelArea$)) * 10^{-6}$
        areas.append(totalArea)
    end for
end for

Chapter 5

Results

5.1 Total Area Plot

In 5.4, We see that there are considerable size differences between the cities. Hyderabad and Bangalore lead the group with surface area of 2230 and 1958 km^2 respectively. Then follows Mumbai with a considerable drop to 1313 km^2. Next is Coimbatore being half its size, and finally we have Indore being nearly half the size of Coimbatore. It is interesting to notice the small size of Mumbai as compared to its fellow Metropolitan cities (i.e Bangalore and Hyderabad).

5.2 Urban Area Plot

In 5.3(b), we notice the same pattern repeating again - Hyderabad and Bangalore lead the group with a large margin, whereas the other cities cluster together with an area less than half the formers. An interesting trend to look with respect to the years is that all cities have either maintained or decreased their urban area, with the sharpest dip observed in Hyderabad.

In 5.5(b), Most of the cities have a stable state of urban area with low variability, with an exception of Bangalore. However, every city's plot is left skewed.

5.3 Urban Percentage Plot

Here is the plot of the percentage of total area which is urban per city across the time period of 2019-2024.

In 5.3(a), we make a striking observation: Coimbatore competes closely with both Hyderabad and Bangalore in terms of Urban percentage, with it exceeding both in the year 2023. Indore on the other hand, fluctuates between 59.8% to 36.9% across the years. And at the last is Mumbai with a stead mean of about 30%.

In 5.5(a), Bangalore shows the greatest scatter among the group with an even dispersion within the inter-quartile range. Indore shows the largest dispersion within the interquartile range. Mumbai shows the least variability in data, with a single outlier at 34.8%. All of the cities' plots are left skewed.

5.4 Inferences

5.4.1 Geographical Location dictates the City Size

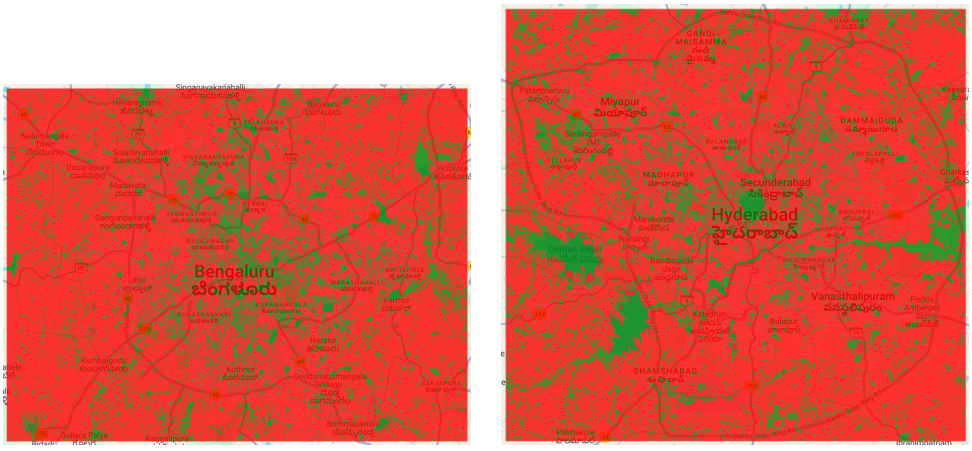
Mumbai, despite being the most populous city in India, falls short in Total Area compared to other Metropolitan cities such as Bengaluru and Hyderabad. This is because Mumbai is situated on an island which offers limited space while other cities are on mainland with ample space to grow. However, Mumbai also has a lowest Urban Area percentage across other cities. This may once again be attributed to the presence of the large Sanjay Gandhi National Park as well as vast swaths of mangroves around the shorelines of the city.

5.4.2 Decreasing trend of City Urban Area

All of the cities showed close to none net gain in urban area across the five year period, except for a few variance in the data. This suggests that the green cover within the cities have been maintained if not increased over the period. An additional factor would be the rise in the preference of vertical housing (apartments and skyscrapers) over traditional housing, which takes up less surface land for the higher number of residents.

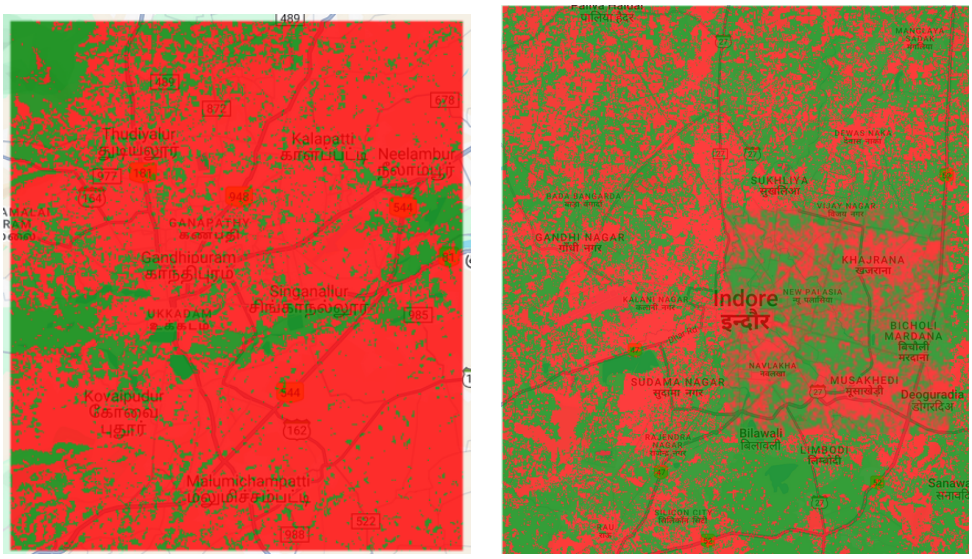
5.4.3 Importance of Trade on Urban Size

What does Mumbai and Coimbatore have in common? Both form an important geographical center for trading. Mumbai is a vital port city for India, dealing with million of tonnes of container cargo each year and being well connected to both other seaports and airports. Coimbatore falls right beside the Palakkad gap in the Western Ghats, which has been a historically important entrance into the Malabar. Thus we see a significantly higher urban presence in Coimbatore compared to its size.



(a) Bangalore (b) Hyderabad

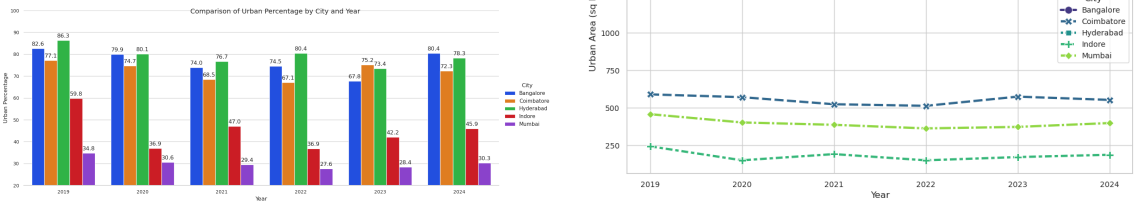
Fig. 5.1: Bangalore & Hyderabad Urban Classification



(a) Coimbatore (b) Indore

(c) Mumbai

Fig. 5.2: Coimbatore, Indore & Mumbai Urban Classification



(a) Urban Percentage Plot (b) Urban Expansion Plot

Fig. 5.3: Plots on Urban Expansion

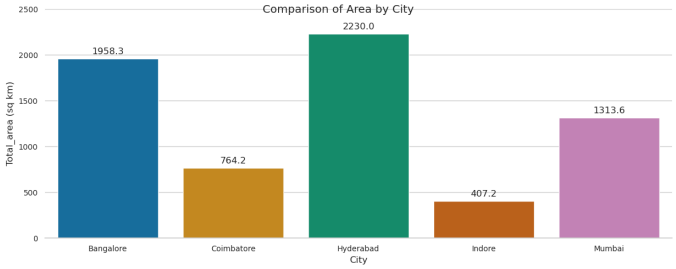
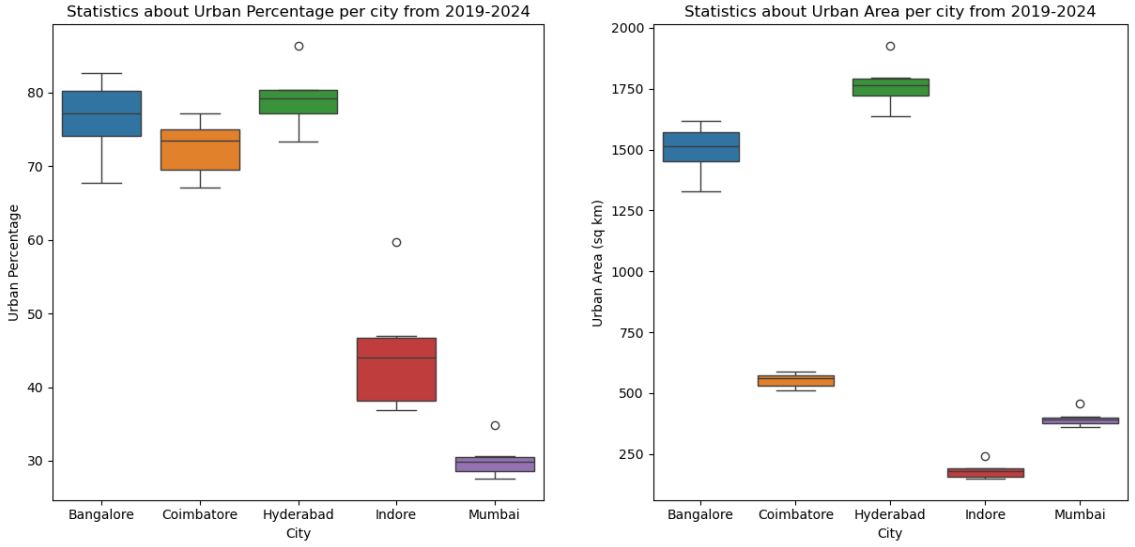


Fig. 5.4: Total Area



(a) Urban Percentage Plot (b) Urban Area Plot

Fig. 5.5: Box Plots

Chapter 6

Conclusion

In this project, we looked at the urban expansion of five cities - Bengaluru, Coimbatore, Hyderabad, Indore and Mumbai using Satellite Imaging. The specific method we chose to do so was by creating a specific spectral index known as NDBI (Normalised Difference in Built-up Index) which utilises Infrared light detected by the satellite Sentinel 2A for marking the region as urban or otherwise. We then calculated the total urban area per sq km for each city and made some interesting observations on the data, such as difference of urban city area in Tier 1 and Tier 2 cities and the anomaly of Mumbai's low area size.

This study provides us with some key insights about the respective cities. It can later on be supplemented with a more comprehensive study of them with greater features such as Census study, Water table study, etc.

Some additional improvements that can be made are:

- **Increased timespan**: Five year period is a short duration for a city's growth. This time span was chosen based on the satellite data quality. Although there exists the Landsat satellites with data of much larger duration, the image quality was a compromise. This study can be replicated with a time span of 20 to 30 years.
- **Increase the Study Areas' scope**: We can include more cities such as Delhi-NCR, Patna, Jaipur, Surat, etc. for analysing a further diversity of data.

Chapter 7

References

1. NDBI
2. Wikipedia for Study Area Information
3. Datasets from Google Earth
4. Sentinel 2A information