VIIRS TO DEFINE EXTENT OF INDIA'S URBANISATION

Parth Khare[1] and Serene Vaid[2]
[1]Consultant to the CEA, Ministry of Finance, Government of India, New Delhi, India
Email: parthakhare@gmail.com
[2]Research Assistant, Indian Council for Research on International Economic Relations, New Delhi, India
Email: serenevaid@gmail.com

KEY WORDS: Rate of urbanisation; VIIRS night-time lights; GHSL Built-up; Supervised clustering

ABSTRACT: India is rapidly urbanizing, but does the census based definition of 31.2% realistically capture the extent and scale of India's urban story? Based on density of human activity over built-up area, this paper attempts to redefine the rate of national urbanisation as well as provide a more granular assessment of state and district level urban sprawl. (Gridded Population: GHSL Human activity is defined as a weighted combination of nightlights (VIIRS) and micro population). Built-up data is extracted from GHSL data. The high volume and largely non-parametric distribution of satellite imagery allows for a robust application of machine learning techniques. The paper explores unsupervised clusters for exploring and defining urban sprawls. Using an ensemble of various supervised clustering methods, the study aims to employ supervised algorithm in strengthening human activity attribute, by training population based attributes through VIIRS. Unlike DMSP, VIIRS does not suffer with top coding and showcases reasonably well local spatial variance. VIIRS is also relatively free of seasonal biases as the release at a monthly frequency helps in accounting for localized spurts in light intensity caused due to event based external shocks like festivals or calamities. Accurately capturing Urban and Rural regions is critical, particularly from an economic policy perspective. In case of India, a sizeable proportion of schemes and grants are disbursed based on an urban-rural bifurcation by center, states and local administrative bodies.

INTRODUCTION

There is a discrepancy in the definition of urban settlements between the various government sources in India. According to Census, towns are categorized by areas with population of at least 5000, density of at least 400 persons per square kilometer and that at least 75% of the male main working population is engaged in non-agricultural activities. Whatever is not covered under this definition is termed as rural. The rate of urbanisation was 31.16% according to Census 2011 with population of 377.1 million, growing only marginally at 2.76% per annum during 2001 and 2011. On the contrary, going by the administrative definition, the population living in areas governed by urban local bodies such as municipal corporations, municipal councils etc is classified as urban. According to this definition, the rate of urbanisation was measured at far less rate of approximately 26% in the year 2011.

The census definition of categorization into urban-rural frame in India has remained unchanged since its inception in 1961.[1] With the dynamic shift in the pattern of spread of the urban sprawl, the barometer for the classification into urban and rural should also be updated to accommodate the varied rapid pace of development. Moreover, the given traditional measures are inadequate to capture the complex phenomenon of urbanisation in light of dynamically updated, real time, micro satellite data. These measures become even more intense when we study urbanisation at the state or local level. Satellite imagery can be exploited in this context to identify the urban spheres in India.

This paper attempts to postulate a dynamically updated satellite imagery driven definition of urbanisation. By redefining and extending Global Human Settlement Layer (GHSL) provided by the Joint Research Center of the European Commission Science Hub, a new contemporary definition of urbanisation can be constructed for India. According to GHSL's benchmarks, settlements are categorized as,[2] a) High-density cluster/urban centre-contiguous grid cells of 1 km^2 with a density of at least 1500 inhabitants per km^2 and a minimum population of 50,000; b) Urban cluster: cluster of contiguous grid cells of 1 km^2 with a density of at least 300 inhabitants per km^2 and a minimum population of 5000. As quantified in this year's Economic Survey, based on the definition of High Density Settlements, India was 63% urbanized in 2015 which is evidently more than twice the figure of that defined in census 2011.

[1] http://www.censusindia.gov.in/2011-prov-results/paper2-vol2/data_files/kerala/Chapter_I.pdf
[2] http://ec.europa.eu/eurostat/statistics-explained/index.php/Degree_of_urbanisation_classification_-_2011_revision

However, population density of 1500 per km^2 might not hold true uniformly for all countries, especially India where the variance in spatial distribution of population is very high. According to World Bank's estimates, population density of European Union on which the GHSL classification is set on, was roughly 120 per km^2 in 2016 and for India it was 445 per km^2, whereas it was 57 for world on an average.[3] Therefore, if becomes vital to refine this definition by an additional filter which will be more stringent than the GHSL urban settlements. For this purpose, GHSL built-up layer can be juxtaposed with Night Light VIIRS (Visible Infrared Imaging Radiometer Suite) data, which illustrates the degree of economic activity at micro-precision. The built-up grid can be filtered by taking a weighted combination of nightlights and micro population. Setting a threshold of economic activity by VIIRS can help in creating an index of urbanisation rather then specifying one binding definition. Redefining urban areas for pan India with a panel of multiple datasets will give it a dynamic edge, which can be molded according to the requirement of the policy maker. There are several advantages of using VIIRS NTL database. For one, unlike DMSP (Defense Meteorological Satellite Program) OLS Nighttime lights, it does not suffer with top coding and showcases reasonably better local spatial variance. It has been documented in the past that the DMSP/OLS NTL data tend to exaggerate the size of urban areas compared to the Landsat analysis, because OLS-derived light features are substantially larger than the lighting sources on the ground (Imhoff et al 1997, Henderson et al 2003, Elvidge et al 2009).

DATA

In order to gauge the transition in the morphology and human footprint signature, we have primarily used raw and processed satellite imagery across the following years 1990, 2000, 2014 for processed built-up data (@1km) is collected from Global Human Settlement Layer (GHSL), to train and validate the feature classification algorithm. To assess the evolution of human induced economic sprawl (human sprawl), Gridded World Population data (@1km resolution) from SEDAC is used. This data is further complemented by nightlights data, to measure the spatial-sprawl at a micro level. Visible Infrared Imaging Radiometer Suite yearly product for the year 2015 is used for this study. Amongst the different versions available for the given year, the VCM-SL-CFG-OTR version proves out to be best suited for the study as it is devoid of disruptions from forest fires and fires associated with swidden agriculture, if any. This is cross-checked with the monthly VIIRS I-Band 375 m Active Fire data. Although the VIIRS DNB detection limits are much lower than the I-4 band that a lot of fires are detected in the DNB only and are missing from the "VIIRS 375 meter" fire product, but is most effective mask available for outlier radiance values.

METHODOLOGY

Figure 1 illustrates the process followed for the study, which can be broadly divided into two sub parts. In the first part, we prepare the data and resample VIIRS to 1km from ~750m resolution to make it comparable with the information available from GHSL (both population and built-up). After running few diagnostic tests, this study attempts to create an index with several permutations of population density to capture more realistic distribution of urbanisation in India. Further, it was also vital for this analysis to explore and segregate the type of settlements to precisely identify urban areas.

The second part of the analysis involves extracting the data at both state level and at district level. District level estimates are used to create an index of urbanisation. A threshold is determined with three components in place by visualizing the raster imagery value for a pilot of six districts. The distinction in the spatial characteristics amongst districts helps in checking for any outliers of value, if any, and create a more universal pertinent threshold. Concomitantly, more sophisticated statistical methods can be put in place for clustering the urban regions in India. After attempting to cluster urban population using unsupervised machine learning tools, we attempt to use night-time lights to train the combination of built-up and population constellations to define extent of urbanisation in India. The objective is to arrive at a rate of urbanisation based on key spatial features, population density and minimum population and more apt definition for India where the spatial variance is relatively high. The high density clusters as defined by GHSL which are essentially the urban centres and are classified by the population density and minimum population can be narrowed down for India. Amongst various supervised clustering algorithm, such as Support Vector Machine (SVM), Gradient Boosting Machine (GBM), random forest, SVM proves to be the best way forward. In the next part of the study, we attempt to implement SVM on VIIRS and GHSL Built-up data and quantify all India and state-wise level of urbanisation.

---

[3] Food and Agriculture Organization and World Bank population estimates

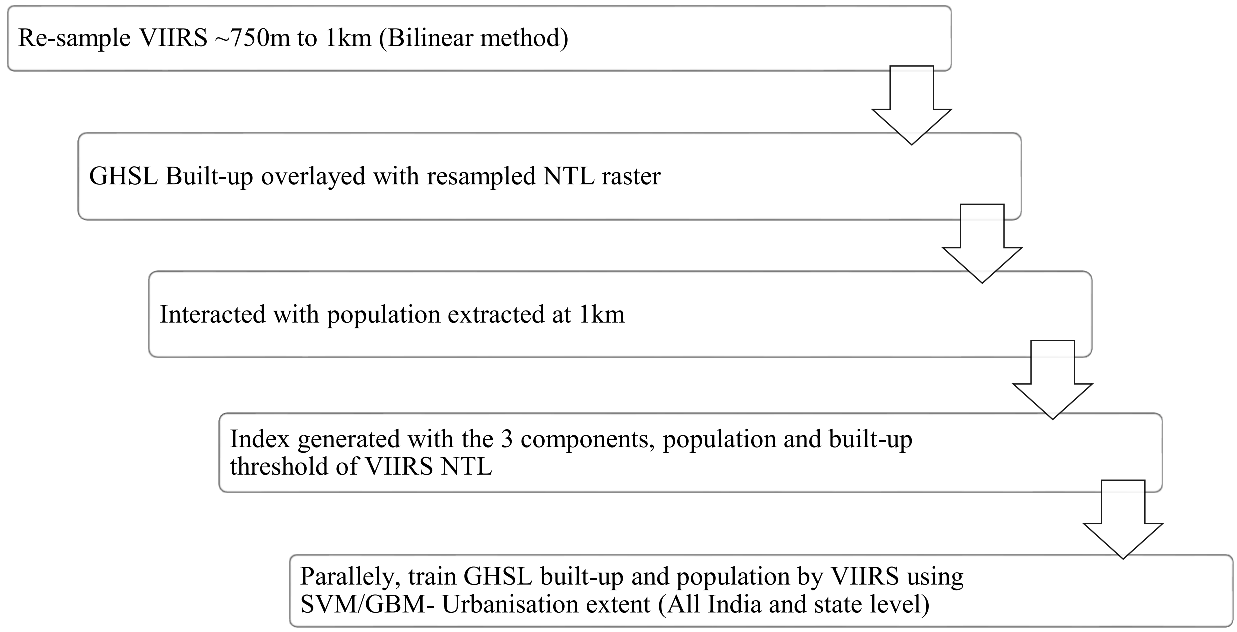


Figure 1: Flowchart exhibiting the methodology of the study

For a more granular analysis of the extent of urbanisation in India, we sift 6 districts spread across regions. The districts are selected on the basis of geographical variance across the country, division in terms of urban agglomeration with population above a defined threshold. The selected districts and their key characteristics are represented in Table 1.

Table 1: Brief description of the selected districts

District	**Census 2011 Population**	**Area (km^2)**	**Constituents of UA with population more than 1 lakh**
Aligarh (Uttar Pradesh)	36,73,889	3700	Aligarh UA (M.Corp + 3 CT)
Jaipur (Rajasthan)	66,26,178	11,152	-
Vellore (Tamil Nadu)	39,36,331	6077	Vellore UA (M.Corp + 6 CT + 5 TP + 2M)
Visakhapatnam (Andhra Pradesh)	42,90,589	11,161	-
Nashik (Maharashtra)	61,07,187	15,582	Nashik UA(M.Corp + CT + CB + M Cl)
Patiala (Punjab)	18,95,686	3430	Patiala UA (M.Corp + 7 OG)

Note: UA – Urban Agglomeration; M.Corp – Municipal Corporation; CT – Census Town; TP – Town Panchayat; M – Municipality; CB – Cantonment Board/Cantonment; M Cl – Municipal Council; OG – Out Growth;

The correlation at district level between the population density derived from GHSL and average night-time radiance value is which is now being widely accepted as a proxy of development, is 85.8% (Figure 2). At the state level, the correlation coefficient is even more at 87.9%. This indicates a robust relationship between population density and economic activity captured by nighttime lights.

Next, we create an index of all three parameters individually at state level to have an inter-state comparison. This index is generated using the maximum and minimum value of respective variable. As indicated earlier, high population density is accompanied by high average value of VIIRS. Greater level of built-up in a state is not always complemented with higher index value of built-up index. This holds true when the analysis is done at district level as well.

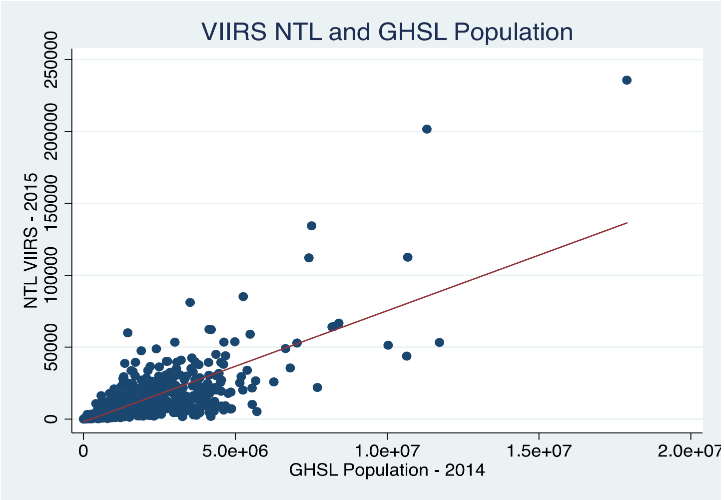


Figure 2: Relationship between GHSL Population and Nighttime lights

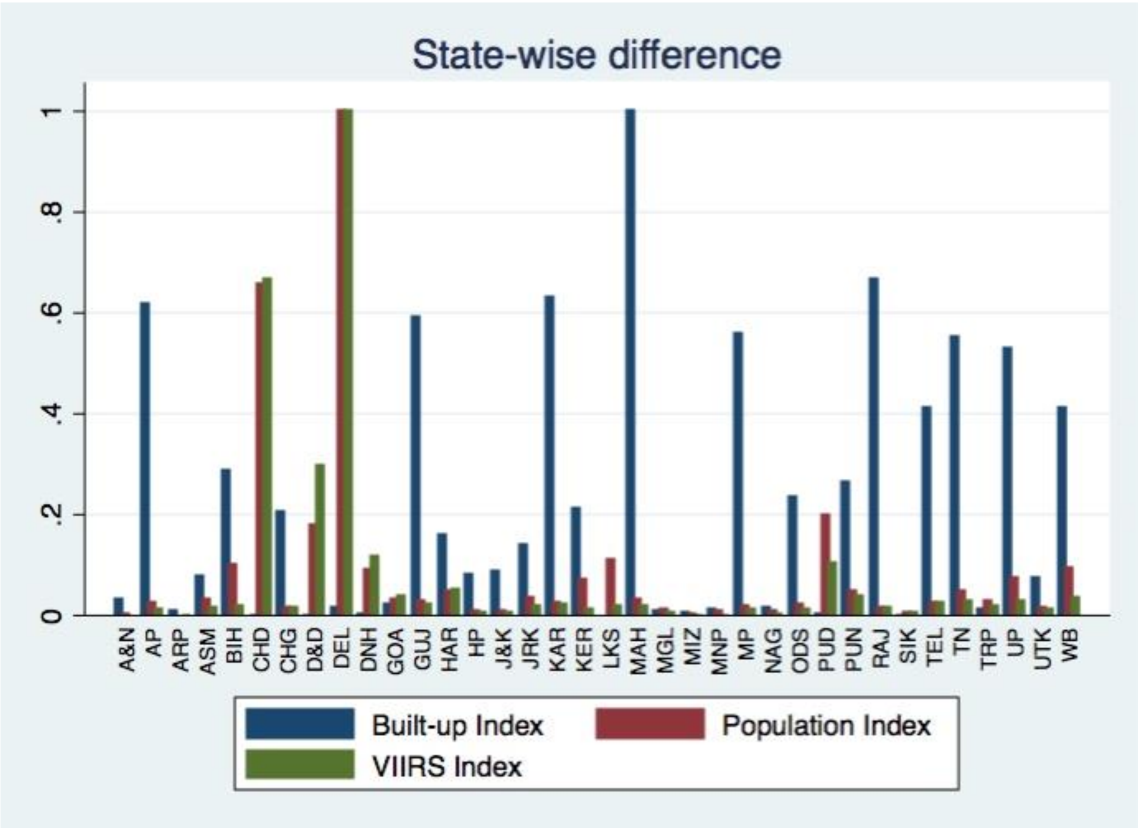
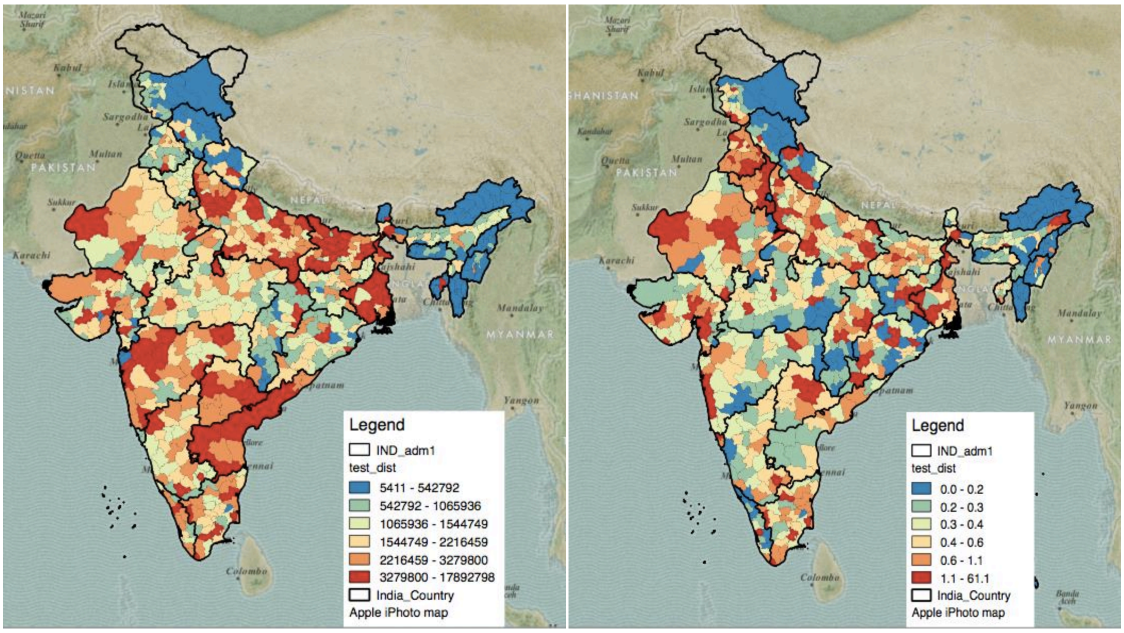


Figure 3: Parameter-wise Index for all Indian states and Union territories

NCT of Delhi has the highest index value for both population density and VIIRS nighttime light index. Another UT which has almost comparable indices is Chandigarh. Low correlation between index value of Built-up and the other two has to be further analyzed and clusters with high radiance value of nighttime lights but low or no built-up have to be sieved out or checked out for other outliers. This has to be particularly checked for states with larger area per sq km.

As indicated earlier, figure 3 displays the district-wise distribution of key variables. Figure 3(a) showcases the distribution of population within each state and 3(b) reveals the spread of nighttime lights across districts. It is worth noting that although the clusters with low population density have relatively comparable low average light intensity, such is not consistently true for the districts with higher value on the spectrum.



(a) (b)

Figure 3: District wise (a) – GHSL population; (b) Average VIIRS radiance value

For the 6 designated districts, namely, Aligarh, Jaipur, Nashik, Patiala, Vellore and Visakhapatnam, we study the spatial distribution of NTL and built-up within a district, i.e. around the central city or urban agglomeration. GHSL built-up imagery has binary value of 0 (no built-up) and 1(built-up) displayed at 1km resolution. VIIRS raw imagery (at ~750m resolution) is categorized into five continuous categories. Figure 4 exhibits three of the selected six districts for both the parameters.

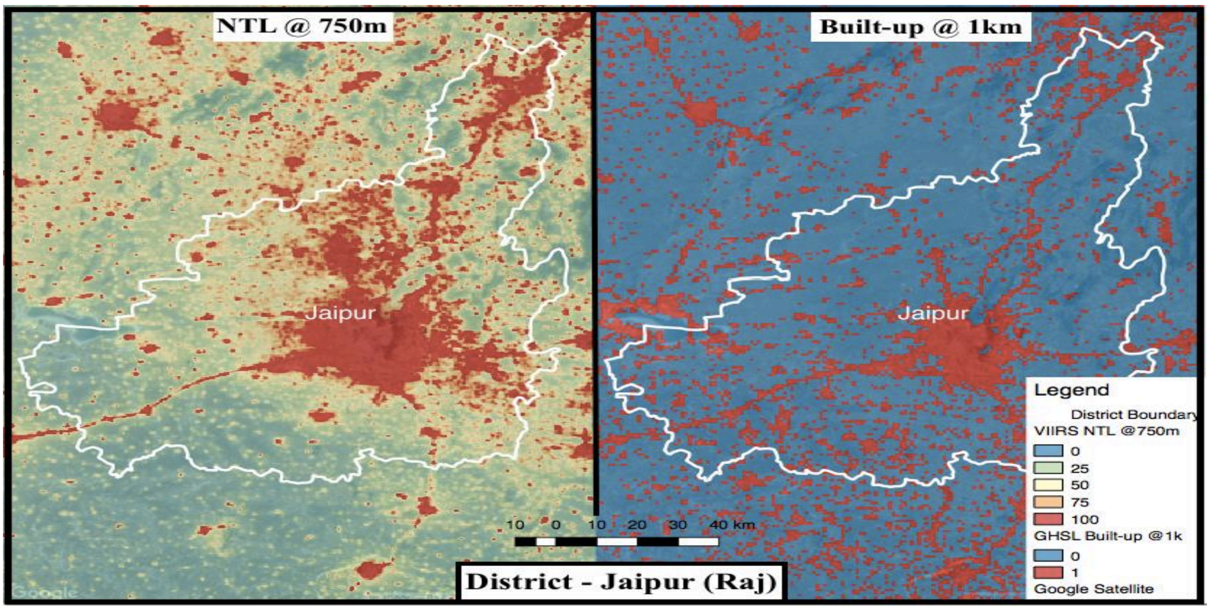


Figure 4 (a): Selected districts – Built up and nighttime radiance - Jaipur (Rajasthan)

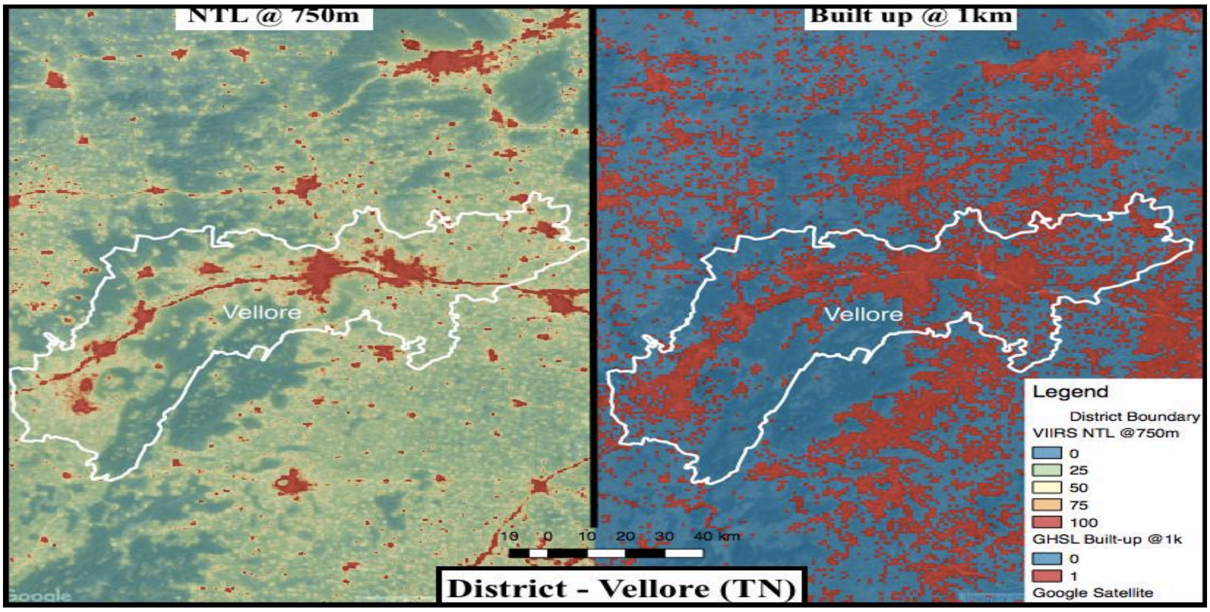


Figure 4 (b): Selected districts – Built up and nighttime radiance - Vellore (Tamil Nadu)

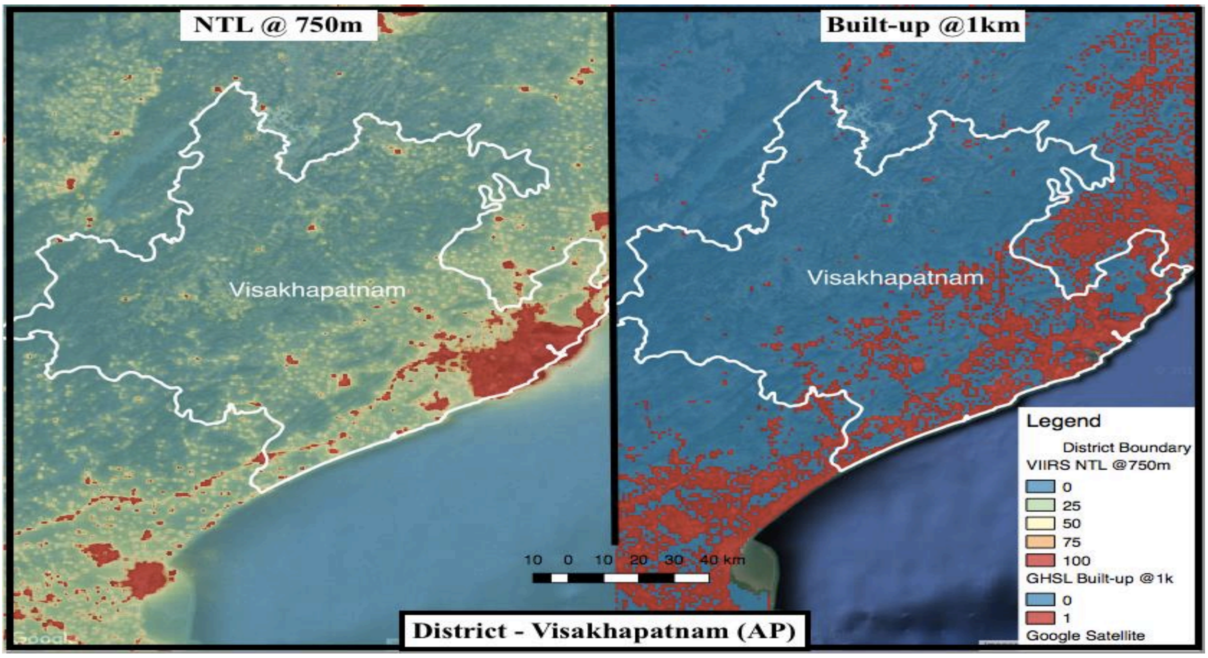


Figure 4 (b): Selected districts – Built up and nighttime radiance - Visakhapatnam (Andhra Pradesh)

Satellite imagery can provide a better definition to apprehend the magnitude of urbanisation, as it would provide an accurate measure of extent of sporadic and inexplicable urban growth in India. The revamped definition would aptly capture the urban areas in the respective states, which would lead to proportionate fund allocations to urban local bodies and municipal corporations for infrastructure development. Under-allocation of finances and resources because of misidentification of urban regions and more specifically 'Rurban' areas, which are often classified as rural, is an existing challenge. In the long run, this misclassification leads to delaying or even worse, denying growth in the nascent stages of development. In India, according to Census 2011, there are some 52 urban agglomerations where the population is more than a million. An urban agglomeration is a continuous urban spread constituting a town and its adjoining outgrowths (OGs), or two or more physically contiguous towns together with or without outgrowths of such towns.[4]

[4] http://censusindia.gov.in/2011-prov-results/paper2/data_files/India2/1.%20Data%20Highlight.pdf

CONCLUSION

The urban region in India contributes more than 60% to Gross Development Product (GDP). By 2020, share of urban population in India's GDP is projected to increase to 70-75% (Barclays, 2014). There is immense imbalance in the distribution in burden of development on cities which houses only about 30% of total population in India according to the official measurements. For more inclusive growth in the coming decades, this discrepancy has to be minimized. A major hurdle in dissolving the divide and capturing untapped potential of urban areas is identifying the extent of urbanisation in India and realising the importance of paying heed to the upcoming peri-urban areas for preparedness by policy makers for urban planning.

REFERENCES

Economic Survey 2016-17, Volume II, How Urban is India?

Elvidge, C., Baugh, K., Zhizhin, M., Hsu, F. C., Ghosh, T. (2017). VIIRS night-time lights. International Journal of Remote Sensing, 38(21).

Eric, D., Kamala, M-G. (2011). Toward a better appraisal of urbanization in India. Cybergeo: European Journal of Geography.

Goldblatt, R., You, W., Hanson, G., Khandelwal. A. (2016) Detecting the Boundaries of Urban Areas in India: A Dataset for Pixel-Based Image Classification in Google Earth Engine

Harari, M. (2016) Cities in Bad Shape: Urban geometry in India.

Mukhopadhyay, P. (2017) Does Administrative Status Matter for Small Towns in India? Subaltern Urbanisation in India: An Introduction to the Dynamics of Ordinary Towns (Chapter – 17)

Schneider et al (2015) A new urban landscape in East–Southeast Asia, 2000–2010. IOPscience.

Zhou, Y., Smith, S.J., et al (2015) A Global Map of Urban Extent from Nightlights. IOPscience.