Policy Brief

89448 | July 2019

Kathryn Baragwanath Vogel, Ran Goldblatt, Gordon Hanson, and Amit Khandelwal

Mapping urbanisation in India with satellite imagery

In brief:

- Satellite images provide a promising source of data for measuring urbanisation: It provides higher spatial and temporal resolution than administrative data.
- This project uses remotely sensed data to detect urban markets in India for 2013 (accessible via the Google Earth Engine).
- We find that landcover-based markets detected from daytime satellite imagery are realistically jagged in shape and reveal substantially more variation in the density of economic activity when compared to nightlight-based markets.
- Nightlight-based markets capture 23.4% of India's total population and 75.3% of its urban population. Landcover-based markets capture 29% of India's total population and 93.2% of its urban population.
- We detect strong correlations between market size and proximity to public infrastructure.
- High resolution day-time satellite data can be used to accurately measure urban footprints, while coarser nighttime satellite data can be used to measure the intensity of economic activity within those footprints.
- Satellite imagery can be used for the study of urban sprawl, urbanisation dynamics, the effects of land use policies, and the effects of infrastructure on urbanisation and economic activity.

This project was funded by IGC India

Measuring urbanisation: Challenges

Urbanisation is a central feature of economic development. As countries grow, cities expand, by absorbing labour and capital released from rural towns and villages. How to deal with the complex outcomes of urbanisation is a significant challenge for policymakers in the developing world.

A common approach to the measurement of urbanisation is to use officially designated administrative units. These may be as large as a metropolitan area (Duranton and Turner, 2012), as small as a town or village (Eeckhout, 2004), or an intermediately sized unit such as a county or a district (Hanson, 2005; Ghani et al., 2014; Donaldson and Hornbeck, 2016). Because administrative boundaries are defined according to pre-existing legal jurisdictions, they may fail to reflect the reality of how land is actually organised.

Data constraints further complicate analysing the urbanisation process. Censuses and surveys provide information at the level of officially designated administrative boundaries that often bear little resemblance to the actual size and reach of cities. Additionally, administrative data are typically captured at low temporal frequencies that often make the data inadequate for real-time policy analysis (e.g. every ten years). Comparing spatial outcomes across regions and countries is also complicated by differences in sampling strategies.

Satellite imagery: A potential solution

Remotely sensed data, including satellite imagery, offer a promising new way to track urbanisation (Donaldson and Storeygard, 2016). Economists have used satellite data on the intensity of light emitted at night to study economic growth and the distribution of economic activity at a 1km resolution (Henderson et al., 2012; Henderson et al., 2018).

Daytime satellite imagery is available at even higher spatial and temporal resolutions, down to 30-metre resolution at a biweekly frequency dating back to the late 1990s. Economists now have access to data at previously unimaginable spatial and temporal resolutions. The challenge is how to translate this high-dimensional spatial data into a form that generates insights for research and policy.

The study: Detecting urban markets in India with satellite imagery

In our study (Vogel et al., 2019), we develop a method to process terabytes of satellite data into a unit of analysis that economists have long studied: Urban markets. The method relies solely on public data (accessible via the Google Earth Engine) and can be scaled across the globe. Our approach constructs markets as a set of contiguous, or near contiguous, pixels of economic activity based on nighttime and/or daytime satellite imagery.

Satellite images of nighttime lights are one source of data that indicates the presence of economic activity. Operationalising this definition using nightlight data requires choosing a minimum threshold of light intensity for the contiguous pixels. Following Rozenfeld et al. (2011), we experiment with buffers that combine contiguous sets of pixels if they lie within a radius of 1km, 2km, 4km, or 8km.

Having to choose a nightlight threshold to define a nightlight-based market immediately reveals a trade-off: A strict threshold only captures major urban agglomerations and lowering the threshold to include smaller cities comes at the expense of exploding the size of larger cities that have satellite towns. This trade-off is in part a consequence of the blooming effect of light, which tends to produce cities whose boundaries are too expansive and too smooth relative to the haphazard shape of actual cities.

We therefore contrast the spatial extent of nightlight-based markets with those formed from high-resolution daytime satellite imagery. These data are available at a much finer resolution than nighttime light data but require image classification techniques to detect the spectral signature of a pixel.

We examine data on built-up land cover from two publicly available layers developed using Landsat data: The MODIS layer (Channan et al., 2014) and the Global Human Settlements Layer (GHSL) developed by Pesaresi et al. (2015). We also examine a layer of built-up landcover for India developed by Goldblatt et al. (2018).

The findings: Comparing landcover- and nightlight-based markets

We find that landcover-based markets, irrespective of the data source, are realistically jagged in shape and reveal substantially more variation in the density of economic activity when compared to nightlight-based markets. Collectively, our MODIS 1km markets capture 29.0% and 93.2% of India's total and urban populations, respectively. Market size correlates strongly with population, and the variance in population for smaller sized landcover markets reflects the fact that these markets include both dense areas within major metropolitan areas and less populated peripheral towns.

Figure 1 illustrates the results for the New Delhi metropolitan area. It compares markets detected from daytime imagery (top row) against various nightlight thresholds (rows 2 - 5). The buffers increase across columns. Consider the 5th row, which detects all pixels above a nightlight value of ten. While this threshold captures satellite cities far from the city's centre (e.g. Meerut), they are swallowed up, and the New Delhi market explodes in size. As the threshold increases (move up the rows), New Delhi takes its more familiar shape, but those definitions fail to detect the satellite towns.

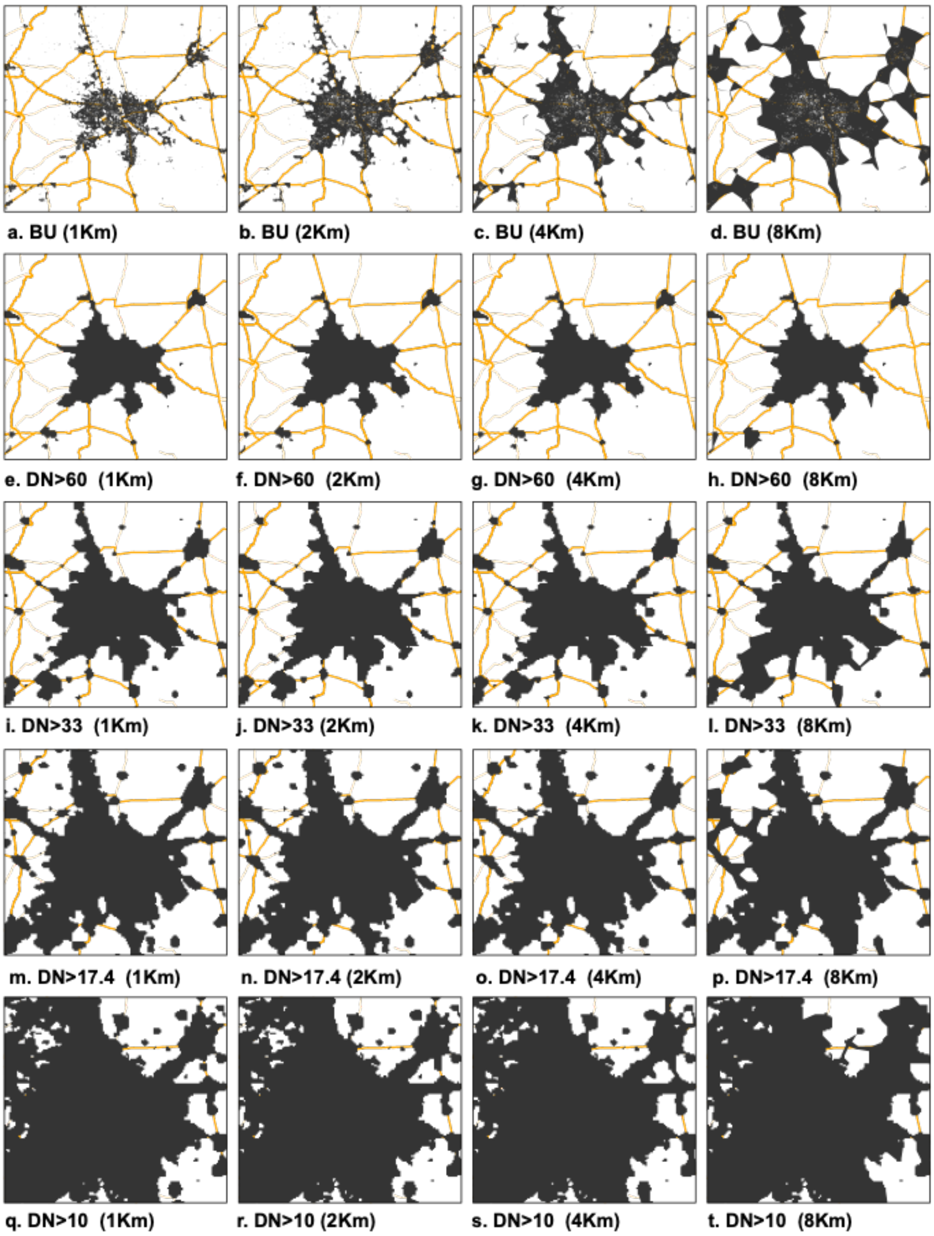
The top row illustrates the markets detected from daytime imagery and reveals several differences with respect to nightlight-based markets:

1. The structure of markets is revealed to be jagged and uneven, unlike the smooth amoeba-like shape of nightlight markets.
2. We detect distinct neighbourhoods close to the city centre, which shows the ability of daytime imagery to capture the evolution of the polycentric city.
3. The imagery delineates many distinct satellite towns that have formed around New Delhi.

Across India, we can detect markets with an average nightlight intensity of just five, which suggests we can capture rural areas of India that do not have reliable access to electricity. While these areas could be observed by lowering the light-intensity threshold, this would come at the cost of vastly increasing the area of above-threshold contiguous pixels around India's large cities. Landcover-based markets are therefore able to capture small cities and towns in India, while preserving the spatial distribution of activity of the largest cities.

We also demonstrate the potential to combine nighttime and daytime imagery to measure the extensive and intensive margins of economic activity across markets. This exercise reveals that daytime imagery is particularly well-suited for defining the extent of market areas, and that nightlight imagery is useful for capturing the intensity of activity within these market boundaries. In other words, *daytime imagery can be used to measure the boundary of a market, while the area's nightlight intensity can measure the economic activity that the market generates*.

Figure 1: Markets around New Delhi



Notes: Row 1 displays landcover-based markets using the Goldblatt et al. 2018 layer. Rows 2 - 5 display nightlight-based markets. The buffer changes across columns.

To explore the potential to use these data to analyse spatial policies, we match our markets to several maps of India's public infrastructure. We detect *strong correlations between market size and proximity to roads, railway stations, and mobile phone towers*. For example, landcover-based markets exhibit a sharp negative elasticity in size and nightlight intensity with respect to nearest-road distance: A landcover-based market, whose nearest paved road is just 2km away, is roughly half as large as one with a road that bisects the market centroid.

Policy implications

- **A new tool to study urbanisation:** By using publicly available data, our methods for identifying urban markets are scalable, replicable, and can be used for analysing urbanisation in any country and timeframe for which satellite data is available.

- **Potential applications:** The method can be applied to the study of urban sprawl, urbanisation dynamics, and effects of land use and urbanisation policies. Although much more work is needed, our research suggests that it may be possible to evaluate the impacts of investments in infrastructure using satellite imagery. The complementarity of day-time and nighttime data allows us to measure the effects of such infrastructure investments by looking at both the intensive and extensive margins.