Use of night lights data as proxy for economic growth: A multi-year Light based growth indicator (LBGI) for China, India and the U.S.

Rajendra Kulkarni, Kingsley Haynes, Roger Stough and James Riggle

School of Public Policy, George Mason University

Abstract

In this paper we explore how satellite images of global night lights from the years 2001 to 2007 can be used to estimate economic activity at the sub-regional level in the U.S., India and China. The night lights based estimates of economic activity are then spatially analyzed and compared with sub-regional economic indicators where available for selected years. We also briefly discuss two theoretical models that shed light on the geospatial patterns reflected by the night light data. The results are presented with a series of maps, charts and tables.

Key Words: Economic growth, Satellite data, night lights, LBGI (Light based growth indicator), India GDP, China GDP, USA

JEL: R11, R12, C02

Introduction

In much of the developed world and in some parts of the developing world, GDP is recognized as the single most important economic indicator from among many that are periodically released. Its primary role in setting economic policies by central bankers and government officials is well known. Even after well documented problems in its ubiquitous use the world over as a comparative economic growth indicator, very few alternatives exist (Deaton, Angus and Alan Heston. 2008). Not only does it indirectly reflect trillions of economic transactions carried out by hundreds of millions of ordinary citizens, it may also have intangible effects on their perception of personal economic well-being. For all of this disproportionate level of influence on society, the GDP indicator is an amalgamation of many economic factors aggregated at the national and state levels, no other statistic exists for measuring growth patterns at the local or regional level. However, the use of a single national/state level growth indicator such as GDP
may not reflect variations in economic activities that are likely to exist at the local and sub-regional level. The lack of a universal growth indicator for small areas suggests a need to create an alternative measure that could be used as a comparative measure of economic activity in time and space. Over the last several years, a growing number of researchers have been exploring alternatives to GDP (Johnson, Simon H., William D. Larson, Chris Papageorgiou, and Arvind Subramanian. 2009). In the past satellite images have been employed to classify land usage for agriculture, shrinkage of forested areas, growth of deserts, growth of urban areas etc. Much of the satellite image data show changes across the globe that are combinations of naturally occurring processes as well as those due to human activity. Therefore using remote sensing to measure changes occurring across the globe that are entirely due to human activity is not easy. For that purpose, the night light images are useful, since the night lights are a direct result of human activity. Once such data has been cleaned and scrubbed to filter out lights due to forest fires, moon shine etc., they may prove useful. The DMSP’s (Defense Meteorological Satellite Program; http://msl.jpl.nasa.gov/Programs/dmsp.html) release of clean and scrubbed night lights images spanning 1992 to 2003 has been employed recently by number of researchers to measure human activity (Ebener, Steeve, Christopher Murray, Ajay Tandon and Christopher D. Elvidge. 2005 ; Doll, Christopher N.H., Jan-Peter Muller and Jeremy G. Morley. 2006). Especially of interest are the efforts (Henderson, Storeygaurd, Weil, 2009 and Chen, Nordhaus, 2010) in developing measures of economic activity at the national and regional level.

In this paper we explore how night lights data may be used to measure economic activity at the sub-regional/local level, for example, 3108 counties in the U.S., 570 plus districts in India and 200 plus prefectures in China. Note that subregional unit of analysis varies in size for each of the countries, in general a typical U.S. county is bigger in size than a typical Indian districts and is smaller than a typical Chinese prefectures. The following section describes briefly the nature of night lights data, and the following section discusses how we used GIS to extract image data into a measure called “Light based growth (LBG)” at the county/district/prefecture level that could then be used as a proxy for economic activity. Next the subregional LBG measure for each of the country is spatially compared to the corresponding subregional unit of analysis of each country: county level income for the U.S., district level GDP for India and prefecture level GDP in case of China for selected years. In our previous paper we used Luc Anselin’s LISA measure (1995) as well as local and global Moran’s Indices, to compare co-clustering characteristics of subregional
night lights for years 1993, 1998 and 2003, in U.S. and China with income for the U.S counties and GDP for Chinese prefectures for the same three years. The bi-variate analyses helped to establish spatial relation between each jurisdiction’s night lights and economic growth of its neighbors, but the explanation was weak in explaining how night lights could be used as proxy for economic growth. In this paper we present a more robust explanation of relation between night lights and economic growth. We modified our methodology from our previous paper not only in carrying out bi-variate analysis but more importantly the bi-directional analysis. All of the jurisdictions that are common to both the bi-directional and bi-variate analyses help establish how night lights and economic growth are proxies for each other. We also pose some challenging hypothetical questions for future research to add to explanations and indications of the available night light data. The last section discusses results, conclusions and directions for future research.

**Night lights data**

NOAA’s Earth Observation Group provides archived DMSP (Defense Meteorological Satellite Program) image data. The DMSP is a Department of Defense program run by Airforce Space and Missile Systems (SMC) (http://www.ngdc.noaa.gov/dmsp/dmsp.html). The DMSP’s night light satellite imagery covers all of the area of the planet between latitudes 65 degrees N and 65 degrees S at the resolution of a 30 second arc, roughly equivalent to an area of 0.81 sq. km (less than 0.31 sq. miles) near the equator. The annual stable night lights data are generated by averaging only the night lights during dark half of lunar cycles and when days are shorter. These data are then filtered to remove light due to forest fires and any other natural source such as northern and southern lights so that the end result is an image at the global level of lights generated mostly by human activity. The luminosity or light intensity is a digital number between 0 and 63 where zero refers to no light while the top coding of 63 stands for maximum light. The top coding of 63 is an artifact of limitations of satellite sensors and it does point to a problem in measuring growth over time, especially in the densely populated urban regions.

**Processing night lights image data to compute Light based growth indicator (LBDG)**

The data from DMSP are available as a TIFF image measuring between 550 to 600 Mbytes. We downloaded the archived DMSP data for all the years between 1992 and 2009. Night light

---

1 Recently DMSP has also made available global night lights data from 2004 to 2009
image data for each year was imported into a GIS software application. This was then used to produce vector point data of nearly 2 Gbytes, where each point had a value between 0 and 63. However, dealing with 2 Gbytes shapefiles proved to be highly cumbersome both in time and computationally. For example it was at times beyond the limit of the GIS software available to the authors. Hence we followed an alternate shortcut method to extract light intensity information. Since our geographic unit of analysis was either a county (for the U.S.), a district (for India) or a prefecture (for China), we used a zonal tabulation routine to assign as a single measure for each of the levels of intensity between 0 and 63 for all image pixels that fell inside a county or a prefecture. This generated a table for each time period (year) where each row stands for a single county (prefecture) and each column stands for light intensity between 0 and 63, giving rise to a matrix of size $N \times M$ where $N$ are number of counties and $M$ are number of light-columns and table was repeated for $Y$ years. For example, in the case of the U.S. there were $Y=3$ matrices of size $3008 \times 64$ ($M \times N$), corresponding to 3008 counties/independent jurisdictions in the U.S. lower 48 states and 64 light columns for years 2001, 2004 and 2007. This process was repeated for same three years for 219 prefectures in China and 570 districts in India resulting in 3 matrices, 219 prefectures x 64 light columns for China and 570 districts x 64 light columns for India. For each of the counties, a weighted row sum of the data in all the light columns based on each column’s pixel value was carried out. The weighted sum for each county, referred to as the Light Based Growth indicator (LBGI), represents the total amount of light due to human activity for that county. The LBGI estimate was computed for three time periods for all of the counties in the lower 48 U.S states. LBGI was also computed for same time periods for 291 Chinese prefectures and 570 Indian districts. Next we need to find out how well the LBGI serves as a proxy for economic activity. For this we compiled county income data from the BEA (Bureau of Economic Analysis, http://www.bea.gov/) for each of the corresponding years (2001, 2004 and 2007). For Chinese prefectures and Indian districts, a GDP equivalent was obtained\(^2\). The choice of time periods was dictated in part by the availability of GDP data for China and India. For example, Chinese prefecture GDP data were available for all years between

Comparing LBGI and local Income/GDP

As part of the exploratory data analysis the scatter charts from figure 2, 3 and 4 show log-log plots of LBGI and income for U.S. counties, Chinese prefectures and Indian districts for selected years. In case of U.S. there is a pattern that can be gleaned from the spread of counties around a trend line. Similarly distribution pattern appears for Indian districts. On the other hand in case of China, the distribution of points appears random.

Since both light and income/GDP data have spatial dimensions, a spatial equivalent of exploratory data analyses was carried out based on well known spatial analysis techniques such as Moran’s index and Luc Anselin’s LISA (Local Indicators of Spatial Association) index.

Moran’s index for a given study area A consisting of n spatially distributed units with attribute value represented by $x_i$ for unit where $i \in [1,n]$, is computed as:

$$I_m = \frac{1}{n} \sum_{i \neq j} W_{ij} \left( \sum_{i,j} (x_i - \bar{x}) (x_j - \bar{x}) / \sum_i (x_i - \bar{x})^2 \right),$$  \hspace{1cm} (1)

where $W_{ij}$ is a contiguity matrix with a value of 1 if spatial units $i$ and $j$ are neighbors and zero otherwise and $\bar{x}$ is mean value of $[x_1 \ldots x_n]$. Rearranging terms in both numerator and denominator and with row standardization the term $\sum_{ij} W_{ij} = n$.

So equation (1) becomes

$$I_m \approx \frac{\sum_i (x_i - \bar{x}) \sum_j W_{ij} (x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2},$$ \hspace{1cm} (2),

where one can see that the term $\sum_j W_{ij} (x_j - \bar{x})$ is equivalent of a weighted spatial lag representing the sum over all the spatial neighbors of $i$ and the formula for $I_m$ can now be viewed

\footnote{Indian Statistical Bureau (2001-2008). India’s GDP data covers 4 quarters between April of each year and March of the following year.}
as spatial equivalent of a linear regression coefficient where a variable is regressed against the spatial lag of itself. This interpretation allows one to draw a scatter plot of a spatial variable and its spatial lag, the slope of which constitutes the Moran’s Index.

Note that \( W_{ij} \) can be a weight matrix based on some other measure of proximity between different spatial units such as number of units within a distance ‘d’. A commonly used method is to compute \( W_{ij} \) as either a “Queen” or “Rook” contiguity matrix where Queen contiguity refers to corner and edge neighbors while Rook contiguity refers to edge sharing neighbors, not unlike moves of Queen and Rook on a chess board. For our analysis we used nearest 4 neighbor contiguity.

The other spatial technique LISA (Local Indicators of Spatial Autocorrelation) is based on spatial association between different neighboring units. For a given area of study A with \( n \) spatial units, each of which with attribute \( x \) is given as follows:

\[
I_{lisa} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} \cdot (x_i - \overline{x}) \cdot (x_j - \overline{x})}{\sum_{i=1}^{n} (x_i - \overline{x})^2} \tag{3}
\]

Where \( i \) and \( j \) take value between 1 and \( n \). The LISA index helps identify similarity between a unit and its neighbors. If \( I_{lisa} \) values for a unit and its neighbor are high then together they give rise to hot-spots, on the other hand if \( I_{lisa} \) values of a unit and its neighbors are low then they give rise to cold-spots, while a unit with a high value of \( I_{lisa} \) surrounded by low value of \( I_{lisa} \) neighbors and vice versa are referred to as outliers. LISA index also offers a significance test such that for all those units below a certain significance level (\(< 0.05 \) for eg.) are identified as having no spatial association for that specific attribute. Note that LISA index summed over all spatial units is equivalent of Moran’s I for those units for a specific attribute of those units.

For our analyses we used bi-variate versions of Moran’s Index and Luc Anselin’s LISA index, where the two variables for each spatial unit were log of LBGI and Local Income (counties in the U. S) or Local GDP (Indian districts and prefectures in China). However, this analysis will be able to identify hot and cold spots where high (low) values of LBGI are surrounded by neighbors that have high (low) values of income/GDP, thus offering a weak or perhaps half explanation of how LBGI for a specific location is spatially related to income/GDP of its neighbors. In order to make LISA based explanations strong, we extended the LISA analysis by
carrying out bi-variate bi-directional analysis. The later part of the analysis refers to computing the so called complement of previous LISA, i.e., hot and cold spots now identify spatial relation between high (low) values of income/GDP locations surrounded by high (low) values of LBGI locations. Note that equations (1) and (3) are appropriately modified such that for each spatial unit, $x_i$ and $x_j$ refer to an attribute and its spatially lagged second attribute respectively. Where the spatially lagged value is computed based on $k$ neighbors determined by the weight matrix $W$. Similarly $x_{imean}$ and $x_{jmean}$ are mean values computed for each of these attributes. Thus equations (1) and and (3) are now as shown below:

$$I_m = \sum_{i,j}^{n} W_{ij} \times \left( \sum_{i,j} (x_i - X_{imean}) (x_j - X_{jmean}) / \sum_{i} (x_i - X_{imean})^2 \right), \quad (4)$$

and

$$I_{lisa} = \sum_{i=1}^{n \times n} W_{ij} (x_i - x_{imean}) (x_j - x_{jmean}) / \sum (x_i - x_{imean})^2 \quad (5)$$

The results for year 2001, 2004 and 2007 are shown for the U.S in figures 3, 4 and 5; for India in figures 6, 7 and 8 and for China in figures , 9, 10 and 11.

**Discussion of the results of Moran’s Index and LISA hot spots for three time periods, Y2001, Y2004 and Y2007**

For each of the three year time periods, we computed bi-variate and bi-directional Moran’s index using log values of county night lights and income/GDP for each of the jurisdictions and for each of the years. In each case we chose 999 permutations and then bi-variate bi-directional LISA for the same variables was computed at a 0.05 significance level. The patterns seen in both bi-variate bi-directional Moran scatter plots and LISA maps show the distribution of counties among the hot spots (High-High), Cold spots (Low-Low) and outliers (Low-High and High-Low). In each figure, the left side maps and Moran scatter plots show analysis of how each jurisdiction’s LBGI is spatially related to its neighbor’s income; while the right side maps and Moran scatter plot show how each jurisdiction’s income is spatially related to its neighbor’s LBGI. The top row are LISA maps while the middle row maps show the level of significance (0.001, 0.01 and 0.05) for the corresponding jurisdictions from the LISA maps in the top row.
The results are also summarized in the tables, each has three columns for each time period, the first two columns show count of jurisdictions associated with bi-variate LISA (LBGI to income/GDP and income/GDP to LBGI). The first two rows show the number of jurisdictions that correspond to hot and cold spots respectively, third and fourth row each has counts for outliers. The last row shows the total number of jurisdictions for each column. These are the number of jurisdictions that are accounted in the LISA analysis at the 0.05 significance level. For each year, the table also has in the third column, the count of jurisdictions that are common to both bi-variate LISA analyses, in other words its shows results of the bi-directional LISA. The count of jurisdictions in the top two rows of the third column corresponding to hotspot and coldspot bi-directional LISA while the 3rd and 4th row entries are for outliers. Since these are common to bi-variate and bi-directional LISA, we can say that at 95% confidence interval (or at 0.05 significance level) night lights data can be used as proxy for income/GDP for these jurisdictions. Each table also shows Global Moran’s index for the two variables and it appears that Moran’s Index show very little variation in either direction.

<table>
<thead>
<tr>
<th>Number of U.S. counties with LISA @ 0.05 significance level</th>
<th>Sat01Inc01</th>
<th>Inc01Sat01</th>
<th>Common01</th>
<th>Sat04Inc04</th>
<th>Inc04Sat04</th>
<th>Common04</th>
<th>Sat07Inc7</th>
<th>Inc07Sat07</th>
<th>Common07</th>
</tr>
</thead>
<tbody>
<tr>
<td>High High</td>
<td>410</td>
<td>490</td>
<td>348</td>
<td>397</td>
<td>466</td>
<td>346</td>
<td>406</td>
<td>475</td>
<td>364</td>
</tr>
<tr>
<td>Low Low</td>
<td>459</td>
<td>470</td>
<td>391</td>
<td>449</td>
<td>459</td>
<td>401</td>
<td>456</td>
<td>460</td>
<td>411</td>
</tr>
<tr>
<td>Low High</td>
<td>12</td>
<td>11</td>
<td>1</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>9</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>High Low</td>
<td>22</td>
<td>21</td>
<td>10</td>
<td>16</td>
<td>20</td>
<td>6</td>
<td>20</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>SubTotal</td>
<td>903</td>
<td>992</td>
<td>750</td>
<td>870</td>
<td>952</td>
<td>754</td>
<td>891</td>
<td>955</td>
<td>786</td>
</tr>
<tr>
<td>Global Clustering Index</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morans Index</td>
<td>Sat01Inc01</td>
<td>Inc01Sat01</td>
<td>Sat04Inc04</td>
<td>Inc04Sat04</td>
<td>Sat07Inc7</td>
<td>Inc07Sat07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global Index</td>
<td>0.6852</td>
<td>0.6811</td>
<td>0.6903</td>
<td>0.6914</td>
<td>0.6819</td>
<td>0.6785</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pval</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e(l)</td>
<td>-0.0003</td>
<td>-0.0003</td>
<td>-0.0003</td>
<td>-0.0003</td>
<td>-0.0003</td>
<td>-0.0003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>-0.0002</td>
<td>0.0002</td>
<td>-0.0003</td>
<td>-0.0006</td>
<td>0.0001</td>
<td>0.0014</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sd</td>
<td>0.0116</td>
<td>0.012</td>
<td>0.0126</td>
<td>0.0119</td>
<td>0.0119</td>
<td>0.0122</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Count of lower 48 state U.S. counties in the hot, cold spots and outlier regions from the bi-variate bi-directional LISA analysis and Global Moran’s index over three different time periods.

Between 730 to 775 counties show bi-variate bi-directional LISA index at the desired significance level of 0.05, and remains fairly constant over the three times periods. It appears that the night lights as proxy for income can be applied to only about 25% counties in the U.S., suggesting there are other factors that may contribute to income for majority of the counties in the U.S. And yet the relatively high values of Global Moran’s Index suggest spatially close clustering patterns.
Table 2. Count of Indian districts in the hot, cold spots and outlier regions from the bi-variate bi-directional LISA analysis and Global Moran’s index over three different time periods.

Less than 100 districts in India show LISA index at the desired significance level with very little variation over the three time periods. However the near constant global Moran’s Index as well as count of districts suggests that night lights data can be used as a proxy in case of between 15 to 20% districts in India.

Table 2. Count of Chinese prefectures in the hot, cold spots and outlier regions from the bi-variate bi-directional LISA analysis and Global Moran’s index over three different time periods.

Compared to the U.S. and even India, very low count of prefectures (between 5 to 10%) that are common to bi-variate bi-directional LISA analysis suggests that night lights are not a good proxy for determining local GDP. Bi-variate Moran’s scatter plots as well the actual value of Moran’s index (the slope of the trend line) shows that the distribution appears nearly random with very
small value for Moran’s index. This weak relation is also confirmed by the presence of very small number of prefectures that were common to all the three time periods at 0.05 signals

**Theoretical discussion of night light data and resulting observable geospatial patterns**

This paper extends our earlier work in important ways. Here we illustrate a three-way cross-national comparison as opposed to two. China and India are respectively the two most populated countries in the world. Modern China and India are very young countries, between roughly 60 and 65 years old. These countries’ youth means that they have had little time to establish their own contemporary institutional and rule-making structures. The U.S. and India are democratic republics, respectively the world’s oldest and largest democracies. China, on the other hand, is an authoritarian socialist state with no (allowed) competition among political parties or opposing interest groups. The U.S. and India both emerged as independent nation-states while British colonies, in both cases carrying influences from British political, social and economic traditions. However the U.S., as opposed to India, has now had roughly 235 years to establish its own social and political institutions and rule-making apparatus. Thus this three-way comparison is embedded with numerous similarities and differences among the subject nations, thereby increasing the complexity of explaining respective patterns of geographic and economic development. Nonetheless, two complementary theories shed some light on patterns of contemporary economic geography.

The American geographer, John Fraser Hart (1991) argued theoretically that every American metropolis pushes a bow wave outward along the periphery of the urbanized area. Hart’s wave analogy is literal, compared to the bow wave of a large ocean-going ship as a “standing wave” that always remains “immediately in front of the bow of a ship moving through water”. As a ship displaces water and imparts energy to it shaping a regular gravitational wave extending outward, so a wave along the urban rural fringe imparts its energy on less intensive land uses. Hart argues that this irregular wave, reflecting the irregular edges of the built-up area, creates a “zone of dissonance” wherein the “least intensive” urban land uses displace the “most intensive agricultural” land uses. Hart’s bow wave theory is incremental, often taking a full generation to realize the conversion of land use and economics from rural to urban (Hart 1991, 35-36). A perimetropolitan bow wave emanating from New York City’s urban fringe might be conceived as pushing to the west and northwest for example. As the wave sweeps over the less intensively
developed side of the urban rural fringe, the irregular shape of the wave representing economic activity may be explained with the help of Percolation theory. Why would the wave effects not reflect the nature of a literal wave, smoother, regular and expanding predictably? Percolation theory has its origins in the study of the flow of fluids in porous media (Kulkarni et al. 2000), and in analysis of the spread of diseases (Newman 2003). The notion of “porous media” is the crucial element for the present study.

The patterns of denser or more intense economic activity shown in the maps of night light data are also irregular at the edges, giving the appearance that some places are more amenable or permeable to the energy or economic activity of the urban bow wave. It is at these edges where perimetropolitan bow wave theory and percolation complement one another. Percolation theory relies on "porous media". The degree of porosity may be seen as either opportunities or constraints. For example less porous, even impermeable, media can be natural, such mountains, oceans, deserts, etc. (note the influence of the Great Lakes on the patterns for the US data, and that coastal urban areas in all three countries expand toward the interior). Constraints on permeability may also be political, such as land use controls and zoning, even designations of national and state parks and wilderness areas that function in the same constraining way, i.e. reducing porosity. Thus we have sets of natural and political (or artificial) constraints on the ability of the bow wave to percolate. Percolation theory also suggests ways to overcome such constraints, or to increase porosity.

Newman (2003, 225) distinguishes between “site percolation” and “bond percolation”. Thus two sites, established urban areas may be separated by non-porous or impermeable features such as terrain features or a wide river valley. The sights might be “bonded” via highways, railroads, bridges or air services. Likewise land use regulations and zoning restrictions might be modified to accept more intensive development, thereby increasing the porosity of the medium opposite an impending perimetropolitan bow wave. In the case of China, as a command economy, the infrastructural “bonds” may be supplied in advance of any current economic demand or social need, which may help explain the almost random appearance of Chinese spatial development patterns outside established larger urbanized areas. That is, long stretches of Chinese highways and railroads may, at least for the time being, be seen as “roads to nowhere.”
Conclusions and Future Research

Our current analysis based on three different time periods suggests that trying to determine whether night lights data can be useful as proxy for local and sub-regional economic activity indicator appear to be promising in case of the U.S but only for 25% to 33% counties in the lower 48 states. In fact among these less than 770 counties of the 3,100 plus total number of counties one could determine with 95% confidence interval that it is highly likely that the night lights can be used as a proxy for county income. In case of India, use of night lights as proxy for local GDP is limited to very small number of India’s 570 plus districts. However this number as well as the spatial pattern stays nearly constant over the three different time periods, suggesting a limited promise of substituting night lights for GDP. For China, it turned out to be a very mixed bag, with our current analysis showing that for less than 10% of prefectures we could say that at 0.05 significance level, it was highly likely that one could use night lights as a proxy for GDP. This low number of prefectures suggests that the main problem is likely due to the limit set due to saturation level of satellite sensors such the top coding value is constant of 64 no matter how much light emitted by a pixel. In the case of China where most of the east coast prefectures have seen explosive growth between 1997 and 2007, this ceiling in the value of top coded pixel cannot capture this growth. Also, it is likely that the huge migration from rural to urban areas have created surplus population in some of the prefectures that may not contribute to that sub-region’s growth. Moreover, China’s internal migration may be fueled in part by trans-national rail lines and highways, enabling rural peasants to move more easily to the large urban cores. These later observations have to be tested with some other data sets which at present are not available.

The process of extracting and analyzing light data has shown the way for potential future research in other areas of regional economics, such as using the night lights data as a footprint for determining urban boundaries, also rates of change of such footprint from one time period to next may help us detect differential growth patterns at a much finer scale than a county or a prefecture. Since DMSP has released night lights data for years 2004 to 2009, we have a fairly long time series data (1992 to 2009) that can be utilized to determine spatial and non-spatial principal component analysis (PCA) of night lights data. Assuming that we have a priori knowledge of economic growth attributable to specific sectors of local economies, we think that components with significant values of PCA will help us identify differential sectoral impact on these jurisdictions. We also would like to explore how we could combine Hart’s bow-wave
model with the ideas from Percolation theory to analyze economic development patterns represented by night lights data at the regional/sub-regional level. Yet another research avenue for future is to use the nightlights data in the rural or non-metropolitan areas of the U.S. for studying issues related to economic development that are dominated by economic sectors other than Service sectors.

***

Acknowledgement

We would like to express our thanks to Prof. K. Ramachandran, Thomas Schmidheiny Chair Professor of Family Business and Wealth Management, Indian School of Business and Prof. Tojo, School of Public Policy, GMU, for providing access to the Indian GDP data and for their valuable suggestions. We would also like to thank Dr. Hauqun Li, an SPP alumni for providing Chinese GDP data and special thanks to Prof. Laurie Schintler, School of Public Policy, GMU for research discussion and valuable suggestions.

References


Figures: Charts and Maps

Figure 1: A schematic of night lights data (country maps not drawn at same scale)
Figure 2: Scatter plots of night lights vis income/GDP

Figure 3: Y2001 U.S. bi-variate bi-directional LBGI vis GDP
Figure 4: Y2004 U.S. bi-variate bi-directional LBGI vis GDP

Figure 5: Y2007 U.S. bi-variate bi-directional LISA: LBGI vis Income
Figure 6. Y2001 India bi-variate bi-directional LISA: LBGI vis GDP

Figure 7. Y2004 India bi-variate bi-directional LISA: LBGI vis GDP
Figure 8. Y2007 India bi-variate bi-directional LISA: LBGI vis GDP

Figure 9. Y2001 China bi-variate bi-directional LISA: LBGI vis GDP
Figure 10. Y2004 China bi-variate bi-directional LISA: LBGI vis GDP

Figure 11. Y2007 China bi-variate bi-directional LISA: LBGI vis GDP