A Report of "Using Nighttime Satellite Imagery as a Proxy Measure of Human Well-Being"

Mengting Lei

December 2018

1 Introduction

In 1973, the Defense Meteorological Satellite Program (DMSP) was established in dealt with meteorology and military weather forecasting. However as time passes by and its data becomes more publicly available, researchers have noticed another use of the DMSP data—nighttime light maps. An necessity of life, lights have been a indicator for human presence since the first day of human history. Moreover, the invention and popularization of electrical lights even make the standards of this indicator more unified and quantifiable. Therefore since 1990s, these maps have spawned hundreds of economic, social science and environmental research projects, including pollution monitor, disaster control and so forth. For example, Figure 2 is a comparison of two nighttime light maps of Wilmington, North Carolina before and after the hurricane Florence in September, 2018. Without any further data analysis, the picture itself delivers a rather clear message in terms of the dimming of lights across the city, indicating a evacuation of considerable size in progress.

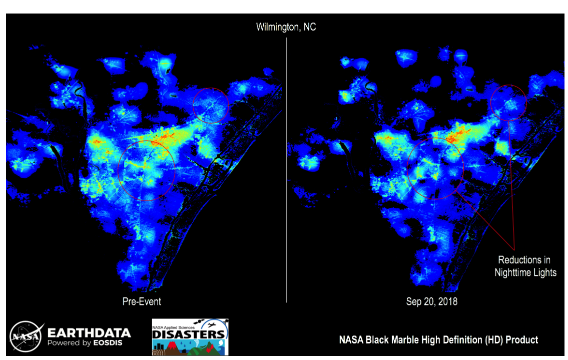

Figure 1: *The nighttime light map of Wilmington, NC before and after Hurricane Florence by NASA.*

2 The study

While there are more than a dozen of world nighttime light map sources available to the public nowadays (e.g. NASA, Light Pollution Map, Night Earth, etc.), the first to start this trend remains unknown. This paper uses the data from DMSP because they are free and more importantly, light intensities data are also published along with the light maps, thus sparing the need for further data mining and image processing. Just as most of the data provided by other sources, DMSP data are pre-processed to remove occasional noises before publishment. These noises include fire, sunlight, moonlight, aurora, and cloud impact. The reported light score ranges from 1 to 63 according to the annual light intensity, with zero representing background lighting.

The paper in examination is a comprehensive review in ways nighttime satellite imagery can be used to measure the human well-being within nations and it touches seven perspectives. We are only discussing about one of them: Gross National Income (GNI). Gross Domestic Product (GDP) is formally defined by the market value of all the final goods and services produced of a country in a certain period of time, usually in a calendar year. While it is a frequent indicator for judging the standard of living of a country in comparison to others, GDP fails to reflect informal transactions outside markets, such as street vendor, voluntary work, and household work. Moreover, it provides no information for remittances, which is sent back by international migrants to their countries of origin. To account for these unattainable transaction amounts, the concept Gross National Income (GNI) was coined, defined by the sum of GDP and incomes a country receives from overseas. Due to the nature of outside-market transactions and remittances, GNI can only be estimated, and thus leading to huge disparities among the amounts from different estimators.

To solve this problem, researchers started to use nighttime light data to estimate GNI in the past decade. This paper particularly elaborates on a study in Mexico GNI estimation. The research goal is to develop a regression model for United States GNI based on nighttime lights and urban population, then train the model with official U.S. data, and lastly use the same model to estimate GNI. The assumption of this study is that economic activity associated with urban populations creates the same spatial patterns of nighttime lights in Mexico as in the U.S..

Since the nighttime light data are available for direct use, a large portion of this study is to determine the urban population. In order to do so, the researchers first identify areas in the 2006 U.S. nighttime light imagery that exceed the empirical brightness threshold of 201.3510^{-10} $watts/cm^2/sr$ as urban regions. Then the 'thresholded' nighttime image was used to mask the Landscan population grid, which is an open-access global population distribution data source with $1km^2$ spatial resolution. This process extracts the urban populations of each U.S. state from the areas demarcated by the brightness threshold. After that, a log-log regression model was used to estimate urban population for each of the 48 contiguous U.S. states for the size of urban areas and corresponding urban populations. The result shows that the natural

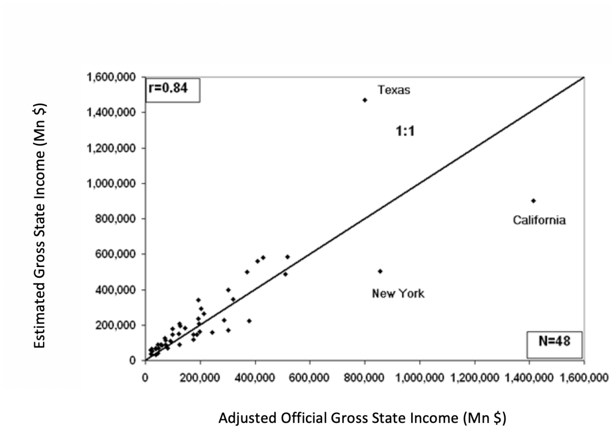

Figure 2: *The actual versus predicted plot of the ln(EGSP) values of the U.S. states derived from the multiple regression model*

log of the size of urban areas of U.S. states (P_{USi}) and the natural log of the population (A_{USi}) of the U.S. states have the following linear relationship:

$$ln(P_{USi}) = \alpha_{1US} + \beta_{1US} + \times ln(A_{USi})$$

The regression parameters α_{1US} and β_{1US} were 5.10 and 1.07, respectively. The model is statistically significant and the coefficient of correlation is 0.87.

Next, a multiple regression model was developed for estimating Gross State Income for each U.S. state based on the estimated urban populations of the 48 contiguous U.S. states. The estimated urban population of each of the 48 U.S. states ($P'US_i$) and the light intensity score for each U.S. state (SUS_i) were the predictors in the regression model. The regression parameters, α_{2US}, β_{2US}, and β_{3US} were determined to be 16.11, 0.62, and 2.110^{-7}, respectively. The estimated Gross State Income ($EGSI_{USi}$) for each U.S. state was subsequently estimated by the exponentiation of the logarithmic equation:

$$ln(EGSP_{USi}) = \alpha_{2US} + \beta_{2US} \times ln(P'_{USi}) + \beta_{3US} \times S_{USi}$$

$$EGSP_{USi} = exp(\alpha_{2US} + \beta_{2US} \times ln(P'_{USi}) + \beta_{3US} \times S_{USi})$$

Figure 1 is the Actual-versus-Predicted plot for the log of the officially reported GSI values. There are three outlier states: Texas, New York, and California, the former one seriously underestimated and the latter two overestimated by official GSI according to the estimated GSI. The paper does not elaborate on the potential reason behind the results. Personally I think the reason may be related to one commonality of these three states—big population, which can affect the estimation for urban population.

After that, the same regression model developed for the U.S. states was used to estimate the estimated Mexico GSI ($EGSI_{Mexi}$) of each Mexican state using

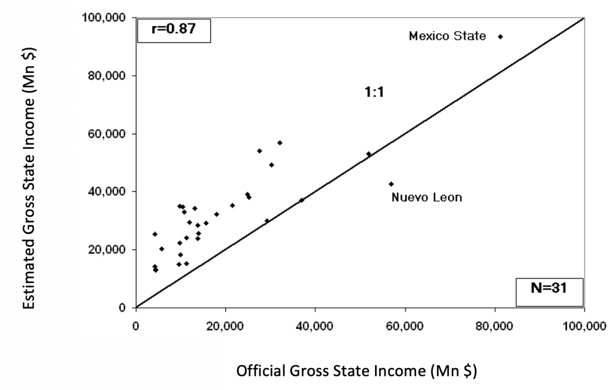

Figure 3: *Official* GSI_{Mex_i} *versus Modeled* $EGSI_{Mex_i}$ *of the Mexican states, excluding Mexico City.*

the light intensity data for each Mexican state (S_{Mexi}) and estimated urban population of each Mexican state (P'_{Mexi}):

$$EGSI_{Mex_i} = exp(\alpha_{2US} + \beta_{2US} \times ln(P'_{Mex_i}) + \beta_{3US} \times S_{Mex_i})$$

In Figure 3, Estimated Mexico GSI ($EGSI_{Mex_i}$) was plotted against the official Mexico GSI (GSI_{Mex_i}) for each Mexican state, excluding Mexico City. The regression model is overall statistically significant and the coefficient of correlation is 0.87. The plot shows that official Mexico GSI (GSI_{Mex_i}) was overestimated for 27 of the Mexican states and underestimated for only one state: Nuevo Leon. And there are in total two outliers: Nuevo Leon and Mexico State. Again, the authors do not provide further explanation. This time I look into the population of these two outliers and found them ranked as the first and eighth in population, respectively, among all 31 Mexico states. Such fact does not offer me any more insight into the possible reasons behind the results.

After summing up the estimated GSI for all Mexican states (US $\$1,041$ billion) and compare it to the officially reported GSI of Mexico(US $886 billion), the authors notice approximately 17.4% underestimation for the official accounts. Such result is reasonable since most states in Mexico have more lighting compared to their officially reported GSI would suggest. Therefore they suggest that such disparity could be attributed to the informal economy and inflow of remittances in Mexico. Since the current official Mexico GSI is calculated as 12% GDP of the same year, the authors suggest that the informal economy in Mexico may be larger than 12% of GDP.

3 Reflection on My Presentation

I was glad to see that nearly all of our classmates commented this topic and the presentation as interesting and innovative. There are certainly flaws inbuilt in the methodology of this study: even though the regression model developed for U.S. was statistically fit for Mexico states, it is possible that it is still meaningful. As the U.S. is usually recognized as the most developed country in the world, Mexico is still an developing country, which will certainly lead to different economic compositions, including informal transactions and remittances. In addition, one classmate offered a thought that different regions may also have different habits in terms of using electrical lights. For instance, the New York City has a nick name of Sleepless City, and we would thus expect much less economic activity-related lighting, which is exactly one of the results.

Admittedly, this general approach of using nighttime lights to investigate human behaviors is certainly an encourageable one. First of all, there are a range of different satellite imageries and data available to the public and up until now, they have already possessed great spatial and temporal resolution. Secondly, I would not be surprised to see many other applications of these imageries in the future. For example, since there is such a bounty reservoir of good data, we can even predict future light intensities worldwide and use it as a proxy to predict nearly all aspects of human activities. Last but not least, Professor CK raised a question about what is another good indicator of human activities. My personal response is that there could be many, such as temperature change (Urban heat island effect), amount of rubbish disposal, amount of staple food (rice, corn, potato...) consumption, etc. However what really matters is that which indicator of human activities is readily available to the public or can be easily obtained. I still do not have an answer to this question.

References

[1] Ghosh, T., et al. [*Estimation of Mexico's Informal Economy and Remittances Using Nighttime Imagery*]. Remote Sens. 2009, 1(3), 418-444; doi:10.3390/rs1030418.

[2] Ghosh, T., et al. [*Using Nighttime Satellite Imagery as a Proxy Measure of Human Well-Being*]. Sustainability 2013, 5, 4988-5019; doi:10.3390/su5124988.