Journal of Development Economics 149 (2021) 102602

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Regular Article

Which night lights data should we use in economics, and where?

John Gibson *, Susan Olivia, Geua Boe-Gibson, Chao Li [1]

ARTICLE INFO

JEL classification:
O15
R12

Keywords:
Density
DMSP
Inequality
Night lights
VIIRS
Indonesia

ABSTRACT

Popular DMSP night lights data are flawed by blurring, top-coding, and lack of calibration. Yet newer and better VIIRS data are rarely used in economics. We compare these two data sources for predicting GDP, especially at the second subnational level, for Indonesia, China and South Africa. The DMSP data are a poor proxy for GDP outside of cities. The gap in predictive performance between DMSP data and VIIRS data is especially apparent at lower levels of the spatial hierarchy, such as for counties, and for lower density areas. The city lights-GDP relationship is twice as noisy with DMSP data than with VIIRS data. Spatial inequality is considerably understated with DMSP data, especially for the urban sector and in higher density areas. A Pareto adjustment to correct for top-coding in DMSP data has a modest effect but still understates spatial inequality and misses key features of economic activity in big cities.

1. Introduction

Night lights, as detected by satellites, are increasingly used by economists, especially to proxy for local economic activity in poor countries. A recent review finds over 150 studies in economics using night lights, almost all of which use the Defense Meteorological Satellite Program (DMSP) data even as there is a rapid switch to using newer and better data from the Visible Infrared Imaging Radiometer Suite (VIIRS) in other disciplines (Gibson et al., 2020). Flaws in DMSP data include blurring, coarse resolution, no calibration, low dynamic range, top-coding, and unrecorded variation in sensor amplification that impairs comparability over time and space (Elvidge et al., 2013; Abrahams et al., 2018; Bluhm and Krause, 2018). Many of these flaws stem from the original purpose of DMSP to detect clouds for short-term Air Force weather forecasts. In contrast, the VIIRS Day-Night Band (DNB) was designed to help researchers consistently measure the radiance of light coming from earth, in a wide range of lighting conditions (covering almost seven orders of magnitude while DMSP covers less than two), with high spatial accuracy and with temporally comparable data.

Continued use of DMSP data by economists reflects several factors. First, there is larger scholarly impact, in terms of citations, from Henderson et al. (2012) who argue DMSP lights data can be used successfully in a wide range of circumstances, than from Chen and Nordhaus (2011) who offer more limited support.[2] Second, the longer DMSP time-series, from 1992 to 2013, is attractive to economists. Two caveats to that potential advantage are that night lights do better at predicting differences in economic variables over space than changes in those variables over time (Nordhaus and Chen, 2015; Goldblatt et al., 2020; Gibson and Boe-Gibson, 2020), and the time-series is old, with DMSP data stopping in 2013 while VIIRS data are available monthly with only a slight lag.[3] Also, economists may have more slowly switched to using VIIRS night lights data, compared to other disciplines, because flaws in DMSP data are rarely highlighted in the economics literature. The remote sensing literature has several comparisons that highlight the superiority of VIIRS (e.g. Elvidge et al., 2013; Shi et al., 2014; Chen and Nordhaus, 2015, 2019) but these studies are rarely cited in economics journals.

In light of the limited comparisons between DMSP and VIIRS within economics, the current paper presents a test of these data for estimating regional GDP and inequality for rural and urban areas. We concentrate especially on Indonesia, as one of the few developing countries with

* Corresponding author.
E-mail address: jkgibson@waikato.ac.nz (J. Gibson).

[1] Gibson, Olivia, Boe-Gibson and Li are with the Department of Economics, University of Waikato. This paper was begun while Gibson visited the Centre for the Study of African Economies, Department of Economics, University of Oxford and he acknowledges their hospitality. Helpful comments from seminar audiences at CERDI, Oxford, and the Tinbergen Institute and from two anonymous referees and the editor are gratefully acknowledged. Financial support is from Marsden Fund grant UOW-1901. These are the views of the authors.

[2] Specifically, Chen and Nordhaus (2011, p.8594) noted that "luminosity data do not allow reliable estimates of low-output-density regions" and only for countries with the worst statistical systems, accounting for under nine percent of world population, are DMSP night lights data likely to add value as a proxy for economic output.

[3] As of November 2020, monthly VIIRS data from April 2012 to May 2020 can be downloaded from: https://eogdata.mines.edu/download_dnb_composites.html.

https://doi.org/10.1016/j.jdeveco.2020.102602
Received 19 July 2020; Received in revised form 19 December 2020; Accepted 27 December 2020
Available online 1 January 2021

reliable GDP statistics for the second sub-national level. Night lights data are potentially most valuable for developing countries, because richer countries often have other sources of information on local economic activity, including more conventional data such as sub-provincial (e.g. district or county) GDP estimates. Indonesia is also of interest in its own right, as an important developing country. Nevertheless, to provide some sense of the external validity of our results, we also report evidence from China and South Africa.

In the main part of the paper we report on the relationship between night lights and Indonesian GDP at the second sub-national level for 497 spatial units. The DMSP data do not appear to be a suitable proxy for GDP outside of cities, with a negative relationship between DMSP lights and the real GDP of non-urban spatial units, while VIIRS data positively relate to GDP. The negative lights-GDP relationship is especially for rural areas of low population density, in line with the conclusion of Chen and Nordhaus (2011) that luminosity data do not provide reliable estimates of economic activity for low output density regions. A related finding with cross-country data comes from Keola et al. (2015); DMSP night lights negatively relate to GDP if agriculture is a large share of GDP, while there is a positive relationship otherwise. In Indonesia's urban sector, the lights-GDP relationship is positive irrespective of the type of night lights data used, but the relationship is twice as noisy if the DMSP data are used rather than the VIIRS data. We also find spatial inequality considerably understated by the DMSP data, especially in the urban sector. A Pareto-based adjustment from Bluhm and Krause (2018) to deal with top-coding in DMSP data has a modest effect but the adjusted data still greatly understate spatial inequality, and miss much of the intra-city heterogeneity in brightness (coming from key economic facilities) for Jakarta.

Similar patterns are evident for the other two countries we study. Spatial inequality is understated if the DMSP data are used as a proxy for local economic activity, especially in the urban sector or in higher density regions. Night lights data have higher predictive power for sub-national GDP when working with more aggregated spatial units, such as provinces, while the gap between DMSP and VIIRS in terms of predictive power gets wider for analyses based on smaller or lower level units within the administrative hierarchy. The data for China allow us to compare lights-GDP relationships in the primary sector with those in industrial and services sectors and this comparison shows that night lights data are a poor proxy for primary sector economic activity. Overall, our results suggest that great caution is needed if development economists use night lights data, not only for low density rural areas but also when studying the spatial patterns of urban development.

2. Background and related literature

Researchers have used night lights data from the Defense Meteorological Satellite Program Operational Linescan System (DMSP for short) for over 40 years, even though these satellites were designed to observe clouds for short-term weather forecasts rather than to give a consistent long-term record of lights on earth. While a few earlier studies by remote sensing researchers had an economics focus, it was not until Henderson et al. (2012), and to a lesser extent Chen and Nordhaus (2011), developed optimal weighting of night lights and reported GDP, in order to predict true GDP, that many economists paid attention to night lights data.[4] These key studies noted that night lights data are noisy but concluded that in a fairly wide range of contexts (Henderson et al., 2012), or, alternatively, in a narrower set of contexts (Chen and Nordhaus, 2011), DMSP lights data could add value to conventional economic statistics like national and regional GDP. Some following studies noted there was far more uncertainty in lights-based time-series of GDP estimates than for cross-sections of GDP (Nordhaus and Chen, 2015), and that unstable relationships made DMSP data a poor proxy for regional economic activity (Bickenbach et al., 2016).

Despite these critical findings, many more applied economics studies – especially in development economics – have used the DMSP data to study a wide range of topics (see Gibson et al., 2020 for a survey). Recently, Abrahams et al. (2018) and Bluhm and Krause (2018) link measurement errors in the DMSP data to inherent flaws in the sensors and data processing, and propose correction methods. We briefly summarize these flaws, and contrast DMSP with features of the Day-Night Band (DNB) of the Visible Infrared Imaging Radiometer Suite (VIIRS), which is a far more accurate source of night lights data from the *Suomi* satellite that launched in 2011.

The DMSP data lack spatial accuracy because the sensor and data processing allocate light to different places than the point of origin. Early studies argued this was from reflection off water or snow and so may matter in only some places.[5] In fact, 'blurring' is an inherent feature of DMSP data (Abrahams et al., 2018). The DMSP satellite altitude is just one-quarter the width of the 3000 km sweep of the sensor, so away from the nadir the earth is viewed at an angle and the field-of-view expands (by 400% at the edge and 240% at the half-sweep, where NOAA stop using the data) but all light from this expanded footprint is attributed to a smaller pixel in the centre. The on-board computers cannot hold all the data so pixels are aggregated to 5 × 5 blocks (of size 2.7 km × 2.7 km at the nadir) prior to the data being sent to earth, further spreading light from point of origin. Random geo-location errors, with a mean of about 3 km, further spread lights data (Tuttle et al., 2014). In contrast, VIIRS has near-constant spatial resolution across the sweep of the sensor, by compensating for the expanded ground footprint as the scan goes towards the edge, and handles finer 0.7 km × 0.7 km pixels due to abundant data storage. Spatial errors in DMSP data show up as exaggerated estimates of lit area: with errors averaging 77% for big cities (Abrahams et al., 2018) and getting up to 500% for smaller towns (Gibson et al., 2020). In overstating the lit area of towns and cities, DMSP data wrongly attribute light to hinterland areas, introducing cross-sectional noise.

The lack of temporal consistency, which causes errors in time-series of DMSP-based GDP estimates, is from two main sources. First, there is no on-board calibration, with on-the-fly changes in sensor amplification over the monthly lunar cycle not recorded in the data. The signal is amplified going into the dark part of the month to keep the brightness of moon-lit cloud tops the same, so lights on earth then appear brighter, but no record is kept to allow *ex post* adjustment to restore consistency (Hsu et al., 2015). Also, the number of nights with images meeting quality controls for inclusion in the DMSP annual composites varies widely over time and space, due to factors like cloudiness (especially near the equator), so convergence to an average amplification level that might provide comparable data over time and space is unlikely (Gibson et al., 2020). A lack of temporal consistency in DMSP data is exacerbated by limited on-board data storage, with continuous measures converted to 6-bit integers (the Digital Number (DN) that ranges from 0 to 63, as $2^6 = 64$), that are subject to censoring from the top and the bottom (Abrahams et al., 2018).

A lack of calibration for DMSP also shows up as inter-satellite differences. For 12 of the 22 years (from 1992 to 2013) with DMSP annual composites, data are from two satellites in orbit that often report very different values. For example, for a place (Sicily) with little temporal variation, Gibson et al. (2020) show that satellite F12 gave 29% higher DN values than F14 in the overlapping years, F15 recorded a 24% decline in the DN value from 2002 to 2003 while F14 showed just a 2% change in

[4] While Chen and Nordhaus (2011) is clearly an economics contribution, it is not in an economics journal and only one-quarter of the ca. 500 citations (in *Google Scholar* as of Sept 9, 2019) are from economics journals or working papers. In contrast, far more of the ca. 1200 citations to Henderson et al. (2012) are from economics, so most economists with exposure to night lights data will have gained this through Henderson et al. (2012).

[5] For example, see the discussion in footnote 3 of Michalopoulos and Papaioannou (2014).

Table 1
Descriptive statistics for second level sub-national spatial units in Indonesia.

	All spatial units			*Kabupaten* (mainly rural)			*Kota* (cities)		
	Mean	Std Dev	Median	Mean	Std Dev	Median	Mean	Std Dev	Median
Population (thousands)	478.0	570.0	254.3	464.4	547.9	261.6	533.1	650.2	229.5
Area (km^2)	4012.1	6044.7	2044.5	4928.7	6419.8	2780.1	280.4	421.2	166.2
Density (persons per km^2)	996.2	2376.8	139.1	279.9	425.6	83.3	3912.6	4168.8	2169.0
Real GDP 2011	62.7	180.8	7.4	64.0	185.9	7.3	57.2	158.9	9.0
Real GDP 2012	55.8	163.5	7.7	57.2	169.4	7.6	50.1	137.5	8.7
Real GDP 2015	54.9	162.1	8.5	56.8	171.8	8.5	47.3	114.5	9.1
Real GDP 2016	56.0	166.3	9.0	57.5	175.8	8.9	50.0	120.6	9.6
Sum of DMSP lights, 2011	5927.4	8143.0	2685.0	6114.4	8678.8	2391.0	5166.3	5411.1	2981.0
Sum of DMSP lights, 2012	6668.7	9328.9	2930.0	6924.0	10020.5	2727.0	5629.2	5627.0	3397.0
Sum of VIIRS lights, 2015	2903.6	4992.4	1034.2	2386.2	4263.8	822.3	5010.2	6875.1	1839.8
Sum of VIIRS lights, 2016	2730.3	4928.0	964.2	2151.8	4096.1	729.4	5086.0	6953.6	1969.4

Notes: There are n = 497 spatial units, with n = 399 *Kabupaten* and n = 98 *Kota*. Std Dev is standard deviation. The DMSP lights data are a Digital Number (0–63) while the VIIRS data are radiances (in nano Watts per cm^2 per steradian). Real GDP is in billions of Rupiah, in spatially and temporally real prices (with a 2010 base).

the same year, and F18 gave a 32% higher DN value in 2010 than what F16 gave in 2009.[6] Inter-satellite differences and unrecorded changes in sensor amplification (with variation in how many nights contribute to annual composites limiting the convergence to some average amplification level) make it doubtful that a DN value from a certain satellite year refers to the same brightness as the same DN in another year (Doll, 2008). In contrast to these temporal consistency problems with DMSP, VIIRS sensors are calibrated radiometers whose instruments measure light intensity (in units of nanoWatts/cm^2/sr). The VIIRS sensors have in-flight calibration to ensure that data are comparable over time and space and the continuous signal is quantized with 14-bit precision (n = 16,384 potential values) compared to the coarse 6-bit Digital Number for DMSP.

Further flaws in DMSP data stem from the limited dynamic range of the sensors. City centers often get the same digital number (usually DN = 63) as less brightly-lit suburbs (Bluhm and Krause, 2018) and smaller towns (Gibson, 2020). This is because the DMSP sensor dynamic range covers less than two orders of magnitude so it cannot simultaneously capture light from brightly lit areas and from dimly lit areas. Under usual conditions, with sensor amplification turned up to view cloud tops, pixels in city centers are saturated with light and get the top-coded DN values of 63.[7] In contrast, the dynamic range of the VIIRS Day-Night Band is about seven orders of magnitude (L_{max}/L_{min} = 6,700,000) so there are no saturation problems with the VIIRS data. Bluhm and Krause (2018) suggest that lights follow a Pareto distribution (and use the more accurate VIIRS data to verify this), and use this fitted distribution to create an adjusted set of DMSP data that correct for top-coding.[8]

In addition to these flaws in the DMSP data, it is becoming clear that satellite data on night lights (including from VIIRS) are poorly suited to the study of areas of low population density, which includes most rural places. One reason is that the sort of lights typically used in rural villages are not the type easily detectable from space. An experiment on accuracy of DMSP data, where researchers lit up previously dark areas, needed a bank of 1000-W high pressure sodium lamps (large lamps of about 25 kg each, usually used in big warehouses), modified with aluminum shields to help direct light skywards (Tuttle et al., 2014) in order to be seen with the DMSP sensors. Such lights are not found in rural villages, but are more like light from concentrated street lamps and industrial facilities, which are typically found in urban areas. While VIIRS can better detect dimly-lit areas, the overpass time when the *Suomi* satellite that hosts VIIRS observes earth is around 1.30am, and lights coming from the household sector in rural areas are unlikely to be switched on then (while urban street lights tend to stay on all night).

There are several examples of this inability to detect low density areas. Nordhaus and Chen (2015) divided the globe into 1 ° × 1 ° grid cells and about one-third of cells with positive population and output are recorded as having zero light in the DMSP data. In a follow-up for Africa, Chen and Nordhaus (2015) found almost half of the cells were recorded as having zero light in the DMSP data, yet all cells recorded light in the VIIRS data. Indeed, even for cells having between 10,000 and 100,000 people, 51% are recorded as having zero light in the DMSP data. Other examples from Africa, Asia and the Pacific of low density areas that are home to up to 70% of the population not being detected by either DMSP or VIIRS, even when more than half of households in those areas use electric light, are given by Gibson et al. (2020). At a more aggregate level, in a cross-country panel study of relationships between night light (annual composites of DMSP lights) and national GDP, Keola et al. (2015) find positive elasticities of light with respect to GDP for countries where the agricultural share of GDP is less than 20% but negative relationships when the agricultural share exceeds that threshold. The authors note that agriculture's value-added can increase without an increase in lights, which is much less true for value-added in the urban sector.

3. Data and methods

In light of the above literature, comparing performance of DMSP data and VIIRS data in predicting regional GDP is useful. Such an analysis should consider urban and rural areas separately, given sectoral differences in types of economic activity, in population density and in lighting types, that likely affects the detection performance of the satellites. One issue with such a test is imperfect temporal overlap of the two data sources. The DMSP data are only available annually, with the time-series ending in 2013. Currently, annual VIIRS composites are only available for 2015 and 2016. While monthly VIIRS data are available from April 2012, there is potential to introduce extraneous elements if the monthly and annual data are compared. For example, not only are there seasonal differences like summer glare, the annual composites (for both DMSP and VIIRS) undergo further processing by scientists at the Earth Observation Group of NOAA (National Oceanic and Atmospheric Administration) to screen out ephemeral sources of light such as aurora, fires and gas flares

[6] Significantly higher values from F18 than F16, for the same light source (an experiment where portable generators powered high-pressure sodium lamps in previously dark areas) are reported by Tuttle et al. (2014).

[7] Based on experiments when researchers had the Air Force fix the DMSP sensor amplification at low levels on a few nights to avoid DN values being top-coded, there are radiance-calibrated lights available for seven years. These rely on pre-flight calibrations of the satellites, rather than their degraded (from exposure to dust and radiation) actual performance, and require merging with the usual DMSP data so as to create annual composites (Hsu et al., 2015). This process seems to create some instability between years (Bluhm and Krause, 2018).

[8] Other approaches to dealing with top-coding (or 'saturation') involve combining DMSP data with other remote sensing data that measure vegetation, as city centers are usually brightly lit but have little vegetation, to create a vegetation adjusted night light index. For example, the index created by Zhang et al. (2013) is widely cited in other disciplines but is not at all cited in economics journals, in contrast to the Bluhm and Krause adjustment.

and to mask background (non-light) noise. An annual sum of the monthly VIIRS data may not be comparable to the annual composite of either VIIRS or DMSP data because the monthly VIIRS data have not undergone this screening process. Therefore, in order to ensure the closest like-with-like comparison, we mainly restrict attention to the annual composites provided by NOAA. This limits the length of the time-series and so the fact that Indonesia has a lot of heterogeneity is valuable in providing a source of cross-sectional variation.[9] As noted above, Indonesia is also one of the few developing countries with reliable second level sub-national GDP data that we can use as our benchmark.

Our research design relies on the fact that the second sub-national level in Indonesia is comprised of two types of spatial units. The first is *Kabupaten* (regencies), that are mainly rural areas and towns, with a mean (median) population density of 280 (83) persons per square kilometre (Table 1). The other type of spatial unit is highly urbanized *Kota* (cities), with a mean (median) population density of 3900 (2200) persons per km^2 in the 2010 census. Both types of spatial unit are quite populous; the average *Kota* had a population of 530,000 and the average *Kabupaten* had 460,00 in the 2010 census, with 22% of Indonesia's population located in *Kota* (of which there are $n = 98$) and the rest in the *Kabupaten* ($n = 399$).[10] If we cannot detect lights, or if lights poorly predict GDP, for such populous units then we would expect at least as poor a performance elsewhere in settings with less populous units.[11]

We use three data sources to test the relationship between night lights and Indonesia's second-level sub-national GDP. The first is VIIRS annual composites for 2015 and 2016, that are the earliest (and currently only) available annual composites. We use the "vcm-orm-ntl" product that, at the pixel level, excludes nights where Day-Night Band images are affected by stray light or by clouds. These annual composites also have outliers removed by the NOAA scientists, where these outliers may be due to ephemeral sources of light, such as fires, gas flares or from fishing boats, and the background (i.e., non-lights) is set to zero. The data are radiance values in units of nano Watts per square cm per steradian (nanoWatt/cm^2/sr) and range from zero to about 1600 for Indonesia.

The second data source is DMSP annual composites for 2011 and 2012, also from NOAA (e.g. for 2012 the file is F182012_v4b_stable_lights.avg_vis.tif). The F18 satellite providing these images has a 4-year time series starting in 2010. A feature of DMSP data is that the first and last year of the time series for each satellite often have fewer nights whose images contribute to the annual composite (Gibson et al., 2020), and so we use the middle two years to provide the most reliable annual estimates. This also helps maintain comparability with the VIIRS data that are also for two years. The DMSP data are Digital Numbers, ranging from 0 to 63, and have no interpretation in terms of radiance values.

The third data source is the Indonesian government's Central Bureau of Statistics (BPS) estimates of Gross Regional Domestic Product (GRDP). The BPS calculate and report GRDP at both the provincial level, and the next level down (*Kabupaten/Kota*). For this paper we use GRDP data from 2011 to 2016, that are in spatially and temporally real terms and have a 2010 base. As a cross-check on the reliability of these regional GDP data we compared them with survey totals (at the second sub-national level) for total household consumption from the national socioeconomic survey (SUSENAS) and with total employment and total wages from the national labour force survey (SAKERNAS). The correlations were 0.73 for consumption, 0.64 for employment and 0.61 for wages.

We present descriptive statistics for GDP and the two types of night lights data in Table 1. According to the VIIRS data, the average *Kota* emits more than twice as much light as the average *Kabupaten* but with DMSP data *Kabupaten* seem to produce about 20% more light than *Kota* (at the means). A likely reason is that DMSP data spread light, due to blurring, and also attenuate the differences in light output between areas, due to top-coding brightly lit city centers, with both flaws causing mean-reverting errors (Gibson, 2020). On average, a *Kabupaten* is over 15-times the area of a *Kota* and so with the mean-reverting errors, when one sums across the pixels in a *Kabupaten* a larger total of light gets accumulated compared to what occurs with the more accurate VIIRS data that do not have blurring or top-coding.

The relationship we estimate is $\ln(\textit{real GRDP})_{it} = \alpha + \beta \ln(\textit{lights})_{it} + \delta_t + \varepsilon_{it}$ where the dependent variable is the log of the regional GDP in each year (we use the terms GDP and GRDP interchangeably). The main right-hand side variable is the log of the sum of lights within each *Kabupaten* or *Kota* in each year, coming from either DMPS or VIIRS and we also use a time dummy, δ_t in regressions when we pool data from two years. Any issues with the BPS deflators, which suggest that real GDP fell in 2012 and was fairly stagnant after that, should also be soaked up by the time dummy. The relationship we estimate is not meant to imply that lights cause GDP because as a production function one would write it the other way around. Instead, the aim is to see how well the two types of night lights data proxy for sub-national GDP, as a measure of local economic activity.

In preparing the data for econometric analysis, we had to deal with 17 *Kabupaten* in 2011, and those same 17 plus three more in 2012, where no light was detected by the DMSP sensors. These areas are in sparsely populated eastern provinces of Indonesia. Two of these *Kabupaten* also had no light detected by VIIRS in 2015, with two more undetected in 2016. We therefore used the inverse-hyperbolic-sine transformation for the lights data, which is identical to using logarithms for the non-zero observations, but also lets us use zeros without resorting to transformations like adding one to all values before logging (Gibson et al., 2017). We use logs of the real GDP values, so that our regression coefficients can be interpreted as elasticities (noting that the units of the DMSP data – DN values – are not comparable to the VIIRS radiance units so we need to use unit-free elasticities).

4. Results

There is no statistically significant relationship between DMSP night lights data and real GDP for Indonesia's second level sub-national regions, with an elasticity of -0.059 from the pooled regression that is surrounded by a wide standard error (Table 2). The same result holds in purely cross-sectional year-by-year analyses where the elasticities are -0.081 and -0.039. The R^2 of these relationships is, at best, 0.01.

In contrast, VIIRS data give precisely estimated elasticities of between 0.17 and 0.19 when using night lights to predict real regional GDP. The standard error of the elasticities is 0.08 in the separate year regressions and 0.05 in the pooled regression and these regressions have R^2 values of 0.05. The variation is dominated by between-unit rather than within-unit components (the within R^2 is zero) with statistically insignificant coefficients on the year 2 dummies in the pooled regressions.

When the regressions are estimated separately for the *Kabupaten* (which covers the rural sector and some towns) and the *Kota* (which covers cities) it shows that the prior results aggregate over very different relationships. Lower density regions administered as *Kabupaten* have real GDP negatively (and statistically significantly) related to DMSP night lights, with an elasticity of -0.11. In contrast, the elasticity of city GDP with respect to DMSP night lights is 0.94 (in the pooled regression or 0.86 and 1.02 in the year-by-year regressions). This gap between elasticities of -0.11 and 0.94 reflects sectoral differences in types of economic activity,

[9] This research design plays to the strength of night lights, which are far better at predicting economic variables in the cross-section, than at predicting time-series changes in these variables (Nordhaus and Chen, 2015; Chen and Nordhaus, 2019; Goldblatt et al., 2020; Gibson and Boe-Gibson, 2020).

[10] We use the administrative geography from 2010, and where spatial units subsequently split we re-aggregate them to have a temporally consistent set of 497 spatial units.

[11] We show below, in Section 4.2 and Section 5, that the predictive performance of night lights data is worse for lower level spatial units, for lower density areas, and for smaller areas. It is especially predictions with DMSP data that get weaker when working with these lower level, smaller, and less dense spatial units.

Table 2
The Predictive Power of Night Lights for Regional GDP is much higher with VIIRS than with DMSP and is much higher for Cities.

	DMSP 'stable lights' for 2011 and 2012			VIIRS Annual Composites for 2015 and 2016		
	All spatial units (Urban & rural)	*Kabupaten* (Mainly rural)	*Kota* (Cities)	All spatial units (Urban & rural)	*Kabupaten* (Mainly rural)	*Kota* (Cities)
Pooled regressions						
Log (sum of lights)$_{it}$	−0.059	−0.107***	0.939***	0.179***	0.086	0.936***
	(0.041)	(0.039)	(0.121)	(0.054)	(0.056)	(0.056)
Year 2 dummy	−0.008	−0.005	−0.123	0.056	0.045	0.056
	(0.099)	(0.109)	(0.174)	(0.092)	(0.105)	(0.119)
Constant	2.786	3.125	−5.734	1.041	1.704	−5.290
	(0.367)	(0.355)	(1.136)	(0.443)	(0.444)	(0.502)
R^2 overall	**0.01**	**0.03**	**0.36**	**0.05**	**0.01**	**0.68**
R^2 within	0.00	0.00	0.00	0.00	0.00	0.00
R^2 between	0.01	0.03	0.38	0.05	0.01	0.69
Number of observations	994	798	196	994	798	196
Year-by-Year Regressions						
Log (sum of lights)$_{it=1}$	−0.081	−0.130**	0.858***	0.185**	0.093	0.945***
	(0.057)	(0.054)	(0.193)	(0.075)	(0.078)	(0.076)
Constant	2.969	3.310	−5.028	0.992	1.658	−5.366
	(0.504)	(0.482)	(1.779)	(0.603)	(0.617)	(0.686)
R^2	**0.01**	**0.04**	**0.30**	**0.05**	**0.01**	**0.69**
Log (sum of lights)$_{it=2}$	−0.039	−0.086	1.024***	0.172**	0.080	0.927***
	(0.059)	(0.057)	(0.139)	(0.077)	(0.080)	(0.082)
Constant	2.607	2.947	−6.618	1.144	1.794	−5.158
	(0.525)	(0.509)	(1.291)	(0.610)	(0.623)	(0.731)
R^2	**0.00**	**0.02**	**0.43**	**0.05**	**0.01**	**0.68**
Number of observations	497	399	98	497	399	98

Notes: The dependent variable is log real GDP for the *Kabupaten* or *Kota* (in 2010 prices and using the administrative divisions from 2010 to account for subsequent splitting of spatial units). Robust standard errors in ( ), ***, **, and * denote statistical significance at 1%, 5% and 10% levels.

in population density and in types of lights used. It is also seen in the differences in the degree of predictive fit, with the between-R^2 values (that greatly exceed the within-R^2, supporting the idea that lights better proxy for economic activity in the cross-section than for short-term changes) being much higher for the urban *Kota* than for the more rural *Kabupaten*.

If the VIIRS data are used to predict GDP, separately for *Kabupaten* and *Kota*, the elasticities are all positive, unlike for DMSP. However, for the *Kabupaten*, the elasticities are only about 0.08–0.09, and are imprecisely estimated. In contrast, for the urban *Kota*, the elasticities are from 0.93 to 0.95, and very precisely estimated. Thus, satellite observation of night lights does not seem to be a suitable source of data to proxy for GDP in non-urban areas in a country like Indonesia, even though it is highly predictive for cities. Another contrast in Table 2 concerns noise in the lights-GDP relationship for the urban sector. The unexplained share of the variation in real GDP is twice as large when using DMSP to predict urban GDP, at 64%, compared to using VIIRS where it is only 32% (for the overall R^2 values in the panel regressions). In the year-by-year regressions for GDP, the predictive power of VIIRS data is 60–130% higher than the predictive power of DMSP data.[12]

4.1. Evidence from an overlapping year

The regressions in Table 2 use annual composites of night lights that were cleaned and processed by NOAA scientists. This allowed like-with-like comparisons in terms of data that were equally cleaned but did not allow overlap in time; DMSP annual composites stop in 2013 and VIIRS annual composites start in 2015. Yet even without the facilities of NOAA scientists, we can partially process monthly VIIRS data to get annual estimates for 2013 so as to predict GDP in a year with DMSP data also available. To remove ephemeral sources of light from the monthly VIIRS data for 2013 we use the 2015 cleaned annual composite as a background noise mask. In other words, if a pixel is recorded as having light in a monthly record from 2013 but shows up as unlit area in 2015 after the NOAA scientists had processed the image to create a cleaned annual composite, then we assume that the light in the monthly file must have been ephemeral and so we ignore it.[13] After doing this partial processing of the VIIRS monthly data, and then summing over months in 2013 and estimating regressions like those in Table 2, we get an elasticity of 0.122 (SE = 0.071) with an R^2 of 0.0234. The elasticity is attenuated, compared to the 0.17 or 0.19 estimated from cleaned annual VIIRS composites in Table 2, as is to be expected if random measurement errors affect our VIIRS estimate for 2013. Yet even with these errors, the VIIRS estimate for 2013 better predicts regional GDP than does the DMSP data (that was cleaned by NOAA scientists), which gives an elasticity of 0.004 (SE = 0.061) with an R^2 of 0.00004. Thus, the higher predictive power of VIIRS data compared to DMSP data seen in Table 2 persists when the analysis is restricted to a year with both types of data available.

4.2. Effects of aggregation and population density

Night lights data are used by economists to examine issues in low-income countries at various levels of aggregation. Early studies following Henderson et al. (2012) mainly focused on the national level or the first sub-national level (provinces, regions or states) but recent studies use DMSP data at much finer scales. For example, Amare et al.

[12] Likewise, for the first sub-national level in the United States, VIIRS data has 60% higher predictive power than DMSP data, while for the second sub-national level in Europe (NUTS2 regions), VIIRS has 80% higher predictive power, for panel and cross-sectional GDP regressions (Gibson and Boe-Gibson, 2020; Gibson, 2020).

[13] Evidence on the effectiveness of this masking procedure is provided by Gibson (2020). An early example of masking VIIRS data to remove effects of ephemeral lighting sources and other noise is Shi et al. (2014), however they rely on the DMSP image from 2012 to mask the unlit area and so may be affected by the blurring of DMSP images that causes lit area to be overstated (Abrahams et al., 2018) and so unlit area will be too small.

Table 3
DMSP data on night lights considerably understate spatial inequality, even after Pareto adjustment for top-coding.

	Gini coefficient			Theil Index		
	All spatial units (Urban & rural)	*Kabupaten* (Mainly rural)	*Kota* (Cities)	All spatial units (Urban & rural)	*Kabupaten* (Mainly rural)	*Kota* (Cities)
DMSP (2011–12)	0.798	0.777	0.575	1.424	1.276	0.565
	(0.017)	(0.016)	(0.059)	(0.089)	(0.076)	(0.123)
DMSP (Pareto-adjusted	0.809	0.781	0.640	1.542	1.305	0.753
for top-coding)	(0.016)	(0.015)	(0.044)	(0.092)	(0.077)	(0.120)
VIIRS (2015–16)	0.860	0.803	0.705	2.185	1.527	0.929
	(0.014)	(0.019)	(0.041)	(0.129)	(0.127)	(0.134)
GDP (2011–12)	0.888	0.874	0.814	2.322	2.125	1.512
	(0.013)	(0.017)	(0.038)	(0.183)	(0.204)	(0.247)
GDP (2015–16)	0.886	0.873	0.833	2.272	2.028	1.749
	(0.013)	(0.012)	(0.038)	(0.183)	(0.131)	(0.269)
Number of observations	497	399	98	497	399	98

Notes: Standard errors in ( ). Inequality statistics are based on the share of total lights (or total GDP) and of total area from each spatial unit.

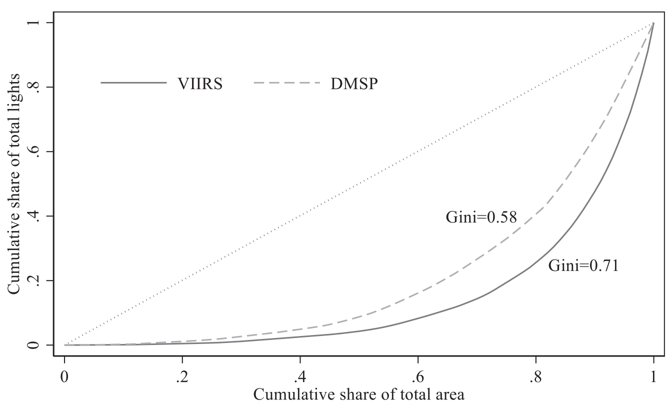

Fig. 1. Lorenz curves for lights: Urban areas, Indonesia.

(2020) use pixel-level DMSP data for the reported latitude and longitude coordinates of each enumeration area in the Demographic and Health Survey for Nigeria, when studying effects of urbanization on child nutrition. At similarly fine scale, Lee (2018) uses DMSP values for a 1.9 km × 1.4 km grid over North Korea to examine local economic impacts of sanctions. A pertinent study for our analysis is Heger and Neumayer (2019) who use sub-district data from Aceh, Indonesia, to examine effects of the Boxing Day tsunami and reconstruction aid on economic activity. The median area of these sub-districts is just 133 km^2 which is just 1/15th the median area of the spatial units used in the Table 2 results. With this use of night lights data for such small spatial units it is worth assessing how the DMSP data and the VIIRS data perform in terms of predicting GDP at different levels of spatial aggregation.

We aggregated real GDP and the sum of DMSP and VIIRS lights to the provincial level (n = 34) by year and re-estimated the pooled regressions from Table 2. The R^2 using either source of night lights data rose to 0.20, with the GDP-VIIRS elasticity rising to 0.50 (SE = 0.12) and the GDP-DMSP elasticity rising to 0.01 (SE = 0.002). Thus, if more spatially aggregated data are used the worse performance of DMSP for predicting economic activity (compared to how VIIRS does) is less apparent. We cannot take the comparison the other way, as there are no GDP data for Indonesia's sub-districts, but we show below, using data from China, that DMSP's predictive performance for lower level and smaller areas is worse than it is for more aggregated units.

A related comparison is to consider lower density and higher density *Kabupaten*, where we split the sample at the median population density of 83.3 persons per km.^{2} In the lower density areas there is a negative relationship between GDP and night lights, with an elasticity of −0.31 (SE = 0.04) using DMSP data and −0.19 (SE = 0.07) using VIIRS data. In contrast, night lights positively relate to GDP for higher density *Kabupaten*; the elasticity is 0.31 (SE = 0.09, R^2 = 0.17) using DMSP data and 0.60 (SE = 0.04, R^2 = 0.62) for the VIIRS data. Thus, neither source of night lights data seems useful as a proxy for GDP in low density rural areas, echoing a conclusion made by Chen and Nordhaus (2011). Notably, 11 of 23 districts in Aceh (and about 45% of the sub-districts) where DMSP data are used to examine impacts of the Boxing Day tsunami are in the low density sub-sample; with a negative relationship between GDP and night lights for such areas, the strength of the inferences that can be drawn from night lights data about impacts on local economic activity may be limited.

4.3. Spatial inequality

In addition to being used as a proxy for local economic activity in regressions, DMSP night lights data are also used to study regional inequality (Gennaioli et al., 2014; Lessmann and Seidel, 2017; Singhal et al., 2020). In this regard, it is relevant that the DMSP data may considerably understate spatial inequality. In Table 3 we report the Gini coefficient and Theil index, estimated over all 497 spatial units, and then separately for *Kabupaten* and *Kota.*[14] The VIIRS data show 53% higher inequality than do the DMSP data (with a Theil index of 2.19 versus 1.42, noting the Theil index is sensitive to inequality at the top of the distribution). The difference comes especially from urban areas, where VIIRS data show 64% higher inequality than what DMSP data show. We expand upon this comparison in Fig. 1, showing Lorenz curves for lights in urban areas calculated using DMSP data and using VIIRS data. With DMSP data, the Lorenz curve is significantly closer to the line of equality at all points, and yields a Gini coefficient of 0.58 compared to the Gini of 0.71 using VIIRS (the difference is statistically significant).

We can rule out rising spatial inequality between 2011 and 2016 as causing the gap between what the DMSP data and what the VIIRS data show because the GDP data show no evidence of any increase in inequality (Table 3). Specifically, the Theil index for GDP fell slightly, by 2% from 2011–12 to 2015–16, with no change in the Gini. Notably, spatial inequality estimates using GDP data are somewhat higher than estimates with VIIRS data (and very much higher than inequality estimates with the DMSP data), especially when disaggregating into *Kabupaten* and *Kota*. Thus, using DMSP data causes the inequality estimates to move further away from what the GDP data show, justifying our claim that using DMSP data would cause one to understate spatial inequality.

[14] For the analysis of spatial inequality we combine the DMSP data for 2011 and 2012. Likewise, the VIIRS data for 2015 and 2016 are combined. This should give more precise estimates of cross-sectional variation.

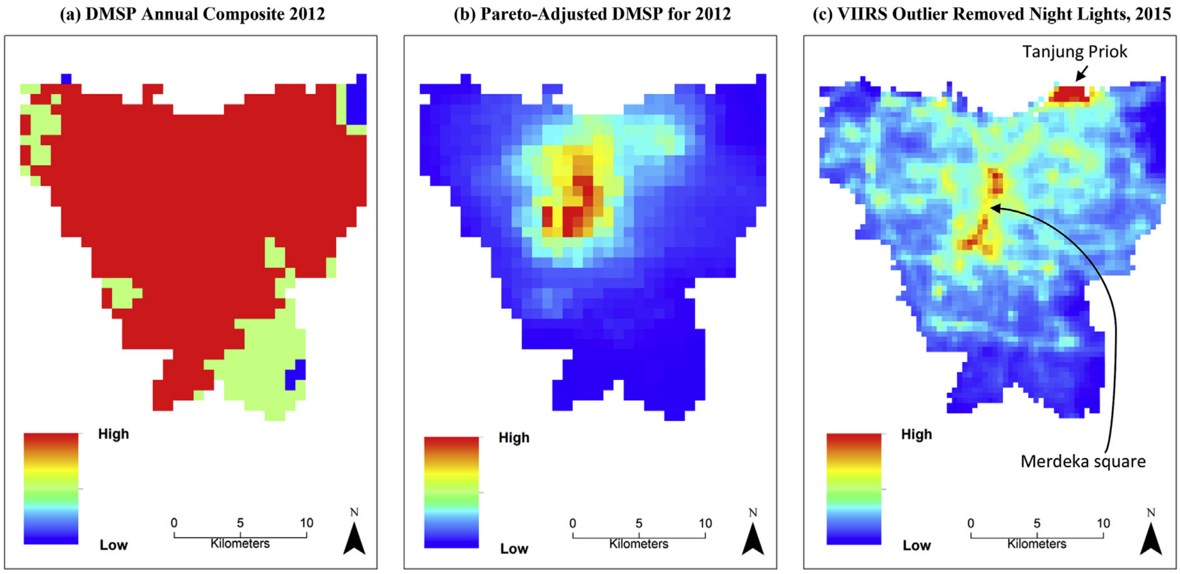

Fig. 2. Intra-Jakarta Heterogeneity in Night Lights is Obscured When Using DMSP Data, Even With Pareto Adjustment for Top-Coding.

Such understatement should be expected, given the mean-reverting measurement errors in DMSP data (Gibson, 2020).

There are at least two flaws in DMSP data that contribute to cause spatial inequality to be understated. Blurring will tend to spread measured lights by attributing light to some unlit places. The top-coding of DN values at 63 will also dampen inequality, especially at the top of the distribution. In order to which see flaw matters most, we use the Pareto-adjusted DMSP values for 2011 and 2012 developed by Bluhm and Krause (2018) and made available at their website: http://lightinequality.com/top-lights.html. These adjusted data are designed to correct for top-coding so if they give inequality estimates close to what VIIRS shows it would suggest that top-coding is the main problem.

After re-estimating Table 3 inequality statistics we find that the Pareto-adjustment closes just one-sixth of the gap in overall inequality between the DMSP and VIIRS results (e.g., raising the Theil index from 1.42 to 1.54, compared to the value of 2.19 with VIIRS). This result implies that blurring is the more important problem. The Pareto adjustment does better for urban areas, where it closes about one-half of the gap. While this is promising, it remains true that inequality is substantially understated, even when using the Pareto-adjusted DMSP data. Thus, use of DMSP data to estimate spatial inequality (e.g. Lessmann and Seidel, 2017; Bluhm and Krause, 2018) likely understates the true extent of the spatial concentration of economic activity, which we also show below at a finer scale.

4.4. Using night lights data to study intra-city variation

The inequality described in Fig. 1 and Table 3 considers the urban sector as a whole, but night lights also are used to study *intra*-city differences (e.g. Kocornik-Mina et al., 2020). The top-coding and limited dynamic range of DMSP may distort understanding of spatial patterns in the development of particular cities by disguising differentiation that occurs in certain places. We illustrate this effect in Fig. 2, which maps night lights in Jakarta (restricting attention to the area within the administrative boundaries of the city). The DMSP stable lights for 2012 in panel (a) show very little intra-city heterogeneity; 82% of the pixels have a digital number (DN) value of 63 (the highest possible) and 17% have DN = 62. With such data, one cannot see where within a city the lights are brightest, and by treating all areas as equally bright the DN values are almost like a dummy variable for whether a pixel is part of the city or not. This limited variation may explain a finding of Gibson et al. (2017) that decomposing the sum of lights from cities into the extensive margin (the lit area) and intensive margin (the brightness of the lit area) shows that only growth on the extensive margin predicts their outcome of interest (poverty rates in rural India). With the DMSP data being so bad at showing brightness variations within cities, a more appropriate research design when using such data may be to use simpler indicators, like binary variables for whether pixels are lit brightly enough or not to be considered part of the city. Indeed, this is how some of the remote-sensing literature (e.g. Small et al., 2005) tends to use DMSP data.

The map in panel (b) of Fig. 2 shows the intra-city patterns in the Pareto-adjusted DMSP data for 2012. With this adjustment, the most brightly-lit core of the city appears as an approximately rectangular area, running about 8 km in a north-south direction and 6 km in an east-west direction, located slightly left of centre, with a moderately lit area to the northeast (going towards Jakarta Bay). The remainder of the city appears as a largely undifferentiated area with lower levels of light recorded.

While Pareto-adjusted data roughly locate the position of the CBD they miss major features of Jakarta, as seen by comparing with the map in panel (c) from the more accurate VIIRS images. First, pixels in the CBD form less of a block because they are interrupted by Merdeka square (one of the largest public squares in the world, at five times larger than Tiananmen square in Beijing) and surrounding parkland. Second, the brightly lit axis of the CBD is smaller and runs more northeasterly towards the port of Tanjung Priok. The third, and most concerning, feature the Pareto-adjusted data miss is the brightness of the port. This is the 22nd busiest in the world (ranked ahead of the port of New York/New Jersey), handling two-thirds of Indonesia's international goods trade. Yet despite this importance, and the fact of the port covering over 900 ha (three times the size of Central Park in Manhattan) and being very bright lit, it is entirely missed in the Pareto-adjusted DMSP images, which do not show any differentiation from the surrounding area.[15] Finally, the Pareto-adjusted DMSP data miss spots of brightly lit areas in east Jakarta, and spots near the northwestern edge of the city, on the way to the Soekarno-Hatta airport (that is just outside the city administrative boundary). Thus, while the Pareto-adjusted images improve upon the usual DMSP data that portray most of the city as an undifferentiated blob of light, they still miss much of the detail and therefore provide a poor guide to patterns of intra-city spatial heterogeneity.

In addition to the Pareto-adjusted data missing key spatial features,

[15] The radiance-calibrated DMSP data for 2010 (the latest year) also cannot distinguish the port from surrounding areas.

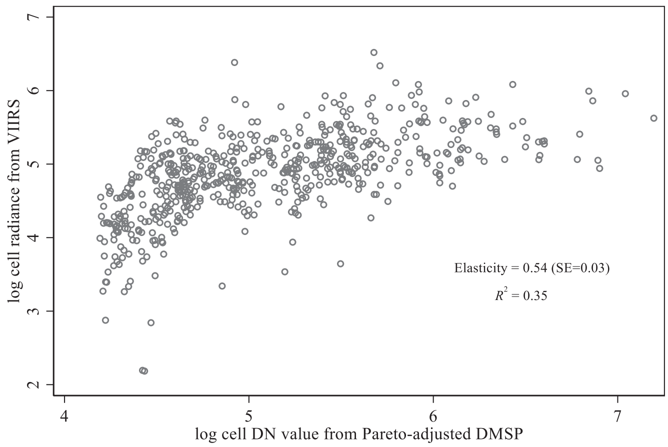

Fig. 3. Pareto-adjusted DMSP lights and VIIRS radiance: Jakarta.

the values used to replace top-coded DN values of 63 may overstate brightness differences. To allow for a quantitative analysis to supplement the visual comparison offered in Fig. 2, we laid a grid of 590 cells over the maps in Fig. 2 and calculated cell-level statistics so that we could compare reports of light coming from the same small areas.[16] The grid-level coefficient of variation (CoV) of the Pareto-adjusted DMSP data is 60% above the CoV for VIIRS data, suggesting that the adjustment introduces more variability – in the sense of a wider range of values – than found in actual measures of radiance. A fairly poor correspondence between the Pareto-adjusted data and the VIIRS radiance measures also shows up in Fig. 3, which provides a scatter plot of the radiance for each cell, from the VIIRS data, against the DN values for the same cell from the Pareto-adjusted DMSP data. The R^2 for this relationship is only 0.35 and the elasticity of cell radiance with respect to cell DN values from the Pareto-adjusted data is only 0.54. While this is a better fit than using the original DMSP data to predict radiance, which has an elasticity of 0.30 and a R^2 of just 0.03, it is still true that there will be a lot of error if Pareto-adjusted DMSP data are used to proxy for actual radiance (or to proxy for the underlying differences in economic density that produce the spatial patterns in radiance).

5. Evidence from other countries

We have emphasized evidence from Indonesia because it is one of the few developing countries with official GDP data at the second sub-national level.[17] It is also an inherently interesting and important country. The main patterns shown in Section 4 are also found for second level sub-national regions in the EU (Gibson, 2020) but evidence from developed countries may be less relevant to development economists interested in using night lights data. An important question for external validity, therefore, is whether evidence from other developing countries can corroborate the results from Indonesia. To answer this question we obtained data from China and South Africa, although for both of these countries the nature of the data available provides less convincing evidence than that from Indonesia.

5.1. China

We obtained GDP data at the second sub-national level for China from the 2013 and 2014 editions of the *City Statistical Yearbook* (National Bureau of Statistics, 2013-2014), covering calendar years 2012 and 2013. The GDP data are broken into two sub-aggregates, for the primary sector (agriculture, forestry, fishing) and the combined secondary and tertiary sector for industry and services. We obtained GDP data for 288 prefectures, which is all prefectures in China in 2013 except two: Haidong in Qinghai province was newly upgraded from region status in 2012 (and as it was not a prefecture in 2012 there are no GDP data for it that year in the *Yearbook*), and we also drop Sansha prefecture, which is an island city of about 1500 people in the South China Sea whose prefectural status likely reflects military and strategic objectives rather than either its economic significance or demographic importance.

An important difference between China and Indonesia in data availability related to sub-national spatial organization limits the extent that results for China may directly inform about the external validity of the Indonesian results. For Indonesia, the second sub-national level, comprised of either *Kabupaten* or *Kota*, covers the entire country, with no other types of spatial units at the second level. In contrast, while China's general administrative hierarchy is Province, then Prefecture, then county for more rural areas or district for more urban areas, there are important exceptions. Large parts of western China, Inner Mongolia and a few other scattered areas have no Prefectures and instead spatial units are either Leagues, Autonomous regions or Provincially Administered areas. The areas that we cover, that are organized into prefectures, contained 93.3% of China's resident population in the 2010 census (1.25 billion people) but covered only 50.1% of the land area.[18] In other words, China's disaggregated GDP data available to us systematically exclude areas with the lowest population densities.

If we were to apply a similar exclusion to the Indonesian data, removing spatial units with the lowest density that contain 6.7% of the population, a far better performance of night lights in predicting GDP and a smaller gap between DMSP and VIIRS in predictive power is apparent. Specifically, the GDP-lights elasticity with DMSP data rises from −0.06 to 0.44 and the R^2 becomes 0.24, while for VIIRS the elasticity rises from 0.18 to 0.67 and the R^2 is 0.61. The dramatic effect from trimming these low density spatial units is another demonstration of the general unsuitability of satellite detected luminosity as a proxy for economic activity in sparsely populated rural areas of developing countries. Moreover, the spatial units we use for China are naturally more aggregated than those used for Indonesia—1.25 billion people are divided into 288 areas while after Indonesia's low density regions are removed a population of just 0.22 billion is divided amongst 386 areas. The results in Section 4.2 suggest that the gap between DMSP data and VIIRS data in terms of predictive performance gets smaller when working with more aggregated spatial units so our expectation is that the results for China will show smaller differences when VIIRS data are used compared to when the DMSP data are used, compared to the differences that are seen when using data from Indonesia.

In Table 4 we report results of regressions predicting GDP for prefectures in 2012 and 2013. The DMSP data used are the cleaned annual composites from NOAA. The VIIRS data are annual estimates for these years that we formed by using the masked monthly data, in the same way that we used data for an overlapping year for Indonesia in Section 4.1. The predictive power when using the VIIRS data is about 25% higher, with an R^2 of 0.76 compared to 0.62 if DMSP data are used. If we aggregate to the province level ($n = 31$), predictive power rises, especially when using the DMSP data, which provides an R^2 of 0.86,

[16] The grid is 30 × 30 but 310 cells fall outside Jakarta's boundaries, given that the city is not perfectly square.

[17] As an example of the rarity of sub-national developing country GDP data, Gennaioli et al. (2014) assembled a database of regional GDP data at the most disaggregated administrative level available for as many countries as possible. Amongst African countries, 45 had no data and even for those with data it was at very aggregated levels; for example, for both Nigeria and South Africa the GDP data were only for four broad regions.

[18] The census data on the resident population of China's sub-national areas are from NBS (2012). Given that the non-prefectural areas are not 'cities' (in the sense of being organized into prefectures with one or more urban districts) they are not covered by the *City Statistical Yearbook*.

Table 4
Night lights and GDP at the second sub-national level: Prefectures in China.

	DMSP 'stable lights' for 2012 and 2013			VIIRS Annual Estimates for 2012 and 2013		
	Prefectures	Merged counties	Merged districts	Prefectures	Merged counties	Merged districts
Log (sum of lights)$_{it}$	0.938***	0.820***	1.136***	0.845***	0.770***	0.960***
	(0.034)	(0.033)	(0.029)	(0.020)	(0.027)	(0.018)
Year 2 dummy	0.066	0.049	0.090	−0.003	−0.016	−0.011
	(0.046)	(0.050)	(0.051)	(0.036)	(0.050)	(0.041)
Constant	−5.778	−4.702	−7.768	−5.768	−4.855	−7.467
	(0.384)	(0.368)	(0.305)	(0.250)	(0.327)	(0.223)
R^2	0.618	0.604	0.712	0.764	0.621	0.815

Notes: The dependent variables are 2012 and 2013 log GDP for each prefecture (with breakdowns into merged counties and into merged districts). The prefectures that are considered ($n = 288$) contained 93.3% of China's resident population in 2010 and covered 50.1% of the land area. Robust standard errors in ( ), ***, **, and * denote statistical significance at 1%, 5% and 10% levels.

compared to $R^2 = 0.90$ if VIIRS data are used. We show in Section 5.2 below that going in the opposite direction, by disaggregating to the third sub-national level (counties and districts) increases the gap in the predictive performance of DMSP and VIIRS. In other words, it is particularly when trying to proxy for local economic activity that the flaws in the DMSP data come to matter, while the flaws are less apparent when working with more aggregated spatial units.

The prefectures used for the results in Table 4 can be further divided into two types of third-level units—districts (*shiqu*) that contain the urban core of the prefecture, and counties that cover the remaining area. While there may be several counties within a prefecture, and a big city like Shanghai may have up to 15 districts, the *City Statistical Yearbook* reports GDP data for an aggregate of all districts within a prefecture, and for an aggregate of all counties within a prefecture. These two sub-prefectural aggregates are more alike than are the *Kota* and *Kabupaten* in Indonesia; for example, the mean population density of the merged districts is only three times as high as that of the merged counties, while the mean density of the *Kota* was 14-times as high as the mean density of the *Kabupaten*. Indeed, the top quintile of the merged counties has a higher population density than the bottom two quintiles of the merged districts, which is an overlap between what is, nominally, urban and rural areas that is much greater than seen in many other countries.[19]

Notwithstanding this caveat about overlapping population densities, we observe the same pattern as in Indonesia, of night lights data better predicting GDP in urban areas than in rural areas. The results in Table 4 show that the R^2 for the regressions using merged districts are up to 30% higher than the regressions for merged counties. A related comparison uses the sector-disaggregated GDP data; with VIIRS data the R^2 for predicting primary sector GDP is only 0.09, while it is 0.78 for predicting industrial and services sector GDP. Thus, it appears that night lights data are a less useful proxy for rural agricultural economic activity.

Another pattern seen for Indonesia that is repeated with the data from China is the understatement of spatial inequality when using DMSP data. We focus on 2013, as that year had a full 12 months of VIIRS data while there were only nine months of data for 2012. The Theil Index and Gini coefficient for GDP in that year are, respectively, 1.18 (SE = 0.09) and 0.72 (SE = 0.02). These are almost exactly the same as the inequality estimates derived from the VIIRS data, with a Theil Index of 1.24 (SE = 0.09) and a Gini of 0.74 (SE = 0.02). However, spatial inequality would be understated by almost one-half if the DMSP data are used, with a Theil Index of 0.62 (SE = 0.05) and a Gini of 0.59 (SE = 0.02).

5.2. Chongqing

Chongqing is a centrally controlled municipality in China, with status equivalent to a province, and thus it is able to issue its own statistical yearbooks. Helpfully for this study, the Chongqing yearbooks report GDP for each individual county and district, and so provide a benchmark for assessing night lights data at the third sub-national level.[20] For other provinces and the other three centrally controlled municipalities–Beijing, Shanghai and Tianjin–there is just an aggregated GDP for all counties and for all districts in each prefecture, which were the data used in Section 5.1. Chongqing is also far larger than the other centrally controlled municipalities, with an area of 82,300 km^2, which is equivalent to Austria, or the US State of Maine. Moreover, the GDP data for the 38 individual counties and districts in Chongqing are available through 2015 so we can benchmark against the VIIRS cleaned annual composite for that year in order to test the validity of the procedure applied to VIIRS monthly data when forming estimates for 2012 and 2013 that overlap in time with the DMSP data.

If we use the VIIRS annual composite for 2015 as a mask to filter ephemeral lights and other background noise from the monthly data (as used in Sections 4.1 and 5.1) we can better match the annual composites created by NOAA scientists. In Fig. 4 we present two scatter plots, where the y-axis in each case is the log of the sum of lights for each county or district in Chongqing using the VIIRS 2015 annual composite from NOAA. In the scatter plot on the left, the x-axis is the average of VIIRS monthly lights in 2015, without filtering or masking. Using the unmasked monthly data to approximate what a cleaned annual composite would look like (where an approximation is needed in most years because, to date, just two years have cleaned annual composites) gives an R^2 of only 0.86, with a relationship well off the 45-degree line (the elasticity is 1.52 (SE = 0.12)). In contrast, the masked monthly data used in the scatter plot on the right closely match the cleaned annual composite; the R^2 is 0.98 and the elasticity is 1.05 (SE = 0.02). This benefit of using masked data carries through to predicting GDP in 2015; using either NOAA's cleaned annual composite or our annual estimate formed from masked monthly data, GDP for the counties and districts is predicted with an R^2 of 0.79. However, if we use the unmasked monthly VIIRS data in a regression to predict GDP the R^2 is just 0.64 (and the estimated GDP-lights elasticity is eight standard errors away from the elasticity based on the annual composite).

In Table 5 we report results of regressions predicting GDP for counties and districts in Chongqing in 2012 and 2013. The DMSP data used are the cleaned annual composites from NOAA while the VIIRS data are annual estimates we formed by using the masked monthly data. The predictive power for total GDP is 53% higher with the VIIRS data, with an R^2 of 0.81 compared to an R^2 of 0.53 with DMSP data. This is a larger gap than was seen when using the more spatially aggregated data in Section 5.1. We also report regressions for breakdowns of GDP into the primary sector, which is largely agriculture, and the combined industry and services sector. There is no relationship between primary sector GDP and VIIRS

[19] Some of this overlap reflects counties that have only recently been upgraded to urban district status, and so their characteristics, including population density, are still largely rural.

[20] Our analyses in this section account for merging into new districts and upgrading of counties to districts when we form a temporally consistent set of spatial units. The data are from NBS (2013–2016).

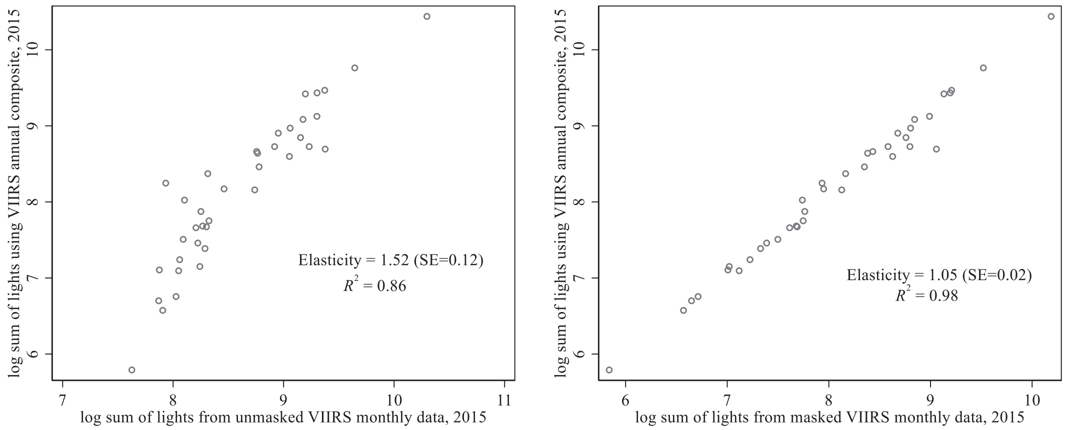

Fig. 4. Comparison of predictive performance of unmasked and masked VIIRS monthly data: Counties and districts of chongqing.

Table 5
Night lights and GDP at the third sub-national level: Counties and districts in Chongqing, China.

	DMSP 'stable lights' for 2012 and 2013			VIIRS Annual Estimates for 2012 and 2013		
	Total GDP	Primary sector GDP	Industry & services GDP	Total GDP	Primary sector GDP	Industry & services GDP
Log (sum of lights)$_{it}$	0.870***	0.545***	0.927***	0.756***	−0.064	0.824***
	(0.164)	(0.180)	(0.178)	(0.032)	(0.115)	(0.034)
Year 2 dummy	0.048	0.026	0.049	−0.100	0.078	−0.114
	(0.127)	(0.212)	(0.140)	(0.082)	(0.230)	(0.086)
Constant	−4.749	−4.302	−5.400	−4.702	1.267	−5.540
	(1.520)	(1.642)	(1.652)	(0.316)	(1.143)	(0.344)
R^2	0.528	0.134	0.512	0.809	0.005	0.820

Notes: The dependent variables are 2012 and 2013 log GDP (with breakdowns into primary sector GDP and secondary and tertiary sector GDP) for each county or district within Chongqing ($n = 76$). Robust standard errors in ( ), ***, **, and * denote statistical significance at 1%, 5% and 10% levels.

night lights, while the DMSP lights relationship with primary sector GDP is only one-quarter as strong as the relationship with industrial and services sector GDP. This pattern found with disaggregated spatial units is consistent with what Keola et al. (2015) find at the country level, and implies that luminosity data are unlikely to be a good proxy for economic activity in places where agriculture is a large share of the local economy.

A further disaggregation of the results for Chongqing, not reported in Table 5, was to split the sample (at the median) according to the area of each spatial unit. The DMSP data are far worse at predicting GDP for small spatial units than for larger ones; for the small units the GDP-lights elasticity is 0.39 (SE = 0.26, $R^2 = 0.18$) while larger units have a far higher elasticity of 1.15 (SE = 0.10, $R^2 = 0.82$). The VIIRS data are also better at predicting the GDP of larger spatial units but the difference is less dramatic; the GDP-lights elasticity is 0.58 (SE = 0.04, $R^2 = 0.66$) for smaller areas and 0.92 (SE = 0.07, $R^2 = 0.86$) for larger areas. Thus, in addition to problems in predicting GDP for low density rural areas, as seen in the results for Indonesia, there may also be problems in using DMSP data to predict GDP for small areas. A possible reason for this is blurring of DMSP images, which creates proportionately larger errors the smaller is the area being studied (Gibson et al., 2020).

Finally, if luminosity data are used to estimate spatial inequality in Chongqing there is a repeat of the pattern from Section 4.3 for Indonesia, of DMSP data understating inequality. On average, the DMSP data show only two-thirds as much inequality as what the VIIRS data show. The understated spatial inequality is especially for districts, which are meant to be the urban core for each prefecture or municipality. For example, the Theil Index for the districts is 0.79, using either GDP or VIIRS night lights data, but it is only 0.42 with DMSP data.

5.3. South Africa

For South Africa we obtained data on total employment in 2013 and GDP in 2011 at the local municipality level ($n = 234$). These are model-based estimates from Statistics South Africa (Morudu, 2014). Specifically, the statistics office used data from the Financial Census of Municipalities and the Census of Population to disaggregate provincial employment and GDP data, based on the assumption that the municipality level data follow a rank-size rule. Validation of the model-based estimates suggested a good fit, except for the lower tail of very small municipalities (e.g. with total employment below ca. 6000 persons) which were smaller than the fitted distribution. These municipality level GDP estimates have been used by the widely respected Southern Africa Labour and Development Research Unit (SALDRU) for estimating municipality level poverty (David et al., 2018), attesting to their usefulness. Nevertheless, we emphasize that these are model-based estimates and so this evidence is less convincing than the Indonesian evidence which uses conventionally measured GDP data.

When DMSP stable lights and VIIRS annual estimates (both for 2013, with the VIIRS estimates based on using the 2015 annual composite as a mask to remove outliers) are used to predict municipality level GDP, the VIIRS data provide about 15% higher predictive power (Table 6). The standard errors surrounding the DMSP elasticities are about one-third higher than the VIIRS standard errors. The patterns are similar but the magnitudes of differences are slightly smaller for predicting total employment of each municipality. The predictive fit in GDP regressions for both DMSP and VIIRS data is about twice as high in high density areas as it is in low density areas. However, in contrast to the Indonesian results, we do not see any negative relationships between GDP and lights.

In terms of inequality, the VIIRS data indicate that inequality amongst

Table 6
Comparison of DMSP and VIIRS for predicting municipality level total employment and GDP and for estimating inequality: Local municipalities in South Africa.

	DMSP 'stable lights' for 2013			VIIRS Annual Estimates for 2013		
	All spatial units	Low density	High density	All spatial units	Low density	High density
Total employment, 2013	0.870	0.748	0.833	0.800	0.692	0.779
	(0.040)	(0.047)	(0.053)	(0.027)	(0.041)	(0.034)
R^2 (for employment regressions)	0.793	0.699	0.783	0.852	0.694	0.882
Municipality GDP, 2011	1.189	0.809	1.491	1.129	0.791	1.411
	(0.066)	(0.081)	(0.084)	(0.050)	(0.073)	(0.057)
R^2 (for GDP regressions)	0.584	0.381	0.679	0.660	0.423	0.783
Gini coefficient	0.527	0.338	0.468	0.608	0.380	0.549
	(0.047)	(0.020)	(0.052)	(0.063)	(0.022)	(0.065)
Theil index	0.486	0.183	0.386	0.696	0.239	0.565
	(0.080)	(0.021)	(0.073)	(0.134)	(0.026)	(0.119)
Number of observations	234	117	117	234	117	117

Notes: VIIRS annual estimates for 2013 based on the sum of masked monthly composites, using the outlier-removed 2015 annual composite as a background noise mask. The split of the sample into low density and high density municipalities is based on the median population density of 28.2 persons per km^2. All regressions use double-log specifications. Robust standard errors in ( ) with standard errors on inequality statistics based on 1000 bootstrapped replications.

the full set of municipalities is 43% higher (for the Theil index, or 15% higher for the Gini) than what the DMSP data suggests. The inequality indices are somewhat higher using the GDP estimates (e.g. the Gini is 0.66 for GDP, 0.61 for VIIRS night lights and 0.53 for DMSP night lights) and so the gap between estimates from DMSP and those from VIIRS indicates that the DMSP data are understating spatial inequality, which was the same pattern observed for Indonesia. The understatement of inequality when using DMSP data is more apparent in higher density areas, which is also a pattern seen in Indonesia.

6. Conclusions

In terms of the first question in our title, VIIRS night lights data are a better proxy for local economic activity than are the more widely used DMSP data. For the second question, of where night lights data should be used, neither DMSP data nor VIIRS data seem to provide a good proxy for economic activity in low density rural areas. Such areas are of natural interest to many development economists, and so our findings suggest that economists should look elsewhere for studying these areas remotely, while also persevering with their traditional survey-based methods of measuring living standards. These other remote sensing sources may include day-time images, such as Landsat, that seem to be better cross-sectional predictors of economic activity than are DMSP data (Goldblatt et al., 2020).[21]

Our results also suggest that early findings in the literature of good performance for DMSP night lights data in predicting GDP may be less appropriate as a justification for some more recent uses of DMSP data. The early studies were mainly using national GDP or data from the first sub-national level but recent uses of DMSP data in developing countries are at much finer scale and sometimes even down to the pixel level. The results reported here show that the predictive performance of DMSP data is worse for lower level spatial units, for lower density areas, and for smaller areas. There is also some decline in the predictive performance of VIIRS data for lower level spatial units but the decline is less marked than with the DMSP data. It seems that it is especially when trying to proxy for local level economic activity in developing countries that flaws in the DMSP sensor and data management became visible, while these flaws are less apparent when working with more aggregated spatial units.

While satellite-detected night lights data are more suitable for studying urban areas, the limitations of the DMSP data in this context are also apparent. Inequality in the urban sector is greatly understated if DMSP data are used, even when a correction method for top-coding has been applied. The inability of the DMSP images to distinguish important infrastructure, such as Jakarta's port, from the surrounding area suggests that the VIIRS data can provide a much better resource for studying spatial patterns of urban development.

There may be some resistance to these conclusions, as much recent applied economics research relies upon DMSP night lights data, and also because the newer and better VIIRS data have a shorter time-series for relating to changes in economic variables. We believe that this second source of resistance to switching to VIIRS data is misplaced because night lights and other remote sensing data are poor predictors of time-series changes in activity, even as they can be good predictors of economic variables cross-sectionally. Thus, assessing which data source is better should be based mainly on their performance in cross-sectional uses, as has been emphasized in this study. Moreover, some of the time-series variation in the DMSP data is noise, due to the lack of calibration and inter-satellite differences, and these data are becoming old due to the 2013 end-date. In contrast, not only does VIIRS provide consistent data over time, it does so with a time-series that will only get longer, with launch of the NOAA-20 satellite in November 2017 that has the identical measuring instruments as on the *Suomi* satellite that hosts VIIRS. It would therefore be an opportune time for economists to follow the lead of other disciplines, and make more use of the VIIRS night lights data.

[21] Whether better predictive performance of Landsat over night lights images would hold if the comparison used the more accurate VIIRS data rather than the noisy DMSP data is an interesting question for future analysis.

Credit author statement

John Gibson: Conceptualization, Estimation, Writing Susan Olivia: Data curation, graphics, editing; Geua Boe-Gibson: GIS and remote sensing, graphics Chao Li: Data curation, editing.

Data availability

We have made a replication dataset and code appendix which will be made available.

References

Abrahams, Alexei, Oram, Christopher, Lozano-Gracia, Nancy, 2018. Deblurring DMSP night-time lights: a new method using Gaussian filters and frequencies of illumination. Rem. Sens. Environ. 210 (1), 242–258.
Amare, Mulubrhan, Arndt, Channing, Abay, Kibrom, Todd, Benson, 2020. Urbanization and child nutritional outcomes. World Bank Econ. Rev. 34 (1), 63–74.
Bickenbach, Frank, Bode, Eckhardt, Nunnenkamp, Peter, Söder, Mareike, 2016. Night lights and regional GDP. Rev. World Econ. 152 (2), 425–447.
Bluhm, Richard, Krause, Melanie, 2018. Top lights - bright cities and their contribution to economic development. CESifo Working Paper No. 7411.
Chen, Xi, Nordhaus, William, 2011. Using luminosity data as a proxy for economic statistics. Proc. Natl. Acad. Sci. Unit. States Am. 108 (21), 8589–8594.
Chen, Xi, Nordhaus, William, 2015. A test of the new VIIRS lights data set: population and economic output in Africa. Rem. Sens. 7 (4), 4937–4947.
Chen, Xi, Nordhaus, William, 2019. VIIRS night time lights in the estimation of cross-sectional and time-series GDP. Rem. Sens. 11 (9), 1057–1068.

David, Anda, Guilbert, Nathalie, Hamaguchi, Nobuaki, Higashi, Yudai, Hino, Hiroyuki, Murray, Leibbrandt, Shifa, Muna, 2018. Spatial Poverty and Inequality in South Africa: A Municipality Level Analysis. Working Paper No. 221. Southern Africa Labour and Development Research Unit.

Doll, Christopher, 2008. CIESIN Thematic Guide to Night-Time Light Remote Sensing and its Applications. Center for International Earth Science Information Network, Columbia University, New York.

Elvidge, Christopher, Baugh, Kimberly, Zhizhin, Mikhail, Hsu, Feng-Chi, 2013. Why VIIRS data are superior to DMSP for mapping night time lights. Proceedings of the Asia-Pacific Advanced Network 35 (1), 62–69.

Gennaioli, Nicola, La Porta, Rafael, Florencio Lopez De Silanes, Shleifer, Andrei, 2014. Growth in regions. J. Econ. Growth 19 (3), 259–309.

Gibson, John, 2020. Better Night Lights Data, for Longer. Oxford Bulletin of Economics and Statistics (forthcoming).

Gibson, John, Boe-Gibson, Geua, 2020. *Three Facts about Night Lights Data. Working Paper* No. 20/03. Department of Economics, University of Waikato.

Gibson, John, Datt, Gaurav, Murgai, Rinku, Ravallion, Martin, 2017. For India's rural poor, growing towns matter more than growing cities. World Dev. 98 (1), 413–429.

Gibson, John, Olivia, Susan, Boe-Gibson, Geua, 2020. Night lights in economics: sources and uses. J. Econ. Surv. 34 (5), 955–980.

Goldblatt, Ran, Heilmann, Kilian, Vaizman, Yonatan, 2020. Can medium-resolution satellite imagery measure economic activity at small geographies? Evidence from Landsat in Vietnam. World Bank Econ. Rev. 34 (3), 635–653.

Heger, Martin Philipp, Neumayer, Eric, 2019. The impact of the Indian Ocean tsunami on Aceh's long-term economic growth. J. Dev. Econ. 141 (1), 102365.

Henderson, Vernon, Adam, Storeygard, Weil, David, 2012. Measuring economic growth from outer space. Am. Econ. Rev. 102 (2), 994–1028.

Hsu, Feng-Chi, Baugh, Kimberly, Ghosh, Tilottama, Zhizhin, Mikhail, Elvidge, Christopher, 2015. DMSP-OLS radiance calibrated night-time lights time series with inter-calibration. Rem. Sens. 7 (2), 1855–1876.

Keola, Souknilanh, Andersson, Magnus, Hall, Ola, 2015. Monitoring economic development from space: using night time light and land cover data to measure economic growth. World Dev. 66 (1), 322–334.

Kocornik-Mina, Adriana, McDermott, Thomas, Michaels, Guy, Rauch, Ferdinand, 2020. Flooded cities. Am. Econ. J. Appl. Econ. 12 (2), 35–66.

Lee, Yong Suk, 2018. International isolation and regional inequality: evidence from sanctions on North Korea. J. Urban Econ. 103 (1), 34–51.

Lessmann, Christian, Seidel, André, 2017. Regional inequality, convergence, and its determinants–A view from outer space. Eur. Econ. Rev. 92 (1), 110–132.

Michalopoulos, Stelios, Papaioannou, Elias, 2014. National institutions and sub-national development in Africa. Q. J. Econ. 129 (1), 151–213.

Morudu, Hlabi, 2014. Simulating Key Estimates for Local Municipality Socio-Economic Development Models: Applications of the Rank-Size Rule from 2007 to 2012. *Discussion Document* No. D9129, Statistics South Africa.

National Bureau of Statistics [NBS], 2013-2014. China City Statistical Yearbook 2013-2014 (Zhongguo Chengshi Tongji Nianjian 2013-2014. China Statistics Press, Beijing.

National Bureau of Statistics [NBS], 2012. Tabulation on the 2010 Population Census of the People's Republic of China by County (*Zhongguo 2010 Nian Renkou Pucha Fenxian Ziliao*. China Statistics Press, Beijing.

National Bureau of Statistics [NBS], 2013-2016. Chongqing Statistical Yearbook 2013-2016 (*Chongqing Tongji Nianjian 2013-2016*. China Statistics Press, Beijing.

Nordhaus, William, Chen, Xi, 2015. A sharper image? Estimates of the precision of night time lights as a proxy for economic statistics. J. Econ. Geogr. 15 (1), 217–246.

Shi, Kaifang, Yu, Bailang, Huang, Yixiu, Hu, Yingjie, Yin, Bing, Chen, Zuoqi, Chen, Liujia, Wu, Jianping, 2014. Evaluating the ability of NPP-VIIRS nighttime light data to estimate the gross domestic product and the electric power consumption of China at multiple scales: a comparison with DMSP-OLS data. Rem. Sens. 6 (2), 1705–1724.

Singhal, Abhishek, Sahu, Sohini, Chattopadhyay, Siddhartha, Mukherjee, Abhijit, Bhanja, Soumendra, 2020. Using night time lights to find regional inequality in India and its relationship with economic development. PloS One 15 (11), e0241907.

Small, Christopher, Pozzi, Francesca, Elvidge, Christopher D., 2005. Spatial analysis of global urban extent from DMSP-OLS night lights. Rem. Sens. Environ. 96 (3–4), 277–291.

Tuttle, Benjamin, Anderson, Sharolyn, Elvidge, Chris, Ghosh, Tilottama, Kim, Baugh, Sutton, Paul, 2014. Aladdin's magic lamp: active target calibration of the DMSP OLS. Rem. Sens. 6 (12), 12708–12722.

Zhang, Qingling, Schaaf, Crystal, Seto, Karen, 2013. The vegetation adjusted NTL urban index: a new approach to reduce saturation and increase variation in nighttime luminosity. Rem. Sens. Environ. 129 (1), 32–41.