

# Whose Power Gets Cut?

Using High-Frequency Satellite Images  
to Measure Power Supply Irregularity

*Brian Min*  
*Zachary O’Keeffe*  
*Fan Zhang*



**WORLD BANK GROUP**

South Asia Region

Office of the Chief Economist

June 2017

## Abstract

In many parts of the developing world, access to electricity is uneven and inconsistent, characterized by frequent and long hours of power outages. Many countries now engage in systematic load shedding because of persistent power shortages. When and where electricity is provided can have important impacts on welfare and growth. But quantifying those impacts is difficult because utility-level data on power outages are rarely available and not always reliable. This paper introduces a new method of tracking power outages from outer space. This measure identifies outage-prone areas by detecting excess fluctuations in light outputs. To develop these measures, the study processed the

complete historical archive of sub-orbital Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) nighttime imagery captured over South Asia on every night since 1993. The analysis computes annual estimates of the Power Supply Irregularity index for all 600,000 villages in India from 1993 to 2013. The Power Supply Irregularity index measures are consistent with ground-based measures of power supply reliability from the Indian Human Development Survey, and with feeder-level outage data from one of the largest utilities in India. The study's methods open new opportunities to study the determinants of power outages as well as their impacts on welfare.

---

This paper is a product of the Office of the Chief Economist, South Asia Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at [fzhang1@worldbank.org](mailto:fzhang1@worldbank.org).

*The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.*

# **Whose Power Gets Cut? Using High-Frequency Satellite Images to Measure Power Supply Irregularity<sup>1</sup>**

Brian Min  
University of Michigan

Zachary O’Keeffe  
University of Michigan

Fan Zhang  
World Bank

**Key words: Power outages, load shedding, nighttime lights, remote sensing, DMSP, India**

---

<sup>1</sup> This paper is part of a broader analytical effort to assess the cost of power sector distortions in South Asia conducted by Office of the Chief Economist of the South Asia Region. The authors are grateful to Chris Elvidge, Kim Baugh, Yue Li and Virgilio Galdo for providing access to the data. We thank Martin Rama and Kwawu Mensan Gaba for helpful comments on the methodology of the study. Priyamvada Trivedi and Fatima Najeeb provided excellent research assistance. Financial support from the Partnership for South Asia Trust Fund is greatly appreciated.

## I. Introduction

Almost 1.2 billion people in the world live without electricity. A billion more only have an unreliable and intermittent supply. Maintenance failures, transmission congestion, and inadequate generation all contribute to frequent and long hours of blackouts in the developing world. According to World Bank Enterprise Surveys, firms in a third of developing countries experience at least 20 hours of power outages in a typical month. The situation is even worse in South Asia, where businesses report almost one outage a day, lasting on average 5.3 hours each.

Power outages impose sizable costs on households, firms, and the environment. Lack of access to reliable electricity is a major barrier to economic advancement for underserved populations (Samad and Zhang, 2016 and 2017). Unreliable power supply also negatively affects firms' productivity, resulting in lost output and revenue (Reinikka and Svensson 2002; Alby, Dethier et al. 2011; Allcott, Collard-Wexler et al. 2014, Andersen and Dalgaard, 2013, Grainger and Zhang, 2017a, b). Blackouts also lead to environmental damages when households have to rely on kerosene for lighting and businesses use diesel-based self-generators. Both kerosene lamps and diesel generators are major sources of black carbon, contributing to climate change and degraded air quality.

Although power outages have severe economic and environmental consequences, tracking when and where power outages occur is often difficult. Data on load-shedding are often not reported consistently and are always difficult to acquire. Household surveys based on a Multi-Tier Framework developed by the World Bank offer an alternative way to monitor quality of power supply. But household surveys are costly to implement, representative at a highly-aggregated level, and only reflect a snapshot in time due to limited frequency.

In this paper, we introduce a new method of using high-frequency satellite imagery of nightly lights data to detect near real-time power outages at a highly disaggregated level. This measure identifies outage-prone areas by detecting excess fluctuations in light outputs. To develop this measure, we process the complete historical archive of sub-orbital DMSP-OLS nighttime imagery captured over South Asia on every night since 1993, including 5 terabytes of data encompassing some 30,000 visible band images. Our approach is a departure from prior research. Most earlier analysis of nighttime lights data uses annual composite images which describe the average brightness of a locality over a calendar year. While annual composites provide a useful estimate of average brightness, they cannot be used to assess volatility of brightness which indicates disruptions to power supply. To the

best of our knowledge, our study is the first to explore nightly variation in light output to detect the stability of electricity provision.

To illustrate the approach, we compute annual estimates of the Power Supply Irregularity (PSI) index for all 600,000 villages in India from 1993 to 2013. We test the validity of the computed PSI using ground-based measures of power supply reliability from Indian Human Development Survey and feeder-level outage data from Maharashtra's largest utility. We find that PSI measures are significantly positively correlated with the duration and frequency of outages reported in the ground-truth data.

Our methods open new opportunities to study the determinants of power outages as well as their impact on welfare. A key strength of our method is that it can be applied to any specific location of interest using the same autonomously collected data. Moreover, since the coefficient of variation is a normalized measure of variability, values on the PSI index can be directly compared across regions. The method also demonstrates the possibility of using near real-time satellite data to monitor power outages and identify problem areas.

The rest of the paper proceeds as follows: Section II describes the significance of power outages in the developing world. Section III presents the new approach to measuring power supply reliability based upon an analysis of a time series of nighttime satellite imagery. Section IV discusses ground-truth validation of our new power supply irregularity index. Section V presents power outage estimates across India using the new measure. Section VI concludes the paper.

## **II. Power Outages in the Developing World**

For many in the developing world, access to electricity is uneven and inconsistent, characterized by frequent and long hours of power outages. Many countries now engage in planned load shedding because of a systemic shortage of supply relative to rapidly increasing demand for electrical power. For example, the World Bank reports that the average Nigerian firm suffers 25 power outages in a typical month, with more than one outage a day reported in many other countries like Albania, Kosovo, Nepal, and the Republic of Yemen. In Bangladesh, firms reported over 100 power outages a month (World Bank Enterprise Surveys).

In India, firms report almost one outage every other day, with an average duration of 2 hours. The situation is almost certainly worse for Indian households. According to the Indian Human Development Survey, 57% of electrified households reported

at least 4 hours of power cuts per day. Moreover, in 2012 fewer homes benefited from uninterrupted power supply than in 2005 (12% compared to 17%), suggesting that outages are becoming more common, not less.

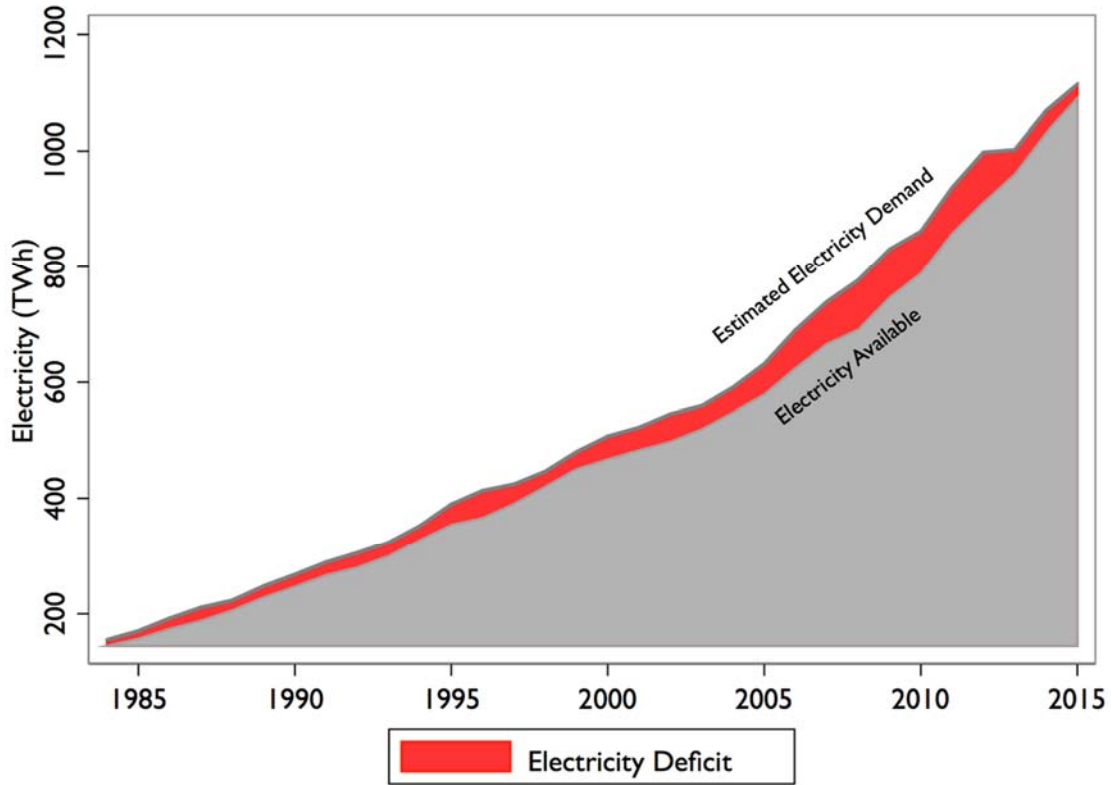
Blackouts can occur for many reasons. Electricity is sent from power generating plants to a country's users through the grid, an intricate network of wires interconnected by transformers and substations. Sometimes the faults are due to downed wires or short circuits, as from a fallen tree during a storm. More commonly in the developing world, when demand exceeds supply or the capacity of a portion of the grid, outages are intentionally imposed to reduce loads and protect the sensitive infrastructure of the grid.

Unlike many other services, electricity is a dynamic system that must be balanced in real time. Every unit of electricity generated must be used at the very same moment. There are no cost effective ways to store electricity, and if too much electricity is generated without an accompanying load, the excess energy can lead to system failure. Not only do wires heat up when placed under extreme loads, but transformers, whose job it is to convert the very high voltages used in transmission into a level more safely used by consumers, are placed under enormous stress. When too much power is drawn, transformers and other critical equipment can blow, with costly consequences.

As a result, utility officials must constantly monitor and predict changing demand as it fluctuates throughout the day and across seasons, and manage supply to ensure the right number of power plants are in operation. When demand exceeds the supply capacity, power companies must preemptively disable the supply of electricity to portions of the grid. These intentional blackouts — usually called load shedding or rostering — often cut off electricity to users when they need power the most.

Official estimates in India indicate that there has been a consistent deficit in electricity supply relative to estimated demand, at least for the last three decades. Figure 1 plots the total supply of electricity in India against estimated demand. In 1984, India experienced an overall shortage of 6.7%, peaking to 11.6% on the highest demand days. By 2012, even after total electricity supply had increased more than sixfold, the overall shortage had nevertheless increased to 8.7%, with a peak shortfall of 9.0%.

**Figure 1: Official electricity shortages in India, 1984–2015**



Source: Central Electricity Authority, *Growth of Electricity Sector in India from 1947–2015*.

There are significant regional differences in demand and supply of power across India. Table 1 shows both the spatial and temporal variation in peak shortages. While all regions experience shortages, the highest shortfalls are in the northeast and in the western region, which includes Gujarat, Madhya Pradesh, and Maharashtra. The lowest shortages are in the eastern region, which includes Bihar and Orissa.

Table 1: Peak Energy Shortages (%) by Region

	2007	2008	2009	2010	2011	2012	2013	2014	2015	Avg (07-15)
Eastern <sup>1</sup>	-11.1	-8.6	-6.3	-5.0	-4.8	-7.4	-1.8	-0.6	-0.6	-5.1
North-Eastern <sup>2</sup>	-22.7	-25.4	-17.9	-18.5	-7.2	-6.7	-5.4	-12.9	-8.0	-13.8
Northern <sup>3</sup>	-9.1	-10.7	-15.4	-8.9	-7.8	-8.9	-6.9	-8.3	-7.1	-9.2
Southern <sup>4</sup>	-9.0	-9.4	-9.7	-6.4	-14.4	-18.5	-7.6	-5.2	-0.4	-9.0
Western <sup>5</sup>	-23.2	-19.0	-17.7	-14.7	-13.8	-1.5	-2.4	-2.3	-0.9	-10.6
All India	-16.6	-11.9	-12.7	-9.8	-10.6	-9.0	-4.5	-4.7	-3.2	-9.2

<sup>1</sup> Bihar, Damodar Valley Corporation, Jharkhand, Orissa, West Bengal, Sikkim, Andaman & Nicobar

<sup>2</sup> Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland, Tripura

<sup>3</sup> Chandigarh, Delhi, Haryana, Himachal Pradesh, Jammu & Kashmir, Punjab, Rajasthan, Uttar Pradesh

<sup>4</sup> Andhra Pradesh, Karnataka, Kerala, Tamil Nadu, Telangana, Puducherry, Lakshadweep

<sup>5</sup> Chhatisgarh, Gujarat, Madhya Pradesh, Maharashtra, Goa, Daman & Diu, Dadra and Nagar Haveli

Source: Central Electricity Authority, *Growth of Electricity Sector in India from 1947–2015*.

While the overall shortage of power is well recognized, we know much less about exactly whose power gets cut most frequently, and how the incidence of power outages varies at lower levels like districts or villages. While many state utilities publish rostering schedules, these do not capture the actual timing of outages, nor their precise geographic coverage.

Some observers allege that load shedding affects urban areas and industries most severely. Yet in recent years, anecdotal evidence suggests that load shedding is substantially worse in rural areas than in urban areas. In 2008, official policy in Uttar Pradesh’s main power utility dictated no more than four hours of daily power cuts in the largest cities, but up to twelve hours of cuts for rural villages (Min 2015). Harish and Tongia (2014) examine minute-by-minute data from hundreds of feeders across 9 days in 2012–13 and find that feeders supplying electricity to rural villages experienced higher load shedding than those that served urban areas. Among urban areas, Bangalore had the lowest rates of load shedding. These patterns are notable given that if the overall objective of load shedding is to remove demand, it could be more efficient to do so in urban areas where the number of consumers and the average per capita loads are higher.

To date, systematic evaluation of patterns of load shedding across all of India has been impossible due to a lack of comprehensive and comparable data on outages.

### III. A New Measure of Power Supply Irregularity (PSI)

We propose a new method to detect power outages by relying upon a time series of nighttime satellite imagery. Our premise is that areas prone to power outages are more likely to have higher variability in light output, since electrified areas will be



illuminated when the power is working but appear dark when there is no power. Across a time series of imagery, areas prone to power outages should have higher variability in light output than areas with no outages.

Most prior research on nighttime light has relied on analysis of annual composite images. While annual composites provide a useful estimate of average brightness across the globe, they depict overly prolonged timespans, smooth away substantial variation in light output over the calendar year, and do not enable precise evaluation of discrete interventions like new village electrification projects, extensions of the power grid into new areas, nor short-term disruptions to the supply of electricity as a result of load shedding or maintenance failures. In this analysis, we return to the original raw nightly data to better study changes in the level of light output, as well as information on the variability of light output, which is a useful indicator of power supply irregularity.

The primary assumption of the approach, building upon a growing body of research (Elvidge et al. 1997b; De Souza Filho 2004; Doll et al. 2006; Sutton et al. 2007; Kiran Chand et al. 2009; Ghosh et al. 2010; Henderson et al. 2012; Min 2015), is that villages that benefit from electrical power appear brighter at night than unelectrified villages because of the use of outdoor lighting and leakage from indoor lighting sources. This relationship has been documented via ground-truthing studies comparing survey-based measures of electricity use to nighttime output in villages in Senegal, Mali, and Vietnam (Min et al. 2013; Min and Gaba 2014).

In partnership with NOAA's National Centers for Environmental Information (NCEI), we acquired the complete archive of raw DMSP-OLS visual band imagery captured over India from 1993 to 2013. The archive comprises more than 30,000 high-resolution image strips captured over nearly 8,000 nights. Each image strip represents a single pass on DMSP's sun-synchronous, near-polar orbit with a swath width of about 3,000 km. Because India has an east-west width of about 3,000 km, image strips from two or three overpasses were often required to capture all locations on the Indian subcontinent.

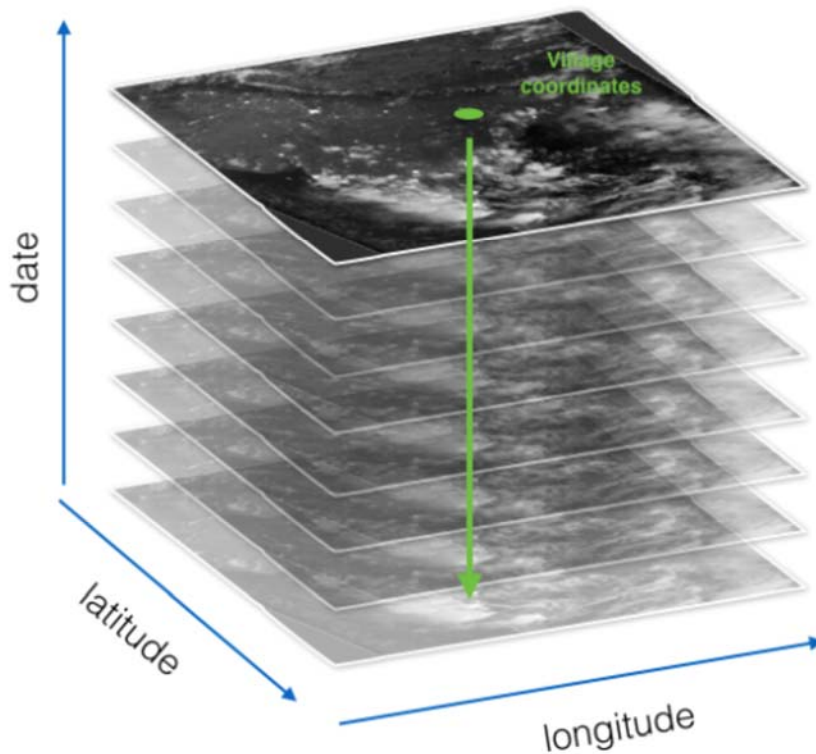
The primary data of interest is the visible (VIS) band data, which records light output in a bandpass covering wavelengths from about 0.4 to 1.0  $\mu\text{m}$ , which overlaps the wavelengths most commonly detectable to the human eye. Brightness on the visible band is recorded on a 6-bit scale, with relative brightness (digital number) values from 0 to 63.

Once the data were acquired and organized, the next step was to extract relevant brightness values from the image products. Since energy access and quality

problems are most severe in rural areas, we focused on extracting brightness values only for villages in India, excluding urban areas and large towns from the analysis.<sup>2</sup> We extracted the brightness observed at the center of each of India's 600,000 villages from every nighttime image provided by NOAA.

Figure 2 illustrates the process by which we extracted the brightness level for individual village points from all 8,000 nights of imagery. This process was run on Flux, the University of Michigan's high performance computing cluster, which allowed us to conduct the extraction in parallel for each village. The final output is a data set of brightness levels for all villages on every night from 1993 to 2013, comprising several billion observations.<sup>3</sup>

**Figure 2: Data extraction from satellite imagery**



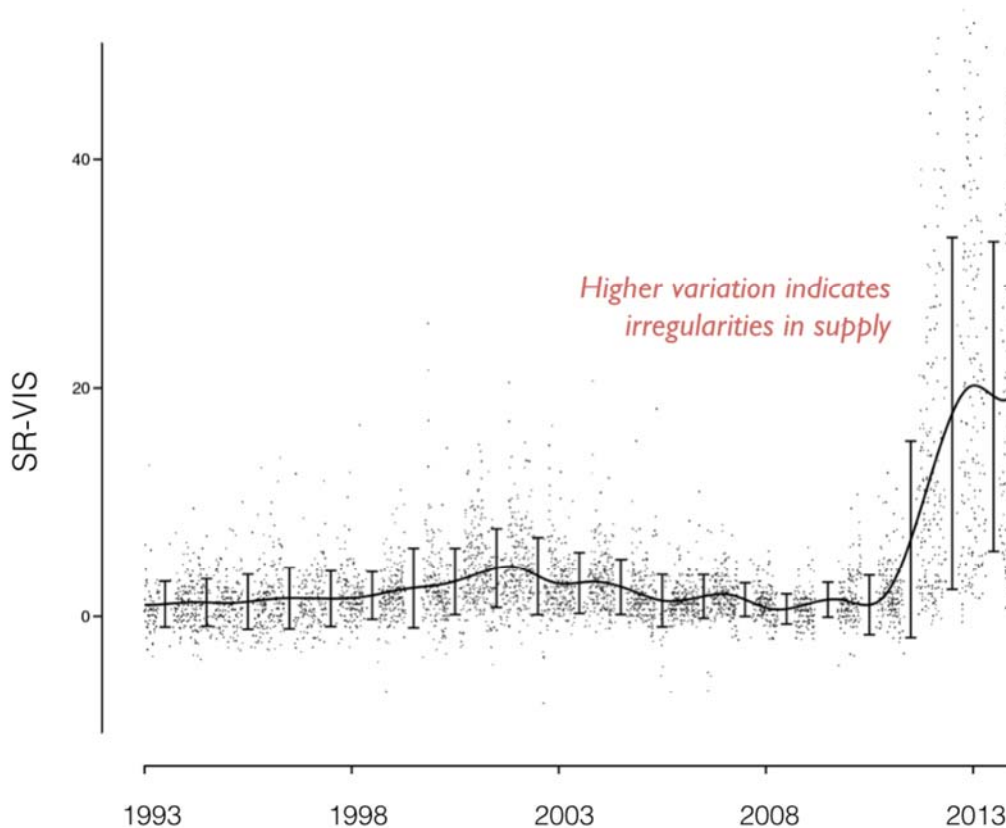
---

<sup>2</sup> Another reason to exclude urban areas is that the DMSP-OLS satellite sensor is not well calibrated, and lacks the dynamic range to accurately measure changes in light in the very bright areas common in urban cores.

<sup>3</sup> An online visualization of this data set is available at <<http://india.nightlights.io>>.

The raw nightly data are not directly comparable over time because the satellite sensors use dynamic gain settings to maximize contrast and image clarity. Since the gain settings are not recorded in the data stream, the 0–63 relative brightness values cannot be converted to an absolute measure of brightness. To overcome this limitation, we use statistical and image processing procedures to create what we refer to as statistically recalibrated visible (SR-VIS) band data. The SR-VIS data transforms the raw light output measures onto a more comparable scale, enabling more reliable statistical comparisons of cross-sectional and within-unit variation (see Appendix for details).

**Figure 3: Variability in light output over time can indicate power supply irregularities**

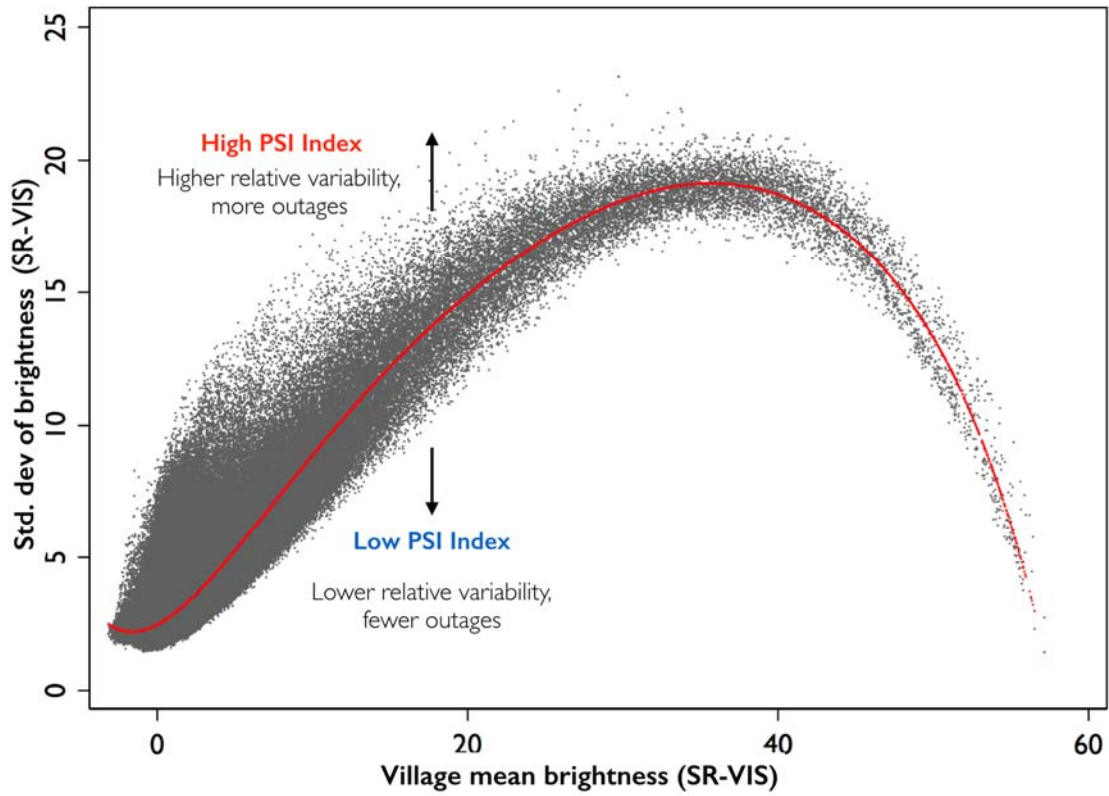


Note: Points represent SR-VIS brightness on all nights with good quality data. Bars indicate plus and minus one standard deviation from the mean brightness of the village calculated for each calendar year.

Examining the time series data of SR-VIS brightness values reveals both how villages vary in their average brightness level, as well as the variability of the brightness values over time. Figure 3 plots the brightness of the village of Sheorajpur in Allahabad district on every night from 1993 to 2013. Electrification of Sheorajpur as part of India's national electrification program, that is Rajiv Gandhi Grameen Vidyutikaran Yojana (RGGVY) program, was completed on 30 June 2007. Light output in the village did not increase much until about 2010, suggesting a lengthy period before electrical power was widely available and in use by residents of the village. The bars indicate one standard deviation above and below the mean of all statistically recalibrated visible band brightness observations in each year. The increase in the size of the bars in recent years shows that variability in light output has increased. Power outages are an obvious reason why a village would appear brightly lit on one night but far dimmer on another. We use this information on the variability of light output over a village to construct a new Power Supply Irregularity (PSI) index.

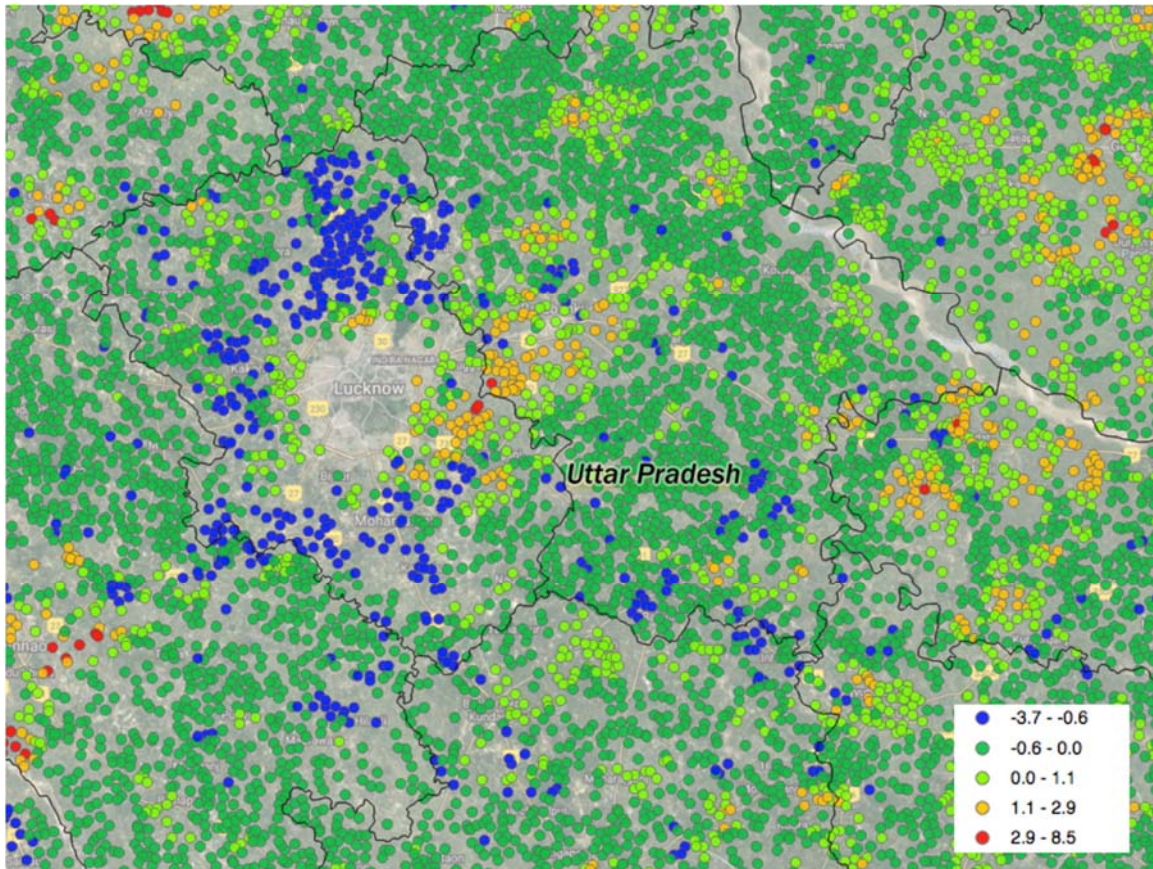
We define PSI as the unexplained variability in light output relative to the predicted variability of light output for a village, calculated using data over a time period of interest. The excess variability is calculated as the residual of a regression using data from all villages that predicts the standard deviation of light output given the brightness of each village. To fit the curvilinear relationship between standard deviation and brightness in the DMSP data, we estimate a regression model including several polynomial terms. More details can be found in the Appendix.

**Figure 4: Calculating the Power Supply Irregularity (PSI) index**



Note: Points represent all rural villages in India. PSI of a village is the excess variability of light output relative to the predicted variability, calculated over the relevant time period using all good quality nightly (SR-VIS) observations.

**Figure 5: Power Supply Irregularity (PSI) of villages**



Note: Village center point locations are from MLInfomap.

Figure 4 is a scatterplot of predicted values from the polynomial regression for all villages in India. Each point represents a village's mean brightness and standard deviation of nightly brightness calculated over a specified time period (in this case, a year). PSI for that village is the difference between the actual standard deviation and the predicted standard deviation. To enable clearer visualization, we transform and normalize the PSI values (see Appendix). Figure 5 plots PSI values in 2013 for individual villages. The map shows how PSI values can be regionally clustered, like the ring of high values in villages surrounding Lucknow, the capital of Uttar Pradesh. Zooming out to the whole country, there is significant variation in PSI values across the country. We discuss these patterns in greater detail in section V.



## IV. Validation of PSI against Ground-Based Data on Outages

### A. Validation in Maharashtra

To evaluate the reliability and consistency of our new measure, we also compare PSI against officially reported power system reliability data from Maharashtra's largest power utility, Maharashtra State Electricity Distribution Co. Ltd, or MahaDiscom. The company provides information on its website on the reliability of its feeder network. The most widely used measure of reliability is SAIFI, or System Average Interruption Frequency Index, which describes the average number of interruptions experienced by a customer over a prescribed period. MahaDiscom reports SAIFI for "circles," which overlap closely with districts in each year over time. However, MahaDiscom's SAIFI reports are not a perfect measure of outages because MahaDiscom does *not* include planned load shedding in their measure, only unplanned outages. Nevertheless, we would expect districts with higher SAIFI levels to be associated with higher PSI.

MahaDiscom provides power reliability indices at the circle level by month from 2009 to 2013. Circles are the largest geographic unit MahaDiscom uses, which roughly correspond to districts. Because PSI is calculated annually, annual averages of the reliability indices were calculated. In the few cases where multiple circles are contained in a district, their values were averaged. Because PSI is calculated for rural villages only, urban circles were omitted from the analysis. There are 33 district-circle matches, and five years of data. There are no data from MahaDiscom for two districts in 2010. The total  $N$  for the regressions is thus 163.

We focus on the System Average Interruption Frequency Index (SAIFI), which is the average number of sustained interruptions a customer would experience in a given period of time. It is  $(\text{total number of customer interruptions})/(\text{total number of customers})$ . It does not include planned outages (e.g., load shedding), or outages due to natural disasters.

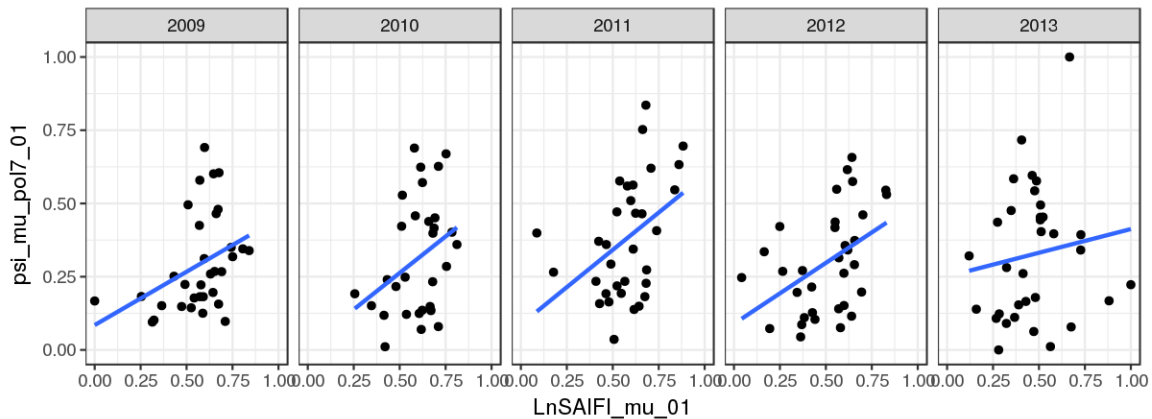
While we expect SAIFI to correlate positively with PSI despite the fact that it does not encompass all outages, there are other factors associated with light output patterns. For example, agricultural electricity usage is less likely to result in light output than residential or industrial use. To control for these confounders, we gather district estimates from the 2011 Indian Census on: logged population; logged area; % illiterate; % work in agriculture; % forested; % population rural; and % area rural. We use these because they might impact actual electricity use (e.g., poverty, as approximated by the literacy rate and proportion of people working in

agricultural jobs) and/or the observed light output (e.g., population density and forests).

Due to the highly skewed nature of SAIFI, a natural logarithmic transformation is applied. For ease of interpretation, both log-SAIFI and PSI are rescaled from 0 to 1. Four OLS regressions are run, with PSI as the outcome variable, and log-SAIFI as the primary predictor of interest. Because the data are panel and observations are likely not independent, all models are run with robust delta method standard errors clustered on district. The controls in models (2) and (4) are: logged population, logged area, % illiterate, % work in agriculture, and % forested. Models (3) and (4) also include year fixed effects.

As can be seen from the regression results in Table 2, the estimated relationship between SAIFI and PSI is consistently positive. This holds across models that add additional controls as well as with the inclusion of year fixed effects.

**Figure 6: Correlations of annual ln(SAIFI) (scaled) and PSI (scaled) by Maharashtra district for 2009-2013,**





**Table 2: PSI and power supply reliability in Maharashtra districts**

DV is PSI measured in each district-year, 2009–13

	(1)	(2)	(3)	(4)
<b>ln(SAIFI)</b>	<b>0.348**</b>	<b>0.224'</b>	<b>0.369**</b>	<b>0.239'</b>
	<b>(0.114)</b>	<b>(0.125)</b>	<b>(0.125)</b>	<b>(0.139)</b>
ln(population)		-0.124		-0.122
		(0.199)		(0.202)
ln(area)		0.189		0.187
		(0.190)		(0.193)
% population illiterate		0.009'		0.009'
		(0.005)		(0.005)
% workers in agriculture		-0.004		-0.004
		(0.005)		(0.005)
% area forest		-0.001		-0.001
		(0.003)		(0.003)
Year FE	No	No	Yes	Yes
intercept	0.133*	0.340	0.082	0.271
	(0.061)	(1.706)	(0.066)	(1.735)
<i>N</i>	163	163	163	163

Robust standard errors clustered on district in parentheses.

' p &lt; .1; \* p &lt; .05; \*\* p &lt; .01.

**B. Validation against IHDS Survey Data**

The second source of ground-truth data is the Indian Human Development Survey (IHDS), which was jointly carried out by University of Maryland and the National Council of Applied Economic Research (NCAER) in New Delhi. The survey covers all of India's key states and union territories except Andaman and Nicobar Islands and Lakshadweep. The first round of the survey was carried out mostly in 2005 and collected information on 41,554 households. The second one, conducted mostly in 2012, interviewed a total of 42,152 households, including 83 percent of the original households. The survey asks a wide range of questions related to households' energy consumption patterns including electrification status and daily average duration of electricity outages. While the survey includes both urban and rural households, we focus on the rural sample for reasons described in the previous section.

We calculate district-level average PSI and district-level average duration of power outages reported by rural households. We also obtain data on total population, total area, % of agriculture employment and forest area all at the district level from the 2011 Indian Census because these variables could also affect PSI values extracted from satellite images. Table 3 summarizes regression results using district-level PSI

as the dependent variable. The results show that average duration of power outages reported in IHDS are significantly and positively correlated with PSI.

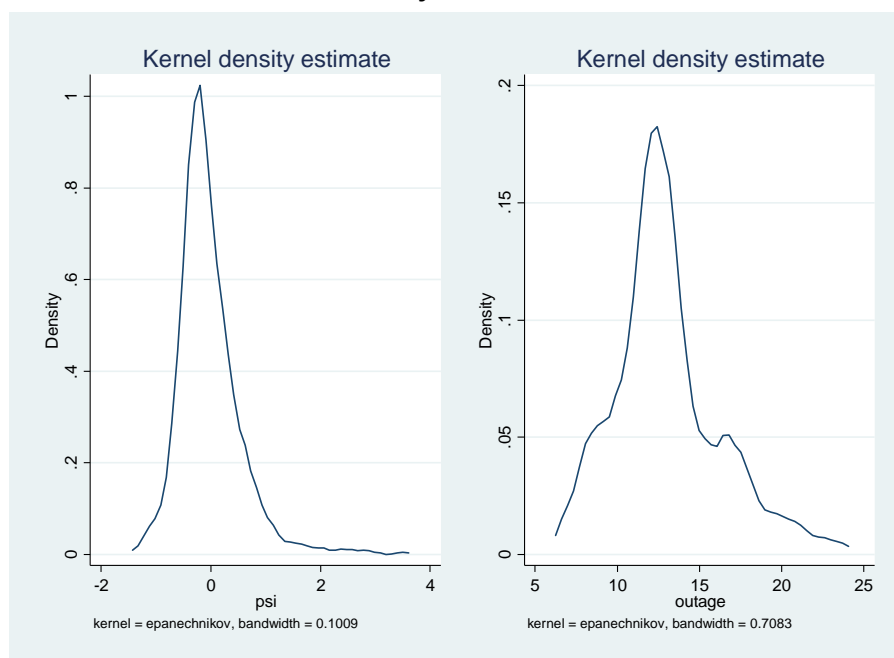
**Table 3 Correlation between PSI and duration of power outages reported by IHDS**

	(1)	(2)
outage_IHDS	0.0358** (0.0140)	0.0503*** (0.0143)
Forest		0.00266 (0.00535)
Area		0.0000160 (0.0000102)
Population		-0.0000920** (0.0000364)
Emp_Agriculture		0.00304 (0.00360)
N	217	206
adj. R-sq	0.025	0.077

Note: The dependent variable is PSI. Standard errors clustered at the district level are reported in parentheses.

Figure 7 depicts the probability density distribution of PSI and the outage variable based on kernel density estimation. It shows that both variables generally follow a normal distribution.

**Figure 7: Probability density distribution of PSI and power outages reported by IHDS**



## V. Variations in Power Outages across India

Table 4 ranks all states by the mean level of PSI in 2013. In this section, we describe anecdotal evidence which comports with these rankings. Power supply reliability as measured by PSI was highest in West Bengal, a state that has been recognized for its progress in reforming its power sector.<sup>4</sup> In 2013, it recorded a power surplus for the first time (though much of it was driven by a drop in industrial demand). Speaking to a reporter, the principal secretary for West Bengal’s department of power, Malay Kumar De, said, “Political interference at the local level is totally absent here. We don’t have a power rationing or shortage situation.”<sup>5</sup>

Areas with severe power reliability problems have also received media attention. For example, a news article on the ongoing load shedding situation in Madhya Pradesh identified particularly poor service in Harda and Hoshangabad districts which had been receiving only 10 and 13 hours of power a day, respectively.<sup>6</sup> These

<sup>4</sup> World Bank. 2009. *West Bengal Power Sector Reforms: Lessons Learnt and Unfinished Agenda*. Washington, DC. <<https://openknowledge.worldbank.org/handle/10986/12375>>

<sup>5</sup> “Power thieves prosper in India’s patronage-based democracy.” *Washington Post*. 4 October 2012.

<sup>6</sup> “No end to state’s power woes.” *Times of India*. 27 August 2014.

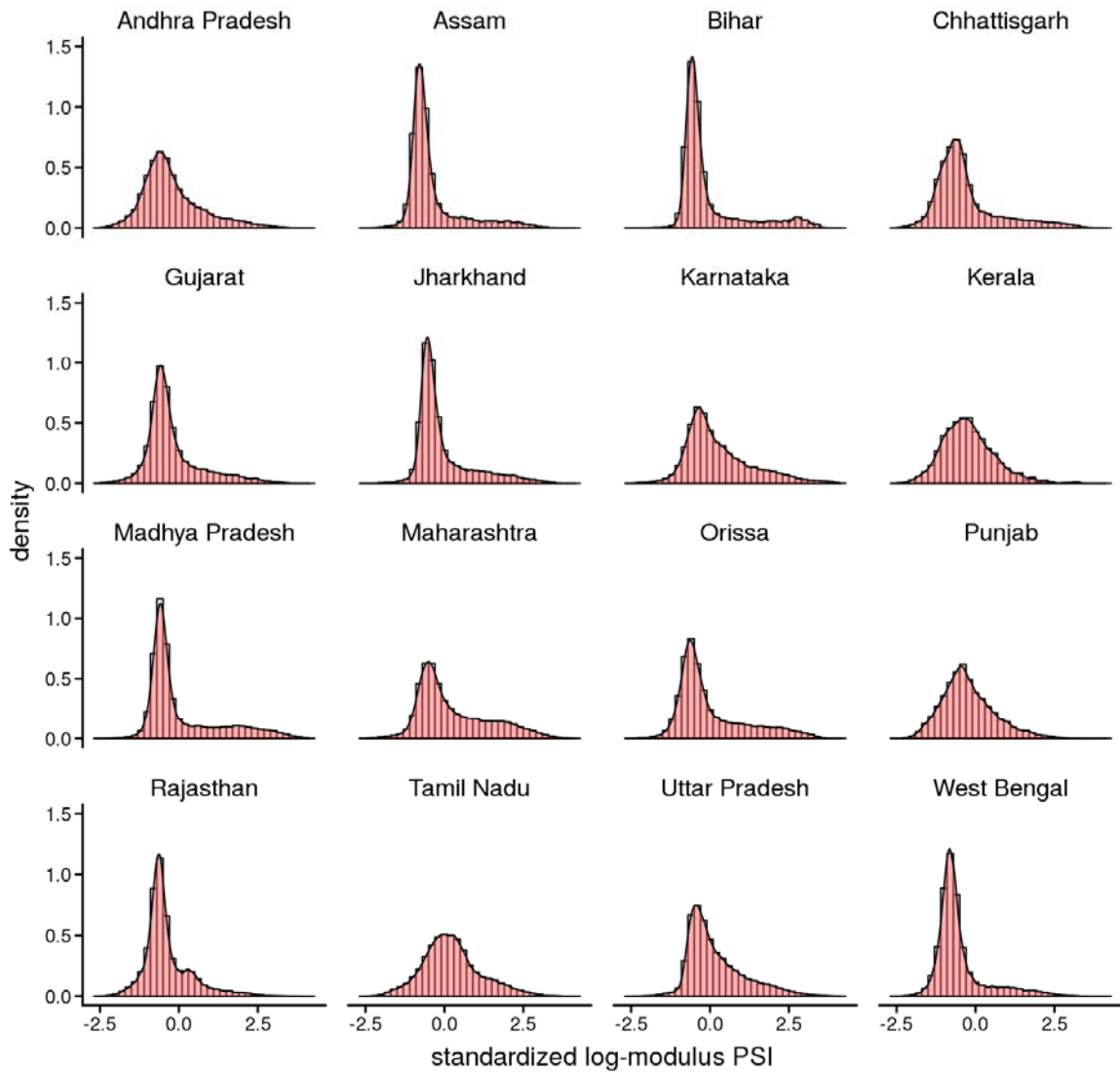
districts had the second and third highest PSI levels in 2013 among all districts in the country.

While the table of state level averages is informative, the map in Figure 9 reveals that in many parts of the country, there is even more local and regional variation within states than across-state variation.

**Table 4: Power Supply Irregularity (PSI) index, 2013  
state-wise rankings (higher values indicate more outages)**

Rank	State	Mean PSI <sub>13</sub>	Std. dev PSI <sub>13</sub>	Village observations
1	Mizoram	1.955	1.989	729
2	Manipur	1.124	1.595	2,199
3	Meghalaya	0.835	1.569	5,782
4	Jammu And Kashmir	0.411	0.538	6,417
5	Uttrakhand	0.257	0.626	15,761
6	Maharashtra	0.257	1.200	41,095
7	Karnataka	0.248	1.179	27,481
8	Nagaland	0.234	1.032	1,278
9	Himachal Pradesh	0.214	0.674	17,495
10	Arunachal Pradesh	0.209	0.708	3,863
11	Madhya Pradesh	0.175	1.354	52,117
12	Tamil Nadu	0.132	0.866	15,215
13	Tripura	0.129	1.108	858
14	Uttar Pradesh	0.128	0.887	97,942
15	Orissa	0.048	1.143	47,529
16	Bihar	0.023	1.149	39,031
17	Jharkhand	-0.039	0.891	29,354
18	Gujarat	-0.198	0.846	18,066
19	Andhra Pradesh	-0.231	0.906	28,108
20	Kerala	-0.261	0.728	1,364
21	Punjab	-0.271	0.720	12,278
22	Chhattisgarh	-0.273	1.046	20,286
23	Assam	-0.336	0.834	25,121
24	Rajasthan	-0.369	0.659	39,753
25	Haryana	-0.439	0.771	6,764
26	West Bengal	-0.472	0.716	37,954
	ALL INDIA	0.000	1.029	594,818

**Figure 8: Distribution of PSI across villages in India's 16 largest states, 2013**



**Figure 9: PSI index in northern India, 2013**

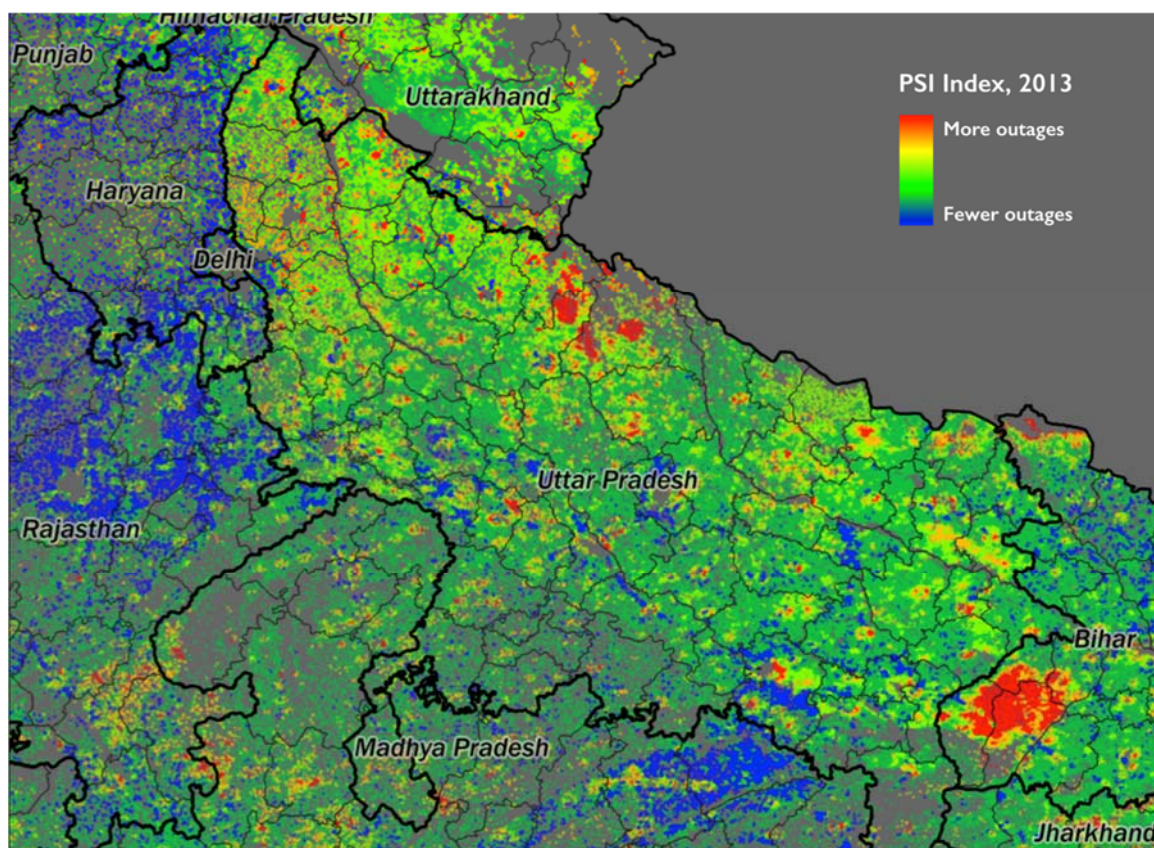


Figure 9 presents a closer look at power supply irregularities in northern India. One striking pattern is the notable difference across state borders, where state-level distribution companies' responsibilities end. The power supply is far more stable in Haryana and Rajasthan than in neighboring Uttar Pradesh.<sup>7</sup> In Uttar Pradesh, India's largest state, reliability varies widely, with the worst conditions in a few northern districts, and better conditions around Lucknow, Etawah, and Varanasi.

In 2013, power supply was especially irregular in a concentrated area of western Bihar, spanning Buxar, Kaimur, Rohtas, and Bhojpur districts. Media reports also describe unrest in these areas due to the irregular supply.<sup>8</sup>

---

<sup>7</sup> The Delhi region is largely unrepresented in our data since we track only rural villages, and exclude cities and large towns.

<sup>8</sup> "Protests continue in Bihar against power crisis." Indo-Asian News Service. 28 March 2011. <<https://www.yahoo.com/news/protests-continue-bihar-against-power-crisis-20110328-012651-751.html>>

## **VI. Conclusion**

In this paper, we develop a new method to monitor the quality of power supply using high-frequency satellite imagery of nightly lights data. We define PSI as the excessive variability in light output relative to the predicted variability in light output. Variability is measured as the deviation from the mean brightness of a locality observed from a lengthy time series of nighttime satellite imagery. To test the validity of the method, we compute the PSI index for all 600,000 villages in India each year from 1993 to 2013. We find that power outages predicted by the PSI index are consistent with those reported by alternative ground-based measures of power supply reliability in India.

Our method of monitoring PSI opens new opportunities for policy research, power system planning, and citizen engagement. First, previous analyses of the impact of power outages have often relied on shortage data at the district or state level. Our research shows that PSI varies significantly across villages. Using the PSI index, researchers can capture the spatial granularity in supply disruption to better estimate the effects of power outages on welfare and growth. Second, the method developed in this study can potentially be used to facilitate near real-time monitoring of power supply disruption by utilities and government agencies. Power system planners can rely on this information to identify outage-prone areas in a cost-effective way. Finally, when information on PSI is made available to consumers, it could help consumers to better engage with utilities to improve the quality of power supply.

In the future, we plan to expand the approach to other parts of the world and to develop a global database of the PSI index for other countries across time.

## References

- Alby, P., J.-J. Dethier, S. Straub, et al. 2011. Let There be Light! Firms Operating under Electricity Constraints in Developing Countries. Technical report, Toulouse School of Economics (TSE).
- Allcott, H., A. Collard-Wexler, and S.D. O'Connell. 2016. How do Electricity Shortages Affect Industry? Evidence from India. *American Economic Review* 106(3): 587–624.
- Andersen, T. B. and C.-J. Dalgaard. 2013. Power Outages and Economic Growth in Africa. *Energy Economics* 38: 19–23.
- De Souza Filho, C.R., J. Zullo Jr., and C.D. Elvidge. 2004. Brazil's 2001 energy crisis monitored from space. *International Journal of Remote Sensing* 25(12): 2475–2482.
- Doll, C.N.H., J.-P. Muller, and J.G. Morley. 2006. Mapping regional economic activity from night-time light satellite imagery. *Ecological Economics* 57: 75–92.
- Elvidge, C.D., K.E. Baugh, E.A. Kihn., H.W. Kroehl, E.R. Davis, and C. Davis. 1997. Relation between satellite observed visible-near infrared emissions, population, economic activity, and power consumption. *International Journal of Remote Sensing* 18(6): 1373–1379.
- Ghosh, T., R. Powell, C. Elvidge, K.E. Baugh, P.C. Sutton, and S. Anderson. 2010. Shedding light on the global distribution of economic activity. *The Open Geography Journal* 3: 148–161.
- Grainger, C. and F. Zhang 2017a. The Impact of Electricity Shortages on Firm Productivity: Evidence from Pakistan. World Bank Policy Research Working Paper.
- Grainger, C. and F. Zhang 2017b. The Impact of Electricity Shortages on Micro- and Small-Enterprises: Evidence from India. World Bank Policy Research Working Paper.
- Harish, S.M. and R. Tongia. 2014. Do rural residential electricity consumers cross-subsidize their urban counterparts? Exploring the inequity in supply in the Indian power sector. Brookings India working paper 04.
- Henderson, J.V., A. Storeygard, and D.N. Weil. 2012. Measuring Economic Growth from Outer Space. *American Economic Review* 102(2): 994–1028.



- Kiran Chand, T.R., K.V.S. Badarinath, C.D. Elvidge, and B.T. Tuttle. 2009. Spatial characterization of electrical power consumption patterns over India using temporal DMSP-OLS night-time satellite data. *International Journal of Remote Sensing* 30(3): 647–661.
- Min, B. and K.M. Gaba. 2014. Tracking Electrification in Vietnam Using Nighttime Lights. *Remote Sensing* 6(10): 9511–9529.
- Min, B. and M. Golden. 2014. Electoral cycles in electricity losses in India. *Energy Policy* 65: 619–625.
- Min, B., K.M. Gaba, O.F. Sarr, and A. Agalassou. 2013. Detection of Rural Electrification in Africa using DMSP-OLS Night Lights Imagery. *International Journal of Remote Sensing* 34(22): 8118–8141.
- Min, B. 2015. *Power and the Vote: Elections and Electricity in the Developing World*. New York: Cambridge University Press.
- Reinikka, R. and J. Svensson. 2002. Coping with poor public capital. *Journal of Development Economics* 69: 51–69.
- Samad, H. and F. Zhang. 2016. Benefits of Electrification and the Role of Reliability: Evidence from India. World Bank Policy Research Working Paper.
- Samad, H. and F. Zhang. 2017. Heterogeneous Effects of Rural Electrification: Evidence from Bangladesh. World Bank Policy Research Working Paper.
- Sutton, P.C., C.D. Elvidge and T. Ghosh. 2007. Estimation of gross domestic product at sub-national scales using nighttime satellite imagery. *International Journal of Ecological Economics and Statistics* 8(S07): 5–21.

## **Appendix: Calculation of the Power Supply Irregularity (PSI) Index**

In partnership with the NOAA National Centers for Environmental Information (NCEI), we extracted and processed the complete historical archive of sub-orbital DMSP-OLS nighttime imagery captured over South Asia. The archive comprises imagery from every night since 1993–2013 spread across 30,000 visible band images and hundreds of thousands of auxiliary data files. In total, the archive contains over 5 terabytes of data. The nightly satellite images from NOAA are georeferenced, such that pixel values correspond to latitude-longitude coordinates. We use the coordinates of the center of each village to determine which light values it should be assigned.

Looking across a sequence of raw nightly images, it is evident that the brightness levels across images varies widely in ways that are unlikely to be due to changes in ground luminosity. There are several well-documented reasons for this (Elvidge et al. 2007). First, the sensor uses dynamic gain calibration to maintain constant cloud reference values under varying conditions of solar and lunar illumination, and thus some images have higher gain settings applied than others. Unfortunately, gain levels are not recorded in the data stream. Second, environmental and atmospheric conditions can significantly distort the brightness of a ground-based points of light through refraction, diffusion, and enshrouding. Finally, digital noise is an unavoidable feature of charge-coupled device (CCD) sensor technology and is more pronounced in high-gain, low-light photography.

To account for these issues, we employ several strategies. First, we filter the data to drop the worst quality observations, including those with cloud cover and those corrupted by stray light. Stray light results from the specific geometry of the DMSP-OLS sensor relative to the sun, and results in rays of sunlight bleeding onto the image strip during certain periods of the year over certain portions of the globe. Removal of these corrupted areas can often result in several weeks of missing data during the summer months at some northern latitudes.

We then apply background noise reduction procedures to construct statistically recalibrated visible band (SR-VIS) values, which corrects for sources of “data stream noise” that can reasonably be attributed to non-ground based factors. This measure of noise is calculated as the average brightness recorded in a sample of unpopulated and unelectrified areas in India that should not be emitting any light at night from the ground. The assumption is that any brightness recorded in these “dark” areas must be due to exogenous factors such as sensor noise or atmospheric conditions. We select “dark” points by identifying land points within the country in the LandScan Population 2013 raster with a zero population count. Within this set, we

keep points that have zero brightness in both the 1992 and 2013 DMSP Nighttime Lights Time Stable Annual Composites. Finally, we take a random sample of 10,000 of these points.

### **Data Selection and Extraction**

Having determined the x-y coordinate pairs of villages and dark points, we move to the data extraction stage. Each of the tens of thousands of GeoTIFFs are successively read into the R software environment. For each satellite observation, the SAM, SLM, TIR, LI, and VIS values for each of the coordinates are extracted. Unreliable and missing data points are filtered out. Specifically, observations are dropped if 1) the SAM (sample) value is one of the six codes of bad data according to NOAA; 2) the CM (cloud mask) is not 0 (which indicates the absence of clouds); 3) the SLM (stray light mask) is not 0 (which indicates the absence of stray light); or 4) the date and latitude are in a set we have determined to be unreliable.

Next, the coordinates, VIS values, satellite ID, date, and timestamp are loaded into Hadoop. For each date, the observation from the newest satellite is kept if more than one observation exists. If more than one observation still remains, the observation from the latest timestamp is kept.

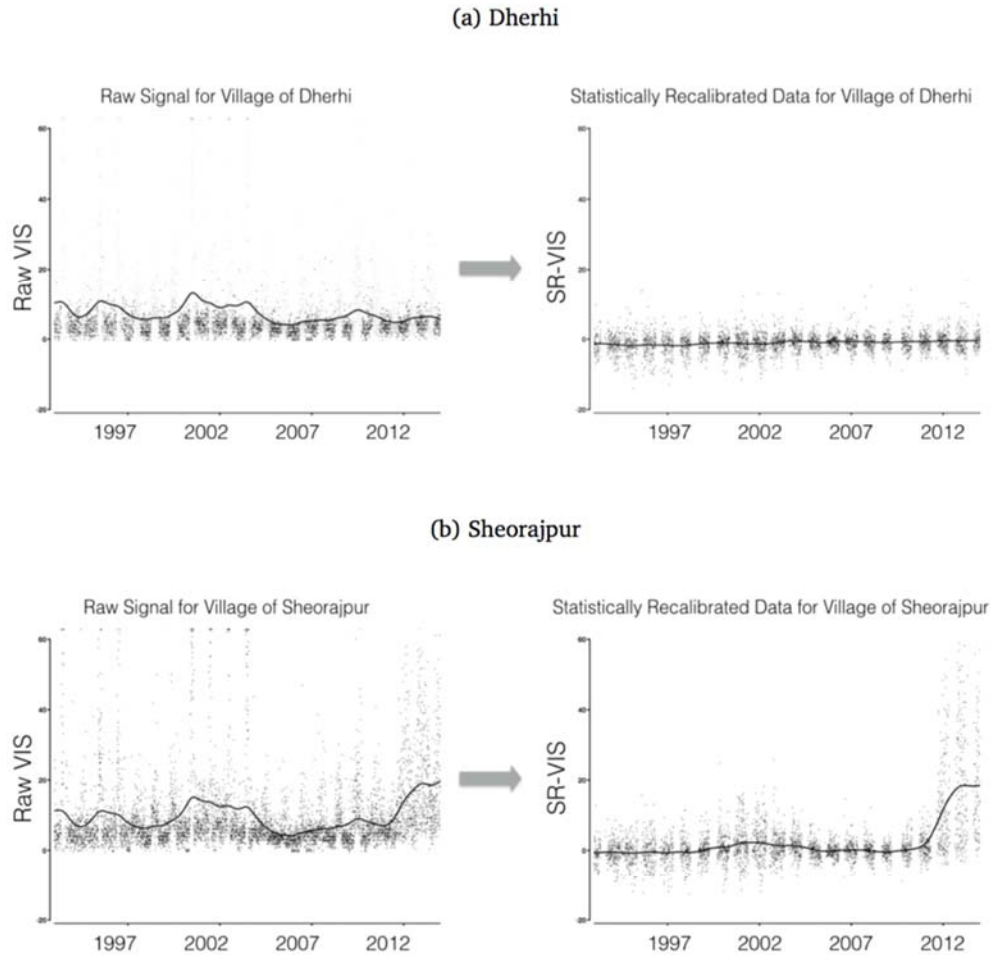
### **Calculation of Statistically Recalibrated Visible Band (SR-VIS) Data**

Villages are matched to dark points in R. For each village coordinate, all dark point coordinates within a 300km radius are considered matches.

For each date, the VIS values of dark point combinations corresponding to villages are averaged (after dropping outliers). If there are fewer than 5 dark VIS observations, the village-date observation is dropped. The remaining mean dark VIS values are then subtracted from their corresponding village VIS values to generate new normalized values that we refer to as statistically recalibrated visible band data, or SR-VIS.

Figure A.1 compares the raw, uncorrected visible band values with the improved SR-VIS values in two villages in Uttar Pradesh. The top panel A.1(a) shows Dherhi, a village in Mirzapur district, and the bottom panel A.1(b), shows Sheorajpur village from Allahabad district. The statistically recalibrated visible band (SR-VIS) data demonstrates less noise and more clearly reveals long-term changes in brightness over time than the raw values.

**Figure A.1: Transforming raw VIS data to Statistically Recalibrated Visible Band (SR-VIS) data**



Note: Transformation involves application of a data quality filter (drop images with clouds, solar glare); background noise removal; and statistical data smoothing to identify temporal trends. Each dot is the brightness value of Sheorajpur, Uttar Pradesh (Allahabad district), on each calendar day. In periods where more than one satellite sensor was recording data, we keep only data from the newer sensor. On days with more than one reading due to multiple evening overpasses by the satellite, the readings are averaged.

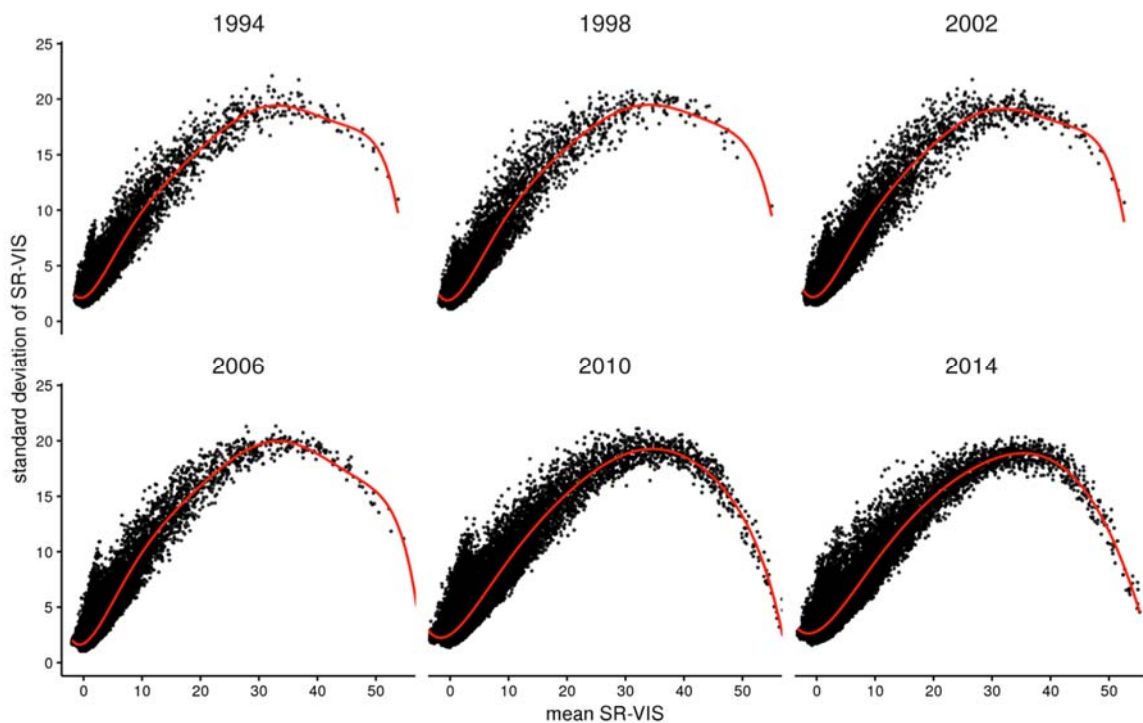
### **Power Supply Irregularity Index**

To calculate the Power Supply Irregularity index for a location, we measure the variability of SR-VIS values across a period of interest. We do so by loading the complete set of all SR-VIS values for all village-days into a Hadoop distributed file system. Using Hive, the mean and standard deviation of SR-VIS is calculated by village for each year. Next, for each year, the standard deviation of village SR-VIS is regressed on mean SR-VIS and additional polynomial terms of SR-VIS. The residuals

of these regressions form the village-year Power Supply Irregularity (PSI) indices. Positive values indicate higher variability in light output than expected given the average light output, while negative values represent more stability in light output.

Because the distribution is highly skewed and there are many outliers, both negative and positive, the PSI values are transformed, primarily for purposes of visualization. First, a log-modulus transformation is applied to PSI. This is done by taking the natural logarithm of 1 plus the absolute value of PSI. This is then multiplied by the sign of the PSI value, such that positive values continue to indicate excess variability, and negative ones represent stability. Finally, this transformed log-modulus PSI is standardized by subtracting the yearly mean from each observation, and dividing the result by the standard deviation, such that the transformed mean is 0 and standard deviation 1.

**Figure A.2: Std. dev. vs. mean SR-VIS of Indian villages, various years**



**Figure A.3: PSI vs. mean SR-VIS in India, 2013**

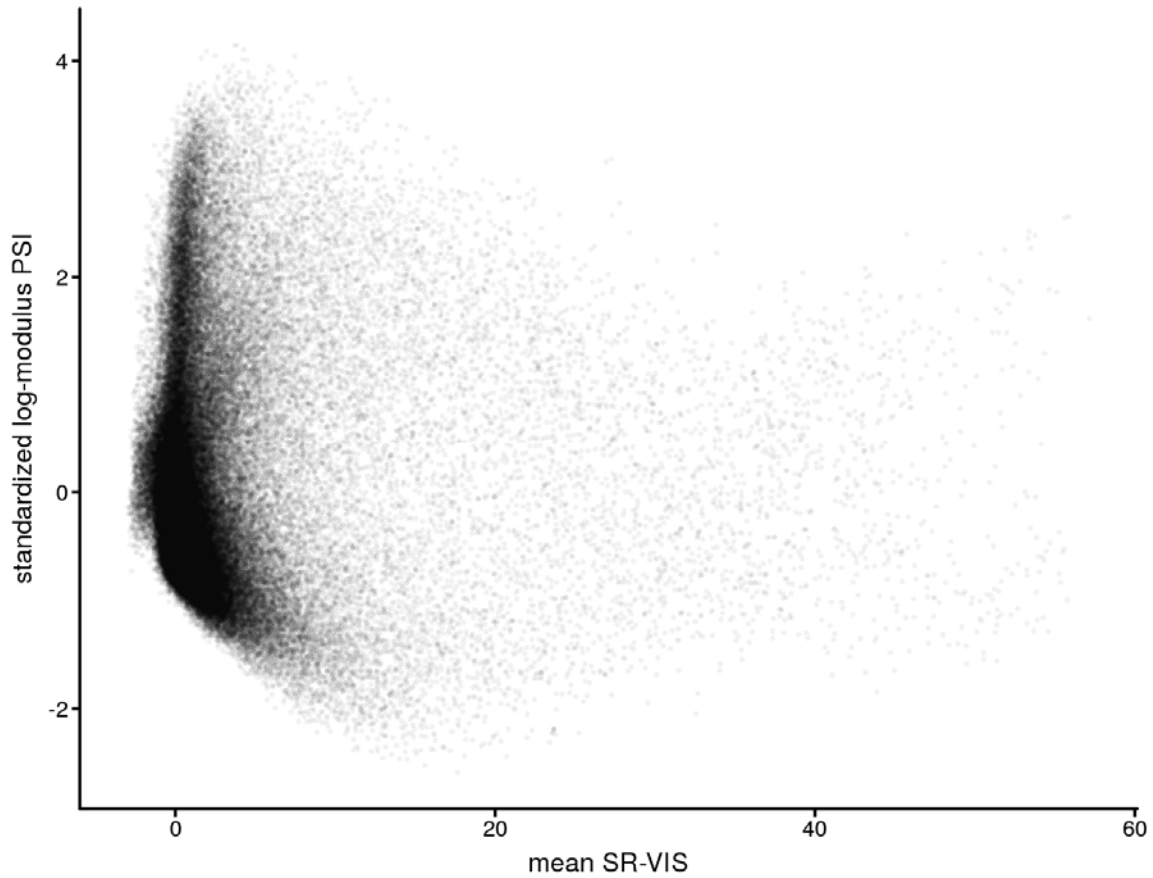


Figure A.3 displays a scatterplot of PSI values versus mean brightness values for villages in India (20% random sample). The data reveal some notable patterns. First, most villages are relatively dim, indicated by the large concentration of data at mean SR-VIS levels below 10. Among these dimly lit villages, it appears that the majority have high PSI values, indicating a high frequency of power outages. As we move to villages with higher levels of brightness (typically due to higher development, more intensive electricity use, larger populations, and proximity to cities), the distribution of PSI is much more evenly distributed across positive and negative values.