

Assessing reliability of electricity grid services from space: The case of Uttar Pradesh, India[☆]

Eugenie Dugoua^{a,*}, Ryan Kennedy^b, Myriam Shiran^b, Johannes Urpelainen^c

^a London School of Economics, United Kingdom of Great Britain and Northern Ireland

^b University of Houston, United States of America

^c Johns Hopkins SAIS, United States of America

ARTICLE INFO

Article history:

Received 15 October 2021

Revised 13 March 2022

Accepted 11 April 2022

Available online 17 May 2022

Keywords:

Electricity

Reliability

India

Nightlights

Remote sensing

ABSTRACT

While most households around the world have access to electricity, the number of hours per day when the grid supplies them with adequate voltage can be low. Improving the reliability of electricity is crucial to make progress on energy poverty but measuring and monitoring it is difficult, especially in lower-income countries where official data is sparse. We develop a transparent method using only easily accessible data to track the reliability of electricity. We train a decision tree model to predict the number of hours with normal electricity in Uttar Pradesh, India, using monthly nighttime luminosity, village characteristics, and voltage data from monitors installed in households. The approach successfully predicts reliability across time and space, and we document that, in Uttar Pradesh, the average number of hours per day with normal electricity has increased by 0.6 h between 2014 and 2019. The predicted number of hours with normal/reliable electricity supply for 2019 remains as low as 8.1 h.

© 2022 The Authors. Published by Elsevier Inc. on behalf of International Energy Initiative. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

The past decade has seen dramatic efforts to address energy poverty throughout the world. The United Nation's Sustainable Development Goal 7 (SDG7) sets a target of ensuring access to affordable, sustainable and modern energy to the world's population by 2030. Significant steps have been taken towards achieving this goal, with developing countries spending billions of dollars to expand access. In 2017, the Indian government promised to electrify every Indian household by 2019. The "Saubhagya" scheme was launched, and 99.9% of households now have access to electricity, according to its online tracking portal (Saubhagya Initiative, 2021). While there are some important caveats to this claim (Aklin et al., 2020; Blankenship et al., 2020; Fairley, 2019; Urpelainen, 2019), there is no doubt about the scale of the achievement in a country that had around 600 million without electricity at the beginning of the 21st century (The Economist, 2018).

If connecting households to national electricity grids is an essential first step towards the goal of universal electrification, the quality of the electricity flowing through the wires is almost as important. When the electricity that reaches households is outside of a reasonable range

of voltage (too high or too low), appliances can quickly lose their utility or be damaged. Previous studies have found that reliability helps determine whether grid connections improve development outcomes. For example, Chakravorty et al. (2014) estimated that, although grid connections raised rural non-agricultural income by about 9%, a higher number of hours of electricity raised income by 28.6%. Two World Bank studies found, in India and Bangladesh, that the scale of economic benefits derived from electrification was highly dependent on the reliability of the electricity provided (Allcott et al., 2016; Samad & Zhang, 2016, 2017).

While data about grid connections is readily available via official records, how much voltage each connection gets throughout the day is typically not. Previous efforts have used crowdsourcing (Kennedy et al., 2020) or surveys (Aklin et al., 2016) to measure reliability, but such efforts remain limited in terms of accuracy and aggregation. Others have used highly disaggregated satellite data (Min et al., 2017). This data, however, is not easily accessible for most scholars and practitioners and requires an extensive infrastructure to store and process.

In this paper, we test a method for monitoring the reliability of electricity in Uttar Pradesh, India, that uses only publicly available and easily accessible nighttime lights data. The intuition is that areas with less reliable electricity should display lower nighttime luminosity, conditional on similar demographic characteristics. We use a regression tree algorithm to develop a predictive model of electricity reliability where the main features are monthly nightlight luminosity and socio-demographic characteristics which we obtain from the Indian census.

[☆] We thank Brian Min and Shweta Kulkarni for helpful comments and feedback.

* Corresponding author.

E-mail address: e.dugoua@lse.ac.uk (E. Dugoua).

We construct a training sample using household voltage data as a ground-truth proxy of electricity reliability.

The results suggest that this approach produces surprisingly accurate estimates of electricity reliability on aggregate and across time and space. We extend the analysis to estimate that the average number of hours per day of normal voltage in Uttar Pradesh has increased by 0.6 h between 2014 and 2019. The expected number of hours of normal voltage remains as low as 8.1 h per day.

The difficulty in measuring reliability

Policy-makers have increasingly recognized the importance of reliability for electrification to achieve its aims in poverty reduction. The importance of electricity reliability in the fight against energy poverty is now fully recognized in policy circles; for example, in the multi-tier framework developed by the World Bank (Bhatia & Angelou, 2014).

If they expect poor quality from the electricity grid, customers have little incentive to connect more devices and use more electricity, resulting in low demand (Grimm et al., 2020; Lee et al., 2017). Kemmler (2007), for example, found that electricity consumption in India correlates strongly with the quality of the electricity supply and only weakly with household expenditures. Others have found that willingness to pay for electricity in Indian villages correlates strongly with the reliability of supply in their area (Graber et al., 2018; Kennedy et al., 2019) and that satisfaction with electricity supply strongly correlates with reliability (Aklin et al., 2016). When poor reliability is combined with the monthly fees imposed by utilities to keep households connected to the grid, customers may find themselves better off by sticking with or returning to traditional energy sources like kerosene or diesel generators.

Unfortunately, monitoring reliability is challenging because it requires almost continuous measurement of currents flowing through the grid. Governments and utilities rarely provide consistent and accurate reporting, and even these reports are often difficult to acquire (Min et al., 2017). Both the utilities (Grimm et al., 2020; Lee et al., 2017) and politicians (Baskaran et al., 2015; Min & Golden, 2014) have little incentive to improve this situation. One may think that utilities and governments would keep comprehensive and accurate data, but this is rarely the case in developing countries. Data on load-shedding is inconsistently reported and challenging to acquire (Min et al., 2017). For example, in 2017, official data in Uttar Pradesh showed nearly no power cuts, while independent monitors indicated regular outages of two to nine hours a day (Sengupta, 2017).

Incentives for both utilities and government officials to accurately report reliability are weak. Utilities have little incentive to release data that would document the poor reliability of the electricity they sell. Extending electricity connections to poorer, more rural areas has already placed their resources under strain (Grimm et al., 2020; Lee et al., 2017). Publicizing problems in deliveries, especially poorly planned load-shedding, is only likely to encourage further requests for higher investments into areas with low purchasing power. Incentives are also weak for politicians to enforce accurate and public reporting of power outages. In fact, previous studies have suggested that they strategically ignore practices that contribute to poor reliability, like electricity theft, to improve their electoral prospects (Baskaran et al., 2015; Min & Golden, 2014). Publicly available data from official sources is therefore often unavailable or not reliable.

In response to the lack of data on electricity reliability, scholars have adopted a range of approaches. One is to directly ask people about the quality of their electricity (Gibson & Olivia, 2010; Kennedy et al., 2019). While this can provide some indication of where reliability problems exist, large-scale surveys are costly to implement, representative only at a highly-aggregated level, and only reflect a snapshot in time (Min et al., 2017). More importantly, respondents often have difficulty remembering how much and when they had good electricity, even over 24 h. One study found, for

example, that there is only a moderate correlation between people's reported and actual hours with normal electricity (measured by monitors) (Kennedy et al., 2020).

Recently, scholars have turned to satellite data. Although nightlight luminosity has a long history of being used to track electricity access (Baskaran et al., 2015; Min et al., 2013; Min & Gaba, 2014), it is an underexploited tool to study the issue of reliability. An exception is Min et al. (2017) who use high-frequency, nightly, DMSP-OLS nighttime imagery in India from 1993 to 2013 to create a power supply irregularity index and show that the index is relatively consistent with ground-based measures. This approach, however, is not designed to be easily reproduced. The highly disaggregated data is challenging to obtain and very large (about five terabytes), and the standard public distribution is at a yearly composite with a one-year lag. This poses some difficult obstacles for those interested in getting near-real-time estimates of reliability using publicly available data.

Some studies have attempted to use VIIRS nighttime luminosity data to measure reliability within specific areas using smart-meter data. Deng et al. (2019), for example, study the relationship between night light and electric power consumption for a small area in upstate New York. Our approach is most similar to Mann et al. (2016) since they also use Prayas monitor data (for the state of Maharashtra). However, they focus on predicting the rate of power outages at each location (that is, the percentage of days when power fell below 100 V) between Jan 2015 and Sept 2015. Our approach differs in two critical ways. First, we predict a quantity that more directly translates into households' living conditions: the number of hours with normal electricity. Second, we track progress over several years with the hope of informing policy discussions.

Estimating electricity reliability with night lights

To assist policy-makers and NGOs in creating baseline measures of reliability, we rely on data that is widely available and relatively easy to process. We also utilize machine learning methods that have a high degree of interpretability and explainability (regression tree algorithms). In this section, we provide an overview of these data and methods.

Ground truth data

As ground truth data, we use household-level voltage data collected by Prayas, a non-governmental and non-profit organization; the data are available at <https://dataverse.harvard.edu/dataverse/esmi>. As part of their Electricity Supply Monitoring Initiative (ESMI), Prayas installs plug-in devices of about the size of a handheld radio that record electric voltages. The monitors record voltage with an accuracy of plus or minus 4 V on a minute-by-minute basis and transmit an update over mobile networks to a central server every minute. A key advantage of this data, beyond its high frequency, is that it is not subject to systematic omission of load-shedding outages or other censoring. This makes it preferable to many other sources of data on electricity reliability. Although Prayas has installed a total of 528 locations across India, we focus on the state of Uttar Pradesh where 179 monitors exist across 16 districts.

We look at Uttar Pradesh for a few reasons. First, as the most populous area of India, with about 200 million residents, it is outsized in its importance for politics. Uttar Pradesh is also one of the poorest areas of India, with approximately 60 million people living below the poverty line, making it disproportionately important when it comes to studying energy poverty. Second, villages and towns have a wide range of sizes, wealth, and composition, including a significant "scheduled caste" (SC) populations which have been traditionally discriminated against in Indian society. Finally, as noted earlier, Uttar Pradesh is, at least on the surface, a real success story for the government's electrification expansion program. According to official statistics, the state went from

having the fourth lowest electrification level in 2017 to purportedly 100% household electricity connections in 2019.

Fig. SI1 maps the villages where at least one Prayas monitor has been installed. While the locations are not randomly selected, we note that they are widely distributed across the area, covering the extent of the state. Prayas installed a total of 179 monitors in the state. However, at times, monitors fail to record or transmit data (because of various malfunctions), and some locations have also discontinued monitoring and have no data for our time period. After dropping locations without any voltage observations, we are left with 154 unique Prayas monitors located in 96 distinct towns or villages of Uttar Pradesh.

Table SI1 provides summary statistics at the village level for all of the variables used for this study. We note that the sample contains good variation. The average village or town with a Prayas monitor has a population close to 100,000 inhabitants and an area of 243,290 km². The average monthly sum of nighttime per area is about 20, but it varies from 0 to almost 600.

Voltage observations are values between 0 and 260 V for every minute of every hour, for a time period spanning from November 2014 to December 2018 (a total of 50 months). Since the predictive feature in this study is monthly nighttime luminosity, we aggregate voltage observations at the monthly level. We calculate a “monthly voltage” by taking the mean of the measurements available in any given month. We also derive a mean “nighttime” voltage by calculating a similar variable but only considering hours between 8 PM and 6 AM.

Although average monthly voltages offer a precise measure of electricity supply, they are not particularly easy to interpret. Instead, we use the number of hours per day with “normal” electricity as a better, easy-to-understand proxy of reliability and define “normal” as an electric voltage between 205 and 270 V. We define voltage below 130 V as “no voltage”; voltage between 131 and 204 V as “low voltage”, and voltage between 271 and 400 V as “high voltage”. We take these definitions from the Prayas website: www.watchyourpower.org/faqs.php. We also use an alternative measure of electricity as voltage higher than 175 V as a robustness check.

Unfortunately, monitor data is not continuously available for all locations. To illustrate the issue, Fig. SI2 provides a snapshot of the number of hours per day with normal electricity over time for seven different Prayas locations. Fig. SI3 also shows periods of coverage and missing data for twelve locations. Many locations have data for only a few months. Although the average span between the first and last recording is 18 months, voltage data is available only for about six months on average. This emphasizes the continued need for increasing the quality of ground truth data. In the end, our training data contains a total of 679 observations (on average 7 monthly observations for each of the 96 villages/towns).

Our primary variable of interest counts the number of minutes in each day of each month with “normal” voltage. When data is not missing (i.e. when data for all hours of the day is available), our measure reports the average number of hours in a day with the corresponding voltage. In a few cases, voltage recordings are incomplete and several hours are missing.¹ For those missing hours, we replace the missing value with the average number of minutes with normal electricity that this monitor typically has in the same hour of the day.² Table SI2 indicates that, in the average month, the average village or town in the sample has about 12.5 h of normal electricity per day. There is much variation, however, since some places have as little as zero hours and others as much 24.

¹ Only 1.3% of all hours are missing, i.e., whenever voltage data is available for a monitor-day, we observe all hours during that monitor-day.

² There is one monitor where we could not implement this rule because data was available for only one day. We could therefore not use data from other days to infer what the missing values may be. In this case, we use a proportional rule: we calculated the percentage of minutes with “normal” electricity for the observed hours and assume the same percentage would have occurred for the other hours of the day that we don't observe.

Nightlights data

For nightlights data, we use data from the Visible Infrared Imaging Radiometer Suite (VIIRS) instrument, located on the Suomi National Polar-orbiting Partnership (NPP) satellite (Miller et al., 2013; NASA, 2020). VIIRS provides global coverage for important environmental products, including imagery of daytime solar and nocturnal lunar reflectance in addition to nighttime natural and artificial visible and near-infrared lights. Specifically, we use the VIIRS Nighttime Day/Night Band Composites product, which is processed and provided by the Earth Observatory Group at the Colorado School of Mines, accessed through the Google Earth Engine (2022).³ This data is easily accessible to policy-makers, and because of its smaller size and professional hosting on Google Earth Engine, it can be much more easily analyzed.

Scholars have noted a number of issues with this data, such as biophysical seasonality, significant inter-temporal variability, and others (Levin, 2017; Levin & Zhang, 2017). Some have also raised the issue that the measurements tend to be taken very early in the morning (between midnight and 2 AM), which do not represent peak use hours (Mann et al., 2016). With all of this said, previous research has demonstrated that models using this data can be quite accurate, even at the pixel or building level (Deng et al., 2019; Mann et al., 2016). Google Earth Engine (2022) also processes the data to, for example, account for cloud cover using the VIIRS Cloud Mask product (VCM).

Monthly average radiance composite images were collected from January 2012 to 2019 for all 98,082 villages and 908 towns of the state of Uttar Pradesh. Following the findings of Dugoua et al. (2018), we use the logged sum of the digital number (DN) of luminosity within the area of the target village. The area of the village is drawn from shapefiles constructed by ML Infomap, a private corporation that provides village-level shapefiles for India. Importantly, the data is filtered to exclude observations impacted by stray light, lightning, lunar illumination, and cloud cover before averaging GEE.

A disadvantage of using the VIIRS DNP data is that it is monthly, compared to the daily data used by Min et al. (2017), and, as a result, it may not detect short-term fluctuations. However, since our goal is to construct an approximate measure of reliability for given villages, monthly aggregate should be sufficient. The higher the reliability of electricity in a particular area, the higher the number of hours at night that will appear as “luminous”. The more of such nights during a particular month, the higher the value in the monthly composite nightlight data.

Demographic data

The satellite data is combined with information about the villages and towns over which the digital number (DN) of luminosity is calculated. To measure these characteristics of villages and towns, we use the 2011 Census of India. We include four specific variables which were found to substantially impact the accuracy of our results. First, we account for the total population and households in the region since this helps account for the potential number of electricity consumers. Second, we look at the level of literacy in the village as a method for accounting for the affluence of the village. Finally, we use the area of the village to take into account the size of the landmass over which the sum of DN in the raster data is summed.

Analysis

The analysis is conducted using a regression tree method. This machine learning method functions by partitioning the data at a “parent” node into two “child” nodes. The split is done based on information gains (IG), which are defined as:

³ https://developers.google.com/earth-engine/datasets/catalog/NOAA_VIIRS_DNB_MONTHLY_V1_VCMSLCFG.

$$IG(D_p, f) = I(D_p) - \left(\frac{N_{left}}{N_p} I(D_{left}) + \frac{N_{right}}{N_p} I(D_{right}) \right) \quad (1)$$

Here f is the feature on which the split is performed, I is the impurity measure, N_p is the total number of samples at the parent node, and N_{left} and N_{right} are the number of samples in the child nodes. For regression problems, which attempt to model continuous variables, the impurity measure used is the weighted mean squared error (MSE), defined as:

$$MSE(t) = \frac{1}{N_t} \sum (y_i - \hat{y}_t)^2 \quad (2)$$

Here, N_t is the number of cases at node t , y_i is the true target value for cases in node t , and $\hat{y}_t$ is the predicted value (the mean for the cases in the child node).

Regression trees increase their performance by adding layers that split child nodes into additional child nodes until such splits cease to yield additional IG. This strategy allows the modeling of complex relationships that include discontinuities and interactions, which we expect in this data. In addition to their flexibility, regression trees are reasonably easy to implement and interpret, making them especially well-suited to our goal of demonstrating a system that can be utilized by policy-makers who may lack resources and expertise for more complex models (Friedman et al., 2001). The flexibility of regression trees can come at the cost of overfitting the training set, introducing bias, and jeopardizing the generalizability and predictive performance on unseen data. For this reason, trees must be “pruned” (Friedman et al., 2001; Witten et al., 2011).

We estimate the optimal number of layers using the training set (i.e., observations for which we have both nightlight and voltage data) using a five-fold cross-validation procedure. The exercise consists of dividing the training data into five equal parts and fitting a model with a number of layers equal to depth l using only four of the splits (out of the five). The mean absolute error (MAE) of the model is then calculated on the fifth split (which was withheld). The procedure is followed for each of the five splits and for as many layers l desired.

We repeat the exercise 1000 times, each time randomly splitting the sample in five, and report the average MAE for all procedures. We find that a tree with ten layers provides optimal error minimization (see Fig. S15), and therefore the best balance between bias and variance in the decision tree model. The ten-layer decision tree is relatively large, with 511 child nodes reflecting the suspected nonlinear and interactive nature of the learning problem we face.

Estimating electricity reliability

We start with examining the data from ESMI monitors as we will use it to construct our ground truth measure of reliability. First, we investigate whether the number of hours with normal electricity has increased over time for the locations for which we have monitor data. As illustrated on Fig. 1, we find that electricity supply became more reliable over time only for a few locations. For most, there was no substantial improvement in reliability.

To provide preliminary support for our approach, we check whether there is any clear correlation between monthly aggregates of nighttime luminosity and the average hours per day of normal electricity (averaged at the monthly level), absent additional socio-demographic information about the villages and towns. In Fig. 2, we see that, indeed, the two variables have a nonlinear relationship, with a particularly strong and positive slope after 10 h per day.

Next, we train a regression tree using eight different features: log of the sum of nightlights, area, log of population, log of the total number of households, scheduled castes population, schedule tribes population, and literate population. We also add polynomials of the log of the sum of nightlights up to the 9th degree. The optimal number of layers in

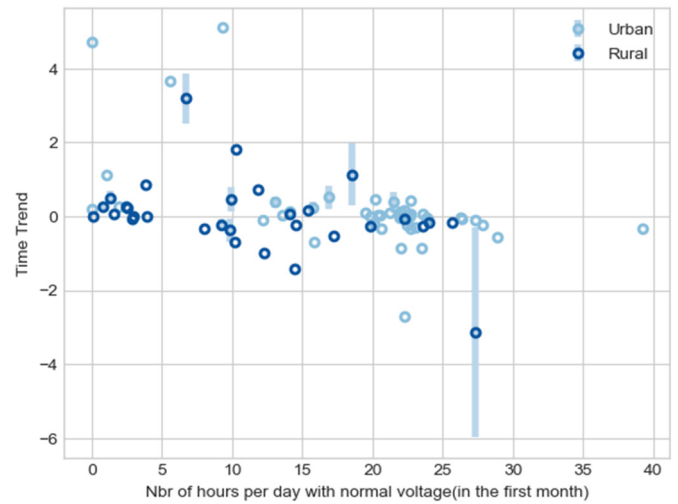


Fig. 1. Electricity supply measured at the Prayas monitors.

Note: Each dot represents a monitor. The x-axis reports the number of hours per day with normal electricity in the first month for which monitor data is available. The y-axis reports the slope resulting from a linear regression of time (measured in months) on the number of hours per day with normal electricity. A positive slope indicates that electricity became more reliable for that location over time in our sample. When the slope is statistically significant from zero (at the 10% level), we add error bars around the dot. All points without error bars are, therefore, not statistically significant from zero.

our training data is estimated using five-fold cross-validation, repeated 1000 times. This results in a ten-layer tree.

To provide more intuition regarding the performance of the model, Fig. 3 describes the error terms obtained for one specific cross-validation split. The prediction here is out-of-sample because the model is trained using only four splits of the data, and predictions are

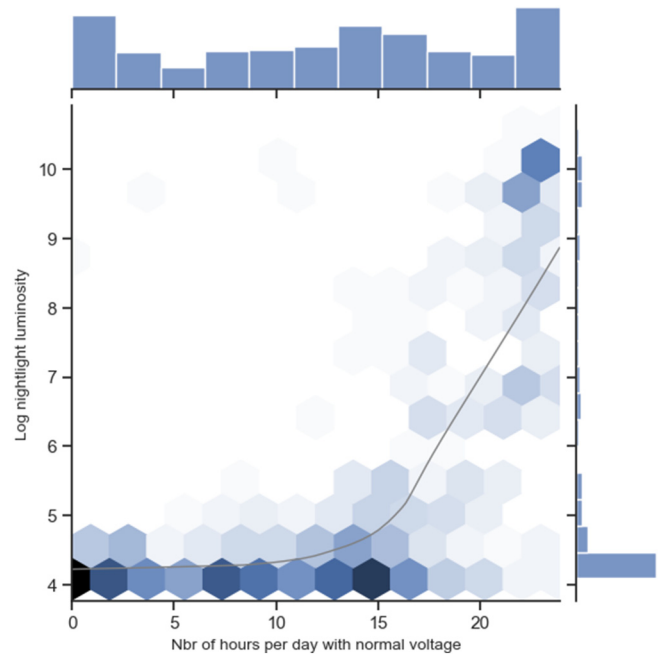


Fig. 2. Relationship between nightlights and average hours per day with normal electricity.

Note: We define “normal” electricity as an electric voltage between 205 and 270 V. Voltage data is sourced from 179 monitors in Uttar Pradesh installed by a non-governmental organization called Prayas. The nightlight luminosity variable corresponds to the sum of the monthly average luminosity of all pixels within the area of the village or town where the monitor was installed. The intensity of color in each hexagonal bin represents the number of village-month observations that fall within the range of the shape.

generated on the last split. On the Panel (a), we plot the Prayas voltage (ground truth) together with the predicted value against nighttime luminosity. We note that the predictions are close to the ground truth observations. They also present a distinctly curvilinear relationship between nightlights and the number of hours per day with electricity: hours increase rapidly at the low ends of nighttime luminosity and level out at higher levels.

Supplementary Fig. SI6 provides a depiction of the first three layers of the model. We find that the first decision rule assesses whether the value of the sixth polynomial of log sum of nightlights is lower than 15,327.989. If so, which is the case for 404 observations, the tree path goes up to the next decision node, which evaluates whether the 4th polynomial is lower than 374.239. Unsurprisingly, we see that the first three layers use the most important features of the model: variables related to nightlights, area, schedule castes population, and literate population.

We also find that nightlight variables are the most important ones for prediction performance. Combining all the polynomials, the sum of nightlight reduces the mean absolute error by about 64%. The following most important features are area (13%), literate population (8%), the total number of households (5%), and the population of scheduled castes (4%). The least important feature is the dummy variable for rural or urban, which does not contribute at all in reducing the mean absolute error.

Panel (b) on Fig. 3 shows that the errors are close to being normally distributed, with a mean at -0.1 and a standard deviation of 4.3 . There is some left-skew in the distribution, suggesting some areas have relatively high predicted reliability but actually low hours of electricity per day. These larger errors, however, seem relatively rare. In sum, the results suggest that the model does a relatively good job of capturing the variance in reliability using relatively static measures of nighttime luminosity combined with structural information about the villages.

Extrapolating results to unmonitored regions

We leverage our predictive model of electricity reliability to provide a more comprehensive picture of trends in reliability over time in the whole state of Uttar Pradesh, beyond areas for which ground truth data exists. To obtain uncertainty bounds around our predictions, we implement a bootstrap procedure, a standard method for estimating

out-of-sample errors (Friedman et al., 2001). Bootstrapping is a fundamental tool in computational statistics for estimating error bounds in models without a closed solution for generating such measures and without relying on strong parametric assumptions. The procedure consists in creating 1000 training samples of the same size as the original training set by sampling with replacement from the original training sample and re-estimating the model for each bootstrapped sample. Regression trees are then trained on each of these samples, such that we obtain 1000 different trees, and therefore 1000 different predictions for each unit. For any particular village or town, the spread of the bootstrap predictions is typically high, with, for example, confidence intervals at the 90% level as wide as 10 h (see Supplementary Fig. SI9). Therefore, a limitation of our approach is that we cannot precisely predict the reliability at any particular location.

We can, however, predict average reliability for Uttar Pradesh as a whole with relatively good precision. Fig. 4 illustrates the distribution of the bootstrap predictions for the average Uttar Pradesh. In Fig. 4 Panel (a), the 90% confidence interval is indicated by the whiskers and spans about 1.5 h. The distribution shifts upward as time goes by, with a particular shift after 2017. This possibly reflects the massive government-led efforts of electrifying villages that used to be off-grid. Fig. 4 Panel (b) further indicates that the change in the average number of hours with normal electricity between 2014 and 2019 for Uttar Pradesh is between 0.2 and 1 additional hour, with 90% confidence (median: 0.6). Although positive, this remains a very small improvement. In 2014, the median bootstrap prediction for reliability in the average location in Uttar Pradesh corresponds to 7.55 h of normal electricity per day; for 2019, it increases to 8.14 h, but that still leaves most of the hours of a day with bad or no electricity.

Although we can't accurately predict reliability in specific locations, our approach is still valuable for discovering general trends and patterns at the village or town level. Fig. 5 Panel (a) maps the median predictions for the month of January 2014 for all locations in Uttar Pradesh. Areas colored in more intense red indicate higher numbers of hours with normal electricity per day. Throughout Uttar Pradesh, there are concentrated areas of darker red, indicating more than 15 h per day of electrification. While far from 24-hour access, this is still quite an achievement, given the extent of the electrification campaigns. Nonetheless, we also note that these red areas are usually surrounded by larger, lighter red zones, which indicate less than 11 h of electricity

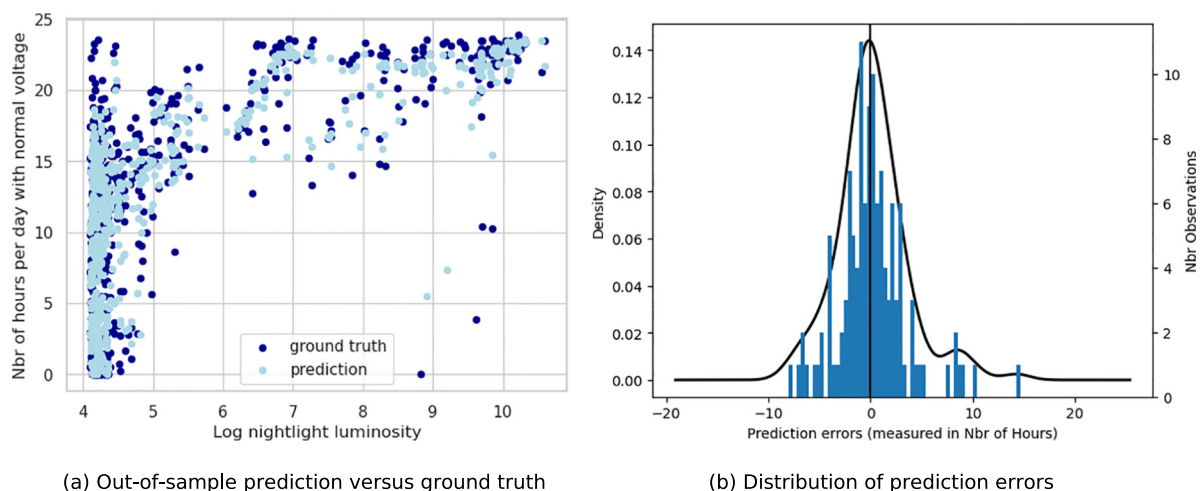


Fig. 3. Performance of the regression tree model.

Note: We train a regression tree using eight different features, including polynomials of the log sum of nightlights and village-level characteristics. The optimal number of layers in our training data is estimated using five-fold cross-validation, repeated 1000 times. Panel (a) plots the out-of-sample predictions against the ground truth values for one of the these cross-validation splits. The prediction here is out-of-sample because the model is trained using only four splits of the data, and predictions are generated on the last split. Panel (b) shows that the distribution of errors (defined as the ground-truth number of hours per day with normal voltage minus the predicted number) across all 1000 splits. The figure shows that errors are close to being normally distributed, with a mean at -0.1 and a standard deviation of 4.3 .

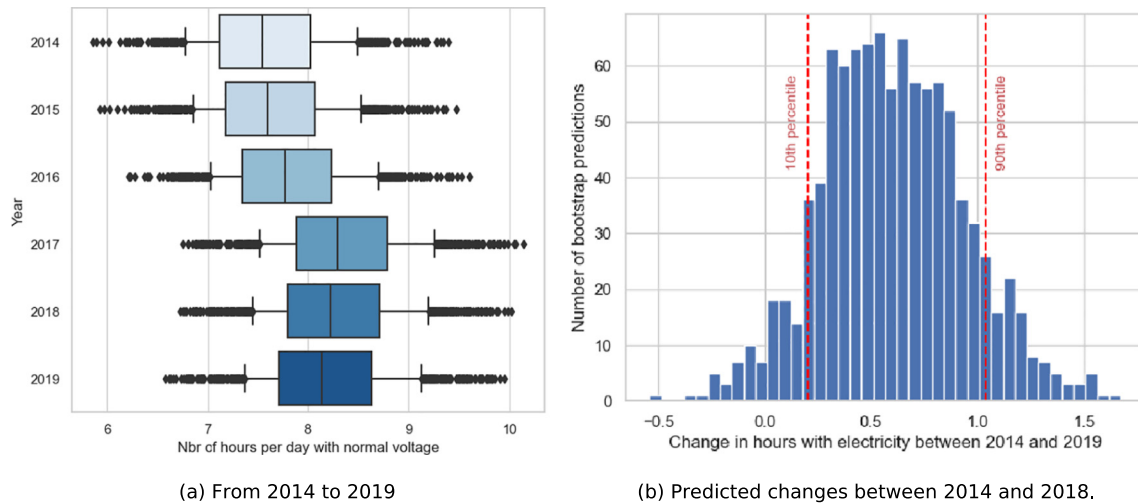


Fig. 4. Uttar Pradesh average availability predictions.

Note: To obtain uncertainty bounds around our predictions, we implement a bootstrap procedure where regression trees are trained on 1000 different bootstrap samples. Panel (a) illustrates the distribution of the bootstrap predictions for average hours of normal electricity in Uttar Pradesh across years in the bootstrapped estimates. The 90% confidence interval is indicated by the whiskers and spans about 1.5 h. The distribution shifts upward as time goes by, with a particular shift after 2017. Panel (b) indicates that the change in the number of hours with normal electricity between 2014 and 2019 for Uttar Pradesh on average is between 0.2 and 1 additional hour, with 90% confidence.

access. There are also substantial islands of white, indicating less than 4 h of access. In these areas, households are unlikely to experience much welfare improvement from the grid connection they received.

Comparing the predictions for 2014 and 2019, we can examine which locations have gained the most. In Fig. 5 Panel (b), the shape of the cloud of points indicates that most locations are likely to have seen a small improvement, and those which started with worse electricity in 2014 seem to observe more progress. However, the local regression line on the graph also reveals a high mass of locations close to zero that seem to have seen no improvement. This stands in comparison

to the scale of expansion that has taken place in terms of connections, albeit not in reliability.

Next, we investigate the relationship between the predicted change in reliability between 2014 and 2019 and location characteristics. We find a strong, positive relationship for population density and the total number of households, indicating that areas with larger populations and higher density have gained the most reliability (see Supplementary Fig. S111). This is consistent with utilities prioritizing areas where demand is high to recoup the high fixed costs associated with infrastructure development. On the other hand, we find a relatively small and

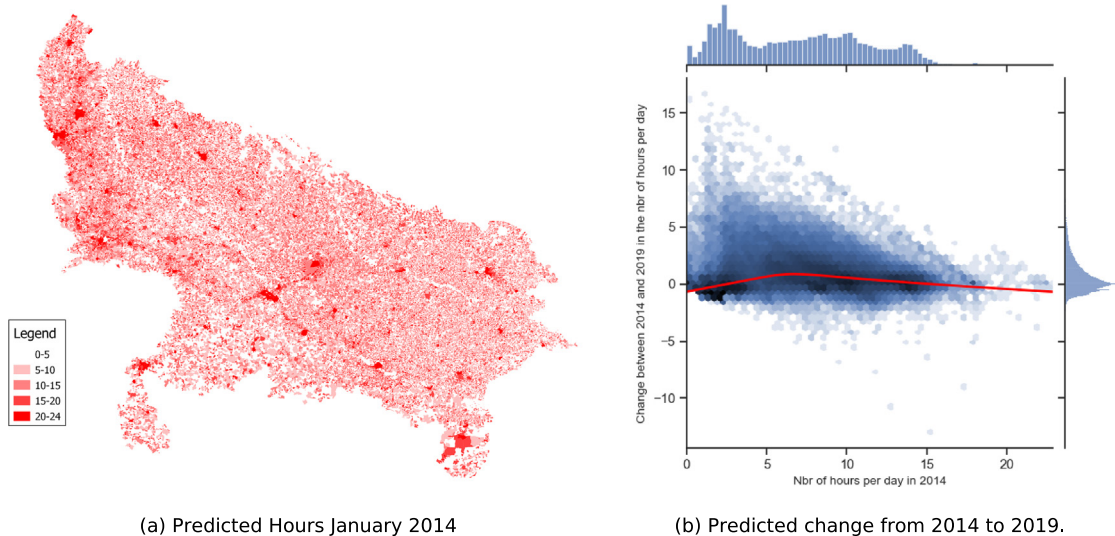


Fig. 5. Predicted reliability of all locations in Uttar Pradesh.

Note: Panel (a) maps the median predictions for the month of January 2014 for all locations in Uttar Pradesh. Areas colored in more intense red indicate higher numbers of hours with normal electricity per day following natural breaks (junks) in the data distribution. Panel (b) plots the median predicted change for each location from 2014 to 2019 as a function of the median predicted number of hours in 2014. In other words, for each village, we compute the median of bootstrap predictions for the entire year of 2014, repeat the same for 2019, and then simply subtract the two median values. The shape of the cloud of points indicates that most locations are likely to have seen a small improvement, and those which started with worse electricity in 2014 seem to observe more progress. Yet, the local regression line also indicates that there is a high mass of locations close to zero that seem to have seen no progress at all.

curvilinear relationship for literacy or the proportion of the population from a disadvantaged community.

Conclusion and policy implications

While connecting households to national electricity grids is an essential first step towards universal electrification, improving reliability is critical to ensure that customers can benefit from it. When poor reliability is combined with the monthly fees imposed by utilities to keep households connected to the grid, customers may find themselves better off by sticking with or returning to traditional energy sources like kerosene or diesel generators, furthering the problem of low demand that has plagued rural electrification in developing countries.

One reason the expansion of grid connections has been so successful in India is that it has been relatively easy to set concrete goals and track progress. Similarly, monitoring reliability will be essential to keep the issue salient in the public debate and mobilize public opinion. Unfortunately, utilities have little incentive to release data that would document the poor reliability of the electricity they sell. Extending electricity connections to poorer, more rural areas have already placed their resources under strain. Publicizing problems in deliveries, especially poorly planned load-shedding, is only likely to encourage further requests for higher investments into areas with low purchasing power.

Incentives are also weak for politicians to enforce accurate and public reporting of power outages. In fact, previous studies have suggested that they strategically ignore practices that contribute to poor reliability, like electricity theft, to improve their electoral prospects (Baskaran et al., 2015; Min & Golden, 2014). Publicly available data from official sources is therefore often unavailable or not reliable. Our results provide strong evidence for the utility of openly available and easily accessible remote sensing tools to monitor electricity reliability. This, in turn, can provide international organizations, domestic activists, and NGOs with critical information to advocate for better electricity supply.

This paper, in particular, shows how much progress has been made but also how much still remains to be seen regarding the quality of electricity. Although our predictive model provides comforting evidence of its reliability and performance through cross-validation and bootstrapping, we shall mention a few important caveats. First, for areas without ground truth data, one cannot guarantee that results are accurate. Moreover, it is not possible to exclude the possibility that the observations used to train the model are not representative of the rest of Uttar Pradesh, introducing some bias and jeopardizing the generalization we are attempting.

Future work could extend the approach used in this paper in several ways. First, to the extent that ground truth data is collected in other areas of Uttar Pradesh, the training sample would increase, which could lead to more precise predictions at the location level. Second, additional ground-truth data also offer the opportunity to implement an actual out-of-sample test and help confirm (or infirm) the usefulness of relatively highly aggregated nighttime luminosity data to provide guidance on electricity reliability. Third, while this paper uses data aggregates that are more easily accessed, stored and analyzed by resource-constrained actors, advances in access to more fine-grained daily composites may lead to further advancements.

Similarly, extending the exercise to other countries can help validate or invalidate the approach in a broader set of contexts. A systematic, more representative placement of monitors over large areas by governments or NGOs would greatly assist in leveraging existing data tools for monitoring electricity reliability, as well as providing directly useful data themselves. Since many developing countries are rapidly expanding their electricity supply and facing the same difficulties with monitoring reliability, satellite data may provide a key element in monitoring access.

Future studies will likely benefit from using various tools and data, such as official data, surveys, and satellite imagery. This paper contributes by documenting the performance of one such tool using easily

accessible data and a transparent method of analysis. Therefore, it provides some hope that, despite the difficulties in measuring electricity reliability, NGOs and researchers can develop monitoring tools to inform policymaking and hold government and utilities accountable at a reasonably low cost in terms of resources and expertise. Such measurement is a first step towards creating incentive for greater reliability, even in lower income areas, and realizing the full potential of national electrification.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary information

Supplementary information to this article can be found online at <https://doi.org/10.1016/j.esd.2022.04.004>.

References

- Aklin, Michael, Blankenship, Brian, Nandan, Vagisha, & Urpelainen, Johannes (2020). *Energy in rural Jharkhand. Technical report*. ISEP and Oak Foundation.
- Aklin, Michael, Cheng, Chao-Yo, Urpelainen, Johannes, Ganesan, Karthik, & Jain, Abhishek (2016). Factors affecting household satisfaction with electricity supply in rural India. *Nature Energy*, 1, 16170.
- Allcott, Hunt, Collard-Wexler, Allan, & O'Connell, Stephen D. (2016). How do electricity shortages affect industry? Evidence from India. *American Economic Review*, 106(3), 587–624.
- Baskaran, Thushyanthan, Min, Brian, & Uppal, Yogesh (2015). Election cycles and electricity provision: Evidence from a Quasi-experiment with Indian special elections. *Journal of Public Economics*, 126, 64–73.
- Bhatia, Mikul, & Angelou, Nicolina (2014). Capturing the multi-dimensionality of energy access. *Livewire*. World Bank.
- Blankenship, Brian, Kennedy, Ryan, Mahajan, Aseem, Wong, Jason Chun Yu, & Urpelainen, Johannes (2020). Increasing rural electrification through connection campaigns. *Energy Policy*, 139, Article 111291.
- Chakravorty, Ujjayant, Pelli, Martino, & Marchand, Beyza Ural (2014). Does the quality of electricity matter? Evidence from rural India. *Journal of Economic Behavior and Organization*, 107, 228–247.
- Deng, Chengbin, Lin, Weiying, & Chen, Shili (2019). Use of smart meter readings and nighttime light images to track pixel-level electricity consumption. *Remote Sensing Letters*, 10(3), 205–213.
- Dugoua, Eugénie, Kennedy, Ryan, & Urpelainen, Johannes (2018). Satellite data for the social sciences: measuring rural electrification with night-time lights. *International Journal of Remote Sensing*, 39(9), 2690–2701.
- Fairley, Peter (2019). A power line to every home: India closes in on universal electrification. *IEEE Spectrum*. <https://spectrum.ieee.org/energy/policy/a-power-line-to-everyhome-india-closes-in-on-universal-electrification>.
- Friedman, Jerome, Hastie, Trevor, & Tibshirani, Robert (2001). *The elements of statistical learning*, Vol. 1. (pp. 10). Springer.
- Gibson, John, & Olivia, Susan (2010). The effect of infrastructure access and quality on non-farm enterprises in rural Indonesia. *World Development*, 38(5), 717–726.
- Google Earth Engine (2022). *Earth Engine Data Catalog: VIIRS Nighttime Day/Night Band Composites Version 1*. <https://developers.google.com/earthengine/datasets/catalog/NOAAVIIRSDNBMONTHLYVIVCMCFG> (Accessed 10 January 2021).
- Graber, Sachiko, Narayanan, Tara, Alfaro, Jose, & Palit, Debajit (2018). Solar microgrids in rural India: Consumers' willingness to pay for attributes of electricity. *Energy for Sustainable Development*, 42, 32–43.
- Grimm, Michael, Lenz, Luciane, Peters, Jörg, & Sievert, Maximiliane (2020). Demand for off-grid solar electricity: Experimental evidence from Rwanda. *Journal of the Association of Environmental and Resource Economists*, 7(3), 417–454.
- Kemmler, Andreas (2007). Factors influencing household access to electricity in India. *Energy for Sustainable Development*, 11(4), 13–20.
- Kennedy, Ryan, Mahajan, Aseem, & Urpelainen, Johannes (2019). Quality of service predicts willingness to pay for household electricity connections in rural India. *Energy Policy*, 129, 319–326.
- Kennedy, Ryan, Mahajan, Aseem, & Urpelainen, Johannes (2020). Crowdsourcing data on the reliability of electricity service: Evidence from a telephone survey in Uttar Pradesh, India. *Energy Policy*, 145, Article 111746.
- Lee, Kenneth, Miguel, Edward, & Wolfram, Catherine (2017). Electrification and economic development: a microeconomic perspective. *EEG State-of-Knowledge Paper Series*. 1. (pp. 2).
- Levin, Noam (2017). The impact of seasonal changes on observed nighttime brightness from 2014 to 2015 monthly VIIRS DNB composites. *Remote Sensing of Environment*, 193, 150–164.
- Levin, Noam, & Zhang, Qingling (2017). A global analysis of factors controlling VIIRS nighttime light levels from densely populated areas. *Remote Sensing of Environment*, 190, 366–382.

- Mann, Michael L., Melaas, Eli K., & Malik, Arun (2016). Using VIIRS day/night band to measure electricity supply reliability: Preliminary results from Maharashtra. *India Remote Sensing*, 8(9), 711.
- Miller, Steven D., Straka, William, Mills, Stephen P., Elvidge, Christopher D., Lee, Thomas F., Solbrig, Jeremy, & Weiss, Stephanie C. (2013). Illuminating the capabilities of the Suomi National Polar-Orbiting Partnership (NPP) visible infrared imaging radiometer suite (VIIRS) day/night band. *Remote Sensing*, 5(12), 6717–6766.
- Min, Brian, & Gaba, Kwawu Mensan (2014). Tracking electrification in Vietnam using nighttime lights. *Remote Sensing*, 6(10), 9511–9529.
- Min, Brian, Gaba, Kwawu Mensan, Sarr, Ousmane Fall, & Agalassou, Alassane (2013). Detection of rural electrification in Africa using DMSP-OLS night lights imagery. *International Journal of Remote Sensing*, 34(22), 8118–8141.
- Min, Brian, & Golden, Miriam (2014). Electoral cycles in electricity losses in India. *Energy Policy*, 65, 619–625.
- Min, Brian, O'Keeffe, Zachary, & Zhang, Fan (2017). Whose power gets cut? Using high-frequency satellite images to measure power supply irregularity. *World Bank policy research working paper*.
- NASA (2020). Visible infrared imaging radiometer suite (VIIRS). <http://earthdata.nasa.gov/earth-observation-data/near-real-time/download-nrt-data/viirs-nrt> (Accessed 10 January 2021).
- Samad, Hussain, & Zhang, Fan (2016). Benefits of electrification and the role of reliability: Evidence from India. *World Bank policy research working paper*.
- Samad, Hussain, & Zhang, Fan (2017). Heterogeneous effects of rural electrification: Evidence from Bangladesh. *World Bank policy research paper*.
- Saubhagya Initiative (2021). <https://saubhagya.gov.in/> (Accessed 10 January 2021).
- Sengupta, Debjoy (2017). Power deficit disparity among Indian States. *Economic times*. <https://economictimes.indiatimes.com/industry/energy/power/wide-power-deficitdiscrepancy-between-ngo-and-govt-data-in-uttar-pradesh/articleshow/57311992.cms?from=mdr>.
- The Economist (2018). Daily chart: India invests in electrification. <https://www.economist.com/graphic-detail/2018/11/14/india-invests-in-electrification> (Accessed 10 January 2021).
- Urpelainen, Johannes (2019). Saubhagya scheme: Millions remain without electricity. *Power for all*. <https://www.powerforall.org/countries/india/universal-rural-electrificationindia-not-so-fast>.
- Witten, Ian H., Franke, Eibe, & Hall, Mark A. (2011). *Data Mining* (Third). Morgan Kaufmann.