Automating Electricity Access Prediction with Satellite Imagery

Fangge Deng, Shamik Hossain, Prithvir Jhaveri, Ashley Meuser, Harshvardhan Sanghi, Joe Squillace, Anuj Thakkar, Brian Wong, Xiaolan You
Faculty Advisors: Dr. Kyle Bradbury (Energy Initiative), Dr. Leslie Collins (Pratt), Dr. T. Robert Fetter (Nicholas Institute), Dr. Marc Jeuland (Sanford), Dr. Timothy Johnson (Nicholas)

Introduction & Overview

Energy access is correlated with improvements in the wellbeing, economic prosperity, and gender equality of a region. Particularly, it is linked to an increase in the number of students enrolled in school, time students spend studying, business hours, agricultural productivity and labor supply, and a reduction of the poverty rate (Khandker, et al., 2012).

Despite these benefits, an estimated 1.2 billion people do not have electricity access, and more have too unreliable electricity to achieve the aforementioned welfare gains (World Energy Outlook, 2017).

This study aims to fill current data gaps on global energy access assessment through producing high resolution geographic energy access metrics. We seek to overcome data unavailability and inaccuracies in existing data by creating a method for continuously monitoring electricity access over time, and to produce higher resolution estimates of electricity access data at the village-level.

Process Summary

Data Collection
Curate dataset of village electrification data and corresponding satellite imagery

Feature Extraction
Extract features from lights at night data for each village to be input into classifier

Classification Models
Run a gradient boosted decision tree classifier to predict whether each village is electrified and evaluate performance using cross validation

Output
Classify electrification status for Bihar, India and produce high-resolution maps of electrification

Results

We demonstrate the performance of our classifier using Receiver Operating Characteristic (ROC) curves in Figure 5. Since smaller villages may not always have sufficient light visible at night to register on the VIIRS instrument (Min and Gaba 2014), we also explore the discriminative abilities of our classifier limited to villages with at least 100 or 400 households, also demonstrating better performance in classifying the electrification rate of larger villages than smaller villages.

Conclusion & Future Steps

This study confirms that lights at night data can be used to estimate village electrification status and quantified the cross-validated performance of our classifier. We also found that villages with larger populations were more accurately classified than villages with smaller populations, since the difference between larger electrified villages and unelectrified villages is much more visible in the lights at night imagery data. In the future, additional features extracted from satellite imagery will be added to explore potential classification performance improvements using information such as vegetation and rainfall data for identifying electrified irrigation, built environment detection (buildings and roads), and other energy access indicators.

Sources


Figure 1: Global population with access to no or inadequate electricity

Figure 2: Energy poverty in East Asia India relative to developed nations and the G8 average.

Figure 3: Process of data collection, feature extraction, village electrification classification and output validation.

Figure 4. Distribution of top four features of importance separated by Electrified and Unelectrified classes.

Figure 5. ROC Curve demonstrating results of energy access projections separated by three models, each by selecting a minimum number of households threshold.

Energy Poverty in East Asia

Our team classified VIIRS data from 16,389 villages in Bihar as either electrified or unelectrified based on ground truth data from the Indian government’s Garv dataset. Here we assume that an electrified village is one where at least 10% of households are electrified (Min and Gaba, 2014). We extracted the lights at night data within each village boundary and for each village calculated the mean, max, and sum radiance values as well as the 10th, 25th, 50th, 75th, and 90th radiance percentiles. We used these values as features to train our classifier to predict the electrification status of each village. We used a gradient boosted decision tree classifier and cross validated with testing data.