Automating Electricity Access Prediction with Satellite Imagery

Duke UNIVERSITY | BASS CONNECTIONS

Fangge Deng, Shamikh 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)

Duke ENERGY INITIATIVE

Introduction & Overview

Energy access is a key resource in improving 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, S.R. et al., 2012).

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

This study aims to fill current data gaps on global energy access, particular in resolving finer-scale geographic access metrics. The study aims to overcome inaccurate or biased data, to provide a method for continuously measuring progress in electricity access over time, and to get more refined electricity access data on a village-to-village basis.

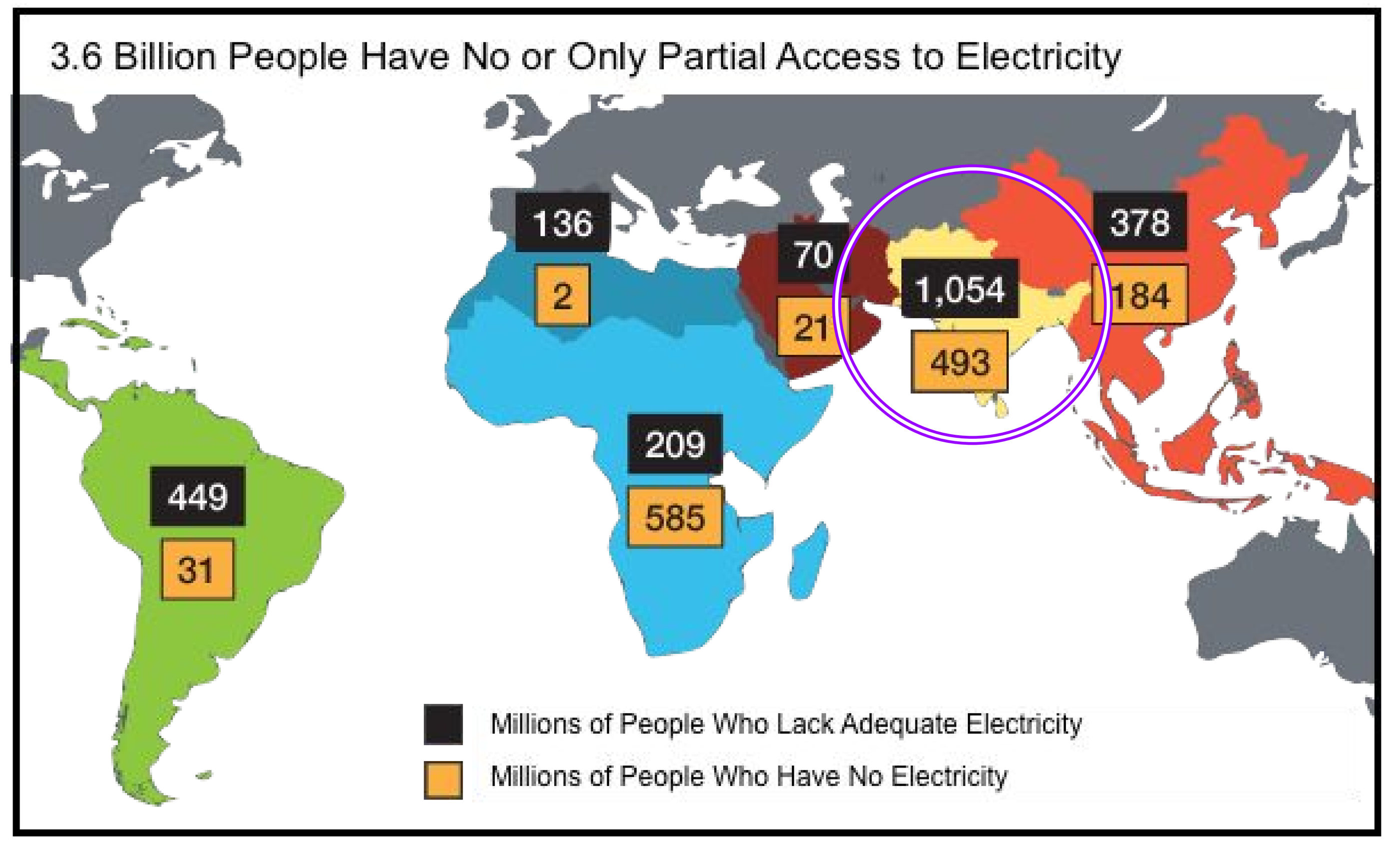

Figure 1: Global population with access to no or inadequate electricity

Source: International Energy Agency World Energy Outlook 2011 and The World Bank World Development Indicators 2011

The primary deliverable of this preliminary study is to produce a functional collaborative machine learning infrastructure with VIIRS Lights at Night data capable of predicting electrification rates at the village level in Bihar, India. In doing so, the study aims to advance energy access mapping at finer resolutions (Shi et al., 2014).

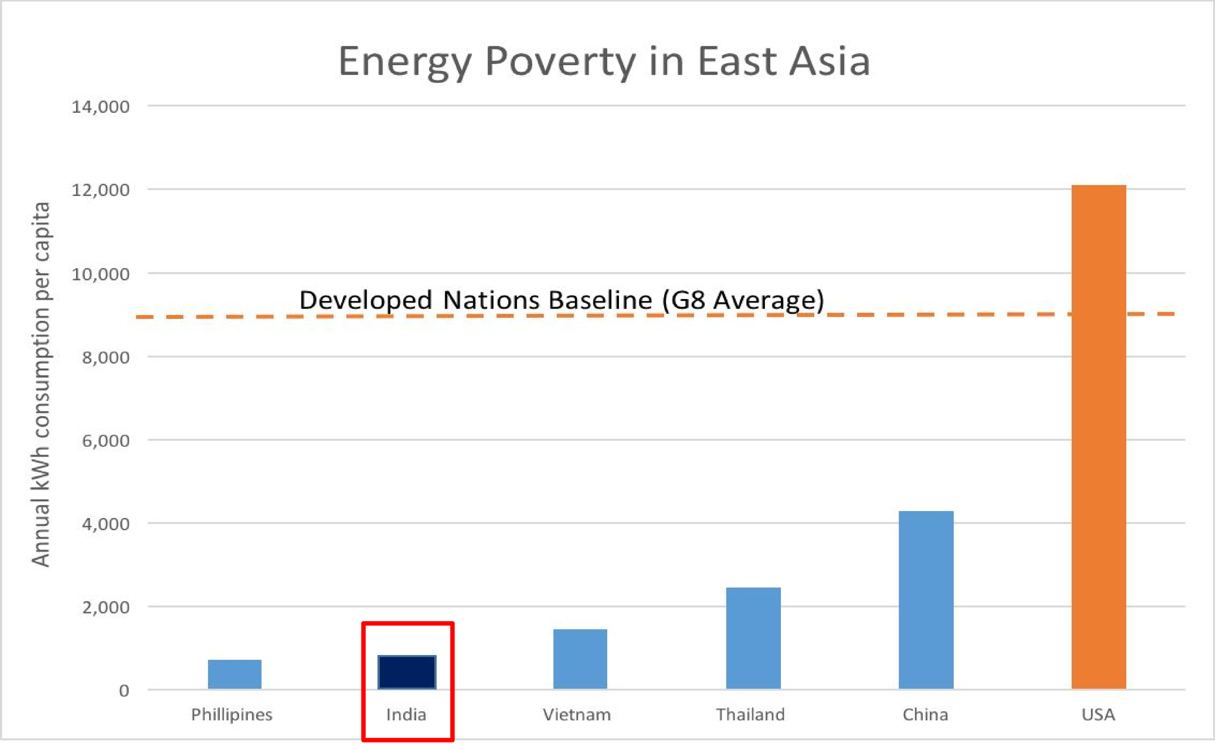

Figure 2. Energy poverty in East Asia India relative to developed nations and the G8 average.

Process Summary

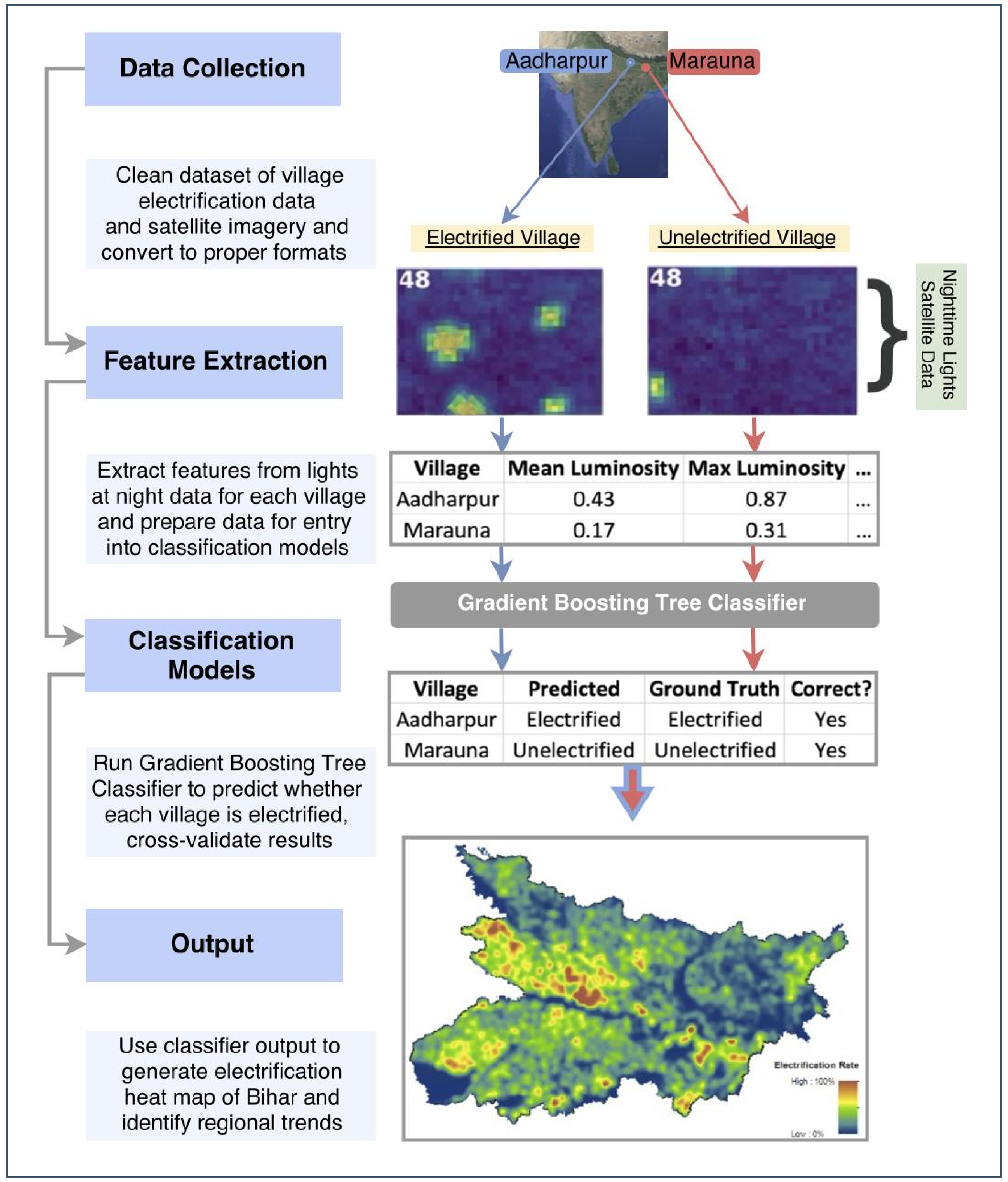

Figure 3. Process of data collection, feature extraction, village electrification classification and output validation.

Our team performed operations on VIIRs band arrays that correspond to villages in Bihar. We first masked each array to include only pixels that are within the boundaries of their corresponding village. From these arrays we calculated the mean, max, and sum radiance values as well as the 10th, 25th, 50th, 75th, and 90th radiance percentiles. We input these values as features to train our classifier to predict for each village whether it belongs to the "electrified" class or the "unelectrified" class, for which the threshold for being considered electrified was if 10% of the households have electricity access. Finally, we produced an electrification heat map of Bihar from the classifier's predictions for all villages.

Results

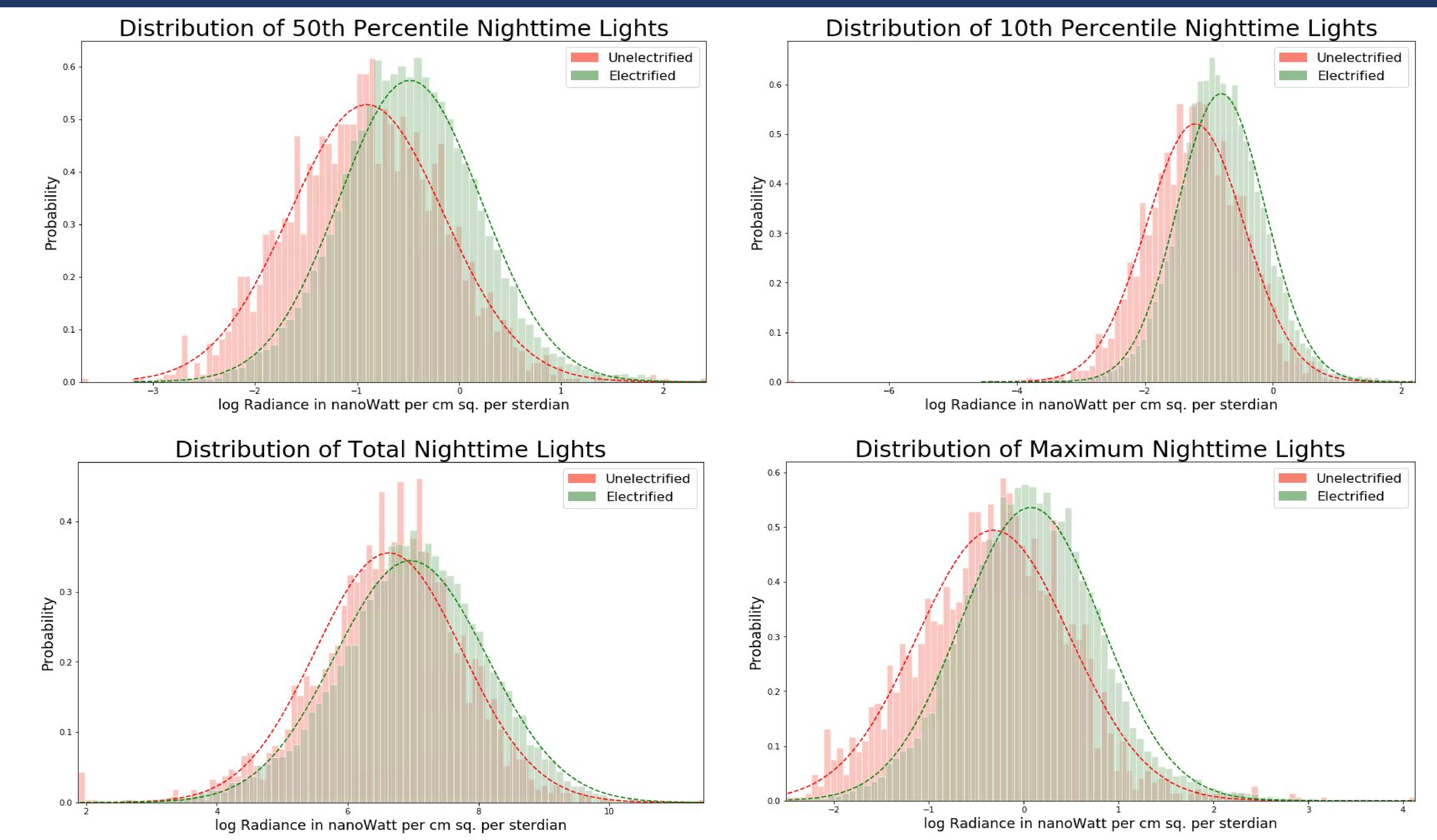

Figure 4. Distribution of top four features of importance separated by Electrified and Unelectrified classes.

Machine learning classifiers were used to predict village data into two categories: un-electrified and electrified. The classifier identified villages with less than 10% access as unelectrified and villages with larger values as electrified, thresholds that were based on the literature (Min and Gaba, 2014). The results of the Gradient Boosting Classifiers are shown to right in the form of Receiver Operating Characteristic (ROC) curves in Figure 5. Min and Gaba 2014 identified difficulties for nighttime lights data to detect small villages, even if electrified, thus three thresholds (minimum number of households in a village) were selected to pre-filter the data.

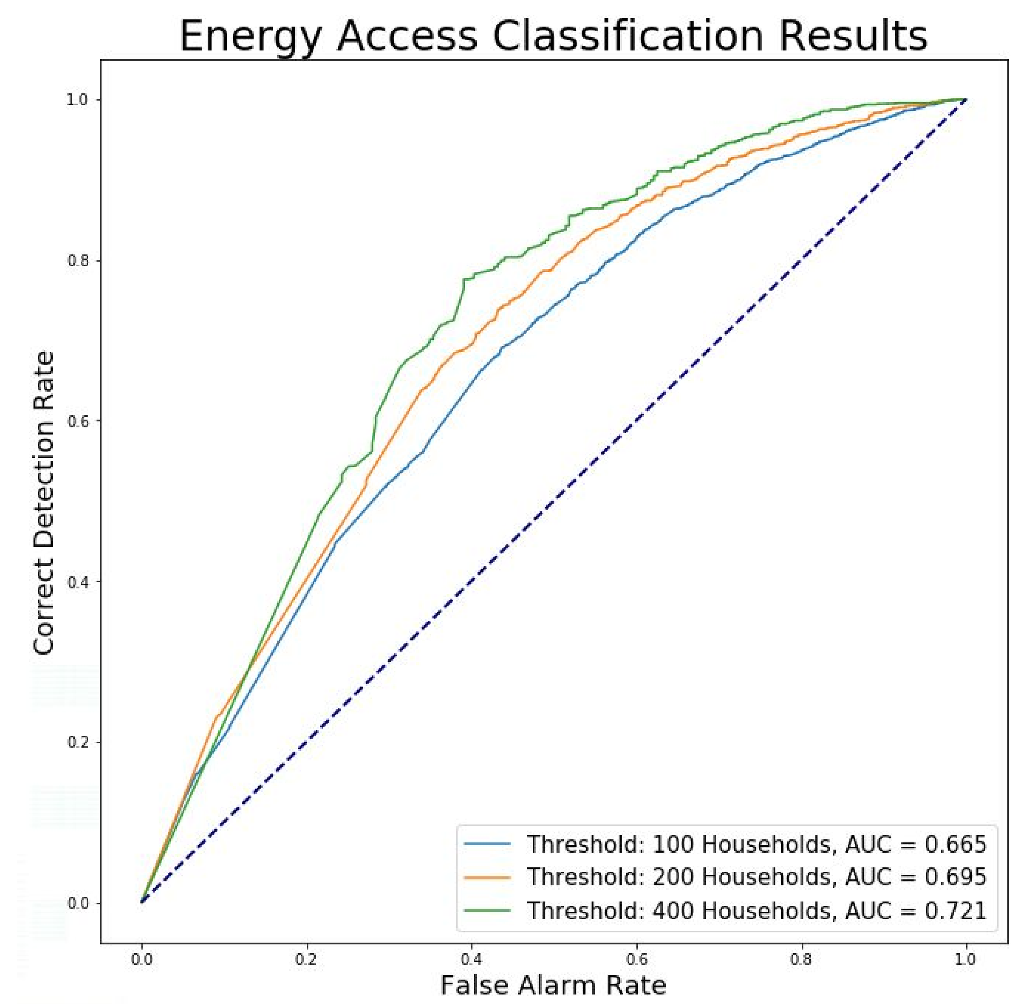

Figure 5. ROC Curve demonstrating results of energy access projections separated by three models, each by selecting a minimum number of households threshold.

Conclusion & Future Steps

This preliminary study demonstrated that the data are nonlinear, and thus a nonlinear classifier will be more successful in differentiating between village electrification status. It also yielded a development environment that is flexible to expanded feature extraction and testing for quicker iterations of modeling in the future. These descriptors may include other features detected from VIIRS or additional satellite imagery, such as the presence of buildings, irrigation, and other energy access indicators. Given the constraints of using only nighttime lights to predict energy access, this preliminary study indicates significant potential for model improvements with the incorporation of additional predictors.

Sources

- Energy access database. (n.d.). Retrieved October 03, 2017, from http://www.worldenergyoutlook.org/resources/energydevelopment/energyaccessdatabase/.
- Khandker, S.R., Samad, H.A., Ali, R., & Barnes, D.F. (2012). Who Benefits Most from Rural Electrification? Evidence in India. *Policy Research Working Papers*. doi:10.1596/1813-9450-6095
- Min, B., & Gaba, K. M. (2014). Tracking Electrification in Vietnam Using Nighttime Lights. Remote Sensing, 6(10), 9511–9529.
- Shi, K., Yu, B., Huang, Y., Hu, Y., Yin, B., Chen, Z., Wu, J. (2014). Evaluating the Ability of NPP-VIIRS Nighttime Light Data to Estimate the Gross Domestic Product and the Electric Power Consumption of China at Multiple Scales: A Comparison with DMSP-OLS Data. Remote Sensing, 6(2), 1705–1724.