Brief Note: Potential Benefits of Higher Resolution Satellite Imagery for Electrification

June Lukuyu, Jay Taneja, and Gunther Bensch

Existing data for informing energy, climate, and development policymaking in sub-Saharan Africa are scarce. Further, traditional data collection methods, such as surveys, are expensive to scale. Satellite imagery, which is rapidly becoming available at higher resolutions, combined with machine learning techniques, is increasingly used in the region to derive information for guiding decision-making where alternative solutions do not exist.

Our initial study aimed to address electricity demand constraints in the electrification planning process to ensure that electrification stimulates economic growth. We used a novel approach to identify areas with existing diesel-powered irrigation in Ethiopia by combining ground data from an agricultural survey with satellite-measured pollution, crop cover, and topography data. The pollution data is at a spatial resolution of 7 by 3.5 km^2 measured nearly every day and the crop cover satellite imagery at a spatial resolution of 250 by 250 m^2, collected every 16 days. Though our study showed promising results, the limited ground-truth sample size collected and the nature of the publicly available satellite data used makes it difficult to generalize our findings beyond the study region of Amhara. It is, therefore, crucial to explore innovative methods to increase the volume and diversity of ground-truth samples and improve the transferability of our models. We argue that using higher resolution satellite imagery could overcome some of the challenges encountered in our initial study and aid in realizing the potential impacts of the work.

We propose to extend the work by using satellite imagery with a higher Spatio-temporal resolution combined with deep learning techniques to perform the same task of identifying areas with existing diesel-powered irrigation. The goal of the model will be to extract relevant features from the daytime satellite imagery and meteorological data, considering both spatial and temporal correlations, that are predictive of diesel-based irrigated farms without first prescribing what features the model should look for. In this brief note, we present some preliminary insights into the advantages of higher resolution satellite imagery in the context of our work.

Higher Resolution Satellite Imagery Could Increase the Ground-Truth Sample Size

Key finding: **Higher resolution imagery results in more extensive and higher quality ground-truth samples, which are expensive and onerous to collect.**

Planet is an unparalleled remote sensing system, simultaneously satisfying high spatial resolution, temporal resolution, and coverage capacity requirements. It collects 3-meter spatial resolution imagery at a one-day revisit cycle worldwide. Our initial study collected irrigation and crop cultivation data for 6,178 plots. Our model extracted patterns of vegetation greenness (vegetation indices) from 250 m resolution MODIS satellite imagery over time to predict plots with irrigation activity. To accomplish this, we only selected 250 m pixels encompassing plots with the same irrigation status. As shown in Table 1, using 250 m resolution imagery reduced our sample size by more than half. With Planet imagery, we would not have to group the plots as we would be able to extract patterns of vegetation greenness at the plot level. Ground truth samples are expensive and onerous to collect. Therefore, the ability to stretch ground-truth samples is very beneficial for model development.

Table 1: Ground truth sample size for satellite imagery of different resolutions

	Ground Truth Sample Size	Pixel coverage with plot area
250 m resolution MODIS satellite imagery	2874	10 – 80 %
3 m resolution Planet satellite imagery	6178	100 %

Higher Resolution Satellite Imagery Could Result in More Detailed Patterns of Information for Model Training

Key finding: **Higher resolution imagery results in higher quality ground-truth samples, which can enable the extraction of more detailed patterns for information for more rigorous model development.**

We downloaded both MODIS and Planet satellite imagery that encompasses the areas in the Amhara and Oromia regions, where we collected ground truth data in our initial study between June 2020 and August 2021. We evaluate the value of higher resolution imagery in capturing detailed patterns of information for model training by comparing the distribution and patterns of vegetation greenness (Normalized Difference Vegetation Index, NDVI) measurements within a 250 m area/pixel between those calculated from low resolution (MODIS) and high resolution (Planet) imagery. As shown in Figure 1, we observe a wide distribution of NDVI values within the 250 m pixel on any given day. We lose this level of detail when we consider the single NDVI value obtained for the pixel from the lower resolution imagery. Further, we also lose the temporal detail we get from the daily NDVI values from the higher resolution imagery when we only consider the NDVI values every 16 days from the MODIS imagery.

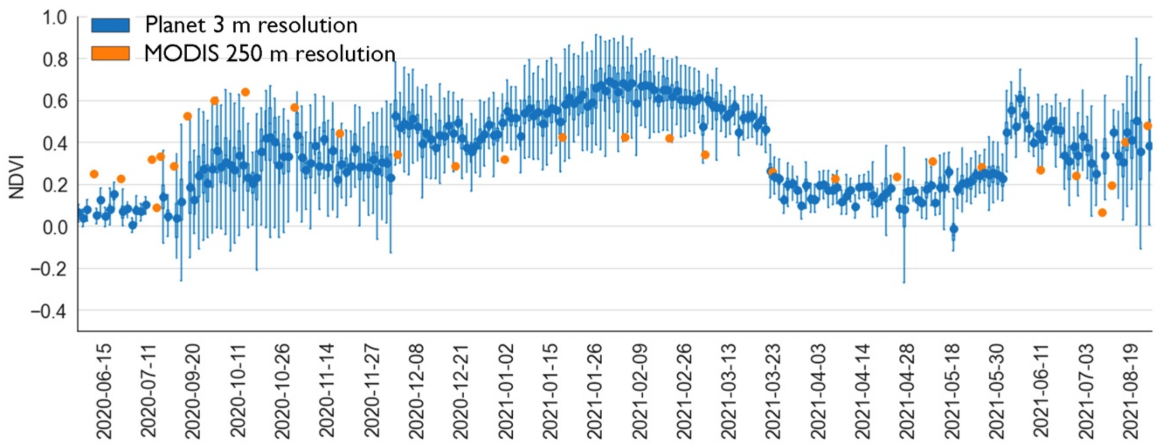

Figure 1: Distribution of NDVI values over a 250 m area in the Amhara region

Sensors of satellite instruments usually trade off temporal and spatial resolution since it is difficult to maximize both. While the benefits of improvement in spatial resolution are pretty evident, as described above, we do not observe significant granular differences in the temporal patterns of the high and low-resolution NDVI patterns, as shown in Figure 1. We see that the general shape of NDVI values follows the same path over the year in both cases. Our study could therefore benefit from imagery with a much higher spatial resolution and sacrifice higher temporal resolution.

To further illuminate the benefit of higher resolution imagery in extracting more detailed information for model development, we consider the change in patterns of NDVI for irrigated and non-irrigated pixels/plots between irrigation and non-irrigation seasons. Higher values of NDVI correspond to denser vegetation. We compare those calculated from low-resolution (MODIS) imagery and those from high-resolution (Planet) imagery. We specifically compare the change in average NDVI values :

1. Between the rainy (non-irrigation) and dry (irrigation) seasons
2. Between irrigated and non-irrigated pixels/plots in our sample
3. Between those calculated from low resolution (MODIS) and high resolution (Planet) imagery

From the precipitation data we obtained for the study area during this period, we observed a dry period in the study area between November 2020 and May 2021, which coincides with the irrigation season (Figure 2).

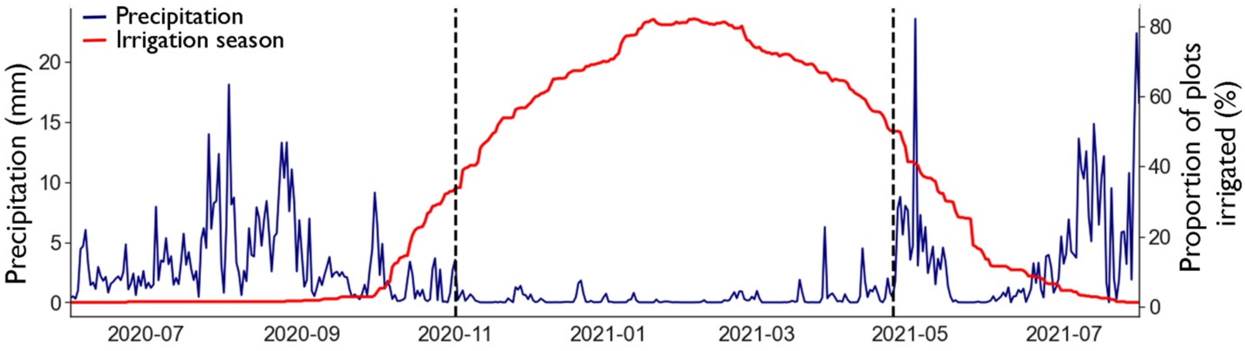

Figure 2: Average daily rainfall and 2020/2021 irrigation season in the Amhara region

From our preliminary analysis (shown in Figure 3), we can observe that the change in average NDVI between the 2020/2021 dry season and the prior 2020 rainy season is much more prominent in the non-irrigated plots than in the irrigated plots (~15% vs. ~3%) in the case of the higher resolution imagery. Therefore, we can more easily deduce that there is more vegetation greenness in the irrigated plots during the dry season than in the non-irrigated plots. On the other hand, we lose this level of detail in the case of the lower resolution imagery, whereby the change in average NDVI between the 2020/2021 dry season and the 2020 rainy season is not significantly different between the non-irrigated plots than the irrigated plots (~ 43% vs. ~44%). We, therefore, lose some detailed vegetation greenness information captured by higher-resolution imagery.

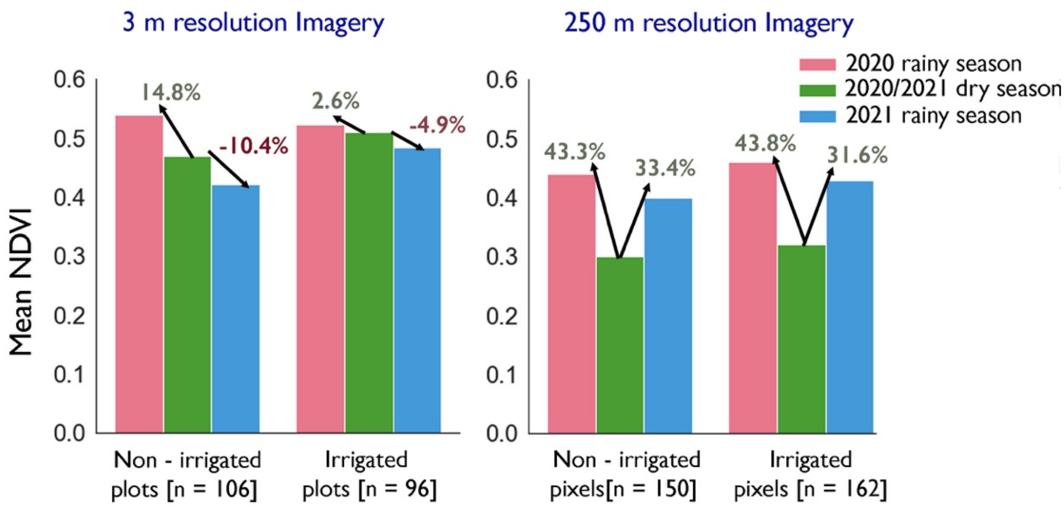

Figure 3: Seasonal variation in average NDVI values calculated from high-resolution and low-resolution imagery

Next Steps

1. We propose to **collect more ground-truth data** to enable us to develop a model that applies to a broader range of settings
2. We propose to **develop, apply and compare the performance of different machine learning models**. For example :
    - We propose to use the 3 m resolution planet imagery and the labels from the previously collected survey data and meteorological data of the study regions to train an unsupervised deep learning model to perform the same task of identifying areas with existing diesel-powered irrigation. The goal is to compare the efficacy of using a different computing approach and leveraging a larger sample size and higher resolution information to work with in guiding electrification planning in Ethiopia.
3. We propose to explore how to **better connect our models with existing planning tools** such as least-cost electrification models, agriculture and energy planning tools, etc. Further, we would like to explore how to

visualize better and present the results of our models so that they are more accessible to policymakers and actionable.