Village
Data
Analytics

e-GUIDE

CASE STUDY

Electricity demand estimation and viability analysis for off-grid villages in Kenya

Table of contents

1 - Project goal

In this project, we set out to test innovative data analytical tools to estimate and validate predictions of energy demand in off-grid areas. This is both difficult and important.

It is difficult, because, so far, there have been few data sources to work with and state-of-the-art data analytics had earlier rarely been used. It is important, because reliably predicting energy demand is key to choosing the most cost-efficient electrification solution and to right-size the chosen electricity access solution, thereby ensuring the best possible electrification economics. This will help us achieve universal access to modern energy both more quickly and in the most cost-efficient manner, enabling scarce infrastructure investment resources to go further.

In this project, we demonstrate the use of the Village Data Analytics (VIDA) together with e-GUIDE's Electricity Consumption Prediction service in underserved areas in Kenya. The electricity consumption predictions provided by e-GUIDE seamlessly fit into the VIDA site selection and village-level analysis and allow for a thorough assessment of the potential for rural electrification at the scale of the entire country and with the precision of individual villages.

The results presented here employ machine learning techniques to draw insights from satellite imagery and electricity consumption data from grid-connected customers. They show how these technologies enable scalable, data-driven energy access planning for government planners, utilities, and off-grid electrification companies.

2 - Problem

In several Kenyan counties, less than a quarter of households have access to electricity. More than half of them, an estimated 1.1 million households, are located more than five kilometres away from the national electricity grid.

To electrify remote settlements requires extending the grid, but also investments into off-grid solutions, such as mini-grids and solar home systems (SHS). In order to select the best electrification scenario, it is necessary to first estimate the energy demand of remote settlements. However, so far, there has been a distinct lack of reliable data and information about the potential consumption of these future electricity consumers. Basic questions, such as: where are villages?, how big are they? and what is their energy demand? are difficult to answer.

This lack and the associated uncertainty is one of the main barriers to effective electrification. It hampers planning for the most cost-effective approach and makes right-sizing solutions hard. As a result, the electrification of over one million off-grid households in Kenya alone, and many more millions globally is too slow, and deployed systems are either too large -- costing money that could have been used to add electricity access elsewhere -- or too small -- hampering nascent growth of economies. This puts the sustainable development goal of universal electrification by 2030 at risk.

3 - The e-GUIDE Electricity Consumption Prediction Service

The e-GUIDE Electricity Consumption Prediction Service introduces an approach that uses daytime satellite imagery to predict individual residential building consumption levels upon connection. Using six years of longitudinal data of electricity consumption for a large cross-section of grid customers, we apply convolutional neural networks (CNNs) to daytime satellite images to predict expected levels of residential electricity consumption for individual customers throughout the country.

Through an API, we share a novel dataset of predicted individual building electricity consumption levels aggregated at 250m, 500m, or 1000m resolution, which can be used for tasks such as electricity system planning and estimating economic development.

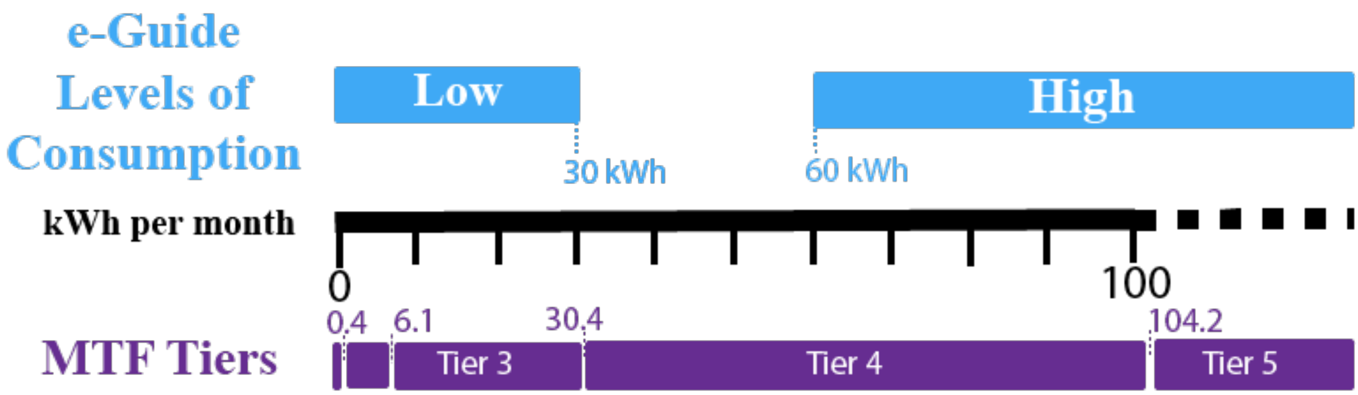


Figure 1
Alignment of e-GUIDE levels of consumption with the MTF Tiers

Individual buildings are classified as either “low” or “high” consumption. Buildings classified as “low” have an estimated monthly grid consumption of less than or equal to 30kWh. Buildings classified as “high” have an estimated monthly grid consumption of more than 60 kWh. For buildings whose estimated monthly consumption falls between 31 and 59 kWh, the model classifies buildings with consumption of less than 40kWh into the low class, while the rest have a random chance of being assigned to either consumption level. Further details justifying this design decision are elaborated in our model documentation.

The “low” and “high” classes draw inspiration from the World Bank's Multi Tier Framework (MTF), which divides electricity consumption into different levels of electricity services. Our classification considers “low” levels of consumption as corresponding to Tiers 0 to 2 of the framework while “high” levels of consumption generally correspond to Tier 4 and above.

The models that underlie the e-GUIDE Consumption Prediction Service were primarily developed by Simone Fobi, a PhD student researcher at Columbia University, and the API for accessing the model results was primarily developed by Joel Mugyenyi from the e-GUIDE team.

4 - Village Data Analytics (VIDA)

VIDA is a modular planning tool that ingests satellite imagery, public data, and non-public, on-ground data. A set of machine learning-based algorithms for image processing and energy modelling are used to identify and evaluate electrification options for investors, planners and companies. The energy modelling is based on TFE Energy's long-standing expertise in electrification in frontier markets.

VIDA's data analysis workflow follows the four steps shown below. Both the extraction, as well as prediction steps are aided by a library of machine learning (ML) enabled algorithms.

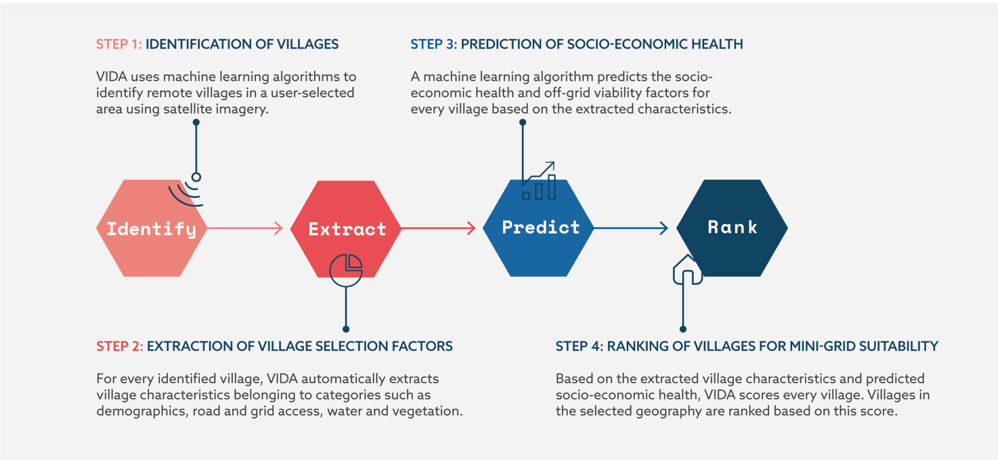


Figure 2
Village Data Analytics' workflow

The result is a "smart map", including a ranking of sites and decision-making information at the level of individual villages and at the scale of entire countries. The comparison and ranking is specifically designed to reflect

5 - Workflow

The VIDA team incorporated e-GUIDE's results in Kenya into its overall analysis to identify rural villages with a promising energy demand profile and assess their mini-grid viability. The analysis starts at a country level and then filters down to a list of attractive villages to engage (i.e. creating a shortlist).

We looked at underserved counties across Kenya for the first few steps and then picked 5 villages for a detailed analysis (steps 4) to showcase how it works.

Step 1- Identify

Firstly, VIDA uses a machine learning based algorithm to automatically detect settlements that are off-grid. For this analysis, Gridfinder's night-light imagery based electrification prediction data was used.

Step 2- Extract

VIDA extracts data about the individual settlements, including size and demographics, quality of road access, distance to the grid, and socio-economic data. e-GUIDE's consumption API is then utilized to predict the number of potential connections in villages as well as the total demand of both "low" and "high" buildings.

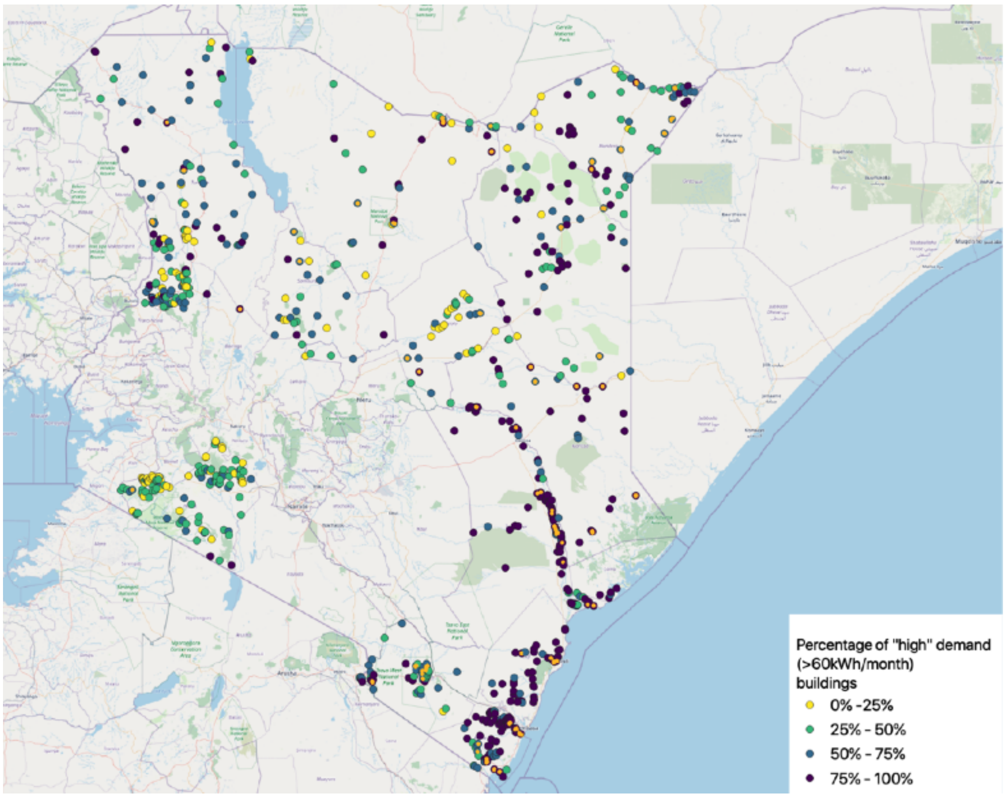


Figure 3
Geolocations of several hundred assessed villages across Kenya

Step 3- Predict

Extracted village level information is then combined to predict potentially viable mini-grid sites. Villages that are best electrified either by grid extension or by SHS are filtered out. After that, we set threshold values for the number of connections, demand, and ratio of “high” demand buildings which are important criteria used by the mini-grid developers. The threshold values were:

- More than 200 customers
- More than 4000 kWh / month predicted demand (assuming a monthly consumption of 15 kWh and 60kWh for “low” and “high” demand buildings, respectively.
- More 50% of the customers have a “high” predicted demand (>60kWh/month)

Based on this analysis, a shortlist of potential mini-grid sites was generated.

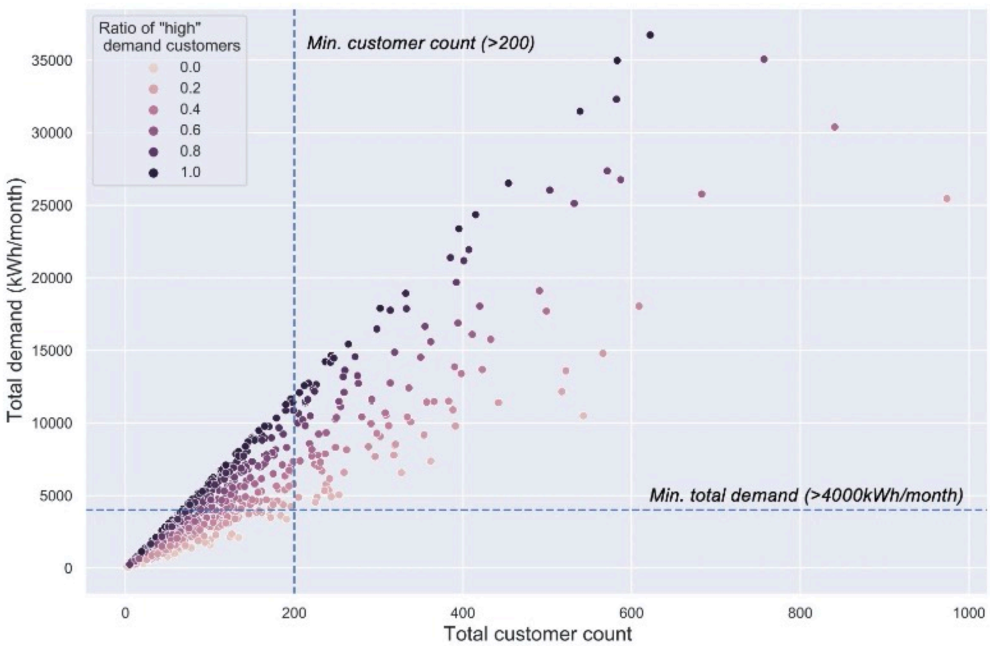


Figure 4
Distribution of sites based on customer count, total monthly demand as well as ratio of “high” demand connections. The dotted lines represent our filters. Villages in the top right quadrant with more than 50% of “high” predicted demand customers are considered viable mini-grid sites.

Step 4- Rank

In the potential mini-grid villages, an in-depth viability assessment is conducted using VIDA's algorithms, complemented by e-GUIDE's demand prediction. For this case study, we chose 5 villages out of the shortlisted villages to demonstrate the process. We followed the following process

Figure 5
Satellite image of one of the identified potential settlements satisfying our filtering conditions on monthly demand estimate, total number of customers and ratio of "high" demand buildings (image credit: Google Earth)

High-value customers: VIDA identifies the village core and outskirts using a cluster detection methodology on a quadratic, 100m raster. Within the core area high-value customers are then identified. This is based on building category, location of buildings, density of building and proximity to road. In addition, potential anchor loads, such as healthcare facilities or schools are taken into consideration. This analysis provides household-level details on total connections and composition of connections by demand.

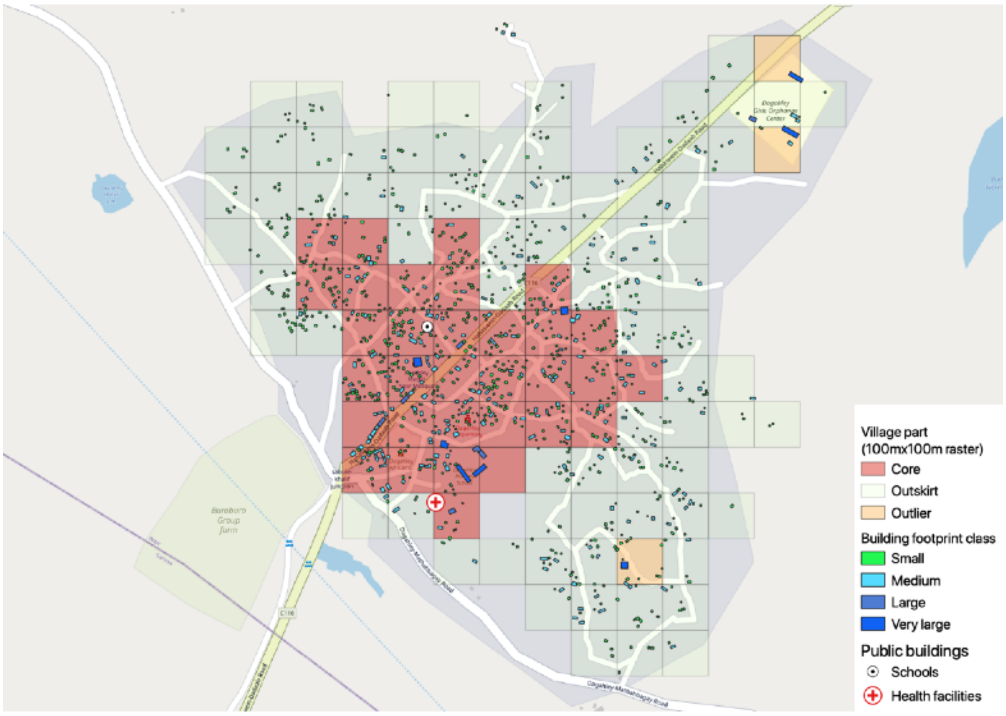


Figure 6
Identified village core (red pixels), outskirts (transparent), and outliers (light red). Building colors indicate their category.

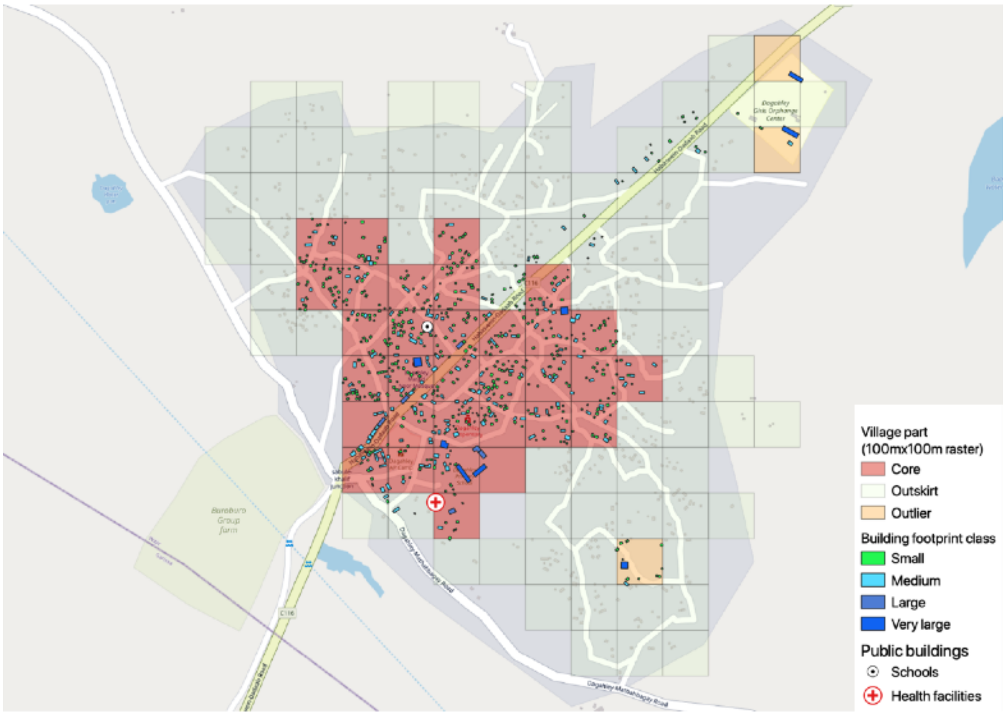


Figure 7
Selection of high value mini-grid customers based on estimated demand and location

Total energy demand: The total energy demand of the identified customers is predicted based on the 250m raster output from the e-GUIDE model, as well as on VIDA's methodology which takes into account size, location and type of buildings. This results in an estimate of kWh/month from which the monthly revenue potential can be inferred. For this report, the e-GUIDE team prepared the model output on a 100m x 100m raster level in order to estimate the total demand of the high value mini-grid customers (see figure 9).

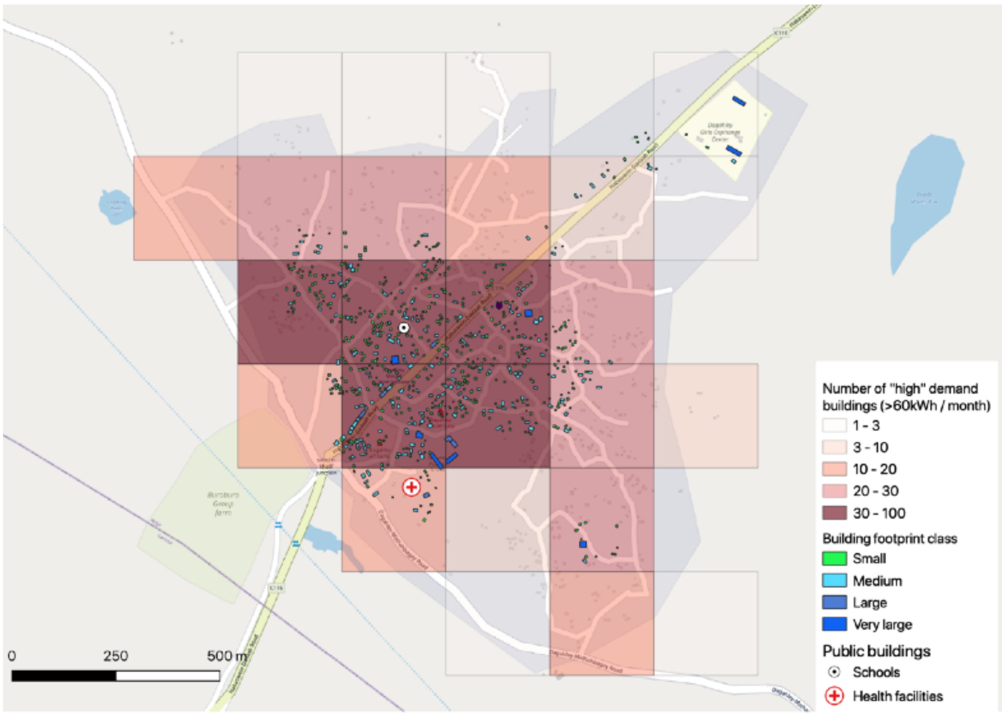


Figure 8
Granular demand prediction with the e-GUIDE model. The coloring of the 250m raster indicates the number of "high" demand buildings identified. A comparison with the previous figure shows a good overlap with VIDA's core and outskirts detection.

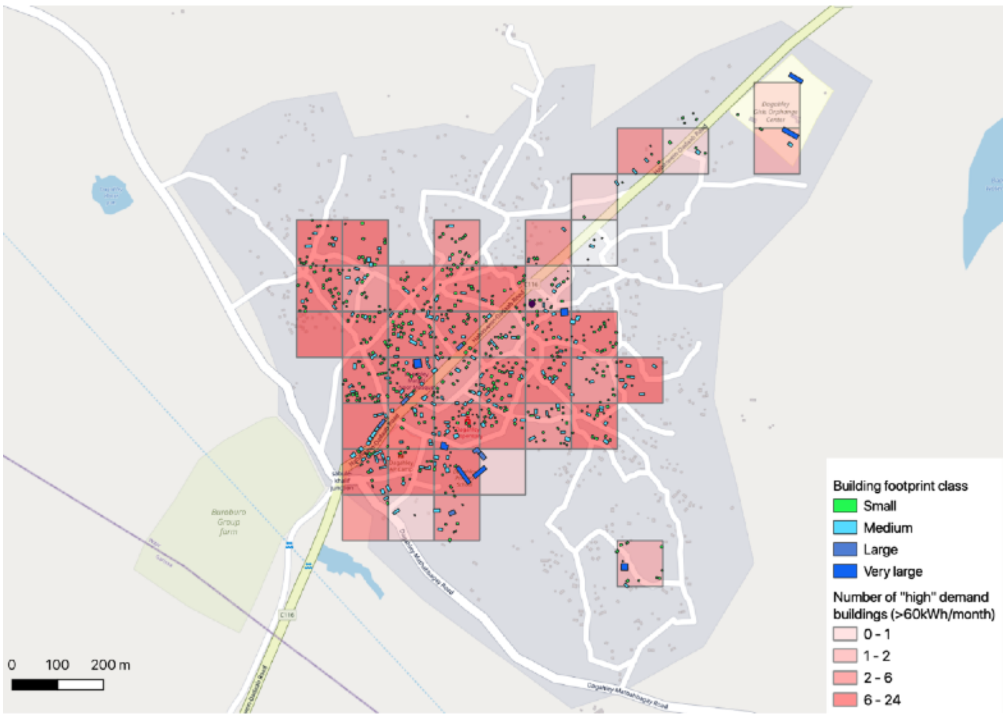


Figure 9
e-GUIDE demand prediction for the high value customers (this 100m x 100m raster is a custom output of the e-GUIDE API)

Distribution layout: VIDA then generates a distribution layout that connects the high-value mini-grid customers. First, we generate a minimum spanning tree to get an upper limit for the connection density.

Next, the trunk lines are generated. These are either single or three phase and typically follow the main roads through the village. VIDA sets the location of the poles at an equal distance and then generates dropdown lines that connect every high value mini-grid customer to the nearby poles. The output is shown in figure 10 below.

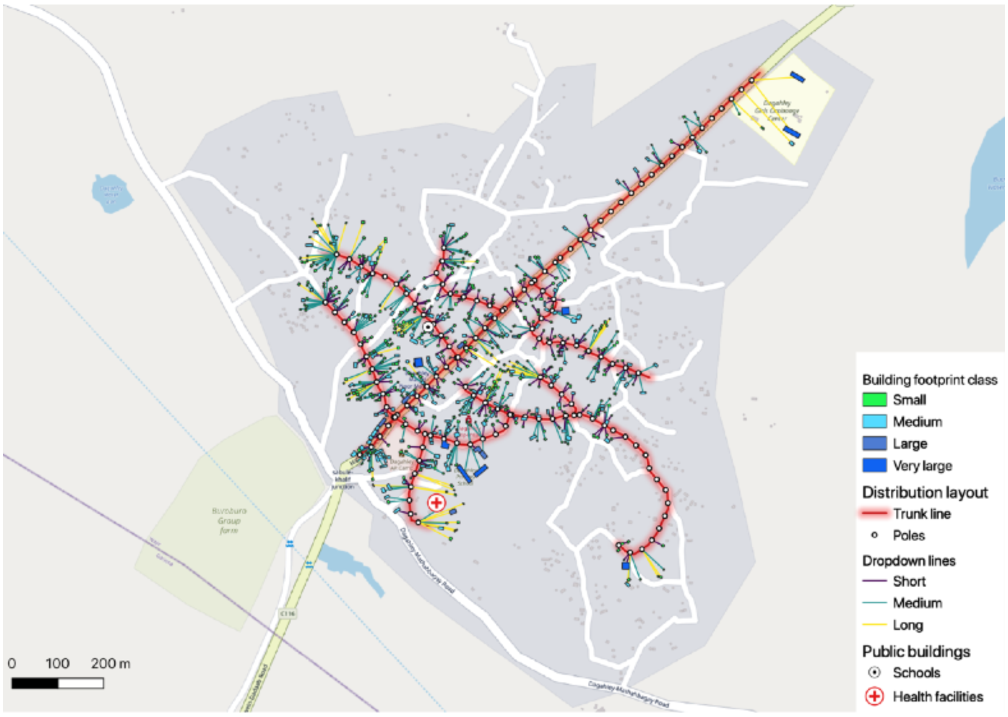


Figure 10
Mini-grid layout with trunk line, poles and dropdown lines connecting previously identified high value mini-grid customers.

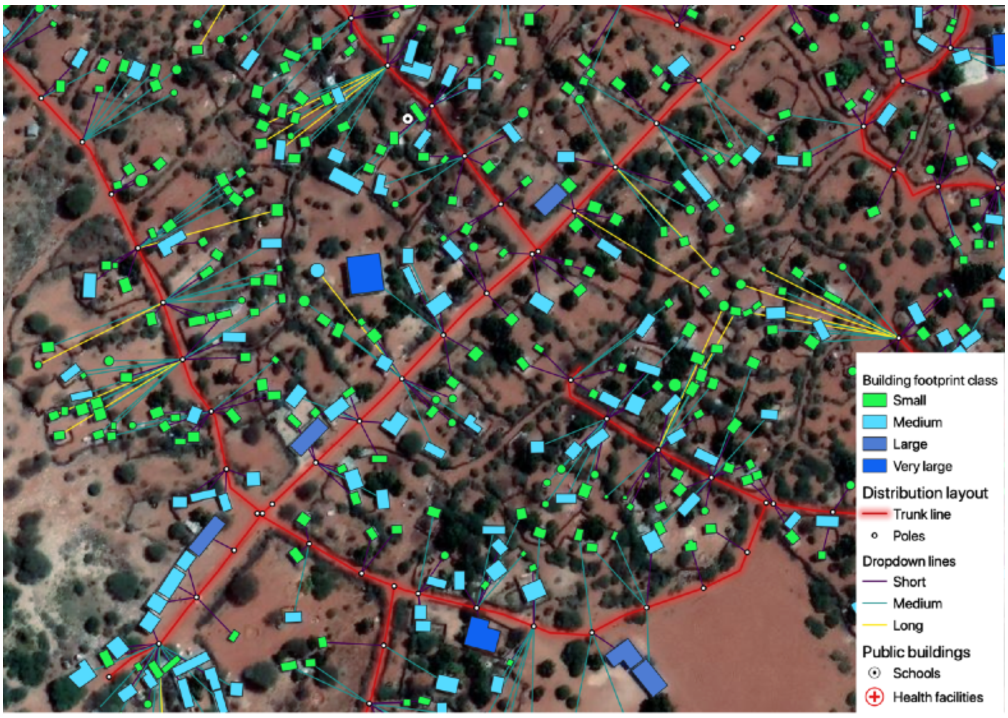
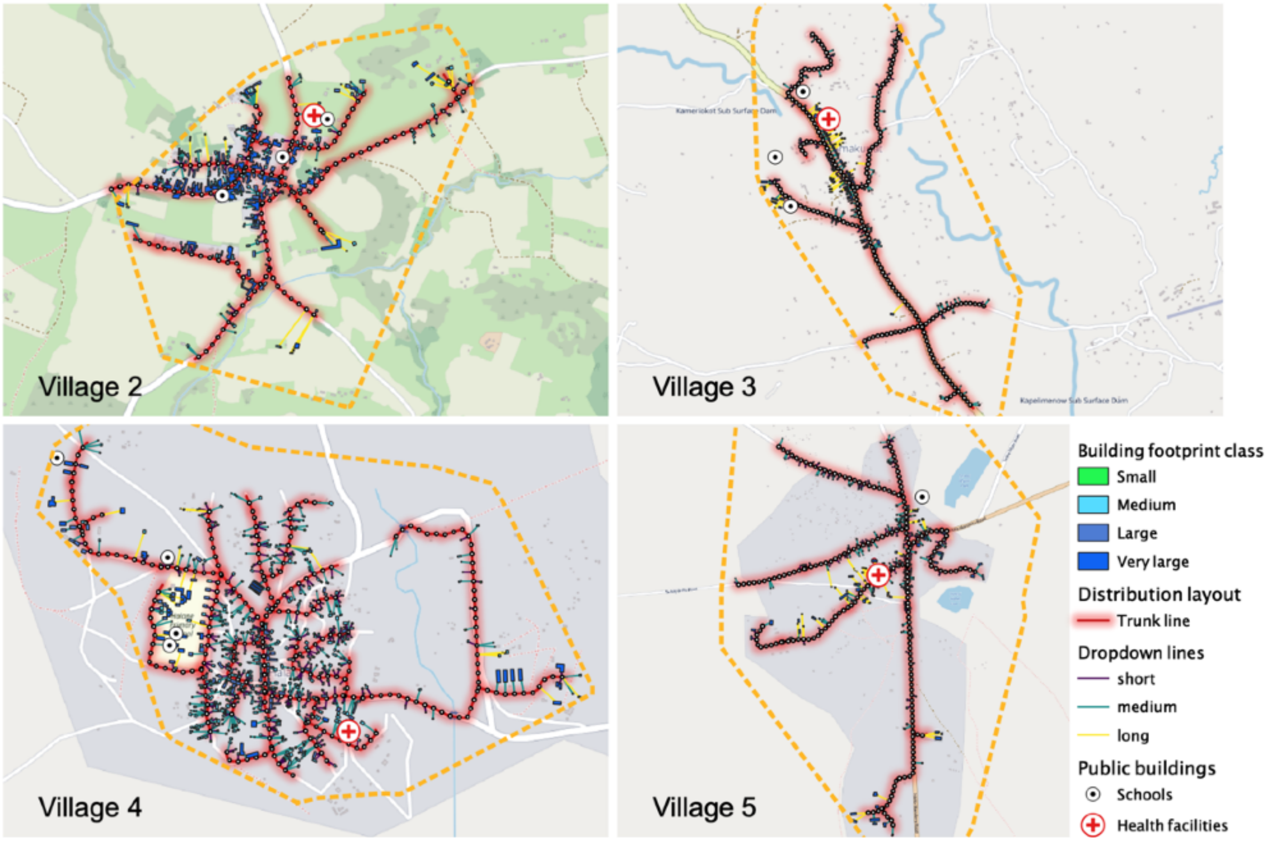


Figure 11
Predicted trunk lines (red), poles (white dots) and dropdown lines connecting the identified high-value mini-grid customers to the poles

Distribution cost: Based on the length of the predicted trunk line, the number of poles and the length of the dropdown lines, VIDA calculates total estimated cost of distribution. This algorithm uses average cost per meter of wiring (different sizes) and the average cost of poles on a village by village basis.

Figure 12
VIDA's predicted distribution layouts



The process above yields both the bill of materials (BoM) for the distribution grid, as well as the energy demand and average connection value. Based on these numbers and the local cost of materials, an initial mini-grid layout, as well as a cost estimate can be created.

Taking the above factors into account, VIDA ranks villages based on mini-grid viability. Furthermore, based on the data extracted by VIDA and shown in the table below, users can select the most viable sites based on their own priorities.

In this analysis, villages 2 and 4 are viable mini-grid sites. As a next step, these settlements can be surveyed.

Table 1
VIDA's decision-making factors for mini-grids

	Village 1	Village 2	Village 3	Village 4	Village 5
Village statstics					
Village area (km2)	1.2	0.9	**3.2**	1.5	3.7
Number of buildings	**1,128**	346	440	850	665
Building density (buildings/km2)	**910**	372	138	557	179
e-GUIDE predicted electricity demand of the whole village (kWh/ month)*	**31,485**	10,260	16,890	23,385	11,115
Mini-grid customer statistics					
Number of mini-grid customers	692	269	215	**715**	381
Distribution of mini-grid customers (small:medium:large:very large buildings)	509:169:6:8	95:95:30:49	72:90:22:31	**350:283:52: 30**	241:89:20:31
Availability of institutional loads (education/health facilities)	1 / 1	3 / 1	3 / 1	**4 / 1**	1 / 1
Distribution layout					
Length of distribution wire (km)	28.5	13.5	15.5	30.0	19.5
Distribution wire per customer (m)	**41.2**	50.2	72.1	42.0	51.2
Number of poles	128	132	203	250	206
Customer per pole	**5.4**	2.0	1.1	2.9	1.8
Overall result					
VIDA's predicted demand for mini-grid (kWh/month)**	6,870	6,270	4,680	**10,890**	5,610
Average demand per customers kWh / month	9.9	**23.3**	21.8	15.2	14.7
Overall connection value (kWh/km/ month)	241	**464**	302	363	288
Mini-grid viability	**Medium**	**High**	**Medium**	**High**	**Medium**

** - e-GUIDE's predicted electricity demand for the whole village is used to filter out villages that have low overall demand for mini-grids (see step 3).*

*** - VIDA's predicted electricity demand is separately calculated by VIDA and represents the demand for electricity of the high-value mini-grid customers within the village.*

6 - Conclusion

This study shows the progress that has been made in estimating off-grid energy demand and identifying viable sites. It shows how the VIDA site selection service can be further enriched with the energy demand predictions supplied by the e-GUIDE model. This offers the electrification industry an highly innovative, tested, rapid, data-driven site selection process from the country level down to individual villages.

This work shows how research organizations such as e-GUIDE can provide crucial research to help the sectors adapt new technologies and methodologies. Furthermore, the use of an open API allows for seamless integration into fully digital tools like VIDA and the reduction of tedious, manual data transfers between data silos from different actors.

This analysis makes site selection more bankable and helps the industry grow at the scale needed to match the electrification challenge. It significantly reduces the number of on-ground surveys required and complements them, thereby increasing their utility and reliability. The data provided by VIDA and enriched by e-GUIDE can help to cross-check and validate survey results and vice versa. This is crucial for electrification companies, investors and planners alike.

7 - About us

Village Data Analytics

Village Data Analytics (VIDA) is an AI-enabled custom software that supports data-driven investment, business and policy decisions in remote, frontier markets. VIDA's goal is to catalyse the required large-scale investment into these markets to meet global development goals and benefit around 2 billion people. VIDA is developed by an international team based out of Munich and Cape Town. The team brings together high-impact technology and deep, on-ground sector expertise. VIDA is supported by the European Space Agency as well as appliedAI, Germany´s largest AI initiative. The software is used by governments and development organisations, by investors and banks and by companies and NGOs.

More examples of VIDA's work include:

- Identification of more than 2,500 potential off-grid villages for mini-grids in Ethiopia, linked to information about anchor loads, including schools, hospitals or telecom towers. Analysis includes a predicted distribution layout for mini-grid design.

- Identification of attractive sales regions for solar home systems in Kenya, based on settlements, grid and road access and potential sales and logistics hubs. Download the PowerAfrica report here.

- Detailed site assessment and evaluation for mini-grid companies in Tanzania, DRC, Nigeria and Kenya, including socio-economic assessments of villages, distribution layout, and assessment of productive use opportunities and anchor loads.

- COVID 19 response / rural health centre electrification in Africa: VIDA analysed over 1,000 unelectrified rural health centers and helped prioritize them for electrification, based on the population they serve, their strategic importance and the suitability of the surrounding village for electrification.

e-GUIDE

The Electricity Growth and Use in Developing Economies (e-GUIDE) Initiative seeks to transform the approaches used for planning and operations of electricity infrastructure in developing regions. We are constructing measurement and data analytics techniques that are scalable, transnational, and verified using real data on electricity consumption and infrastructure. We partner with electricity service companies to develop our techniques, deploy them at scale, and build capacity for data and analytics in the electricity sector.

The e-GUIDE Electricity Consumption Prediction Service is presently available for residential consumption predictions throughout Kenya. The team is currently enhancing the model with additional training data in Rwanda and Uganda; the expanded training dataset will improve the performance of the model in those additional countries, but also enable higher confidence predictions in countries where training data are unavailable. Further, the team is also extending beyond residential consumption predictions into small and medium enterprise (SME) consumption predictions. These extensions will be released as future versions of the Electricity Consumption Prediction Service.

8 - References

This case-study built upon earlier VIDA work for PowerAfrica. For a more detailed overview, see the Off-grid Solar Market Assessment brief for 14 underserved countries of Kenya, 2020 report by USAID Power Africa (link).

For this case study, VIDA used a number of publicly available data-sets. These include Gridfinder (link), a preliminary version of the World Settlement Footprint 2019 (WSF2019) provided by DLR (link), the high resolution settlement layer (HRSL) from Facebook (link), OpenStreetMap (link), the Energy Access Explorer (link), and other relevant data-sets from the Humanitarian Data Exchange (link).

Contact us

Village Data Analytics
Franz-Joseph Str. 10
Munich 80801, Germany

nabin.gaihre@villagedata.io
philippe.raisin@villagedata.io

www.villagedata.io

e-GUIDE Initiative
c/o Jay Taneja
151 Holdsworth Way
Amherst MA 01002, USA

jtnaeja@umass.edu
modi@columbia.edu

www.eguide.io