ViDA Village Data Analytics

PowerGen

TFE Energy

IMPROVED MINI-GRID SITE SELECTION IN WEST AFRICA USING **VILLAGE DATA ANALYTICS (VIDA)**

> **VIDA provides a very high-quality analysis of remote areas. For a company entering new markets, with limited ground data, VIDA can dramatically influence site selection and generate a quality pipeline of sites quickly.**
>
> – **DANIEL ANASTOS**, PowerGen

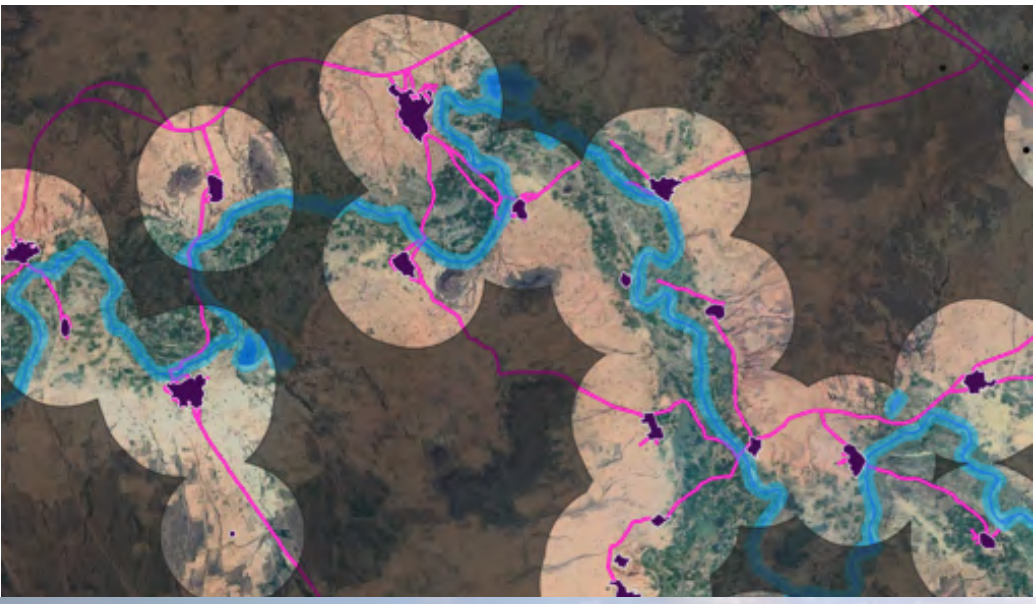


TABLE OF CONTENTS

1. INTRODUCTION

Over 2 billion people live without grid electricity or receive very unreliable power. That puts their safety, productive and aspirational needs at risk. Most affected people live in remote villages in Africa and Asia. To electrify them, more than 210,000 mini-grids are required by 2030, at an investment of 220 billion US (**link**).

However, locating remote villages and gathering information about them for making impactful investment choices is a major challenge. Mini-grid developers often manually analyze maps to identify villages and then conduct on-site surveys. This time-consuming, error-prone and costly process is a significant barrier for scaling up mini-grids to meet universal energy access (SDG7). TFE Energy's Village Data Analytics (VIDA) offers a solution to this challenge.

VIDA is an artificial intelligence (AI) powered, data-enabled service that utilizes satellite imagery, publicly available geospatial data, on-ground survey data, and energy modeling to identify and extract insights about rural villages, anywhere in the world, and to assess their suitability for off-grid electrification, including mini-grids.

VIDA has been developed with support from the European Space Agency (ESA) and is designed as a decision-making tool that provides map visualizations, analysis and reports tailored to individual users.

PowerGen is one of the largest mini-grid developers in the world, operating a rapidly expanding portfolio of more than 100 mini-grids across Sub-Saharan Africa. PowerGen is a collaboration partner of TFE Energy and has been crucial in testing and validating VIDA in many parts of West and East Africa.
PowerGen is expanding to new markets in West Africa. TFE Energy offered to provide a VIDA analysis of one of their areas of interest (AOI's). The PowerGen team had previously conducted manual GIS analysis and on-site surveys in the same AOI. As a result of this analysis, a number of villages were shortlisted. VIDA's task was to analyze the shortlisted as well as other promising villages to assess their suitability for a mini-grid installation. The AOI had an area of 5,500 km^2. The goal of this work was to validate the VIDA analysis, its key metrics, and the ranking of villages with the on-ground survey data collected by PowerGen.

This report outlines the methodology of the analysis, the validation process, and its results.

2. THE VIDA APPROACH

VIDA is a highly modular tool that ingests satellite imagery, publicly available geospatial datasets, customer data and non-public on-ground data. A set of machine learning-based image processing and energy modelling algorithms are used to identify and rank viable mini-grid sites. The energy modelling is based on TFE Energy's long-standing expertise in electrification in frontier markets.

The result is a "smart map", including a ranking of sites and decision-making information at village level, specifically designed to reflect a user's site selection criteria.

2.1 DATA AND USER REQUIREMENT COLLECTION

In a first step, the TFE Energy team engaged with the PowerGen team to understand the area of interest (AOI), the number and types of mini-grid PowerGen plans to build, their specific process and criteria for selecting sites, as well as any other preferences and limitations.

We then collected radar, visible and near-infrared satellite imagery from the Copernicus fleet operated by the European Space Agency. In addition, products derived from nightlight and very high-resolution imagery were used. Publicly available GIS data such as Open Street Map (OSM) road data, Humanitarian Open Street Map's household data layers, and other available GIS/statistical data were collected.

2.2 DATA ANALYSIS

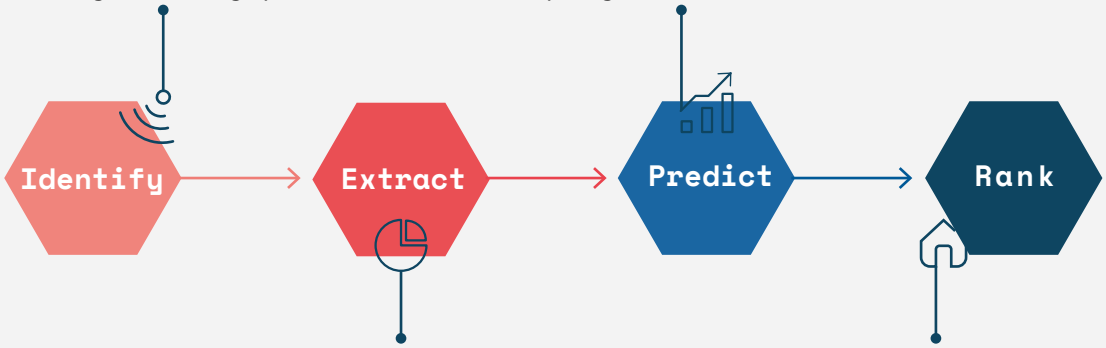
VIDA's data analysis workflow follows the four-step methodology shown below. Both the extraction as well as prediction steps are aided by a suite of machine learning (ML) enabled algorithms.

STEP 1: IDENTIFICATION OF VILLAGES

VIDA uses machine learning algorithms to identify remote villages in a user-selected area using satellite imagery.

STEP 3: PREDICTION OF SOCIO-ECONOMIC HEALTH

A machine learning algorithm predicts the socio-economic health and off-grid viability factors for every village based on the extracted characteristics.



STEP 2: EXTRACTION OF VILLAGE SELECTION FACTORS

For every identified village, VIDA automatically extracts village characteristics belonging to categories such as demographics, road and grid access, water and vegetation.

STEP 4: RANKING OF VILLAGES FOR MINI-GRID SUITABILITY

Based on the extracted village characteristics and predicted socio-economic health, VIDA scores every village. Villages in the selected geography are ranked based on this score.

Step 1 Identification of villages

VIDA used a machine learning (ML) algorithm to automatically identify villages in the AOI using Sentinel-2 imagery. The model was trained on a set of 12,000 expert labelled images.

Step 2 Extraction of village selection factors

An array of image processing algorithms was used to extract selection factors from the villages. These include:

- Grid-access: Using existing medium voltage distribution data and a recently published algorithm developed by Facebook, VIDA classifies villages into on- and off-grid. On-grid villages were filtered out for the remainder of the analysis since they are not of interest for PowerGen in this AOI.
- Demographics: Indicators of size, shape, population and other demographic data are extracted using a machine learning algorithm. From these, we estimate the number of potential mini-grid customers (also called the "connection count") - one of the key factors for PowerGen.
- Connectivity: Using road network data, characteristics of the connectivity of the village and distances from points of interest (towns, grid, etc.) were extracted.
- Waterbodies: Access to water is an important factor for economic productivity of a village. VIDA determines year-round water availability as well as seasonal water access with a methodology based on Sentinel-1 imagery.
- Agriculture: Using visible and NIR imagery, VIDA detects vegetation and predicts agricultural productivity in the villages. This is a valuable proxy for local wealth as well as productive uses of electricity such as mills or pumps.
- Topography and clusters: Topography and clusters of settlements help understand the local context of a village and determine groups of villages to be served with mini-grids.

Step 3 Prediction of socio-economic health

Using TFE Energy's energy modeling tool, the 10-15 selection factors from every village were analyzed to predict a socio-economic score. It serves as a proxy for the socio-economic vibrancy of the village, as well as for practical considerations of building and operating a mini-grid.

The score is calculated by an energy model that normalizes and mixes the various selection factors. It is based on PowerGen's requirements and TFE Energy's experience in rural energy access projects. The modularity of the tool allows VIDA's customers to set and change preferences for a tailored prediction.

10-15 selection factors were extracted per village

- Area
- Population
- House count
- Growth percentage
- Road connectivity
- Number of roads
- Distance from town
- Waterbody distance
- Waterbody seasonality
- Vitality index
- Terrain data
- Cluster

Furthermore, for this specific use case, we assigned categories to the off-grid villages that represent different sources of income. These categories are:

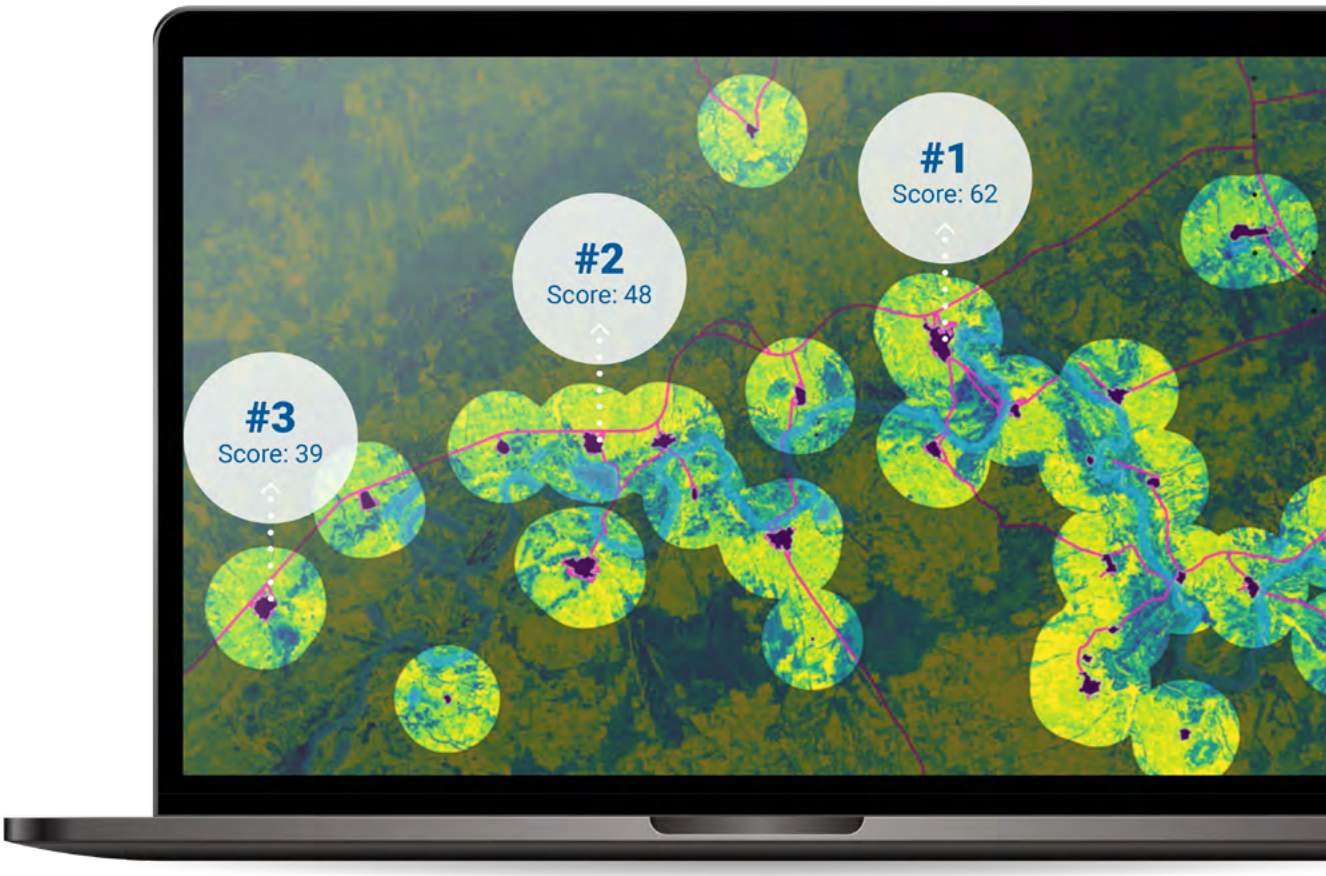
- Agricultural communities
- Fishing communities
- Communities with good road access to larger towns

Step 4 Ranking of villages for mini-grid suitability

Based on the scores of the villages and the labels in the three categories, the villages were ranked. The villages with the highest rank are the most suited for mini-grid development, according to the VIDA analysis.

2.3 DATA VISUALIZATION

The result is a "Smart Map" of the entire area of interest with a ranked list of the most promising off-grid villages, and a database of meaningful village selection factors extracted from every off-grid village in the area.



3. VALIDATION OF RESULTS

PowerGen had already analyzed and collected on-ground data from site visits as well as public sources to create its own independent list of preferred sites. We compared their findings with the findings provided by VIDA.

3.1 VALIDATION OF VILLAGE CHARACTERISTICS

VIDA automatically detected 1,119 villages within the 5,500 km^2 of area using machine learning algorithms. The list included all 60 preselected villages from PowerGen.

3.1.1 Validation of off-grid settlement detection

We compared the predictions from VIDA with on-ground information from 44 villages across the AOI collected by PowerGen. Out of the 44 surveyed villages, 37 were off-grid and 7 on-grid.

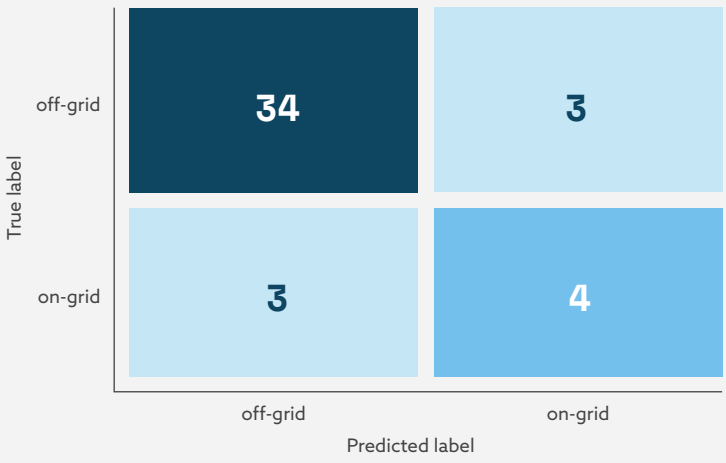
The VIDA methodology predicted grid connectivity well: It correctly identified 34 off-grid villages and incorrectly suggested 3 villages were off-grid. This translates into a true-negative rate (i.e. a village being correctly classified as off-grid) of 0.92 and an overall accuracy of 0.82.

A closer look at the incorrectly classified samples reveals that these villages were in very close vicinity to the estimated grid.

In conclusion, this methodology is well suited to detect off-grid villages, especially when applying a generous buffer zone around estimated grid lines.

Key metrics (left) and confusion matrix (right) for VIDA's off-grid settlement detection

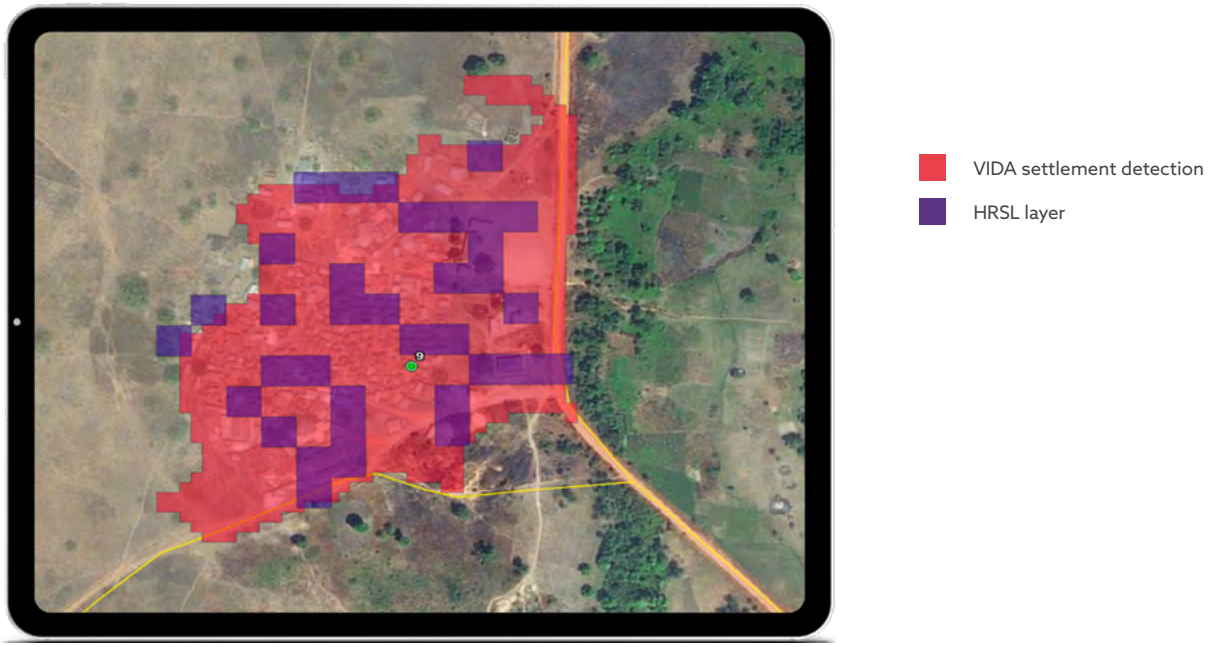
METRIC	VALUE
Sensitivity (recall)	0.57
Specificity (true negative rate)	0.92
Precision (positive predictive value)	0.57
Fallout (false positive rate)	0.08
F-Score	0.57
Overall accuracy	**0.86**



3.1.2 Validation of settlement demographics

The TFE Energy team compared the village size detected by VIDA's algorithm to the state-of-the art High Resolution Settlement Layer (HRSL) published by **Facebook and Columbia University.** The HRSL offers 30m spatial resolution and includes population estimates. We compared the two layers both across the whole AOI as well as on the subset of the 60 pre-selected villages. An example is shown in the image below.

Two example villages showing VIDA settlement detection (red) and the HRSL layer (blue).



Comparison of detected settlement area between VIDA and HRSL

VALUE	RANKED VILLAGES ONLY (N=60)	FULL AOI (N=1,119)
Area detected: VIDA (km^2)	4.54	67.03
Area detected: HRSL (km^2)	3.08	46.05
Intersection (km^2)	2.05	16.7
Jaccard index (intersection over union)	**0.37**	**0.17**

VIDA's algorithm identified a larger settlement area than the HRSL: While the HRSL identifies the villages, it often doesn't fill in the gaps to capture the whole extent and shape of the settlement. VIDAs 10m pixel-based mask captures the village shape and size better.

At the same time, the HRSL is more sensitive to detecting remote, isolated buildings which were either not detected or filtered out by VIDA. While these individual houses might be interesting for other interventions (such as solar home systems), they are not of importance for mini-grid planning.

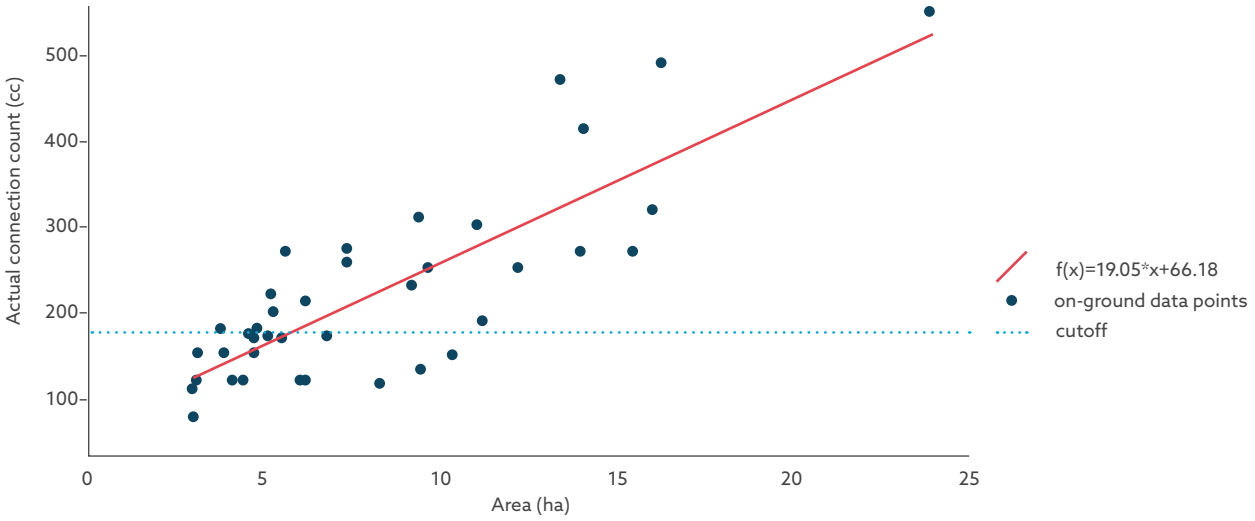
Overall, VIDA's algorithm seems to be well suited for the purpose of assessing the shape and size of the villages in this AOI.

3.1.3 Estimating connection count

To validate the connection count estimate, the TFE Energy team compared VIDA's analysis with the actual connection count provided by PowerGen. We found a linear relationship between the two datasets, as displayed in the figure below. This means that VIDA's algorithm was able to fairly accurately predict the connection count based on the area.

> **VIDA helped significantly with estimating structure count, which is a better predictor of site revenue and connection count than other estimation tools.**
>
> **DANIEL ANASTOS,**
> PowerGen Nigeria

Linear fit for measured area with VIDA algorithm versus actual connection count provided by PowerGen.



3.1.4 Estimating presence of water body

VIDA uses an algorithm to understand the presence of water bodies around the villages. This analysis helps identify areas which are swampy, fertile or prone to flooding.

Most of the rivers and larger lakes are also indicated on Open Street Maps (OSM). The power of VIDA lies in the additional 247km^2 of seasonal water bodies that were identified. Those were not included in OSM but yield insights about agricultural and fishing activities.

Detected area with either year-round or seasonal surface water

VALUE	AREA (KM^2)
Total permanent water surface area (visible on >80% of images over the year)	90
Total additional area detected with seasonal water access	247

3.2 VALIDATION OF RANKED VILLAGES

PowerGen is planning to develop 21 sites out of the 60 preselected sites. A comparison with VIDA's ranking shows that 17 of the selected sites were also in the top 25 list provided by VIDA. This includes the first site that PowerGen has already built, which ranked in the top-10 of the VIDA ranking.

Detected area with either year-round or seasonal surface water

Map layers

- Settlements
- Roads
- Year-round surface water
- Seasonal water bodies / wetlands
- Agricultural areas

Village analysis

- Fishing community
- Agricultural community
- Satellite town to larger settlement / access to good road
- High-ranking site

This suggests a good overlap between PowerGen's in-house assessment and the VIDA ranking. The partial overlap means that VIDA found a number of other valuable sites across the AOI that had not been picked up in the earlier process.

Venn diagram showing the overlap between VIDA's ranking and the sites selected by PowerGen for mini-grid development.

4. OUTCOME

The site selection project with PowerGen has shown that VIDA is able to identify the most promising sites for mini-grids within an area of interest.

This real-world case study shows how the application of VIDA in the off-grid sector can create a distinct commercial advantage and increase the investment-readiness of electrification companies.

Dedicated, data driven tools not only edge the emerging mini-grid sector closer to profitability and scale. They are also essential to providing 1 billion unconnected people with modern energy services by 2030, a target enshrined by SDG7. It is an exciting time to be working in the energy sector. In the words of PowerGen CEO Sam Slaughter, "we are just scratching the surface when it comes to new data-driven tools that enable the mini-grid space. It's going to be interesting to see what we and other mini-grid develoers are going to move into in the future".

TFE Energy will continue to work with PowerGen to help them reach their target of 100 million customers.

Aerial imagery of a mini-grid site developed by PowerGen

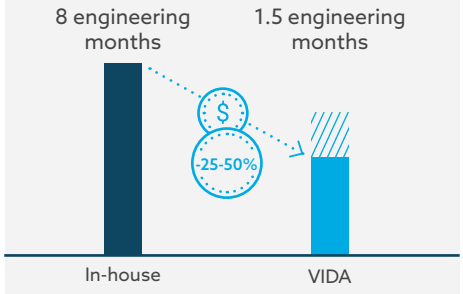
> **We are just scratching the surface when it comes to new data-driven tools that enable the mini-grid space.**
>
> – **SAM SLAUGHTER**, CEO PowerGen

5. USER BENEFITS

The goal of VIDA is to decrease the cost and project risk while increasing the speed of site selection at scale. VIDA is a swift decision-making tool that can fit into the workflow of different developers. In this road-test with PowerGen we measured the following key performance indicators (KPIs) and added values compared to the traditional (purely in-house) site selection process.

1

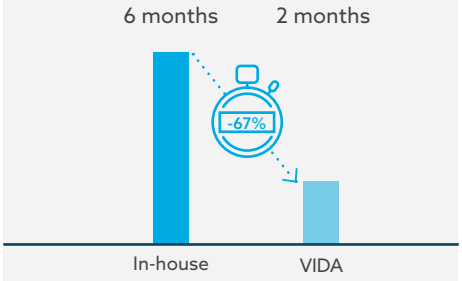
COST REDUCTION



VIDA saves 25-50% of the cost of man-hours during the site selection.

2

TIME TO MARKET



VIDA reduces the time to find a site by 67% and site visit time by 80% through better scheduling.

3

SCALE

VIDA allows for a large-scale, data-driven site selection process.

4

DE-RISKING INVESTMENTS

Additional information layers strengthen decision-making

VIDA finds more sites that are viable

VIDA leads to indirect cost savings and higher revenue generation through better site selection.

6. OTHER VIDA USERS

VIDA can substantially help other stakeholders in the electrification market as well:

Governments and development finance institutions...
...often perform country-wide analyses and surveys to identify optimal electrification strategies and make data-driven decisions about bundling sites for tender processes or designing subsidy schemes, such as performance-based grants. VIDA can help prepare tenders and ensure that chosen sites are viable and bankable. VIDA can also help design required subsidy levels across different village types in a country.

VIDA benefits
- Development of least-cost electrification strategies for off-grid areas
- Identification of mini-grid sites at scale for country-wide electrification tenders
- Support in determining required subsidy levels specific to different villages types
- Cross checking of survey results to improve planning quality

Off-grid energy product companies...
...(e.g. Solar Home Systems, solar lanterns), want to plan profitable growth into new sales regions, plan for optimal logistics, find sales partners, and reduce customer credit risks.

VIDA benefits
- Validation of customer credit analysis
- Provision of deep market intelligence on off-grid markets for expansion
- Location of underserved areas within existing markets
- Planning of supply chains based on expected demand

Investors and banks...
...are looking for financially viable energy access projects. Many, however, struggle with the size of investments and the level of market and customer information made available.

VIDA benefits
- De-risking of investments by providing information and data on existing or potential sites or customers areas
- Aggregation of mini-grid sites through scaling of the site selection process, using an independent, data-driven service

ABOUT

VIDA

Village Data Analytics (VIDA) is an AI-enabled solution that supports data-driven investment, business and policy decisions in remote, frontier markets. VIDA helps catalyze the required large-scale investment into frontier markets in Africa and Asia to meet global development goals and benefit around 2 billion people. A key barrier to effective development is the lack of information and transparency about remote, frontier markets, where many of the poorest people live. That makes planning and risk assessments imprecise, impact unmeasured and scale difficult to achieve. VIDA addresses this challenge with a combination of high-impact technology and deep, on-ground sector expertise. VIDA is used by governments and development organizations, by investors and banks and by the companies and NGOs offering impact products and services.

PARTNERS

esa European Space Agency

INITIATIVE FOR APPLIED ARTIFICIAL INTELLIGENCE

POWERGEN

PowerGen Renewable Energy develops, constructs and operated mini-grids It is considered by many to be the largest and most successful private minigrid developer operating in Africa today and has shown impressive growth. PowerGen has recently announced another significant inward investment from Shell and Sumitomo. In a clear signal of confidence in the minigrid sector, Shell CEO Ben van Beurden has set PowerGen the target of connecting 100 million people to their modern energy services by 2030. PowerGen is dedicated to digital solutions to improve its planning and operations.

TFE ENERGY

TFE Energy is a values-led business based in Germany and South Africa. We use digital technologies and hard-won practical experience of frontier markets and electrification to support the goal of universal energy access. Village Data Analytics is a product developed by TFE Energy with support from the European Space Agency (ESA) and its AI partner appliedAI.

CONTACT US

10 Franz-Joseph Str.
Munich 80801
Germany

contact@tfe.energy
www.tfe.energy
www.villagedata.io

Publisher

TFE Energy GmbH
10 Franz-Joseph Str.
Munich 80801
Germany

Design
www.simpelplus.de