

# Machine learning of redundant energy of a solar PV Mini-grid system for cooking applications

Richard Opoku<sup>a,b,\*</sup>, Gidphil Mensah<sup>a</sup>, Eunice A. Adjei<sup>a</sup>, John Bosco Dramani<sup>c</sup>, Oliver Kornyo<sup>d</sup>, Rajvant Nijjhar<sup>e</sup>, Michael Addai<sup>b</sup>, Daniel Marfo<sup>b</sup>, Francis Davis<sup>a</sup>, George Yaw Obeng<sup>a</sup>

<sup>a</sup> Department of Mechanical Engineering, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

<sup>b</sup> Sustainable Energy Service Centre, College of Engineering, KNUST, Kumasi, Ghana

<sup>c</sup> Department of Economics, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

<sup>d</sup> Department of Computer Science, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

<sup>e</sup> Consortio Limited, RG1 5RE Berkshire, United Kingdom

## ARTICLE INFO

### Keywords:

Solar PV mini-grids  
Household energy demand  
Redundant energy  
Machine learning  
Cooking applications  
Thermal batteries

## ABSTRACT

Solar PV mini-grids are increasingly being deployed in off-grid and island communities especially in sub-Saharan Africa (SSA) countries to meet household energy demand. However, one challenge of solar PV mini-grids for community energy supply is the mismatch between the PV energy generation and household energy demand. PV mini-grid energy generation is highest in the afternoon whilst household energy demand is highest in the mornings and evenings, but lowest in the afternoons. This mismatch creates redundant energy generation during peak sunshine hours when battery energy storage is full, leading to low profitability for mini-grid systems. In this study, four machine learning models have been applied on an installed 30.6 kW mini-grid system in Ghana to ascertain the level of the redundant energy. The study has revealed that redundant energy exists on the mini-grid, in the range of 56.98 – 119.86 kWh/day. Further analysis has shown that the redundant energy can support household cooking energy demand through sustainable thermal batteries. With the four machine learning (ML) models applied in predicting the redundant energy, the most accurate ML model, K-nearest Neighbour Regressor, had a root mean square error (RMSE) of 0.148 and a coefficient of determination ( $R^2$ ) value of 0.998.

## 1. Introduction

For many energy-deprived communities in developing countries [44] and sub-Saharan Africa (SSA) in particular, solar PV mini-grids have been found to be cost-effective and practical in providing energy access [42,5,47] to meet the 2030 target of clean and affordable energy for all (SDG 7) [33,50]. Investment in mini-grids in SSA has therefore increased for the past decade [49], targeting electricity supply for household energy needs such as light bulbs, ventilation fans, television set, refrigerators and phone charging. This has increased electrification rates for rural communities, usually for those far-removed from the national grid due to long distances.

In SSA, it is reported that about 1% of energy demand is met by the use of solar PV mini-grids [14]. In countries like Chad, where 75% of the population lives in rural communities [7,8,21], studies have shown that their electrification rate of 10% [24] could be substantially increased by use of solar PV mini-grids [20]. In a study on power generation in

Uganda, [17] found out that mini-grids can provide grid-like services at a relatively cheaper cost for off-grid rural communities, compared to if the grid was to be extended to those areas.

Solar PV mini-grids have been used to electrify very remote rural communities, both off-shore and islands communities, where there is no possibility of national grid extension [22,35]. In Ghana for example, as at the end of the year 2022, 5 mini-grids had been developed for off-grid island communities, beyond the Lake Volta. The Government is planning additional 35 mini-grids under the project: “*Scaling-Up Renewable Energy Program (SREP)*” to meet its renewable energy policy target of 10% RE penetration by 2030 [27].

For community-based solar PV mini-grids, battery energy storage (BES) needs to be installed to store energy during the daytime for continuous energy supply to the households during the evenings and early mornings when the sun is not available [40]. BES takes significant portion of community solar PV mini-grid investment (usually more than 30% of total capital cost) [6]. For many community-based mini-grids

\* Corresponding author at: Department of Mechanical Engineering, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana.

E-mail address: [ropoku.coe@knust.edu.gh](mailto:ropoku.coe@knust.edu.gh) (R. Opoku).

<https://doi.org/10.1016/j.solener.2023.06.008>

Received 19 April 2023; Received in revised form 31 May 2023; Accepted 5 June 2023

Available online 21 June 2023

0038-092X/© 2023 International Solar Energy Society. Published by Elsevier Ltd. All rights reserved.

**Table 1**  
Specification of the mini-grid.

Component	Size
Solar PV array	30.6 kW
Battery capacity	101.95 kWh
Number of connections	57 Households

that have been installed in SSA and Ghana in particular, lead-acid batteries have been used due to their relatively cheaper capital cost compared to lithium-ion batteries [50].

For community-based PV mini-grid systems, the BES is the most critical component for sustainability. Huge replacement cost for BES after every 3–5 years [9] when it has reached its end-of-life is a big challenge for sustainability of mini-grids. This is because payments made by customers (which is heavily subsidized, due to their economic status) is not able to match BES replacement cost. Many community-based mini-grids have therefore been found to be less profitable and unsustainable after battery end-of-life [12]. Enhancing the profitability and sustainability of community mini-grids is therefore of high interest to energy service companies (ESCOs), project developers, national governments and international donor agencies [13].

Generally, one technical challenge of solar PV systems for household energy supply is the mismatch between household energy demand and energy generation from the PV system. Household energy demand is highest in the mornings and evenings, whilst PV energy generation is highest in the afternoon [40,28]. This mismatch creates redundant energy generation and low profitability for solar PV systems. In a study conducted by [40] on a household, they defined a new parameter called “solar system redundancy factor (RF)”, which is the ratio of unused PV energy to the PV energy generation available after the batteries are fully charged. In their study, they highlighted that the redundant energy associated with the solar PV system of the household could be harnessed and used for productive purposes such as water pumping, cooking, water heating, etc.

From available literature, there is dearth of knowledge on the level of

redundancy with community-based solar PV mini-grid systems. Knowing the magnitude of the hourly and daily redundant energy of community-based solar PV mini-grid system can help to harness it for other productive means to increase its profitability. This study investigates through machine learning, the level of redundant energy from a 30.6 kW mini-grid system in an island community in Ghana. The metering system of the mini-grid, consisting of power analyzers and data loggers were used to monitor the total community energy demand and PV energy generation from the mini-grid. Machine learning of the energy demand and generation was then conducted to determine the magnitude of redundant energy and the possibility of adding cooking on the mini-grid. The findings of this study highlight the possibility of adding cooking onto the mini-grid during daytime peak sunshine hours to increase its profitability, at no extra cost and no mini-grid expansion.

The research methodology used for the study is presented in section 2. The results and discussion are also presented in section 3, and finally the conclusion of the study is presented in section 4.

## 2. Research methodology

### 2.1. Description of the mini-grid

The case study mini-grid comprises of a 30.6 kW solar PV array, 48 V–2124 Ah battery and a distribution grid that connects the community households to the solar PV system. The mini-grid has a total of 57 connected households. Table 1 presents the specification of the mini-grid. The mini-grid is one of the five mini-grids developed under the Ghana Energy Development and Access Project (GEDAP) as part of Ghana’s goal of increasing renewable energy penetration to 10% by 2030, which is in congruence with the Sustainable Development Goal (SDG) 7.

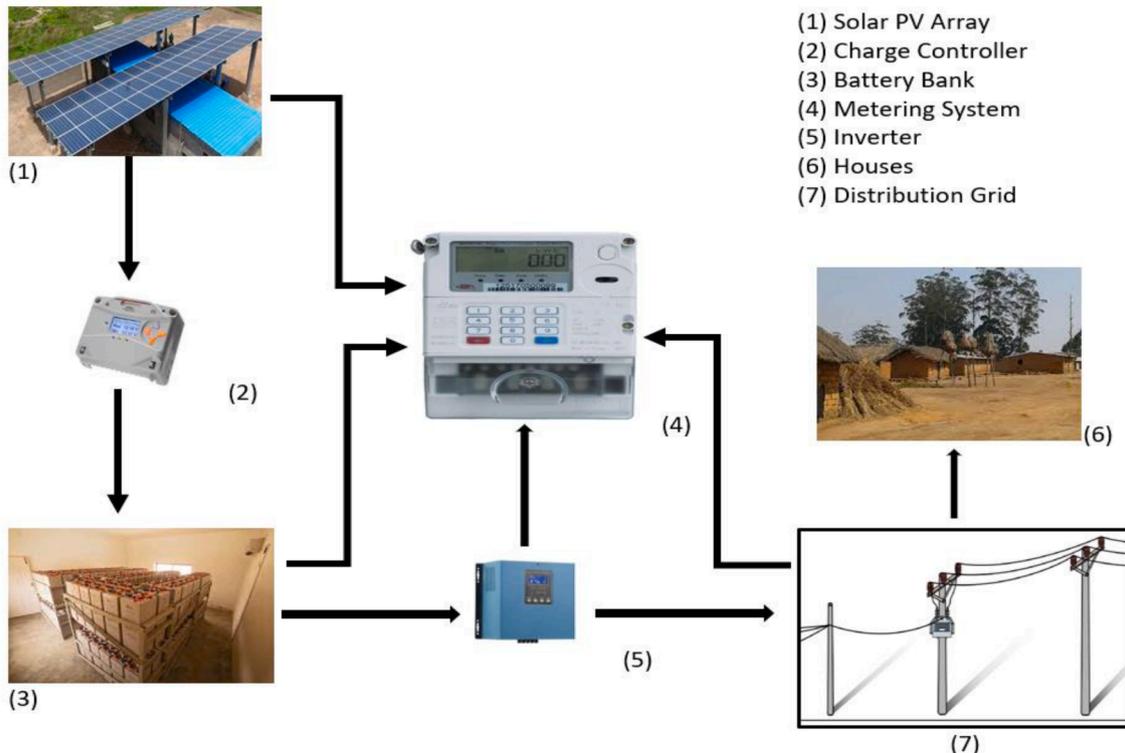


Fig. 1. Schematic diagram for data collection.

## 2.2. Data collection and analysis

### 2.2.1. Redundant energy from solar PV mini-grids

Community energy demand, which is typically for household usage, is usually out-of-phase (mismatch) with PV mini-grid energy generation. Household energy usage is highest in the mornings and evenings, whilst PV mini-grid energy generation is highest in the afternoon [40]. This mismatch creates redundant energy generation and low profitability for mini-grids. The redundant energy ( $E_{red}$ ) can be computed using Eq. (1).

$$E_{red} = E_{nom} - E_{gen,full} - \text{battery} - \text{charge} \quad (1)$$

where  $E_{nom}$  is the nominal PV generation from the mini-grid and  $E_{gen}$  is the actual energy generation from the mini-grid. It is important to emphasize that the redundant energy occurs during peak sunshine hours when the battery is fully charged and the consumer energy demand is lower than the mini-grid PV energy generation. When the battery is full, the actual PV energy generation is dictated by the amount of load on it (the energy demand). That is, when the battery is fully charged during peak sunshine hours, a solar PV system can only generate just enough to meet the instantaneous demand on it, even if it can generate more. The redundancy can also be expressed as a ratio, known as the ‘‘solar system redundancy factor (RF)’’ [40]. The RF is the ratio of the redundant energy (or unused PV energy) to the total available PV energy generation [40]:

$$RF = \frac{E_{red}}{E_{mini-grid}} \quad (2)$$

### 2.2.2. Nominal and actual PV generation

The nominal PV generation is computed using Eq. (3), where  $G$  is solar irradiance of the location,  $A$  is the area of a single PV panel,  $\eta$  is the rated efficiency,  $PR$  is the performance ratio, and  $n$  is the number of panels for the mini-grid system.

$$E_{nom} = G \times A \times \eta \times PR \times n \quad (3)$$

Data was collected on the actual PV generation (kW) and consumption (kW), all recorded at a time interval of one minute. It is important to mention that when the battery is not full, the actual PV generation is equal to the nominal PV generation, however, when the battery is full, the PV panel produces power just enough to meet the demand, such that the actual PV generation is equivalent to the demand (or the consumption on the mini-grid). In the data collection process, the mini-grid’s metering infrastructure, made up of power analyzers and data loggers, were used to collect data for a span of one year (365 days). The collected data was then used to compute the redundant energy of the solar PV system using Eq. (1). Fig. 1 shows a schematic diagram for the data collection. Radiation data that was used to compute the nominal power (Eq. (3)) was accessed from the National Solar Radiation Database of the National Renewable Energy Laboratory [38,46]. A detailed consumption profile of the mini-grid, based on the collected data, is presented in section 3.1.

## 2.3. Machine learning of redundant energy

The data obtained were analyzed using 4 different machine learning (ML) algorithms: Artificial Neural Network, K-nearest Neighbour Regressor, Random Forest Regressor, and XG Boost Regressor. Root Mean Square Error (RMSE) was employed as the metric to evaluate the performance of the machine learning models (Table 2). Six input parameters, namely, nominal power, actual generated power, consumption, month, day and minute, were used to train the various machine learning models to make predictions on redundant energy. The total count of the datapoints was 525,600, which corresponds to the number of minutes in 365 days of the year. Zero points in the dataset, which corresponds to times of the day when the sun is usually down, were removed, leaving

**Table 2**

Machine learning models and performance evaluation metric.

S/ N	ML Model	Evaluation Metric	Formula
1	Artificial Neural Network (ANN)	RMSE	$\sqrt{\frac{1}{n} \sum_{i=1}^n (Y_{p,i} - Y_{a,i})^2}$
2	Random Forest Regressor		
3	XG Boost Regressor		
4	K-nearest Neighbour Regressor		

only data from 7am to 5 pm for the machine learning of the redundant energy. 70% of the data was used to train the ML models while 30% was used for validation.

Where  $n$  = total number of data points,  $Y_{p,i}$  = predicted value,  $Y_{a,i}$  = actual value.

### 2.3.1. Data preprocessing

The collected data was preprocessed using python in Google Colaboratory and Microsoft Excel 2021. The data preprocessing was crucial in organizing the data to obtain accurate results. MATLAB and Google Colaboratory were used to run the machine learning models. There were no missing data points. The input data were normalized using the min–max normalization method to scale all input parameter values down to a range of 0 to 1. This was done to facilitate easy machine learning of the data. Equation (4) presents the min–max normalization Eq. [41,48].

$$\text{Min} - \text{max} (x) = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (4)$$

where  $x$  is the input data.

### 2.3.2. Artificial neural network

Artificial neural network (ANN) is a machine learning model that mimics the configuration and working of biological neural networks, like the human brain. Artificial neurons, sometimes known as ‘‘neurons,’’ are interconnected nodes that are arranged into layers to form ANNs [39]. A perceptron, commonly referred to as an artificial neuron, is the fundamental component of a neural network. Each neuron receives various inputs, gives them weights, and then runs the weighted sum through an activation function as shown in Fig. 2a. Based on the weighted total, the activation function chooses the neuron’s output. The neurons in the following layer are then given this output as input [34]. Eq. (5) presents the governing equation for this operation of the neurons. The activation function ‘ $F(y)$ ’ is presented as Eq. (6).

$$y = \sum_{i=1}^n (w_i x_i) + b \quad (5)$$

$$F(y) = f\left(\sum_{i=1}^n (w_i x_i) + b\right) \quad (6)$$

where  $y$  is the output/predicted value,  $n$  is the total number of input variables,  $w_i$  is the weight assigned to a particular neuron,  $x_i$  is the input variable and  $b$  is the bias.

Neurons are arranged into layers, having an input layer, one or more hidden layers, and an output layer as depicted in Fig. 2b. The output layer generates the final outcome or forecast, whereas the input layer receives the original data. Data processing and transformation occur during network transmission in the hidden layers. The network is trained with a set of data, to learn and make prediction. In training the neural network, a process called backpropagation is employed. Based on the error or difference between the desired output and the projected output, the network in this phase modifies the weights assigned to each connection between neurons. This adjustment aids the network’s continual learning and development of its forecasts [43,11].

Several types of neural networks exist, such as feedforward neural

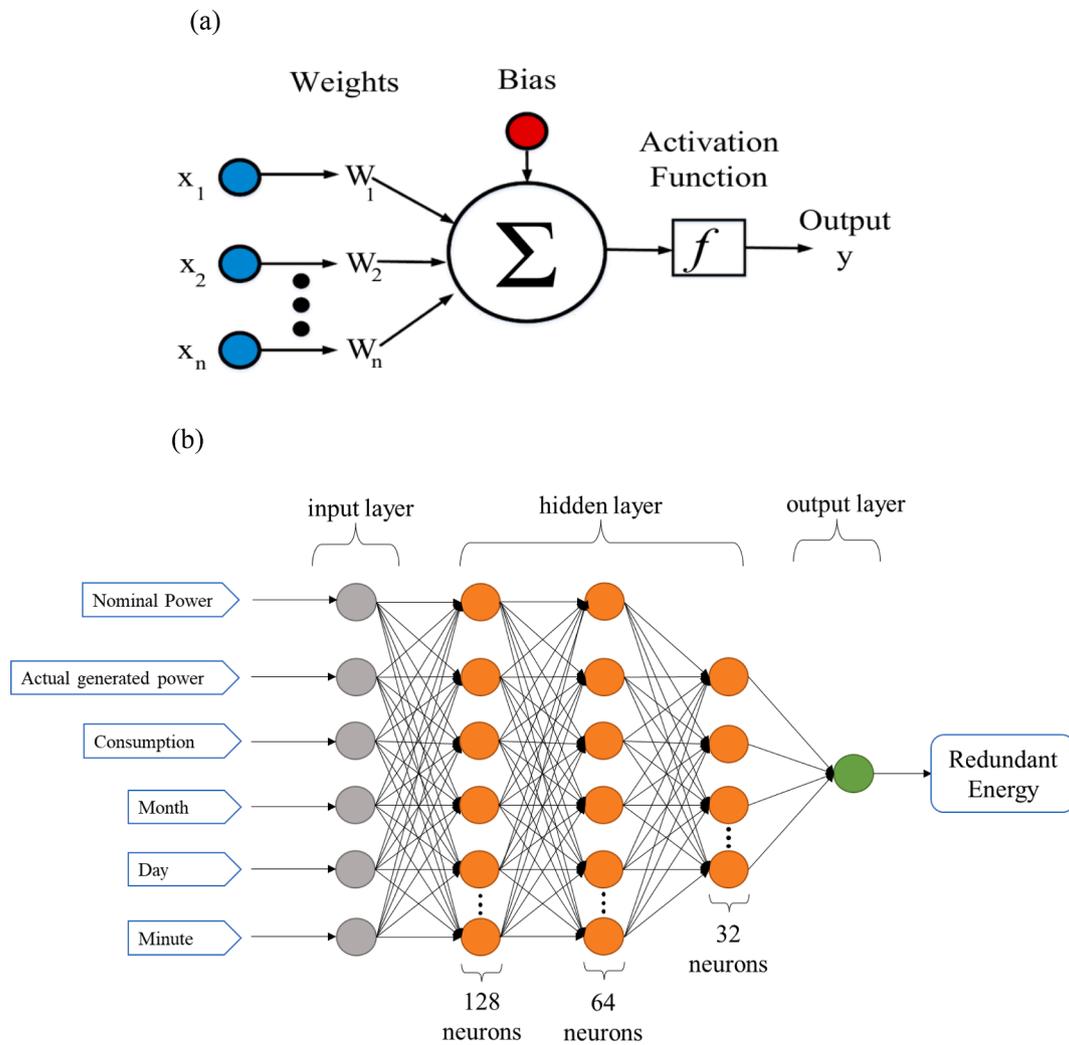


Fig. 2. (a) Schematic of an artificial neuron [34]. (b) Feedforward neural network model with 128, 64 and 32 neurons in hidden layers 1, 2 and 3, respectively.

networks, recurrent neural networks (RNNs), convolutional neural networks (CNNs), and generative adversarial networks (GANs). Each type has its own architectural modifications and training algorithms and is designed for particular tasks. As a key component in the field of artificial intelligence, neural networks have been successfully applied in numerous fields such as image and speech recognition, natural language processing, recommendation systems, and autonomous driving. They have demonstrated remarkable abilities to learn from large amounts of data and make complex predictions and decisions based on patterns and relationships in the data [36].

In this study, a feedforward neural network with three hidden layers was employed to process the six input parameters and make predictions of the redundant energy of the mini-grid. Fig. 2b is a schematic representation of a feedforward neural network as adopted in this study, showing the 6 inputs, three hidden layers (224 neurons = 128 + 64 + 32) and one output. The activation function used is the Sigmoid and Relu activation functions. The Levenberg Marquardt training function was used in the training of the model. 70% of the data set was used to train the model, while 30% was used for prediction and validation of the model.

### 2.3.3. Random Forest Regressor

A Random Forest Regressor is a machine learning model that belongs to the family of ensemble methods [19]. It is basically employed in regression problems where the objective is to forecast a continuous numerical value instead of a categorical label. The algorithm is based on

the concept of a random forest, which combines multiple decision trees to make predictions [48]. To use this algorithm, the data is first properly cleaned and formatted, to facilitate easy learning. The algorithm then creates an ensemble of trees to build a forest. Each tree, also termed base learners, is trained using a bootstrap sample, which is a randomly chosen subset of the training data [1]. Additionally, during the training process of each tree, only a random subset of features is considered for each split, which helps to introduce diversity among the trees. A process called recursive partitioning is used to train each decision tree on its corresponding bootstrap sample. It involves minimizing the prediction error by recursively splitting the data based on features and their thresholds [3,51]. The splitting process goes on until a stopping requirement is satisfied, such as when the tree has grown to its deepest point or when more splits do not significantly increase the predicted accuracy. Equation (7) presents the governing equation for this operation.

$$y(x) = \frac{1}{J} \sum_{j=1}^J h_j(x) \tag{7}$$

where  $y(x)$  is the ensemble predictor, and  $h_j$  is the base learner of the  $j$ th term.

To make predictions with the trained Random Forest Regressor, the predictions of each individual tree in the forest are combined by the algorithm. The final prediction for regression tasks is frequently calculated as the average or median of the predictions produced by each tree.

This aggregation helps to reduce the impact of individual noisy or biased predictions, leading to a more robust overall prediction. Random Forest Regressor is extensively used in several domains, like finance, healthcare, and retail, for tasks such as forecasting stock prices, valuing house prices, and forecasting customer demand. This study used 100 decision trees, and a maximum depth of 3, while the random state was set to 100.

#### 2.3.4. XG Boost Regressor

XGBoost (eXtreme Gradient Boosting) Regressor is a powerful machine learning algorithm that belongs to the gradient boosting family. It is particularly known for its efficiency, speed, and high predictive performance in regression tasks. XGBoost is designed to handle both numerical and categorical features and has gained popularity in various domains due to its effectiveness [16,31]. The mathematical representation of the XG Boost algorithm is presented in Eq. (8).

$$y_i = y_o + \eta \sum_{k=1}^n f_k(U_i) \quad (8)$$

where  $y_i$  is the predicted output for the parameter vector  $U_i$ ,  $n$  denotes the number of estimators,  $y_o$  denotes the mean of the parameters in the training data and  $\eta$  represents the learning rate of the model. XG Boost creates binary trees and assigns similarity weights (Eq. (9)) for the training of the model.

$$\text{Similarityweight} = \frac{\sum (\text{residuals})^2}{\text{numberofresiduals} + \alpha} \quad (9)$$

where  $\alpha$  is a hyperparameter, and the residuals is the difference between the mean and the output parameters.

XGBoost Regressor is based on the gradient boosting framework, where weak regression models (typically decision trees) are iteratively trained to correct the mistakes made by previous models [37]). The algorithm optimizes a loss function by minimizing the residuals or errors at each iteration, gradually improving the overall predictive accuracy. Regularization techniques are used in XGBoost to reduce overfitting and improve the model's generalizability. It controls the complexity of the model and lessens the impact of irrelevant features by using both L1 (Lasso) and L2 (Ridge) regularization terms in the objective function [4,29,45]. XGBoost provides insights into feature importance, allowing the contribution of each feature to be assessed in the regression task. Feature importance is computed based on the number of times a feature is used in the tree ensemble across all iterations, as well as the average gain in the loss function attributable to each feature. XGBoost Regressor has built-in capabilities to handle missing values. During the training process, it automatically learns how to handle missing data by creating default directions for missing values in the decision trees. XGBoost supports parallel processing, utilizing multiple CPU cores to accelerate the training process. This feature makes it particularly efficient when dealing with large datasets and complex models. XGBoost allows for early stopping, which means the training process can be halted when the performance on a validation set no longer improves. This helps prevent overfitting and saves computational resources. XGBoost Regressor has been widely adopted in various domains as it is effective in solving regression problems such as predicting housing prices, forecasting sales, and estimating customer demand. The combination of its boosting framework, regularization techniques, and efficient implementation makes XGBoost a powerful tool for regression tasks [2,32,18,30]. In this study, 100 n-estimators with a maximum depth of 3 were used.

#### 2.3.5. K-nearest Neighbour Regressor

The k-nearest neighbors (KNN) regressor is a supervised machine learning algorithm used for regression problems. The average or weighted average values of the target variable's k-nearest neighbors in the feature space is used to predict the value of the target variable [23,25]. During the training phase, the algorithm stores the feature vectors and corresponding target values from the training dataset. You

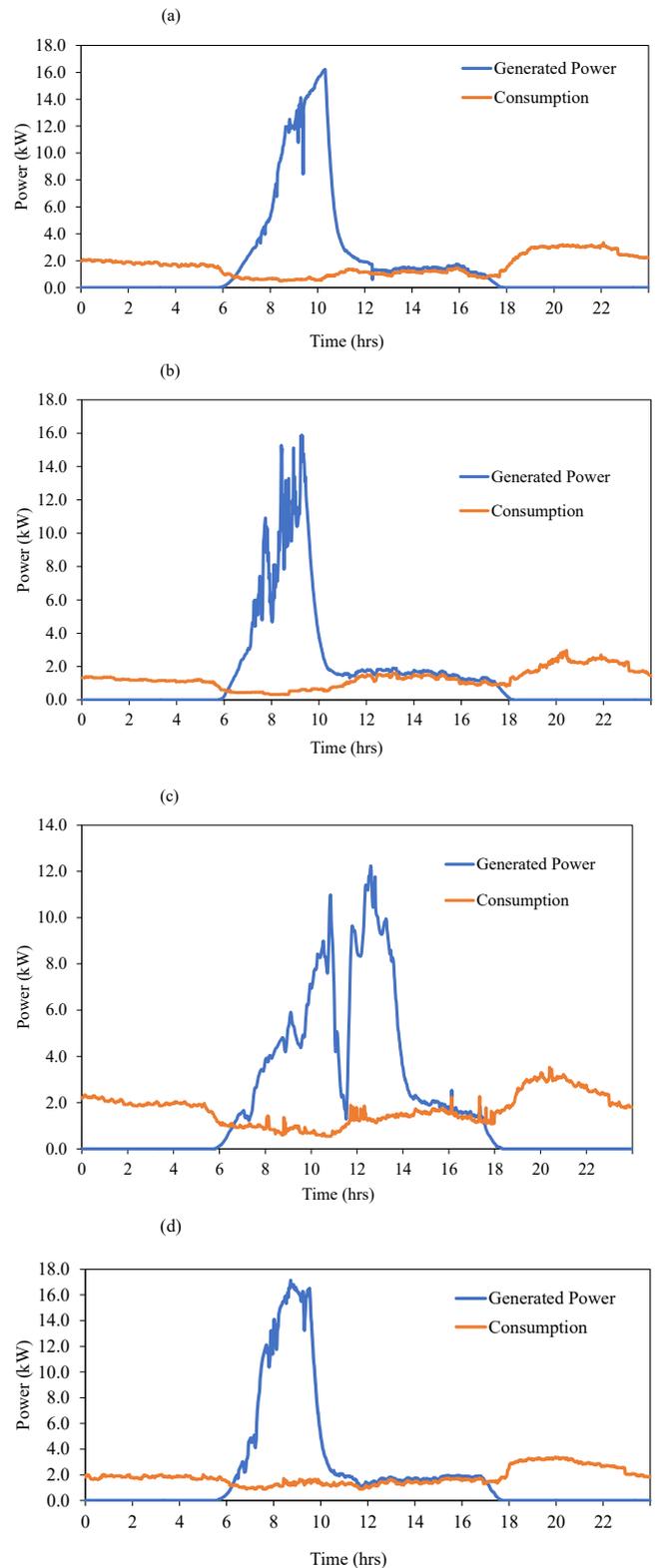


Fig. 3. System power (a) January 15th (b) April 15th (c) July 15th (d) October 15th.

need to specify the value of  $k$ , which represents the number of neighbors to consider for making predictions. The choice of  $k$  is an important decision and can affect the performance of the algorithm. A smaller value of  $k$  tends to give more flexible predictions, while a larger value of  $k$  provides smoother predictions. To make a prediction for a new input sample, the KNN regressor measures the distance between the sample

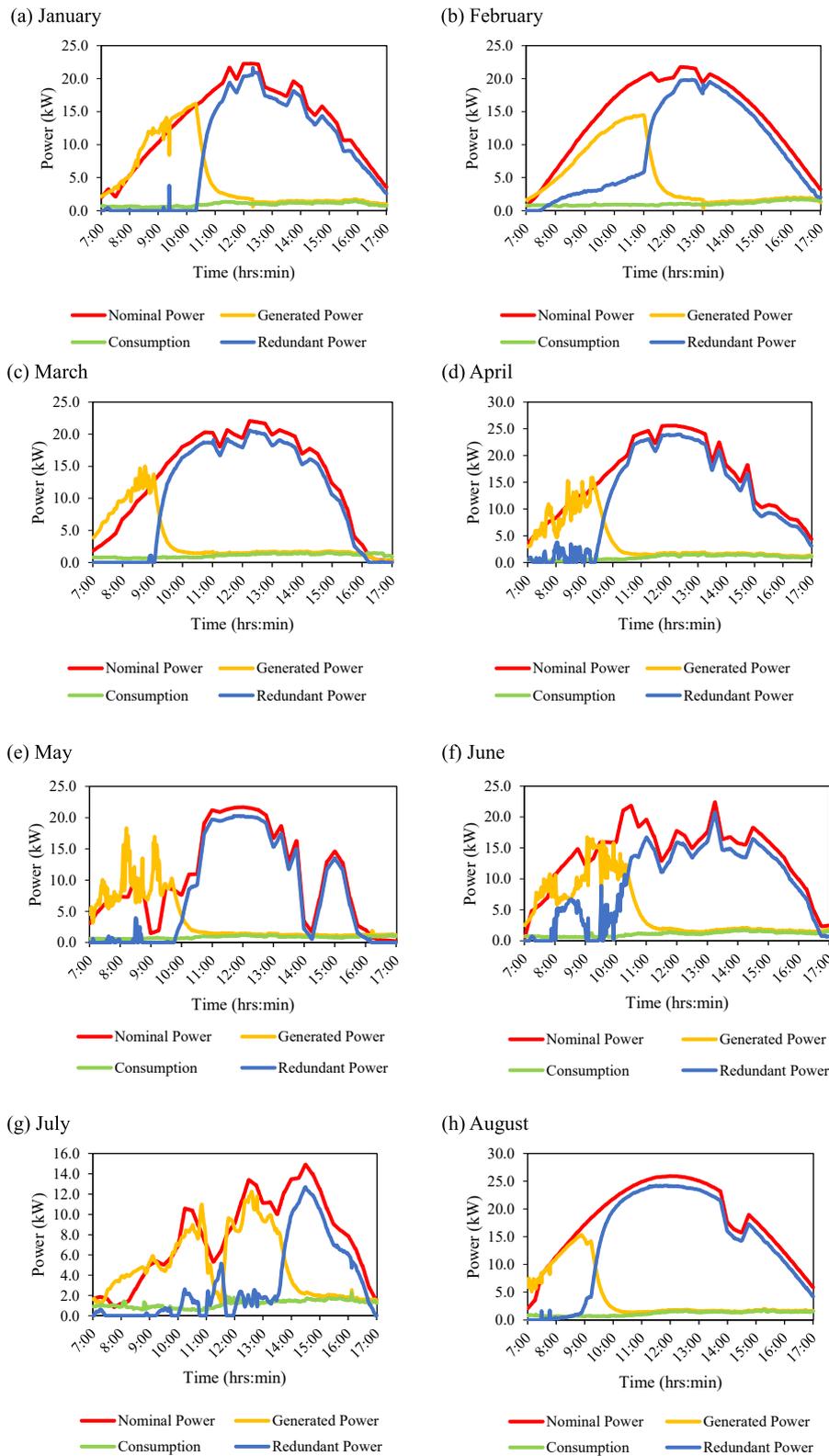


Fig. 4. Nominal generation, actual generation, consumption and redundant energy of the PV mini-grid.

and all the training samples using a distance metric, usually the Euclidean distance [15]. It then selects the k nearest neighbors based on the shortest distances. For regression tasks, the KNN regressor typically calculates the average or weighted average of the target values of the k

nearest neighbors. The weights can be assigned based on the distance or similarity of each neighbor to the input sample. KNN's mathematical model is presented as equation (10). This study used 5 k-neighbours in the training model.

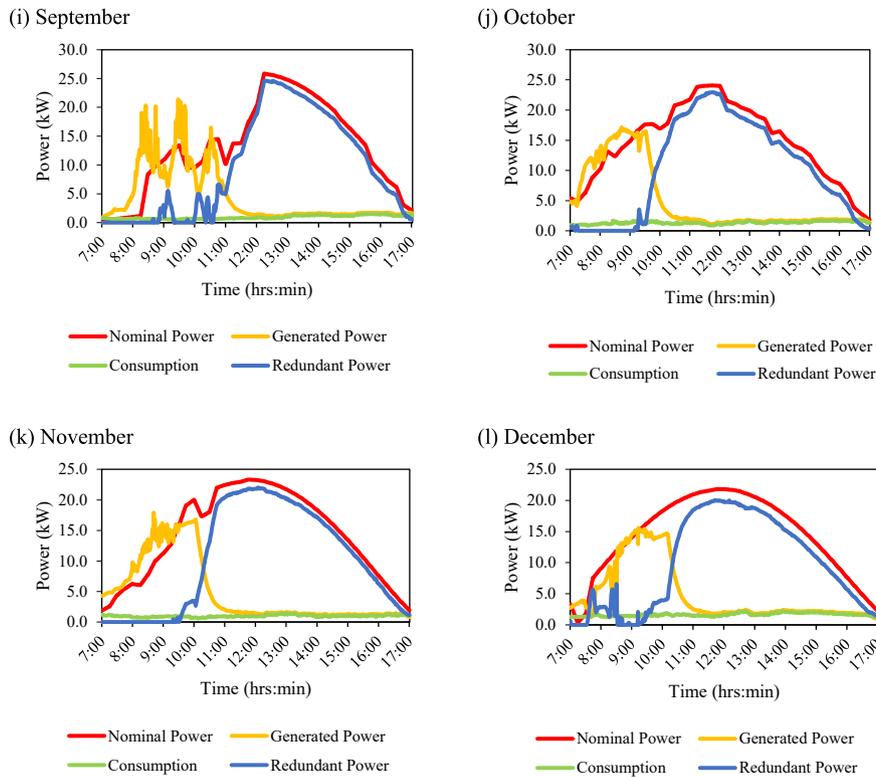


Fig. 4. (continued).

$$y_o = \frac{1}{k} \sum_{i=1}^k g_i \tag{10}$$

where  $y_o$  is the output and  $g_n$  is the neighbours,  $k$  is the total number of neighbours.

### 3. Results and discussion

#### 3.1. Daily PV energy generation and households' energy consumption

(a-d) presents graphs of the actual PV generation profile and load consumption (power demand) on the mini-grid for 4 selected days in January, April, July and October. For the purpose of analysis and comparison, the graphs are presented for the 15th day of each month. From Fig. 3, it is observed that there is a sharp increase in actual PV generation from around 6am, due to the rising of the sun at this hour of the day. The actual PV generation peaks between 10am and 11am, and sharply drops to the level of the load consumption at about 1 pm, and maintains that level till sunset at 6 pm. It is important to highlight that before the sharp drop in PV energy generation, that is from sunrise to about 1 pm, the actual PV generation is equal to the nominal generation.

The sharp drop in actual PV generation at about 1 pm is due to the fact that the battery gets fully charged at that time, so the PV generation drops to the level of the instantaneous energy demand of the community. During the time when the actual generation of the solar PV rises and peaks (between 6am and 11am), the solar PV supplies power both to the batteries and the community. However, at full battery charge, the power being drawn from the solar PV to the battery gets cut off by the charge controller, and only the community is supplied with power. A critical observation of the energy demand of the community reveals that the demand is highest after 5 pm because that is the time that the community inhabitants have returned from their businesses and workplaces. For the mini-grid under study, there is a diesel generator (back-up generator) connected to the system, but it is usually off, and no power

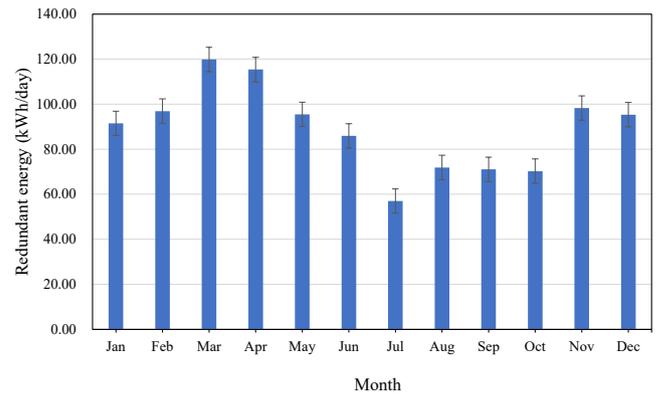
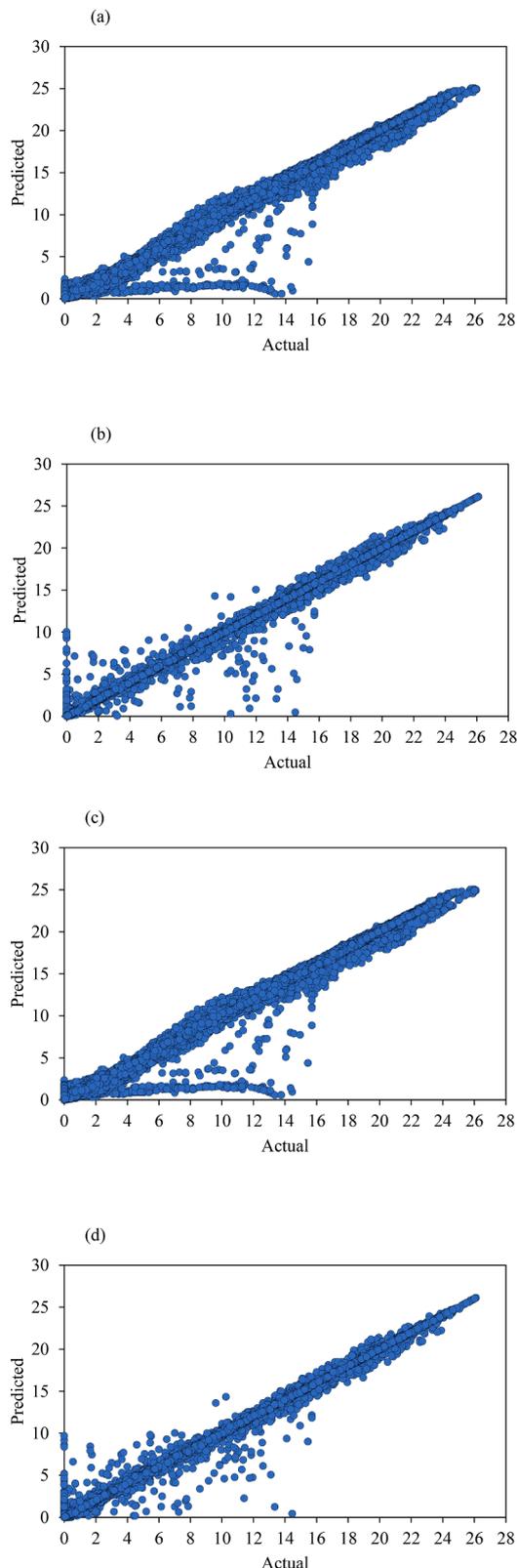


Fig. 5. Daily average redundant energy for each month of the year.

is generated from it.

#### 3.2. Redundant energy

Fig. 4 presents graphs of the nominal PV generation, actual PV generation, energy consumption and redundant energy of the mini-grid for the 15th day of each month of the year. As explained in section 3.1, the actual PV generation peaks at around 11am and drops sharply to the level of the consumer energy demand at about 1 pm due to full battery charge. If there were a place to push the PV energy generation to, the PV system would have continued to generate more energy (at nominal energy: red color in graphs) above the community energy demand during the sunshine hours. The nominal PV generation extends beyond 11am and peaks at about 12 pm and steadily decreases to the minimum at sunset (between 5 pm and 6 pm). The area under the nominal PV generation curve minus the area under the actual PV generation depicts the redundant energy of the solar PV mini-grid (the blue line).



**Fig. 6.** Regression plot of machine learning models (a) Artificial Neural Networks (b) Random Forest Regressor (c) XGBoost Regressor (d) K-nearest Neighbour Regressor.

**Table 3**

Root mean square error (RMSE) and R-squared values of the machine learning models.

ML models	RMSE	R <sup>2</sup>
K-nearest Neighbour Regressor	0.148	0.998
Random Forest Regressor	0.279	0.998
XG Boost Regressor	0.565	0.988
Artificial Neural Network	0.542	0.994

From the results of Fig. 4, it can be observed that generally, there is substantial redundant energy generation peaking at around 12 pm and gradually declining to about 4–5 pm when the sun sets. Analysis of the data showed that the redundant energy can be as high as 80–90% of the nominal generation during peak generation periods between the hours 12–3 pm. This redundant energy can be used to support other household energy demand such as cooking.

### 3.3. Redundant energy for e-cooking

Fig. 5 presents the daily average redundant energy available on the mini-grid for each month.

From the result of Fig. 5, it is observed that the highest daily average redundant energy was recorded in March (119.86 kWh/day) and the lowest in July (56.98 kWh/day). The available redundant energy could potentially be used as a means to increase the profitability of the mini-grid through e-cooking as modern energy service [10]. Using the redundant energy from the mini-grid for cooking in households could make mini-grids more profitable, without additional energy cost for cooking.

Typically, cooking a staple food (usually rice, yam, cocoyam, cassava, plantain, etc., with vegetable sauce, or beans stew) in a household of 3–5 people requires 0.8–2.22 kWh of energy [26]. With the available redundant energy from the mini-grid, 26–54 households cooking could be supported per day.

### 3.4. Results on machine learning of the redundant energy

A scatter plot of predicted redundant energy values versus actual redundant energy values were plotted for each ML model to visualize the performance (Fig. 6). It can be seen from all four graphs that a straight inclined regression line can be drawn through the points such that a large proportion of the points falls on the line or are very close to the line. This depicts that, generally, all four models perform well. However, it can be observed that among the four plots, the K-nearest Neighbour Regressor plot has the least points that do not align to the regression line. This puts the K-nearest Neighbour Regressor as the best performing model with the lowest root mean square error (RMSE).

Selection of the best model for predicting the redundant energy was based on two performance criteria as discussed in Section 2.3. Table 3 presents the root mean square error (RMSE) and coefficient of determination (R<sup>2</sup>) values of the four machine learning models used in this study. Gleaning from the table, K-nearest Neighbour Regressor had the lowest RMSE of 0.148 and the highest R<sup>2</sup> value of 0.998, denoting that it is the best performing model among the four for predicting redundant energy.

## 4. Conclusion

In this study, analysis has been conducted to ascertain the level of redundant energy that is available for e-cooking on a community-based solar PV mini-grid system, with a case study on a 30.6 kW system in Ghana. Four machine learning models have been employed to predict the redundant energy. From our study the following conclusions are made:

1. Redundant energy exists for community solar PV mini-grids during peak sunshine hours when the battery is fully charged. Average redundant energy in the range of 56.98–119.86 kWh/day was found to be available for ecooking, which can support 26–54 household cooking load.
2. Based on the performance metrics, the K-nearest Neighbour Regressor was found to be the most accurate predictor for redundant energy, with a root mean square error (RMSE) of 0.148 and a coefficient of determination ( $R^2$ ) value of 0.998.

## 5. Further studies

This study has ascertained the level of redundant energy of community-based solar PV mini-grid system. Using an installed 30.6 kW system, the study has shown that significant redundant energy exists for community mini-grid after 12 pm till sunset. It is recommended that future work should consider integrating ecooking on mini-grid to see how it minimizes the redundant energy and increases the profitability of the mini-grid.

## Funding

This research project was supported with funding from Modern Energy Cooking Services (MECS) / Loughborough University. MECS is funded by UKAid through the Foreign, Commonwealth and Development Office and is led by Loughborough University, UK in partnership with ESMAP. Please note, the views expressed in this report do not necessarily reflect the UK government's official policies.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

- [1] Ahmad, M.W., Reynolds, J. Rezgui, Y., 2018. Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees, *J. Clean. Prod.*, 203, 810–821. Available at: 10.1016/j.jclepro.2018.08.207.
- [2] Aler, R. et al. (2017) 'Improving the separation of direct and diffuse solar radiation components using machine learning by gradient boosting', *Solar Energy*, 150, pp. 558–569. Available at: 10.1016/j.solener.2017.05.018.
- [3] Babar, B. et al. (2020) 'Random forest regression for improved mapping of solar irradiance at high latitudes', *Solar Energy*, 198(March 2019), pp. 81–92. Available at: 10.1016/j.solener.2020.01.034.
- [4] Bae, D.-J., Kwon, B.-S., Song, K.-B., 2021. XGBoost-based day-ahead load forecasting algorithm considering behind-the-meter solar PV generation. *Energies* 15 (1), 128. Available at: 10.3390/en15010128.
- [5] Bukari, D. et al. (2022) 'Ex-post design, operations and financial cost-benefit analysis of mini-grids in Ghana: What can we learn?', *Energy for Sustainable Development*, 68, pp. 390–409. Available at: 10.1016/j.esd.2022.04.009.
- [6] Chaurasia, R., Gairola, S., Pal, Y., 2022. Technical, economic feasibility and sensitivity analysis of solar photovoltaic/battery energy storage off-grid integrated renewable energy system. *Energy Storage* 4 (1).
- [7] Didane, D.H., et al., 2016. Assessment of wind energy potential in the capital city of Chad, N' Djamena. In: *AIP Conference Proceedings*, p. 20049.
- [8] Didane, D.H., Rosly, N., Zulkafli, M.F., Shamsudin, S.S., 2017. (2017) 'Evaluation of wind energy potential as a power generation source in Chad'. *Int. J. Rotating Mach.* 2017, 1–10.
- [9] Diouf, B. and Pode, R. (2015) 'Potential of lithium-ion batteries in renewable energy', *Renewable Energy*, 76, pp. 375–380. Available at: 10.1016/j.renene.2014.11.058.
- [10] Eales, A., et al., 2022. 'Opportunities and Challenges for eCooking on Mini-grids in Malawi. *Case Study Insight* 8, 8–12.
- [11] Elsheikh, A.H. et al. (2019) 'Modeling of solar energy systems using artificial neural network: A comprehensive review', *Solar Energy*, 180(January), pp. 622–639. Available at: 10.1016/j.solener.2019.01.037.
- [12] Esan, A.B. et al. (2019) 'Heliyon Reliability assessments of an islanded hybrid PV-diesel-battery system for a typical rural community in Nigeria', *Heliyon*, 5(April), p. e01632. Available at: 10.1016/j.heliyon.2019.e01632.
- [13] ESMAP (2017) 'BENCHMARKING STUDY OF SOLAR PV MINI GRIDS INVESTMENT COSTS', (December).
- [14] ESMAP (2019) 'Mini Grids for Half a Billion People: Market Outlook and Handbook for Decision Makers. ESMAP Technical Report;014/19', The World Bank Group [Preprint].
- [15] Feng, L., Zhao, C., Huang, B., 2021. Adversarial smoothing tri-regression for robust semi-supervised industrial soft sensor. *J. Process Control* 108, 86–97.
- [16] Ge, J., et al., 2022. Prediction of greenhouse tomato crop evapotranspiration using XGBoost machine learning model. *Plants* 11 (15), 1923.
- [17] González-García, A. et al. (2022) 'A Rising Role for Decentralized Solar Minigrids in Integrated Rural Electrification Planning? Large-Scale, Least-Cost and Customer-Wise Design of Grid and Off-Grid Supply Systems in Uganda', *Energies*, 15(13). Available at: 10.3390/en15134517.
- [18] Grzebyk, D. et al. (2023) 'Individual yield nowcasting for residential PV systems', 251(January), pp. 325–336. Available at: 10.1016/j.solener.2023.01.036.
- [19] Han, S., Kim, H., 2021. Optimal feature set size in random forest regression. *Applied Sciences (Switzerland)* 11 (8), 1–13. Available at: 10.3390/app11083428.
- [20] Hassane, A.I. et al. (2022) 'Techno-economic feasibility of a remote PV mini-grid electrification system for five localities in Chad', *International Journal of Sustainable Engineering*, 15(1), pp. 179–193. Available at: 10.1080/19397038.2022.2101707.
- [21] Hassane, A.I., Hauglustaine, J.-M., Tahir, A.M., 2016. the Promotion of Renewable Energies: A Sustainable Answer to the Energy Problems of the Rural Households in Chad. *Int. J. Renew. Energy Resources* 6 (2), 45–52.
- [22] Hellmuth, M. et al. (2019) 'Integrated Resource and Resilience Planning (IRRP) Project: Climate Resilience Assessment of Mini-Grids in Ghana', (November).
- [23] Hu, C., Jain, G., Zhang, P., Schmidt, C., Gomadam, P., Gorka, T., 2014. Data-driven method based on particle swarm optimization and k-nearest neighbor regression for estimating capacity of lithium-ion battery. *Appl. Energy* 129, 49–55.
- [24] IEA (2019) Africa Energy Outlook 2019 – Analysis - IEA. Available at: <https://www.iea.org/reports/africa-energy-outlook-2019> (Accessed: 7 April 2023).
- [25] Johannesen, N.J., Kolhe, M., Goodwin, M., 2019. Relative evaluation of regression tools for urban area electrical energy demand forecasting. *J. Clean. Prod.* 218, 555–564.
- [26] Kajumba, P.K. et al. (2022) 'Assessment of the energy needs for cooking local food in Uganda: A strategy for sizing thermal energy storage with cooker system', *Energy for Sustainable Development*, 67, pp. 67–80. Available at: 10.1016/j.esd.2022.01.005.
- [27] Karembu, A., Tettey-cofie, J. and Abdoulaye, T. (2021) 'Ghana mini grid and solar pv net metering - appraisal report'.
- [28] Keddar, S. et al. (2021) 'An overview of the technical challenges facing the deployment of electric cooking on hybrid pv/diesel mini-grid in rural tanzania—a case study simulation', *Energies*, 14(13). Available at: 10.3390/en14133761.
- [29] Li, X. et al. (2022) 'Probabilistic solar irradiance forecasting based on XGBoost', *Energy Reports*, 8, pp. 1087–1095. Available at: 10.1016/j.egyr.2022.02.251.
- [30] Lopez-lorente, J. et al. (2023) 'Characterizing soiling losses for photovoltaic systems in dry climates : A case study in Cyprus', *Solar Energy*, 255(October 2022), pp. 243–256. Available at: 10.1016/j.solener.2023.03.034.
- [31] Lv, Q., et al., 2023. Modelling CO2 diffusion coefficient in heavy crude oils and bitumen using extreme gradient boosting and Gaussian process regression. *Energy*, 127396. Available at: 10.1016/j.energy.2023.127396.
- [32] Mahendiran, T.V., 2020. A color harmony algorithm and extreme gradient boosting control topology to cascaded multilevel inverter for grid connected wind and photovoltaic generation subsystems. *Sol. Energy* 211 (October), 633–653. Available at: 10.1016/j.solener.2020.09.079.
- [33] Maso, M.D. et al. (2019) 'Sustainable development impacts of nationally determined contributions: assessing the case of mini-grids in Kenya', *Climate Policy*, 20(7), pp. 815–831. Available at: 10.1080/14693062.2019.1644987.
- [34] Masoumi, A.P., Tajalli-Ardekani, E., Golneshan, A.A., 2020. Investigation on performance of an asphalt solar collector: CFD analysis, experimental validation and neural network modeling. *Sol. Energy* 207 (April), 703–719. Available at: 10.1016/j.solener.2020.06.045.
- [35] Mensah, G., et al., 2022. Stimulating green energy potential in Sub-Saharan Africa: An analysis of copper – copper sulphate thermogalvanic cell architecture. *Thermal Sci. Eng. Prog.* 34 (June), 101448. Available at: 10.1016/j.tsep.2022.101448.
- [36] Mohandes, S.R., Zhang, X., Mahdiyar, A., 2019. A comprehensive review on the application of artificial neural networks in building energy analysis. *Neurocomputing* 340, 55–75. Available at: 10.1016/j.neucom.2019.02.040.
- [37] Natras, R., Soja, B., Schmidt, M., 2022. Ensemble Machine Learning of Random Forest, AdaBoost and XGBoost for Vertical Total Electron Content Forecasting. *Remote Sens. (Basel)* 14 (15), 3547. Available at: 10.3390/rs14153547.
- [38] NREL (2019) *NSRDB: National Solar Radiation Database*. Available at: <https://nsrdb.nrel.gov/data-viewer> (Accessed: 7 April 2023).
- [39] Okwabi, R., et al., 2023. Towards the estimation of quantity of fuel consumed in steam generation through predictive modelling of feedwater temperature. *Scientific African* 20, e01650. Available at: 10.1016/j.sciaf.2023.e01650.
- [40] Opoku, R., Obeng, G.Y., Adjei, E.A., Davis, F., Akuffo, F.O., 2020. Integrated system efficiency in reducing redundancy and promoting residential renewable energy in countries without net-metering: A case study of a SHS in Ghana. *Renew. Energy* 155, 65–78.
- [41] Opoku, R. et al. (2022) 'Optimization of industrial energy consumption for sustainability using time-series regression and gradient descent algorithm based on historical electricity consumption data', *Sustainability Analytics and Modeling*, 2 (April 2021), p. 100004. Available at: 10.1016/j.samod.2022.100004.
- [42] Prinsloo, G., Dobson, R., Mammoli, A., 2016. Model based design of a novel Stirling solar micro-cogeneration system with performance and fuel transition analysis for rural African village locations. *Sol. Energy* 133, 315–330. Available at: 10.1016/j.solener.2016.04.014.

- [43] Qazi, A., et al., 2015. The artificial neural network for solar radiation prediction and designing solar systems: A systematic literature review. *J. Clean. Prod.* 104, 1–12. Available at: [10.1016/j.jclepro.2015.04.041](https://doi.org/10.1016/j.jclepro.2015.04.041).
- [44] Ranaboldo, M., et al., 2015. Off-grid community electrification projects based on wind and solar energies: A case study in Nicaragua. *Sol. Energy* 117 (2015), 268–281. Available at: [10.1016/j.solener.2015.05.005](https://doi.org/10.1016/j.solener.2015.05.005).
- [45] Rocha, P.A.C., Santos, V.O., 2022. 'Global horizontal and direct normal solar irradiance modeling by the machine learning methods XGBoost and deep neural networks with CNN-LSTM layers: A case study using the GOES-16 satellite imagery. *Int. J. Energy Environ. Eng.* 13 (4), 1271–1286. Available at: [10.1007/s40095-022-00493-6](https://doi.org/10.1007/s40095-022-00493-6).
- [46] Sengupta, M., et al., 2018. The National Solar Radiation Data Base (NSRDB). *Renew. Sustain. Energy Rev.* 89, 51–60. Available at: [10.1016/J.RSER.2018.03.003](https://doi.org/10.1016/J.RSER.2018.03.003).
- [47] Sevillano-bendezú, M.A. et al. (2023) 'Predictability and interrelations of spectral indicators for PV performance in multiple latitudes and climates', 259(May), pp. 174–187. Available at: [10.1016/j.solener.2023.04.067](https://doi.org/10.1016/j.solener.2023.04.067).
- [48] Talayero, A.P., et al., 2023. Machine Learning models for the estimation of the production of large utility-scale photovoltaic plants. *Sol. Energy* 254 (March), 88–101. Available at: [10.1016/j.solener.2023.03.007](https://doi.org/10.1016/j.solener.2023.03.007).
- [49] Troost, A.P., Musango, J.K. and Brent, A.C. (2018) 'Strategic Investment to Increase Access to Finance among Mini-Grid ESCOs : Perspectives from sub-Saharan Africa', *Proceedings - 2018 2nd International Conference on Green Energy and Applications, ICGEA 2018*, pp. 229–237. Available at: [10.1109/ICGEA.2018.8356268](https://doi.org/10.1109/ICGEA.2018.8356268).
- [50] Zebra, E.I.C. et al. (2021) 'A review of hybrid renewable energy systems in mini-grids for off-grid electrification in developing countries', *Renewable and Sustainable Energy Reviews*, 144(December 2020). Available at: [10.1016/j.rser.2021.111036](https://doi.org/10.1016/j.rser.2021.111036).
- [51] Ziame, A., et al., 2021. Photovoltaic output power performance assessment and forecasting: Impact of meteorological variables. *Sol. Energy* 220 (October 2020), 745–757. Available at: [10.1016/j.solener.2021.04.004](https://doi.org/10.1016/j.solener.2021.04.004).