Evaluation of solar radiation obtained from NASA and satellite imagery-based prediction models adjusted for microgrid sizing in Homer Pro

Evaluación de la radiación solar obtenida de la NASA y de modelos de predicción basados en imágenes satelitales ajustados para el dimensionamiento de microrredes en Homer Pro

DOI: http://dx.doi.org/10.17981/ingecuc.20.1.2024.13
Scientific Research Article. Date of Receipt:, Date of Acceptance:.

Luis Eduardo Ordoñez-Palacios https://orcid.org/0000-0001-5154-9472
Universidad del Valle – GUIA Research Group. Cali, (Colombia)
luis.ordonez.palacios@correounivalle.edu.co

Víctor Andrés Bucheli-Guerrero https://orcid.org/0000-0002-0885-8699
Universidad del Valle - GUIA Research Group. Cali, (Colombia)
victor.bucheli@correounivalle.edu.co

Eduardo Caicedo-Bravo https://orcid.org/0000-0003-0727-2917
Universidad del Cauca - PSI Research Group. Cali, (Colombia)
eduardo.caicedo@correounivalle.edu.co

To cite this paper

L. Ordoñez-Palacios, V. Bucheli-Guerrero & E. Caicedo-Bravo "Evaluation of solar radiation obtained from NASA and satellite imagery-based prediction models adjusted for microgrid sizing in Homer Pro" DOI: http://doi.org/10.17981/ingecuc.20.1.2024.13

Abstract

Introduction: The optimization of renewable energy resources is transcendental to satisfy the world energy demand and to avoid the adverse effects produced by the burning of fossil fuels. Therefore, there are several studies that seek to estimate the capacity of renewable energy sources in a geographical location. Likewise, there are several software applications that seek a balance between the investment and the installed capacity of an electric power generating plant.
Objective: This work uses the results of the Random Forest algorithm to predict solar radiation from satellite images. This technique achieved a performance in R^2 of 0.82 and in RMSE of 107.05. The purpose of this study is to evaluate the results of 2 models of photovoltaic systems designed for 10 different locations in the Colombian territory. Model M1 uses solar radiation data from NASA. The M2 model uses solar radiation data generated by Random Forest.
Methodology: The evaluation of solar radiation from NASA and the Random Forest algorithm is based on simulations provided by the energy resource optimization tool Homer Pro.
Results: The simulations of both models in Homer Pro show a difference in the capacity of the system components of between 0.0% and 47.31%. The difference between electric power generation ranges from 0.0% to 11.99%. Similarly, the difference between system costs is between 1.34% and 23.64% respectively.
Conclusions: The solar radiation data estimated by Random Forest is constituted as an alternative to the solar radiation data provided by NASA, given that the differences in the capacity of system components, electric power generation and total system costs are on average at around 27%.

Keywords

Electrical microgrids; Homer Pro; Renewable energy; Solar energy; Photovoltaic systems.

Resumen

Introducción: La optimización de los recursos de energía renovable es trascendental para satisfacer la demanda energética mundial y para evitar los efectos adversos producidos por la quema de combustibles fósiles. Por lo tanto, existen distintos estudios que procuran estimar la capacidad de las fuentes de energía renovable en una ubicación geográfica. Asimismo, existen diversas aplicaciones de software que buscan un equilibrio entre la inversión y la capacidad instalada de una central generadora de energía eléctrica.
Objetivo: Este trabajo utiliza los resultados del algoritmo Random Forest para predecir la radiación solar a partir de imágenes satelitales. Esta técnica alcanzó un desempeño en R^2 de 0.82 y en RMSE de 107.05. El objeto de este estudio es evaluar los resultados de 2 modelos de sistemas fotovoltaicos diseñados para 10 lugares distintos del territorio colombiano. El modelo M1 utiliza datos de radiación solar de la NASA. El modelo M2 utiliza datos de radiación solar generados por Random Forest.
Metodología: La evaluación de la radiación solar proveniente de la NASA y del algoritmo Random Forest está basada en las simulaciones proporcionadas por la herramienta de optimización de recursos energéticos Homer Pro.
Resultados: Las simulaciones de ambos modelos en Homer Pro, arrojan una diferencia en la capacidad de los componentes del sistema de entre 0.0% y 47.31%. La diferencia entre la generación de energía eléctrica oscila entre 0.0% y 11.99%. De igual manera, la diferencia entre los costos del sistema está entre 1.34% y 23.64% respectivamente.
Conclusiones: Los datos de radiación solar estimados por Random Forest se constituyen como una alternativa a los datos de radiación solar proporcionados por la NASA, dado que las diferencias en la capacidad de los componentes del sistema, la generación de energía eléctrica y los costos totales del sistema están en promedio, en alrededor del 27%.

Palabras clave

Microrredes eléctricas; Homer Pro; Energía renovable; Energía Solar; Sistemas fotovoltaicos.

I. INTRODUCTION

The global trend in electricity generation is related to the construction of large-scale hydroelectric power plants and the use of fossil fuels. In contrast, energy generation based on unconventional renewable energy sources remains insufficient, according to the Latin American Energy Organization (OLADE) [1]. Similarly, greenhouse gas emissions continue to increase, despite significant efforts being made to mitigate them [2].

The increasing energy demand and concern for global warming have led various governments and private organizations to consider energy transformation. In this regard, various renewable energy sources (small-scale hydroelectric, photovoltaic, solar thermal, wind, tidal, biomass, geothermal, among others) have become viable options due to being considered clean and not depleting their generating source [3].

According to the CONPES document [4] in Colombia, non-conventional renewable energy sources only represent 6% of the energy mix, which includes small-scale hydroelectric, photovoltaic systems, wind farms, and biomass power plants that process bagasse. In contrast, 63.3% of the installed capacity corresponds to large and small-scale hydroelectric power plants, and 30.7% is thermal energy. However, both energy sources are affected by phenomena such as El Niño, climate change, price volatility, and eventual depletion at local and global levels [5].

Although there are various renewable energy sources [6], the construction of power plants that enable their utilization depends on resource availability and electricity demand requirements. Therefore, monitoring stations exist to assess the potential of the source. However, the number of measurement points and available sensors is limited, which hinders the study of their behavior and leads researchers to build mathematical, statistical, and predictive models.

Predictive models applied to the electricity resource [7] [8] [9] [10] help determine the behavior of available energy sources in a specific location. However, the sizing, operation, and performance of electric power generation systems require the optimization of different input variables and the characteristics inherent to energy generation to contribute to the efficient utilization of the energy source. Consequently, finding a balance for power generation in terms of reliability, flexibility, and energy generation capacity has led to the design of various proprietary and open-source software tools that use optimization algorithms for the design of power generation plants.

This work arises from the research conducted by Ordoñez Palacios et al. [11], where a Machine Learning model is implemented using variables obtained from satellite images to estimate solar radiation at any location on the planet. The data generated by the mentioned model serves as one of the sources of information for the current project, aiming to build a microgrid model for the Colombian case.

In this research, the results of simulating different microgrid models designed in Homer Pro were evaluated. Each microgrid model was fed with data obtained from NASA and data generated by the Machine Learning model. The contribution of this work includes evaluating the simulation results in Homer Pro based on the two sources of solar radiation data. Additionally, it is possible to confirm that the data generated by the Random Forest algorithm are considered a valid alternative to the data provided by NASA.

It is also important to highlight that in this work: (i) The M1 microgrid model was designed using solar radiation data obtained from NASA. (ii) The M2 microgrid model was designed using solar radiation data generated by the Random Forest algorithm. (iii) Comparisons of the results obtained from each model were performed for a 25-year project.

This document comprises the following sections: Related Works, Materials and Methods, Results, Discussion, and finally, Conclusions.

II. RELATED WORKS

The research in [12] presents a review of the methods used for multi-objective optimization in hybrid power generation. The study considers design parameters such as the optimal configuration and location of the plant, operating costs for hybrid power generation, energy demand, emission reduction, total and social costs, cost-effective sizing of energy storage systems, and biomass system capacity. It utilizes multicriteria analysis as a tool for optimizing problems related to minimizing diesel fuel consumption, determining the optimal number of turbines, sizing the upper water reservoir, and load/discharge rates of the hydraulic pump storage system. This research employs techniques such as Neural Networks, Evolutionary Computing, Swarm Intelligence, and Fuzzy Systems.

A search conducted in the GitHub source code repository yielded information about related projects that build algorithms for optimizing power generation. The work of OffGridEnergy [13] presents an algorithm designed in Objective-C for estimating the size of a photovoltaic system and energy storage capacity for an off-grid renewable energy system. The algorithm calculates the cost of each off-grid configuration and determines the most cost-effective system. On the other hand, the project designed by SiyueZoe [14] includes optimization methods for scheduling household loads in smart grids, aiming to manage and control all elements of the smart grid. The main objective is to increase reliability and economic benefits, while also considering operational and probabilistic constraints on energy availability and consumption.

Other works, such as the project by jgarcia211 [15], implement a Recurrent Neural Network in Keras to forecast the energy consumption of a smart home. It optimizes the climatic parameters and considers the relationship between electrical energy consumption and the appropriate supply level. Understanding that excess supplied electricity cannot be stored unless it is converted into other forms of energy, generating additional costs and resources. Additionally, underestimating energy consumption could be detrimental, as excess demand leads to overload on the supply line, causing blackouts. The project by grhervas [16] aims to optimize electricity generation for the Spanish Electric Grid, adjusting the energy produced by different installations grouped by technology and region.

Currently, there are various software tools that facilitate the optimization of energy generation from various sources. Among them, RETScreen [17] is a clean energy management system for analyzing the feasibility of energy efficiency, renewable energy, and cogeneration projects, as well as continuous energy performance analysis. HOMER PRO [18] is a software tool for optimizing the design of isolated microgrids or those connected to public distribution grids. Its sensitivity analysis and optimization algorithms facilitate the evaluation of the technical and economic viability of energy projects.

On the other hand, the Model for Electric Technology Assessment (META) [19] facilitates comparative evaluation of the economic costs of various electricity generation and delivery technologies, including conventional (thermal, hydroelectric, etc.) and unconventional (renewable) options, as well as emerging options such as energy storage and carbon capture and storage (CCS). ENPEP-BALANCE [20] is a model for optimizing the balance of energy demand with available resources and technologies. It is based on a decentralized decision-making process in the energy sector and can be adapted to different preferences of energy users and providers.

HOMER Pro is a commercially used tool for hybrid optimization of multiple energy resources. Some studies, such as Ali et al. [21], Khalil et al. [22], Deshmukh and Singh [23], and Singh et al. [24], utilize the software for conducting techno-economic assessments of hybrid energy systems, optimizing, and designing hybrid energy systems, modeling the energy performance of autonomous photovoltaic systems, and computationally simulating the optimization of hybrid solar energy, fuel cell, and biomass energy systems, respectively.

III. MATERIALS AND METHODS

In this section, the research's guiding questions of interest, the sources of information that provided the data, how the data was processed, and the tools used to compare the microgrid models using data from NASA, machine learning algorithms, and climate monitoring stations in a specific location are formulated.

A. Research questions

Renewable energy sources can be transformed into electrical energy through systems designed for their utilization. While it is crucial to understand their behavior in a specific location using measurement instruments, mathematical or predictive models, it is also important to determine the optimal parameters for sizing power plants that guarantee reliability, flexibility, and energy generation capacity.

Considering the diversity of energy sources available in Colombia, it is important to analyze the optimal parameters that power generation plants must meet to ensure their proper functioning. Therefore, it is necessary to address questions such as: Q1: What tools are available for optimizing energy generation? Q2: Does the solar radiation generated by machine learning algorithms approximate that provided by NASA? Q3: Are solar energy sources in Colombia sufficient to generate electrical energy to meet the power demand? These questions are addressed throughout the entire document.

B. Sources of Information

In this work, a dataset from the ERA monitoring station of the Department of Environmental Management, DAGMA (Table 1), from the Municipality of Cali, was used. Additionally, 10 datasets were generated from satellite images from the years 2012, 2013, and 2014 for different regions of Colombia. The process of extracting image features is based on research by Ordoñez Palacios et al. [11]. Table 2 provides information about each location. Monthly average solar radiation data generated from predictive models and solar radiation resources from NASA were also used, as shown in Table 3.

ID	Station	Latitude	Longitude	Years	Hourly Records	Variables
1	ERA	3.44779	-76.51918	2012 - 2014	18705	Year, Month, Day, Hour, Wind Speed, Wind Direction, Temperature, Humidity, Rainfall, and Solar Radiation.

Table I. ERA weather station dataset from DAGMA

A Jupyter notebook with Python code was used to extract variables from satellite images from the years 2012, 2013, and 2014. The notebook and the data used in the model are available in a GitHub repository [25]. Ten locations were selected in different regions of the country, and for each case, 5821 hourly records were generated. The extracted characteristics include: 1. Reflectance: Represents the value of solar radiation reflected by clouds. 2. Cloudiness Index (nc): Related to cloud conditions, indicating clear sky, partly cloudy, or cloudy. 3. Number of Sunshine Hours (N): The duration in hours during which the sun has effective sunshine. 4. Extraterrestrial Solar Radiation (Hext): The value of electromagnetic radiation emitted by the Sun before entering the atmosphere.

ID	**Location**	**Municipality**	**Department**	**Latitude**	**Longitude**	**Population**
L1	Block A of Bosques de San Joaquín	Cali	Valle del Cauca	3.3730149	-76.547825	112
L2	San José del Guineo	Villagarzón	Putumayo	0.9451371	-76.6347924	90
L3	Santa Rosa de Juanambú Indigenous Reservation	Puerto Caicedo	Putumayo	0.7295021	-76.5913502	163
L4	Alto Lorenzo Indigenous Reservation	Puerto Asís	Putumayo	0.3651866	-76.5500819	101
L5	Alto Peñol	El Peñol	Nariño	1.4649621	-77.4768385	237
L6	El Nilo	Caloto	Cauca	3.0626769	-76.3683521	161
L7	La Victoria	Acevedo	Huila	1.7886154	-75.8959628	212
L8	San Luis del Plan	San Juanito	Meta	4.4706509	-73.6804890	137
L9	Caimital	Malambo	Atlántico	10.8673140	-74.7494810	154
L10	Puerto Alegría	Tarairа	Vaupés	-0.5384325	-69.6115196	56

Table II. Information about the selected locations in Colombia

Two microgrid models were designed for each location. Both models were fed with solar radiation and temperature resources. M1 Model uses solar energy data from NASA, while M2 Model uses solar energy data estimated by the Random Forest regression algorithm. Tables 3 and 4 present the monthly average solar radiation values from NASA and the predictive model, respectively. Solar radiation is represented in kWh/m^2/día.

Month	**L1**	**L2**	**L3**	**L4**	**L5**	**L6**	**L7**	**L8**	**L9**	**L10**
January	4.060	3.940	3.940	3.940	4.060	4.060	4.240	4.720	5.630	4.470
February	4.280	3.680	3.680	3.680	4.250	4.280	4.020	4.560	5.700	4.410
March	4.370	3.500	3.500	3.500	4.440	4.370	3.780	4.400	5.890	4.520
April	4.210	3.630	3.630	3.630	4.280	4.210	3.780	4.060	5.510	4.380
May	4.100	3.580	3.580	3.580	4.060	4.100	3.790	4.160	5.080	4.170
June	4.050	3.380	3.380	3.380	4.030	4.050	3.570	4.160	5.240	3.920
July	4.340	3.400	3.400	3.400	4.220	4.340	3.550	4.090	5.420	4.020
August	4.310	3.630	3.630	3.630	4.210	4.310	3.730	4.150	5.360	4.540
September	4.260	4.090	4.090	4.090	4.110	4.260	4.180	4.500	4.970	4.900
October	3.990	4.290	4.290	4.290	4.100	3.990	4.300	4.330	4.680	4.760
November	3.890	4.110	4.110	4.110	3.900	3.890	4.210	4.270	4.720	4.650
December	3.820	3.930	3.930	3.930	3.840	3.820	4.180	4.430	5.040	4.460
Annual average	**4.14**	**3.76**	**3.76**	**3.76**	**4.13**	**4.14**	**3.94**	**4.32**	**5.27**	**4.43**

Table III. Monthly average solar energy from NASA

Month	**L1**	**L2**	**L3**	**L4**	**L5**	**L6**	**L7**	**L8**	**L9**	**L10**
January	3.959	3.997	3.764	3.947	3.884	3.871	3.801	3.721	3.807	3.981
February	3.706	3.597	3.705	3.746	3.813	3.734	4.056	4.423	4.070	3.785

March	3.446	3.847	3.624	3.617	3.464	3.618	3.603	3.631	3.722	4.008
April	3.669	3.834	3.812	3.697	3.677	3.769	3.818	4.031	3.773	3.725
May	3.806	3.552	3.496	3.467	3.793	3.904	3.775	3.938	3.794	3.685
June	3.796	3.488	3.642	3.503	3.606	3.820	3.631	3.812	3.799	3.420
July	3.797	3.445	3.495	3.418	3.501	3.803	3.862	3.833	4.128	3.364
August	3.980	3.910	3.898	3.897	3.908	3.804	4.035	4.017	3.597	4.063
September	4.450	4.179	4.357	4.201	4.140	4.432	4.262	4.246	3.730	4.269
October	4.194	4.578	4.623	4.583	4.239	4.310	4.400	4.428	3.829	4.611
November	3.669	3.877	4.282	4.047	4.036	4.394	4.111	4.420	3.786	4.680
December	4.062	4.121	4.530	4.119	4.048	4.459	4.572	3.683	4.249	4.342
Annual average	**3.88**	**3.87**	**3.94**	**3.85**	**3.84**	**3.99**	**3.99**	**4.02**	**3.86**	**3.99**

Table IV. Monthly average solar energy estimated by the Random Forest algorithm

It is important to highlight that the monthly average solar radiation from NASA corresponds to a period of 22 years between June 1983 and June 2005. In contrast, the monthly average solar radiation estimated by the predictive model corresponds to the years 2012, 2013, and 2014.

C. Basic parameters of the models designed in Homer Pro

All models have a discount rate of 8%, two values for the inflation rate of 5% and 6% respectively, considering that the inflation rate in Colombia in 2021 was 5.62% according to La Republica [26], an annual capacity deficit of 0%, and a project lifespan of 25 years.

D. Estimation of the consumer load

The Solartex electricity consumption calculator [27] was used. This tool allows the input of the quantity of electrical appliances, their wattage, and the number of hours they are used per day. The calculations for the required electric load were estimated based on families of 4 members. Table 5 presents the data used by the calculator to estimate the average electric load for supplying the population of each location, considering the consumption differences between urban and rural areas. Table 6 presents the required electric load to supply the population of each selected location for the study.

Id.	Appliance	Urban Sector Quantity	Rural Sector Quantity	Consumption (watts)	Hours of Use	Urban Sector Energy	Rural Sector Energy
1	Fridge	1	1	80	24	80	80
2	Microwave	0.2	0	800	0.25	60	-
3	Oven	0.4	0	1500	0.6	600	-
4	Rice cooker	0.2	0.2	700	0.4	140	140
5	Juicer (Blender)	1	0.9	150	0.1	150	135
6	Extractor	0.2	0	150	0.3	30	-
7	Dishwasher	0.1	0	650	1	65	-
8	Kettle	0.1	0	850	0.2	85	-
9	Light bulbs	5	4	5	5	25	20
10	Fluorescent tube	0.5	0	18	5	9	-
11	Hair dryer	0.7	0.1	800	0.3	560	80
12	Washing machine	1	0.4	700	1	700	280

13	Dryer	0.1	0	2800	1	280	-
14	Clothes iron	0.9	0.1	1200	0.2	1080	120
15	Hair straightener	0.9	0.1	1200	0.1	1080	120
16	Electric coffee maker	0.7	0	10	0.3	7	-
17	LED TV	1.5	1.5	40	4	60	60
18	Audio system	0.8	0.8	70	4	56	56
19	Computer	0.4	0.1	200	4	80	20
20	Laptop	0.6	0.2	70	4	42	14
21	Printer	0.5	0.2	30	0.1	15	6
22	Router	1	0.5	25	24	25	13
23	Cell phone charger	3	3	30	2	90	90
24	Air conditioner	0.1	0	1700	4	170	-
25	Ceiling fan	0.2	0	600	2	120	-
26	Floor fan	0.6	0.3	900	2	540	270
Total Energy in 24 Hours (Watts/day) per family						7860	4050

Table V. Estimated consumption load by the Solartex calculator

ID	Location	Population	Number of Families	Consumption per Family (Watts/day)	Electric Load (kW/day)
L1	Block A of Bosques de San Joaquín	112	28.00	7860	220.08
L2	San José del Guineo	90	22.50	4050	91.13
L3	Santa Rosa de Juanambú Indigenous Reservation	163	40.75	4050	165.04
L4	Alto Lorenzo Indigenous Reservation	101	25.25	4050	102.26
L5	Alto Peñol	237	59.25	4050	239.96
L6	El Nilo	161	40.25	4050	163.01
L7	La Victoria	212	53.00	4050	214.65
L8	San Luis del Plan	137	34.25	4050	138.71
L9	Caimital	154	38.50	4050	155.93
L10	Puerto Alegría	56	14.00	4050	56.70

Table VI. Electric consumption load

All designed models include generic photovoltaic modules to produce 1 kW of electrical energy, a lifespan of 25 years, and a power reduction factor of 80%, using a direct current (DC) electrical bus. Additionally, a generic Li-Ion battery bank was configured to support 1 kWh of electrical energy generated by the solar farm, with a lifespan of 15 years and a capacity of 3 kWh, considering an initial state of charge of 100% and a minimum state of charge of 20%. Furthermore, a DC to AC converter with a capacity of 1 kWh and an efficiency of 95% was configured. In each case, the capital investment, approximate replacement costs, and annual operation and maintenance (O&M) costs were considered, as shown in Table 7.

ID	Component	Capacity	Capital	Replacement Costs	Operation and Maintenance Costs

1	Photovoltaic Modules	1 kW	$ 2000	$ 0	$ 250
2	Battery Bank	1 kWh	$ 2000	$ 1000	$ 100
3	Converter	1 kW	$ 700	$ 0	$ 0

Table VII. Costos de la microrred

E. Machine Learning Algorithms

In the research conducted by Ordoñez Palacios et al. [11], satellite images were obtained and solar radiation data from DAGMA and the Institute of Hydrology, Meteorology, and Environmental Studies (IDEAM) were used. The images were processed in Python, and the generated datasets were integrated with the solar radiation data. Machine learning algorithms were then trained using the extracted data from the images and the solar radiation as the target variable. The predictions made by the models were evaluated using different metrics.

The study employed machine learning algorithms such as Multiple Linear Regression, Support Vector Regression (SVR), Random Forest (RF), and Artificial Neural Networks (ANN). RandomizedSearchCV was used for hyperparameter tuning, and RepeatedKFold cross-validation was performed to improve the estimated performance of each model and prevent overfitting.

Based on the metrics of coefficient of determination (R^2) and root mean square error (RMSE), the Random Forest algorithm achieved the highest performance. It obtained an R^2 of 0.82 and an RMSE of 107.05. On average, it provided 5% more confidence and up to 10 points less error in predictions compared to the other evaluated algorithms. Therefore, this model was chosen as the base for predicting solar radiation in different geographical locations across Colombia using data extracted from satellite images.

F. Model Architecture

The model construction is based on the initial sources of information to design the microgrid, according to the tool's analysis requirements, the automatically generated simulations, and the results obtained from the application, classified according to resource optimization.

Figure 1 illustrates the flow of data, from the information sources to the generation of optimized models of microgrids, encompassing the definition of the components of each model's photovoltaic system, the calculation of the electrical consumption load, and the conducted simulations.

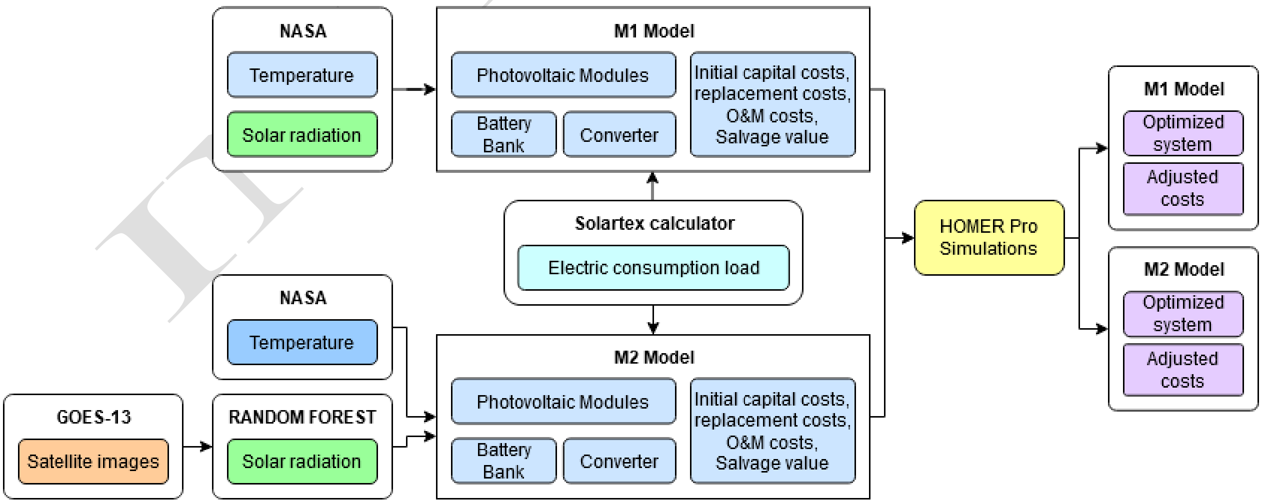

Fig. 1 General Architecture of the Model

IV. RESULTS

In Tables 3 and 4, the differences between the monthly average solar radiation from NASA and the monthly average estimated by the Random Forest predictive model can be observed. Based on the simulations conducted, the optimized structure of the photovoltaic system to meet the electrical consumption load in each model is shown in Table 8. Table 9 presents the results in terms of electrical generation for each model and the total system costs. Both tables include a percentage column representing the difference between M1 and M2 models.

ID	Location	Photovoltaic modules (kW)			Battery bank (kWh)			Converter (kW)		
		M1	**M2**	**%**	**M1**	**M2**	**%**	**M1**	**M2**	**%**
L1	Block A of Bosques de San Joaquín	237	264	**10.23**	343	334	**2.62**	62.5	72.9	**14.27**
L2	San José del Guineo	119	112	**5.88**	139	140	**0.71**	17.6	33.4	**47.31**
L3	Santa Rosa de Juanambú Indigenous Reservation	195	196	**0.51**	288	260	**9.72**	42.1	59.5	**29.24**
L4	Alto Lorenzo Indigenous Reservation	137	128	**6.57**	153	158	**3.16**	22.2	25.4	**12.60**
L5	Alto Peñol	257	280	**8.21**	372	380	**2.11**	67.5	78.6	**14.12**
L6	El Nilo	184	195	**5.64**	237	245	**3.27**	58.2	33.3	**42.78**
L7	La Victoria	270	248	**8.15**	323	340	**5.00**	52.3	97.3	**46.25**
L8	San Luis del Plan	153	157	**2.55**	200	219	**8.68**	44.1	29.8	**32.43**
L9	Caimital	134	190	**29.47**	211	257	**17.90**	38.4	33.8	**11.98**
L10	Puerto Alegría	56.6	68.1	**16.89**	93	93	**0.00**	12.4	14.3	**13.29**

Table VIII. Components of the system for each model

Table 8 presents the required capacity of the system components to supply the electrical load in the M1 and M2 models for each selected location. In the case of photovoltaic modules, the difference ranges from 0.51% to 29.47%. Regarding the battery bank capacity, the difference fluctuates between 0.0% and 17.90%. On the other hand, the converter capacity shows a difference ranging from 11.98% to 47.31%, respectively.

ID	Model	Energy Genertaion		Energy Consumption		Excess Energy		Total System Cost	
		Value	%	Value	%	Value	%	Value	%
L1	M1	268.42	**4.82**	80.29	**0.00**	179.36	**7.05**	$ 632.75	**4.66**
	M2	282.02		80.29		192.96		$ 663.69	
L2	M1	122.48	**3.30**	33.24	**0.00**	85.61	**4.72**	$ 286.03	**2.12**
	M2	118.44		33.24		81.57		$ 279.97	
L3	M1	201.23	**0.13**	60.20	**0.00**	134.44	**6.76**	$ 525.03	**4.04**
	M2	200.96		60.20		144.18		$ 503.80	
L4	M1	138.85	**4.60**	37.30	**0.00**	97.48	**6.56**	$ 323.56	**2.30**
	M2	132.46		37.30		91.09		$ 316.12	
L5	M1	297.25	**1.70**	87.55	**0.02**	200.12	**2.51**	$ 686.67	**5.41**
	M2	302.38		87.53		205.27		$ 725.95	
L6	M1	213.86	**2.46**	59.48	**0.02**	147.90	**3.53**	$ 466.56	**3.78**
	M2	219.26		59.47		153.31		$ 484.87	
L7	M1	293.57	**7.29**	78.30	**0.00**	206.75	**10.37**	$ 658.78	**1.34**
	M2	272.16		78.30		185.30		$ 649.98	
L8	M1	185.00	**4.43**	50.61	**0.00**	128.89	**6.38**	$ 390.72	**4.89**
	M2	176.80		50.61		120.67		$ 410.82	
L9	M1	186.82	**4.88**	56.90	**0.04**	123.77	**7.15**	$ 373.78	**23.64**
	M2	196.41		56.88		133.30		$ 489.50	
L10	M1	66.57	**8.18**	20.68	**0.00**	43.60	**11.99**	$ 160.80	**8.92**
	M2	72.50		20.68		49.54		$ 176,56	

Table IX. Electricity generation (kWh/year) and system costs

On the other hand, Table 9 shows the differences in the results of the Homer Pro simulations in terms of electricity generation, consumption, and excess energy for the models designed for the 10 geographical locations across the country. The results for electricity generation differ between 0.13% and 8.18%. Regarding energy consumption, the difference ranges from 0.0% to 0.04%. As for the excess energy, the results vary between 2.51% and 11.99%. Finally, the total cost of the system for both models range from 1.34% to 23.64% respectively.

V. DISCUSSION

It is important to highlight that in the literature review, several studies were found that design algorithmic models to optimize resources in hybrid energy generation. Additionally, various software tools were found that can be used as alternatives to Homer Pro. However, numerous studies employ Homer Pro to conduct evaluations of hybrid energy systems, optimize hybrid energy systems, model the energy performance of photovoltaic systems, among other case studies. In this regard, the contributions of this work are valuable because, in addition to analyzing the results of the simulations conducted in Homer Pro, it uniquely evaluates the relevance of the solar radiation data generated by the Random Forest algorithm compared to the data provided by NASA.

This research builds upon the work of Ordoñez Palacios et al. [11], in which the performance of 7 predictive models of solar radiation was evaluated using 4 different metrics. The best-performing model included data extracted from satellite images of locations below 800 meters above sea level. The Random Forest algorithm achieved an R^2 of 0.82 and an RMSE of 107.05. Subsequently, the aim was to design photovoltaic system models to evaluate the optimized results of costs and electricity generation for different locations in Colombian territory. In this regard, solar radiation provided by NASA and predictive models was considered.

As can be seen in Table 3, the monthly average solar energy values from NASA for geographical points L2, L3, and L4 located in the Putumayo department are identical. Similarly, locations L1 and L6 located in the Valle del Cauca and Cauca departments, respectively, are the same. This indicates a low level of detail in the solar radiation values provided by NASA. In the case of the estimated solar energy data from predictive models, it is possible to notice a greater granularity in the information, even if the sites are within the same region.

It is important to highlight that for the L9 site, located in Malambo, Atlántico department, the difference in the annual average solar radiation estimated by NASA and the Random Forest predictive model is over 1 kW/m^2. This indicates that the machine learning model is somewhat discreet in estimating solar radiation values in regions where solar energy is more concentrated.

It is important to note that the design of electric microgrids is sensitive to solar radiation resources. Therefore, differences in monthly average solar energy result in different outcomes for each optimized model. In this regard, simulations conducted by the tool indicate that, in some cases, the M2 model requires more investment; however, it produces a higher excess of energy compared to the M1 model. This enables the sale of surplus energy in the daily electricity market.

VI. CONCLUSIONS

This work compares the results obtained from the electric microgrid models designed in the Homer Pro analysis tool. The first model (M1) was fed with information from NASA, while the second model (M2) was fed with data estimated by the Random Forest algorithm. The annual average solar radiation estimated by NASA in all selected locations is 4.17 kWh, while the annual average estimated by the predictive model is 3.92 kWh.

In the estimation of the electricity consumption load for each location, carried out using the Solartex calculator, it was found that for rural areas, there is lower energy consumption per family due to the general use of fewer electrical appliances and less usage time for each one.

The optimization and sizing of electric microgrids require the selection of a small population in a specific geographical location, an approximate electricity consumption load, and the necessary components to configure the microgrid. Additionally, solar modules for the photovoltaic system, battery banks, and AC to DC power converters must be configured. For each system component, the capital costs, replacement costs, and operation and maintenance costs must be defined, along with their lifespan, power, and efficiency, as applicable.

In response to question Q1, it is worth mentioning that there are currently various software tools that facilitate the optimization of energy generation from different sources. Among them, RETScreen is a clean energy management system for the analysis of energy project feasibility. The Model for Evaluation of Electric Technology (META) enables comparative evaluation of the economic costs of various electricity generation and delivery technologies. ENPEP-BALANCE is a model for optimizing the energy demand balance with available resources and technologies. Homer Pro, the tool used in this study, allows for the optimization of the design of isolated microgrids or those connected to public power distribution networks.

Although there are differences in the results between the M1 and M2 models for each site located in different regions of Colombia, in most cases, and to answer question Q2, the differences in values are less than 20%. Therefore, it can be said that the solar radiation estimated by the Random Forest predictive model approximates the values provided by NASA.

Even though the Caribbean Coast is a strategic region with ideal conditions for solar generation in Colombia, as stated in the report by Lupe Mouthón [28], the study provides a positive response to question Q3. For all selected locations in the research, there is a viable configuration of electric microgrids to supply the electricity consumption load, according to the defined population in each geographical location.

VII. CRediT AUTHORSHIP CONTRIBUTION STATEMENT

Luis Eduardo Ordoñez-Palacios conceptualization, data curation, formal analysis, research, methodology, writing - original draft, writing - revision and editing**, Víctor Andrés Bucheli-Guerrero** supervision, validation, methodology, writing - proofreading and editing, **Eduardo Caicedo-Bravo** supervision, validation, methodology, writing - proofreading and editing.

ACKNOWLEDGEMENTS

We would like to thank the Universidad del Valle, the Facultad de Ingeniería, the Escuela de Ingeniería de Sistemas y Computación (EISC) and the Escuela de Ingeniería Eléctrica y Electrónica (EIEE).

REFERENCES

[1] OLADE, «Panorama Energético de América Latina y el Caribe». 2019. [Online]. Available in: http://biblioteca.olade.org/opac-tmpl/Documentos/old0434b.pdf

[2] X. Zheng, D. Streimikiene, T. Balezentis, A. Mardani, F. Cavallaro, and H. Liao, «A review of greenhouse gas emission profiles, dynamics, and climate change mitigation efforts across the key climate change players», *Journal of Cleaner Production*, vol. 234, pp. 1113-1133, Oct. 2019, doi: 10.1016/j.jclepro.2019.06.140.

[3] Energy Journal, «Energías Renovables No Convencionales (ERNC)», *Energy Journal*, August 10, 2012. https://www.revistaenergia.com/1135/ (Accessed April 27, 2021).

[4] CONPES, «Documento CONPES 3969». 2019. [Online]. Available in: https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/3969.pdf

[5] UPME, «Integración de las energías renovables no convencionales en Colombia». 2015. [Online]. Available in: http://www.upme.gov.co/Estudios/2015/Integracion_Energias_Renovables/INTEGRACION_ENERGIAS_RENOVANLES_WEB.pdf

[6] M. Fernandez, «Energías Renovables: qué son y tipos», *Otovo Blog*, May 25, 2021. https://www.otovo.es/blog/energia/que-son-las-energias-renovables/ (Accessed June 7, 2021).

[7] D. Yang, W. Wang, and X. Xia, «A Concise Overview on Solar Resource Assessment and Forecasting», *Adv. Atmos. Sci.*, Jan. 2022, doi: 10.1007/s00376-021-1372-8.

[8] S. Pereira, E. F. M. Abreu, M. Iakunin, A. Cavaco, R. Salgado, and P. Canhoto, «Method for solar resource assessment using numerical weather prediction and artificial neural network models based on typical meteorological data: Application to the south of Portugal», *Solar Energy*, vol. 236, pp. 225-238, Apr. 2022, doi: 10.1016/j.solener.2022.03.003.

[9] L. E. Ordoñez-Palacios, D. A. León-Vargas, V. A. Bucheli-Guerrero, and H. A. Ordoñez-Eraso, «Predicción de radiación solar en sistemas fotovoltaicos utilizando técnicas de aprendizaje automático», *Journal Facultad de Ingeniería*, vol. 29, n.o 54, Art. n.o 54, Sep. 2020, doi: 10.19053/01211129.v29.n54.2020.11751.

[10] A. Geetha et al., «Prediction of hourly solar radiation in Tamil Nadu using ANN model with different learning algorithms», *Energy Reports*, vol. 8, pp. 664-671, abr. 2022, doi: 10.1016/j.egyr.2021.11.190.

[11] L. E. Ordoñez Palacios, V. Bucheli Guerrero, and H. Ordoñez, «Machine Learning for Solar Resource Assessment Using Satellite Images», *Energies*, vol. 15, n.o 11, Art. n.o 11, ene. 2022, doi: 10.3390/en15113985.

[12] Syafaruddin, «Review on Multi-Objectives Optimization Methods in Hybrid Power Generation», *Journal of Engineering Science and Technology Review*, vol. 12, pp. 143-152, Feb. 2019, doi: 10.25103/jestr.121.17.

[13] OffGridEnergy, «OffGridEnergy/offgrid-solar-power-system». January 23, 2018. Accessed: June 7, 2021. [Online]. Available in: https://github.com/OffGridEnergy/offgrid-solar-power-system

[14] SiyueZoe, «SiyueZoe/FYP-Project». May 17, 2018. Accessed: June 7, 2021. [Online]. Available in: https://github.com/SiyueZoe/FYP-Project

[15] jgarcia2411, «jgarcia2411/Energy-Supply-Smart-Home». May 14, 2021. Accessed: June 7, 2021. [Online]. Available in: https://github.com/jgarcia2411/Energy-Supply-Smart-Home

[16] grhervas, «grhervas/Energy-production-Spain». June 22, 2020. Accessed: June 7, 2021. [Online]. Available in: https://github.com/grhervas/Energy-production-Spain
[17] RETScreen, «RETScreen», March 10, 2010. https://www.nrcan.gc.ca/maps-tools-and-publications/tools/modelling-tools/retscreen/7465 (accessed on June 9th, 2021).
[18] HOMER, «HOMER Pro - Microgrid Software for Designing Optimized Hybrid Microgrids», 2021. https://www.homerenergy.com/products/pro/index.html (accessed on June 9th, 2021).
[19] ESMAP, «Model for Electricity Technology Assessment (META) | ESMAP», 2021. https://www.esmap.org/node/3051 (accessed on June 9th, 2021).
[20] Argonne, «Energy Systems - Energy and Power Evaluation Program (ENPEP-BALANCE)», 2021. https://ceeesa.es.anl.gov/projects/Enpepwin.html (accessed on June 9th, 2021).
[21] F. Ali, M. Ahmar, Y. Jiang, y M. AlAhmad, «A techno-economic assessment of hybrid energy systems in rural Pakistan», *Energy*, vol. 215, p. 119103, Jan. 2021, doi: 10.1016/j.energy.2020.119103.
[22] L. Khalil, K. Liaquat Bhatti, M. Arslan Iqbal Awan, M. Riaz, K. Khalil, y N. Alwaz, «Optimization and designing of hybrid power system using HOMER pro», *Materials Today: Proceedings*, vol. 47, pp. S110-S115, Jan. 2021, doi: 10.1016/j.matpr.2020.06.054.
[23] M. K. Deshmukh y A. B. Singh, «Modeling of Energy Performance of Stand-Alone SPV System Using HOMER Pro», *Energy Procedia*, vol. 156, pp. 90-94, Jan. 2019, doi: 10.1016/j.egypro.2018.11.100.
[24] A. Singh, P. Baredar, y B. Gupta, «Computational Simulation & Optimization of a Solar, Fuel Cell and Biomass Hybrid Energy System Using HOMER Pro Software», *Procedia Engineering*, vol. 127, pp. 743-750, Jan. 2015, doi: 10.1016/j.proeng.2015.11.408.
[25] Ordoñez Palacios, «Notebook en Python y datos utilizados en el modelo de evaluación de los resultados de las simulaciones en Homer Pro, basados en dos fuentes de datos de radiación solar.» Accessed: April 24, 2023. [Online]. Available in: https://github.com/luise-phd/evalDatosAjustadosDimenMicrogids
[26] Editorial La República, «Colombia cierra 2021 con inflación de 5,62%, la más alta registrada en cinco años», *Diario La República*, January 5, 2022. https://www.larepublica.co/economia/colombia-cierra-2021-con-inflacion-de-5-62-la-mas-alta-registrada-en-cinco-anos-3283454 (Accessed on November 21, 2022).
[27] Solartex, «Calculadora de Consumo Eléctrico», *Solartex Colombia*, 2021. https://www.solartex.co/calculadora-de-consumo-electrico/ (Accessed on August 29, 2022).
[28] L. Mouthón, «La Costa Caribe, estratégica para desarrollar la generación solar», *EL HERALDO*, November 22, 2020. https://www.elheraldo.co/economia/la-costa-caribe-estrategica-para-desarrollar-la-generacion-solar-775110 (Accessed on September 1, 2022).

Luis Eduardo Ordoñez-Palacios Systems Engineer, Esp. in Project Management, Ph.D. Candidate at Universidad del Valle. Cali, Colombia. Member of the Artificial Intelligence Research Group (GUIA). https://orcid.org/0000-0001-5154-9472

Víctor Andrés Bucheli-Guerrero Systems Engineer, M.Sc. in Engineering and Computing, Ph.D. in Engineering. Full Professor at Universidad del Valle. Cali, Colombia. Director of the Artificial Intelligence Research Group (GUIA), Senior Researcher. https://orcid.org/0000-0002-0885-8699

Eduardo Caicedo-Bravo Electrical Engineer, M.Sc. in Information Technologies for Manufacturing, Ph.D. in Engineering. Full Professor at Universidad del Valle. Cali, Colombia. Director of the Perception and Intelligent Systems Research Group (PSI), Senior Researcher. https://orcid.org/0000-0003-0727-2917