Predictive Modeling of Photovoltaic Solar Power Generation

GIL-VERA V. D.
SISCO Research Group,
Luis Amigó Catholic University,
Trans. 51A N° 67 B-90, Medellín,
COLOMBIA

QUINTERO-LÓPEZ C.
NBA Research Group,
Luis Amigó Catholic University,
Trans. 51A N° 67 B-90, Medellín,
COLOMBIA

Abstract: - Photovoltaic solar power referred to as solar power using photovoltaic cells, is a renewable energy source. The solar cells' electricity may be utilized to power buildings, neighborhoods, and even entire cities. A stable and low-maintenance technology, photovoltaic solar power is an appealing alternative for generating energy since it emits no greenhouse gases and has no moving components. This paper aimed to provide a photovoltaic solar power generation forecasting model developed with machine learning approaches and historical data. In conclusion, this type of predictive model enables the evaluation of additional non-traditional sources of renewable energy, in this case, photovoltaic solar power, which facilitates the planning process for the diversification of the energy matrix. Random Forests obtain the highest performance, with this knowledge power systems operators may forecast outcomes more precisely, this is the main contribution of this work.

Key-Words: Forecasting, Generation, Machine Learning, Predictive Modeling, Solar Power.

Received: July 16, 2022. Revised: March 11, 2023. Accepted: April 9, 2023. Published: May 3, 2023.

1 Introduction

The utilization of renewable energy sources instead of fossil fuels has been emphasized as a way to reduce the carbon footprint globally. Global population growth is closely related to rising energy consumption and while old energy sources are becoming depleted, new energy sources are being investigated to fill the void, [1]. As a substitute treatment, more promotion and use of environmentally friendly energy sources are desired.

Solar photovoltaic (PV) energy is presented by this group. In 2010, solar energy generated less than 1% of the world's electricity, but by 2022, that percentage is projected to rise to over 28%, [2], throughout the last three decades. The local climate in the area where the system will be deployed has a strong correlation with the solar power-producing capacity of PV systems, [3], [4]. Among the performance models used to estimate generation, the models of [5], [6], [7], stand out.

These models illustrate the direct link between environmental parameters (solar radiation, air temperature, and wind speed) and the electrical output of a solar system. Even though these models are initial approximations, [8], [9], they do not sufficiently take into consideration radiation's changing and nonlinear character.

Because meteorological data variations and intermittencies affect energy output and the performance index of solar systems, [10], solutions that foresee and assess these changes are necessary. Predictive models are one of these alternatives since they focus on using data processing and analysis to find relationships, patterns, and/or trends. Models may be divided into two categories and applied to the development of prediction models. In the first, machine learning (ML) and artificial intelligence (AI) are utilized, whereas classical statistics are used in the second, [11], [12].

This work aimed to create a forecasting model for electricity generation for the next few days. For 34 days in a row (68,788 entries), a database was utilized to collect data on energy production from a solar photovoltaic power plant every 15 minutes.

The database's variables included the date and time of each observation, the amount of DC power the inverter produced in 15 minutes (in kW), the amount of AC power the inverter (source key) produced in 15 minutes (kW), the total amount of energy produced throughout the day, and the total return on investment.

The other sections of this work are: section 2 provides a broad background on photovoltaic solar power, section 3 offers the problem formulation, section 4 generalization of ML, section 5 problem solution, section 6 discussion, and finally conclusions.

In conclusion, predictive models provide improved network management, detect the need for panel cleaning/maintenance, and locate faulty or inefficient equipment. Predictive modeling may be used to forecast the long-term power reserve for PV system design and size as well as minimize generation uncertainty.

2 Photovoltaic Solar Power

A set of electrical and electronic components called a photovoltaic array uses sun radiation to generate electricity. The photovoltaic module, which is made up of cells capable of converting incident light energy into direct electrical energy, is the main component of this system, [13], [14].

The rest of the equipment in a photovoltaic system is primarily determined by the system's intended use. Grid-connected, off-grid, and pumped storage systems are the three basic categories under which photovoltaic systems may be categorized.

Grid-connected systems generate electricity that is supplied to the traditional grid; they are exempt from the requirement for energy storage as they are not directly responsible for meeting customer demand or ensuring consumption, [15], [16]. These systems, which may be separated into ground-mounted systems and building-mounted systems, comprise inverter equipment that adjusts the electricity supplied by the solar generator to the circumstances of the traditional grid to permit the proper linkage with the electrical grid, [17].

Above-ground systems often have more power than 100 kW and are entirely intended for energy generation and related economic efficiency. In-building systems perform tasks in addition to generating energy, such as replacing architectural elements, creating aesthetically pleasing effects, shading glass, etc.

They generally have power ratings of less than 100 kW and are smaller than ground-mounted systems, [18], [19].

The requirement to satisfy a particular energy demand unites the many uses for stand-alone systems. Because of this, almost all standalone systems have energy storage technology, [20].

According to their related applications, these systems may be divided into three groups: professional, rural electrification, and modest consumption. Little photovoltaic modules, frequently built of amorphous silicon, are used in modest consumption applications to power electronic devices like calculators or watches, mobile phone chargers, tiny power tools, household beacons, etc., [21], [22].

There are many professional applications, including radio links, cathodic protection of gas pipelines, hotels, traffic signals, air navigation, refrigeration of vaccines, equipment for remote data acquisition and transmission, and even power supply for satellites and other space equipment, [23].

Due to the extremely high costs associated with power failures in all of these applications, it is typically chosen to add solar generators and electrochemical accumulators that are bigger than technically necessary, [24], [25]. This reduces the likelihood of failure.

Often included in development cooperation projects and funded by nonprofit organizations or institutions like the World Bank or the European Union, rural electrification systems provide energy to rural communities that are located distant from traditional power lines, [26], [27]. Solar home systems (SHS), hybrid power plants, and pumped-storage systems are the most common types of rural electrification systems. Lighting devices, radio, television, and small power tools may all be powered by home systems and hybrid plants, respectively, [28], [29].

Domestic systems with 100 W or 200 W power ratings are often found in a family home, however, occasionally they can also be found in community centers or medical facilities. A rural village's electrical grid is provided by hybrid power plants with a solar generator, an electrochemical accumulator, and a generator set or wind turbine. The size of these plants depends on the population they serve, with capacity ranging from 10 kW to 100 kW, [30], [31]. Pumping systems employ a motor pump to raise and move water from an aquifer to a reservoir or distribution system by using the electrical energy generated by the solar generator, [32]. These systems often store energy in the form of potential energy from the water stored in the raised reservoir to minimize costs and boost dependability. Pumping systems can be used to desalinate water that has been extracted using reverse osmosis systems, deliver water for human or animal use, and irrigate private or public plantations, [33], [34].

An electrical configuration designed to employ photovoltaics to generate useable solar power is known as a photovoltaic system. A photovoltaic cell is a type of electrical device that directly transforms light energy into electricity by harnessing the physical and chemical phenomena known as the photovoltaic effect, [35], [36]. Moreover, it is the fundamental photovoltaic component that serves as the foundation for solar modules.

When a substance is exposed to light, the photovoltaic effect occurs, which produces voltage and electric current, [37]. Several solar cells linked in series and/or parallel and enclosed in an ecologically friendly laminate make up a photovoltaic module, [38]. A solar array's fundamental building component is a photovoltaic panel, a collection of modules, [39], [40]. A collection of solar panels that together form the entire photovoltaic-producing unit is called a photovoltaic array, [41], [58].

Photovoltaic inverters convert DC electricity from batteries or solar arrays to AC power for use with standard utility-powered appliances. The inverter acts as the brain of photovoltaic systems since the solar array is a DC source and it takes one to convert DC electricity to the common AC power used in our homes and offices, [42]. It is crucial to understand how weather conditions might affect the output of the two solar power plants since photovoltaic systems are heavily impacted by the weather; in good weather, we receive the most yield, while in bad weather, we get the least yield, [43], [44], [45]. The transition from a solar cell to a photovoltaic system is shown in Figure 1.

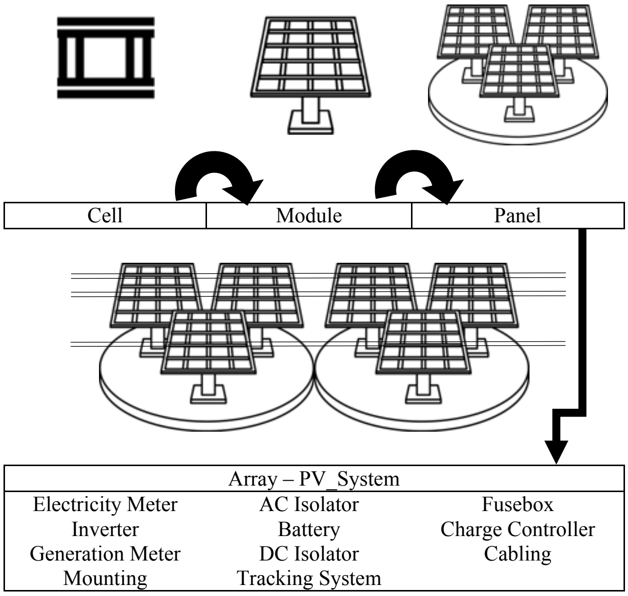

Fig. 1: From a solar cell to a PV system

3 Problem Formulation

Throughout 34 days, generation data were acquired at 15-minute intervals at the inverter level, where each inverter was connected to several lines of solar panels. Table 1 presents the variables that make up the power plant database.

Table 1. Database description

Variable	Description
Date_Time	Each observation's date and time. Observations were made and recorded every 15 minutes.
Plant_ID	Plant identification.
Source_Key	Inverter identification.
DC_Power	Amount of DC power produced over 15 minutes by the inverter (source key) (kW).
AC_Power	Amount of AC power produced over 15 minutes by the inverter (source key) (kW).
Daily_Yield	The total amount of energy produced in a day.
Total_Yield	Total investor return.

In the following GitHub link are available the generation database of the PV plant was analyzed for 34 consecutive days: https://acortar.link/w8yUqp

Sunlight is the cause of the Plant's Direct Current (DC) power production between 05:33:20 and 18:00:00, but else there is none. There are 22 inverters in the facility, each linked to several PV arrays. Each inverter captures its data every 15 minutes. Hence, to determine how much electricity the plant produced in an hour, we just compute the contribution of the 22 inverters. There are 22 inverters for data time on May 15, 2020, at 0:00. Except for the curve of May 20 and 25, which provides a consistent shape, nearly all the curves are the same despite some variation between 11 am and 2 pm. DC electricity is only at its peak on 2020-05-25. Data provides us with a logistics-like function, but after 18:00 the energy progressively declines until breaking down completely at 00:00. As you can see, certain daily yield dates (2020-02-06, 2020-05-19, etc.) have a logistic shape with missing values, but not others (2020-02-06, 2020-05-19, etc.). Data are logged every 15 minutes, and then we receive a fresh yield. Figure 2, presents the DC Power Plot.

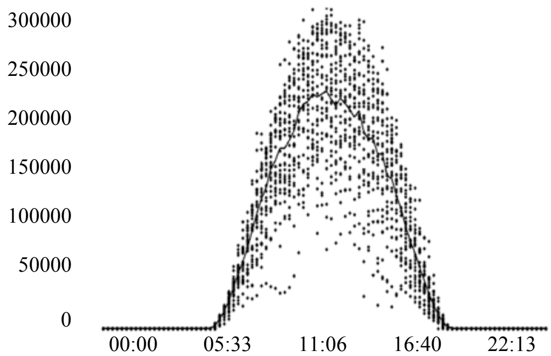

Fig. 2: DC Power Vs Time

Figure 3 presents the Daily DC Power on each day of the considered period (34 days).

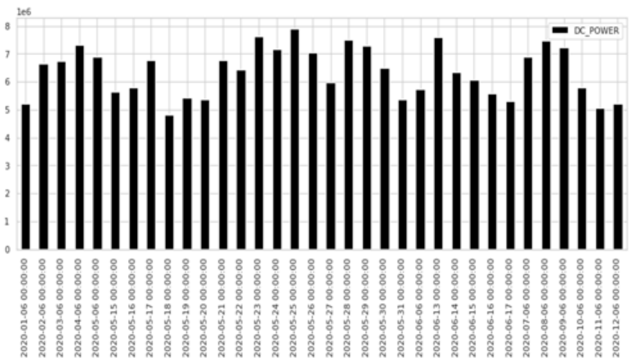

Fig. 3: Daily DC Power

4 Machine Learning

This computer science area is characterized by an artificial intelligence (AI) method, which is applied in many different disciplines of study, including biology, economics, and the energy sector, [46]. ML enables the creation of models that can make judgments that are challenging for explicit methods, such as straightforward numerical and analytical techniques, to describe, [47]. In [48], the authors assert that if representation is feasible, ML models can identify correlations between predictor variables and target variables.

Three steps make up ML: the first stage is the pre-processing and categorization of the data, the second stage is the data input, and the third stage is the handling of the discrepancy between the estimated and measured data, [49]. By taking into account the fact that the deviation and forecasting abilities of the models depend not only on the climatic conditions but also on the prediction horizon, criteria like the treatment of non-linearity, the behavior when using multiple inputs, the prediction horizon, the treatment of the deviation associated with the prediction, and flexibility, provide guidelines in model development, [50]. Next, we provide a summary of the ML models examined in this work:

4.1 Naïve-Bayes

This probabilistic ML method is frequently employed for classification problems. It is founded on the Bayes theorem, which calculates an event probability using information about prior confounding variables. This model assumes that the features used to categorize instances are independent of one another in the context of classification, i.e., one feature's existence or absence does not affect the presence or absence of any other feature, [51]. This model is based on the Bayes Theorem.

$$\mathrm{P(A|B)} = \frac{P(B|A).P(A)}{P(B)} \quad (1)$$

Where P(A) and P(B) are the odds of seeing A and B in the absence of any provided circumstances. P(A | B) is the probability that event A occurs given that B is true, and P(B | A) is the probability of the contrary case. A and B are events, and P(B) is different from zero.

4.2 Artificial Neural Network

ANN was developed based on how the human brain functions. Neurons, the linked layers of nodes that make up ANNs, process and transfer information via mathematical operations. A neuron is the fundamental unit of an ANN. It receives input from other neurons or external sources, computes an output using the weighted sum of the inputs, and then applies a nonlinear activation function. Using methods like backpropagation, the weights of the connections between neurons are learned using training data, [52]. Equation (2), presents the general equation of this model.

$$output = \mathrm{a}(z) = \mathrm{a}\left[\left(\sum_{\mathrm{i}=1}^{n} x_{i.}w_i\right) + b\right] \quad (2)$$

4.3 Support Vector Machine

This model is employed for classification and regression problems. When dealing with high-dimensional datasets and non-linear decision boundaries, SVMs are especially successful. Finding the hyperplane that optimally separates the various classes in the input data is fundamental in this model. When there are more than two classes, the hyperplane may be a plane or a higher-dimensional manifold, [53]. In a two-class classification issue, the hyperplane is a line that divides the two classes. Equation (3) presents a feature vector.

$$g(x) = w^t x + b \quad (3)$$

For a binary classification problem C_1, C_2, if $g(x) > 0$ then $x \in C_1$, else $x \in C_2$.

4.4 Logistic Regression

This model is a typical statistical learning approach for binary classification problems, where the objective is to predict the likelihood that an event will belong to one of two classes. A probability score may be the threshold for a binary logistic regression prediction result. The logistic function, on which this model is built, converts an output resulting from a linear combination of input data into a probability score between 0 and 1.

Weights are used to model the linear combination of input characteristics, and they are learned from training data via maximum likelihood estimation, [54]. Equation (4) presents the general equation of this model, where μ is the midpoint of the curve and s is a scale parameter.

$$\mathrm{p}(x) = \frac{1}{1 + e^{-(x-\mu)/s}} \quad (4)$$

4.5 Decision Tree

This model is employed for classification and regression problems. Each leaf node represents a class label or a numeric value, and each interior node represents a decision based on a feature or attribute, [55]. Recursively dividing the input space into subsets based on the values of the input characteristics is how decision trees are built, the data is divided into two or more subsets via top-down partitioning, where the most informative feature is chosen at each internal node, [55].

A decision tree may be easily converted into a collection of rules by mapping from the root node to the leaf nodes one by one. This model may be trained using many techniques, including ID3, C4.5, CART, and Random Forests, [55]. The objective is to reduce the impurity or entropy of the subsets, which quantifies the level of homogeneity of the class labels or values, [55]. In this model, entropy is a measure of the randomness in the information being processed and information gain (IG) is a decrease in entropy (equation 5).

$$\mathrm{IG} = \mathrm{Entropy}(before) - \sum_{\mathrm{j}=1}^{\mathrm{k}} \mathrm{Entropy}(j, after) \quad (5)$$

4.6 Random Forest

This model is used for supervised learning tasks including classification, regression, and others. An extensive number of decision trees are constructed during training using this ensemble learning approach, which results in a class that reflects the average of the predictions (regression) or classifications produced by the individual trees. This technique works by building a collection of decision trees, each of which is trained using a portion of the input features and training data that is randomly chosen. The random forest aggregates all of the individual trees' predictions throughout the prediction phase to provide a final prediction, [56]. In this model, the Gini Index (equation 6) is used to identify how much impurity has a particular node.

$$Gini\ Index = 1 - \sum_{j=1}^{c} (p_j)^2 \quad (6)$$

Where p_j is the proportion of samples belonging to class c for a given node.

4.7 K-Nearest Neighbors

In supervised learning, this model is used for classification and regression applications. It is a non-parametric approach because it makes no assumptions about the distribution of the underlying data. Based on a selected distance metric, such as Euclidean distance, the K-NN method finds the K data points in the training set that is closest to a given data point. The majority class or mean value of the K nearest neighbors in the training set is then used to forecast the output of the supplied data point, [57], [58]. The distance functions used in this model can be Euclidean (equation 7), Manhattan (equation 8), or Minkowski (equation 9).

$$d = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad (7)$$

$$d = \sum_{i=1}^{n} | x_i - y_i | \quad (8)$$

$$d = \left(\sum_{i=1}^{n}(|x_i - y_i |)^q \right)^{1/q} \quad (9)$$

Two current strategies utilized in photovoltaic solar power generation forecasting models are deep learning and statistical approaches. The authors of [59], propose a hybrid model that blends machine-learning approaches with the Theta statistical method to more accurately anticipate future solar power output from renewable energy facilities. Long short-term memory (LSTM), gate recurrent unit (GRU), AutoEncoder LSTM (Auto-LSTM), and a recently suggested Auto-GRU are among the ML models.

In [60], the authors suggest a fast-track methodology to handle two essential issues: long-term solar resource assessment and photovoltaic energy forecasting when investigating prospective locations for PV plant construction. These writers employed data clustering and probability techniques while exploring potential sites for PV plant construction.

In [61], the authors used numerous time-series algorithms to predict PV power generation output to respond swiftly to equipment and panel problems. The Long Short-Term Memory (LSTM) model exhibited the lowest error rate when compared to other models for quick PV power generation estimates, according to the study's findings.

5 Problem Solution

Many variables, including geographic location, sunshine intensity, solar panel efficiency, and meteorological conditions, can have an impact on the production of photovoltaic energy. However, even within the same place, it might change from day to day.

There are, however, a few approaches to foresee PV power generation at a certain site and time. One method is to employ solar radiation prediction models, which calculate the quantity of solar radiation that will arrive at a certain location at a specific time using meteorological and satellite data.

Based on the amount of sunshine and the present weather, real-time PV monitors may also be used to measure current generations and forecast future generations. Another approach is to estimate using previous PV generation data from the same site and season. In this work, the latter is utilized. After cross-validation. Table 2 lists the optimal parameters for each model examined.

Table 2. Optimal libraries and parameters

Model	Library & Optimal Parameters
Random Forest	RandomForestClassifier: {'n_estimators': 49}
Decision Tree	DecisionTreeClassifier: {'criterion':'gini','class_weight': 'balanced', 'max_depth': 5, 'max_features': 'log2, 'splitter': 'best'}
ANN-MLP	MLPRegressor: {'activation': 'relu', 'hidden_layer_sizes': 4, 'learning_rate': 'constant', 'solver': 'adam', 'learning_rate_init': 0.5}
K-NN	KNeighborsClassifier: {'n_neighbors':6}
Naïve Bayes	GaussianNB: {'max_features': 'auto', 'var_smoothing':1e-8}
Logistic Regression	LogisticRegression: {'C': 15, 'max_iter':6800, 'penalty': 'l2', 'tol': 1e-7}
SVM	SVC: {'C': 88, 'kernel': 'RBF', 'tol': 0.001}

Table 3 displays the findings for each of the examined accuracy measures.

Table 3. Training results

Model	Accuracy	Precision	Recall	Specificity	F1-Score
Random-Forest	.843	.848	.875	.793	.854
Decision Tree	.837	.827	.869	.773	.843
ANN-MLP	.707	.787	.861	.798	.882
SVM	.780	.765	.844	.690	.703
K-NN	.697	.675	.785	.566	.722
Naïve Bayes	.588	.537	.622	.476	.672
Logistic Regression	.543	.532	.521	.432	.563

This finding enables us to determine that Random Forest (Accuracy=.843, F1-Score=.854) was the model that performed the best. Decision Trees (Accuracy=.837, F1-Score=.843) and ANN-MLP (Accuracy=.707, F1-Score=.882) were two more models that did well. The accuracy of the positive predictions made by the ANN-MLP is lower than that of the first two models.

The models with the lowest ability to recognize negative instances (Specificity) were Naive Bayes and Logistic Regression; for both of these models, this measure was below .50, making them ineffective prediction models. The Logistic Regression model has the lowest efficiency, with an Accuracy of .543.

6 Discussion

Solar power projections will have a huge impact on the future of large-scale renewable energy installations. Predicting solar electricity generation is strongly reliant on changing weather patterns. Beyond climatic and altitude factors, the total amount of energy produced by a solar station depends on its capacity; thus, the total amount of energy produced by each solar power plant depends on its capacity.

A solar plant with a higher kilowatt peak (kWp) capacity will produce more energy than a plant with a lower kWp capacity. Because the variables that affect PV power generation are unpredictable, it can be difficult to predict it with any degree of accuracy.

With the use of machine learning (ML), it is now feasible to forecast power generation for the upcoming few days to manage the grid better, recognize the need for panel cleaning and maintenance, and spot broken or underperforming machinery.

Future studies might concentrate on estimating the yearly solar energy that a solar power plant is anticipated to yield based on site attributes and local meteorological data. Other environmental variables that may impact daily solar power output, in addition to the height of a solar power station's location, include temperature, wind speed, vapor pressure, solar radiation, day length, precipitation, and snowfall.

7 Conclusion

Forecasting photovoltaic power output is crucial for grid planning and management because it enables operators to better plan and manage energy supply and demand, which leads to lower costs and higher system efficiency. Due to weather fluctuation, the difficulty of detecting solar radiation, system capacity uncertainty, and a lack of historical data, this task may be challenging. It helps to increase prediction accuracy to utilize sensors and modeling.

The Random Forests method demonstrated the highest performance (Accuracy=.843, Precision=.848, Recall=.875, Specificity=.793, and F1-Score=.854); with this knowledge, power systems operators may forecast outcomes with more precision.

This ML model is a robust algorithm that can handle a large number of features, missing values, and noisy data. It is less prone to overfitting than other algorithms, such as decision trees, thanks to the use of bagging and feature randomness. Random Forests is known for its high accuracy in predicting outcomes, making it a popular choice for many applications.

This ML model is adaptable and may be used for both classification and regression problems. It can offer crucial insights into the link between the characteristics and the outcome, increasing the model's interpretability. Overall, ML integration can result in improved efficiency, reliability, and cost reductions in the photovoltaic power generation sector.

References:

[1] El-Hadary, M. I., Senthilraja, S., & Zayed, M. E., "A hybrid system coupling spiral type solar photovoltaic thermal collector and electrocatalytic hydrogen production cell: Experimental investigation and numerical modeling," Process Safety and Environmental Protection, vol. 170, p. 1101–1120, 2023. DOI: 10.1016/j.psep.2022.12.079.

[2] Mamatha, G., & Kulkarni, P. S., "Assessment of floating solar photovoltaic potential in India's existing hydropower reservoirs," Energy for Sustainable Development, vol. 69, p. 64–76, 2022. DOI: 10.1016/j.esd.2022.05.011.

[3] Shi, M., Lu, X., Jiang, H., Mu, Q., Chen, S., Fleming, R.M., Zhang, N., Wu, Y., & Foley, A.M., "Opportunity of rooftop solar photovoltaic as a cost-effective and environment-friendly power source in megacities," iScience, vol. 25, No. 9, p. 104890, 2022. DOI: 10.1016/j.isci.2022.104890.

[4] Yang, H., & Wang, H., "Numerical simulation of the dust particles deposition on solar photovoltaic panels and its effect on power generation efficiency," Renew Energy, vol. 201, p. 1111–1126, 2022. DOI: 10.1016/j.renene.2022.11.043.

[5] Veeramanikandan, M., Arjunan, T. V., & Jidhesh, P., "Experimental investigation of sandwich glazed solar photovoltaic module," Mater Today Proc, vol. 27, p. 136–139, 2020. DOI: 10.1016/j.matpr.2019.09.069.

[6] Lonergan, K.E., & Sansavini, G., "Business structure of electricity distribution system operator and effect on solar photovoltaic uptake: An empirical case study for Switzerland," Energy Policy, vol. 160, p. 112683, 2022. DOI: 10.1016/j.enpol.2021.112683.

[7] Kumar, N., & Pal, N., "Location and orientation-based performance analysis of 4.98 kW solar photovoltaic system for isolated Indian islands," Sustainable Energy Technologies and Assessments, vol. 52, p. 102138, 2022. DOI: 10.1016/j.seta.2022.102138.

[8] Wang, J., Chen, H., Deng, H., & Dong, F., "A combined power and steam system integrated with solar photovoltaic/thermal collector: Thermodynamic characteristics and cost-benefit analyses," Case Studies in Thermal Engineering, vol. 39, p. 102477, 2022. DOI: 10.1016/j.csite.2022.102477.

[9] Chirwa, D., Goyal, R., & Mulenga, E., "Floating solar photovoltaic (FSPV) potential in Zambia: Case studies on six hydropower power plant reservoirs," Renewable Energy Focus, vol. 44, p. 344–356, 2023. DOI: 10.1016/j.ref.2023.01.007.

[10] Özçelep, Y., Bekdaş, G., & Apak, S., "Meeting the electricity demand for the heating of greenhouses with hydrogen: Solar photovoltaic-hydrogen-heat pump system application in Turkey," Int J Hydrogen Energy, vol. 48, No. 7, p. 2510–2517, 2023. DOI: 10.1016/j.ijhydene.2022.10.125.

[11] Hasheem, M.J., Wang, S., Ye, N., Farooq, M.Z., & Shahid, H. M., "Factors influencing purchase intention of solar photovoltaic technology: An extended perspective of technology readiness index and theory of planned behavior," Cleaner and Responsible Consumption, vol. 7, p. 100079, 2022. DOI: 10.1016/j.clrc.2022.100079.

[12] Pimpalkar, R., Sahu, A., Patil, R. B., & Roy, A., "A comprehensive review on failure modes and effect analysis of the solar photovoltaic system," Mater Today Proc, 2022. DOI: 10.1016/j.matpr.2022.11.353.

[13] Manoj Kumar, N., Chakraborty, S., Kumar Yadav, S., Singh, J., & Chopra, S.S., "Advancing simulation tools specific to floating solar photovoltaic systems – Comparative analysis of field-measured and simulated energy performance," Sustainable Energy Technologies and Assessments, vol. 52, p. 102168, 2022. DOI: 10.1016/j.seta.2022.102168.

[14] Al-Ghussain, L., Taylan, O., Abujubbeh, M., & Hassan, M.A., "Optimizing the orientation of solar photovoltaic systems considering the effects of irradiation and cell temperature models with dust accumulation," Solar Energy, vol. 249, p. 67–80, 2023. DOI: 10.1016/j.solener.2022.11.029.

[15] Nej, S.K., Sreejith, S., & Chakraborty, I., "Dual-Output Multistage Switched-Capacitor Quadratic Boost (MSC-QBC) DC-DC Converter for Solar Photovoltaic Application," IFAC-PapersOnLine, vol. 55, No. 1, p. 965–970, 2022. DOI: 10.1016/j.ifacol.2022.04.159.

[16] He, B., Lu, H., Zheng, C., & Wang, Y., "Characteristics and cleaning methods of dust deposition on solar photovoltaic modules-A review," Energy, vol. 263, p. 126083, 2023. DOI: 10.1016/j.energy.2022.126083.

[17] Xu, Z., Kong, Q., Qu, H., & Wang, C., "Cooling characteristics of solar photovoltaic panels based on phase change materials," Case Studies in Thermal Engineering, vol. 41, p. 102667, 2023. DOI: 10.1016/j.csite.2022.102667.

[18] Macías, R.J., Ceballos, C., Ordoñez-Loza, J., & Ortiz, M., "Evaluation of the performance of a solar photovoltaic - Biomass gasifier system as an electricity supplier," Energy, vol. 260, p. 125046, 2022. DOI: 10.1016/j.energy.2022.125046.

[19] Dhass, A.D., Patel, D., & Patel, B., "Estimation of power losses in single-junction gallium-arsenide solar photovoltaic cells," International Journal of Thermofluids, vol. 17, p. 100303, 2023. DOI: 10.1016/j.ijft.2023.100303.

[20] Kulikov, A.L., Shepovalova, O.V., Ilyushin, P.V., Filippov, S.P., & Chirkov, S.V., "Control of electric power quality indicators in distribution networks comprising a high share of solar photovoltaic and wind power stations," Energy Reports, vol. 8, p. 1501–1514, 2022. DOI: 10.1016/j.egyr.2022.08.217.

[21] Gómez-Calvet, R., Martínez-Duart, J.M., & Gómez-Calvet, A.R., "The 2030 power sector transition in Spain: Too little storage for so many planned solar photovoltaics?" Renewable and Sustainable Energy Reviews, vol. 174, p. 113094, 2023. DOI: 10.1016/j.rser.2022.113094.

[22] Mohammed, C., Mohamed, M., Larbi, E.M., Manale, B., Hassan, Z., Jalal, B., & Smail, Z.,

"Extended method for the sizing, energy management, and techno-economic optimization of autonomous solar Photovoltaic/Battery systems: Experimental validation and analysis," Energy Convers Manag, vol. 270, p. 116267, 2022. DOI: 10.1016/j.enconman.2022.116267.

[23] Nain, P., & Kumar, A., "A state-of-art review on end-of-life solar photovoltaics," J Clean Prod, vol. 343, p. 130978, 2022. DOI: 10.1016/j.jclepro.2022.130978.

[24] Kinsey, G.S., Riedel-Lyngskær, N.C., Miguel, A.A., Boyd, M., Braga, M., Shou, Ch., Cordero, R.R., Duck, B.C., Fell, Ch. J., Feron, S., Georghiou, G.E., Habryl, N., John, J.J., Ketjoy, N., López, G., Louwen, A., Maweza, E.L., Minemoto, T., Mittal, A., Molto, C., Neves, G., Garrido, G.N., Norton, M., Paudyal, B.R., Pereira, E.B., Poissant, Y., Pratt, L., Shen, Q., Reindl, T., Rennhofer, M., Rodríguez-Gallegos, C.D., Rüther, R., van Sark, W., Sevillano-Bendezú, M.A., Seigneur, H., Tejero, J.A., Theristis, M., Töfflinger, J.A., Ulbrich, C., Vilela, W.A., Xia, X., & Yamasoe, M.A., "Impact of measured spectrum variation on solar photovoltaic efficiencies worldwide," Renew Energy, vol. 196, p. 995–1016, 2022. DOI: 10.1016/j.renene.2022.07.011.

[25] Liu, D., Qi, S., & Xu, T., "Visual observation or oral communication? The effect of social learning on solar photovoltaic adoption intention in rural China," Energy Res Soc Sci, vol. 97, p. 102950, 2023. DOI: 10.1016/j.erss.2023.102950.

[26] Chandel, R., Chandel, S.S., & Malik, P., "Perspective of new distributed grid connected rooftop solar photovoltaic power generation policy interventions in India," Energy Policy, vol. 168, p. 113122, 2022. DOI: 10.1016/j.enpol.2022.113122.

[27] Prasad, M., & Prasad, R., "Bifacial vs monofacial grid-connected solar photovoltaic for small islands: A case study of Fiji," Renew Energy, vol. 203, p. 686–702, 2023. DOI: 10.1016/j.renene.2022.12.068.

[28] Ndzibah, E., Pinilla-De La Cruz, G.A., & Shamsuzzoha, A., "Collaboration towards value creation for end-of-life solar photovoltaic panel in Ghana," J Clean Prod, vol. 333, p. 129969, 2022. DOI: 10.1016/j.jclepro.2021.129969.

[29] Boretti, A., & Castelletto, S., "Lacking energy storage, and nuclear contribution, wind, and solar photovoltaic electricity is expensive and scarce," The Electricity Journal, vol. 35, No. 10, p. 107222, 2022. DOI: 10.1016/j.tej.2022.107222

[30] Luan, R., & Lin, B., "Positive or negative? Study on the impact of government subsidy on the business performance of China's solar photovoltaic industry," Renew Energy, vol. 189, p. 1145–1153, 2022. DOI: 10.1016/j.renene.2022.03.082.

[31] Le, M., Luong, V.S., Nguyen, D.K., Dao, V.D., Vu, N.H., & Vu, H.H.T., "Remote anomaly detection and classification of solar photovoltaic modules based on deep neural network," Sustainable Energy Technologies and Assessments, vol. 48, p. 101545, 2021. DOI: 10.1016/j.seta.2021.101545.

[32] Oteng, D., Zuo, J., & Sharifi, E., "An expert-based evaluation on end-of-life solar photovoltaic management: An application of Fuzzy Delphi Technique," Sustainable Horizons, vol. 4, p. 100036, 2022. DOI: 10.1016/j.horiz.2022.100036.

[33] Singh, S.K., & Chander, N., "Mid-life degradation evaluation of polycrystalline Si solar photovoltaic modules in a 100 kW grid-tied system in east-central India," Renew Energy, vol. 199, p. 351–367, 2022. DOI: 10.1016/j.renene.2022.09.013.

[34] Ilyushin, P.V., Shepovalova, O.V., Filippov, S.P., & Nekrasov, A.A., "The effect of complex load on the reliable operation of solar photovoltaic and wind power stations integrated into energy systems and into off-grid energy areas," Energy Reports, vol. 8, p. 1515–1529, 2022. DOI: 10.1016/j.egyr.2022.08.218.

[35] Choi, J., Lee, I.W., & Cha, S.W., "Analysis of data errors in the solar photovoltaic monitoring system database: An overview of nationwide power plants in Korea," Renewable and Sustainable Energy Reviews, vol. 156, p. 112007, 2022. DOI: 10.1016/j.rser.2021.112007.

[36] Majji, R.K., Mishra, J.P., & Dongre, A.A., "Model predictive control based autonomous DC microgrid integrated with solar photovoltaic system and composite energy storage," Sustainable Energy Technologies and Assessments, vol. 54, p. 102862, 2022. DOI: 10.1016/j.seta.2022.102862.

[37] Sadanand, S., & Dwivedi, D.K., "Numerical modeling for earth-abundant highly efficient solar photovoltaic cell of non-toxic buffer layer," Opt Mater (Amst), vol. 109, p.

110409, 2020. DOI: 10.1016/j.optmat.2020.110409.

[38] Gholami, A., Ameri, M., Zandi, M., Ghoachani, R.G., Gerashi, S. J., Kazem, H.A., & Al-Waeli, A. H. A., "Impact of harsh weather conditions on solar photovoltaic cell temperature: Experimental analysis and thermal-optical modeling," Solar Energy, vol. 252, p. 176–194, 2023. DOI: 10.1016/j.solener.2023.01.039.

[39] Chindamani, M., & Ravichandran, C.S., "A hybrid DDAO-RBFNN strategy for fault tolerant operation in fifteen-level cascaded H-bridge (15L-CHB) inverter with solar photovoltaic (SPV) system," Solar Energy, vol. 244, p. 1–18, 2022. DOI: 10.1016/j.solener.2022.08.015.

[40] Kërçi, T., Tzounas, G., & Milano, F., "A dynamic behavioral model of the long-term development of solar photovoltaic generation driven by feed-in tariffs," Energy, vol. 256, p. 124506, 2022. DOI: 10.1016/j.energy.2022.124506.

[41] Kamil, K., Chong, K. H., & Hashim, H., "Excess power rerouting in the grid system during high penetration solar photovoltaic," Electric Power Systems Research, vol. 214, p. 108871, 2023. DOI: 10.1016/j.epsr.2022.108871.

[42] Gassar, A.A.A., & Cha, S.H., "Review of geographic information systems-based rooftop solar photovoltaic potential estimation approaches at urban scales," Appl Energy, vol. 291, p. 116817, 2021. DOI: 10.1016/j.apenergy.2021.116817.

[43] Mia, S., Kumer Podder, A., Manoj Kumar, N., Bhatt, A., & Kumar, K., "Experimental verification of a dynamic programming and IoT-based simultaneous load-sharing controller for residential homes powered with grid and onsite solar photovoltaic electricity," Sustainable Energy Technologies and Assessments, vol. 55, p. 102964, 2023. DOI: 10.1016/j.seta.2022.102964.

[44] Oteng, D., Zuo, J., & Sharifi, E., "A scientometric review of trends in solar photovoltaic waste management research," Solar Energy, vol. 224, p. 545–562, 2021. DOI: 10.1016/j.solener.2021.06.036.

[45] Premkumar, M., Jangir, P., Ramakrishnan, C., Kumar, C., Sowmya, R., Deb, S., & Kumar, N.M., "An enhanced Gradient-based Optimizer for parameter estimation of various solar photovoltaic models," Energy Reports, vol. 8, p. 15249–15285, 2022. DOI: 10.1016/j.egyr.2022.11.092.

[46] Lin, B., & Shi, L., "New understanding of power generation structure transformation, based on a machine learning predictive model," Sustain. Energy Technol. Assessments, vol. 51, p. 101962, 2022. DOI: 10.1016/j.seta.2022.101962

[47] Zhang, Y., Liu, K., Qin, L., & An, X., "Deterministic and probabilistic interval prediction for short-term wind power generation based on variational mode decomposition and machine learning methods," Energy Convers. Manag., vol. 112, p. 208–219, 2016. DOI: 10.1016/j.enconman.2016.01.023

[48] Sharma, H., Marinovici, L., Adetola, V., & Schaef, H.T., "Data-driven modeling of power generation for a coal power plant under cycling," Energy AI, vol. 11, p. 100214, 2023. DOI: 10.1016/j.egyai.2022.100214

[49] Chen, H., Birkelund, Y., & Yuan, F., "Examination of turbulence impacts on ultra-short-term wind power and speed forecasts with machine learning," Energy Reports, vol. 7, p. 332–338, 2021. DOI: 10.1016/j.egyr.2021.08.040

[50] Chen, Z., Xiao, F., Guo, F., & Yan, J., "Interpretable machine learning for building energy management: A state-of-the-art review," Adv. Appl. Energy, vol. 9, p. 100123, 2023. DOI: 10.1016/j.adapen.2023.100123

[51] Li, L., Zhou, Z., Bai, N., Wang, T., Xue, K.H., Sun, H., He, Q., Cheng, W., & Miao, X., "Naive Bayes classifier based on memristor nonlinear conductance," Microelectronics J., vol. 129, p. 105574, 2022. DOI: 10.1016/j.mejo.2022.105574

[52] Monteiro, R.V.A., Guimarães, G.C., Moura, F.A.M., Albertini, M.R.M.C., & Albertini, M.K., "Estimating photovoltaic power generation: Performance analysis of artificial neural networks, Support Vector Machine and Kalman filter," Electr. Power Syst. Res., vol. 143, p. 643–656, 2017. DOI: 10.1016/j.epsr.2016.10.050

[53] Roy, A., & Chakraborty, S., "Support vector machine in structural reliability analysis: A review," Reliab. Eng. Syst. Saf., vol. 233, p. 109126, 2023. DOI: 10.1016/j.ress.2023.109126

[54] Maresch, K., Marchesan, G., & Cardoso, G., "A logistic regression approach for improved safety of the under-frequency load shedding

scheme owing to feeder machine inertia," Electr. Power Syst. Res., vol. 218, p. 109189, 2023. DOI: 10.1016/j.epsr.2023.109189

[55] Asvapoositkul, S., & Preece, R., "Decision tree-based prediction model for small signal stability and generation-rescheduling preventive control," Electr. Power Syst. Res., vol. 196, p. 107200, 2021. DOI: 10.1016/j.epsr.2021.107200

[56] Liu, D., & Sun, K., "Random Forest solar power forecast based on classification optimization," Energy, vol. 187, p. 115940, 2019. DOI: 10.1016/j.energy.2019.115940

[57] Amonkar, Y., Farnham, D.J., & Lall, U., "A k-nearest neighbor space-time simulator with applications to large-scale wind and solar power modeling," Patterns, vol. 3, No. 3, p. 100454, 2022. DOI: 10.1016/j.patter.2022.100454

[58] Abushgair, K., "Enhancement of Poly-Crystal PV Panels Performance by Air-to-Air Heat Exchanger Cooling System", WSEAS Transactions on Power Systems, vol. 16, pp. 157-163, 2021. DOI: 10.37394/232016.2021.16.16

[59] AlKandari, M., & Ahmad, I., "Solar power generation forecasting using ensemble approach based on deep learning and statistical methods", Applied Computing and Informatics, vol. ahead of print, pp. 1-20, 2020. DOI: 10.1016/j.aci.2019.11.002

[60] Borunda, M., Ramírez, A., Garduno, R., Ruíz, G., Hernandez, S., Jaramillo, O.A., "Photovoltaic Power Generation Forecasting for Regional Assessment Using Machine Learning", Energies, vol. 15, No. 8895, pp. 1-25, 2022. DOI:10.3390/en15238895

[61] Kim, E., Akhtar, M. S., & Yang, O. B., "Designing solar power generation output forecasting methods using time series algorithms", Electric Power Systems Research, vol. 216, pp. 1-12, 2023. DOI: 10.1016/j.epsr.2022.109073.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

-Victor Daniel Gil Vera has performed the literature review, the normalization of the database, training the predictive models in Python, and the statistical analysis.

-Catalina Quintero López has performed the literature review and the analysis of the results.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

This research was funded by the Luis Amigó Catholic University and was one of the results of the research project entitled "Implementation of Smart Grids in Colombia: a multidimensional analysis" - Cost Center [0502020950].

Conflict of Interest

The authors have no conflict of interest to declare.