International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 9 Sep 2023, pp: 122-125 www.ijaem.net ISSN: 2395-5252

Forecasting Solar Power Generation: An Evaluation of Regression Algorithms

Arsh Dayal

Date of Submission: 05-09-2023 Date of Acceptance: 15-09-2023

ABSTRACT - This research will explore machine learning and statistical techniques to effectively predict solar power generation by leveraging historical weather data and National Weather Service forecasts. Solar energy generation is facilitated by photovoltaic systems, which are employed globally. The variability and dependence on prevailing conditions are inherent characteristics of solar power, leading to the unpredictable nature of photovoltaic system output.
Humidity, Radiation, wind, and PV surface temperature are among the factors that influence solar power generation. Given the unpredictable nature of photovoltaic power generation, it is crucial to make advanced plans for solar power and rely on solar forecasting to meet grid requirements. Solar power generation is heavily influenced by weather conditions, which makes accurate forecasting a major challenge. This research explores the effects of different environmental factors on photovoltaic power generation. The prediction of solar energy output relies on diverse methodologies such as neural deep learning, network approaches, and machine learning techniques.

Keywords– regression, solar power, error, mean, data, forecasting, Machine learning.

I. INTRODUCTION

The energy output of a photovoltaic produced plays a significant role in determining the size of a photovoltaic system. This power output is influenced by various factors, including the time of day, the prevailing weather conditions, and the system's location. Machine learning methods have been utilized to effectively gauge the power output of solar panels by considering elements like time, location, and weather conditions to tackle this problem.

During the summer season, the panels absorb a significantly larger amount of energy from the sun. Weather conditions play a crucial role in determining power generation, prompting the consideration of weather forecasting. Consequently, the quantity of electricity produced is impacted by the amount of solar energy received during a specific day, which is determined by geographical position, time of day, and prevailing weather conditions.

Our main objective is to develop automated and precise models that use data from the National Weather Service to predict the amount of renewable energy generated. Regression methods play a crucial role in the prediction of solar power generation.

II. LITERATURE REVIEW

The most basic approach, the smart persistence model, analyzes previously recorded data to anticipate the amount of power generated within a short period of time, usually 2-3 hours. To begin with, an NWP is developed for a certain time period and area. To address this issue, machine learning techniques are utilized to explore advanced solar radiation prediction models. We employ day-ahead radiation from solar data predictions in these studies to demonstrate that a machine-learning technique can properly predict short-term solar energy [2]. Ahybrid method is built by integrating classification, clustering, and regression methodologies to produce a forecasting model. This model is condition-based and selects the closest weather condition (based on the weather forecast for the next day) to predict power output using clustering regression [3]. Because electricity consumption is proportional to other energy sources, such as oil and natural gas, we employ a variety of data mining approaches, such as compiling previous load data while evaluating the properties of the load time series.
Artificial intelligence technologies are used to investigate the relationship between projected weather conditions and the amount of power generated as a historical time series. Instead of formal statistical analysis, AI techniques employ algorithms that may effectively characterize the nonlinear and extremely detailed relationship between power output and the data used. In the supervised learning approach, artificial neural networks (ANNs) learn from data by being trained to estimate and approximate the underlying relationship [6]. These models have improved the ability to forecast solar power firms' power generations [4-7]. Long Short-Term Memory

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 9 Sep 2023, pp: 122-125 www.ijaem.net ISSN: 2395-5252

(LSTM) network, a recurrent neural network (RNN),can learn a function by taking a sequence of past solar irradiance values as input and generating a solar irradiance value as output. Similarly, DBNs, which are deep neural networks, can also learn a function by taking a sequence of historical solar irradiance values as input and producing a solar irradiance value as output. This approach allows the network to capture and analyze the patterns and dependencies present in the historical solar irradiance data, enabling accurate predictions of future solar irradiance values.

III. PROPOSED WORK

From 2019 to 2020, we collected and examined multiple weather factors to establish the correlation between mean meteorological data and solar irradiance to enhance accuracy while predicting the power generated. The system architecture of the suggested work begins by taking into account the dataset, preprocessing the data, dividing it into test and training information, using algorithmic methods for classification, and predicting the outcomes. The data gathered is typically raw and has missing and erroneous data, and may be in a format inappropriate for direct usage in Machine Learning.

The model is trained using a fraction of data collected as training data, and its predictions are validated using the testing data. The first set of data is used to create a model for prediction, while the second set is used to evaluate the model's effectiveness. Following a normal procedure, the dataset is split into training and testing sets in a 3:1 ratio, with 75% of the data being put aside for training and 25% for testing and validation. Before training the model to forecast solar power generation, the dataset undergoes a data preprocessing step, which involves examining it for the outliers and null values. Subsequently, the model is trained using a three-hour dataset.

IV. METHODOLOGY

The dataset consists of hourly weather parameter values averaged in three hours to compute daily mean values. Multiple weather features were collected to investigate the relationship between meteorological data and mean solar irradiance to estimate mean solar irradiance properly. The data sets gathered include the average daily ambient temperature readings, wind speed, humidity, wind direction, and average pressure. This paper suggests using Regression techniques, namely Random Forest and Linear Regression, and Support Vector Machine as the Machine Learning (ML) models used to predict the amount of solar energy produced after analyzing the data.

1. Forecasting models

The variable we use as a prediction element is valued continuously; we estimate that the unknown test model's output will be the mean of K or the nearest value. Due to the non-linear nature of the dataset, linear models were not used. The undetermined parameters of the model were estimated from the data set using the least possible squares technique [10]. The function and selected kernel, among other variables, affect the efficiency of the Support Vector Regression Model. Root-mean-square deviation and R-squared values were calculated to evaluate the models' test dataset's performance. We fine-tuned the hyper parameters of our models before choosing the models with the lowest RMSE and greatest R squared values. Linear regression is a widely used statistical method to analyze the relationship between a dependent variable and one or more independent variables. By fitting a best-fit linear curve, it seeks to determine the optimal values for the curve and intercept that align with the observed data points. After thorough analysis, we found that random forests, support vector regression and linear regression yielded the best results when applied to our dataset. To assess the accuracy of our models, we tested those using independent data.

V. RESULTS

Table 1: Prediction of Models

Technique	RMSE	MAE	MSE	Accuracy
Support vector machine regressor	132.4	76.96	173.06	-43.92
Linear regression	59.27	47.99	343.08	73.4
Random forest regressor	28.02	12.55	746.48	93.81

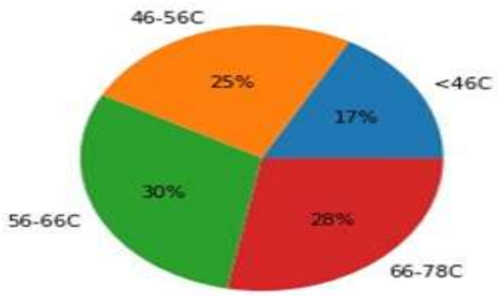

Fig 1. The temperature difference and the proportion of power generated by each.

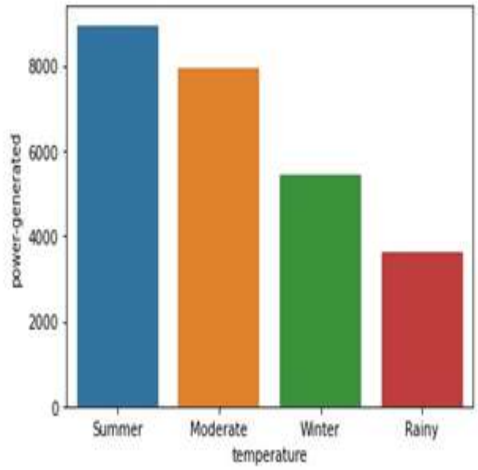

Fig 2. Several Weather condition analyses.

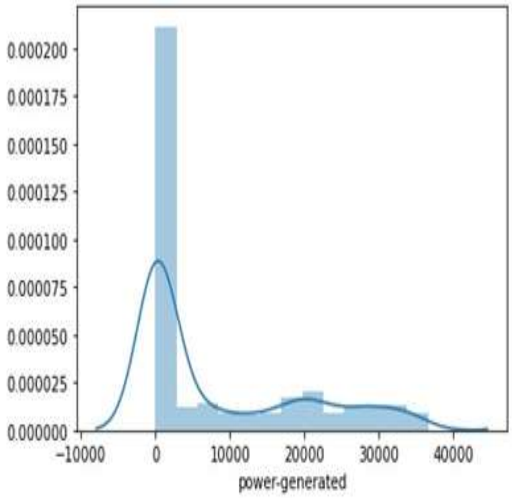

Fig. 3General season Analysis

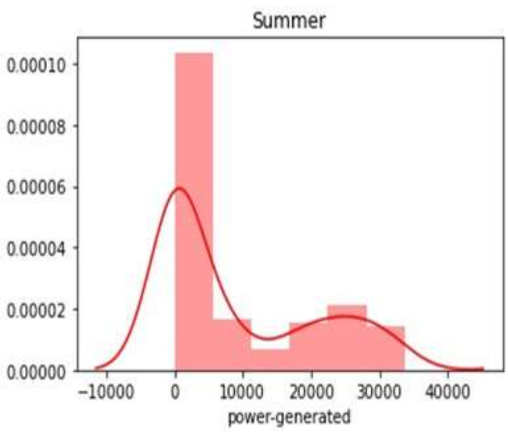

Fig. 4 Summer season Analysis

VI. CONCLUSION

In this research, we have introduced machine learning techniques to analyze solar power generation. Our method accurately predicts the power generated in different states of India by utilizing environmental data. Machine learning techniques in this regression model are used to evaluate the appropriate regression model. However, random forest regression, linear regression, and support vector machine regression models were implemented in this study, and the random forest regression model outperformed the other two models after evaluating their effectiveness and accuracy.

Our analysis observed that power generation increases by 35% when the temperature ranges from 57-56°F, compared to other temperatures. Temperatures below 46°F contribute to an average of 15% of the overall power generation. Based on these findings, it is evident that the Random Forest Regressor model exhibits superior performance with an accuracy of 91.01%, indicating its effectiveness in predicting solar power generation.

REFERENCE:

[1]. P. A. G. M. Amarasinghe and S. K. Abeygunawardane, "Application of Machine Learning Algorithms for Solar Power Forecasting in Sri Lanka" (2nd International Conference On Electrical Engineering (EECon), Colombo, Sri Lanka, 87 2018).

[2]. M. Z. Hassan, M. E. K. Ali, A. B. M. S. Ali and J. Kumar, "Forecasting Day-Ahead Solar Radiation

[3]. Using Machine Learning Approach" (4th AsiaPacific World Congress on Computer Science and Engineering (APWC on CSE), Mana Island, Fiji, 252 2017).

[4]. A. Bajpai and M. Duchon, "A Hybrid Approach of Solar Power Forecasting Using Machine Learning" (3rd International Conference on Smart Grid and Smart Cities (ICSGSC), 108 2019).

[5]. A. Khan, R. Bhatnagar, V. Masrani and V. B. Lobo, "A Comparative Study on Solar Power Forecasting using Ensemble Learning," (4th International Conference on Trends in Electronics and Informatics (ICOEI), 224 2020).

[6]. Khan, P.W.; Byun, Y.-C.; Lee, S.-J.; Kang, D.-H.; Kang, J.-Y.; Park, H.-S. Energies, **13**, 4870 (2020).

[7]. Faquir, Sanaa & Yahyaouy, Ali & Tairi, H. & Sabor, Jalal. International Journal of Fuzzy System Applications. **4**, 10 (2015).

[8]. Aler R., Martín R., Valls J.M., Galván I.M. IntelligentDistributed Computing VIII.

Studies in Computational Intelligence, vol 570 (2015).

[9]. Y. Wang, G. Cao, S. Mao and R. M. Nelms, "Analysis of solar generation and weather data in smart grid with simultaneous inference of nonlinear time series," (IEEE Conference on Computer Communications Workshops (INFOCOMWKSHPS), 600 2015).

[10]. Carrera B, Kim K. Sensors (Basel). **20**, 3129 (2020).

[11]. Jawaid F, NazirJunejo K. Predicting daily mean solar power using machine learning regression techniques. (Sixth International Conference on Innovative Computing Technology (INTECH) 355 2016).

[12]. Batcha RR, Geetha MK. A survey on IOT based onrenewable energy for efficient energy conservation using machine learning approaches. (3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE) 123 2020).