Predicting Solar Energy Generation using Time Series and Machine Learning Models

by

Ayse B Sengul

Solar energy, a renewable and abundant energy source, is a promising alternative to fossil fuel-generated energy due to its capacity to reduce greenhouse gas emissions. Incorporating solar energy into energy production processes is estimated to reduce approximately 32% of carbon dioxide (CO_2) emissions [1]. Photovoltaic (PV) cells or panels, key contributors to renewable energy generation, convert sunlight into electricity through the photovoltaic effect [2]. This electricity can be used for a variety of applications, ranging from powering individual homes and businesses to supplying electricity to entire communities [3]. A notable advantage of PV cells is their ability to generate electricity without emitting greenhouse gases or other pollutants. Unlike fossil fuel-based energy generation methods, which release harmful emissions such as CO_2, sulfur dioxide (SO_2), and nitrogen oxides (NO_x), PV cells harness sunlight directly, making solar energy a clean and environmentally friendly energy source. By reducing air pollution and combating climate change, solar energy plays a crucial role in improving air quality and safeguarding human health and the environment.

The construction and operation of solar energy generation systems depend on several environmental factors, including solar irradiance, temperature, relative humidity, and wind speed. These factors can vary significantly over time, leading to fluctuations in the power output of PV systems. This variability poses challenges for accurately predicting power production and can impact various aspects of the electric grid, including reliability, stability, planning, scheduling tasks, and market operations [4-8]. Consequently, forecasting power generation from PV systems has become a critical area of research interest. Developing accurate forecasting models is essential to address the challenges posed by the variability of environmental factors and ensure the reliable and efficient operation of the electric grid. Additionally, even a small improvement in prediction accuracy can yield significant cost savings. For instance, a 25% improvement in accuracy could result in a 1.56% reduction in generation cost, translating to approximately US$ 46.5 million [9]. This highlights the substantial financial impact of enhancing the accuracy of solar power generation predictions

Various methods have been documented in the literature for forecasting PV energy, which can be categorized into four classes: (i) physical, (ii) empirical, (iii) statistical and (iv) machine learning models. Physical models are based on numerical weather prediction and satellite imagery. Empirical models develop linear or nonlinear regression equations. Statistical models, like autoregressive moving average (ARMA), the autoregressive integrated moving average (ARIMA), are developed based on statistical correlations. Machine learning models, like artificial neural networks (ANNs) and support vector machines (SVM), based on machine learning approaches [6, 10, 11].

The primary objective of this project is to predict solar energy generation with high accuracy using advanced analytical methods. Specific objectives include:

1. Analyzing historical data on solar energy production, weather conditions, and other relevant factors.
2. Developing forecasting models based on time series analysis techniques such as ARIMA and SARIMA to capture temporal patterns.
3. Incorporating machine learning algorithms such as random forests, and gradient boosting to capture complex relationships in the data.
4. Evaluating the performance of different prediction models using metrics such as root mean squared error and mean absolute percentage error and selecting the most accurate and reliable approach.

The dataset that is used in this study taken from Solar Power plant Dataset. Solar power generation data for one plant gathered at 15 minutes intervals over a 34-day period.

References

1. Pelaez-Samaniego, M.R., et al., *Biomass carbonization technologies.* Sustainable Biochar for Water and Wastewater Treatment, 2022: p. 39-92.
2. Kalogirou, S.A., *Introduction to renewable energy powered desalination*, in *Renewable energy powered desalination handbook*. 2018, Elsevier. p. 3-46.
3. Ledmaoui, Y., et al., *Forecasting solar energy production: A comparative study of machine learning algorithms.* Energy Reports, 2023. **10**: p. 1004-1012.
4. Gutiérrez, L., J. Patiño, and E. Duque-Grisales, *A Comparison of the Performance of Supervised Learning Algorithms for Solar Power Prediction.* Energies, 2021. **14**(15).
5. Carrera, B. and K. Kim, *Comparison Analysis of Machine Learning Techniques for Photovoltaic Prediction Using Weather Sensor Data.* Sensors (Basel), 2020. **20**(11).
6. Mishra, S., et al. *An efficient machine learning approach for accurate short term solar power prediction*. in *2020 International Conference on Computational Intelligence for Smart Power System and Sustainable Energy (CISPSSE)*. 2020. IEEE.
7. Shawon, M.M.H., et al. *Forecasting PV panel output using Prophet time series machine learning model*. in *2020 IEEE region 10 conference (TENCON)*. 2020. IEEE.
8. Abubakar, M., et al., *Performance Analysis of Energy Production of Large-Scale Solar Plants Based on Artificial Intelligence (Machine Learning) Technique.* Processes, 2022. **10**(9).
9. Lee, C.-H., H.-C. Yang, and G.-B. Ye, *Predicting the Performance of Solar Power Generation Using Deep Learning Methods.* Applied Sciences, 2021. **11**(15).
10. Fara, L., et al., *Forecasting of Energy Production for Photovoltaic Systems Based on ARIMA and ANN Advanced Models.* International Journal of Photoenergy, 2021. **2021**: p. 1-19.
11. Zhou, Y., et al., *A review on global solar radiation prediction with machine learning models in a comprehensive perspective.* Energy Conversion and Management, 2021. **235**.