

Erdos Data Science Bootcamp Summer 2024 Executive Summary

Team: Md Mehedi Hasan, Kamlesh Sarkar

Github: <https://github.com/mmh52/Forecast-Solar-Energy>

Overview: Solar energy, a renewable source of power, plays an important role in reducing greenhouse gases, mitigating climate change, and protecting ecosystems. Nowadays, the adoption of solar energy into the power grid has increased, and Direct Normal Irradiance (DNI) is particularly important in forecasting the performance of concentrating solar power (CSP) systems. Photovoltaic panels track the sun to receive more DNI, which accounts for a large portion of solar energy from PV.

Objective: Global horizontal irradiance is the total solar radiation per unit area measured at a horizontal surface on the earth and can be broken down into two components: direct normal irradiance (DNI) and diffuse horizontal irradiance (DHI). Most of the solar radiation comes from DNI. So, forecasting DNI is crucial for the effective operation and maintenance of solar power systems, ensuring their ability to harness solar energy effectively. The objective of our project is to select an effective time series model that can forecast Direct Normal Irradiance from solar power.

Stakeholders: Accurate DNI predictions help Operation and Maintenance (O&M) Service Providers in optimizing solar panel orientation for CSP systems and schedule maintenance during low DNI periods to minimize disruption. Also, forecasting DNI aids grid operators in managing solar power integration, balancing supply and demand, maintaining grid stability, and reducing fossil fuel reliance. These predictions are also crucial for assessing the financial viability of solar projects, enabling investors to estimate returns and make informed funding decisions.

Key Performance Indicator: We used RMSE (Root Mean Square Error) metrics to measure the performance of all models based on short-term and long-term forecasting of DNI.

Modeling Approach: We use a dataset from Lowery Power Station, Denver, Colorado, covering the period from June 2008 to December 2013. This dataset contains Direct Normal irradiance, Zenith angle, and other important features of Solar Energy

- This dataset contains minute data of direct normal irradiance and has some anomalies. We conducted a thorough data cleaning process, fixing anomalies, filling in missing values, and converting it to an hourly dataset.
- Check the seasonality and trend of the data. Plot autocorrelation and partial autocorrelation to find the seasonal pattern of direct normal irradiance.
- We take DNI's data from the last week of December 2013 as a test set and the remaining data as a train set.
- To forecast a day and a week ahead DNI, we choose some time series forecasting models such as Naïve Seasonal, Average Seasonal, Exponential smoothing, ARIMA, SARIMA, and SARIMAX

Table 1: Model Accuracy Comparison

Model	RMSE of a day ahead forecast	RMSE of a week ahead forecast
Naive	147.5306	360.6987
Seasonal Average	197.7084	193.3614
Exponential Smoothing	144.4954	202.2607
ARIMA	170.4819	167.7565
SARIMA	156.6338	126.3694
SARIMAX	154.1319	181.8193

Results: Among all the time series forecasting models we applied in our project, Exponential Smoothing, yield better performance for a day ahead DNI forecasting. For a week ahead SARIMA showed the best performance compared to its exogenous counterparts and all other models.

Future Work:

- Incorporate weather and cloud cover data, which can be important features for forecasting direct normal irradiance.
- Use multivariate deep learning models that can incorporate all important features of solar energy and forecast direct normal irradiance more accurately.