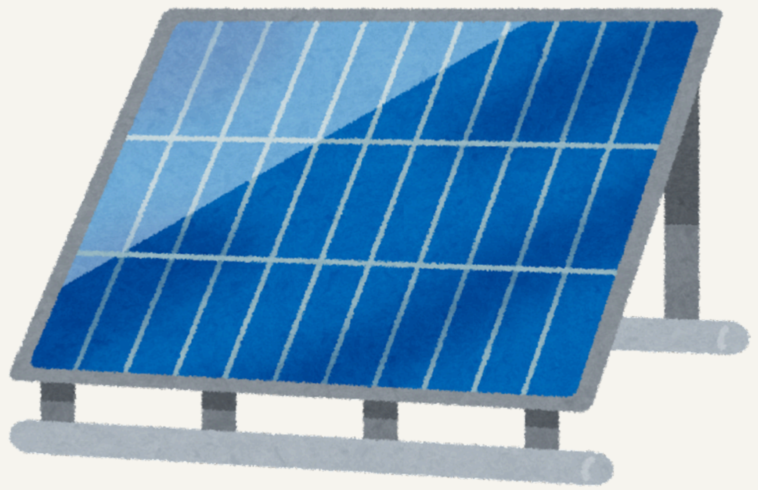
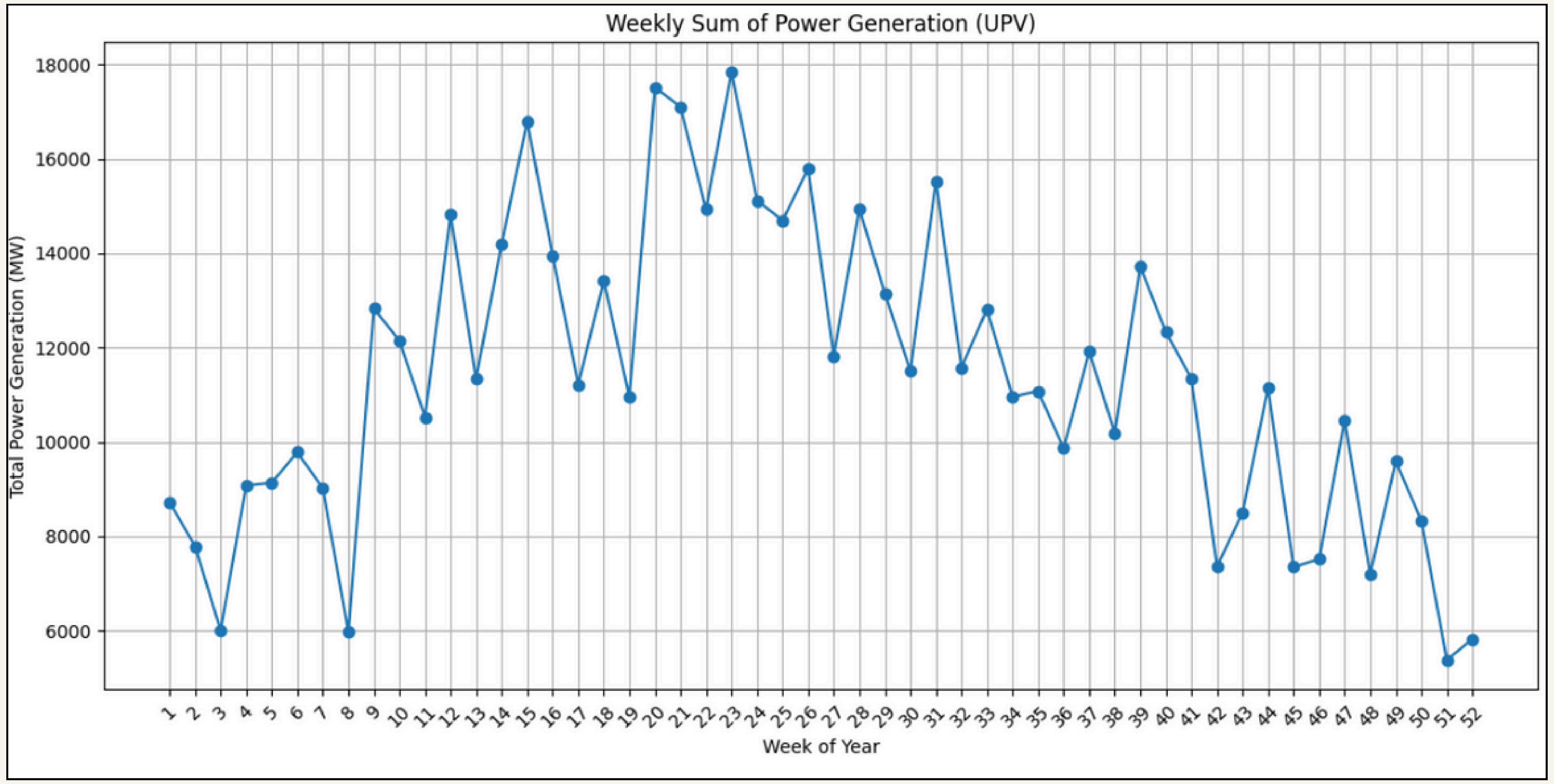
Solar Power Generation Prediction

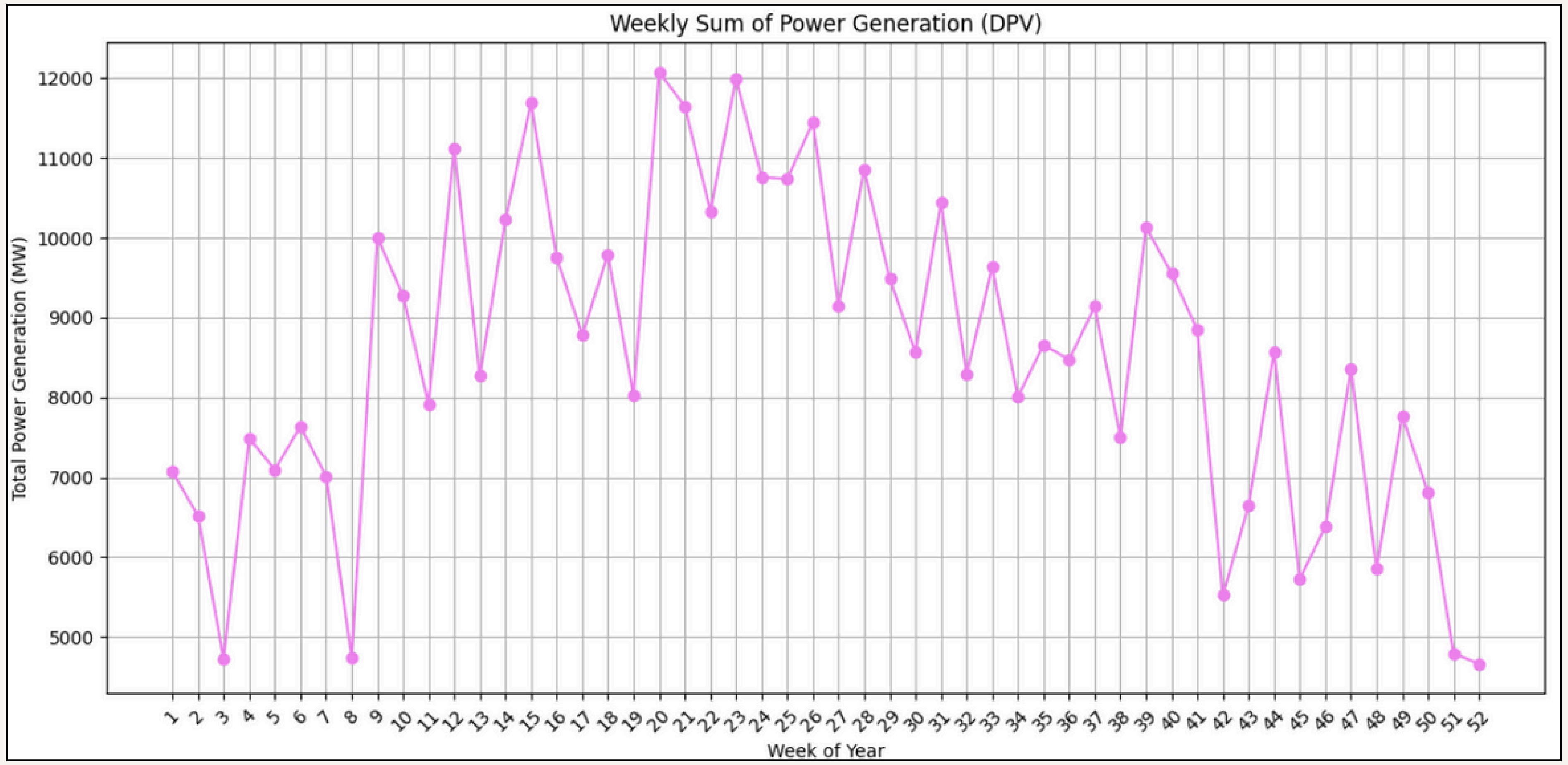


Prateek Majumder
PGP in Artificial Intelligence & Data Science
Jio Institute 2024-2025

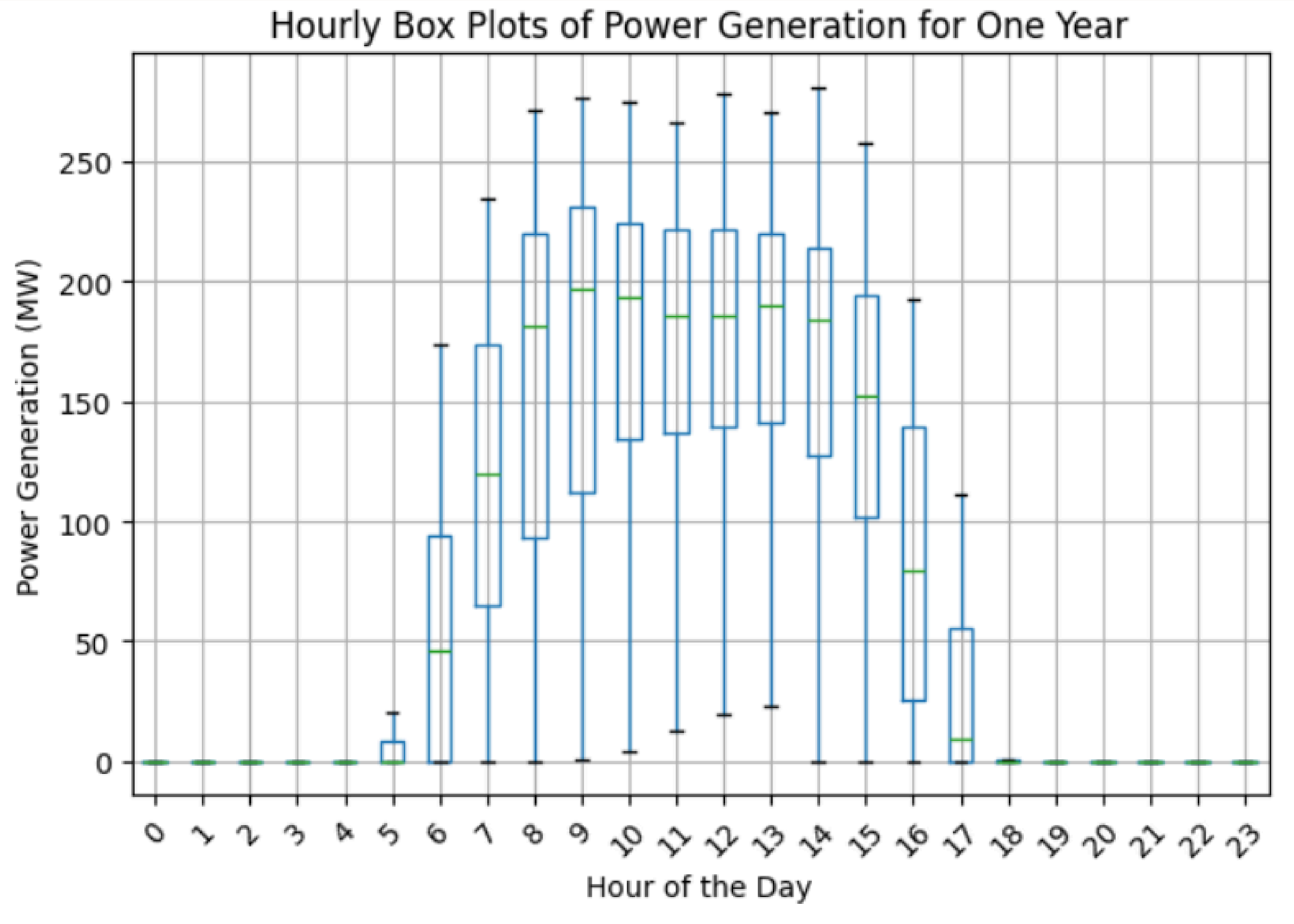
Data

Seasonal Trends	Both UPV and DPV datasets show higher power generation in summer due to longer daylight hours and increased solar irradiance, and lower generation in winter.
Diurnal Patterns	Both datasets exhibit peak power output around midday and minimal generation at night, reflecting the sun's position and sunlight availability.
Correlated Variations	Despite minor differences in power generation magnitudes, UPV and DPV trends are highly correlated, indicating similar influences from solar irradiance and weather conditions





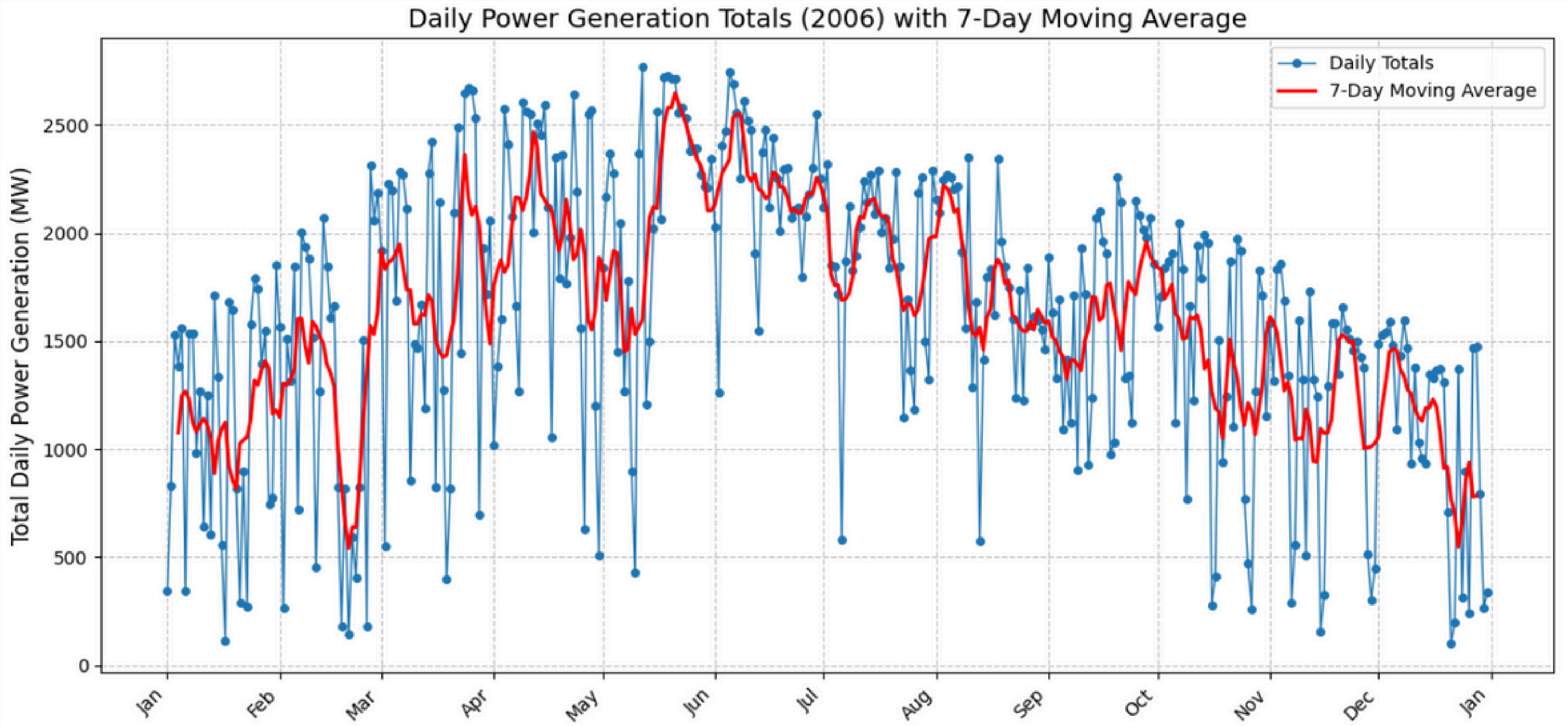
Exploratory Data Analysis (EDA):



Power generation is highest during the middle of the day.

Generation is lowest in the early morning and late evening.

There's a lot of fluctuation in power output, especially during the peak hours.



The graph shows a clear seasonal pattern in power generation. We can see a significant increase in power generation during the summer months and a decrease during the winter months.

This is likely due to the increased solar radiation and longer daylight hours during summer.

Even within the summer months, there are fluctuations in daily power generation. This could be attributed to factors such as cloud cover, weather conditions, and variations in solar irradiance.

Additional Data and Features

	LocalTime	Power(MW)	Hour	temp	rhum	SolarIrradiance	Month
0	2006-01-01 00:00:00	0.0	0	11.0	62.0	0.0	1
1	2006-01-01 01:00:00	0.0	1	8.0	81.0	0.0	1
2	2006-01-01 02:00:00	0.0	2	7.0	87.0	0.0	1
3	2006-01-01 03:00:00	0.0	3	6.0	87.0	0.0	1
4	2006-01-01 04:00:00	0.0	4	5.0	93.0	0.0	1

Given the Latitude and Longitude, fetched the Data for Temperature, Relative Humidity and Solar Irradiance to check how other climatic factors influence the Solar Power Production.

As we have seen clear relation between Month of year & Hour of day to Power production, both of them have been added as Features to the final data, which will be used for Regression.

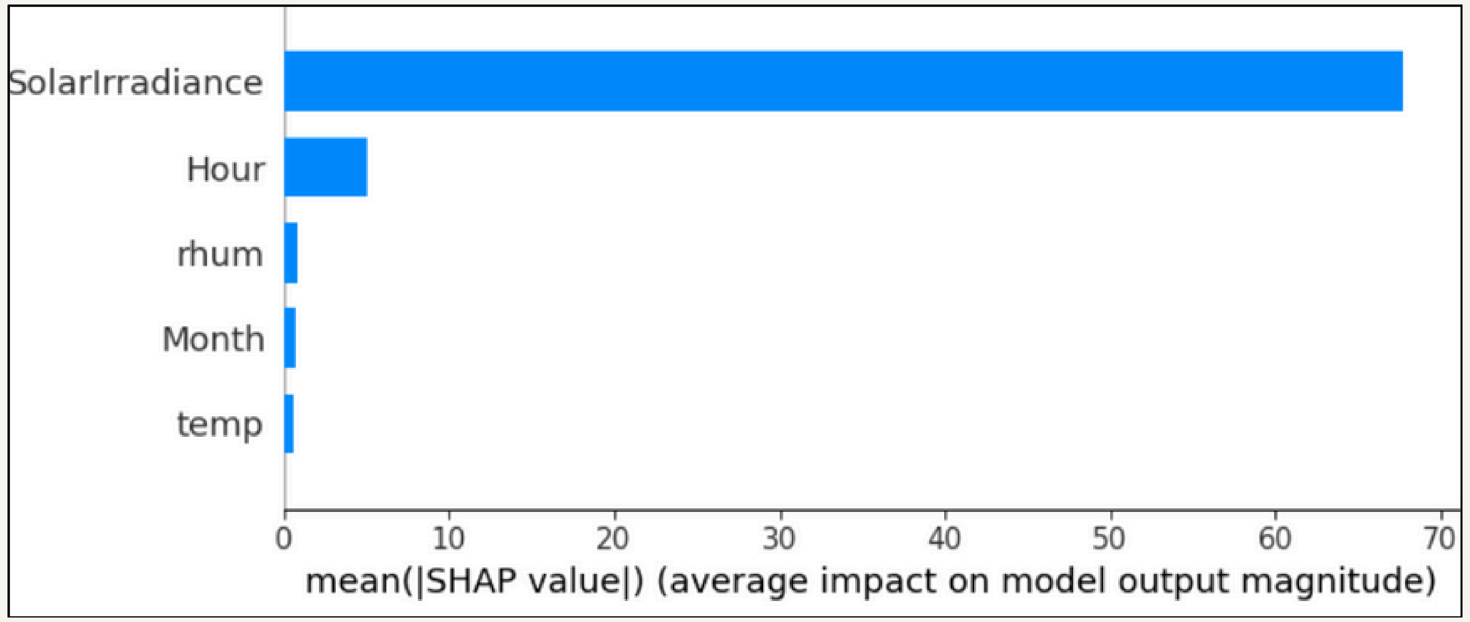
Random Forest Regression

	Testing	Cross-Validation
(MAE)	11.33	13.28
(MSE)	464.32	610.17
(RMSE)	21.54	24.20
R_2	0.94	0.92

A Random Forest Regressor model is trained using a set of numerical features (temperature, humidity, solar irradiance, hour, month, etc.) to predict power generation ('Power(MW)').

Model performed well in normal testing, as well as in Cross Validation, so we can say that, with the addition of features, the model has become robust

	feature	importance
3	SolarIrradiance	0.975498
0	Hour	0.019174
2	rhum	0.001952
4	Month	0.001880
1	temp	0.001495



Neural Network

A sequential neural network with three dense layers is created. The first two layers have ReLU activation functions, which introduce non-linearity. The final layer is a linear output layer (no activation function) as it's predicting a continuous value (power generation).

After training, the model is evaluated on the scaled test data. The model's predictions are then inverse-transformed to the original scale of the target variable (power generation), and traditional regression metrics are calculated evaluate the model.

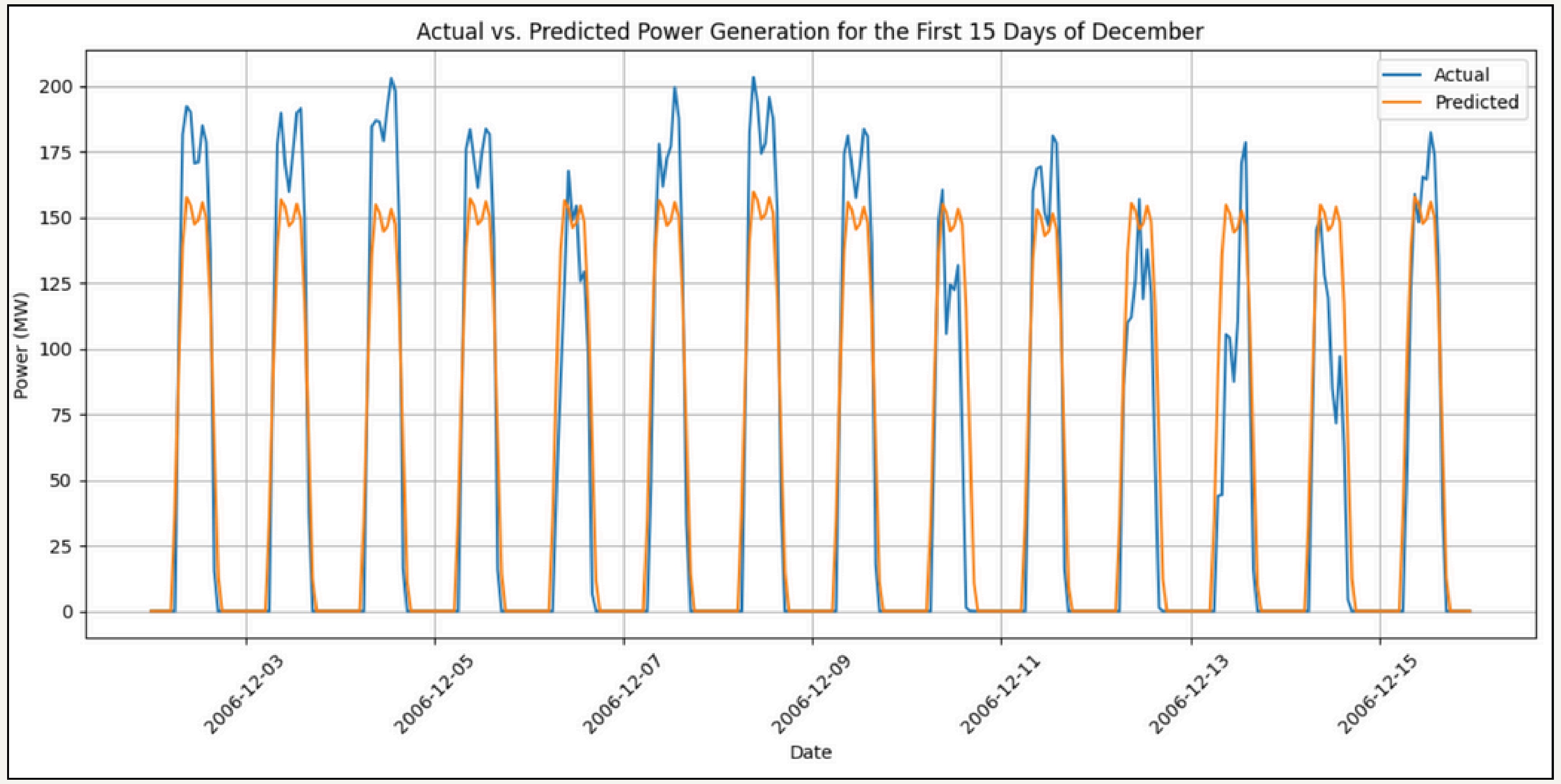
(MAE)	12.08
(MSE)	435.88
(RMSE)	20.87
R_2	0.94

Time Series (Prophet)

Facebook Prophet is utilized to predict hourly power generation by modeling strong seasonality and trend components in time series data.

The data is split into training (11 months) and testing (1 month) sets to evaluate the model's generalization. After training, the model predicts future values, ensuring non-negative outputs. Predictions are assessed using the R-squared metric and visualized against actual data for insights into accuracy and areas for improvement.

(MAE)	20.72
(MSE)	1476.2
(RMSE)	38.42
R_2	0.67

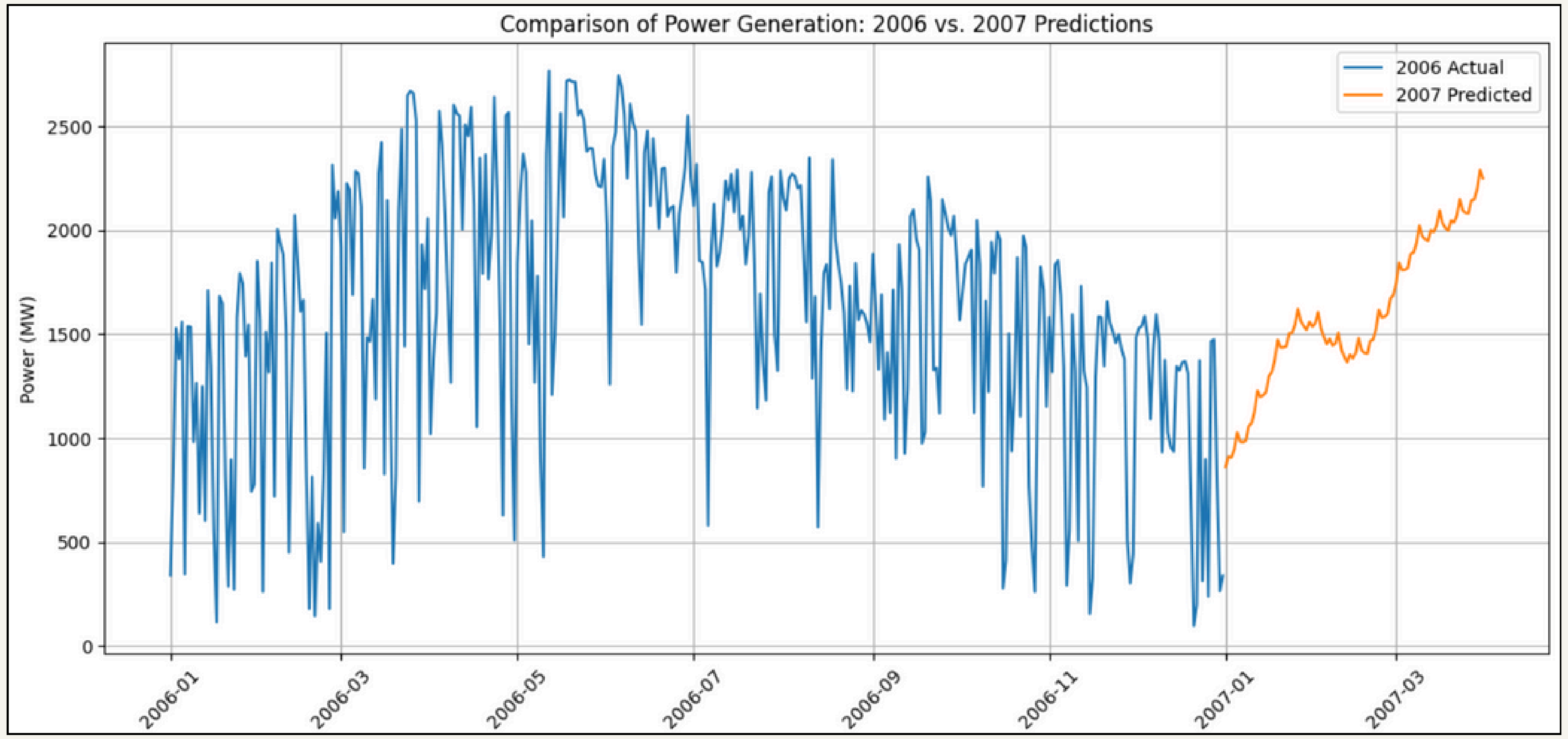


Future Predictions

A Prophet model is trained on the daily power generation data from the full year. The model learns the seasonal patterns and trends in the data.

The trained model is then used to make predictions for the period from January 1, 2007 to March 31, 2007.

The future predictions of Jan to March 2007 is plotted and compared against the actual data from 2006, and it shows significant similarity.



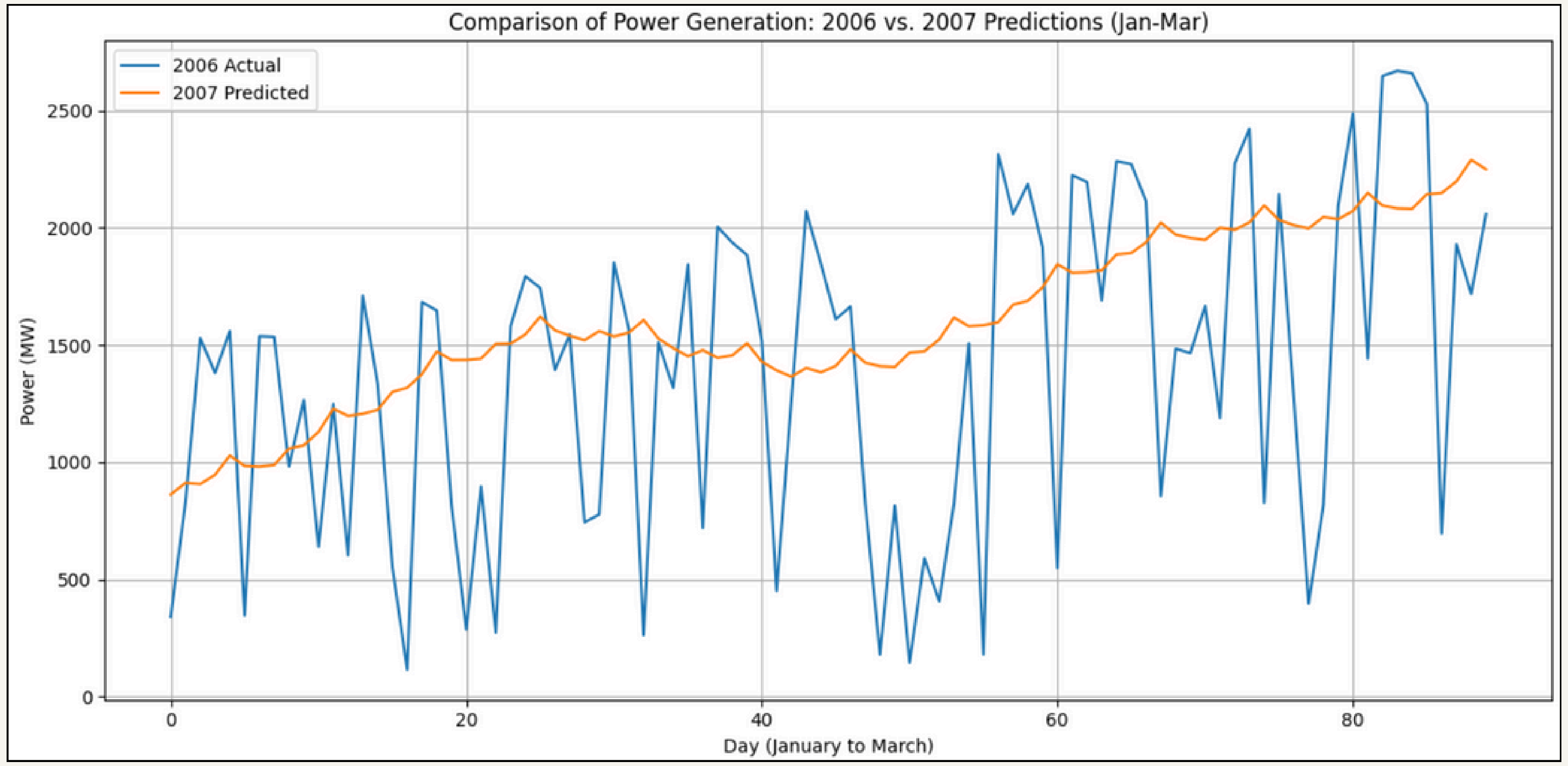
Recommendations:

Consider optimizing solar panel orientation and tilt angles to maximize solar irradiance capture throughout the day and seasons.

Implement robust energy storage systems, such as batteries, to store excess energy during peak hours and supplement power generation during low solar irradiance, while evaluating advanced storage technologies like pumped hydro for long-term needs.

Use the trained model to forecast low solar power periods for predictive maintenance, adjusting schedules proactively and integrating weather forecasts for better planning.

Periodically retrain models like RandomForest, Neural Network, and Prophet with updated data to enhance accuracy and adapt to changing environmental conditions and seasonal variations.



Thank You!