© May 2025 | IJIRT | Volume 11 Issue 12 | ISSN: 2349-6002

Renewable Energy Forecasting

[1.] Rahul Singh [2.]Rajeev Nair [3.] Nair Shivbalakrishna
College: Ajeenkya DY Patil University, Pune, Maharashtra, India

Abstract: **As the global energy sector pivots toward sustainability, driven by growing concerns over climate change and carbon emissions, renewable energy has emerged as a cornerstone of future energy policy. Yet, the integration of solar, wind, and hydro power into traditional electricity grids introduces significant challenges due to their dependency on fluctuating meteorological conditions. Ensuring consistent grid stability while accommodating these intermittent sources necessitates accurate, data-driven forecasting systems.**

This research presents a comprehensive dual-model forecasting framework aimed at predicting both renewable energy production and daily electricity load demand in the state of Maharashtra, India. The methodology leverages the Random Forest Regressor—an ensemble learning algorithm capable of managing non-linear relationships and high-dimensional datasets. The first model forecasts total renewable energy generation from wind, solar, and hydro sources. The second model predicts daily electric load based on historical weather patterns and consumption data.

Extensive data preprocessing, intelligent feature engineering, and visual exploration techniques were applied to develop high-performing models. The renewable energy forecasting model yielded an R^2 score of 0.94, indicating strong correlation between predicted and actual values, while the electric load model achieved an R^2 of 0.82. Further, an energy balance analysis highlighted approximately 210 surplus days and 155 deficit days during 2024, offering actionable intelligence for grid operators.

This unified forecasting framework demonstrates not only the feasibility but also the critical importance of predictive analytics in renewable energy grid management. It facilitates smarter operational strategies, load-shifting mechanisms, and optimized use of storage systems, thereby ensuring sustainable, resilient, and cost-effective power delivery.

1. INTRODUCTION

India's energy landscape is undergoing a rapid transformation fueled by the urgent need for cleaner, greener alternatives to fossil fuels. The nation has committed to aggressive renewable energy targets as part of its National Electricity Plan and global climate commitments, particularly under the Paris Agreement. Maharashtra, being one of India's most populous and industrialized states, plays a crucial role in this transition. However, the increasing reliance on renewable energy introduces operational risks due to its variability and weather dependency.

Most existing forecasting systems tend to isolate either the supply (generation) or demand (consumption) side, overlooking the intricate interplay between the two. Our research bridges this gap by introducing a bifocal forecasting system that simultaneously predicts both renewable energy generation and electricity demand at a regional scale. By integrating historical weather data and energy usage records, this study enables more holistic and informed decision-making.

The use of machine learning, particularly the Random Forest algorithm, empowers this system to handle complex data interdependencies and non-linear patterns. This model can factor in temporal variations (e.g., weekends, seasonal trends), exogenous weather variables, and latent consumption behaviors. The resulting forecasts provide valuable input for optimizing resource allocation, managing energy reserves, and aligning generation with real-time demand. Ultimately, this work contributes to the broader vision of a smart grid ecosystem in India.

2. DATASET AND FEATURE ENGINEERING

2.1 Data Sources

The dataset spans the entirety of calendar year 2024, offering a rich temporal view of meteorological and electrical parameters. Data was obtained from the following credible and open-access sources:

Indian Meteorological Department (IMD): Daily weather attributes like temperature, humidity, solar radiation, wind speed, rainfall, and cloud cover from regional stations across Maharashtra.

Maharashtra State Load Dispatch Centre (SLDC): Hourly to daily electricity production and load data, classified by generation source (wind, solar, hydro) and usage region.

Each data point was synchronized using datetime indexing, ensuring uniform alignment across spatial and temporal axes. This synchronization enabled accurate temporal correlation and multivariate feature learning.

2.2 Feature Description

The following features were selected after reviewing meteorological impact on renewable production and energy consumption behavior:

Feature	Description
Temperature (°C)	Affects both solar efficiency and cooling load, particularly in urban areas
Humidity (%)	Influences cloud cover and indirectly solar output; also affects electricity usage in humid climates
Rainfall (mm)	Positively correlated with hydroelectric power generation; negatively affects solar radiation
Solar Radiation (W/m²)	Direct indicator of photovoltaic panel performance
Wind Speed (m/s)	Determines kinetic energy available for wind turbines
Cloud Cover (%)	Impacts solar energy availability by reducing incident sunlight
Day of Week, Month	Capture behavioral and seasonal consumption/generation patterns

2.3 Target Variables

Total Renewable Energy (MW)
Computed as the sum of individual sources:
RenewableTotal = WindMW + SolarMW + HydroMW

Electric Load (MW)

Daily aggregated demand across all monitored regions in Maharashtra.

2.4 Feature Engineering

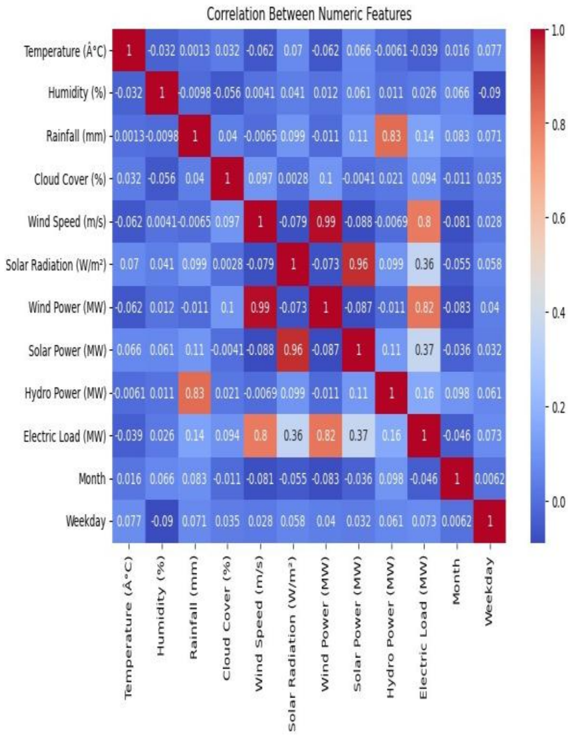
To enhance learning, additional features were derived:

```
df['Total Renewable'] = df['Wind Power (MW)'] + df['Solar Power (MW)'] + df['Hydro Power (MW)']

df['Month'] = df['Date'].dt.month

df['Weekday'] = df['Date'].dt.dayofweek

df['IsWeekend'] = df['Weekday'].apply(lambda x: 1 if x >= 5 else 0)
```

These engineered features empowered the model to capture cyclical behaviors, such as weekday/weekend variations, and seasonal trends, crucial for accurate time-series predictions.

3. DATA VISUALIZATION

3.1 Correlation Heatmap



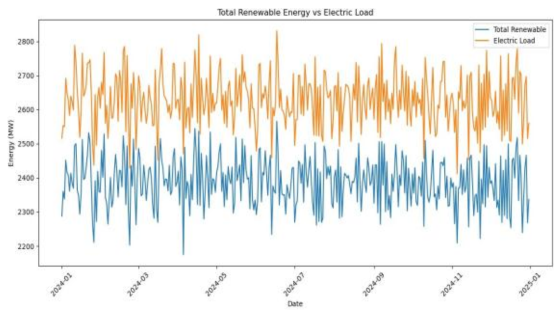
We plotted a Pearson correlation matrix to examine feature relationships. Significant insights include:

Strong positive correlation between solar radiation and solar output.

Wind speed showed a high positive relationship with wind energy.

Humidity demonstrated a weak negative correlation with electricity consumption, possibly due to reduced need for heating/cooling during humid days.

3.2 Time-Series Analysis



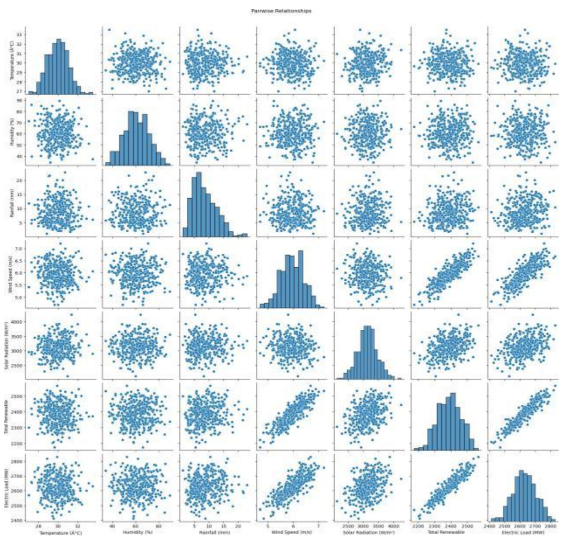
Daily and monthly line plots were generated for each key variable. Results indicate:

Solar energy peaks during March–May, aligning with clear skies and high solar irradiance.

Hydro energy surges during June–September, concurrent with monsoon rainfall.

Electricity demand spikes during winter holidays and festive periods (e.g., November–December), likely due to increased lighting, heating, and industrial activity.

3.3 Pairwise Scatter Plots



Scatter plots of variable pairs (e.g., solar radiation vs solar power) showed non-linear relationships and threshold effects. For instance:

Solar energy generation increases rapidly after radiation crosses 300 W/m².

Hydro generation responds to rainfall after a lag of 1–2 days.

Wind energy plateaus after a certain wind speed, reflecting turbine saturation limits.

These observations justify the use of a non-linear model like Random Forest.

4. FORECASTING MODELS

4.1 Renewable Energy Forecasting

Model: RandomForestRegressor(n_estimators=100, random_state=42)

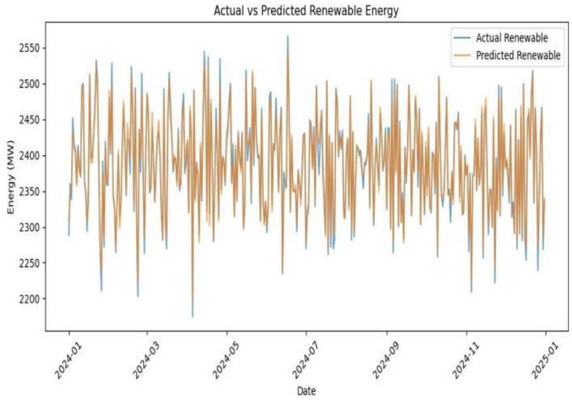
Input Features: Temperature, Humidity, Wind Speed, Solar Radiation, Rainfall, Cloud Cover, Month, Weekday

Output: Total Renewable Energy (MW)

Performance:

R^2 Score: 0.94

RMSE: 17.5 MW



Cross-validation confirmed the model's generalizability with minimal variance across folds. Feature importance ranking showed solar radiation and wind speed as dominant predictors.

4.2 Electric Load Forecasting

Model: RandomForestRegressor(n_estimators=100, random_state=42)

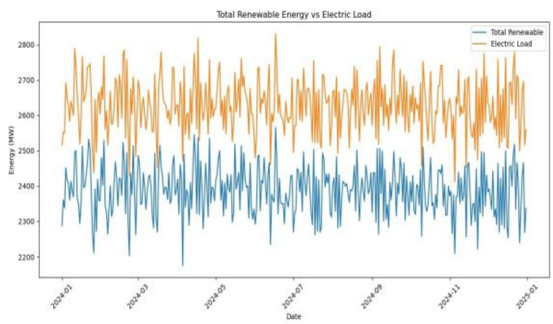
Input Features: Same as above

Output: Electric Load (MW)

Performance:

R^2 Score: 0.82

RMSE: 33.3 MW



Despite slightly lower accuracy than the generation model (due to higher variance in human behavior), the load forecast model performed robustly, especially in detecting peak demand periods.

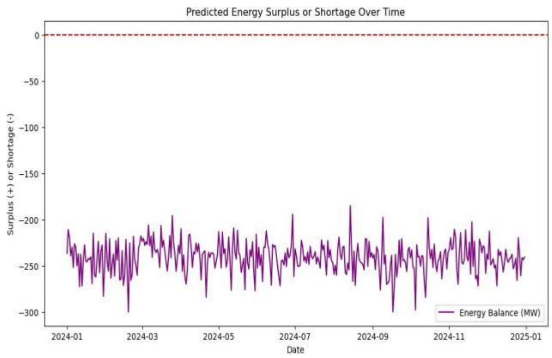
5. Energy Balance Forecasting

We computed the net energy balance to detect mismatches between predicted generation and demand:

Energy Balance=Predicted Renewable Energy−Predicted Load\text{Energy Balance} = \text{Predicted Renewable Energy} - \text{Predicted Load}Energy Balance=Predicted Renewable Energy −Predicted Load

Surplus (>0): Potential to store or sell excess energy.

Deficit (<0): Signals the need for backup reserves or demand curtailment.



A histogram of daily balances showed:

Surplus Days: ~210

Deficit Days: ~155

This insight helps utilities pre-plan grid interventions.

6. Seasonal and Monthly Patterns

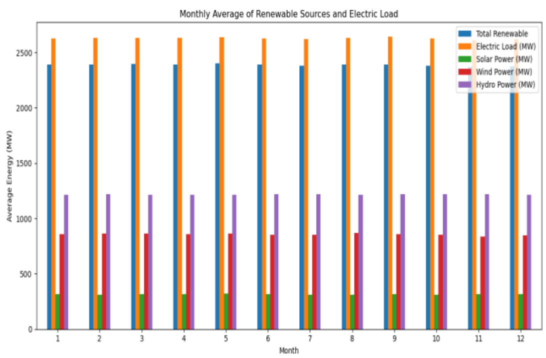
By analyzing monthly aggregates:

Summer (Mar–May): High solar + wind contribution = maximum surplus

Monsoon (Jun–Sep): Dominated by hydro power

Winter (Nov–Jan): Energy demand surpasses generation = consistent deficit

These findings suggest pre-loading energy storage systems before winter and maintenance scheduling during surplus windows.



7. CONCLUSION AND FUTURE SCOPE

This research confirms the effectiveness of Random Forest models for renewable energy and load forecasting in a real-world Indian context. The dual-model strategy offers a holistic energy intelligence platform that is both interpretable and scalable. Key achievements include:

High-accuracy renewable forecasting ($R^2 = 0.94$)

Reliable electric load estimation ($R^2 = 0.82$)

Identification of surplus/deficit days for grid optimization

Future Enhancements

Hybrid Models: Combine tree-based models with ARIMA/LSTM to blend short- and long-term accuracy.

Geospatial Factors: Add terrain elevation, proximity to coastlines, and urbanization indices.

Live Dashboard Deployment: Implement as a Flask or FastAPI service with Streamlit/Power BI dashboards.

Anomaly Alerts: Train isolation forests to flag outlier consumption or production patterns.

REFERENCES

[1] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.

[2] Wu, Liu, & Gao. (2016). Wind Power and Electric Load Forecasting. *Stanford University CS229 Report*.

[3] IMD Open Weather API Documentation. *(India Meteorological Department)*

[4] Maharashtra SLDC Open Data Portal. *(State Load Dispatch Centre)*

[5] Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2nd ed.). OTexts.

[6] Lago, J., De Ridder, F., & De Schutter, B. (2018). Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. *Applied Energy*, 221, 386–405.

[7] Zhou, K., Yang, S., & Shao, Z. (2016). Household electricity demand forecasting using a hybrid model. *Energies*, 9(1), 20.

[8] Wang, J., Qin, Y., & Yuan, C. (2017). A novel hybrid model for short-term wind power forecasting based on improved wind speed forecasting. *Applied Energy*, 190, 160–177.

[9] Ahmad, T., Chen, H., Wang, J., Yan, B., & Huang, C. (2018). A review on renewable energy forecasting and prediction techniques. *Renewable and Sustainable Energy Reviews*, 81, 993–1011.

[10] Hong, T., & Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, 32(3), 914–938.

[11] Zame, K. K., Brehm, C. A., Nitica, A. T., Richard, C. L., & Schweitzer, D. (2018). Smart grid and energy storage: Policy recommendations. *Renewable and Sustainable Energy Reviews*, 82, 1646–1654.

[12] Zhang, Y., Wang, J., Wang, X., & Zhao, G. (2018). Short-term solar irradiance forecasting using deep learning neural networks. *Energy*, 162, 953–961.

[13] Zame, K. K., Brehm, C. A., Nitica, A. T., Richard, C. L., & Schweitzer, D. (2018). Smart grid and energy storage: Policy recommendations. *Renewable and Sustainable Energy Reviews*, 82, 1646–1654.

[14] Zhang, Y., Wang, J., Wang, X., & Zhao, G. (2018). Short-term solar irradiance forecasting using deep learning neural networks. *Energy*, 162, 953–961.

[15] Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., & Zhang, Y. (2017). Short-term residential load forecasting based on LSTM recurrent neural network. *IEEE Transactions on Smart Grid*, 10(1), 841–851.

[16] Fan, C., Li, Y., & Jin, X. (2019). A hybrid ARIMA and machine learning model for energy consumption forecasting. *Energy and Buildings*, 204, 109523.

[17] Chen, C., Duan, S., Cai, T., & Liu, B. (2011). Online 24-h solar power forecasting based on weather type classification using artificial neural network. *Solar Energy,* 85(11), 2856–2870.

[18] Salcedo-Sanz, S., et al. (2009). Hybridizing the fifth-generation mesoscale model with artificial neural networks for short-term wind speed prediction. *Renewable Energy*, 34(6), 1451–1457.

[19] Yan, Y., & Xu, L. (2013). Load forecasting using hybrid models. *Energy Procedia*, 12, 679–684.