

## **Literature Review: Machine Learning-Based Prediction and Optimization of Renewable Energy Production**

The rapid global transition toward sustainable energy has intensified the demand for accurate and reliable renewable energy forecasting. Among renewable sources, solar and wind energy are particularly dependent on atmospheric conditions, making them difficult to predict using conventional statistical methods. Traditional time-series models such as Autoregressive Integrated Moving Average (ARIMA) have been commonly employed due to their simplicity and interpretability. However, these models are limited in capturing non-linear relationships, multi-variable dependencies, and sudden fluctuations inherent in weather-influenced energy production.

To overcome these limitations, modern research has increasingly turned toward machine learning (ML) approaches, which have demonstrated superior accuracy and adaptability. Machine learning models such as Long Short-Term Memory (LSTM) networks and eXtreme Gradient Boosting (XGBoost) are capable of modeling complex, non-linear relationships and learning from vast historical datasets. LSTM networks, a type of recurrent neural network (RNN), are particularly suitable for time-series forecasting due to their ability to retain long-term dependencies. According to Rahman & Saha (2022), ML models like LSTM significantly outperform ARIMA models in forecasting renewable energy outputs due to their robustness in handling missing data, non-stationary series, and seasonality.

XGBoost, a powerful gradient boosting framework, is known for its predictive power and efficiency. It has been successfully applied to various energy forecasting tasks, including solar irradiance estimation, wind power prediction, and hybrid renewable system optimization. In studies comparing XGBoost with traditional regression and tree-based models, XGBoost consistently achieved lower error rates, as highlighted in comparative model assessments (Makridakis et al., 2020).

Prophet, developed by Facebook, is another emerging tool that has found favor in energy forecasting due to its capability to decompose time-series data into trend, seasonality, and holiday effects. Prophet's strength lies in its user-friendliness and ability to produce reliable forecasts even with missing data and irregular time intervals. While it does not match the complexity of deep learning models, its transparency and interpretability make it an excellent addition for cross-verification of model predictions.

Hybrid and ensemble modeling techniques have also garnered attention. By combining statistical models with machine learning frameworks, researchers have achieved higher accuracy and generalizability. The M4 Competition (Makridakis et al., 2020) validated the superiority of such ensemble approaches in time-series forecasting. Similarly, Hyndman & Athanasopoulos (2018) emphasized the need to balance interpretability and accuracy in forecasting tasks, a challenge best addressed through hybrid models.

In terms of data integration, recent literature underscores the importance of including weather parameters—such as global horizontal irradiance (GHI), diffuse and direct irradiance (DHI and DNI), temperature, wind speed, humidity, and cloud cover—as key features in predictive models. These environmental attributes are highly correlated with energy production and help the models adapt to short-term and seasonal variations. Advanced feature engineering, such as

cyclical time representations (e.g., sine and cosine encoding of day or month), and lag features have been shown to significantly enhance model performance.

Furthermore, the open availability of high-quality datasets, such as those provided by the National Renewable Energy Laboratory (NREL), has facilitated comprehensive training and evaluation of machine learning models. These datasets offer granular temporal and geographical data, enabling precise modeling of site-specific renewable energy trends.

In conclusion, the reviewed literature indicates a clear shift from simplistic, linear forecasting methods to sophisticated, data-driven, and hybrid machine learning models. The application of ARIMA, LSTM, XGBoost, and Prophet in renewable energy forecasting allows for better prediction accuracy, model adaptability, and actionable insights. This multifaceted approach not only improves grid reliability but also aids policymakers and utility providers in planning sustainable energy strategies.