**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 (Hyndman and Athanasopoulos, 2018).

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 and 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 (Chen and Guestrin, 2016; 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 (Taylor and Letham, 2018).

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 and 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 (Zhang et al., 2018).

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 (U.S. Department of Energy, 2023).

Other studies have also examined the role of deep learning techniques like Convolutional Neural Networks (CNNs) for spatial-temporal forecasting (Khan et al., 2021), the combination of LSTM and attention mechanisms for improved sequence learning (Wang et al., 2021), and the use of reinforcement learning in optimizing renewable energy systems (Liu et al., 2022).

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.

**References**


- Chen, T. and Guestrin, C., 2016. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp.785–794.
- Hyndman, R.J. and Athanasopoulos, G., 2018. Forecasting: Principles and Practice. 2nd ed. Melbourne: OTexts.
- Khan, Z., Ahmad, M.W., and Chen, L., 2021. A deep learning-based spatiotemporal model for short-term solar irradiance forecasting. Renewable and Sustainable Energy Reviews, 135, p.110120.
- Liu, Y., Qin, H., and Wang, L., 2022. Reinforcement learning for optimal energy dispatching in smart grids with renewable energy. Energy Reports, 8, pp.1645–1656.
- Makridakis, S., Spiliotis, E. and Assimakopoulos, V., 2020. The M4 Competition: Results, findings, conclusion and way forward. International Journal of Forecasting, 36(1), pp.54–74.
- Rahman, M.M. and Saha, T.K., 2022. Machine learning-based forecasting of renewable energy production: A review. Renewable Energy Reports, 8, pp.229–248.
- Taylor, S.J. and Letham, B., 2018. Forecasting at scale. The American Statistician, 72(1), pp.37–45.
- U.S. Department of Energy, 2023. National Renewable Energy Laboratory Solar Radiation Data. [online] Available at: https://www.nrel.gov/grid/solar-resource.html [Accessed 10 February 2025].
- Wang, J., Wang, C., and Zhang, Y., 2021. Attention-based LSTM model for solar power forecasting. IEEE Access, 9, pp.12335–12344.
- Zhang, Y., Wang, J. and Zhang, X., 2018. Short-term solar power forecasting using a hybrid model based on neural networks. Energy, 158, pp.792–806.
- Chen, Y., Zhang, Z., and Wang, X., 2020. Hybrid forecasting approaches for renewable energy generation: A review. Renewable and Sustainable Energy Reviews, 133, p.110287.
- Dong, Z., Yu, R., and Wang, L., 2019. Short-term wind power forecasting based on deep neural networks. Renewable Energy, 140, pp.196–206.

- Gensler, A., Henze, J., Sick, B., and Raabe, N., 2016. Deep learning for solar power forecasting—An approach using autoencoder and LSTM neural networks. In: 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, pp.2858–2865.
- Hong, T., Pinson, P., and Fan, S., 2014. Global Energy Forecasting Competition 2012. International Journal of Forecasting, 30(2), pp.357–363.
- Khosravi, A., Nahavandi, S., Creighton, D., and Atiya, A.F., 2013. Comprehensive review of neural network-based prediction intervals and new advances. IEEE Transactions on Neural Networks and Learning Systems, 22(9), pp.1341–1356.
- Li, J., Han, D., and Zhang, Y., 2021. A hybrid deep learning model for predicting renewable energy output using LSTM and attention mechanism. Applied Energy, 285, p.116370.
- Martinez, A., Valencia, F., and Gil, J.A., 2018. Wind power forecasting using hybrid intelligent models. Renewable Energy, 125, pp.1046–1053.
- Panapakidis, I.P., and Dagoumas, A.S., 2017. Day-ahead electricity price forecasting via the application of artificial neural network based models. Applied Energy, 172, pp.132–151.
- Suganthi, L., and Samuel, A.A., 2012. Energy models for demand forecasting—A review. Renewable and Sustainable Energy Reviews, 16(2), pp.1223–1240.
- Yang, D., Kleissl, J., Gueymard, C., Pedro, H.T., and Coimbra, C.F., 2018. History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. Solar Energy, 168, pp.60–101.