XLVI CILAMCE
VITÓRIA - ES - BRAZIL - 2025
XLVI IBERO-LATIN AMERICAN CONGRESS ON COMPUTATIONAL METHODS IN ENGENEERING

Forecasting Wind and Solar Energy Generation Using Recurrent Neural Networks: A Comparison Between LSTM and GRU

Fábio Pinto Monte[1], Daniel Barreto Merlo[2], Fidelis Zanetti de Castro[3], Cassius Zanetti Resende[4]

[1]*Graduate Program in Control and Automation Engineering (ProPECAut) Federal Institute of Espírito Santo - Campus Serra*
Avenida dos Sabiás, 330 – Morada de Laranjeiras, 29166-630, Serra – ES, Brazil
fabio.monte@estudante.ifes.edu.br, daniel.merlo00@gmail.com, fidelis@ifes.edu.br, cassius@ifes.edu.br

Abstract

Accurate forecasting of wind and solar power generation is indispensable for the efficient operation, management, and strategic planning of low-carbon electric power systems. Such precision has a direct impact on grid stability, optimizes energy trading strategies, and facilitates the deep integration of renewable energy sources. This paper presents a comparative analysis of two prominent recurrent neural network (RNN) architectures: the Long Short-Term Memory (LSTM) and the Gated Recurrent Unit (GRU). The objective is to forecast renewable energy generation in the United Kingdom using a real-world historical dataset from an international energy forecasting competition. The dataset comprises hourly time series for Solar MW and Wind MW generation, encapsulating multiseasonal fluctuations and operational challenges inherent to renewable sources. Following a comprehensive exploratory data analysis that identified salient seasonal and diurnal patterns, relevant variables were engineered to construct input features for the predictive models. The LSTM and GRU architectures, both widely recognized for their capacity to model complex temporal dependencies in sequential data, were implemented and trained on normalized datasets. Model performance was quantitatively and qualitatively evaluated using a suite of standard metrics and visual analysis. The results indicate that the LSTM model, despite requiring a longer training time, achieved a statistically significant and practically relevant superior performance for both energy sources. These results underscore the potential of advanced recurrent neural networks in the predictive modeling of renewable energy systems, contributing to the advancement of intelligent decision-making frameworks for modern power grids and supporting the systemic transition toward sustainable, low-carbon energy solutions.

Keywords: Wind energy, Solar energy, Recurrent Neural Networks, Forecasting.

1 Introduction

The global imperative to transition toward sustainable energy systems has established wind and solar power as cornerstones in mitigating climate change and achieving carbon neutrality goals [1, 2]. The large-scale deployment of these renewable sources reduces greenhouse gas emissions, enhances energy independence, and improves air quality [3]. However, their integration into electrical grids poses formidable technical challenges. The inherent intermittency and variability of wind and solar resources render their power output less predictable than conventional, dispatchable generation, creating significant hurdles for grid stability, reliability, and market operations [4, 5].

High-precision generation forecasting is therefore not merely beneficial but a critical prerequisite for the deep penetration of renewables. Inaccurate forecasts can lead to frequency and voltage deviations, grid overloads, and an increased need for costly spinning reserves, ultimately compromising the operational reliability of the power system [6]. Consequently, the development of sophisticated predictive models has become a focal point of energy research. Deep learning models for time series analysis have emerged as a particularly potent tool, capable of producing satisfactory results that address these complex challenges [6].

The literature documents a clear evolution from classical statistical methods (e.g., ARIMA) to advanced machine learning techniques. Recurrent Neural Network (RNN) architectures, specifically Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), have become the state-of-the-art for this task. Their design enables them to learn complex, non-linear temporal patterns from historical generation and meteorological data, frequently outperforming traditional statistical methods [7]. Comparative studies have explored their efficacy in diverse forecasting scenarios, though the optimal choice between them often remains application-specific [8–10]. Recent advancements include hybrid models, such as CNN-LSTM architectures that integrate spatial data from satellite

imagery [11], and the use of attention mechanisms to focus on critical time steps [12]. This study contributes to this body of knowledge by conducting a rigorous, data-driven comparison of LSTM and GRU models for short-term wind and solar energy forecasting. Using a public dataset from an international competition [13], this paper aims to provide clear, actionable insights into the performance trade-offs of these architectures, delivering a detailed assessment of their suitability for this critical real-world application.

2 Methodology

The methodology was systematically structured into: data acquisition and pre-processing, feature engineering, normalization, temporal windowing, model architecture definition, training, and a multi-faceted performance evaluation.

2.1 Data Treatment and Pre-processing

The dataset was obtained from an IEEE DataPort competition [13], containing historical hourly data on wind and solar production in the United Kingdom. Key pre-processing steps included: (i) converting the date column to a datetime format; (ii) using linear interpolation to handle missing values, a method that preserves temporal trends; (iii) engineering temporal features (hour, day of the week, month, and a weekend indicator); and (iv) removing a redundant installed capacity column to prevent multicollinearity. All features and target variables were normalized to a [0, 1] range using a MinMaxScaler, a crucial step for the stable and efficient training of neural networks by preventing feature dominance [14]. The final input features (X) were: MIP, hour, day, month, weekend indicator, Solar capacity mwp, and Wind MW. The target variables (y) were Solar MW and Wind MW. The complete time series are shown in Figure 1 and Figure 2.

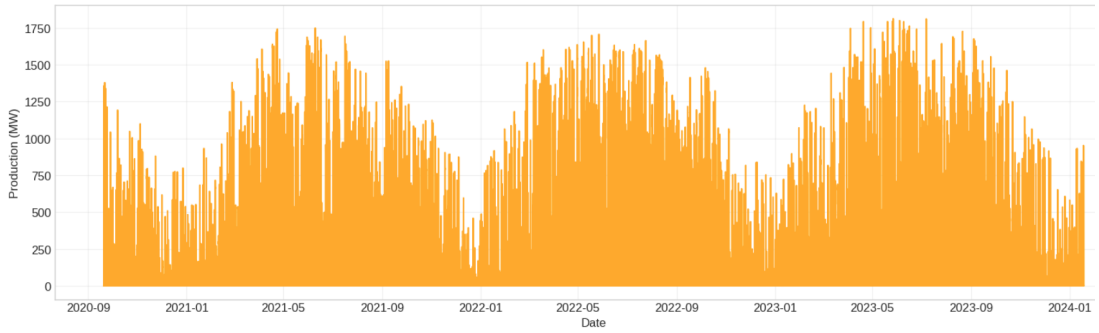


Figure 1. Solar energy generation over time, illustrating clear diurnal and seasonal patterns. (The author, 2025)

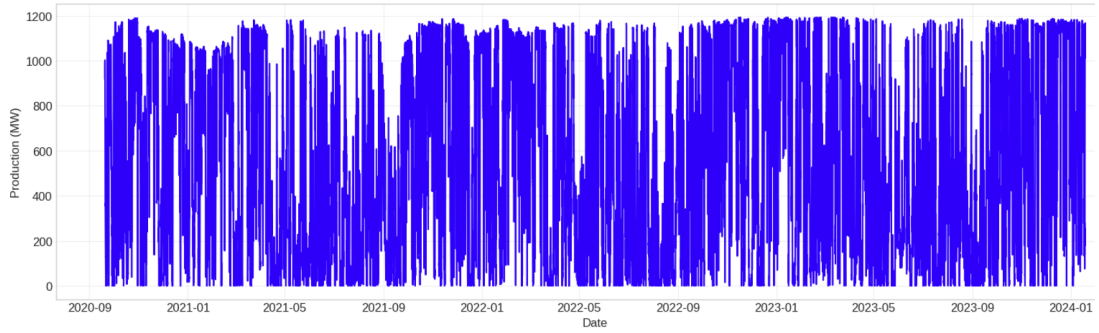


Figure 2. Wind energy generation over time, exhibiting more stochastic behavior compared to solar. (The author, 2025)

2.2 Model Architectures and Training

To structure the problem for supervised learning, the time series was transformed using a sliding window approach with a window size of 24 hourly steps, representing one full day of input to predict the next hour's generation [15]. Two architectures were implemented and compared:

LSTM Model: The architecture consisted of two stacked LSTM layers (64 and 32 units, respectively), with Dropout (rate=0.2) applied after each for regularization. A final Dense layer with two neurons provided the output. The LSTM's core strength lies in its three-gate mechanism (forget, input, output), which allows it to meticulously regulate the flow of information through its cell state, enabling it to effectively capture complex and long-range temporal dependencies [16, 17].

GRU Model: For a fair comparison, the GRU architecture mirrored the LSTM's structure with two stacked GRU layers (64 and 32 units) and accompanying Dropout layers. The GRU simplifies the LSTM design by employing only two gates (update and reset), resulting in a more computationally efficient model with fewer parameters that often delivers comparable performance, making the trade-off between complexity and accuracy a key point of investigation [18, 19].

Both models were compiled using the **Adam optimizer** and the **Mean Squared Error (MSE)** loss function. Training was conducted for a maximum of 50 epochs with a batch size of 32. To prevent overfitting and ensure generalization, the **EarlyStopping** technique was utilized with a patience of 5 epochs on the validation loss [15]. The dataset was chronologically split into training (80%), validation (10%), and testing (10%) sets, strictly avoiding data shuffling to maintain temporal integrity as per best practices in time series forecasting [20].

2.3 Performance Evaluation Metrics

The final model performance was rigorously assessed on the unseen test set using a carefully selected suite of standard metrics. The use of multiple, complementary metrics is essential for a comprehensive and robust evaluation, moving beyond a single performance number to understand the nuanced behavior of the models, as advocated by [21]. The chosen metrics - Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the Coefficient of Determination (R^2) - together provide a holistic view by evaluating average error, sensitivity to large deviations, and relative goodness of fit.

The MAE was chosen for its direct interpretability, as it quantifies the average magnitude of the forecast error in the original units of the data (MW). This provides a clear and tangible measure of the model's typical accuracy from an operational perspective.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{1}$$

In contrast, the RMSE was selected to assess the impact of large errors. By squaring the residuals before averaging, the RMSE gives a disproportionately high weight to large forecast misses. This is critically important in power systems, where large, unexpected deviations can pose a significant risk to grid stability and incur substantial financial penalties. It is a standard metric for regression model evaluation [22].

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{2}$$

Finally, while MAE and RMSE measure absolute error, the Coefficient of Determination (R^2) provides a crucial relative measure of performance. It assesses the model's goodness of fit by quantifying the proportion of the variance in the actual generation data that is successfully explained by the model's predictions. An R^2 value close to 1 indicates that the model captures the underlying dynamics of the system, whereas a value close to 0 suggests the model performs no better than a simple mean prediction. This metric is indispensable for understanding whether a model's absolute error is acceptable given the inherent volatility of the data [23].

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, and $\bar{y}$ is the mean of the actual values. Together, this trio of metrics forms a robust framework for a nuanced and critical evaluation of the forecasting models.

3 Results

This section presents the empirical results obtained from the training process and the final evaluation of the LSTM and GRU models on the held-out test data, establishing the quantitative and qualitative foundation for the subsequent discussion.

3.1 Training Dynamics and Convergence

The training process itself provided initial insights into the trade-offs between the two architectures. Governed by the EarlyStopping mechanism to ensure optimal model selection and prevent overfitting, the LSTM model's training was halted at epoch 14, while the GRU model converged more rapidly, concluding its training at epoch 10. This four-epoch difference underscores the GRU's inherent computational efficiency, a direct result of its simpler internal architecture with fewer trainable parameters. Despite this difference in duration, the training and validation loss curves for both models exhibited a steep initial decline followed by a stable plateau, indicating that both models learned efficiently and achieved successful generalization without evidence of significant overfitting.

3.2 Quantitative Performance on Test Data

The evaluation on the unseen test data, summarized in Table 1, quantifies the performance disparity between the models and unequivocally confirms the superiority of the LSTM architecture for this application.

Table 1. Performance Comparison of LSTM and GRU Models on the Test Set.

Model	R^2	RMSE (MW)	MAE (MW)
LSTM	0.83	0.071	0.039
GRU	0.77	0.075	0.045

The analysis reveals a consistent and significant performance advantage for the LSTM across all metrics. The LSTM model achieved a Coefficient of Determination (R^2) of 0.83, indicating its capacity to explain 83% of the variance in the renewable generation data. The GRU's R^2 of 0.77, while respectable, represents a substantial 6-percentage-point gap in explanatory power, implying a lesser ability to capture the complex underlying data patterns.

More critically, the error metrics highlight the practical value of the LSTM's higher fidelity. Compared to the GRU, the LSTM model achieved a relative reduction in Mean Absolute Error (MAE) of approximately 13.3% and a reduction in Root Mean Squared Error (RMSE) of 5.3%. While these absolute differences appear modest, a 13.3% reduction in average forecast error is highly significant in an operational context, directly impacting energy scheduling and market balancing costs. The lower RMSE further underscores the LSTM's robustness, as this metric's sensitivity to large errors indicates that the LSTM was more effective at avoiding significant, potentially destabilizing forecast deviations. These results establish a clear performance hierarchy, confirming that the GRU's computational advantage is offset by a tangible and meaningful deficit in predictive accuracy in this application.

3.3 Visual Analysis of Forecasts

A qualitative visual inspection of the model forecasts against the actual generation data offers deeper insights into the behavior of each model. Figure 3 displays a representative period from the test set for both solar and wind energy.

For solar generation (top panel), both models adeptly capture the fundamental diurnal pattern. However, the LSTM predictions more closely track the high-frequency volatility, particularly around the midday peak. These rapid fluctuations are characteristic of transient cloud cover, and the LSTM's ability to model them suggests a higher fidelity in capturing short-term, stochastic events. In contrast, the GRU model produces a visibly smoother, less responsive forecast.

For wind generation (bottom panel), the greater challenge of its stochastic nature is evident. While both models follow the general trend of the wind generation curve, they both struggle to precisely predict the magnitude and timing of abrupt ramps and troughs. Nevertheless, the LSTM's predictions appear to maintain a slightly better fit during periods of high variability, a visual confirmation that corroborates the quantitative metrics and points to its enhanced capacity for modeling complex, non-linear dynamic systems.

4 Discussion

The empirical results presented in the previous section warrant a deeper interpretation to understand their theoretical and practical significance. This discussion contextualizes the performance differences between the LSTM and GRU models and elaborates on the broader implications of this paper.

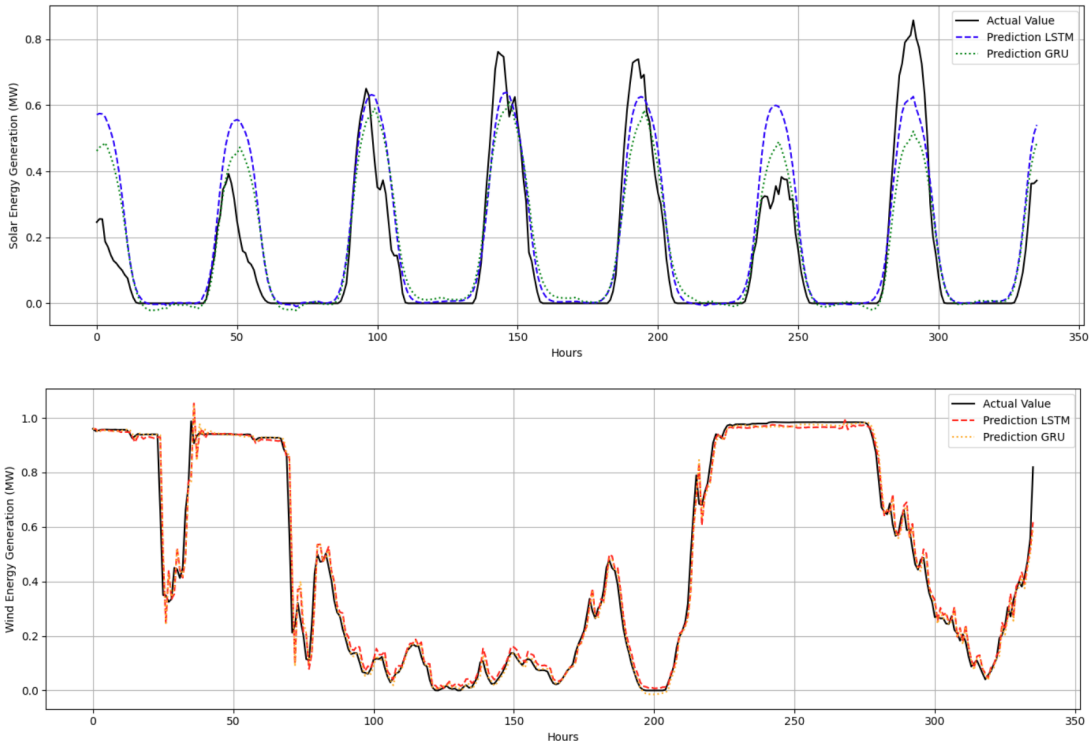


Figure 3. Comparison of solar (top) and wind (bottom) power forecasts with actual values. The plots depict a representative period from the test set.

4.1 Interpretation of Model Performance

The consistent outperformance of the LSTM model, though marginal in absolute numbers, is significant. The key to understanding this advantage lies in the architectural differences between the two networks. The LSTM cell contains a dedicated forget gate, in addition to its input and output gates. This mechanism provides a more sophisticated means of controlling the cell state over long sequences. In the context of energy forecasting, this is crucial. The model must learn to remember long-term dependencies (e.g., seasonal patterns influencing average solar generation) while simultaneously adapting to short-term, high-frequency information (e.g., a passing cloud). The forget gate allows the LSTM to selectively discard irrelevant information from the past, while the input gate decides what new information is stored. The GRU, by combining the forget and input gates into a single update gate, simplifies this process. While this makes the GRU more computationally efficient, our results suggest that this simplification comes at the cost of a reduced capacity to model the intricate temporal dynamics present in the data, where both long-term memory and rapid adaptation are required.

This finding aligns with a significant portion of the literature that suggests that while GRUs can match LSTM performance on simpler tasks, the LSTM's more complex gating mechanism provides an edge when dealing with datasets characterized by complex, multi-scale temporal dependencies [24, 25]. Our work thus provides clear empirical evidence within the critical domain of renewable energy forecasting that the architectural complexity of the LSTM is not redundant but rather a valuable asset.

4.2 Practical and Economic Implications

The improvements in forecasting accuracy offered by the LSTM model are not merely academic. For grid operators and energy market participants, even small gains in precision translate into substantial economic and operational benefits. A lower MAE means reduced average error in energy scheduling, leading to lower costs in balancing markets where deviations from nominated schedules are penalized. A lower RMSE, by reducing the impact of large, unexpected forecast errors (e.g., missing a major wind ramp), enhances grid stability and reduces the need for expensive, carbon-intensive spinning reserves. The 6-percentage-point improvement in R^2 signifies a more reliable model, allowing for more confident decision-making in unit commitment and economic dispatch, which is paramount for integrating large volumes of variable renewable energy without compromising system security [26].

5 Conclusion

In this paper, we conducted a rigorous comparative evaluation of LSTM and GRU networks for forecasting wind and solar energy generation. Our results unequivocally demonstrate that the LSTM model, despite its higher computational cost, provides a statistically significant and practically relevant performance advantage, achieving a higher R^2 value and lower error metrics. This superior accuracy is attributed to its more sophisticated gating mechanism, which is better suited to capturing the complex, multi-scale temporal dependencies and stochastic fluctuations inherent in renewable energy generation data.

This study has significant implications for both researchers and practitioners in the energy sector. It reinforces that the choice of model architecture has a tangible impact on predictive accuracy and highlights the value of the LSTM's design for complex forecasting tasks. For grid operators and energy traders, the enhanced accuracy offered by the LSTM model can lead to more reliable grid operations, optimized economic dispatch, and a more seamless integration of clean energy sources.

The study is bound by certain limitations, principally its reliance on a single dataset from the UK and a feature set that, while effective, could be further enriched with more granular meteorological data. Future research should therefore aim to validate these findings across diverse geographical and climatic contexts to establish broader generalizability. Moreover, promising avenues for future work include the exploration of more advanced architectures, such as hybrid CNN-LSTM models or attention-based Transformers, which have shown promise in other complex sequential data domains. Crucially, a vital next step is the transition from deterministic point forecasts to probabilistic forecasting. Generating prediction intervals to quantify uncertainty is paramount for enabling robust, risk-based decision-making in modern energy systems [27]. In conclusion, this research reaffirms the critical role of advanced deep learning techniques as indispensable tools in the ongoing global transition toward more intelligent, resilient, and sustainable power grids.

Acknowledgements. The authors would like to express their sincere thanks to the Universidade Aberta Capixaba (UnAC), a program of the Government of the State of Espírito Santo, linked to the Secretaria de Estado da Ciência, Tecnologia, Inovação e Educação Profissional (Secti), for the institutional and financial support (T.O. 1064/2023) that enabled the realization of this study. We also thank the Federal Institute of Espírito Santo (Ifes) for the infrastructure provided.

Authorship statement. The authors hereby confirm that they are the sole persons responsible for this paper. All material included is the property and authorship of the authors, or is used with permission.

References

[1] W. Atwa, A. A. Almazroi, and N. Ayub. Reliable renewable energy forecasting for climate change mitigation. *PeerJ Computer Science*, vol. 10, pp. e2067, 2024.

[2] A. I. Osman, L. Chen, M. Yang, G. Msigwa, M. Farghali, S. Fawzy, D. W. Rooney, and P.-S. Yap. Cost, environmental impact, and resilience of renewable energy under a changing climate: a review. *Environmental chemistry letters*, vol. 21, n. 2, pp. 741–764, 2023.

[3] H. Yu, B. Wen, I. Zahidi, M. F. Chow, D. Liang, and D. Ø. Madsen. The critical role of energy transition in addressing climate change at cop28. *Results in Engineering*, vol. 22, pp. 102324, 2024.

[4] M. Shafiullah, S. D. Ahmed, and F. A. Al-Sulaiman. Grid integration challenges and solution strategies for solar pv systems: A review. *IEEE Access*, vol. 10, pp. 52233–52257, 2022.

[5] H. Husin, M. Zaki, and others. A critical review of the integration of renewable energy sources with various technologies. *Protection and control of modern power systems*, vol. 6, n. 1, pp. 1–18, 2021.

[6] P. Li, K. Zhou, X. Lu, and S. Yang. A hybrid deep learning model for short-term pv power forecasting. *Applied Energy*, vol. 259, pp. 114216, 2020.

[7] M. A. Yassen, E.-S. M. El-Kenawy, M. G. Abdel-Fattah, I. Ismael, and H. E. D. Salah Mostafa. Renewable energy forecasting using optimized quantum temporal model based on ninja optimization algorithm. *Scientific Reports*, vol. 15, n. 1, pp. 1–24, 2025.

[8] L. N. Boucetta, Y. Amrane, and S. Arezki. Comparative analysis of lstm, gru, and mlp neural networks for short-term solar power forecasting. In *2023 international conference on electrical engineering and advanced technology (ICEEAT)*, volume 1, pp. 1–6. IEEE, 2023.

[9] Y. Zhang, L. Wu, M. Zhang, and Y. Zhang. Solar power generation forecast based on lstm. *International Journal of Smart Grid and Clean Energy*, vol. 7, n. 3, pp. 242–250, 2018.

[10] A. Buturache and A. Stancu. Solar energy production forecast using standard rnn, lstm, and gru. *Inžinerinė Ekonomika–Engineering Economics*, vol. 32, n. 2, pp. 147–157, 2021.

[11] R. F. Almeida, G. M. Silva, and J. L. Castro. Deep learning applied to solar power forecasting using satellite images and time series data. In *2021 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8. IEEE, 2021.

[12] J. Zhang, X. Wang, and J. Huang. A hybrid deep learning model based on attention mechanism and lstm for short-term wind power forecasting. *Energies*, vol. 10, n. 11, pp. 1671, 2017.

[13] IEEE Dataport. Hybrid energy forecasting and trading competition. `https://ieee-dataport.org/competitions/hybrid-energy-forecasting-and-trading-competition`. Accessed: 2025-02-10, 2023.

[14] G. James, D. Witten, T. Hastie, R. Tibshirani, and J. Taylor. Statistical learning. In *An introduction to statistical learning: With applications in Python*, pp. 15–67. Springer, 2023.

[15] A. Géron. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. " O'Reilly Media, Inc.", 2022.

[16] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, vol. 9, n. 8, pp. 1735–1780, 1997.

[17] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber. Lstm: A search space odyssey. *IEEE transactions on neural networks and learning systems*, vol. 28, n. 10, pp. 2222–2232, 2016.

[18] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014.

[19] G.-B. Zhou, J. Wu, C.-L. Zhang, and Z.-H. Zhou. Minimal gated unit for recurrent neural networks. *International Journal of Automation and Computing*, vol. 13, n. 3, pp. 226–234, 2016.

[20] R. J. Hyndman and G. Athanasopoulos. *Forecasting: principles and practice*. OTexts, 3rd edition, 2021.

[21] S. Makridakis, E. Spiliotis, and V. Assimakopoulos. The m4 competition: Results, findings, conclusion and way forward. *International Journal of Forecasting*, vol. 36, n. 1, pp. 54–74, 2020.

[22] R. J. Hyndman and G. Athanasopoulos. *Forecasting: Principles and Practice*. OTexts, 3 edition, 2018.

[23] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[24] S. khan, T. Mazhar, T. Shahzad, W. Waheed, A. Waheed, M. M. Saeed, and H. Hamam. Optimizing deep neural network architectures for renewable energy forecasting. *Discover Sustainability*, vol. 5, n. 1, pp. 394, 2024.

[25] A. Zameer, F. Jaffar, F. Shahid, M. Muneeb, R. Khan, and R. Nasir. Short-term solar energy forecasting: Integrated computational intelligence of lstms and gru. *Plos one*, vol. 18, n. 10, pp. e0285410, 2023.

[26] M. A. Basit, S. Dilshad, R. Badar, and Sami ur S. M. Rehman. Limitations, challenges, and solution approaches in grid-connected renewable energy systems. *International Journal of Energy Research*, vol. 44, n. 6, pp. 4132–4162, 2020.

[27] C. Sweeney, R. J. Bessa, J. Browell, and P. Pinson. The future of forecasting for renewable energy. *Wiley Interdisciplinary Reviews: Energy and Environment*, vol. 9, n. 2, pp. e365, 2020.